BEAMFORMING AND USER SELECTION IN SDMA SYSTEMS UTILIZING CHANNEL STATISTICS AND INSTANTANEOUS SNR FEEDBACK

David Hammarwall, Mats Bengtsson, and Björn Ottersten

Signal Processing, School of Electrical Engineering, Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden (firstname.lastname@ee.kth.se)

ABSTRACT

Spatial division multiple access (SDMA) systems efficiently take advantage of the spatial dimensions of the channel to increase the performance of the system. A major difficulty, common to all SDMA systems, is the requirement of channel knowledge at the transmitter to enable transmission of multiple streams without catastrophic interference.

Herein we show that, in wide area scenarios, statistical channel information combined with the Euclidean norm of the channel realization, fed back from the users, provide sufficient information for SDMA systems to efficiently allocate users in time and space. A joint beamforming and scheduling algorithm is proposed for the downlink, which extends the proportional fair scheduling criterion to an SDMA setting, resulting in a weighted sum rate maximization.

***Index Terms*—** Spatial division multiple access, array signal processing, mobile communication, and feedback.

1. INTRODUCTION

Multiple antenna transmission schemes have many advantages over single antenna systems. The multiple antennas result in a multiple-input multiple-output (MIMO) channel which allows the transmitter to also allocate resources in the spatial dimensions, as well as in time and frequency. In particular spatial division multiple access (SDMA) [1, 2] may be utilized to allocate multiple user terminals in the same time-frequency slot, as long as they are spatially separated.

The focus of this article will be on showing that, in wide and metropolitan area scenarios, the spatial information provided by an instantaneous SNR feedback is enough for efficient SDMA beamforming and scheduling. To show that this channel state information (CSI) is sufficient, an algorithm for joint SDMA scheduling, beamforming and power control is proposed. The design goal of the proposed algorithm is thus to maximize the system performance, rather than focusing on computational complexity. However, it is shown that the complexity of the resulting weighted sum rate maximization may be reduced considerably by changing coordinates.

Herein we consider the downlink of a system where the user terminals are equipped with a single receiving antenna, whereas the base station has multiple transmit antennas. Each user terminal feeds back the instantaneous SNR, which here refers to the Euclidean norm of the channel vector, $\|\mathbf{h}\|^2$. It is further assumed that the channel statistics are known at the transmitter, either from a low rate feedback or estimated directly from the reverse link (uplink). The CSI provided by the channel norm, when combined with statistics of the channel, was first characterized for Rayleigh fading channels and later extended to Ricean fading [3]. In particular it was shown that, for large channel norms, the channel is almost fully determined.

An efficient scheduler ensures that in each time slot a set of users that currently experience favorable channel realizations are scheduled, as to take full advantage of the multi user diversity principle [4, 5]. In a cell with many users, only the users with particularly strong channel norms are thus candidates for scheduling. For these users, the channel norm thus contains substantial CSI.

Combining the information contained in the correlation matrices with that of the instantaneous channel norm, in an SDMA setting, has also been addressed in [6, 7], where the authors propose to compute the maximum likelihood (ML) estimates of each user's channel. These estimates are however not accurate enough for data transmission, and the scheduled users therefore feed back full CSI. Herein, we utilize the framework of minimum mean square error (MMSE) estimates [3]. These estimates have several advantages over ML estimation; in particular they model the interference more accurately and furthermore the mean square error (MSE) of each estimate may be computed explicitly which allows for adaptive error margins.

2. SYSTEM MODEL

Herein we model the channel as Ricean fading. The M-antenna vector channel of user k, is distributed as $\mathbf{h}_k \in \mathcal{CN}(\overline{\mathbf{h}}_k, \mathbf{R}_k)$, where $\overline{\mathbf{h}}_k \in \mathbb{C}^M$ is the channel mean and $\mathbf{R}_k \in \mathbb{C}^{M\times M}$ is the covariance matrix. We make the assumption that $\overline{\mathbf{h}}_k$ and $\mathbf{R}_k$ are perfectly known at the transmitter, whereas the realization, $\mathbf{h}_k$, is only known at the receiver. The symbol sampled, complex base band equivalent, of the received signal is modeled as

$$y_k(t) = \mathbf{h}_k^{\mathrm{H}}\mathbf{x}(t) + n_k(t), \tag{1}$$

where $n_k(t)$ is additive white Gaussian noise (AWGN) with power σ_k^2 and $\mathbf{x}(t) \in \mathbb{C}^M$ is the vector of transmitted signals. Let $\mathcal{S}$ be the set of users that are scheduled for transmission in the current time slot. We consider here a beamforming system where the signal $s_k(t)$, intended for user k, is mapped onto the antenna array with the beamforming vector, $\mathbf{u}_k$,

$$\mathbf{x}(t) = \sum_{k\in\mathcal{S}} \sqrt{p_k}\mathbf{u}_k s_k(t), \tag{2}$$

where the scalar $s_k(t)$ and $\mathbf{u}_k$ are normalized to unit power, and p_k is the power allocated to user k. For notational convenience the beamformers, powers and noise levels, σ_k^2 are collected as the columns of $\mathbf{U}$ and elements of $\mathbf{p}$ and $\boldsymbol{\sigma}$, respectively. The instantaneous signal to interference and noise ratio (SINR) for user k is, ignoring inter-cell interference, obtained by combining (1) and (2) as

$$\mathrm{SINR}_k = \frac{\mathbf{w}_k^{\mathrm{H}}\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k}{\sum\limits_{i\in\mathcal{S}\setminus\{k\}} \mathbf{w}_i^{\mathrm{H}}\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_i + \sigma_k^2}. \tag{3}$$

The transmission rate of user $k \in \mathcal{S}$, that is supported, is directly related to SINR_k. This relation is described by the rate function

This work has been performed in the framework of the IST project IST-4-027756 WINNER II, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues.

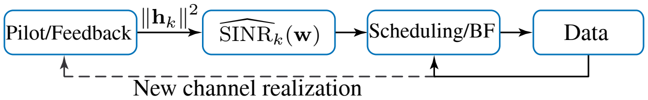

Fig. 1. Illustration of the system operation

$r_k = r(\mathrm{SINR}_k)$, which herein is assumed to be non decreasing. In the simulations we take the rate function as the Shannon capacity of the channel, $r(\mathrm{SINR}) = \log(1 + \mathrm{SINR})$, but the results derived herein are not restricted to this choice. The optimization is however greatly simplified if $r(\mathrm{SINR})$ is smooth, i.e. everywhere differentiable with respect to SINR. The commonly used gap approximation, $r(\mathrm{SINR}) = \log(1 + \mathrm{SINR}/\mathrm{Gap})$, also satisfies these requirements.

2.1. System Operation

The system operation is illustrated in Fig. 1. Initially the base station transmits orthogonal pilot sequences on all antennas, which allows each user to estimate the channel norm, $\rho_k = \|\mathbf{h}_k\|^2$. This channel quality index, ρ_k, is next fed back to the base station, which combines the information with the channel statistics, known to the transmitter, and estimates how the SINRs depend on the beamformers and power allocation. These estimates are utilized by the scheduler which allocates the resources to an appropriate set of users.

2.2. Channel Knowledge

The SNR feedback, ρ_k, is used at the transmitter to estimate the instantaneous channel correlation as, $\hat{\mathbf{R}}_k \triangleq \mathrm{E}\left\{\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}} \mid \rho_k\right\}$. These MMSE estimates are derived and analyzed for both Rayleigh and Ricean fading channels in [3, 8], where it is shown how to efficiently compute them. Using $\hat{\mathbf{R}}_k$ the MMSE estimate of the signal (or interference) power, $p_i\left|\mathbf{u}_i^{\mathrm{H}}\mathbf{h}_k\right|^2$, is obtained as $p_i\mathbf{u}_i^{\mathrm{H}}\hat{\mathbf{R}}_k\mathbf{u}_i$. The rate and SINR is thus estimated as

$$\hat{r}_k = r(\widehat{\mathrm{SINR}}_k), \quad \widehat{\mathrm{SINR}}_k = \frac{p_k\mathbf{u}_k^{\mathrm{H}}\hat{\mathbf{R}}_k\mathbf{u}_k}{\sum_{i\in\mathcal{I}\setminus\{k\}} p_i\mathbf{u}_i^{\mathrm{H}}\hat{\mathbf{R}}_k\mathbf{u}_i + \sigma_k^2}. \qquad (4)$$

2.3. Pessimistic SINR estimation

When allocating a rate to a user it is important not to overestimate the SINR, since this leads to unacceptable frame error rates. The SINR should therefore be estimated pessimisticly.

One of the major causes of overestimating the SINR is destructive cancellation of the signal power. Since $\hat{\mathbf{R}}_k$ is dominated by a single eigenvalue (by the wide area scenario assumption) the beamformer should almost be aligned with the corresponding eigenvector. In particular for the Rayleigh fading case, it is impossible to determine if the signal power projected on the remaining eigenvectors will add constructively or destructively based on the norm feedback. We therefore define the indefinite matrix $\underline{\hat{\mathbf{R}}}_k$, having the same eigenvectors as $\hat{\mathbf{R}}_k$ but the sign reversed on all eigenvalues, except the largest. In the optimization, destructive signal cancellation is avoided by optimizing with respect to the $\underline{\widehat{\mathrm{SINR}}}_k$ given by (4), but with $\underline{\hat{\mathbf{R}}}_k$ in place of $\hat{\mathbf{R}}_k$ in the numerator.

When the optimal beamformers and power allocation have been computed with respect to $\underline{\widehat{\mathrm{SINR}}}$, the SINRs are adjusted according to the actual MSE of the estimates as

$$\mathrm{SINR}_k^{\mathrm{MSE}} = \frac{p_k\left[\mathbf{u}_k^{\mathrm{H}}\hat{\mathbf{R}}_k\mathbf{u}_k - \alpha\sqrt{\mathrm{MSE}\left(|\mathbf{u}_k^{\mathrm{H}}\mathbf{h}_k|^2\middle|\rho_k\right)}\right]}{\sum_{i\in\mathcal{I}\setminus\{k\}} p_i\left[\mathbf{u}_i^{\mathrm{H}}\hat{\mathbf{R}}_k\mathbf{u}_i + \alpha\sqrt{\mathrm{MSE}\left(|\mathbf{u}_i^{\mathrm{H}}\mathbf{h}_k|^2\middle|\rho_k\right)}\right] + \sigma_k^2}.$$

These MSEs are readily computable for both Rayleigh fading and Ricean fading channels [8]. By adjusting α the block error rate may be kept at any level of choice. In the simulations α was set to $\alpha = 2$, which resulted in approximately 5% probability of overestimating the SINR.

2.4. Scheduling and Beamforming

Allocating system resources to the users is a balance between total system throughput and fairness among the users. Also, the multiuser diversity should be exploited to increase the throughput while ensuring that the delays are kept at a reasonable level.

In order to take all these factors into account the design of the joint scheduling, power control and beamforming is based on a weighted sum rate criterion

$$\mathrm{R}_\Sigma(\mathbf{U},\mathbf{p}) = \sum_{i=1}^{d}\alpha_i\, r(\mathrm{SINR}_i) = \sum_{i\in\mathcal{S}}\alpha_i\, r(\mathrm{SINR}_i), \qquad (5)$$

where the weights, α_i, determines the priority of users and may be chosen according to, for instance, the proportional fair scheduling criterion [5]. Optimally the beamformers and power allocation should be chosen as to maximize the weighted sum rate,

$$(\mathbf{U}^\star, \mathbf{p}^\star) = \arg\max_{(\mathbf{U},\mathbf{p})\in\mathcal{W}} \mathrm{R}_\Sigma(\mathbf{U},\mathbf{p}), \qquad (6)$$

where $\mathcal{W}$ is the set of feasible beamformers and power allocations, i.e. beamformers satisfying $\|\mathbf{u}\| = 1$ and $\mathbf{p}$ satisfying $p_i \geq 0$ and the sum power constraint, $\mathbf{1}^{\mathrm{T}}\mathbf{p} \leq \mathrm{P}_{\max}$. The optimal set of scheduled users is implicitly determined as $\mathcal{S}^\star = \{i \mid p_i^\star > 0\}$.

The optimization problem in (6) is however non-convex and thus non-trivial to optimize, in particular since the non-convex optimization involves many parameters. In Section 3 a more structured form of the optimization problem is derived, where the number of non-convex optimization parameters is reduced.

3. PROPOSED OPTIMIZATION FOR JOINT SCHEDULING AND BEAMFORMING

3.1. Re-parametrization of the weighted sum rate optimization

In order to simplify the weighted sum rate maximization, the SINRs are expressed in terms of a common gain factor, Γ, and the relative gains, γ_i, as

$$\mathrm{SINR}_i = \Gamma\gamma_i,\ \forall i \quad \text{with} \quad \mathbf{1}^{\mathrm{T}}\boldsymbol{\gamma} = 1,\ \Gamma \in \mathcal{G}_{\boldsymbol{\gamma}} \qquad (7)$$

where $\mathcal{G}_{\boldsymbol{\gamma}} \triangleq \{\Gamma \mid \exists\ (\mathbf{U},\mathbf{p}) \in \mathcal{W},\ \text{s.t.}\ \mathrm{SINR}_i(\mathbf{U},\mathbf{p}) = \Gamma\gamma_i\ \forall i\}$ is the set of achievable Γ for a given $\boldsymbol{\gamma}$. By the assumption of non-decreasing rate functions it follows that the maximization in (6) may be expressed as

$$\mathrm{R}_\Sigma{}^\star = \max_{\boldsymbol{\gamma},\ \mathbf{1}^{\mathrm{T}}\boldsymbol{\gamma}=1} \sum_i \alpha_i\, r\big(\Gamma^\star(\boldsymbol{\gamma})\gamma_i\big), \qquad (8)$$

where $\Gamma^\star(\boldsymbol{\gamma}) = \max\mathcal{G}_{\boldsymbol{\gamma}}$. For a given $\boldsymbol{\gamma}$, $\Gamma^\star(\boldsymbol{\gamma})$ is given by the max-min problem,

$$\Gamma^\star(\boldsymbol{\gamma}) \triangleq \max_{(\mathbf{U},\mathbf{p})\in\mathcal{W}} \min_i \frac{\mathrm{SINR}_i(\mathbf{U},\mathbf{p})}{\gamma_i}, \qquad (9)$$

which is thoroughly analyzed in [1], where an efficient algorithm for the optimization is derived. It was further shown that at the optimum, all the SINRs are balanced, i.e. $\mathrm{SINR}_i/\gamma_i = \Gamma^\star(\boldsymbol{\gamma})\ \forall i$, from which it follows that the max-min problem is equivalent to $\max\mathcal{G}_{\boldsymbol{\gamma}}$.

The maximization in (8) does however remain non-convex, but the number of optimization parameters is substantially reduced. The non-convex optimization is implemented as a gradient search starting at a heuristically chosen $\boldsymbol{\gamma}_0$. Next it is shown that this search may be efficiently implemented since also the gradient may be computed from the optimization parameters of (9) with negligible additional complexity.

Table 1. Outline of the proposed algorithm for scheduling, beamforming, and power control, with a weighted sum rate criterion.

1: $\forall i : \mathbf{u}_i \leftarrow \arg\max_{\mathbf{u}:\ \|\mathbf{u}\|=1} \mathbf{u}^{\mathrm{H}} \hat{\mathbf{R}}_i \mathbf{u}$
2: $\mathcal{C} \leftarrow$ set of all users
3: $\mathcal{S}_{\text{new}} \leftarrow \emptyset$, $\mathbf{p}_{\text{new}} \leftarrow \mathbf{0}$, $\mathrm{R}_{\Sigma\text{new}} \leftarrow 0$, $\boldsymbol{\gamma}_{\text{new}} \leftarrow \mathbf{0}$
4: **repeat**
5: $\quad \mathcal{S} \leftarrow \mathcal{S}_{\text{new}}$, $\mathbf{p} \leftarrow \mathbf{p}_{\text{new}}$, $\mathrm{R}_{\Sigma} \leftarrow \mathrm{R}_{\Sigma\text{new}}$, $\boldsymbol{\gamma} \leftarrow \boldsymbol{\gamma}_{\text{new}}$
6: $\quad \forall i \in \mathcal{C}:\ \mathrm{R}_{\Sigma}[i] \leftarrow \mathrm{R}_{\Sigma}(\mathbf{U}, \mathrm{P}_{\max}\Theta_i\{\mathbf{p}\})$
7: $\quad i^\star = \arg\max_{i\in\mathcal{C}} \mathrm{R}_{\Sigma}[i]$
8: $\quad \mathcal{S}_{\text{new}} \leftarrow \mathcal{S} \cup \{i^\star\}$
9: $\quad \mathcal{C} \leftarrow \left\{ i \,\middle|\, i \in \mathcal{C}, i \neq i^\star, \mathrm{R}_{\Sigma}[i] > \beta\, \mathrm{R}_{\Sigma}[i^\star] \right\}$
10: $\quad \gamma_0 \leftarrow \Theta_{i^\star}\{\boldsymbol{\gamma}\}$
11: $\quad (\mathrm{R}_{\Sigma\text{new}}, \boldsymbol{\gamma}_{\text{new}}, \mathbf{p}_{\text{new}}, \mathbf{U}_{\text{new}}) \leftarrow f(\mathcal{S}_{\text{new}}, \gamma_0)$
12: **until** $\mathrm{R}_{\Sigma\text{new}} \leq \mathrm{R}_{\Sigma}$
13: **return** $\mathcal{S}$, $\mathbf{U}$, $\mathbf{p}$

3.2. Gradients

Define the optimal weighted sum rate for given relative gains, $\boldsymbol{\gamma}$, as $\mathrm{R}_{\Sigma}(\boldsymbol{\gamma}) \triangleq \sum_i \alpha_i\, r\big(\Gamma^\star(\boldsymbol{\gamma})\gamma_i\big)$. The gradient of $\mathrm{R}_{\Sigma}(\boldsymbol{\gamma})$ is thus given by

$$\nabla \mathrm{R}_{\Sigma}(\boldsymbol{\gamma}) = \Gamma^\star(\boldsymbol{\gamma})\mathbf{r}'(\boldsymbol{\gamma}) + \left[\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{r}'(\boldsymbol{\gamma})\right]\nabla\Gamma^\star(\boldsymbol{\gamma}), \qquad (10)$$

where $[\mathbf{r}'(\boldsymbol{\gamma})]_i = \alpha_i\, r'\Big(\Gamma^\star(\boldsymbol{\gamma})\gamma_i\Big)$. The gradient of the weighted sum rate function is thus obtained from the gradient of $\Gamma^\star(\boldsymbol{\gamma})$. As it turns out, this gradient may be expressed explicitly in terms of the optimization variables of (9). The solver of (9) utilizes virtual uplink duality and first solves for $\mathbf{U}$ and the virtual uplink powers $\mathbf{q}$, from which the optimal $\mathbf{p}$ is obtained [1, 2]. The next theorem states that the gradient, $\nabla\Gamma^\star(\boldsymbol{\gamma})$, may be computed from $\mathbf{U}$, $\mathbf{p}$ and $\mathbf{q}$.

Theorem 1. *The gradient of $\Gamma^\star(\boldsymbol{\gamma})$, can be expressed in terms of the optimal beamformers $\mathbf{U}^\star$, power allocation $\mathbf{p}^\star$, and virtual uplink powers $\mathbf{q}^\star$, as defined in [2]. The gradient is given by*

$$\nabla\Gamma^\star(\boldsymbol{\gamma}) = -\Gamma^\star(\boldsymbol{\gamma})\frac{\mathit{diag}\{\boldsymbol{\gamma}\}^{-1}\,\mathit{diag}\{\mathbf{q}^\star\}\left(\boldsymbol{\Psi}\mathbf{p}^\star + \boldsymbol{\sigma}\right)}{\mathbf{q}^{\star T}\boldsymbol{\Psi}\mathbf{p}^\star + P_{\max}},$$

where $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\mathbf{U}^\star)$ is the cross-coupling matrix, as defined in [2]. The gradient is defined if the optimal point, $(\mathbf{U}^\star, \mathbf{p}^\star, \mathbf{q}^\star)$, is unique.

Proof. See Appendix A. □

4. REDUCED COMPLEXITY USING GREEDY USER SELECTION

Even though it is possible, in principle, to optimize the scheduling, power control and beamformers directly using (8) this is not reasonable from a practical point of view. A large set of users cause overwhelming computational complexity and finding a good initial guess becomes increasingly hard. We therefore propose a greedy approach where the optimization is not explicitly used for scheduling. Instead we propose a user selection, also based on the weighted sum rate, and restrict the optimization in (8) to just a few selected users at a time. The algorithm is given in Table 1. In summary, each iteration of the algorithm is made up of three stages,

1. Add the most compatible user to the set of scheduled users, $\mathcal{S}$
2. Remove incompatible users from the set of candidates, $\mathcal{C}$
3. Optimize $\mathbf{U}$ and $\mathbf{p}$ for users in $\mathcal{S}$.

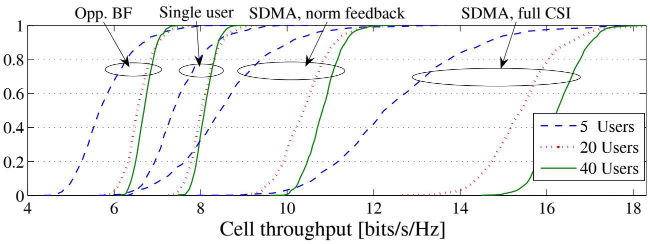

Fig. 2. The CDF of the average cell throughput achieved in a scenario. The statistics where collected in 1500 random scenarios.

This process is continued until adding another user results in a decreased weighted sum rate, or until the set of candidate users, $\mathcal{C}$, is empty. The operator $\Theta_i\{\cdot\}$ is defined as

$$\mathbf{y} = \Theta_i\{\mathbf{x}\}, \text{ where } [\mathbf{y}]_j = \begin{cases} \frac{1}{S+1}, & j = i, \\ \frac{S}{S+1}\frac{[\mathbf{x}]_j}{\mathbf{1}^{\mathrm{T}}\mathbf{x}}, & j \neq i, \end{cases}$$

where $S \triangleq |\{j\,|j \neq i,\ [\mathbf{x}]_j \neq 0\}|$ is the number of nonzero elements in $\mathbf{x}$. For $\mathbf{x} = \mathbf{0}$, the interpretation is that $[\mathbf{y}]_i = 1$ and $[\mathbf{y}]_j = 0$ for $j \neq i$. For instance, if power is allocated to S users, then $\Theta_i\{\mathbf{p}\}$ reallocates a fraction $1/(S+1)$ of the total power to user i. Also note that the elements of $\mathbf{y}$ by construction sum to 1 and an updated power allocation is thus obtained as $\mathbf{p}_{\text{new}} = \mathrm{P}_{\max}\Theta_i\{\mathbf{p}\}$, as on line 6 in Table 1.

In order to simplify the notation of the algorithm we define the function,

$$(\mathrm{R}_{\Sigma}{}^\star, \boldsymbol{\gamma}^\star, \mathbf{p}^\star, \mathbf{U}^\star) = f(\mathcal{S}, \gamma_0)$$

as the optimal weighted sum rate and optimal points obtained in the optimization in (8) and (9). In the optimization, only users in $\mathcal{S}$ are considered and the gradient search starts at the point γ_0. Note that it is necessary to also state the initial guess since the optimization is non-convex.

Furthermore, in each iteration all users that are no longer spatially compatible are eliminated. The set $\mathcal{C}$ contains the list of candidates for scheduling. In each iteration the users that are no longer spatially compatible are eliminated, see line 9 in the algorithm. Only users that, if selected, result in a weighted sum rate larger than the fraction β ($= .5$ in the simulations) of the best choice are kept in the optimization.

5. PERFORMANCE EVALUATION

The performance of the algorithms is evaluated by simulating a single cell of a communication system. The scheduling fairness is determined by the proportional fairness criterion [5] with a rate averageing window of 100 blocks. The users are distributed uniformly in a circular cell, centered at the base station equipped with a 4 element, uniform circular array (UCA). The antennas are separated by half a wavelength, $\lambda/2$. The path loss is modeled as proportional to r^{-2}, where r is the distance to the base station and at the cell edge the single antenna SNR, $\mathrm{P}_{\max}\mathrm{E}\left\{|[\mathbf{h}]_i|^2\right\}/\sigma^2$, is set to 10dB. When not otherwise stated, the angular spread was set to 10 degrees. In the simulations the outage (probability of overestimating the SINR) was $\leq 5\%$. The presented figures are not adjusted for this, since a packet sent in outage should not necessarily be considered as lost, especially in system utilizing hybrid ARQ.

In Fig. 2 the cumulative distribution functions (CDF) of the average cell throughput, in randomly generated scenarios, for 4 different schemes are compared for different numbers of active users in the cell. The considered schemes are (from left to right): Opportunistic Beamforming [5]; Single user scheduling with norm feedback [3]; The proposed algorithm, Table 1, with norm feedback; and the proposed algorithm with perfect CSI, i.e. $\hat{\mathbf{R}}_k = \mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}$. The performance in each scenario was estimated using 300 channel realizations. The channel is modeled as Rayleigh block fading where a single scheduling decision is made in each block.

As can be seen in Fig. 2 there is a significant throughput gain for the proposed SDMA scheme over other schemes with comparable feedback. It is clear from the simulations that the information provided by second order statistics and a norm feedback is indeed sufficint for efficient SDMA beamforming and scheduling. The gain of the proposed SDMA scheme is even greater in systems utilizing more than 4 antenna elements.

The performance gain for the SDMA scheme with norm feedback is achievable in a large range of angular spreads. The performance was simulated for spreads varying from 0 to 30 degrees. Note that the angular spread is related to the eigenvalue distribution of $\mathbf{R}_k$, and is a measure of the information known at the transmitter.

In summary, for 0 degrees angular spread, the channel is fully known (up to a phase rotation) at the transmitter and the performance is identical to that of perfect CSI. As the angular spread increases, the performance gradually decreases and for angular spreads higher than 20 degrees, the performance of the SDMA scheme drops below that of single user scheduling, which is due to increased sensitivity to estimation errors. For spreads higher than 25 degrees the performance also drops below that of opportunistic beamforming. The performance of opportunistic beamforming and the schemes with perfect CSI do not change notably with angular spread. But, as illustrated in Fig. 2, the performance of SDMA with norm feedback is significantly higher than those of the comparable schemes for angular spreads below 20 degrees.

6. CONCLUSIONS

An algorithm for joint SDMA scheduling, beamforming and power control has been proposed, based on the weighted sum rate criterion which includes fairness among users. The proposed algorithm was used to show that a system with only channel norm information and statistical channel information at the transmitter may achieve significant gains by SDMA over other schemes with similar feedback. The performance gain is achieved for angular spreads up to 20 degrees, which makes the scheme ideal for wide area and metropolitan area scenarios.

A. APPENDIX, PROOF OF THEOREM 1

Due to limited space, only an outline of the proof of Theorem 1 is given. For simplicity we use the same notation as in [2], where the following optimality conditions of (9) are reviewed:

$$\mathbf{Z}_i\mathbf{u}_i = \mathbf{0},\ \mathbf{Z}_i \triangleq \frac{q_i}{\Gamma\gamma_i}\mathbf{R}_i - \sum_{j\neq i} q_j\mathbf{R}_j - \mathbf{I} \succeq \mathbf{0},\ \forall i \tag{11}$$

$$\boldsymbol{\Upsilon}(\mathbf{U}, \mathrm{P}_{\max})\mathbf{p}_{\text{ext}} = \frac{1}{\Gamma}\mathbf{p}_{\text{ext}}, \quad \mathbf{p}_{\text{ext}} = [\mathbf{p}, 1]^{\mathrm{T}}, \tag{12}$$

$$\boldsymbol{\Lambda}(\mathbf{U}, \mathrm{P}_{\max})\mathbf{q}_{\text{ext}} = \frac{1}{\Gamma}\mathbf{q}_{\text{ext}}, \quad \mathbf{q}_{\text{ext}} = [\mathbf{q}, 1]^{\mathrm{T}}, \tag{13}$$

where $\boldsymbol{\Upsilon}(\mathbf{U}, \mathrm{P}_{\max})$ and $\boldsymbol{\Lambda}(\mathbf{U}, \mathrm{P}_{\max})$ are the extended coupling matrix and extended virtual uplink coupling matrix, respectively [2]. Next observe that the left eigenvector of $\boldsymbol{\Lambda}$ is given by

$$\boldsymbol{\Lambda}^{\mathrm{T}}\widetilde{\mathbf{p}}_{\text{ext}} = \frac{1}{\Gamma}\widetilde{\mathbf{p}}_{\text{ext}}, \quad \widetilde{\mathbf{p}}_{\text{ext}} = \mathbf{T}\mathbf{p}_{\text{ext}} \triangleq \begin{bmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \mathrm{P}_{\max} \end{bmatrix} \mathbf{p}_{\text{ext}},$$

which follows from the factorization $\boldsymbol{\Lambda}^{\mathrm{T}} = \mathbf{T}\boldsymbol{\Upsilon}\mathbf{T}^{-1}$ and (12).

Next consider the derivative of $\Gamma^\star(\boldsymbol{\gamma})$ with respect to γ_n. Define $\lambda(\gamma_n)$ as the eigenvalue of optimality condition (13), $\lambda(\gamma_n) \triangleq 1/\Gamma \triangleq \Gamma^\star(\boldsymbol{\gamma})^{-1}$. Then,

$$\frac{\partial}{\partial\gamma_n}\Gamma^\star(\boldsymbol{\gamma}) = -\frac{\lambda'(\gamma_n)}{\lambda^2(\gamma_n)} = -\Gamma^2\lambda'(\gamma_n) = -\Gamma^2\frac{\widetilde{\mathbf{p}}_{\text{ext}}^{\mathrm{T}}\boldsymbol{\Lambda}'\mathbf{q}_{\text{ext}}}{\widetilde{\mathbf{p}}_{\text{ext}}^{\mathrm{T}}\mathbf{q}_{\text{ext}}},$$

where $\boldsymbol{\Lambda}' \triangleq \partial/\partial\gamma_n\,\boldsymbol{\Lambda}$. The last identity follows from [9, Theorem 6.3.12] since $\widetilde{\mathbf{p}}_{\text{ext}}$ is a left eigenvector of $\boldsymbol{\Lambda}$. This holds whenever $\lambda(\gamma_n)$ is a simple eigenvalue, i.e. the solution is unique. The derivative is obtained by straight forward derivations as

$$\widetilde{\mathbf{p}}_{\text{ext}}^{\mathrm{T}}\boldsymbol{\Lambda}'\mathbf{q}_{\text{ext}} = \widetilde{\mathbf{p}}_{\text{ext}}^{\mathrm{T}}\begin{bmatrix} \mathbf{I} \\ \frac{\boldsymbol{\sigma}^{\mathrm{T}}}{\mathrm{P}_{\max}} \end{bmatrix}\frac{\partial}{\partial\lambda_n}\left[\mathbf{D}\boldsymbol{\Psi}^{\mathrm{T}}\,\mathbf{D}\mathbf{1}\right]\mathbf{q}_{\text{ext}} = (\boldsymbol{\Psi}\mathbf{p} + \boldsymbol{\sigma})^{\mathrm{T}}\mathbf{a}$$

where $\mathbf{D}$ is defined in [2] and the vector $\mathbf{a}$, obtained by taking the derivative, is given by

$$[\mathbf{a}]_i = \delta_{in}\frac{\mathbf{u}_i^{\mathrm{H}}\big(\sum_{j\neq i} q_j\mathbf{R}_j + \mathbf{I}\big)\mathbf{u}_i}{\mathbf{u}_i^{\mathrm{H}}\mathbf{R}_i\mathbf{u}_i} + 2\gamma_i\Re\left\{\frac{\mathbf{u}_i^{\mathrm{H}}\big(\sum_{j\neq i} q_j\mathbf{R}_j\big)\mathbf{u}_i'}{\mathbf{u}_i^{\mathrm{H}}\mathbf{R}_i\mathbf{u}_i} - \frac{\mathbf{u}_i^{\mathrm{H}}\mathbf{R}_i\mathbf{u}_i'}{\mathbf{u}_i^{\mathrm{H}}\mathbf{R}_i\mathbf{u}_i}\frac{\mathbf{u}_i^{\mathrm{H}}\big(\sum_{j\neq i} q_j\mathbf{R}_j + \mathbf{I}\big)\mathbf{u}_i}{\mathbf{u}_i^{\mathrm{H}}\mathbf{R}_i\mathbf{u}_i}\right\} = \delta_{in}\frac{q_i}{\Gamma\gamma_i},$$

where the last identity follows by applying (11) repeatedly. Since the Kronecker delta function, $\delta_{in} = 1$ for $i = n$ and zero otherwise, the derivative with respect to γ_n simplifies to

$$\frac{\partial}{\partial\gamma_n}\Gamma^\star(\boldsymbol{\gamma}) = -\frac{\Gamma q_n^\star}{\gamma_n}\frac{\left[\boldsymbol{\Psi}\mathbf{p}^\star\right]_n + \sigma_n^2}{\mathbf{q}^{\star\mathrm{T}}\boldsymbol{\Psi}\mathbf{p}^\star + \mathrm{P}_{\max}}.$$

Since n is arbitrary, the theorem follows.

B. REFERENCES

[1] M. Schubert and H. Boche, "Solution of the multi-user downlink beamforming problem with individual SINR constraints," *IEEE Trans. Vehicular Tech.*, vol. 53, no. 1, Jan. 2004.

[2] D. Hammarwall, M. Bengtsson, and B. Ottersten, "On downlink beamforming with indefinite shaping constraints," *IEEE Trans. on Signal Processing*, vol. 54, pp. 3566–3580, September 2006.

[3] D. Hammarwall and B. Ottersten, "Exploiting the spatial information provided by channel statistics and SNR feedback," in *IEEE Proc. Int. Workshop on Signal Proc. Advances for Wireless Comm. (SPAWC)*, July 2006.

[4] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE Int. Conf. on Comm. (ICC'95)*, June 1995, vol. 1, pp. 331–335.

[5] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[6] D. Gesbert, L. Pittman, and M. Kountouris, "Transmit correlation-aided scheduling in multiuser MIMO networks," in *Proc. IEEE, ICASSP*, May 2006.

[7] M. Kountouris, D. Gesbert, and L. Pittman, "Transmit correlation-aided opportunistic beamforming and scheduling," in *Proc. European Signal Processing Conf. (EUSIPCO)*, 2006.

[8] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. on Sig. Proc.*, Submitted.

[9] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.