Towards a joint optimization of scheduling and beamforming for MIMO downlink

G. Primolevo[1], O. Simeone[2] and U. Spagnolini[1]
(1) Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milano, Italy
(2) CWCSPR, ECE Dept., NJIT, Newark, NJ 07102, USA

***Abstract*— In the downlink of a multi-user MIMO system, the base station can multiplex signals intended to different users on the same spectral resource. Joint optimization of channel assignment (scheduling) and beamforming in such a scenario is an open problem. Assuming zero-focing linear beamforming at the base station, this paper addresses this task by investigating a novel greedy solution that approximately maximizes the sum-rate. Performance of the proposed method is studied through numerical simulation and compared with known solutions based on combinatorial search.**

I. Introduction

In the downlink of a multi-user system, deployment of an antenna array at the base station (or access point) opens up the possibility to multiplex signals intended for different users in the same time/frequency resource. If the transmitter is equipped with N_T antenna and employs optimal non-linear interference pre-subtraction (Dirty Paper Coding [1]), it has been recently shown that up to N_T^2 spatial streams can be multiplexed and up to N_R^2 reserved to each user [2] (N_R is the number of receiving antennas). This contrasts with the case of base station and users equipped with a single antenna, where it has been shown that transmission to the user with the strongest channel is a strategy that achieves channel capacity [5].

Due to the large complexity of Dirty Paper Coding, sub-optimal techniques have been investigated for the downlink of a multiantenna system. In particular, (linear) zero-forcing beamforming has been studied in [4] for the case $N_R = 1$. With zero-forcing techniques, only up to N_T data streams can be spatially multiplexed. Zero-forcing beamforming for a multiuser MIMO scenario (i.e., with $N_R > 1$) has been investigated in [6] with a simplified channel allocation (scheduling) over the users. In this latter case, joint optimization of scheduling and beamforming is known to be a challenging task due to the large set of possible transmitting strategies that is necessary to explore. In [8] this issue was tackled by maximizing a performance criterion related to the sum-rate, for a given number of data streams assigned to each user. Obtaining the optimal solution then requires a combinatorial search over the possible channel assignments.

In this paper, a novel approximate greedy solution to the problem of maximizing the sum-rate is proposed, that circumvents the need of the above-mentioned combinatorial search with moderate performance loss. Performance of the proposed methods is studied through numerical simulation and compared to existing solutions and combinatorial search approaches, showing the desirable features of the presented techniques. For further analysis of the proposed scheme in terms of fairness criteria, the reader is referred to [10].

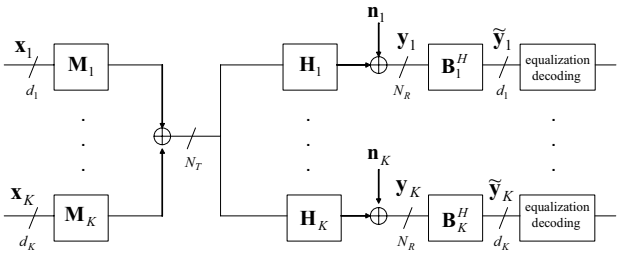

Fig. 1. Block diagram of a broadcast channel with linear interfaces at the transmitter (base station) and receivers (users).

II. Signal model

A. Channel aware scheduling and linear precoding

A MIMO broadcast channel with beamforming matrices at the transmitter and receivers is depicted in fig. 1. Let $\bar{\mathcal{K}}$ be the set of $\bar{K}$ available users. The base station is equipped with an antenna array of N_T elements, whereas each user has N_R antennas. The subset of $K(t)$ users that are served by the base station within the tth time slot is denoted as $\mathcal{K}(t) \subseteq \bar{\mathcal{K}}$ and its elements are indexed by $k = 1, 2, ..., K(t)$. The scheduler distributes the N_T available spatial channels to the users in $\mathcal{K}(t)$ based on the knowledge of the channel state information at the transmitter (*channel aware scheduling*). In particular, the scheduler allocates $d_k(t) \leq N_R$ spatial channels to the kth user so that all the N_T available spatial channels are used[1]:

$$\sum_{k=1}^{K(t)} d_k(t) = N_T. \tag{1}$$

The signal intended for the kth user is collected into the $d_k(t) \times 1$ zero-mean complex Gaussian vector $\mathbf{x}_k(t)$ with $E[\mathbf{x}_k(t)\mathbf{x}_k^H(t)] = \mathbf{I}_{d_k(t)}$ and is linearly precoded by the

[1]Notice that in certain scenarios the equality constraint (1) might lead to a degraded solution as compared to the case where a smaller number of spatial streams is assigned (i.e., $\sum_{k=1}^{K(t)} d_k(t) < N_T$). However, throughout the paper, we assume that the scheduler assigns all the N_T streams as in (1) in order to simplify the presentation.

0-7803-9780-0/06/$20.00 ©2006 IEEE

$N_T \times d_k(t)$ matrix $\mathbf{M}_k(t)$. Following the conventional notation for block fading channels (see, e.g., [3]) and referring to fig. 1, the signal received by the kth user across its N_R receiving antennas within the tth time slot can be written as the $N_R \times 1$ vector $\mathbf{y}_k(t)$

$$\mathbf{y}_k(t) = \mathbf{H}_k(t)\mathbf{M}_k(t)\mathbf{x}_k(t) + \sum_{\substack{i \neq k \\ i \in \mathcal{K}(t)}} \mathbf{H}_k(t)\mathbf{M}_i(t)\mathbf{x}_i(t) + \mathbf{n}_k(t), \tag{2}$$

where $\mathbf{H}_k(t)$ is the $N_R \times N_T$ channel matrix of the kth user and the zero mean additive complex Gaussian noise is $\mathbf{n}_k(t) \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_R})$. The channel matrices $\mathbf{H}_k(t)$ are assumed to be instantaneously known at the transmitter, e.g., through feedback channels. Moreover, they are assumed to be constant in each transmission slot and varying independently across different slots (ergodic block fading).

The received signal $\mathbf{y}_k(t)$ lies in a N_R-dimensional linear space. However, only $d_k(t) \leq N_R$ spatial channels are assigned to the kth user by the scheduler. Therefore, the useful part of the received signal spans a $d_k(t)$-dimensional subspace $\mathrm{span}\{\mathbf{H}_k(t)\mathbf{M}_k(t)\}$ that we refer to as *receiving subspace*. The receiving subspace can be described as the subspace spanned by a $N_R \times d_k(t)$ orthonormal matrix $\mathbf{B}_k(t)$

$$\mathbf{B}_k^H(t)\mathbf{B}_k(t) = \mathbf{I}_{d_k(t)}, \tag{3}$$

i.e., $\mathrm{span}(\mathbf{B}_k(t))$ is the receiving subspace for user k. Now, the goal of the scheduler can be described as that of selecting at each time-slot t the set $\mathcal{B}(t) = \{\mathbf{B}_k(t)\}_{k=1}^{\bar{K}}$, where $\mathbf{B}_k(t) = \mathbf{0}$ if the kth user does not belong to the set of active users.

According to the discussion above, at the kth terminal, sufficient statistics for the estimation of the $d_k(t) \times 1$ transmitted vector $\mathbf{x}_k(t)$ can be obtained by projecting the received signal $\mathbf{y}_k(t)$ within the receiving subspace by the $d_k(t) \times N_R$ matrix $\mathbf{B}_k^H(t)$

$$\begin{aligned} \tilde{\mathbf{y}}_k(t) &= \mathbf{B}_k^H(t)\mathbf{y}_k(t) = \tilde{\mathbf{H}}_k(t)\mathbf{M}_k(t)\mathbf{x}_k(t) + \\ &+ \sum_{\substack{i \neq k \\ i \in \mathcal{K}(t)}} \tilde{\mathbf{H}}_k(t)\mathbf{M}_i(t)\mathbf{x}_i(t) + \tilde{\mathbf{n}}_k(t), \end{aligned} \tag{4}$$

where we have defined the $d_k(t) \times N_T$ equivalent channel $\tilde{\mathbf{H}}_k(t) = \mathbf{B}_k^H(t)\mathbf{H}_k(t)$ and the noise $\tilde{\mathbf{n}}_k(t) = \mathbf{B}_k^H(t)\mathbf{n}_k(t)$. Notice that from (3) the noise correlation matrix reads $E[\tilde{\mathbf{n}}_k(t)\tilde{\mathbf{n}}_k^H(t)] = \sigma^2 \mathbf{I}_{d_k(t)}$. In this paper, we consider the problem of joint designing of channel aware scheduling (i.e., design of the set $\mathcal{B}(t) = \{\mathbf{B}_k(t)\}_{k=1}^{\bar{K}}$) and precoding (i.e., design of the set $\mathcal{M}(t) = \{\mathbf{M}_k(t)\}_{k=1}^{\bar{K}}$).

B. Zero-forcing linear precoding

In this work, linear precoding is assumed to be designed according to the zero-forcing criterion. Therefore, precoding matrices $\mathcal{M}(t) = \{\mathbf{M}_k(t)\}_{k=1}^{\bar{K}}$ are selected so that the following condition is guaranteed:

$$\tilde{\mathbf{H}}_i(t)\mathbf{M}_j(t) = \mathbf{0} \quad \text{for } i \neq j, \tag{5}$$

which implies that the received signal (4) can be written as

$$\tilde{\mathbf{y}}_k(t) = \tilde{\mathbf{H}}_k(t)\mathbf{M}_k(t)\mathbf{x}_k(t) + \tilde{\mathbf{n}}_k(t). \tag{6}$$

In other words, zero-forcing linear precoding equivalently forms a set of parallel MIMO channels free from inter-user interference.

III. Maximization of the sum-rate

In this Section, we address the problem of joint channel aware scheduling and zero-forcing precoding optimization aimed at maximizing the sum-rate in each time-slot. The argument t is dropped in the following to simplify the notation. The problem can be formulated as:

$$\begin{aligned} \{\mathcal{B}, \mathcal{M}\} &= \underset{\mathcal{B},\mathcal{M}}{\mathrm{argmax}} \sum_{i=1}^{\bar{K}} C_i(\mathcal{B}, \mathcal{M}) \\ &\text{s.t.} \begin{cases} \text{(1), (3), (5)} \\ \sum_{i=1}^{\bar{K}} \mathrm{tr}(\mathbf{M}_i\mathbf{M}_i^H) \leq P \end{cases} \end{aligned} \tag{7}$$

where $C_i(\mathcal{B}, \mathcal{M})$ is the link rate for the ith user [3]

$$C_i(\mathcal{B}, \mathcal{M}) = \log_2 |\mathbf{I}_{d_i} + \frac{1}{\sigma^2}\tilde{\mathbf{H}}_i\mathbf{M}_i\mathbf{M}_i^H\tilde{\mathbf{H}}_i^H| \tag{8}$$

and the last constraint in (7) limits the total transmitted power at the base station in each time slot. Note that the rate for each user (8) does not account for interference because of the zero-forcing constraint (5). In the remaining of this section, we review existing approaches to the solution of problem (7). This discussion sets the ground for the presentation of a novel technique in Sec. IV, that will be proved to have desirable performance as compared to known solutions in Sec. V by numerical results.

A. Review of existing approaches to the solution of (7)

1) Separate optimization of zero-forcing precoding and scheduling (LSV) [6]: In [6], problem (7) is tackled by decoupling the optimization of precoding $\mathcal{M}$ and scheduling $\mathcal{B}$. Let us at first consider the optimization of $\mathcal{M}$ for a *given* scheduling $\mathcal{B}$. As a consequence of the constraint (5), any precoding matrix can be written as

$$\mathbf{M}_k = \bar{\mathbf{V}}_k^{\perp}\mathbf{Q}_k, \tag{9}$$

where $\bar{\mathbf{V}}_k^{\perp}$ is a $N_T \times d_k$ matrix with orthonormal columns selected so as to null the inter-user interference as in (5), i.e., (see Appendix-A for an explicit computation of $\bar{\mathbf{V}}_j^{\perp}$)

$$\tilde{\mathbf{H}}_i\bar{\mathbf{V}}_j^{\perp} = \mathbf{0} \quad \text{for} \quad i \neq j \tag{10}$$

and $\mathbf{Q}_k$ is a $d_k \times d_k$ matrix that performs beamforming and power allocation on the interference-free single-user MIMO channels (6) created by zero-forcing precoding. In particular, from (6) and (10) we have

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{H}}_k\bar{\mathbf{V}}_k^{\perp}\mathbf{Q}_k\mathbf{x}_k + \tilde{\mathbf{n}}_k, \tag{11}$$

so that we can use the well known results on single-user MIMO channels in order to design matrix $\mathbf{Q}_k$ based on the singular value decomposition (SVD) of matrix $\tilde{\mathbf{H}}_k\bar{\mathbf{V}}_k^{\perp}$ [3].

If the condition on the transmit antennas $N_T \geq \bar{K}N_R$ is satisfied, there is no need for the scheduling step since all the $\breve{K}N_R$ deployable spatial channels can be allocated. In this

case, optimality with respect to (7) is guaranteed by setting the number of spatial channels for each user to $d_k = N_R$ and $\mathbf{B}_k = \mathbf{I}_{N_R}$, and by computing the precoding matrices as explained above. However, in general, the number of transmit antennas is not large enough ($N_T < \bar{K}N_R$).

In this case, [6] proposes to set as active the spatial channels corresponding to the largest singular values of the channel matrices $\{\mathbf{H}_i\}_{i=1}^{\bar{K}}$. Therefore, scheduling can be equivalently stated as the solution of the following optimization problem: find the set $\mathcal{B}$ so that

$$\mathcal{B} = \underset{\mathcal{B}}{\operatorname{argmax}} \sum_{i=1}^{\bar{K}} \left\| \mathbf{B}_i^H \mathbf{H}_i \right\|^2, \quad \text{s.t. (1) and (3).} \tag{12}$$

We will refer to this technique that performs separate optimization of precoding and scheduling as the Largest Singular Value (LSV) method.

2) Approximate joint optimization of scheduling and precoding with given $\{d_i\}_{i=1}^{\bar{K}}$ (GM-OSDMA)[8]: The approximate solution of the problem (7) proposed by [6] and reviewed in the previous section suffers from degraded performance (as it will be shown in Sec. V) mainly because the precoding matrices $\mathcal{M}$ and the scheduling matrices $\mathcal{B}$ are optimized separately. In [8] a technique is propose that partially circumvents this inconvenient. The approach is based on the approximation of the objective function $C_i(\mathcal{B}, \mathcal{M})$ by its first term of the Taylor expansion: $C_i(\mathcal{B}, \mathcal{M}) \approx \left\| \mathbf{B}_i^H \mathbf{H}_i \mathbf{M}_i \right\|^2 / \sigma^2$ (see also [9] for a similiar approach). Moreover, it exploits the fact that, due to the zero-forcing assumption, the optimization of the sets $\mathcal{B}$ and $\bar{\mathcal{V}}^{\perp} = \{\bar{\mathbf{V}}_k^{\perp}\}_{k=1}^{\bar{K}}$ for the management of the inter-user interference on one hand, and $\{\mathbf{Q}_k\}_{k=1}^{\bar{K}}$ for transmission over the single-user MIMO channels on the other, can be conveniently decoupled (recall (9)). Therefore, it follows that the tackled problem has the form

$$\{\mathcal{B}, \bar{\mathcal{V}}^{\perp}\} = \underset{\mathcal{B}, \bar{\mathcal{V}}^{\perp}}{\operatorname{argmax}} \sum_{i=1}^{\bar{K}} N_i(\mathcal{B}, \bar{\mathcal{V}}^{\perp}), \quad \text{s.t. (1), (3) and (10),} \tag{13}$$

where we defined $N_i(\mathcal{B}, \bar{\mathcal{V}}^{\perp}) = \left\| \mathbf{B}_i^H \mathbf{H}_i \bar{\mathbf{V}}_i^{\perp} \right\|^2$ as the norm of the equivalent single-user MIMO matrices in (11). Once problem (13) is solved, matrices $\{\mathbf{Q}_k\}_{k=1}^{K}$ are computed as explained above.

The technique proposed in [8] for the solution of (13) is limited in that it assumes the knowledge of the number of spatial channel per users $\{d_i\}_{i=1}^{\bar{K}}$ (which is in general a variable of the optimization problem). The method starts by letting $\mathbf{B}_i = \mathbf{I}_{N_R \times d_i}$, and then proceeds to optimize the matrices $\mathbf{B}_i$ and $\bar{\mathbf{V}}_i^{\perp}$ of each user, given those of all other users. The iteration stops when orthogonality between the users is achieved, i.e. when $\left\| \mathbf{B}_i^H \mathbf{H}_i \bar{\mathbf{V}}_j^{\perp} \right\|^2 \simeq 0$ for each $i \neq j$. According to the denomination used by the its authors, we refer to this technique as Generalized Multiuser Orthogonal Space Division Multiple Access (GM-OSDMA).

IV. Approximate joint optimization of scheduling and precoding (SVS)

In this section, a technique that performs joint optimization of scheduling and zero-forcing precoding according to problem (13) is proposed. As compared to the solution of [8], pre-determination of the number of spatial channels per user $\{d_i\}_{i=1}^{\bar{K}}$ is not assumed. The methods adopts a greedy approach as detailed in the following. Notice that suboptimality of this technique with respect to the original sum-rate maximization problem (7) is twofold: (*i*) as for the method in [8], the more tractable merit function (13) is considered; (*ii*) a suboptimal greedy approach to the solution of (13) is proposed. Suboptimality due to point (*ii*) can be quantified by numerical simulations through performance comparison with the GM-OSDMA method applied with a combinatorial search over *every* possible choice of parameters $\{d_i\}_{i=1}^{\bar{K}}$, as discussed in Sec. V.

A. The SVS algorithm

The idea is to select at each step the spatial channel that yields the largest increase of the objective function (13). Let us denote with the superscript (n) the quantities of interest as computed at the nth iteration. At each iteration a spatial channel (out of the N_T available) is allocated to a specific user so that a total number of N_T iterations are needed according to the constraint (1). In particular, at the nth iteration, the number of assigned spatial channels is $\sum_{i=1}^{\bar{K}} d_i^{(n)} = n$. Moreover, at each iteration, we are interested in updating the basis of the receiving subspaces $\mathbf{B}_i$ (initialized as $\mathbf{B}_i^{(0)}$ equal to an empty matrix) and the basis of the transmitting subspaces $\bar{\mathbf{V}}_i^{\perp}$, or equivalently its orthogonal complement $\bar{\mathbf{V}}_i$ (see (23), initialization: $\bar{\mathbf{V}}_i^{\perp(0)} = \mathbf{I}_{N_T}$).

Let $\mathbf{b}_j$ be a possible candidate vector to be included in the basis $\mathbf{B}_j^{(n)}$ of user j at the nth iteration ($j = 1, ..., \bar{K}$). As a result of the choice of $\mathbf{b}_j$ at the nth iteration, the objective function (13) can be written as the sum (dropping the functional dependence on $\mathcal{B}$, $\bar{\mathcal{V}}^{\perp}$ for simplicity of notation)

$$\sum_{i=1}^{\bar{K}} N_i^{(n)}(\mathbf{b}_j) = \sum_{i=1}^{\bar{K}} N_i^{(n-1)} + \sum_{i=1}^{\bar{K}} \Delta N_i^{(n)}(\mathbf{b}_j). \tag{14}$$

Among all the possible vectors $\mathbf{b}_j$ for all users $j = 1, ..., \bar{K}$, the vector $\mathbf{b}_j$ is selected so as to maximize the increase of the objective function $\sum_{i=1}^{\bar{K}} \Delta N_i^{(n)}(\mathbf{b}_j)$. In the following, the computation of $\Delta N_i^{(n)}(\mathbf{b}_j)$ is carried out.

To elaborate, we need to define for each user a basis $\mathbf{U}_j^{(n)}$ that spans the range space of the channel matrix $\mathbf{H}_j$ that, at the nth iteration, has not been assigned to any receiving subspace yet. Formally, it is: $\operatorname{span}\{\mathbf{U}_j^{(n)}\} = \operatorname{span}\{\mathbf{U}_j\} \cap \operatorname{null}\{\mathbf{B}_j^{(n)}\}$. Therefore, the corresponding initialization is $\mathbf{U}_j^{(0)} = \mathbf{U}_j$. At the nth iteration, the candidate vectors to be included in the receiving subspace of the jth user are linear combinations of the columns of $\mathbf{U}_j^{(n)}$:

$$\mathbf{b}_j = \mathbf{U}_j^{(n)} \mathbf{a}_j, \text{ with } \|\mathbf{a}_j\|^2 = 1. \tag{15}$$

Therefore, according to (15), maximization of the increase $\sum_{i=1}^{\bar{K}} \Delta N_i^{(n)}(\mathbf{b}_j)$ has to be carried out with respect to vector $\mathbf{a}_j$. In order to stress this point, in the following, with a slight abuse of notation, we will write $\sum_{i=1}^{\bar{K}} \Delta N_i^{(n)}(\mathbf{a}_j)$.

With the selection of $\mathbf{a}_j$ in (15), the corresponding basis of the receiving subspace is updated as $\mathbf{B}_j^{(n)} = [\mathbf{B}_j^{(n-1)}, \mathbf{b}_j]$ while its transmit subspace remains unchanged, $\bar{\mathbf{V}}_j^{\perp(n)} = \bar{\mathbf{V}}_j^{\perp(n-1)}$, since no new constraint (10) is imposed on it. It follows that the increase in (14) reads

$$\Delta N_j^{(n)}(\mathbf{a}_j) = \left\| \mathbf{b}_j^H \mathbf{H}_j \bar{\mathbf{V}}_j^{\perp(n)} \right\|^2. \tag{16}$$

However, the choice of $\mathbf{a}_j$ for user j results in an additional zero-interference constraint for any other user $i \neq j$ since the constraint (10) reduces to:

$$\mathbf{b}_j^H \mathbf{H}_j \bar{\mathbf{V}}_i^{\perp(n)} = \mathbf{0}, \tag{17}$$

or equivalently the condition $\mathbf{v}_j = \mathbf{H}_j^H \mathbf{b}_j \in \text{null}(\bar{\mathbf{V}}_i^{\perp(n)})$ or $\mathbf{v}_j \in \text{span}(\bar{\mathbf{V}}_i^{(n)})$ holds true. The above formulation of constraint (17) leads directly to the update of $\bar{\mathbf{V}}_i^{(n)}$ as

$$\bar{\mathbf{V}}_i^{(n)} = \left[\bar{\mathbf{V}}_i^{(n-1)} \ \mathbf{w}_i \right], \tag{18}$$

where, in order to preserve the orthonormality of the updated precoding set $\bar{\mathbf{V}}_i^{(n)}$, $\mathbf{w}_i$ is the projection of $\mathbf{v}_j$ onto $\text{span}(\bar{\mathbf{V}}_i^{\perp(n-1)})$ scaled to unit norm

$$\mathbf{w}_i = (\bar{\mathbf{V}}_i^{\perp(n-1)} \bar{\mathbf{V}}_i^{\perp(n-1)H} \mathbf{v}_j) / || \bar{\mathbf{V}}_i^{\perp(n-1)} \bar{\mathbf{V}}_i^{\perp(n-1)H} \mathbf{v}_j ||. \tag{19}$$

$\bar{\mathbf{V}}_i^{\perp(n)}$ is updated as well, so that $\text{span}(\bar{\mathbf{V}}_i^{\perp(n)}) = \text{null}(\bar{\mathbf{V}}_i^{(n)})$. This step can be performed, e.g., by updating the QR decomposition of (18) [7]. On the other hand, nothing changes at the receiver side of the ith user as $\mathbf{B}_i^{(n)} = \mathbf{B}_i^{(n-1)}$ so that the decrease in the objective function (14) due to increased interference reads:

$$\Delta N_i^{(n)}(\mathbf{a}_j) = - \left\| \mathbf{B}_i^{(n)H} \mathbf{H}_i \mathbf{w}_i \right\|^2 \quad i \neq j. \tag{20}$$

To sum up, from (16) and (20) the increase of objective function (13) due to the choice of vector $\mathbf{b}_j$ at the nth iteration is

$$\sum_{i=1}^{\bar{K}} \Delta N_i^{(n)}(\mathbf{a}_j) = \left\| \mathbf{b}_j^H \mathbf{H}_j \bar{\mathbf{V}}_j^{\perp(n)} \right\|^2 - \sum_{i \neq j} \left\| \mathbf{B}_i^{(n)H} \mathbf{H}_i \mathbf{w}_i \right\|^2. \tag{21}$$

The first term accounts for the increased useful power received by user j on the assigned spatial channel, whereas the other terms represent the power loss suffered from the other users from being prevented to transmit over $\mathbf{w}_i$ in order to guarantee the zero-forcing constraint (10).

Recalling (15) and (19), function (21) can be recognized to be a sum of Rayleigh quotients in terms of vector $\mathbf{a}_j$:

$$\sum_{i=1}^{\bar{K}} \Delta N_i^{(n)}(\mathbf{a}_j) = \frac{\mathbf{a}_j^H \mathbf{R}_j^{(n)} \mathbf{a}_j}{\mathbf{a}_j^H \mathbf{a}_j} - \sum_{i \neq j} \frac{\mathbf{a}_j^H \mathbf{R}_i^{(n)} \mathbf{a}_j^H}{\mathbf{a}_j^H \bar{\mathbf{R}}_i^{(n)} \mathbf{a}_j^H}, \tag{22}$$

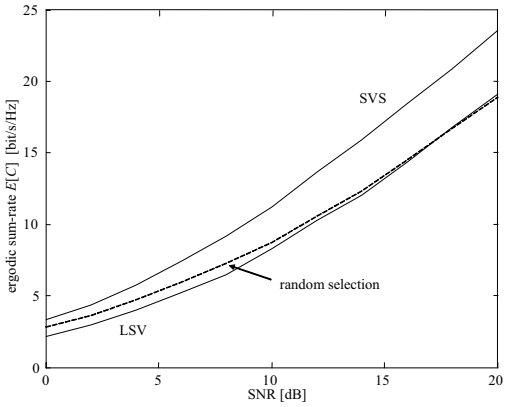

Fig. 2. Ergodic sum-rate $E[C]$ versus SNR ($\bar{K} = 4$, $N_T = 4$, $N_R = 2$, $\sigma_k^2 = 1$).

where for the definition of correlation matrices $\mathbf{R}_j^{(n)}$, $\mathbf{R}_i^{(n)}$ and $\bar{\mathbf{R}}_i^{(n)}$ and a derivation of (22) the reader is referred to Appendix-B. While the maximization of a single Rayleigh quotient is analytically feasible since it corresponds to the solution of a generalized eigenvalue problem, maximizing a sum of Rayleigh quotients requires sophisticated techniques in the context of global optimization. Here, we resort to a suboptimal approach that exploits the structure of the considered problem. In particular, we constraint $\mathbf{a}_j$ to be a column of an identity matrix, which translates to restricting our search of the optimal $\mathbf{u}_j$ to the columns of $\mathbf{U}_j^{(n)}$.

V. Numerical results

In this section, numerical analysis is presented for the downlink of a MIMO system with $N_T = 4$ transmitting antenna at the base station and $N_R = 2$ antennas at each terminal. The channel matrix $\mathbf{H}_k(t)$ is assumed to be zero mean complex Gaussian with independent identically distributed entries and variance σ_k^2. Matrices $\mathbf{H}_k(t)$ are independent over users (k) and time (t).

A. LVS versus SVS scheduling

At first, a comparison of the performance of the algorithms LSV and SVS is presented. As discussed in Sec. III, both algorithms aim at maximizing the sum-rate but while the first performs a separate optimization of scheduling and zero-forcing precoding, the latter approximate the problem of joint optimization of the two stages. Since GM-OSDM assumes the knowledge of the number of spatial channel to be granted to each user, it will be considered separately below. We consider $\bar{K} = 4$ users and a homogeneous scenario with $\sigma_k^2 = 1$ for $k = 1, .., \bar{K}$. As reference performance, a random user selection algorithm that chooses randomly a set $\mathcal{K}$ of users so as to satisfy (1) with $\mathbf{B}_k = \mathbf{I}_{N_R}$ for $k = 1, ..., K$ is considered. On this randomly selected subset, orthogonal precoding is applied as detailed in Sec. III. The ergodic sum-rate $E[C] = \sum_{i=1}^{\bar{K}} E[C_i]$ is shown versus the signal to noise

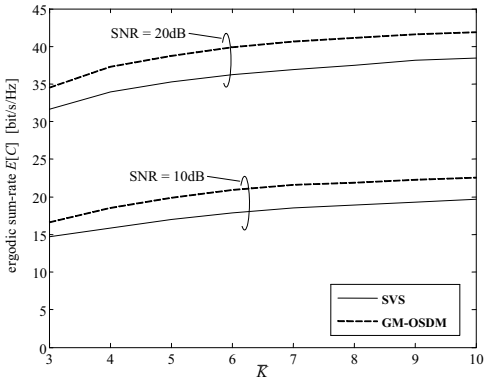

Fig. 3. Ergodic sum-rate $E[C]$ versus the number of users $\bar{K}$ for SVS and GM-OSDM ($N_T = 6$, $N_R = 3$, $\sigma_k^2 = 1$, $SNR = 10dB, 20dB$).

ratio $SNR = P/\sigma^2$ in fig. 2. The proposed SVS algorithm yields a gain of about $4dB$ in SNR as compared to the LSV algorithm and the random users selection.

B. SVS scheduling versus GM-OSDM with combinatorial search over $\{d_i(t)\}_{i=1}^{\bar{K}}$

In order to evaluate the suboptimality of SVS with respect to the approximate sum-rate optimization problem (13), a comparison between SVS and GM-OSDM is presented here. In particular, as discussed in Sec. III-A.2, for GM-OSDM an exhaustive search is performed over all possible allocations of spatial channels $\{d_i(t)\}_{i=1}^{\bar{K}}$. Here we consider $N_T = 6$ transmitting antennas at the base station and $N_R = 3$ antennas at each terminal for varying number of users K. The fading scenario is assumed to be homogeneous with $\sigma_k^2 = 1$. The ergodic sum-rate $E[C]$ is shown versus K for differerent signal-to-noise ratios $SNR = P/\sigma^2$ (10 and $20dB$) in fig. 3. It can be noticed that the gain of GM-OSDM over SVS is approximately constant with respect to K and slightly increasing for larger SNR (around $2.5dB$ for $SNR = 10dB$ and $3dB$ for $SNR = 20dB$). On the other side, the computational cost of GM-OSDM is substantially higher since the method occasionally requires a large number of iterations to converge (as opposed to the fixed number of N_T iterations needed by the SVS algorithm), and, above all, it demands an exhaustive search over the assignment of the number of spatial channels $\{d_i(t)\}_{i=1}^{\bar{K}}$.

VI. Conclusion

Joint optimization of channel-aware scheduling and zero-forcing linear precoding has been addressed for the downlink of a MIMO system. A novel approximate greedy algorithm is proposed for the maximization of the sum-rate. This method avoids the need of a combinatorial search over the number of spatial channels to be assigned to each user with moderate performance loss.

VII. Appendix-A: computation of matrix $\bar{\mathbf{V}}_k^{\perp}$

The precoding matrix $\bar{\mathbf{V}}_k^{\perp}$ in (9) has to ensure the zero-forcing constraint (10). Therefore, it should be computed as a $N_T \times d_k$ basis of the d_k-dimensional null space of the $N_T \times (N_T - d_k)$ matrix $\bar{\mathbf{H}}_k^H$ with the singular value decomposition

$$\begin{aligned}\bar{\mathbf{H}}_k &= [\tilde{\mathbf{H}}_1^H \cdots \tilde{\mathbf{H}}_{k-1}^H \tilde{\mathbf{H}}_k^H \cdots \tilde{\mathbf{H}}_K^H]^H = \\ &= \left[\bar{\mathbf{U}}_i\ \bar{\mathbf{U}}_i^{\perp}\right] \begin{bmatrix} \bar{\mathbf{\Lambda}}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \left[\bar{\mathbf{V}}_i\ \bar{\mathbf{V}}_i^{\perp}\right]^H. \end{aligned} \quad (23)$$

VIII. Appendix-B: proof of (22)

According to (15) and (19), the increment function (21) for $\mathbf{b}_j = \mathbf{U}_j^{(n)}\mathbf{a}_j$ becomes

$$\sum_{i=1}^{\bar{K}} \Delta N_i^{(n)}(\mathbf{a}_j) = \left\| \mathbf{a}_j^H \mathbf{U}_j^{(n)H} \mathbf{H}_j \bar{\mathbf{V}}_j^{\perp(n)} \right\|^2 - \sum_{i \neq j} \frac{\left\| \mathbf{B}_i^{(n)H} \mathbf{H}_i \mathbf{P}_i^{\perp(n-1)} \mathbf{H}_j^H \mathbf{U}_j^{(n)} \mathbf{a}_j \right\|^2}{\left\| \mathbf{P}_i^{\perp(n-1)} \mathbf{H}_j^H \mathbf{U}_j^{(n)} \mathbf{a}_j \right\|^2}, \quad (24)$$

where $\mathbf{P}_i^{\perp(n-1)} = \bar{\mathbf{V}}_i^{\perp(n-1)} \bar{\mathbf{V}}_i^{\perp(n-1)H}$. Equation (24) can be written in terms of the sum of Rayleigh quotients (22) by defining the correlation matrices

$$\begin{aligned}\mathbf{R}_j^{(n)} &= \mathbf{U}_j^{(n)H} \mathbf{H}_j \mathbf{P}_j^{\perp(n)} \mathbf{H}_j^H \mathbf{U}_j^{(n)}, \\ \mathbf{R}_i^{(n)} &= \mathbf{U}_j^{(n)H} \mathbf{H}_j \mathbf{P}_i^{\perp(n-1)} \mathbf{H}_i^H \mathbf{B}_i^{(n)} \mathbf{B}_i^{(n)H} \mathbf{H}_i \mathbf{P}_i^{\perp(n-1)} \mathbf{H}_j^H \mathbf{U}_j^{(n)}, \\ \bar{\mathbf{R}}_i^{(n)} &= \mathbf{U}_j^{(n)H} \mathbf{H}_j \mathbf{P}_i^{\perp(n-1)} \mathbf{H}_j^H \mathbf{U}_j^{(n)}. \end{aligned}$$

References

[1] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[2] Wei Yu and Wonjong Rhee: "Degrees of Freedom in Multi-user Spatial Multiplex Systems with Multiple Antennas", submitted to *IEEE Transactions on Communications*, March 2004.

[3] A. Goldsmith, S. A. Jafar, N. Jindal and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684-702, June 2003.

[4] Y. Taesang, A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journ. Selected Areas Commun.*, vol. 24, no. 3, pp. 528-541, March 2006.

[5] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels-Part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1083-1102, March 2001.

[6] Q. H. Spencer, A. Lee Swindlehurst and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels", *IEEE Trans. Signal Processing*, vol. 52 , no. 2 , pp. 461-471, Feb. 2004.

[7] G. H. Golub and C. F. van Loan, *Matrix Computations*, Johns Hopkins, 1996, 3rd edition.

[8] Zhengang Pan, Kai-Kit Wong and Tung-Sang Ng, "Generalized Multiuser Orthogonal Space-Division Multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969-1973, November 2004.

[9] H. Boche, M. Schubert, and E. Jorswieck, "Trace balancing for multiuser mimo downlink transmission," in *Proc. Asilomar Conf.*, 2002.

[10] G. Primolevo, O. Simeone, U. Spagnolini, "Channel aware scheduling for broadcast MIMO systems with orthogonal linear precoding and fairness constraints", in *Proc. IEEE ICC*, 2005.