DISTRIBUTED PRECODER OPTIMIZATION FOR INTERFERING MISO CHANNELS

David A. Schmidt, Andreas Gründinger, Wolfgang Utschick

Associate Institute for Signal Processing
Technische Universität München, 80290 Munich, Germany
+49 89 289-28508, {dschmidt|utschick}@tum.de

Michael L. Honig

Dept. of Electrical Engineering and Computer Science
Northwestern University, Evanston, IL 60208, USA
mh@eecs.northwestern.edu

ABSTRACT

In this paper, we examine the problem of two *multiple-input-single-output* (MISO) links that interfere with each other. We discuss some straightforward strategies for the users to optimize their rates and show how a certain degree of cooperation between the users can lead to an improvement of both individual links. We propose an algorithm that is based on the users exchanging 'interference prices' and is realizable in a distributed manner. The algorithm converges towards a locally sum-rate optimal solution and exhibits nearly identical performance as a non-distributed gradient ascent algorithm, outperforming the straightforward strategies over the complete range of noise powers.

1. INTRODUCTION

Consider a system with multiple communication links consisting of one transmitter and one receiver each. The receivers are assumed to receive not only the signal from the desired transmitter, but also from at least one other transmitter, thus being subject to interference. A scenario like this could occur in the downlink of a cellular system in which one or more mobile stations are located close to the edge of their cell. Distributed power control algorithms aimed at fairness or a guaranteed link quality for such system models have been explored in [1, 2]. In [3, 4] it has been investigated how some *sum-utility* of the system (such as sum-throughput) can be increased if the transmitters are able to cooperate in adjusting their transmit powers.

If the transmitters have more than one antenna at their disposal, the interference can also be decreased by making use of the spatial degrees of freedom. In [5], the authors introduced an iterative waterfilling scheme in which the transmitters alternatingly optimize their own covariance matrices taking into account the interference from the other users, but not cooperating with them. A general gradient ascent algorithm aimed at the *sum-rate* was proposed in [6]; this approach, however, is not suited for distributed implementation.

In this contribution, we restrict ourselves to the *multiple input single output* (MISO) case, where the receivers have only one antenna. Furthermore, we examine only two interfering communication links. This elementary setting, however, introduced in detail in Section 2, already allows us to show some fundamental effects of such interference scenarios.

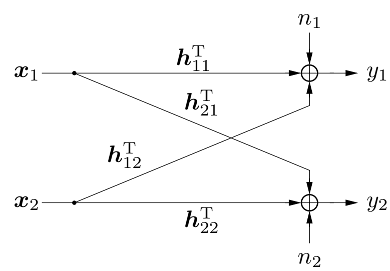


Fig. 1. System Model for Two Interfering MISO Links

By discussing some very simple non-cooperative and cooperative transmit strategies in Section 3, we are able to demonstrate how egoistical behavior limits the possible system throughput and why cooperation is necessary. In Section 4, we formulate the criterion that allows us to search for an optimal strategy, and in Section 5, we extend the distributed interference pricing scheme from [4] to our scenario. Finally, we compare all discussed schemes through numerical simulations in Section 6 and show how the achievable rate region can be non-convex.

Notation: Throughout the paper, vectors are typeset in lower-case boldface letters and matrices in upper-case boldface letters. We use $(\bullet)^{\mathrm{T}}$ for the transpose of a vector or matrix, $(\bullet)^*$ for the complex conjugate, and $\mathrm{E}[\bullet]$ for the expected value.

2. SYSTEM MODEL

We consider a system with two transmitters and two receivers, as depicted in Fig. 1. The i-th transmitter has N_i antennas, both receivers have one antenna. The vectors $\boldsymbol{h}_{ij} \in \mathbb{C}^{N_j}$ contain the channel coefficients between the N_j antennas of the j-th transmitter and the single antenna of the i-th receiver. The channels are assumed to be frequency flat.

The transmit vectors $\boldsymbol{x}_i \in \mathbb{C}^{N_i}$ that are to be transmitted

978-1-4244-1757-5/08/$25.00 © 2008 IEEE

from the i-th transmitter to the i-th receiver experience additive interference from the other link and additive circularly symmetric white Gaussian noise n_i. The received symbols y_i consequently read as

$$y_1 = \boldsymbol{h}_{11}^{\mathrm{T}} \boldsymbol{x}_1 + \boldsymbol{h}_{12}^{\mathrm{T}} \boldsymbol{x}_2 + n_1 \quad \text{and}$$
$$y_2 = \boldsymbol{h}_{22}^{\mathrm{T}} \boldsymbol{x}_2 + \boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{x}_1 + n_2.$$

The power of the noise n_i is $\sigma_i^2 = \mathrm{E}\left[|n_i|^2\right]$.

We express the transmit vectors as $\boldsymbol{x}_i = \boldsymbol{u}_i s_i$, where s_i is the data symbol and $\boldsymbol{u}_i$ is the *precoder* or *beamforming vector*. We assume the data symbols s_i to be unit variance circularly symmetric Gaussian, so that we can express the achievable rates for the two users as

$$\begin{aligned} R_1 &= \log\left(1 + \frac{|\boldsymbol{h}_{11}^{\mathrm{T}} \boldsymbol{u}_1|^2}{|\boldsymbol{h}_{12}^{\mathrm{T}} \boldsymbol{u}_2|^2 + \sigma_1^2}\right) \quad \text{and} \\ R_2 &= \log\left(1 + \frac{|\boldsymbol{h}_{22}^{\mathrm{T}} \boldsymbol{u}_2|^2}{|\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1|^2 + \sigma_2^2}\right). \end{aligned} \tag{1}$$

Furthermore, we impose a transmit power constraint P_i on the i-th link, i. e.

$$\begin{aligned} \|\boldsymbol{u}_1\|_2^2 &\leq P_1 \quad \text{and} \\ \|\boldsymbol{u}_2\|_2^2 &\leq P_2. \end{aligned} \tag{2}$$

Note that, in contrast to the interference channel in information theory (e. g. [7]), we have simply treated the interference as additional noise, corresponding to the assumption that the receivers do not know the codebook of the respective interferer. Furthermore, note that Gaussian input symbols need not necessarily be information-theoretically optimal, i. e. perhaps higher mutual information between $\boldsymbol{x}_i$ and y_i can be achieved by allowing non-Gaussian distributions of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.

The goal of this work is to design the precoders $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, subject to the transmit power constraints (2), so that R_1 and R_2 are maximized in some way.

3. STRAIGHTFORWARD TRANSMIT STRATEGIES

3.1. Per-User Optimal Beamforming

From the perspective of the first user, the optimal choice of the precoder $\boldsymbol{u}_1$ always is $\sqrt{P_1}/\|\boldsymbol{h}_{11}\|_2 \cdot \boldsymbol{h}_{11}^*$. Similarly, the second user, if he is only concerned about his own rate R_2, will choose $\boldsymbol{u}_2 = \sqrt{P_2}/\|\boldsymbol{h}_{22}\|_2 \cdot \boldsymbol{h}_{22}^*$. For asymptotically low noise power σ_1^2, this means that R_1 converges towards

$$\log\left(1 + \frac{P_1 \|\boldsymbol{h}_{11}\|_2^2 \|\boldsymbol{h}_{22}\|_2^2}{P_2 |\boldsymbol{h}_{12}^{\mathrm{T}} \boldsymbol{h}_{22}^*|^2}\right)$$

assuming that $\boldsymbol{h}_{22}$ and $\boldsymbol{h}_{12}$ are not orthogonal. Similarly, R_2 converges towards a constant value.

For this scheme, the only knowledge required at the i-th transmitter is $\boldsymbol{h}_{ii}$. Note that this strategy is the MISO case of the 'Nash Equilibrium' scheme in [6], as well as the scheme proposed in [5].

3.2. Time-Sharing

Now assume that user two is inactive and user one employs its optimal beamformer. Clearly, R_1 will tend towards infinity with vanishing noise σ_1^2. More precisely, the *multiplexing gain*, defined as

$$\lim_{\sigma_1^{-2} \to \infty} \frac{R_1}{\log \sigma_1^{-2}}$$

can be easily shown to be one.

Therefore, if for very low noise power the two users time-share, i. e. alternatingly transmit and remain silent, they can both attain higher rates than through 'egoistical' behavior (i. e. per-user optimal beamforming). This simple example already shows how cooperation can lead to an improvement of both individual rates.

In addition to the knowledge of the respective desired channel vector $\boldsymbol{h}_{ii}$, the two transmitters also have to have some agreement on how to share the available time.

3.3. Orthogonal Transmission

Next, we construct $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ so that $\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1 = 0$ and $\boldsymbol{h}_{12}^{\mathrm{T}} \boldsymbol{u}_2 = 0$, thereby altruistically sacrificing a part of the own beamforming gain in favor of completely forcing the interference to the other user to zero.

To achieve orthogonality, we first project the optimal beamformer into the nullspace of the interference channel vector

$$\boldsymbol{u}_1' = \boldsymbol{h}_{11}^* - \frac{\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{h}_{11}^*}{\|\boldsymbol{h}_{21}\|_2^2} \boldsymbol{h}_{21}^*$$

and then rescale the result so that the power constraint is fulfilled with equality

$$\boldsymbol{u}_1 = \frac{\sqrt{P_1}}{\|\boldsymbol{u}_1'\|_2} \boldsymbol{u}_1'.$$

The second precoder $\boldsymbol{u}_2$ is constructed the same way.

In contrast to the time-sharing scheme, both users simultaneously achieve multiplexing gain one. Thus, for very low noise power, the individual rates are twice as high as with equal time-sharing. When the noise clearly dominates the interference, however, both the orthogonal scheme and the time-sharing scheme are outperformed by the per-user optimal scheme, as they focus solely on interference-avoidance, even though interference is not the limiting factor. This can also be seen in the simulation results in Section 6.

For the orthogonal scheme, the first user needs to know both $\boldsymbol{h}_{11}$ and $\boldsymbol{h}_{21}$, and the second user must have knowledge of $\boldsymbol{h}_{22}$ and $\boldsymbol{h}_{12}$.

4. THE SUM-RATE CRITERION

So far, we have a scheme that performs well for high noise by egoistically ignoring the interference and a scheme that performs well for low noise by altruistically completely removing the interference. In this section, we will attempt to find an optimal tradeoff between the egoistical and the altruistic strategy that works well regardless of the noise power.

To this end, we introduce the *sum-rate*

$$R_{\text{sum}} = R_1 + R_2$$

as our optimization criterion. This way, the rate degradation caused to the other user is penalized just as much as the improvement of the own rate is rewarded.

The optimization problem reads as

$$\max_{\boldsymbol{u}_1, \boldsymbol{u}_2} R_{\text{sum}} \qquad \text{s.t.:} \quad \|\boldsymbol{u}_1\|_2^2 \leq P_1 \quad \text{and} \quad \|\boldsymbol{u}_2\|_2^2 \leq P_2.$$

and has the KKT optimality conditions

$$\begin{aligned} \frac{\partial R_1}{\partial \boldsymbol{u}_1^*} + \frac{\partial R_2}{\partial \boldsymbol{u}_1^*} &= \lambda_1 \boldsymbol{u}_1 \quad \text{and} \\ \frac{\partial R_1}{\partial \boldsymbol{u}_2^*} + \frac{\partial R_2}{\partial \boldsymbol{u}_2^*} &= \lambda_2 \boldsymbol{u}_2 \end{aligned} \tag{3}$$

with $\lambda_i \geq 0$, $\lambda_i(\|\boldsymbol{u}_i\|_2^2 - P_i) = 0$, and $\|\boldsymbol{u}_i\|_2^2 \leq P_i$, for $i \in \{1, 2\}$. The derivatives can be expressed as

$$\frac{\partial R_1}{\partial \boldsymbol{u}_1^*} = \frac{1}{|\boldsymbol{h}_{11}^{\mathrm{T}} \boldsymbol{u}_1|^2 + |\boldsymbol{h}_{12}^{\mathrm{T}} \boldsymbol{u}_2|^2 + \sigma_1^2} \boldsymbol{h}_{11}^* \boldsymbol{h}_{11}^{\mathrm{T}} \boldsymbol{u}_1 \tag{4}$$

and

$$\begin{aligned} \frac{\partial R_2}{\partial \boldsymbol{u}_1^*} = &\frac{1}{|\boldsymbol{h}_{22}^{\mathrm{T}} \boldsymbol{u}_2|^2 + |\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1|^2 + \sigma_2^2} \boldsymbol{h}_{21}^* \boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1 \\ &- \frac{1}{|\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1|^2 + \sigma_2^2} \boldsymbol{h}_{21}^* \boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1. \end{aligned} \tag{5}$$

with the derivatives with respect to $\boldsymbol{u}_2^*$ resulting from exchanging the indices 1 and 2.

It appears that in general it is not possible to find a closed form expression for a pair of $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ fulfilling the optimality conditions. Moreover, cases can be constructed in which there exist multiple such pairs.

One possibility of dealing with such a constrained optimization problem is to apply a standard gradient projection algorithm. The basic idea is to start with some initialization for $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, such as the zero vectors or the precoders resulting from one of the above straightforward schemes, and to take steps of a certain length in the direction of the steepest ascent of the utility function, i. e. following the gradient of R_{sum} with respect to both $\boldsymbol{u}_1^*$ and $\boldsymbol{u}_2^*$, evaluated at the current value of the precoders. Whenever such a step leads outside the region allowed by the transmit power constraints (as will be the case most of the time, as increasing the transmit power increases the rate), an 'orthogonal projection' maps the precoders back into the allowed region. If the step-size is chosen properly, convergence at least towards a local optimum can be ensured.

Such an algorithm has been discussed in [6] for the more general case of multiple receive antennas. Also, the algorithm works on the covariance matrices, i. e. on the outer products of the precoders. We will use this algorithm as a performance benchmark in Section 6. The practical implementation of such an algorithm, however, requires a centralized instance, connected to both transmitters, that has knowledge of all channel vectors and noise powers.

In the following, we will present an algorithm that also converges towards a pair of $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ that fulfill above optimality conditions, but is realizable in a distributed manner, i. e. by performing alternating computations at the two transmitters without full system knowledge. Moreover, the proposed scheme does not require any step-size considerations.

5. PROPOSED PRICING ALGORITHM

The principle behind iteratively optimizing R_{sum} in a distributed manner is to alternate between an optimization of R_{sum} over $\boldsymbol{u}_1$ while keeping $\boldsymbol{u}_2$ fixed, and vice versa. In the following, we examine only the former optimization, as the latter follows simply from exchanging the indices.

The optimization problem for fixed $\boldsymbol{u}_2 = \bar{\boldsymbol{u}}_2$ reads as

$$\max_{\boldsymbol{u}_1} R_{\text{sum}}\Big|_{\boldsymbol{u}_2 = \bar{\boldsymbol{u}}_2} \qquad \text{s.t.:} \quad \|\boldsymbol{u}_1\|_2^2 \leq P_1$$

and has the optimality condition

$$\left.\frac{\partial R_1}{\partial \boldsymbol{u}_1^*}\right|_{\boldsymbol{u}_2 = \bar{\boldsymbol{u}}_2} + \left.\frac{\partial R_2}{\partial \boldsymbol{u}_1^*}\right|_{\boldsymbol{u}_2 = \bar{\boldsymbol{u}}_2} = \lambda \boldsymbol{u}_1$$

with $\lambda \geq 0$, $\lambda(\|\boldsymbol{u}_1\|_2^2 - P_1) = 0$, and $\|\boldsymbol{u}_1\|_2^2 \leq P_1$. From the expressions for the derivatives in (4) and (5), we see that knowledge of the system parameters of the second link, such as $|\boldsymbol{h}_{22}^{\mathrm{T}} \bar{\boldsymbol{u}}_2|^2$ and σ_2^2, is necessary. In general, there appears to be no closed form solution to this problem either.

In order to simplify the problem, we now employ a linear approximation for the rate contribution of the second link R_2 at the current operating point $(\bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)$:

$$\begin{aligned} R_{\text{sum}} &\approx R_1 + \left.\frac{\partial R_2}{\partial \left(|\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1|^2\right)}\right|_{\substack{\boldsymbol{u}_1 = \bar{\boldsymbol{u}}_1 \\ \boldsymbol{u}_2 = \bar{\boldsymbol{u}}_2}} \cdot \left|\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1\right|^2 + c \\ &= R_1 - \pi_2 \left|\boldsymbol{h}_{21}^{\mathrm{T}} \boldsymbol{u}_1\right|^2 + c \end{aligned}$$

where c is a constant that only depends on the operating point $(\bar{\boldsymbol{u}}_1, \bar{\boldsymbol{u}}_2)$, but not on $\boldsymbol{u}_1$, and the scalar

$$\pi_2 = \frac{1}{|\boldsymbol{h}_{21}^{\mathrm{T}} \bar{\boldsymbol{u}}_1|^2 + \sigma_2^2} - \frac{1}{|\boldsymbol{h}_{22}^{\mathrm{T}} \bar{\boldsymbol{u}}_2|^2 + |\boldsymbol{h}_{21}^{\mathrm{T}} \bar{\boldsymbol{u}}_1|^2 + \sigma_2^2} \tag{6}$$

can be interpreted as the degradation of R_2 per unit increase in interference, i. e. the 'price' for causing interference to the second link. The price π_2 is the only information that needs to be communicated from the second to the first user.

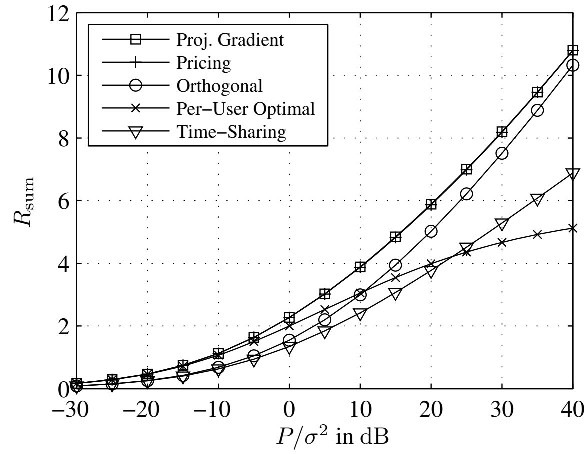


Fig. 2. Sum-Throughput for Two Interfering MISO Channels with $N_1 = N_2 = 2$ Transmit Antennas

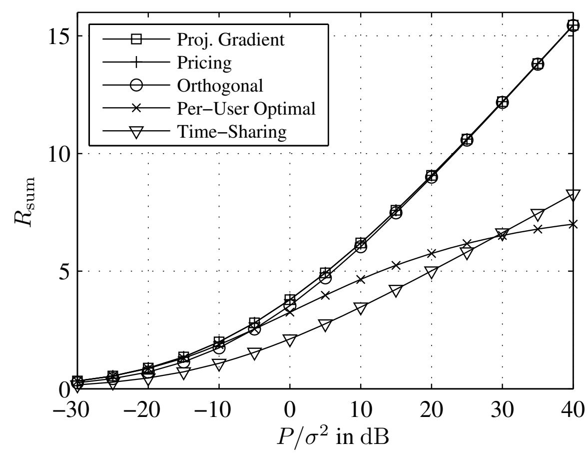


Fig. 3. Sum-Throughput for Two Interfering MISO Channels with $N_1 = N_2 = 4$ Transmit Antennas

The first user now must solve the following optimization problem:

$$\max_{\boldsymbol{u}_1} \log\left(1 + \frac{|\boldsymbol{h}_{11}^{\mathrm{T}}\boldsymbol{u}_1|^2}{|\boldsymbol{h}_{12}^{\mathrm{T}}\bar{\boldsymbol{u}}_2|^2 + \sigma_1^2}\right) - \pi_2 \left|\boldsymbol{h}_{21}^{\mathrm{T}}\boldsymbol{u}_1\right|^2 \quad \text{s. t.: } \|\boldsymbol{u}_1\|_2^2 \le P_1 \tag{7}$$

Assuming linearly independent vectors $\boldsymbol{h}_{11}$ and $\boldsymbol{h}_{21}$, the optimality condition for this problem reads as

$$\left(\frac{1}{|\boldsymbol{h}_{11}^{\mathrm{T}}\boldsymbol{u}_1|^2 + |\boldsymbol{h}_{12}^{\mathrm{T}}\bar{\boldsymbol{u}}_2|^2 + \sigma_1^2}\boldsymbol{h}_{11}^*\boldsymbol{h}_{11}^{\mathrm{T}} - \pi_2\boldsymbol{h}_{21}^*\boldsymbol{h}_{21}^{\mathrm{T}}\right)\boldsymbol{u}_1 = \lambda\boldsymbol{u}_1$$

with $\lambda \ge 0$, $\lambda(\|\boldsymbol{u}_1\|_2^2 - P_1) = 0$, and $\|\boldsymbol{u}_1\|_2^2 \le P_1$. If the power constraint was not fulfilled with equality, we would have $\lambda = 0$, and consequently $\boldsymbol{h}_{11}^{\mathrm{T}}\boldsymbol{u}_1 = 0$ and $\boldsymbol{h}_{21}^{\mathrm{T}}\boldsymbol{u}_1 = 0$. This would lead to zero utility, which cannot be better than choosing a precoder $\boldsymbol{u}_1$ that uses full power and is orthogonal to $\boldsymbol{h}_{21}^*$. We therefore only need to consider the full-power case $\|\boldsymbol{u}_1\|_2^2 = P_1$.

The optimal $\boldsymbol{u}_1$ is the eigenvector corresponding to a positive eigenvalue of a rank two matrix that in turn depends on $|\boldsymbol{h}_{11}^{\mathrm{T}}\boldsymbol{u}_1|^2$. For any value of $|\boldsymbol{h}_{11}^{\mathrm{T}}\boldsymbol{u}_1|^2$ and for $\pi_2 > 0$, it can be seen that the matrix is surely indefinite, i. e. it has exactly one positive and one negative eigenvalue.

In general, there is again no closed form solution fulfilling the optimality conditions. We can, however, construct a fixed point algorithm that may lead to a solution: for some initialization $\boldsymbol{u}_1$, compute the matrix

$$\boldsymbol{A}_1 = \frac{1}{|\boldsymbol{h}_{11}^{\mathrm{T}}\boldsymbol{u}_1|^2 + |\boldsymbol{h}_{12}^{\mathrm{T}}\bar{\boldsymbol{u}}_2|^2 + \sigma_1^2}\boldsymbol{h}_{11}^*\boldsymbol{h}_{11}^{\mathrm{T}} - \pi_2\boldsymbol{h}_{21}^*\boldsymbol{h}_{21}^{\mathrm{T}} \tag{8}$$

and find the eigenvector corresponding to the single positive eigenvector of this matrix. Scale this vector to have norm $\sqrt{P_1}$, resulting in the new $\boldsymbol{u}_1$, and repeat the procedure. When $\boldsymbol{u}_1$ has converged, we have found a solution to the optimization problem (7).

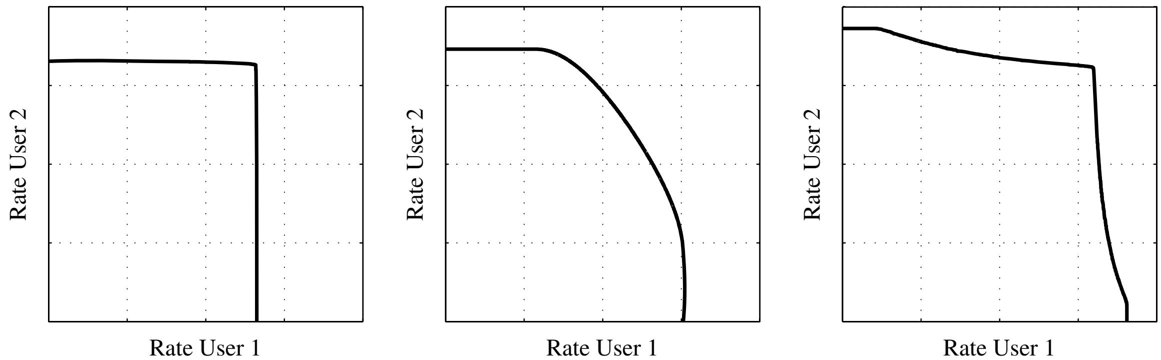
For our distributed algorithm, we only perform one such iteration, as we are not interested in exactly solving (7), which already contains an approximation, but are content with taking a step ‘in the right direction.’

After initializing π_1, π_2, $\boldsymbol{u}_1$, and $\boldsymbol{u}_2$, the pricing algorithm repeats the following steps:

1. transmitter 1 computes $\boldsymbol{A}_1$ using π_2 (cf. Eq. 8)
2. $\boldsymbol{u}_1$ is the eigenvector corresponding to the positive eigenvalue of $\boldsymbol{A}_1$, scaled to have norm $\sqrt{P_1}$
3. π_1 is updated using the new $\boldsymbol{u}_1$ (Eq. 6 with indices 1 and 2 exchanged) and communicated to transmitter 2
4. transmitter 2 computes $\boldsymbol{A}_2$ using π_1 (Eq. 8 with indices 1 and 2 exchanged)
5. $\boldsymbol{u}_2$ is the eigenvector corresponding to the positive eigenvalue of $\boldsymbol{A}_2$, scaled to have norm $\sqrt{P_2}$
6. π_2 is updated using the new $\boldsymbol{u}_2$ (cf. Eq. 6) and communicated to transmitter 1

The value of $|\boldsymbol{h}_{12}^{\mathrm{T}}\bar{\boldsymbol{u}}_2|^2 + \sigma_1^2$, which is needed for the computation of $\boldsymbol{A}_1$, is measured by receiver 1 and reported to transmitter 1. Similarly, $|\boldsymbol{h}_{21}^{\mathrm{T}}\bar{\boldsymbol{u}}_1|^2 + \sigma_2^2$ is reported to transmitter 2 by receiver 2.

When the pricing algorithm reaches a stationary point, the eigenvector of $\boldsymbol{A}_1$ does not change, i. e. the operating point is a solution of (7). At the same time, the eigenvector of $\boldsymbol{A}_2$ does not change, and therefore the second user's equivalent to (7) is also solved. Due to the definition of the π_i as the derivatives evaluated at the current operating point, when problem (7) and the second user's equivalent are simultaneously solved,



(a) High Noise Power (b) Moderate Noise Power (c) Low Noise Power

Fig. 4. Achievable Rate Regions for a Representative Channel Realization with $N_1 = N_2 = 4$ Transmit Antennas

the optimality conditions of the original problem (3) are fulfilled. Therefore, upon convergence, we have found a candidate for a local optimum of the sum-rate maximization problem, just as with the standard gradient algorithm.

6. NUMERICAL RESULTS

For our numerical simulations, we assumed independent, unit-variance, complex Gaussian elements of all channel vectors $\boldsymbol{h}_{ij}$. Therefore, the interference channels with $i \neq j$ were on average as strong as the desired channels with $i = j$. We plotted the sum-rate R_{sum} over the ratio of transmit power to noise P/σ^2, where $P = P_1 = P_2$ and $\sigma^2 = \sigma_1^2 = \sigma_2^2$, averaged over 10000 channel realizations.

In addition to the proposed pricing algorithm, we included the projected gradient search from [6] in our simulations, as well as the schemes discussed in Section 3. As initialization for both algorithms, we used the precoders of the ‘Orthogonal’ scheme. In Fig. 2, the number of transmit antennas is $N_1 = N_2 = 2$, in Fig. 3, it is $N_1 = N_2 = 4$.

We observe that when the noise is dominant and the interference becomes negligible ($P/\sigma^2 < -10\,\mathrm{dB}$), there is not much to be gained by cooperation and the ‘Per-User Optimal’ scheme performs comparatively well. For decreasing noise, however, this scheme converges towards a constant value, and the ‘Time-Sharing’ and ‘Orthogonal’ schemes perform better, as discussed in Section 3. In particular, the multiplexing gain, i. e. the slope of the curve, of the orthogonal scheme is twice as high as for the time-sharing scheme.

The sum-rate oriented schemes (‘Proj. Gradient’ and ‘Pricing’) consistently outperform the straightforward transmission strategies. Remarkably, the performance of the proposed pricing scheme seems to be nearly identical to the performance achieved by the gradient search, which suggests that the two schemes usually converge towards the same local optimum. With an increasing number of transmit antennas, the orthogonal scheme approaches the sum-rate oriented schemes: in Fig. 3, a difference can be observed only for moderate noise levels. This is a consequence of the i. i. d. channel model, where the channel vectors are close to orthogonal anyway most of the time.

Concludingly, the proposed pricing algorithm is the only scheme that performs consistently well regardless of antenna configuration or noise power, and is nonetheless suitable for distributed implementation.

In Fig. 4, a collection of achievable rate regions for one representative channel realization with $N_1 = N_2 = 4$ transmit antennas can be seen. When the noise dominates the interference, the two users hardly influence each other at all, leading to a nearly rectangular region. For decreasing noise powers, we can clearly see that there are straight segments at the edges of the region. These correspond to the case in which one transmitter uses the egoistically optimal beamformer, while the other user employs some precoder that is orthogonal to the interference channel. Finally, for very low noise, we can observe a distinct non-convexity of the rate region.

7. REFERENCES

[1] J. Zander, “Distributed Cochannel Interference Control in Cellular Radio Systems,” *IEEE Trans. Veh. Technol.*, vol. 41, no. 3, pp. 305–311, Aug. 1992.

[2] G. J. Foschini and Z. Miljanic, “A Simple Distributed Autonomous Power Control Algorithm and its Convergence,” *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, Nov. 1993.

[3] S. Stańczak, M. Wiczanowski, and H. Boche, "Distributed Power Control for Optimizing a Weighted Sum of Link-Layer QoS Levels," in *Proc. Globecom '05*, Nov./Dec. 2005, pp. 3539–3543.

[4] J. Huang, R. A. Berry, and M. L. Honig, "Distributed Interference Compensation for Wireless Networks," *IEEE J. Select. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[5] M. F. Demirkol and M. A. Ingram, "Power-Controlled Capacity for Interfering MIMO Links," in *Proc. VTC Fall '01*, Oct. 2001, vol. 1, pp. 187–191.

[6] S. Ye and R. S. Blum, "Optimized Signaling for MIMO Interference Systems With Feedback," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[7] M. Costa, "On the Gaussian Interference Channel," *IEEE Trans. Inform. Theory*, vol. 31, no. 5, pp. 607–615, Sept. 1985.