Receiver-Enhanced Cooperative Spatial Multiplexing with Hybrid Channel Knowledge

Hilde Skjevling
Department of Informatics
University of Oslo
P. O. Box 1080, Blindern
N-0316 Oslo, Norway
Email: hildesk@ifi.uio.no

David Gesbert
Mobile Communications Department
Eurécom Institute
2229 Route des Crêtes, BP 193
F-06904 Sophia Antipolis Cédex, France
Email: gesbert@eurecom.fr

Are Hjørungnes
UniK - University Graduate Center
University of Oslo
Instituttveien 25, P. O. Box 70
N-2027 Kjeller, Norway
Email: arehj@unik.no

***Abstract*— This paper explores the idea of cooperative spatial multiplexing for use in MIMO multicell networks. We imagine applying this cooperation for several multiple antenna access-points to jointly transmit streams towards multiple single-antenna user terminals in neighbouring cells. We make the setting more realistic by introducing a constraint on the *hybrid channel state information* (HCSI), assuming that each transmitter has full CSI for its own channel, but only *statistical* information about other transmitters' channels. Each cooperating transmitter then makes guesses about the behaviour of the other transmitters, using the statistical CSI. We show two of several possible transmission strategies under this setting, and include simple optimization at the receiver to improve performance. Comparisons are made with fully cooperative (full CSI) and non-cooperative schemes. Simulation results show a substantial cooperation gain despite the lack of instantaneous information.**

I. Introduction

The class of so-called cooperative schemes, where two or more transmitters collaborate to improve the quality of transmission toward a common destination, currently sees increased interest. A key scenario has been cooperative diversity, where the devices collaborate to combat the detrimental effects of fading at any one particular device. Typically, the devices are single-antenna user terminals relaying data between a source terminal and the target destination [1], [2], [3]. A specific signalling scheme is distributed Space-Time Block Codes (STBC), where the spatial elements of the codewords are distributed over the antennas of the collaborating transmitters [3], [4], [5]. With space-time coding, the transmitters can operate with limited or no CSI, but it was recently shown that statistical channel information can be very useful here too [6].

We may also employ distributed *spatial multiplexing*, cooperating by using distributed antennas to jointly transmit independent or correlated flows of data. On the receive side, the flows are captured by one multiple-antenna receiver or several distributed (possibly single-antenna) receivers. A typical scenario is in a downlink multicell setting, where multiple base stations want to transmit data to multiple user terminals at once, a problem that is connected to multiuser MIMO. Work in this area include [7], [8], [9], [10]. Finally, one may let each user be served by one transmitter only, using a non-cooperative game approach to mitigate the interference [11].

Unlike with cooperative diversity, cooperative spatial multiplexing or downlink multiuser MIMO in general requires full CSI at the transmitter(s) when the receivers have a single antenna (transmit beamforming only). This means sharing the full, joint multi-user CSI for all transmitters, demanding serious cell-to-cell signalling. Here, we consider instead a *hybrid* CSI scenario where one transmitter has full knowledge of its own CSI, and only statistical knowledge about the other transmitters' channel conditions. We limit ourselves to a two-cell scenario, more general cases are considered in [12]. Our setting is equivalent to a cooperative game, where each base station optimizes a linear spatial filter based on a guess of what the collaborating base is doing simultaneously. We investigate two possible guess strategies and obtain the corresponding linear minimum mean square error (MMSE) transmission schemes. The cooperation gains are substantial, even when using statistical CSI.

This work was supported by the Research Council of Norway and the French Ministry of Foreign Affairs through the Aurora project entitled "Optimization of Broadband Wireless Communications Network".

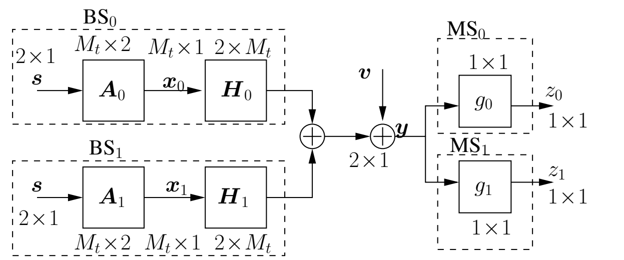

Fig. 1. System studied with two base stations BS$_i$ having M_t antennas and two mobile stations MS$_i$ having one antenna each.

This scenario was first presented in [13], but is extended here to include optimized processing at the receive side. For single-antenna receivers, this processing reduces to a simple scaling, but is beneficial when investigating the MSE performance for high transmit powers. This assumes that the receiving user terminals have access to full CSI, trading a slight increase in complexity for the performance benefit.

The introduction of processing at the receive side means optimizing w.r.t. a joint transmit/receive MMSE criterion. In the transmitters, we iterate between finding the optimal transmit filter given the receive filters and vice versa, until convergence is reached.

II. System Description

We consider downlink communication in a two-cell setting, where two base stations BS$_0$ and BS$_1$, each with M_t antennas, communicate with two single-antenna mobile stations MS$_0$ and MS$_1$. We imagine that MS$_i$ is associated with BS$_i$, it is located within BS$_i$'s cell. However, MS$_i$ also receives part of its data multiplexed from the BS in the neighbouring cell, BS$_j$, $j \neq i$.

In total, two symbols $\{s_0,\, s_1\}$ are sent in the symbol vector

$$\boldsymbol{s} = [s_0,\, s_1]^T, \tag{1}$$

such that $\boldsymbol{s} \in \mathbb{C}^{2\times1}$, with $s_i \in \mathcal{A}$, where $\mathcal{A}$ is constellation alphabet. Symbol s_i is intended for MS$_i$ only, $i \in \{0, 1\}$, but for cooperation purposes, we assume that each BS owns a copy of the global symbol vector $\boldsymbol{s}$ (obtained through relay or fibers etc.). The autocorrelation matrix of the vector $\boldsymbol{s}$ is given by:

$$\boldsymbol{\Phi}_{\boldsymbol{s}} = \mathbb{E}\left[\boldsymbol{s}\boldsymbol{s}^H\right], \tag{2}$$

where $\boldsymbol{\Phi}_{\boldsymbol{s}} \in \mathbb{C}^{2\times 2}$. For independent unit-variance data, $\boldsymbol{\Phi}_{\boldsymbol{s}} = \boldsymbol{I}_2$.

A. Hybrid Channel State Information

We assume flat-fading channel conditions. The channel from BS_i to both MS is given by matrix $\boldsymbol{H}_i \in \mathbb{C}^{2\times M_t}$. *Hybrid channel state information* at the transmitters is considered in the sense that BS_i knows $\boldsymbol{H}_i$ perfectly, but only has access to the *statistics* of $\boldsymbol{H}_j$, $j \neq i$, communicated to it via a low-rate feedback channel. This scenario can be considered realistic, because statistical channel information varies much more slowly than Rayleigh fading, and is easily broadcasted to the various cells.

III. Linear Transmit Filtering and Receive Scaling

Before transmission, each BS_i filters the data vector $\boldsymbol{s}$ with the matrix $\boldsymbol{A}_i \in \mathbb{C}^{M_t\times 2}$, yielding the vector $\boldsymbol{x}_i \in \mathbb{C}^{M_t\times 1}$:

$$\boldsymbol{x}_i = \boldsymbol{A}_i \boldsymbol{s}\,, \quad i \in \{0, 1\}. \tag{3}$$

MS_i receives the scalar y_i, yielding the total signal vector, $\boldsymbol{y} \in \mathbb{C}^{2\times 1}$

$$\boldsymbol{y} = [y_0,\, y_1]^T = \boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{s} + \boldsymbol{H}_1\boldsymbol{A}_1\boldsymbol{s} + \boldsymbol{v}. \tag{4}$$

$\boldsymbol{v} \in \mathbb{C}^{2\times 1}$ represents additive, white, signal-independent noise with

$$\boldsymbol{\Phi}_{\boldsymbol{v}} = \mathbb{E}\left[\boldsymbol{v}\boldsymbol{v}^H\right] \in \mathbb{C}^{2\times 2}. \tag{5}$$

In addition to the transmit filtering [13], we now extend to receiver optimization for each MS_i separately, through use of a scalar coefficient g_i, $i \in \{0, 1\}$. This receive scaling is constructive both under full and hybrid CSI, ensuring a continuous decrease in the MSE when increasing the power used. We define the diagonal matrix $\boldsymbol{G} \triangleq \mathrm{diag}([g_0,\ g_1])$, and the final vector $\boldsymbol{z} \in \mathbb{C}^{2\times 1}$ becomes

$$\boldsymbol{z} = \boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{s} + \boldsymbol{G}\boldsymbol{H}_1\boldsymbol{A}_1\boldsymbol{s} + \boldsymbol{G}\boldsymbol{v}. \tag{6}$$

MS_i is assumed to have full CSI of its own channels, $(\boldsymbol{H}_0)_{i,:}$ and $(\boldsymbol{H}_1)_{i,:}$, but does not need to know the other MSs channels. The transmit filters used in the BSs are assumed known in both MSs, so the optimization of g_i is performed with perfect knowledge of $\boldsymbol{A}_i$, $i \in \{0, 1\}$. The total system in (6) is depicted in Fig. 1.

We consider per-base transmit power constraints on the power P_i:

$$\mathrm{Tr}\left\{\boldsymbol{A}_i\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_i\right\} = P_i, \quad i \in \{0, 1\} \tag{7}$$

IV. Optimal Spatial Filtering

We consider strategies for linear filtering, based on the joint transmit/receive MMSE criterion. As a reference performance bound, we use the optimal case, where full CSI is shared by all BSs.

A. Full CSI Optimal Filtering

Here, $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ and $\boldsymbol{G}$ are all optimized jointly, based on access to full, instantaneous CSI at both BS_i, $i \in \{0, 1\}$. The filters are optimal in the joint MMSE sense where the mean square error (MSE) is

$$\mathrm{MSE} = \mathbb{E}_{\boldsymbol{s},\boldsymbol{v}}\left[\|\boldsymbol{z} - \boldsymbol{s}\|^2\right], \tag{8}$$

which in the full CSI case yields

$$\mathrm{MSE} = \mathrm{Tr}\Big\{\boldsymbol{G}\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}^H + \boldsymbol{G}\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_1^H\boldsymbol{H}_1^H\boldsymbol{G}^H - \boldsymbol{G}\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi}_{\boldsymbol{s}} + \boldsymbol{G}\boldsymbol{H}_1\boldsymbol{A}_1\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}^H + \boldsymbol{G}\boldsymbol{H}_1\boldsymbol{A}_1\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_1^H\boldsymbol{H}_1^H\boldsymbol{G}^H - \boldsymbol{G}\boldsymbol{H}_1\boldsymbol{A}_1\boldsymbol{\Phi}_{\boldsymbol{s}} + \boldsymbol{G}\boldsymbol{\Phi}_{\boldsymbol{v}}\boldsymbol{G}^H - \boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}^H - \boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_1^H\boldsymbol{H}_1^H\boldsymbol{G}^H + \boldsymbol{\Phi}_{\boldsymbol{s}}\Big\}.$$

To optimize the filters under the distributed power constraints, we use the Lagrangian method with the objective function given by

$$L = \mathrm{MSE} + \mu_0\,\mathrm{Tr}\left\{\boldsymbol{A}_0\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_0^H\right\} + \mu_1\,\mathrm{Tr}\left\{\boldsymbol{A}_1\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_1^H\right\}. \tag{9}$$

The final equation for optimal precoders can be shown to be the following (independent of symbol correlation):

$$\begin{bmatrix} \boldsymbol{H}_0^H\boldsymbol{G}^H\boldsymbol{G}\boldsymbol{H}_0 + \mu_0\boldsymbol{I}_{M_t} & \boldsymbol{H}_0^H\boldsymbol{G}^H\boldsymbol{G}\boldsymbol{H}_1 \\ \boldsymbol{H}_1^H\boldsymbol{G}^H\boldsymbol{G}\boldsymbol{H}_0 & \boldsymbol{H}_1^H\boldsymbol{G}^H\boldsymbol{G}\boldsymbol{H}_1 + \mu_1\boldsymbol{I}_{M_t} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{A}_0 \\ \boldsymbol{A}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_0^H\boldsymbol{G}^H \\ \boldsymbol{H}_1^H\boldsymbol{G}^H \end{bmatrix}. \tag{10}$$

Using $\boldsymbol{G}$, we find the matrices $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$, so that the power constraints in (7) are satisfied. The next step is to optimize the g_0 and g_1, found from $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$.

When differentiating L with respect to g_0^* and g_1^*, only the parts of L that contains $\boldsymbol{G}^H$ are of interest, and we create the stripped expression $L_{\boldsymbol{G}}$.

$$L_{\boldsymbol{G}} = \mathrm{Tr}\Big\{\big(\boldsymbol{G}\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_0^H\boldsymbol{H}_0^H + \boldsymbol{G}\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_1^H\boldsymbol{H}_1^H + \boldsymbol{G}\boldsymbol{H}_1\boldsymbol{A}_1\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_0^H\boldsymbol{H}_0^H + \boldsymbol{G}\boldsymbol{H}_1\boldsymbol{A}_1\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_1^H\boldsymbol{H}_1^H + \boldsymbol{G}\boldsymbol{\Phi}_{\boldsymbol{v}} - \boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_0^H\boldsymbol{H}_0^H - \boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_1^H\boldsymbol{H}_1^H\big)\boldsymbol{G}^H\Big\} \triangleq \mathrm{Tr}\left\{\boldsymbol{K}\boldsymbol{G}^H\right\}, \tag{11}$$

and we have

$$\frac{\partial L}{\partial g_i^*} = \frac{\partial L_{\boldsymbol{G}}}{\partial g_i^*} = \frac{\partial}{\partial g_i^*}\Big(g_0^*(\boldsymbol{K})_{0,0} + g_1^*(\boldsymbol{K})_{1,1}\Big). \tag{12}$$

The coefficients g_i, $i \in \{0, 1\}$, $j \in \{0, 1\}$, $j \neq i$, are given by

$$g_i = \left(\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_j^H\boldsymbol{H}_j^H\right)_{i,i} \Big(\big(\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_j^H\boldsymbol{H}_j^H + \boldsymbol{H}_j\boldsymbol{A}_j\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \boldsymbol{H}_j\boldsymbol{A}_j\boldsymbol{\Phi}_{\boldsymbol{s}}\boldsymbol{A}_j^H\boldsymbol{H}_j^H + \boldsymbol{\Phi}_{\boldsymbol{v}}\big)_{i,i}\Big)^{-1}. \tag{13}$$

B. Iterative Search and Receive Optimization

Starting from an initial guess, the two-step process in each BS_i, $i \in \{0, 1\}$ of finding 1) the filtering matrix $\boldsymbol{A}_i$, from (10) and 2) the corresponding $\{g_0, g_1\}$, from (13), iterates until it converges to a solution. The matrix $\boldsymbol{A}_i$ is used as in (3), while discarding the g_i.

Recall that for all cooperation scenarios, the receivers have full CSI and knowledge of all filters $\boldsymbol{A}_i$, $i \in \{0, 1\}$. It is important to note that each MS_i then only needs to know its own channel, $(\boldsymbol{H}_0)_{i,:}$ and $(\boldsymbol{H}_1)_{i,:}$, in order to compute $\{g_0, g_1\}$ from (13). Note that these g_i are, in general, not the same as those found by the BSs.

C. Hybrid CSI Optimal Filtering

In the case of hybrid CSI, the optimal joint MMSE beamformer cannot be obtained. Now, each BS optimizes its linear filter in the MMSE sense, from limited knowledge. We let each BS guess what precoder filter type the other BS uses. We incorporate statistical CSI for two guesses; the transmit MRC and the zero-forcing criteria. In this case with single-antenna receivers, we do not include the receive filter $\boldsymbol{G}$ in the guesses. This choice allows us to keep a simple optimization, while still obtaining good results.

The introduction of guesses means that the iteration described in section IV-B, is performed separately in each BS, yielding different receive scaling matrices $\boldsymbol{G}_i = \mathrm{diag}([g_i^{(0)},\ g_1^{(i)}])$, $i \in \{0, 1\}$. Each receiving MS_i is assumed to have full CSI for its own channel and access to both $\boldsymbol{A}_i$. Then, from (13), the receive coefficient g_i is found.

D. Transmit Maximum Ratio Combining (MRC) Guess

Here, BS_j guesses that BS_j, $i, j \in \{0, 1\}, i \neq j$ uses a scaled matched filter for transmission.

$$\boldsymbol{A}_1 = \sqrt{P_1}\,\frac{\boldsymbol{H}_1^H}{\|\boldsymbol{H}_1\|_F}. \tag{14}$$

Viewed from BS_0, given (8) and (14), the MSE averaged over realizations of the unknown $\boldsymbol{H}_1$ yields

$$\begin{aligned}\mathbb{E}_{\boldsymbol{H}_1}[\text{MSE}] = \text{Tr}\Big\{&\boldsymbol{G}_0\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi_s}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}_0^H - \boldsymbol{G}_0\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi_s} -\\ &\boldsymbol{\Phi_s}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}_0^H + \sqrt{P_1}\boldsymbol{G}_0\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi_s}\boldsymbol{S}_1\boldsymbol{G}_0^H +\\ &\sqrt{P_1}\boldsymbol{G}_0\boldsymbol{S}_1\boldsymbol{\Phi_s}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}_0^H - \sqrt{P_1}\boldsymbol{\Phi_s}\boldsymbol{S}_1\boldsymbol{G}_0^H +\\ &P_1\boldsymbol{G}_0\boldsymbol{W}_1\boldsymbol{G}_0^H - \sqrt{P_1}\boldsymbol{G}_0\boldsymbol{S}_1\boldsymbol{\Phi_s} + \boldsymbol{G}_0\boldsymbol{\Phi_v}\boldsymbol{G}_0^H + \boldsymbol{\Phi_s}\Big\} \end{aligned} \quad (15)$$

where the statistics $\boldsymbol{S}_i$ and $\boldsymbol{W}_i$ are defined as $\boldsymbol{S}_i \triangleq \mathbb{E}\left[\frac{\boldsymbol{H}_i\boldsymbol{H}_i^H}{\|\boldsymbol{H}_i\|_F}\right]$ and $\boldsymbol{W}_i \triangleq \mathbb{E}\left[\frac{\boldsymbol{H}_i\boldsymbol{H}_i^H\boldsymbol{\Phi_s}\boldsymbol{H}_i\boldsymbol{H}_i^H}{\|\boldsymbol{H}_i\|_F^2}\right]$. Estimates of these matrices $\hat{\boldsymbol{S}}_i$ and $\hat{\boldsymbol{W}}_i$ can be found by the estimators:

$$\hat{\boldsymbol{S}}_i = \frac{1}{Q}\sum_{q=0}^{Q-1}\frac{\boldsymbol{H}_i^{(q)}\left(\boldsymbol{H}_i^{(q)}\right)^H}{\|\boldsymbol{H}_i^{(q)}\|_F}$$

$$\hat{\boldsymbol{W}}_i = \frac{1}{Q}\sum_{q=0}^{Q-1}\frac{\boldsymbol{H}_i^{(q)}\left(\boldsymbol{H}_i^{(q)}\right)^H\boldsymbol{\Phi_s}\boldsymbol{H}_i^{(q)}\left(\boldsymbol{H}_i^{(q)}\right)^H}{\|\boldsymbol{H}_i^{(q)}\|_F^2},$$

where $\boldsymbol{H}_i^{(q)}$ is the q-th realization of the channel, and found as:

$$\text{vec}\left(\boldsymbol{H}_i^{(q)}\right) = \boldsymbol{R}_{\boldsymbol{H}_i}^{1/2}\,\text{vec}\left(\boldsymbol{H}_w^{(q)}\right), \quad (16)$$

where $\text{vec}\left(\boldsymbol{H}_w^{(q)}\right) \sim \mathcal{CN}\left(\boldsymbol{0}_{M_tM_r\times 1}, \boldsymbol{1}_{M_tM_r\times 1}\right)$ and $\boldsymbol{R}_{\boldsymbol{H}_i} = \mathbb{E}\left[\text{vec}\left(\boldsymbol{H}_i\right)\text{vec}^H\left(\boldsymbol{H}_i\right)\right]$ is the covariance matrix.

Two Lagrangian multipliers are introduced for the two power constraints and then the optimization with respect to the precoder in BS_0 is performed on the objective function

$$L_0 = \mathbb{E}_{\boldsymbol{H}_1}[\text{MSE}] + \mu_0\,\text{Tr}\left\{\boldsymbol{A}_0\boldsymbol{\Phi_s}\boldsymbol{A}_0^H\right\}. \quad (17)$$

Under the considered cooperation strategy, assuming that the matrix $\boldsymbol{\Phi_s}$ is invertible and using symmetry for BS_1, the optimal $\boldsymbol{A}_i$ at BS_i, $i \in \{0,1\}$ can be written as

$$\boldsymbol{A}_i = \left[\boldsymbol{H}_i^H\boldsymbol{G}_i^H\boldsymbol{G}_i\boldsymbol{H}_i + \mu_i\boldsymbol{I}_{M_t}\right]^{-1}\boldsymbol{H}_i^H\boldsymbol{G}_i^H\left[\boldsymbol{I}_2 - \sqrt{P_j}\boldsymbol{G}_i\boldsymbol{S}_j\right], \quad (18)$$

where we have that $j \in \{0,1\}$ and $j \neq i$. The equations for $\boldsymbol{A}_i$, $i \in \{0,1\}$, can be interpreted as modified MMSE transmit filters, where the modification makes the use of the cooperating base's channel statistics.

When finding the optimal coefficients $g_k^{(i)}$, $i,k \in \{0,1\}$, we see that

$$\frac{\partial L_0}{\partial (g_k^{(i)})^*} = \frac{\partial}{\partial (g_k^{(i)})^*}\left((g_0^{(i)})^*(\boldsymbol{K}^{(0)})_{0,0} + (g_1^{(i)})^*(\boldsymbol{K}^{(0)})_{1,1}\right), \quad (19)$$

where $\boldsymbol{G}_i = \text{diag}([g_0^{(i)},\ g_1^{(i)}])$ and $\boldsymbol{K}^{(i)}$ is such that $\boldsymbol{K}^{(i)}\boldsymbol{G}_i^H$ contains the parts of (15) with $\boldsymbol{G}_i^H$ as the rightmost factor:

$$\begin{aligned}\boldsymbol{K}^{(i)} = &\boldsymbol{G}_i\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \sqrt{P_j}\boldsymbol{G}_j\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi_s}\boldsymbol{S}_j +\\ &\sqrt{P_j}\boldsymbol{G}_i\boldsymbol{S}_j\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + P_j\boldsymbol{G}_i\boldsymbol{W}_j - \boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H -\\ &\sqrt{P_j}\boldsymbol{\Phi_s}\boldsymbol{S}_j + \boldsymbol{G}_i\boldsymbol{\Phi_v}\end{aligned}$$

The above and symmetry together yield what the base stations perceive as the optimal coefficients in $\boldsymbol{G}_i = \text{diag}([g_0^{(i)},\ g_0^{(i)}])$, $i,j,k \in \{0,1\}$, and $j \neq i$:

$$\begin{aligned}g_k^{(i)} = &\left(\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \boldsymbol{\Phi_s}\sqrt{P_j}\boldsymbol{S}_j\right)_{k,k}\cdot\Bigg(\Big[(\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi_s})\\ &(\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \sqrt{P_j}\boldsymbol{S}_j) + \boldsymbol{\Phi_v} + \sqrt{P_j}\boldsymbol{S}_j\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + P_j\boldsymbol{W}_j\Big]_{k,k}\Bigg)^{-1}.\end{aligned}$$

The iterative search and how the receiving MSs find the $\boldsymbol{G}$ that is used, is described in subsection IV-B.

E. Transmit Zero-Forcing Guess

Here, we look at the case where BS_0 assumes that BS_1 uses a scaled-down zero-forcing (ZF), on the form of

$$\boldsymbol{A}_1 = \sqrt{P_1}\frac{\boldsymbol{H}_1^H(\boldsymbol{H}_1\boldsymbol{H}_1^H)^{-1}}{\|\boldsymbol{H}_1^H(\boldsymbol{H}_1\boldsymbol{H}_1^H)^{-1}\|_F}, \quad (20)$$

and vice versa for BS_1's assumptions on BS_0. From (8) and (20), the MSE averaged over all realizations of the unknown $\mathbf{H}_1$, becomes:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{H}_1}[\text{MSE}] = \text{Tr}\big\{&\boldsymbol{G}_0\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi_s}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}_0^H + \\ &\sqrt{P_1}U_1\boldsymbol{G}_0\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi_s}\boldsymbol{G}_0^H - \boldsymbol{G}_0\boldsymbol{H}_0\boldsymbol{A}_0\boldsymbol{\Phi_s} +\\ &\sqrt{P_1}U_1\boldsymbol{G}_0\boldsymbol{\Phi_s}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}_0^H + P_1T_1\boldsymbol{G}_0\boldsymbol{\Phi_s}\boldsymbol{G}_0^H + \boldsymbol{G}_0\boldsymbol{\Phi}_v\boldsymbol{G}_0^H -\\ &\sqrt{P_1}U_1\boldsymbol{G}_0\boldsymbol{\Phi_s} - \boldsymbol{\Phi_s}\boldsymbol{A}_0^H\boldsymbol{H}_0^H\boldsymbol{G}_0^H - \sqrt{P_1}U_1\boldsymbol{\Phi_s}\boldsymbol{G}_0^H + \boldsymbol{\Phi_s}\big\},\end{aligned} \quad (21)$$

where $U_i \triangleq \mathbb{E}\left[\frac{1}{\|\boldsymbol{H}_i^\dagger\|_F}\right]$ and $T_i \triangleq \mathbb{E}\left[\frac{1}{\|\boldsymbol{H}_i^\dagger\|_F^2}\right]$, and $\dagger$ denotes the Moore-Penrose pseudoinverse [14]. In an analogous way as before, finding U_i and T_i can be estimated using

$$\hat{U}_i = \frac{1}{Q}\sum_{q=0}^{Q-1}\frac{1}{\left\|(\boldsymbol{H}_i^{(q)})^\dagger\right\|_F}, \quad \hat{T}_i = \frac{1}{Q}\sum_{q=0}^{Q-1}\frac{1}{\left\|(\boldsymbol{H}_i^{(q)})^\dagger\right\|_F^2} \quad (22)$$

We obtain an objective function L similar to (17). Assuming an invertible $\boldsymbol{\Phi}_s$ and symmetry, the optimal $\boldsymbol{A}_i$, $i,j \in \{0,1\}$, $j \neq i$, is:

$$\boldsymbol{A}_i = \left[\boldsymbol{H}_i^H\boldsymbol{G}_i^H\boldsymbol{G}_i\boldsymbol{H}_i + \mu_i\boldsymbol{I}_{M_t}\right]^{-1}\boldsymbol{H}_i^H\boldsymbol{G}_i^H\left(\boldsymbol{I}_2 - \sqrt{P_j}U_j\boldsymbol{G}_i\right)$$

These results again correspond to *modified* MMSE filters; incorporarting the statistics of the unknown channel and the guess on the filtering strategy of the cooperating base. We find the optimal coefficients $g_k^{(i)}$, $i,k \in \{0,1\}$, as in (19) using a $\boldsymbol{K}^{(i)}$ such that $\boldsymbol{K}^{(i)}\boldsymbol{G}_i^H$ contains the parts of (21) with $\boldsymbol{G}_i^H$ as the rightmost factor.

$$\begin{aligned}\boldsymbol{K}^{(i)} = &\boldsymbol{G}_i\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H - \boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \sqrt{P_j}U_j\boldsymbol{G}_i\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi_s} +\\ &\sqrt{P_j}U_j\boldsymbol{G}_i\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + P_jT_j\boldsymbol{G}_i\boldsymbol{\Phi_s} + \boldsymbol{G}_i\boldsymbol{\Phi}_v - \sqrt{P_j}U_j\boldsymbol{\Phi_s}.\end{aligned}$$

Then, for $i,j,k \in \{0,1\}$, $j \neq i$ and $\boldsymbol{G}_i = \text{diag}([g_0^{(i)},\ g_1^{(i)}])$, we find the optimal receive scaling coefficients

$$\begin{aligned}g_k^{(i)} = &\left(\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + \sqrt{P_j}U_j\boldsymbol{\Phi_s}\right)_{k,k}\big(\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H +\\ &\sqrt{P_j}U_j\boldsymbol{H}_i\boldsymbol{A}_i\boldsymbol{\Phi_s} + \sqrt{P_j}U_j\boldsymbol{\Phi_s}\boldsymbol{A}_i^H\boldsymbol{H}_i^H + P_jT_j\boldsymbol{\Phi_s} + \boldsymbol{\Phi}_v\big)_{k,k}^{-1}.\end{aligned}$$

Again, the iterative search and how the receiving MSs find the scaling matrix $\boldsymbol{G}$, is described in subsection IV-B.

F. Non-Cooperative Case

We assume partial CSI such that BS_i knows $(\boldsymbol{H}_i)_{i,:}$, the 1×2 channel from BS_i to MS_i, $i \in \{0,1\}$. Seen from BS_0, after reception and scaling, the total signal $\boldsymbol{z}$ is assumed to be

$$\boldsymbol{z} = \begin{bmatrix} g_0^{(0)}(\boldsymbol{H}_0)_{0,:}\boldsymbol{A}_0\boldsymbol{s} + g_0^{(0)}v_0 \\ g_1^{(0)}(\boldsymbol{H}_1)_{1,:}\boldsymbol{A}_1\boldsymbol{s} + g_1^{(0)}v_1\,. \end{bmatrix}, \quad (23)$$

The MSE seen from BS_0 is then found from (8) and we optimize the filters under the distributed power constraints, using an objective function on the form of (17). From this, BS_i, $i \in \{0,1\}$ arrive at the following conditions for optimality:

$$\boldsymbol{A}_i = \left(|g_i^{(i)}|^2(\boldsymbol{H}_i)_{i,:}^H(\boldsymbol{H}_i)_{i,:} + \mu_i\boldsymbol{I}_{M_t}\right)^{-1}(g_i^{(i)})^*(\boldsymbol{H}_i)_{i,:}^H\boldsymbol{\Phi}_{s_i\boldsymbol{s}}\boldsymbol{\Phi_s}^{-1},$$

Correspondingly, each BS_i, arrives at a scaling matrix $\boldsymbol{G}_i = \text{diag}([g_0^{(i)},\ g_1^{(i)}])$, assumed to be used in the receivers, $i,k \in \{0,1\}$:

$$g_k^{(i)} = \left(\boldsymbol{\Phi}_{s_i\boldsymbol{s}}\boldsymbol{A}_i^H(\boldsymbol{H}_i)_{i,:}^H\right)_{k,k}\left((\boldsymbol{H}_i)_{i,:}\boldsymbol{A}_i\boldsymbol{\Phi_s}\boldsymbol{A}_i^H(\boldsymbol{H}_i)_{i,:}^H + \sigma_v^2\right)_{k,k}^{-1},$$

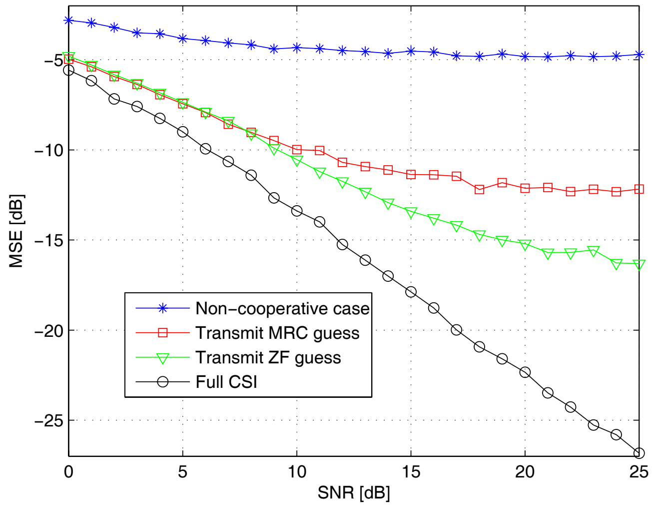

Fig. 2. MSE versus SNR for cooperative transmission, using two two-antenna BSs and two single-antenna MSs, using noise variance $\sigma_v^2 = 0.05$ and $r_{\text{ipl}} = 0$ dB, averaged over 1000 channel realizations.

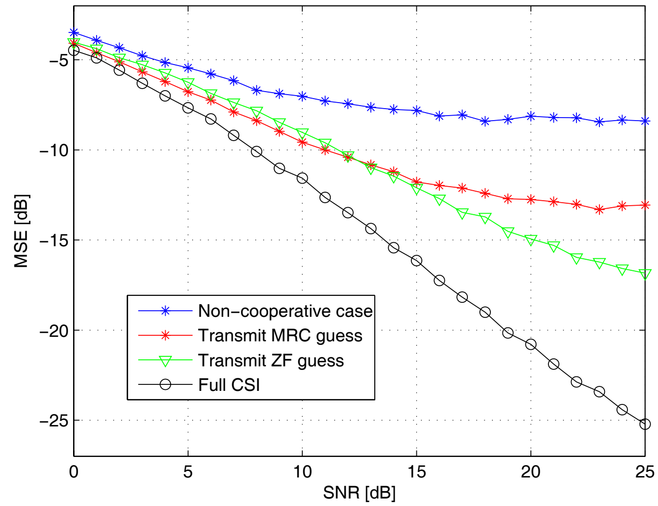

Fig. 3. MSE versus SNR for cooperative transmission, using two two-antenna BSs and two single-antenna MSs, using noise variance $\sigma_v^2 = 0.05$ and $r_{\text{ipl}} = 3$ dB, averaged over 1000 channel realizations.

As before, the iterative search and how the receiving MSs find the $\boldsymbol{G}$ that is used, is described in subsection IV-B.

V. Simulations

Our numerical simulation system consists of two transmitting BSs and two receiving MSs. Each BS is equipped with $M_t = 2$ antennas, while the MSs have only $M_r = 1$ antenna.

We obtain the MSE versus signal-to-noise ratio (SNR), for the four described approaches at optimizing the transmit and receive filters $\boldsymbol{A}_i$, $i \in \{0, 1\}$ and $\boldsymbol{G}$, each using different levels of CSI at the BSs. We define the SNR as SNR $\triangleq P/\sigma_v^2$, where $P = P_0 = P_1$ is the power available at each BS and the noise variance is fixed as $\sigma_v^2 = 0.05$, $\boldsymbol{\Phi_v} = \sigma_v^2 \boldsymbol{I}_2$. The MSE per source signal at each SNR is averaged over 1000 channel realizations.

The intercell loss ratio r_{ipl} is defined as the ratio between own-cell and inter-cell average channel gain. In Fig. 2, $r_{\text{ipl}} = 0$ dB, signals from both BSs experience the same average large-scale path loss on the way to *both* MS. In Fig. 3, we have $r_{\text{ipl}} = 3$ dB.

In Fig. 2, the best MSE-results are obtained by the approach using full CSI. As expected the proposed hybrid CSI schemes yield performance in between the optimal and the non-cooperative scenarios. In the high SNR region, playing the cooperative game using the ZF assumption yields better result than assuming transmit MRC, as the optimal MMSE filter is closer to the ZF in moderate to high SNR levels.

In Fig. 3, we see that more intercell loss hurts the full CSI scheme, because this means less total power received. The approaches using partial CSI, with the transmit MRC or transmit ZF assumptions, are not affected much, we observe a slight improvement at high SNR. The non-cooperative approach, however, clearly benefits from this increased path loss. This makes intuitive sense, as increased intercell path loss means that cooperation between cells is less important.

VI. Conclusions

We investigate spatial multiplexing signalling between two cooperating base stations communicating each with users in two neighbouring cells. We propose a practical transmission strategy that exploit hybrid (mixed instantaneous and statistical) channel state information at the transmitters.

We show that cooperation is possible thanks to a game scenario where each base station makes certain assumptions about the behaviour of the cooperating base in terms of the spatial filter used.

References

[1] T. M. Cover and A. A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, Sept. 1979.

[2] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity–Part I: System Description, Part II: Implementation Aspects and Performance Analysis," *IEEE Trans. Commun.*

[3] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, Oct. 2003.

[4] S. Barbarossa, L. Pescosolido, D. Ludovici, L. Barbetta, and G. Scutari, "Cooperative wireless networks based on distributed space-time coding." Proc. of Int. workshop on wireless ad-hoc netw. (IWWAN), June 2004.

[5] X. Li, "Space-time coded multi-transmission among distributed transmitters without perfect synchronization." *IEEE Signal Processing Letters*, vol. 11, no. 12, 2004.

[6] H. Skjevling, D. Gesbert, and A. Hjørungnes, "Precoding for distributed space-time codes in cooperative diversity-based downlink," to appear in the International Conference on Communication (ICC) 2006.

[7] C. B. Peel, Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "An Introduction to the Multi-User MIMO Downlink," *IEEE Commun. Mag.*, Oct. 2004.

[8] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end." Vehicular Technology Conference (VTC), Spring 2001.

[9] H. Zhang, H. Dai, and Q. Zhou, "Base station cooperation for multiuser MIMO: Joint transmission and BS selection." Conference on information sciences and systems, Mar. 2004.

[10] B. L. Ng, J. S. Evans, S. V. Hanly, and D. Aktas, "Transmit beamforming with cooperative base stations." IEEE Int. Symp. on Inf. Theory (ISIT), Sept. 2005.

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

[12] H. Skjevling, D. Gesbert, and A. Hjørungnes, "Cooperative Spatial Multiplexing with Hybrid Channel Knowledge," in preparation, to be submitted to *IEEE Trans. Wireless Communications*.

[13] D. Gesbert, A. Hjørungnes, and H. Skjevling, "Cooperative spatial multiplexing with hybrid channel knowledge." To appear in the International Zürich Seminar on Broadband Communications (IZS) 2006.

[14] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge Univ. Press, 1985.