Adaptive Data Transmission in Downlink MIMO-OFDM Systems with Pre-equalization

P. Trifonov
St.Petersburg State Polytechnic University and University of Ulm
194021, Polytechnicheskaya str. 21, St.Petersburg, Russia
e-mail:petert@dcn.infos.ru

Elena Costa
Siemens AG, Com MN PG NT RI4
Sankt-Martin Str. 76, D-81541 München, Germany
e-mail: elena.costa@siemens.com

Abstract— **An optimization algorithm for finding user allocation, bit and power loading in the downlink of MIMO-OFDM systems is proposed. The algorithm represents a generalization of the well-known concepts of LQ decomposition with row pivoting and waterfilling.**

I. Introduction

Link adaptation is known to be an essential component of wideband and broadband OFDM-based systems, such as the recently standardized IEEE 802.16 and the currently developing EUTRAN. The parameters to be adapted are modulation/coding formats and power gains for each sub-carrier or group of sub-carriers. Adaptive user-to-subcarrier assignment has been shown to enable significant power gain in frequency division multiple access (FDMA) systems [1]. If multiple antennas are present, sub-carrier sharing among different users can be accomplished in spatial domain, so implementing space division multiple access (SDMA). In this case each sub-carrier can be considered as a set of spatial layers, and the resource scheduler has to allocate users onto layers. Depending on the spatial techniques used and the quality of the channel state information, different layers of the same sub-carrier may induce different amount of crosstalk onto each other. This has to be taken into account by the resource scheduler.

The problem of link adaptation in MIMO-OFDM systems is still far from being completely solved. The capacity region for MIMO broadcast systems was derived only recently [2]. The capacity of such systems can be achieved by means of dirty paper coding [3], but finding proper signal covariance matrices requires solving a non-convex optimization problem [4], which appears to be a difficult computational task. Another way to solve it is to exploit the duality between broadcast and multiple access channels [4], [5]. In principle, these techniques can be extended to the MIMO-OFDM case, by considering the multi-carrier system as a single-carrier multi-antenna one with a block-diagonal channel matrix. However, this would considerably increase the computational complexity of the associated optimization algorithms. Hence, specialized link adaptation algorithms are needed, which can exploit the properties of OFDM systems. If the number of mobile users does not exceed the number of antennas at the base station, one can assume that zero-forcing preequalization is used, and derive bit and power allocation for each user as a solution of a relatively simple single-user optimization problem [6].

Otherwise, one has to find for each subcarrier the set of active users. Each combination of users requires different amount of transmit power in order to implement zero forcing pre-equalization. This leads to a combinatorial optimization problem described in [7], whose solution requires enormous amount of computational power.

In this paper, we propose a resource scheduling algorithm for the downlink of MIMO-OFDMA sytems, in which SDMA is accomplished through linear or non-linear precoding at the transmitter. The problem of finding a set of users most suitable for each subcarrier is solved using the well-known row pivoting method from computational linear algebra.

The paper is organized as follows. Section II presents an overview of preequalization and related linear algebra concepts. Section III describes a novel link adaptation algorithm. Numeric results are presented in Section IV. Finally, some conclusions are drawn.

II. Zero-forcing precoding in MIMO systems

A. Precoding in downlink MIMO systems

Let us consider a multi-user downlink transmission in a MISO system with K active users, in which the base station (BS) is equipped with M antennas and each user has one antenna. Users having more than one antenna can be modeled as a collection of single-antenna sub-users. Moreover, for the sake of simplicity, let us consider the transmission over a single sub-carrier, so that the index of the sub-carrier can be omitted in this section. The vector r of received symbols is given by $r = Hs + \eta$, where s is the vector of symbols transmitted by the different antennas at the BS, H is the $K \times M$ channel matrix, and η is the vector of additive white Gaussian noise (AWGN) samples. If perfect CSI is available to the transmitter, zero-forcing (ZF) linear or non-linear pre-equalization can be employed. If H is square, linear pre-equalization can be performed by constructing s as $H^{-1}u$, where u is the vector of data symbols to be transmitted. If the channel matrix is factored as $H = GLQ$, where G is a diagonal matrix, L is a lower-triangular matrix with 1's on its main diagonal, and Q is an unitary matrix, this can be rewritten as $s = Q^H L^{-1} G^{-1} u$. Then the received signals can be represented as $r = u + \eta$. From this expression it can be seen that if the matrix G contains small values on its main diagonal, the transmit power may increase considerably. However, all

receivers obtain the transmitted symbols distorted only by AWGN. Alternatively, one can generate the transmitted signals as $s = Q^H L^{-1} u$. In this case, the received signals are given by $r = Gu + \eta$, i.e. different users experience different fading, but are still free of multi-user interference. The channel-to-noise ratios (CNR) for the k-th user can be defined as $\chi_k = \frac{|G_{kk}|^2}{\sigma^2}$. This approach may also cause transmit power boosting depending on the off-diagonal values of the matrix L, which in turn depend on the correlation of the channel matrices of different users. This can be avoided by employing non-linear pre-equalization, also known as Tomlinson-Harashima precoding [8].

B. LQ decomposition with row pivoting

As it was shown above, multi-user interference cancellation can be done via LQ decomposition of the channel matrix. In the literature, Gram-Schmidt orthogonalization is commonly used for this purpose. Despite of its simplicity, however, it is possible to show [9] that the roundoff errors cause severe loss of orthogonality of the matrix Q, i.e. produce a matrix such that $||I - QQ^H|| = \delta$ with δ being proportional to the square of the channel matrix condition number (if it is square), or, after some modifications of the algorithm, to its first degree [10]. In terms of MIMO systems, high condition number of the matrix corresponds to spatially close users.

A more numerically robust method is based on Householder reflections [10]. An $n \times n$ matrix P of the form

$$P = I - \frac{2}{v^H v} v v^H \qquad (1)$$

is called a Householder matrix (Householder reflection), and the vector v is called a Householder vector. If a vector x is multiplied by P, then it is reflected in the hyperplane given by $\mathrm{span}(v)^{\perp}$. By constructing a proper vector v, one can zero some of the components of the product xP.

Householder reflections can be used to construct LQ factorization of a matrix, which can be also combined with row pivoting, i.e. obtain a factorization $\Pi H = LQ$, which recursively maximizes the diagonal entries of the lower-trapezoidal L matrix, where Π is a permutation matrix. The following description of this algorithm illustrates its physical meaning in the context of multi-antenna systems.

The algorithm consists of three sub-routines, which will be used as the building blocks for the adaptive resource allocation algorithm proposed in Section III-B. Function $LQInit$ computes the norms c_j of the rows of the channel matrix H. The norm of the k-th row is proportional to the CNR experienced by the k-th user if all other users are inactive, or if the precoder is used with the permutation matrix Π s.t. the k-th user is assigned to the first spatial layer. Function $LQIter$ performs the assignment of user π_k to layer r. It accordingly permutes the rows of the channel matrix, and applies to it an appropriate Householder transformation, so that entries $r+1..M$ of the r-th row become zero. This can be considered as steering the r-th beam at the BS in the direction of the π_k-th user. As a consequence of the channel matrix update, the norms of its

$\mathrm{LQInit}(H)$
1   **for** $j = 1 : K$
2   **do** $c_j = H_{j,1:M}^H H_{j,1:M}$
3   $\pi = (1, 2, \ldots, K)$
4   **return** (π, c)

$\mathrm{LQIter}(H, c, \pi, k, r)$
1   $swap(\pi(r), \pi(k))$
2   $swap(H_{r,1:M}, H_{k,1:M})$
3   $swap(c_r, c_k)$
4   $x = H_{r,r:M}$
5   $v = x \pm e^{i\theta}||x||_2 e_1,$ where $x_1 = \alpha e^{i\theta}$
6   $v = v/v(1); \beta = \frac{2}{v^H v}$
7   $H_{r:K,r:M} = H_{r:K,r:M}(I - \beta v v^H)$
8   $H_{r,r+1:M} = v_{2:M-r+1}$
9   **for** $i = r + 1 : K$
10 **do** $c_i = c_i - |H_{i,r}|^2$
11 **return** (β, H, c, π)

$\mathrm{HOUSEHOLDERLQ}(H)$
1   $(\pi, c) = LQInit(H)$
2   $r = 0, \tau = \max(c_1, \ldots, c_K)$
3   $t = \arg\max_{t'}(c_{t'})$
4   **while** $\tau > 0$
5   **do** $r = r + 1$
6     $(\beta_r, H, c, \pi) = LQIter(H, c, \pi, t, r)$
7   **if** $r < m$
8   **then** $\tau = \max(c_{r+1}, \ldots, c_K)$
9        $t = \arg\max(c_{r+1}, \ldots, c_K)$
10   **else** $\tau = 0$
11 **return** (H, β, π)

Fig. 1.   Householder LQ factorization with row pivoting

rows yield now the values of the effective CNRs experienced by other users conditioned on that layer assignment. Function $HouseholderLQ$ recursively optimizes the allocation of all layers of a given subcarrier. Based on the CNR values provided either by the function $LQInit$ at the initialization, or by $LQIter$, in further iterations, it assigns a layer to the user experiencing the highest CNR, i.e. it performs row pivoting. It returns a modified matrix H containing matrix L in its lower triangle and Householder vectors in the upper one, vector of constants β, and row permutation π. Matrix Q can be derived from the obtained set of Householder vectors and β.

The norms of matrix L diagonal entries after decomposition (i.e. G_{kk} values) are equal to c_k values. Observe, that at each iteraton of the $HouseholderLQ$ algorithm the row k (i.e. the user) is selected so that c_k is maximized. This is equivalent to finding a set of users with maximal χ_k values, which are the key parameters affecting the system performance. In other words, the $HouseholderLQ$ function finds the users which are as orthogonal to each other as possible.

III. LINK ADAPTATION IN PRECODED OFDM SYSTEMS

This section presents a method for user allocation, bit and power loading in multi-user MIMO-OFDM downlink systems. On each spatial layer, the method applies the well-known waterfilling approach [11], which is briefly reviewed in section

III-A. For the user to layer allocation, an extention of the Householder based algorithm is proposed in section III-B.

A. Waterfilling

Let us consider a single user system transmitting data over the vector Gaussian channel consisting of n subchannels, with channel-to-noise ratios $\chi_1, \ldots, \chi_n$. Let $f(c) = \Gamma(2^c - 1)$ be a function specifying the SNR needed to maintain data rate c on a subchannel, where Γ is the capacity gap of the coding/modulation scheme being used, at some fixed error probability [12]. Then the transmit power gain needed to achieve data rate c_i on subchannel i can be computed as $V_i^2 = \Gamma(2^{c_i} - 1)/\chi_i, i = 1..n$. For the sake of simplicity, let c_i be continuous variables. Consider the problem of finding data rate allocation for individual subchannels minimizing the total transmit power $\mathcal{P} = \sum_{i=1}^{n} V_i^2$ under the constraint of fixed data rate $R = \sum_{i=1}^{n} c_i$. The Lagrangian for this problem is

$$L(c_1, \ldots, c_n, \lambda) = \sum_{i=1}^{n} \frac{\Gamma(2^{c_i} - 1)}{\chi_i} - \lambda \left(\sum_{i=1}^{n} c_i - R \right), \quad (2)$$

where λ is the Lagrangian multiplier. Simple derivations lead to the following equations relating c_i, R and $\mathcal{P}$:

$$c_i = \log_2 \left(\frac{\lambda \chi_i}{\Gamma \ln 2} \right)$$

$$R = n \log_2 \left(\frac{\lambda}{\Gamma \ln 2} \right) + \log_2 \left(\prod_{j=1}^{n} \chi_j \right)$$

$$\mathcal{P} = \frac{\lambda n}{\ln 2} - \Gamma \sum_{i=1}^{n} \frac{1}{\chi_i}$$

Depending on the value of λ, these equations may lead to negative c_i values for some subchannels. These subchannels should not carry any data, and the corresponding χ_i values should be excluded from consideration in order to get correct results for other subchannels, i.e. n has to be optimized. Assuming that this has been done, the optimal values of c_i and the corresponding transmit power can be computed as

$$c_i = c_i(\chi_i, \chi) = \frac{R}{n} + \log_2 \left(\frac{\chi_i}{\sqrt[n]{\chi}} \right) \quad (3)$$

$$\mathcal{P} = P(\chi, \{\chi_1, \ldots, \chi_n\}) = \frac{n \Gamma 2^{R/n}}{\sqrt[n]{\chi}} - \Gamma \sum_{i=1}^{n} \frac{1}{\chi_i}, \quad (4)$$

where $\chi = \prod_{j=1}^{n} \chi_j$. These closed-form expressions provide one with an efficient method to determine if assigning a particular subchannel to a given user would improve the performance. Namely, let us suppose that a user is given one more subchannel with CNR χ'. This subchannel is assigned a coding/modulation scheme with rate

$$c' = \mathcal{C}(R, n, \chi, \chi') = \frac{R}{n+1} + \log_2 \left(\frac{\chi'}{\sqrt[n+1]{\chi \cdot \chi'}} \right). \quad (5)$$

If this value appears to be negative, the subchannel should not be used, and one should consider different ways to utilize it,

e.g. give it to another user. If this value is positive, one can estimate the decrease of required transmit power as

$$\mathcal{P}^-(R, n, \chi, \chi') = \frac{n \Gamma 2^{R/n}}{\sqrt[n]{\chi}} - \frac{(n+1) \Gamma 2^{R/(n+1)}}{\sqrt[n+1]{\chi \cdot \chi'}} + \frac{\Gamma}{\chi'}. \quad (6)$$

This expression does not take into account the possibility of other subchannels becoming inactive after adding one more subchannel. Obviously, such subchannels must have $\chi_i < \chi'$. This can be avoided by adding subchannels in the descending order of their CNRs. The power decrease should be compared with the cost of subchannel allocation. This cost may have different nature. For example, if the system being considered is a part of a multi-user system, this can be the increase of power due to revoking this subchannel from another user. It can be estimated as

$$\mathcal{P}^+(R, n, \chi, \chi') = -\frac{n \Gamma 2^{R/n}}{\sqrt[n]{\chi}} + \frac{(n-1) \Gamma 2^{R/(n-1)}}{\sqrt[n-1]{\chi/\chi'}} + \frac{\Gamma}{\chi'}. \quad (7)$$

This allows one to optimize the subchannel-to-user asignment. These derivations may produce non-integer c_i values, which can be corrected by employing any single-user bit and power loading algorithm [13], [12] after the user-to-subchannel allocation has been performed.

B. Adaptive LQ decomposition

Let us consider the problem of adaptive user-to-subcarrier allocation, bit and power loading in the downlink of a precoded MIMO OFDM system with N sub-carriers. Let $H^{(i)}$ be the combined $SK \times M$ channel matrix for subcarrier i, where S is the number of antennas at each mobile terminal. For the sake of simplicity, the derivations below will be presented for the case $S = 1$, and later extended to the more general case. The optimization problem considered here is that of finding $u_{i,j}$, the user to be assigned to the j-th layer, $j = 1..M$, of the i-th sub-carrier, $i = 1..N$, and the coding/modulation formats $c_{i,j}$ minimizing the total transmit power while maintaining the individual user data rates $R_k, k = 1..K$, i.e.

$$\min_{u_{i,j}, c_{i,j}} \sum_{i=1}^{N} \sum_{j=1}^{M} \Gamma \frac{2^{c_{i,j}} - 1}{\chi_{i,j}} \quad (8)$$

under the constraints

$$u_{i,j_1} = u_{i,j_2} \Leftrightarrow j_1 = j_2$$
$$\sum_{i,j : u_{i,j} = k} c_{i,j} = R_k, k = 1..K.$$

The first constraint ensures that no user can employ more than one layer of one subcarrier. Observe also that the values $\chi_{i,j}$ do depend on $u_{i,1}, \ldots, u_{i,j}$. It must be recognized that (8) is not exactly equal to the actual transmit power. In the case of linear pre-equalization there may be power boosting due to off-diagonal entries of L matrices for each subcarrier. However, since link adaptation effectively reduces to transmission over the best part of the channel, these inaccuracies should not adversarily affect the optimization algorithm.

The main difficulty in designing a multi-user link adaptation algorithm for multi-carrier multi-antenna systems is finding a proper subset of users to be assigned onto each sub-carrier. If the number of antennas is equal to the total number of users in the system, i.e. $H^{(i)}$ is square, the total system capacity depends on the determinant of $H^{(i)}$, which does not depend on user-to-layer assignment. But if there are more users than antennas, there are $\binom{K}{M}$ different ways to select a subset of active users for each subcarrier, and these subsets correspond to different sub-matrices of $H^{(i)}$ having different determinants. Furthermore, for each subset of active users there are $M!$ ways to perform actual user-to-layer assignment. Hence, the total search space for each sub-carrier consists of $\frac{K!}{(K-M)!}$ possible configurations. Straightforward solution of the associated optimization problem results in combinatorial explosion of its dimension and complexity [7].

In Section II-B, the LQ decomposition with row pivoting was applied to find the sequence of users best suited for transmission over a given sub-carrier. The algorithm requires selecting at each step the user with the highest gain factor. Instead of performing gain factor maximization independently on each subcarrier, we now jointly consider all subcarriers and data rate constraints of all users. This is implemented by integrating the pivoted LQ decomposition algorithm with the waterfilling principle. Figure 2 illustrates the algorithm. It is composed of an initialization phase (lines 3–5), in which it constructs some initial user and data rate allocation according to some optimization criterion, and of an iteration phase (lines 9–28), where it gradually improves it using the reallocation benefit expressions (6) and (7). This can be considered as a generalization of steps 8 and 9 of the $HouseholderLQ$ algorithm. As soon as optimal assignment of all users for the current layer is determined, $LQIter$ function explained in Section II-B is called for each subcarrier in order to carry out the actual reallocations and derive their impact on the performance of subsequent layers.

The algorithm takes the following inputs: K, the number of users, N, the number of subcarriers, M, the number of base station antennas, T, the number of reallocation attempts, $H^{(i)}$, the subcarrier matrices, and R_k, the user data rate constraints. Similarly to pivoted LQ decomposition, it determines t_i, leading elements for each sub-carrier i, for each layer $j = 1..M$. This algorithm assumes that the number of users K is greater than the number of antennas at the base station. Hence, the $WHILE(\tau > 0)$ loop is replaced with a loop over all base station antennas. The algorithm maintains lists S_k of subcarrier layers assigned to each user. Each subcarrier layer is characterized by layer number j, subcarrier number i, and channel-to-noise ratio χ, represented as a triple (j, i, χ). Recall, that the latter parameter depends on the user allocation.

The algorithm starts from constructing some initial allocation of users onto the first layer of all sub-carriers. In the simplest case, which resembles classical LQ factorization, the first layer of each subcarrier is assigned to the user with the highest CNR. In the iteration phase, the processing of each layer starts with conventional single-user bit and power

$\text{ALLOCATE}(K, N, M, T, (H^{(1)}, \dots, H^{(N)}), (R_1, \dots, R_K))$
1   $S_k = \emptyset, k = 1..K$
2   $\textbf{for } i = 1 : N$
3   $\textbf{do } (\pi^{(i)}, \chi^{(i)}) = LQInit(H^{(i)})$
4      $k' = \arg\max_k(\chi_k^{(i)}); u_{i,1} = k'; t_i = k'$
5      $S_{k'} = S_{k'} + (1, i, \chi_{k'}^{(i)})$
6   $\textbf{for } j = 1 : M$
7   $\textbf{do } (\chi_k, C_k, n_k) = BITLOADING(S_k, R_k), k = 1..K$
8      $\Delta = 1; l = 1$
9      $\textbf{while } (\Delta > 0) \wedge (l < T)$
10     $\textbf{do } \Delta = 0; l = l + 1$
11        $\textbf{for } i = 1 : N$
12        $\textbf{do } k = t_i$
13           $\textbf{if } C_k[i, j] > 0$
14            $\textbf{then } \Delta_1 = \mathcal{P}^+(R_k, n_k, \chi_k, \chi_{t_i}^{(i)})$
15            $\textbf{else } \Delta_1 = 0$
16           $\textbf{for } s = j + 1 : K$
17           $\textbf{do } k' = \pi_s^{(i)}$
18            $\Delta_2 = \mathcal{P}^-(R_{k'}, n_{k'}, \chi_{k'}, \chi_s^{(i)})$
19            $c' = \mathcal{C}(R_{k'}, n_{k'}, \chi_{k'}, \chi_s^{(i)})$
20            $\textbf{if } (c' > 0) \wedge (\Delta_2 - \Delta_1 > \Delta)$
21              $\textbf{then } i_0 = i; \tilde{t} = s; \hat{t} = t_i; \Delta = \Delta_2 - \Delta_1$
22        $\textbf{if } \Delta > 0$
23           $\textbf{then } \hat{k} = \pi_{\hat{t}}^{(i_0)}; \tilde{k} = \pi_{\tilde{t}}^{(i_0)}$
24            $S_{\hat{k}} = S_{\hat{k}} - (j, i_0, \chi_{\hat{t}}^{(i)})$
25            $S_{\tilde{k}} = S_{\tilde{k}} + (j, i_0, \chi_{\tilde{t}}^{(i)})$
26            $u_{i_0,j} = \tilde{k}; t_{i_0} = \tilde{t}$
27            $(\chi_{\hat{k}}, C_{\hat{k}}) = BITLOADING(S_{\hat{k}}, R_{\hat{k}})$
28            $(\chi_{\tilde{k}}, C_{\tilde{k}}) = BITLOADING(S_{\tilde{k}}, R_{\tilde{k}})$
29      $\textbf{for } i = 1 : N$
30      $\textbf{do } (\beta^{(i)}, H^{(i)}, \chi^{(i)}, \pi^{(i)}) = LQIter(H^{(i)}, \chi^{(i)}, \pi^{(i)}, t_i, j)$
31        $\textbf{if } j < M$
32           $\textbf{then } t_i = \arg\max_{j'>j}(\chi_{j'}^{(i)})$
33            $S_{k'} = S_{k'} + \{(j + 1, i, \chi_{t_i}^{(i)})\}$
34            $u_{i,j+1} = \pi^{(i)}(t_i)$
35 $\textbf{return } (u_{ij}, C_k, \beta^{(i)}, H^{(i)})$

Fig. 2.   Subcarrier allocation combined with LQ decomposition

loading based on the existing user allocation. The input to the $BITLOADING$ function is the list S_k of subcarrier layers assigned to a user, and the corresponding data rate constraint R_k. This function computes n_k, the number of subcarrier layers being actually used for data transmission, the corresponding coding/modulation rates C_k and the product χ_k of the active layer CNRs. Having determined the initial user and bit allocation, the algorithm attempts then to improve it. For each subcarrier the algorithm considers revoking it from the currently assigned user and giving it to another user. The benefits of such reallocation are estimated using the expressions (5), (7) and (6). The decision is based on R_k, the target user data rate, n_k, the number of subcarrier layers being actually used, $\chi_k^{(i)}$, the CNR of the current leading element k for the i-th sub-carrier, as well as on $\chi_{k'}^{(i)}$, the CNR of the same layer *if it would be assigned to user k'*. Namely, for each pair (k, k') of users the increase of required transmit power Δ_1 caused by revoking sub-carrier layer (i, j) from user k is computed, as well as the decrease of power Δ_2 due to giving

it to user k'. These values are computed using the expressions obtained for the case of continuous rate waterfilling, while the algorithm should produce a discretized output. This introduces some inaccuracies, which are unavoidable in order to obtain an algorithm with practical complexity. Furthermore, the coding/modulation rate c' for this subcarrier layer *if it would be assigned to user k'* is computed. If $c' \leq 0$, performing such reassignment is useless. Reallocation is performed for the subcarrier i_0 and pair of users $(\hat{k}, \tilde{k})$ maximizing the difference $\Delta_2 - \Delta_1$, i.e. providing the highest reduction of the total required transmit power. This is followed by re-evaluation of bit allocation for these two users. These steps are performed until either no progress is observed, or the specified number of iterations T is exceeded. This results in a vector $(t_1, \ldots, t_N)$ of optimized leading elements, or the best users $\tilde{k}$ for layer j of each subcarrier, which are given as input to the function $LQIterate$. This computes the values of the effective CNRs for all users on subsequent layers, which are again used to select the best user for the next layer of each subcarrier.

Despite of the apparent complexity of the algorithm, the underlying idea is extremely simple. The algorithm assigns subcarrier layers to users in a greedy way, and performs reallocation using the waterfilling identities. However, it is not guaranteed to find the optimal allocation, since after processing of j layers some subchannels corresponding to previous layers may become unused. In this case it might be better to reallocate them, but this would trigger a lot of other reallocations. Computer experiments indicate that such events are quite rare in practice, so they are ignored in order to keep the complexity low. The complexity can be further reduced by reducing the number of subcarriers which are considered for possible reallocation.

The convergence speed of this algorithm does depend on the initial layer allocation. It appears that assigning a certain sub-carrier layer to the user that experiences the maximum CNR over it may lead to unfair configuration, i.e. some users may get a lot of subcarriers, while others just a few ones. This would require a lot of reallocations in order to achieve a reasonably good configuration. The number of iterations needed for convergence can be substantially reduced by constructing a fair initial allocation using the following approximate approach, which is again based on waterfilling identities. Approximating the geometric average of $1/\chi_i$ values with their arithmetic average, (4) can be rewritten as

$$\mathcal{P}(n) \approx \Gamma \left(2^{\frac{R}{n}} - 1\right) \sum_{i=1}^{n} \frac{1}{\chi_i}.$$

The total required transmit power of a multi-user system is given as a sum of $\mathcal{P}(n)$ values of individual users, ignoring the precoding power boost. Since one is interested in relative values of $\mathcal{P}(n)$, the factor Γ can be dropped. Assigning to the user an additional subcarrier with CNR χ' would reduce the
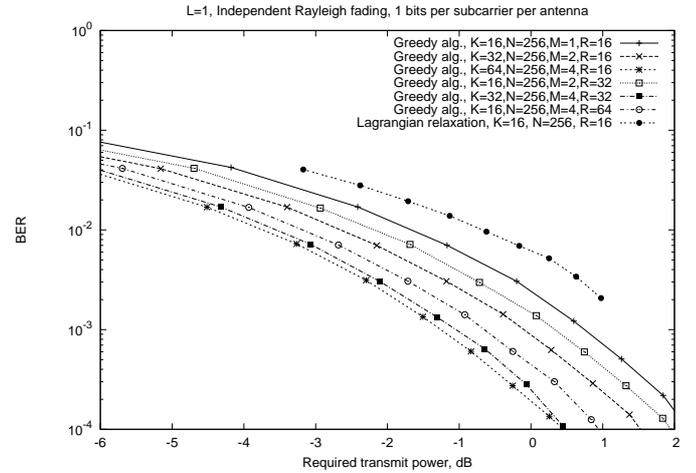

Fig. 3. Performance of systems in IR fading environment

required transmit power by

$$
\begin{aligned}
\Delta(n, \chi') &= \left(2^{\frac{R}{n}} - 1\right) \sum_{i=1}^{n} \frac{1}{\chi_i} - \left(2^{\frac{R}{n+1}} - 1\right) \left(\frac{1}{\chi'} + \sum_{i=1}^{n} \frac{1}{\chi_i}\right) \\
&= \sum_{i=1}^{n} \frac{1}{\chi_i} \left(2^{\frac{R}{n}} - 2^{\frac{R}{n+1}}\right) - \frac{2^{\frac{R}{n+1}} - 1}{\chi'}.
\end{aligned}
\tag{9}
$$

This allows one to perform layer initialization as follows:

1) Calculate for each user $\Delta(n_k, \chi_k^{(i)})$, the decrease of its required transmit power due to assigning him the current layer of the i-th subcarrier, using the expression (9).
2) Assign the current layer of the i-th subcarrier to the user $k' = \arg\max_k \Delta(n_k, \chi_k^{(i)})$.

This ensures fairness of the initial subcarrier allocation, simplifying thus the job of further optimization.

The presented approach can be extended to the case of $S > 1$ by considering each user as a collection of S subusers. S subuser data rate constraints should be replaced with a single constraint. Reallocation benefit expressions of the algorithm should be modified in order to take into account possible reallocation between two subusers corresponding to the same user. That is, the power gain of replacing the subchannel with CNR χ' with the one having CNR χ'' should be calculated. Observe that the proposed approach does not limit the maximal number of users utilizing the same subcarrier to M/S, as opposed to e.g. [7].

IV. NUMERIC RESULTS

Figure 3 illustrates the performance of the THP-based adaptive system in the independent Rayleigh fading channel. Simulations were performed for different values of K, M and R_k, so that the total data rate equals 1 bit per subcarrier per antenna. For $M = 1$ the performance of the algorithm based on Lagrangian relaxation [14] is also reported, which was shown to outperform the one presented in [1] even in the pure OFDMA case [15]. In all considered cases the actual transmit
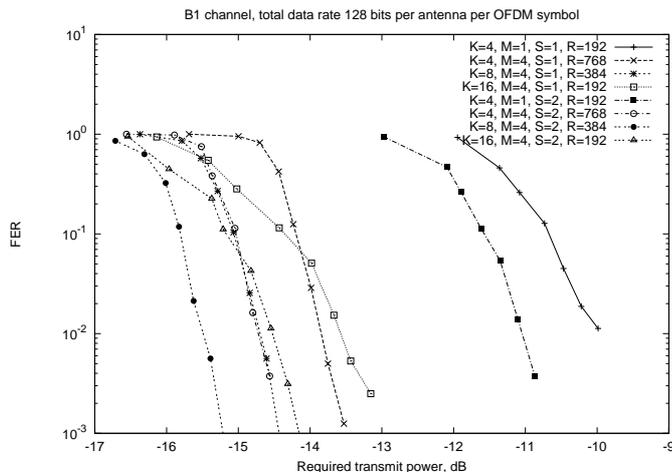
B1 channel, total data rate 128 bits per antenna per OFDM symbol

K=4, M=1, S=1, R=192	—+—
K=4, M=4, S=1, R=768	--×--
K=8, M=4, S=1, R=384	··*··
K=16, M=4, S=1, R=192	·····□·····
K=4, M=1, S=2, R=192	--■--
K=4, M=4, S=2, R=768	--○-·
K=8, M=4, S=2, R=384	--●-··
K=16, M=4, S=2, R=192	··△··

Fig. 4.   Adaptive semi-BICM system performance

power was recorded, i.e. the presented results take into account possible power boosting due to linear pre-equalization.

It can be seen that for $M = 1$ the proposed greedy algorithm outperforms the one based on Lagrangian relaxation. The reason is that the latter one suffers from inaccuracies introduced by the discreteness of the underying optimization problem. The method proposed in this paper is similar to genetic algorithms, which are much more suitable for such problems. Furthermore, increasing the number of base station antennas allows one either to increase the number of users, or their data rate. In both cases the required transmit power decreases. In the first case the gain appears to be 1 dB more than in the second one. Depending on the system parameters, 50–100 iterations were required for convergence of the algorithm.

Figure 4 presents the performance of a system implementing semi-BICM adaptive coded modulation scheme [16]. The single-user link adaptation algorithm proposed there was employed as $BITLOADING$ subroutine in the proposed algorithm. The systems is assumed to have 2048 subcarriers. The time-frequency resource units are grouped into chunks consisting of 16 adjacent subcarriers and 6 OFDM symbols, employing the same modulation format. SCME B1 channel model [17] with 100 MHz bandwidth was used. As it may be expected, increasing the number S of antennas at the mobile terminal always improves the performance. Furthermore, simultaneousely increasing the number of antennas at the base station M and the user data rate R provides up to 3.5 dB performance gain. This is due to spatial diversity gain combined with the gain of longer modulation codes being used. Alternatively, higher number of antennas at BS allows one to accomodate more users. In this case the performance gain achieves 4 dB due to increased spatial diversity.

V. CONCLUSIONS

In this paper a novel adaptive resource allocation algorithm has been proposed for the downlink of OFDMA/SDMA systems in which SDMA is based on linear precoding. The algorithm can be considered as an extension of the pivoted LQ decomposition algorithm widely used in computational linear algebra. The proposed algorithm avoids combinatorial explosion of the optimization complexity, suffered by the method in [7], and can be used in systems with the number of users exceeding the number of antennas, as opposed to the algorithm in [6].

REFERENCES

[1] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, October 1999.

[2] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions On Information Theory*, vol. 52, September 2006.

[3] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439–441, May 1983.

[4] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian mimo broadcast channels," *IEEE Transactions On Information Theory*, vol. 49, October 2003.

[5] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. J. Goldsmith, "Sum power iterative water-filling for multi-antenna gaussian broadcast channels," *IEEE Transactions On Information Theory*, vol. 51, April 2005.

[6] S. Thoen, L. V. der Perre, M. Engels, and H. D. Man, "Adaptive loading for ofdm/sdma-basedwireless networks," *IEEE Transactions On Communications*, vol. 50, November 2002.

[7] I.-S. Park, Y.-H. Jung, I. Min, and Y. Lee, "Dynamic subchannel and bit allocation in multiuser MIMO/OFDMA systems," in *Proceedings of IEEE VTC'2004-Spring*, vol. 2, pp. 884– 888, 2004.

[8] C. Windpassinger, R. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Transactions On Wireless Communications*, vol. 3, July 2004.

[9] L. Giraud, J. Langou, M. Rozloznik, and J. van den Eshof, "Rounding error analysis of the classical gram-schmidt orthogonalization process," *Numerische Mathematik*, vol. 101, pp. 87–100, July 2005.

[10] G. Golub and C. van Loan, *Matrix computations*. The Johns Hopkins University Press, 3rd ed., 1996.

[11] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[12] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Transactions On Communications*, vol. 43, February/March/April 1995.

[13] B. Krongold, K. Ramchandran, and D. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," *IEEE Transactions on Communications*, vol. 48, pp. 23–27, January 2000.

[14] P. Trifonov, E. Costa, and E. Schulz, "Adaptive user allocation, bit and power loading in multi-carrier systems," in *Proceedings of the 9th International OFDM-Workshop*, 2004.

[15] P. Trifonov, *Adaptive data transmission in multi-carrier systems*. PhD thesis, Saint-Petersburg State Polytechnic University, May 2005. in Russian.

[16] P. Trifonov, E. Costa, and E. Schulz, "Joint adaptation of code length and modulation formats in OFDM systems," in *Proceedings of the 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2006.

[17] D. S. Baum, J. Hansen, G. D. Galdo, M. Milojevic, J. Salo, and P. Kyösti, "An interim channel model for beyond-3G systems: Extending the 3GPP spatial channel model (SCM)," in *Proceedings of VTC'05-Spring*, 2005.