Energy-Efficient Zero-Forcing with User Selection in Parallel Vector Broadcast Channels

Christoph Hellings, Nesrine Damak, and Wolfgang Utschick

2012 International ITG Workshop on Smart Antennas (WSA)

Dresden, Germany, 7th - 8th March, 2012

Technische Universität München
Associate Institute for Signal Processing
http://www.msv.ei.tum.de

Energy-Efficient Zero-Forcing with User Selection in Parallel Vector Broadcast Channels

Christoph Hellings, Nesrine Damak, and Wolfgang Utschick
Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany
Telephone: +49 89 289-28516, Fax: +49 89 289-28504, Email: {hellings,utschick}@tum.de, nesrinedamak@mytum.de

***Abstract*—We consider the maximization of the energy efficiency (or, equivalently, the minimization of the energy per bit) in a set of parallel multiple-input single-output (MISO) broadcast channels (e.g., in the downlink of a multiuser multicarrier system). For a given allocation of users to subchannels, the problem is strictly-quasiconvex and can be efficiently solved by introducing a continuous, differentiable, and concave waterfilling-rate function. To optimize the user allocation, we adopt successive allocation schemes that have been formerly applied to the problem of sum rate maximization. In numerical simulations, we compare different allocation schemes, and we demonstrate that a fixed sum transmit power can lead to a significant loss in energy efficiency if the fixed value is either too high or too low.**

I. Introduction

Energy efficiency has become an important design criterion for wireless communication systems (e.g., [1] and [2]). While early research in this area had mainly focused on wireless sensor networks (e.g., [3] and the references therein), the interest in the optimization of the energy efficiency of cellular wireless systems has recently been growing. Potential to increase the energy efficiency has been identified on all abstraction layers of communication systems (e.g., [1]).

On the circuit level, studies on the energy efficiency were performed, e.g., in [4] and [5]. However, in our work, we want to focus rather on the particularities of multiuser systems with beamforming instead of on physical aspects. Thus, we model all power consumed by the circuit electronics in addition to the transmit power as an affine function of the sum rate as was done in [6] and [7]. The case of constant circuit power (e.g., [1] and [8]–[12]) is a special case of this model. Moreover, we assume that the total bandwidth available for transmission is fixed. Consequently, we do not face the problem discussed in [5] that the energy efficiency can be made arbitrarily large by taking the bandwidth to infinity if the model used for the circuit level power consumption is not accurate enough.

For single-user systems, the energy efficiency in the multicarrier case was studied in [6] and [12]–[14]. Part of the optimization that will be performed in our work, namely the minimization of the energy per bit for a given allocation of streams to users (cf. Section IV), is equivalent to such a single-user optimization. However, instead of adopting one of the various solution methods presented in the abovementioned works, we will propose a new approach whose appeal is that it has a nice intuitive interpretation. More precisely, the problem will be decomposed in finding the maximum sum rate for a given sum transmit power and in optimizing the sum transmit power as an outer problem.

In [15], the energy efficiency of a single-user multiple-input multiple-output (MIMO) system was optimized by performing a singular value decomposition (SVD) in order to diagonalize the channel and by applying the method from [14] to the resulting equivalent channel.

Optimization of the energy efficiency was also considered in multiuser communication systems. The authors of [9] applied methods from non-cooperative game theory to study a multicarrier interference channel. However, as this scenario is qualitatively different from the broadcast channel under consideration (cf. Section II), the approach will not be discussed further here.

For a single-carrier vector broadcast channel, the energy efficiency under the application of random beamforming was studied in [10]. The setup considered in our work is more general in that we consider multiple orthogonal carriers. A MIMO broadcast channel which was also limited to a single carrier was considered in [11], where a heuristic selection criterion was derived based on uniform power allocation.

Research on multiuser multicarrier uplink or downlink systems was performed in [7], [8], [16] and [17] under the assumption of orthogonal frequency division multiple access (OFDMA), i.e., with exclusive assignment of subcarriers to users. In our work, however, we will allow that multiple users are allocated to each carrier as long as the number of users per carrier does not exceed the number of antennas at the base station. Therefore, it is necessary to perform a selection procedure to decide which users shall be served on the various carriers (cf. Section V). To this end, we will apply the greedy user allocation scheme from [18] and the semiorthogonal user selection from [19]. After this process, the streams of the selected users are separated from each other by means of zero-forcing beamforming (cf. Section III), and, finally, a power allocation is performed in a way that minimizes the energy per bit (cf. Section IV).

Notation: In this work, we use boldface lowercase letters for vectors, boldface uppercase letters for matrices, and $\bullet^{\mathrm{H}}$ for the conjugate transpose of a vector or matrix.

II. System Model and Problem Formulation

We consider a set of N orthogonal parallel broadcast channels (e.g., N carriers in a multicarrier downlink system) with an M-antenna base station and K single-antenna receivers,

978-1-4577-1924-0/12/$26.00 ©2012 IEEE

where the data transmission is described by

$$y_k^{(n)} = \boldsymbol{h}_k^{(n),\mathrm{H}} \sum_{k'=1}^{K} \boldsymbol{u}_{k'}^{(n)} \sqrt{p_{k'}^{(n)}} s_{k'}^{(n)} + w_k^{(n)}.$$

The channels $\boldsymbol{h}_k^{(n),\mathrm{H}} \in \mathbb{C}^{1\times M}$ of all users k on all subchannels n are assumed to be frequency flat and perfectly known, and the additive noise is assumed to be circularly symmetric complex Gaussian, i.e., $w_k^{(n)} \sim \mathcal{CN}(0, \sigma_k^{(n),2})$. The beamforming vectors $\boldsymbol{u}_{k'}^{(n)} \in \mathbb{C}^M$ have unit norm, and the transmit symbols $s_{k'}^{(n)}$ are circularly symmetric complex Gaussian with unit variance, i.e., $s_k^{(n)} \sim \mathcal{CN}(0,1)$. Thus, $p_{k'}^{(n)} \geq 0$ are the transmit powers. We will restrict ourselves to linear zero-forcing beamforming which is performed on each carrier separately. Note that due to the user selection performed in Section V, the total number of users K can be arbitrary.

The aim of this work is to minimize the energy per bit

$$E_\mathrm{b} = \frac{\eta^{-1}P + P_c}{R} \tag{1}$$

or, equivalently, to maximize the energy efficiency $\frac{1}{E_\mathrm{b}}$. Here, η is the efficiency of the power amplifier,

$$P = \sum_{k=1}^{K}\sum_{n=1}^{N} p_k^{(n)} \tag{2}$$

is the sum transmit power, and P_c is the power consumed by the circuit electronics apart from the power amplifier, which is assumed to be an affine function $P_c = \alpha + \beta R$ of the sum rate

$$R = \sum_{k=1}^{K}\sum_{n=1}^{N} r_k^{(n)}. \tag{3}$$

Modeling P_c as affine function of the rate is motivated by the fact that part of the circuit power is proportional to the clock frequency, which could be dynamically scaled with the sum rate [6]. On the other hand, this model includes the very common assumption of a constant circuit power (e.g., [1] and [8]–[12]) as the special case $\beta = 0$.

With $c = \alpha\eta$, the optimization problem can be written as

$$\min_{(\boldsymbol{u}_k^{(n)},\ p_k^{(n)} \geq 0)\forall n, \forall k} \frac{P + c}{R} \tag{4}$$

$$\text{s.t.} \quad \boldsymbol{u}_k^{(n),\mathrm{H}} \boldsymbol{u}_j^{(n)} = \delta_{kj} \quad \forall n, k, j \tag{5}$$

where δ_{kj} is the Kronecker delta, which is 1 whenever $i = j$ and 0 otherwise. In (4), we have dropped the multiplicative constant η^{-1} and the additive constant β from the objective function.

As the only matter of interest in (4) is the sum of power and rate over all users and carriers, there is no coupling between the different streams of a user, and we can consider the system as a broadcast channel with KN virtual users denoted by ${}_k^{(n)}$.

III. Zero-Forcing Beamforming

To be able to fulfill the zero-forcing constraints (5), the number of users $S^{(n)}$ scheduled on a carrier n has to be smaller than or equal to the number of transmit antennas ($S^{(n)} \leq M$). Thus, part of the optimization is a scheduling problem, where we have to decide for a set $\mathcal{S}$ which contains all virtual users ${}_k^{(n)}$ that are meant to receive data. After the scheduling decision, the beamformers $\boldsymbol{u}_k^{(n)}$ have to be chosen as the scaled columns of the Moore-Penrose pseudo-inverse $\boldsymbol{H}^{(n),+}$ of the joint channel matrix

$$\boldsymbol{H}^{(n)} = [\boldsymbol{h}_{k^{(n)}(1)}^{(n)}, \ldots, \boldsymbol{h}_{k^{(n)}(S^{(n)})}^{(n)}]^\mathrm{H} \in \mathbb{C}^{S^{(n)}\times M} \tag{6}$$

of the users scheduled on carrier n, where $k^{(n)}(\ell)$ is the user corresponding to the ℓ-th stream on subcarrier n. This leads to $S_\text{tot} = \sum_{n=1}^{N} S^{(n)}$ orthogonal scalar subchannels with channel gains

$$g_{k^{(n)}(\ell)}^{(n)} = \sigma_{k^{(n)}(\ell)}^{(n),-2} \left[\left(\boldsymbol{H}^{(n)}\boldsymbol{H}^{(n),\mathrm{H}}\right)^{-1}\right]_{\ell,\ell}^{-1} \tag{7}$$

where $[\boldsymbol{A}]_{\ell,\ell}$ is the ℓ-th diagonal element of $\boldsymbol{A}$ [20].

For the objective function of the energy efficiency optimization (4), it is not important to which user the streams contributing to the sum rate and sum power belong. We therefore reindex all scheduled data streams with a stream index s such that the channel gains g_s are in decreasing order, i.e., $g_1 \geq g_2 \geq \cdots \geq g_{S_\text{tot}}$. Using the new stream index s, we get the optimization

$$\min_{\mathcal{S}} \min_{(p_s \geq 0)\forall s} \frac{c + \sum_{s=1}^{S_\text{tot}} p_s}{\sum_{s=1}^{S_\text{tot}} r_s} \tag{8}$$

where

$$r_s = W \log_2(1 + p_s g_s) \tag{9}$$

and g_s is a function of $\mathcal{S}$. Here, W is the bandwidth available for each carrier, i.e., the fixed total bandwidth divided by the number of carriers N.[1] The two parts of this optimization problem will be treated in the following sections.

IV. Minimum Energy per Bit for Fixed Scheduling

The inner optimization is equivalent to the optimization in the single-user case considered in [6], [12] and [14]. Therein, the problem was classified as a convex-concave fractional program in S_tot variables, and was converted to a parametric convex program by means of the so-called Dinkelbach method. An equivalent problem was also treated in [13], where a gradient algorithm was applied to a rate space formulation (per-stream rates as optimization variables).

We propose an alternative method to efficiently solve the inner problem, which has a nicer intuitive interpretation: by

[1] Note that in a practical system, the noise powers are $\sigma_k^{(n),2} = W N_{0,k}^{(n)}$, where $N_{0,k}^{(n)}$ is the noise power spectral density of user k on carrier n. Thus, the gains g_s are proportional to $\frac{1}{W}$.

making use of the well-known waterfilling solution [21] of the sum rate maximization problem

$$R_{\text{opt}}(P) = \left(\max_{(p_s \geq 0)_{\forall s}} \quad \sum_{s=1}^{S_{\text{tot}}} r_s \quad \text{s.t.} \quad \sum_{s=1}^{S_{\text{tot}}} p_s \leq P \right) \qquad (10)$$

the inner problem can be reduced to the scalar problem

$$\min_{P \geq 0} \quad \frac{P + c}{R_{\text{opt}}(P)} \qquad (11)$$

where the only optimization variable is the sum transmit power.

The optimal solution to the maximization in (10) is given by [21] as

$$p_s = \max\left\{0,\ \mu - g_s^{-1}\right\} \qquad (12)$$

where the water level μ has to be chosen according to

$$\mu = \frac{P + \sum_{s=1}^{m} g_s^{-1}}{m} \qquad (13)$$

with m being the number of active streams with non-zero power $p_s > 0$. Note that the streams are ordered according to decreasing channel gains (cf. Section III).

The number of active streams m increases with increasing transmit power, and the mth stream becomes active for $\mu = g_m^{-1}$, i.e., for

$$P = P_m = m\, g_m^{-1} - \sum_{s=1}^{m} g_s^{-1}. \qquad (14)$$

Inserting (12) and (13) into (9), we get

$$R_{\text{opt}}(P) = \sum_{s=1}^{m} W \log_2 \left(\frac{P + \sum_{t=1}^{m} g_t^{-1}}{m\, g_s^{-1}} \right) \quad \text{if } P_m \leq P < P_{m+1}. \qquad (15)$$

This means, instead of defining $R_{\text{opt}}(P)$ by means of a maximization as was done in (10), we can write the waterfilling-rate function $R_{\text{opt}}(P)$ explicitly as a piece-wise defined function of P. In the following, we will show that this function is continuous, continuously differentiable, strictly increasing, and concave, and we will explicitly calculate its derivative.

Proposition 1. *$R_{\text{opt}}(P)$ is continuous.*

Proof: Within any interval $]P_m, P_{m+1}[$, continuity is obvious. Near $P = P_m$, we have

$$\begin{aligned} &\lim_{\epsilon \to 0} R_{\text{opt}}(P_m - \epsilon) \\ &= \lim_{\epsilon \to 0} \sum_{s=1}^{m} W \log_2 \left(\frac{m\, g_m^{-1} - \sum_{t=1}^{m} g_t^{-1} - \epsilon + \sum_{t=1}^{m-1} g_t^{-1}}{(m-1) g_s^{-1}} \right) \\ &= \sum_{s=1}^{m} W \log_2 \left(\frac{g_s}{g_m} \right) \end{aligned} \qquad (16)$$

and

$$\begin{aligned} &\lim_{\epsilon \to 0} R_{\text{opt}}(P_m + \epsilon) \\ &= \lim_{\epsilon \to 0} \sum_{s=1}^{m} W \log_2 \left(\frac{m\, g_m^{-1} - \sum_{t=1}^{m} g_t^{-1} + \epsilon + \sum_{t=1}^{m} g_t^{-1}}{m\, g_s^{-1}} \right) \\ &= \sum_{s=1}^{m} W \log_2 \left(\frac{g_s}{g_m} \right) \end{aligned} \qquad (17)$$

which completes the proof. ■

Proposition 2. *$R_{\text{opt}}(P)$ is continuously differentiable.*

Proof: Within any interval $]P_m, P_{m+1}[$, we have

$$\frac{\partial R_{\text{opt}}}{\partial P} = \frac{W m}{(\ln 2) \left(P + \sum_{s=1}^{m} g_s^{-1}\right)}. \qquad (18)$$

Near $P = P_m$, we have

$$\begin{aligned} &\lim_{\epsilon \to 0} \frac{\partial R_{\text{opt}}}{\partial P}(P_m - \epsilon) \\ &= \lim_{\epsilon \to 0} \frac{W(m-1)}{(\ln 2) \left(m\, g_m^{-1} - \sum_{s=1}^{m} g_s^{-1} - \epsilon + \sum_{s=1}^{m-1} g_s^{-1}\right)} \\ &= \frac{W g_m}{\ln 2} \end{aligned} \qquad (19)$$

and

$$\begin{aligned} &\lim_{\epsilon \to 0} \frac{\partial R_{\text{opt}}}{\partial P}(P_m + \epsilon) \\ &= \lim_{\epsilon \to 0} \frac{W m}{(\ln 2) \left(m\, g_m^{-1} - \sum_{s=1}^{m} g_s^{-1} + \epsilon + \sum_{s=1}^{m} g_s^{-1}\right)} \\ &= \frac{W g_m}{\ln 2} \end{aligned} \qquad (20)$$

which completes the proof. ■

Proposition 3. *$R_{\text{opt}}(P)$ is strictly increasing and strictly concave.*

Proof: As can be easily verified from (15) and (18), respectively, $R_{\text{opt}}(P)$ is strictly increasing in all intervals, and its first derivative is strictly decreasing in all intervals. ■

Note that unlike the first derivative, the second derivative jumps at the interval boundaries. Apart from this property, the optimal rate function $R_{\text{opt}}(P)$ has the same qualitative behavior as the rate r_s of a single stream [cf. (9)] when considered as a function of p_s. Therefore, disregarding the fact that we use the slightly modified rate function $R_{\text{opt}}(P)$, the optimization (11) is equivalent to the energy efficiency optimization in the single-stream case, which was discussed, e.g., in [1].

To illustrate this, we have included Fig. 1, which shows a small example with $W = \frac{1}{4}$, $g_1 = 4$, $g_2 = 2$, $g_3 = \frac{4}{3}$, $g_4 = 1$, and $c = 0.2$. In the plot, it can be seen that the function $R_{\text{opt}}(P)$ is increasing and concave and has indeed the logarithmic shape of a conventional rate function. At $P \in \{\frac{1}{4}, \frac{3}{4}, \frac{3}{2}\}$, new streams are activated, i.e., m increases, which is in compliance with (14). The discontinuity of the second derivative at these points can be observed in the plot as

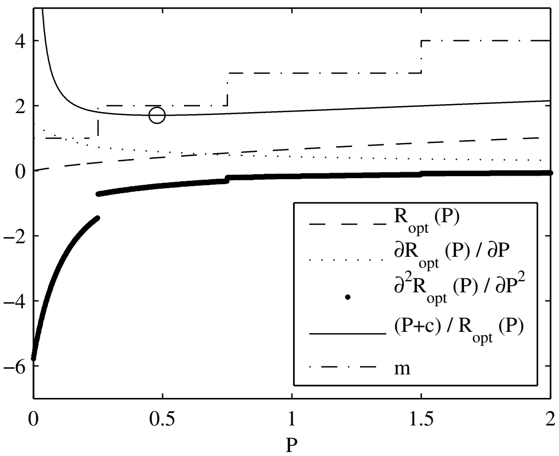

Fig. 1. The optimal rate function $R_{\text{opt}}(P)$ with its first and second derivative, the objective function of (11) with its optimum (circle), and the number of active streams m.

well as the continuity and monotonicity of the first derivative. The objective function $(P+c)/R_{\text{opt}}(P)$ of the energy-per-bit minimization (11) has the same shape as the single-stream version in [1]. The aim is now to find the minimum of this function, which is marked in the plot with a blank circle.

As the numerator of the cost function in (11) is affine in P and the denominator is concave in P, the problem is a convex-concave fractional program. The same was true for the formulation chosen in [6], [12] and [14]. However, therein, the considered convex-concave fractional program was an optimization in many variables while (11) has only one optimization variable, i.e., it is a scalar problem.

According to [22], the cost function of a convex-concave fractional program is strictly quasiconvex, which implies that any local optimum is a global one. Moreover, as is shown in Appendix A, there is a unique point with vanishing derivative. Therefore, to find the global optimum, we have to find the root of the derivative

$$\frac{\partial f}{\partial P} = \frac{R_{\text{opt}}(P) - (P+c)\frac{\partial R_{\text{opt}}(P)}{\partial P}}{R^2_{\text{opt}}(P)} \tag{21}$$

of $f(P) = (P+c)/R_{\text{opt}}(P)$. Due to the fact that the only stationary point of $f(P)$ is a minimum, a sign change of $\frac{\partial f}{\partial P}$ must occur at this root of $\frac{\partial f}{\partial P}$, which means that we can find it by applying the bisection method.

V. User Selection

In the last section, we have derived a method to optimize the energy per bit for a given user selection $\mathcal{S}$. The aim is now to find a good choice for $\mathcal{S}$ that leads to a low energy per bit.

The globally optimal solution $\mathcal{S}_{\text{opt}}$ of the combinatorial user allocation problem could be obtained by an exhaustive search, which is not feasible for large systems due to the exponential complexity of evaluating the energy per bit for the $\left(\sum_{m=1}^{\min\{M,K\}} \binom{K}{m}\right)^N$ possible allocations.

Therefore, we propose to perform a successive user allocation by means of the the greedy user selection (GUS) scheme from [18] or the semiorthogonal user selection (SUS) from [19]. These methods were originally proposed for the problem of sum rate maximization, but can also be applied for the energy efficiency optimization.

A. Greedy User Selection (GUS)

The greedy user allocation from [18] successively allocates streams to users by a series of locally optimal decisions, leading to a globally suboptimal solution in general.

In step i, for each carrier n that still has available spatial degrees of freedom (i.e., the number of users $S^{(n)}$ scheduled on carrier n is smaller than the number of transmit antennas M) and for each user k that is not yet scheduled on n,[2] the user allocation $\mathcal{S}(\binom{(n)}{k}) = {}^{(i-1)}\mathcal{S} \cup \{\binom{(n)}{k}\}$ resulting from adding the virtual user $\binom{(n)}{k}$ to the current allocation ${}^{(i-1)}\mathcal{S}$ is created. Here, the left superscript is the step index. For each $\mathcal{S}(\binom{(n)}{k})$, the resulting energy per bit $E_{\text{b}}(\binom{(n)}{k})$ is computed by solving the inner problem as described in Section IV, and we set ${}^{i}\mathcal{S}$ to the best allocation, i.e.,

$${}^{i}\mathcal{S} \leftarrow \mathcal{S}(\operatorname{argmin}_{\binom{(n)}{k}} E_{\text{b}}(\binom{(n)}{k})). \tag{22}$$

The algorithm terminates when the energy per bit increases from one step to the next or when M streams have been allocated on each carrier.

Note that for the energy efficiency optimization—just like for the sum rate maximization in [18]—there is no need for a special initialization phase. Such a phase was, e.g., necessary for the quality-of-service-constrained optimizations considered in [23] and [24]. In case of a problem with quality of service constraints (e.g., minimum rate constraints), the inner problem is not feasible unless each user has obtained at least one data stream. Therefore, it cannot be used as a decision criterion in the first steps of the greedy algorithm. As there are no such constraints in the energy efficiency optimization (4), the algorithm could be started with an empty set ${}^{0}\mathcal{S} = \varnothing$. However, to reduce the computational complexity, we propose the following initialization.

Since the energy per bit is small if the channel gains of the allocated streams are high, it is clear that the first user to be allocated on a carrier is the one with maximal channel norm $\|\boldsymbol{h}_k^{(n)}\|_2$. Therefore we can start with an initial allocation

$${}^{0}\mathcal{S} = \left\{ \binom{(n)}{k} \,\middle|\, n \in \{1,\ldots,N\} \wedge k = \operatorname*{argmax}_{k'} \left\| \boldsymbol{h}_{k'}^{(n)} \right\|_2 \right\} \tag{23}$$

i.e., N streams are already allocated before the actual greedy algorithm starts.

By doing so, the number of inner optimizations is at most $(KN-N)+(KN-N-1)+\cdots+(KN-MN+1) < KMN^2$. Therefore, the algorithm can be executed in reasonable time even for large numbers of users K. However, the number of evaluations of the inner problem grows quadratically in the number of carriers N.

[2]The last condition is specific to single-antenna receivers.

To avoid this, we also propose an alternative algorithm where the greedy allocation is performed separately on each carrier as follows. For each carrier n, the greedy algorithm is applied in order to find a set $\mathcal{S}^{(n)}$ of virtual users ${}_{k}^{(n)}$ scheduled on carrier n that leads to a low energy per bit on carrier n. Note that this reduces the number of possible choices in each step of the greedy scheme as well as the problem dimension of the inner problem. Moreover, the greedy algorithm can be executed on the various carriers in a parallelized manner. Afterwards, we set $\mathcal{S} = \bigcup_{n=1}^{N} \mathcal{S}^{(n)}$, and we perform the inner optimization from Section IV in order to find the optimal power allocation for the resulting $\mathcal{S}$. In total, the inner optimization is performed up to $N\left((K-1)+(K-2)+\cdots+(K-M+1)\right) < KMN$ times for a single carrier (i.e., with reduced problem dimension) and only once for the overall system. This reduces the complexity of the method significantly in case of a high number of carriers. In the numerical simulations in Section VI, it will turn out that the per-carrier version performs close to the original greedy scheme.

B. Semiorthogonal User Selection (SUS)

While the greedy user allocation scheme from [18] makes decisions based on evaluations of the objective function, the decisions of the semiorthogonal user selection (SUS) from [19] only rely on the channel conditions.

The method is based on the observation that the achieved energy per bit is small whenever the joint channel matrix $\boldsymbol{H}^{(n)}$ defined in (6) is well conditioned, which is the case if the channels of the selected users are nearly orthogonal. As the method is based only on the channel conditions, it can be performed separately on each carrier. As before, this could again be done in a parallelized manner.

In the ith step, the channels of all candidate users ${}_{k}^{(n)}$ on carrier n are projected onto the orthogonal complement of $\mathrm{span}[{}^{(i-1)}\boldsymbol{H}^{(n)}]$, where ${}^{(i-1)}\boldsymbol{H}^{(n)}$ is the joint channel matrix of the virtual users ${}_{k}^{(n)} \in {}^{(i-1)}\mathcal{S}^{(n)}$ selected in step $i-1$ on carrier n. Then, the user whose projected channel vector has the biggest norm is chosen to be included in the new set of scheduled users ${}^{i}\mathcal{S}^{(n)}$.

The complexity is significantly lower than the complexity of the greedy scheme since for each candidate, only the norm of a projected vector has to be computed while the greedy scheme performs the whole inner optimization for each candidate. Moreover, the complexity is further reduced by only considering users whose channel vectors are semiorthogonal to $\mathrm{span}[{}^{(i-1)}\boldsymbol{H}^{(n)}]$, i.e., whose component in $\mathrm{span}[{}^{(i-1)}\boldsymbol{H}^{(n)}]$ is smaller than a parameter α. If this criterion is not fulfilled for a user in a step, the user is permanently removed from the set of candidate users. This not only reduces the complexity, but also ensures that the resulting joint channel matrix $\boldsymbol{H}^{(n)}$ after the execution of the algorithm is well conditioned. The algorithm for carrier n terminates when M users have been allocated on that carrier or when there are no further candidates.

Finally, the overall user allocation is obtained from $\mathcal{S} = \bigcup_{n=1}^{N} \mathcal{S}^{(n)}$, and the overall energy efficiency is optimized for this user allocation by performing the inner optimization once.

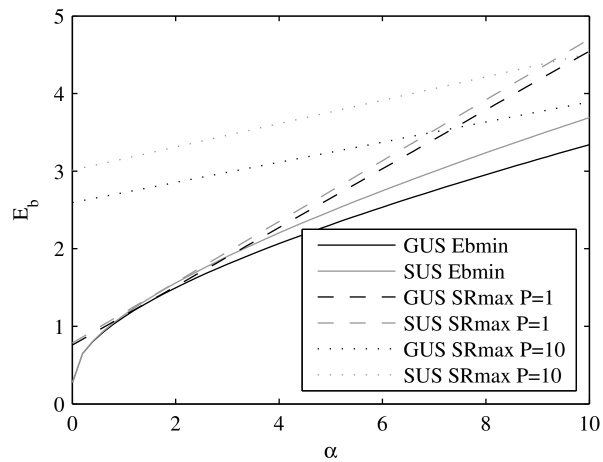

Fig. 2. Energy per bit achieved in a system with $M = 2$ transmit antennas, $K = 10$ users, $N = 2$ subchannels, circuit power $P_c = \alpha$, and amplifier efficiency $\eta = 0.5$ (averaged over 1000 channel realizations).

As the SUS method makes its decisions only based on the channel properties and not on the actual objective function, we expect it to perform worse than the greedy algorithm. This can indeed be observed in the numerical results in Section VI. However, the increase in energy per bit comes with a reduced computation time.

VI. Numerical Results and Discussion

For the numerical simulations that will be presented in this section, we have used 1000 realizations of i.i.d. circularly symmetric complex Gaussian channel coefficients with zero mean and unit variance. We have set the per-carrier bandwidth to $W = \frac{1}{N}$ and the noise power to $\sigma_{k}^{(n),^2} = W = \frac{1}{N}$ for all users and all carriers.

The system considered for the simulations presented in Fig. 2 is a small example with $N = 2$ carriers, $M = 2$ transmit antennas, $K = 10$ users, and a constant circuit power $P_c = \alpha$, i.e., $\beta = 0$. For the efficiency of the power amplifier, we have chosen $\eta = \frac{1}{2}$, but other values of η can be read from the same plot by scaling both the α and the E_{b} axis since

$$\eta E_{\mathrm{b}} = \frac{P + \eta\alpha}{R} \tag{24}$$

if $\beta = 0$ (cf. Section II).

In the figure, we plot the energy per bit resulting from the application of the proposed GUS-based and SUS-based energy efficiency optimization as well as from the application of the respective sum rate maximization for a fixed transmit power P. As can be seen, the greedy user selection outperforms the semiorthogonal user selection in all cases, and the difference is most pronounced for large values of the circuit power $P_c = \alpha$ and the sum transmit power P. On the other hand, as discussed in Section V, the semiorthogonal user selection has a lower computational complexity.

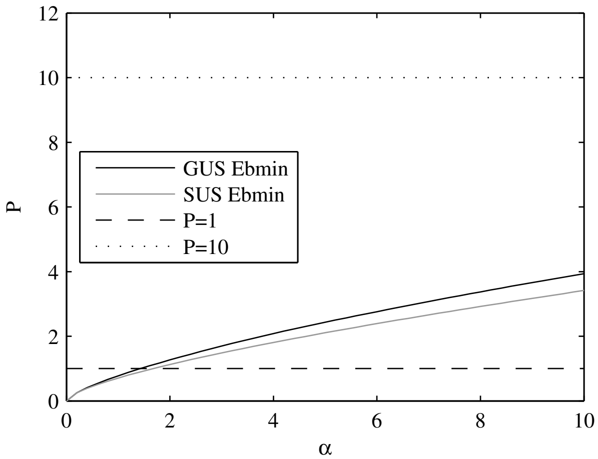

Fig. 3. Resulting sum transmit power in a system with $M = 2$ transmit antennas, $K = 10$ users, $N = 2$ subchannels, circuit power $P_c = \alpha$, and amplifier efficiency $\eta = 0.5$ (averaged over 1000 channel realizations).

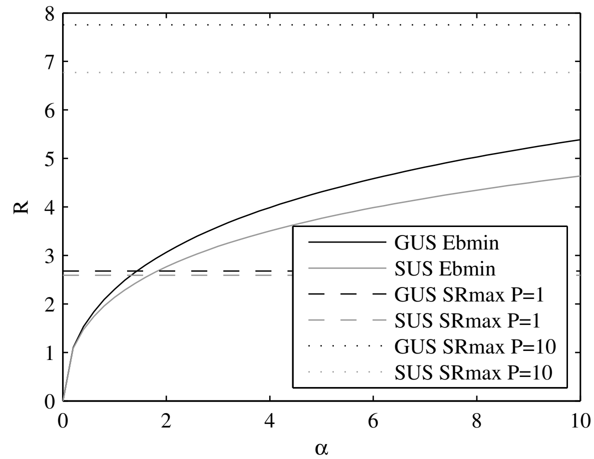

Fig. 4. Resulting sum rate in a system with $M = 2$ transmit antennas, $K = 10$ users, $N = 2$ subchannels, circuit power $P_c = \alpha$, and amplifier efficiency $\eta = 0.5$ (averaged over 1000 channel realizations).

If the sum rate is maximized for fixed sum transmit power P, the strategy resulting from the optimization does not depend on the actual value of the circuit power, but the obtained energy per bit does. In this case, the achieved energy per bit is an affine function of the circuit power, as can be seen from (1) and from Fig. 2. Due to the suboptimality of a fixed sum transmit power, the affine curve always lies above the curve of the corresponding energy efficiency optimization and touches it at a certain point as can be seen for the case $P = 1$ in Fig. 2. At this point the circuit power is such that the fixed sum transmit power is by chance the optimal choice. This can also be observed in conjunction with Fig. 3, where we have plotted the sum transmit power P for the same scenario. The value of α where the curve of the optimized sum transmit power intersects the horizontal line of the fixed sum transmit power $P = 1$ corresponds to the point of tangency in Fig. 2.

However, for other values of α, the energy per bit resulting from the sum rate maximization with fixed transmit power is much higher than the value obtained when performing the energy efficiency optimization as can be clearly seen in the plots. In particular, for very low transmit power P, we get a very low sum rate $R(P)$ leading to a high energy per bit due to the term $\frac{P_c}{R(P)} \overset{P\to 0}{\to} \infty$ in the definition of E_b [cf. (1)]. On the other hand, for very high transmit power P, the sum rate $R(P)$ grows much slower than P due to its logarithmic dependence on P. In this case, we get again a high energy per bit since we are penalized by the term $\frac{P}{R(P)} \overset{P\to\infty}{\to} \infty$. The optimal sum transmit power P lies in between and can be found as proposed in Section IV. Performing such an optimization can significantly improve the energy efficiency of the considered communication system.

In Fig. 3, we can also see that the utilized sum transmit power is increasing in α, which is the constant part of the circuit power. To understand this behavior, it helps to look at Fig. 4, where the resulting sum rate is plotted. If the circuit power is high, there is a strong penalty for slow transmission. Therefore, the sum rate has to grow with α. However, to achieve a higher sum rate, an increased sum transmit power is needed (cf. Proposition 3). On the other hand, if the circuit power tends to zero, it is optimal to transmit infinitely slow, i.e., with rate and transmit power tending to zero. This effect was discussed, e.g., in [1].

In Fig. 5, we show the optimal energy per bit for a larger system with $K = 20$ users, $N = 16$ carriers, and $M = 2$ base station antennas. In this case, we have assumed a non-zero parameter β, i.e., the circuit power now grows linearly with the sum rate. However, from (1), we have

$$E_\mathrm{b} = \frac{\eta^{-1}P + \alpha + \beta R}{R} = \frac{\eta^{-1}P + \alpha}{R} + \beta \qquad (25)$$

which means that the parameter β only yields an offset on the E_b-axis as can also be observed in the figure. The behavior of the optimal energy per bit as a function of the parameter α as well as the different performance of the two considered user selection algorithms are qualitatively the same as in the smaller system.

In Fig. 5, we have also included the per-carrier version of the greedy user selection (cf. Section V-A), which performs so close to the original version that the difference can hardly be noticed in the plot. A slightly larger difference between the conventional and the per-carrier GUS can, however, be observed in the same system if the number of users is decreased to $K = 10$ as can be seen in Fig. 6, where we have plotted the achieved energy per bit for different numbers of users $K \in \{10, 20, \ldots, 100\}$.

In that figure, we can also observe that the energy per bit is lower if the system has a larger number of users. This effect is called multiuser diversity and was discussed, e.g., in [18] and [19] for the sum rate maximization problem: the more users we have in the system, the higher is the probability that on each carrier, we can find M users whose channels are strong

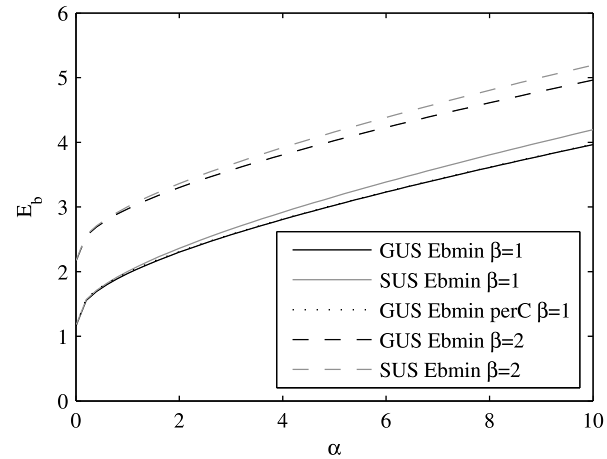

Fig. 5. Energy per bit achieved in a system with $M = 2$ transmit antennas, $K = 20$ users, $N = 16$ subchannels, circuit power $P_c = \alpha + \beta R$, and amplifier efficiency $\eta = 0.5$ (averaged over 1000 channel realizations).

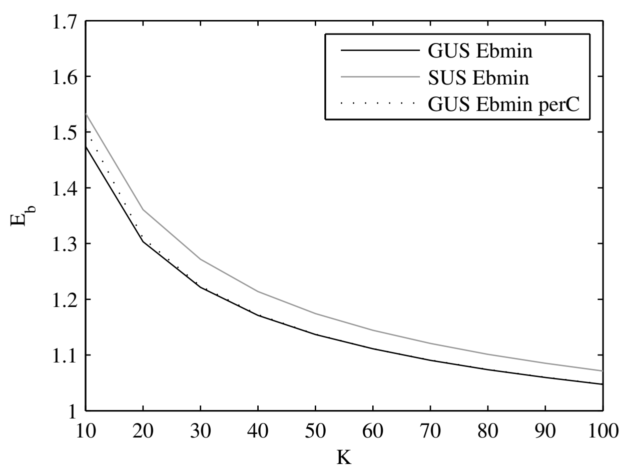

Fig. 6. Energy per bit achieved in a system with $M = 2$ transmit antennas, $N = 2$ subchannels, circuit power $P_c = \alpha = 2$, $\eta = 0.5$, and various numbers of users (averaged over 1000 channel realizations).

and nearly orthogonal to each other. Since this leads to higher rates for a given sum transmit power, it also yields a lower energy per bit or, equivalently, an increased energy efficiency. This behavior was also shown for the random beamforming method for single-carrier vector broadcast channels in [10].

VII. Conclusion

Optimizing the sum rate of a communication system for a fixed sum transmit power can lead to a high energy per bit, i.e., a bad energy efficiency. On the other hand, allowing a variable sum transmit power that is adapted such that the energy per bit is optimized yields a much better energy efficiency.

In the parallel vector broadcast channels considered in this paper, a simple way of implementing such an energy efficient transmission is the application of zero-forcing beamforming in combination with a successive user allocation algorithm. The optimization problem that remains to be solved for a given user allocation is then equivalent to the energy efficiency optimization in a single-user multicarrier scenario. For this inner optimization, we have proposed a new method which uses the sum transmit power as the only optimization variable and encapsulates the detailed power allocation in a so-called waterfilling-rate function. This allows an intuitive understanding of the energy efficiency optimization and reveals similarities to the case with a single data stream.

Applying the proposed method, a significant reduction of the energy per bit compared to the case of transmission with a fixed sum transmit power could be observed in numerical simulations.

Appendix A

Proposition 4. *There is only one stationary point of $f(P) = (P + c)/R_{\text{opt}}(P)$, and this point is the global minimum.*

Sketch of proof: The second derivative of $f(P)$ is

$$\frac{\partial^2 f(P)}{\partial P^2} = -\frac{R_{\text{opt}}(P)(P+c)\frac{\partial^2 R_{\text{opt}}(P)}{\partial P^2}}{R^3_{\text{opt}}(P)} - \frac{\left(R(P) - (P+c)\frac{\partial R_{\text{opt}}(P)}{\partial P}\right) 2\frac{\partial R_{\text{opt}}(P)}{\partial P}}{R^3_{\text{opt}}(P)} \tag{26}$$

at points where $\frac{\partial^2 R_{\text{opt}}(P)}{\partial P^2}$ exists. Note that the second summand is zero whenever the first derivative of $f(P)$ is zero [cf. (21)]. Moreover, as $P > 0$, we have the following inequalities at stationary points of $f(P)$: $R_{\text{opt}}(P) > 0$, $(P + c) > 0$, and $\frac{\partial^2 R_{\text{opt}}(P)}{\partial P^2} < 0$. Therefore, all stationary points of $f(P)$ where $\frac{\partial^2 R_{\text{opt}}(P)}{\partial P^2}$ exists are minima. If $\frac{\partial^2 R_{\text{opt}}(P)}{\partial P^2}$ does not exist at a stationary point of $f(P)$, the same can be shown by considering the right and left second derivative of $f(P)$ and $R_{\text{opt}}(P)$.

Note that above observations also imply that the function cannot be constant on any interval as the first and second derivatives cannot be zero at the same time. On the other hand, due to strict quasiconvexity of $f(P)$, local minima can only be right next to each other. Therefore, there must be a unique local minimum which is also the global one, and there are no other stationary points. ■

References

[1] G. Miao, N. Himayat, Y. G. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: a survey," *Wireless Commun. and Mobile Computing*, vol. 9, no. 4, pp. 529–542, Apr. 2009.

[2] G. Y. Li, Z.-K. Xu, C. Xiong, C.-Y. Yang, S.-Q. Zhang, Y. Chen, and S.-G. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," to appear in *IEEE Wireless Commun. Mag.*

[3] L. Wang and Y. Xiao, "A survey of energy-efficient scheduling mechanisms in sensor networks," *Mobile Netw. and Appl. Archive*, vol. 11, no. 5, pp. 723–740, Oct. 2006.

[4] A. Wang and C. Sodini, "On the Energy Efficiency of Wireless Transceivers," in *Proc. Int. Conf. Commun. (ICC)*, Jun. 2006, pp. 3783–3788.

[5] A. Mezghani, N. Damak, and J. Nossek, "Circuit aware design of power-efficient short range communication systems," in *Proc. 7th Int. Symp. on Wireless Commun. Syst. (ISWCS)*, Sep. 2010, pp. 869–873.

[6] C. Isheden and G. P. Fettweis, "Energy-Efficient Multi-Carrier Link Adaptation with Sum Rate-Dependent Circuit Power," in *Proc. IEEE GLOBECOM 2010*, Miami, FL, USA, Dec. 2010.
[7] C. Xiong, G. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and Spectral-Efficiency Tradeoff in Downlink OFDMA Networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3874–3886, Nov. 2011.
[8] G. Miao, N. Himayat, Y. Li, and D. Bormann, "Energy Efficient Design in Wireless OFDMA," in *Proc. Int. Conf. Commun. (ICC)*, May 2008, pp. 3307–3312.
[9] G. Miao, N. Himayat, G. Y. Li, A. T. Koc, and S. Talwary, "Distributed Interference-Aware Energy-Efficient Power Optimization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1323–1333, Apr. 2011.
[10] Z. Chong and E. Jorswieck, "Energy Efficiency in Random Opportunistic Beamforming," in *Proc. Veh. Technol. Conf. (VTC) 2011-Spring*, Budapest, Hungary, May 2011.
[11] J. Xu, L. Qiu, and C. Yu, "Link Adaption and Mode Switching for the Energy Efficient MIMO Systems," to appear in *IEICE Trans.*
[12] C. Isheden and G. Fettweis, "Energy-Efficient Link Adaption on Parallel Channels," in *Proc. 19th European Signal Process. Conf. (EUSIPCO)*, Aug./Sep. 2011, pp. 874–878.
[13] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 545–554, Feb. 2010.
[14] R. Prabhu and B. Daneshrad, "An Energy-Efficient Water-Filling Algorithm for OFDM Systems," in *Proc. Int. Conf. Commun. (ICC)*, Cape Town, South Africa, May 2010.
[15] R. S. Prabhu and B. Daneshrad, "Energy-Efficient Power Loading for a MIMO-SVD System and its Performance in Flat Fading," in *Proc. IEEE GLOBECOM 2010*, Miami, FL, USA, Dec. 2010.
[16] G. Miao, N. Himayat, G. Li, and S. Talwa, "Low-Complexity Energy-Efficient OFDMA," in *Proc. Int. Conf. Commun. (ICC)*, Dresden, Germany, Jun. 2009.
[17] X. Ge, J. Hu, C.-X. Wang, C.-H. Youn, J. Zhang, and X. Yang, "Energy Efficiency Analysis of MISO-OFDM Communication Systems Considering Power and Capacity Constraints," 2011, to appear in *Mobile Netw. and Appl.*
[18] G. Dimic and N. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.
[19] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
[20] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
[21] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.
[22] S. Schaible, "Minimization of ratios," *J. Optim. Theory and Appl.*, vol. 19, no. 2, pp. 347–352, Jun. 1976.
[23] C. Guthy, W. Utschick, and G. Dietl, "Spatial Resource Allocation for the Multiuser Multicarrier MIMO Broadcast Channel – A QoS Optimization Perspective," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2010, pp. 3166–3169.
[24] C. Hellings, M. Joham, and W. Utschick, "Power Minimization in Parallel Vector Broadcast Channels with Zero-Forcing Beamforming," in *Proc. IEEE GLOBECOM 2010*, Miami, FL, USA, Dec. 2010.