

# From MU Massive MISO to Pathwise MU Massive MIMO

Yohan Lejosne<sup>‡¶</sup>, Manijeh Bashar<sup>‡</sup>, Dirk Slock<sup>‡</sup>, Yi Yuan-Wu<sup>¶</sup>

<sup>‡</sup>EURECOM, Sophia-Antipolis, France, Email: {bashar,lejosne,slock}@eurecom.fr

<sup>¶</sup>Orange Labs, Issy-les-Moulineaux, France, Email: {yohan.lejosne,yi.yuan}@orange.com

**Abstract**—The Interfering Broadcast Channel (IBC) applies to the downlink of cellular and heterogeneous networks, which are limited by multi-user (MU) interference. The interference alignment (IA) concept has shown that interference does not need to be inevitable. In particular spatial IA in MIMO IBC allows for low latency. However, IA requires perfect and typically global Channel State Information at the Transmitter(s) (CSIT), whose acquisition does not scale with network size. Hence, designs that are optimal in terms of Degrees of Freedom (DoF) may not be so in terms of more relevant net DoF, accounting for CSI acquisition or at finite SNR. Also, the design of transmitters (Tx) and receivers (Rx) is coupled and hence needs to be centralized or duplicated. Here we propose to take advantage of Massive MIMO simplifications, esp. for mmWave, by considering (multi-)path CSIT for crosslinks, which can be obtained without feedback. We consider a hierarchical cross/direct link beamformer design, maximizing Weighted Sum Rate (WSR) with partial CSIT at finite SNR, requiring on local CSIT. We also point out the use of receive antennas in genuine MU Massive MIMO.

**Index Terms**—Interfering Broadcast Channel (IBC), Channel State Information at the Transmitter (CSIT), Massive MIMO, mmWave, pathwise channel models

## I. INTRODUCTION

In this paper, Tx may denote transmit/transmitter/transmission and Rx may denote receive/receiver/reception. Interference is the main limiting factor in wireless transmission. Base stations (BSs) disposing of multiple antennas are able to serve multiple Mobile Terminals (MTs) simultaneously, which is called Spatial Division Multiple Access (SDMA) or Multi-User (MU) MIMO. However, MU systems have precise requirements for Channel State Information at the Tx (CSIT) which is more difficult to acquire than CSI at the Rx (CSIR). Hence we focus here on the more challenging downlink (DL). In cellular systems, one can distinguish between the cell center where a single cell design is appropriate (due to high SIR) and the cell edge where a multi-cell (MC) approach is mandatory. The MU MIMO DL problem for the cell center users is called the (MIMO) Broadcast Channel (BC). For the cell edge users, the recent introduction of Interference Alignment (IA) has shown that approaching high system capacity through aggressive frequency reuse should in principle be possible. Whereas precise capacities for cellular systems remain unknown, IA allows to reach the optimal high SNR rate prelog, called Degree of Freedom (DoF) (or spatial multiplexing factor), that is before accounting for CSI acquisition. The transmitter coordination required for DL IA in a MU multi-cell (MC) setting corresponds to the Interfering Broadcast Channel (IBC).

The main difficulty in realizing linear IA for MIMO I(B)C is that the design of any BS Tx filter depends on all Rx filters whereas in turn each Rx filter depends on all Tx filters [1]. As a result, all Tx/Rx filters are globally coupled and their design requires global CSIT. To carry out this Tx/Rx design in a distributed fashion, global CSIT is required at all BS [2]. The overhead required for this global distributed CSIT is substantial, even if done optimally, leading to substantially reduced Net DoF [3]. In [4] the simplified SIMO uplink (UL) problem with only CSIR acquisition is considered and it is shown that it is impossible to maintain any positive Net DoF if one wants to design a cellular network that extends infinitely far (the problem arises already at a finite network size that depends on the Doppler bandwidth). However, all these DoF considerations may be of limited relevance for operation at any finite SNR, in which case interfering Tx that are sufficiently far away can be ignored, leading to the topic of *topological* IA [5]. This is one of the angles we shall consider here as in [6] to allow to decompose the global Tx design problem.

The recent development of Massive MIMO [7] opens new possibilities for increased system capacity while at the same time simplifying system design. From a DoF point of view it may seem like a suboptimal use of antennas. However, as shown in [8], section V, Fig. 6, the (Massive MIMO asymptotics based analytical expression for the) optimal number of users decreases below the DoF as the SNR decreases. Furthermore, Net DoF considerations and CSI acquisition make the optimal number of users decrease further. In [9], [10], MISO was considered in a single cell. Statistical CSIT between user groups was considered and instantaneous CSIT within user groups. The hypothesis is that some users overlap strongly in terms of covariance subspaces but not in terms of instantaneous CSIT. However, users can be differentiated with instantaneous CSIT iff they can be differentiated pathwise. Also, we advocate the use of 2D antenna arrays for 3D beamforming (BF) which should enhance path differentiability. In any case, some paths coinciding between users can be tolerated as discussed later. In [11] a hierarchical approach is considered that will be taken up here also. MISO is considered and (high SNR based) user selection also. Intercell zero-forcing (ZF) BF is considered based on statistical CSIT, treating interfering links in a binary fashion (either ZF or ignore). Intracell BF is based on instantaneous CSIT and performs Regularized-ZF, which is claimed to be asymptotically optimal (which is only true for uniform user power profile). In [12], following up

on work in [13], beamspace processing is proposed, which is the basic form of hierarchical BF. As argued in [10] also, mmWave communications, which we target here also, facilitate Massive MIMO, and lead to a limited number of dominant paths as they approach optics. The issue with beamspace may appear to be the proper aligning of the discretized main beams. However, the problem is more the limited sidelobe attenuation towards cross links. Nevertheless, beamspace may present an interesting low complexity approach.

What is known as Massive MIMO is more appropriately called MU Massive MISO whereas here we consider actual MU MC Massive MIMO. Here we consider path CSIT for the cross links. The handling of the cross links allows to transform the system into parallel single-user MIMO links. For the direct links we either consider instantaneous CSIT, or also path CSIT with then space-time coding for diversity exploitation. Path CSIT can be obtained without any feedback! It only requires calibration, be it for channel reciprocity in TDD or for the antenna array response in terms of DoA when in FDD. And of course uplink channel estimation and decomposition in terms of path contributions. Whereas we show that path CSIT by itself may allow zero-forcing, which is of interest at high SNR, we are particularly concerned here with maximum Weighted Sum Rate (WSR) designs accounting for finite SNR and possibly for residual Rayleigh channel components beyond the dominant paths accounted for, through partial CSIT. The Massive MIMO setting leads to simple mutual information limits for WSR with partial CSIT. In summary, Massive MIMO makes the pathwise approach viable, with the following implications: the (cross-link) BF can be updated at a reduced (slow fading) rate, parsimonious channel representation facilitates not only uplink but especially downlink channel estimation, the cross-link BF can be used to significantly improve the downlink direct link channel estimates, minimal feedback can be introduced to perform meaningful WSR optimization at a finite SNR (whereas zero-forcing requires much less coordination).

## II. CHANNEL (INFORMATION) MODELS

In this section we drop the user index  $k$  for simplicity.

### A. Specular Wireless MIMO Channel Model

We then get for the MIMO channel transfer matrix at any particular subcarrier of a given OFDM symbol

$$\mathbf{H} = \sum_{i=1}^{N_p} A_i e^{j\psi_i} \mathbf{h}_r(\phi_i) \mathbf{h}_t^T(\theta_i) = \mathbf{B} \mathbf{A}^H \quad (1)$$

where there are  $N_p$  (specular) pathwise contributions with

- $A_i > 0$ : path amplitude
- $\theta_i$ : direction of departure (AoD)
- $\phi_i$ : direction of arrival (AoA)
- $\mathbf{h}_t(\cdot)$ ,  $\mathbf{h}_r(\cdot)$ :  $M/N \times 1$  Tx/Rx antenna array response

and

$$\mathbf{B} = [\mathbf{h}_r(\phi_1) \mathbf{h}_r(\phi_1) \cdots] \begin{bmatrix} e^{j\psi_1} \\ e^{j\psi_2} \\ \vdots \end{bmatrix}, \mathbf{A}^H = \begin{bmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{h}_t^T(\theta_1) \\ \mathbf{h}_t^T(\theta_2) \\ \vdots \end{bmatrix} \quad (2)$$

The antenna array responses are just functions of angles AoD, AoA in the case of standard antenna arrays with scatterers in the far field. In the case of distributed antenna systems, the array responses become a function of all position parameters of the path scatterers. The fast variation of the phases  $\psi_i$  (due to Doppler) and possibly the variation of the  $A_i$  (when the nominal path represents in fact a superposition of paths with similar parameters) correspond to the fast fading. All the other parameters vary on a slower time scale and correspond to slow fading.

### B. Mean and Covariance Gaussian CSIT

Mean information about the channel can come from channel feedback or reciprocity, and prediction, or it may correspond to the non fading (e.g. LoS) part of the channel (note that an unknown phase factor  $e^{j\phi}$  in the overall channel mean does not affect the BF design). Covariance information may correspond to channel estimation (feedback, prediction) errors and/or to information about spatial correlations. Given only mean and (separable) covariance information, the fitting maximum entropy distribution is Gaussian. Hence consider  $\text{vec}(\mathbf{H}) \sim \mathcal{CN}(\text{vec}(\bar{\mathbf{H}}), \mathbf{C}_t^T \otimes \mathbf{C}_r)$  which can be rewritten as

$$\mathbf{H} = \bar{\mathbf{H}} + \mathbf{C}_r^{1/2} \tilde{\mathbf{H}} \mathbf{C}_t^{1/2} \quad (3)$$

where  $\mathbf{C}_r^{1/2}$ ,  $\mathbf{C}_t^{1/2}$  are Hermitian square-roots of the Rx and Tx side covariance matrices

$$\begin{aligned} \mathbf{E}(\mathbf{H} - \bar{\mathbf{H}})(\mathbf{H} - \bar{\mathbf{H}})^H &= \text{tr}\{\mathbf{C}_t\} \mathbf{C}_r \\ \mathbf{E}(\mathbf{H} - \bar{\mathbf{H}})^H(\mathbf{H} - \bar{\mathbf{H}}) &= \text{tr}\{\mathbf{C}_r\} \mathbf{C}_t \end{aligned} \quad (4)$$

and the elements of  $\tilde{\mathbf{H}}$  are i.i.d.  $\sim \mathcal{CN}(0, 1)$ . In what follows, it will also be of interest to consider the total Tx side correlation matrix

$$\mathbf{R}_t = \mathbf{E} \mathbf{H}^H \mathbf{H} = \bar{\mathbf{H}}^H \bar{\mathbf{H}} + \text{tr}\{\mathbf{C}_r\} \mathbf{C}_t. \quad (5)$$

Note that the Gaussian CSIT model could be considered an instance of Ricean fading in which the ratio  $\text{tr}\{\bar{\mathbf{H}}^H \bar{\mathbf{H}}\} / (\text{tr}\{\mathbf{C}_r\} \text{tr}\{\mathbf{C}_t\})$  could be considered the Ricean factor.

### C. Dominant Paths Partial CSIT Channel Model

Assuming the Tx disposes of not much more than the information about  $r$  dominant path AoDs, we shall consider the following MIMO (Ricean) channel model

$$\mathbf{H} = \mathbf{B} \mathbf{A}^H(\theta) + \sqrt{\beta} \tilde{\mathbf{H}}' \quad (6)$$

which follows from (1), (2) except restricted to the  $r$  strongest paths, with the rest modeled by  $\sqrt{\beta} \tilde{\mathbf{H}}'$  (elements i.i.d.  $\sim \mathcal{CN}(0, \beta)$ , independent of the  $\psi_i$ ). Averaging of the path phases  $\psi_i$ , we get for the Tx side covariance matrix

$$\mathbf{C}_t = \mathbf{A} \mathbf{A}^H + N\beta \mathbf{I}_M \quad (7)$$

since due to the normalization of the antenna array responses,  $\mathbf{E} \mathbf{B}^H \mathbf{B} = \text{diag}\{[\mathbf{h}_r(\phi_1) \mathbf{h}_r(\phi_1) \cdots]^H [\mathbf{h}_r(\phi_1) \mathbf{h}_r(\phi_1) \cdots]\} = \mathbf{I}$ . Note that  $\mu = \text{tr}\{\mathbf{A} \mathbf{A}^H\} / \beta N M$  could be considered a Ricean factor. When needed, we may also consider the  $\mathbf{h}_r$ ,

the columns of  $\mathbf{B}$ , to be isotropically distributed. Note that the rank of  $\mathbf{A}\mathbf{A}^H$  can be substantially less than the number of paths. Consider e.g. a cluster of paths with narrow AoD spread, then we have  $\theta_i = \theta + \Delta\theta_i$  where  $\theta$  is the nominal AoD and  $\Delta\theta_i$  is small. Hence

$$\mathbf{h}_t(\theta_i) \approx \mathbf{h}_t(\theta) + \Delta\theta_i \dot{\mathbf{h}}_t(\theta). \quad (8)$$

Such a cluster of paths only adds a rank 2 contribution to  $\mathbf{A}\mathbf{A}^H$ . This Taylor series modeling of clusters is in contrast to the uniform DoA profile used in [10], [14].

### III. STREAMWISE IBC SIGNAL MODEL

In the rest of this paper we shall consider a per stream approach (which in the perfect CSI case would be equivalent to per user). In an IBC formulation, one stream per user can be expected to be the usual scenario. In the development below, in the case of more than one stream per user, treat each stream as an individual user. So, consider again an IBC with  $C$  cells with a total of  $K$  users. We shall consider a system-wide numbering of the users. User  $k$  is served by BS  $b_k$ . The  $N_k \times 1$  received signal at user  $k$  in cell  $b_k$  is

$$\mathbf{y}_k = \underbrace{\mathbf{H}_{k,b_k} \mathbf{g}_k x_k}_{\text{signal}} + \underbrace{\sum_{\substack{i \neq k \\ b_i = b_k}} \mathbf{H}_{k,b_k} \mathbf{g}_i x_i}_{\text{intracell interf.}} + \underbrace{\sum_{j \neq b_k} \sum_{i: b_i = j} \mathbf{H}_{k,j} \mathbf{g}_i x_i}_{\text{intercell interf.}} + \mathbf{v}_k \quad (9)$$

where  $x_k$  is the intended (white, unit variance) scalar signal stream,  $\mathbf{H}_{k,b_k}$  is the  $N_k \times M_{b_k}$  channel from BS  $b_k$  to user  $k$ . We considering a noise whitened signal representation so that we get for the noise  $\mathbf{v}_k \sim \mathcal{CN}(0, I_{N_k})$ . The  $M_{b_k} \times 1$  spatial Tx filter of beamformer (BF) is  $\mathbf{g}_k$ . Treating interference as noise, user  $k$  will apply a linear Rx filter  $\mathbf{f}_k$  to maximize the signal power (diversity) while reducing any residual interference that would not have been (sufficiently) suppressed by the BS Tx. The Rx filter output is  $\hat{x}_k = \mathbf{f}_k^H \mathbf{y}_k$

$$\begin{aligned} \hat{x}_k &= \mathbf{f}_k^H \mathbf{H}_{k,b_k} \mathbf{g}_k x_k + \sum_{\substack{(i,j)=(1,1), \\ \neq (k,b_k)}}^{(K,C)} \mathbf{f}_k^H \mathbf{H}_{k,j} \mathbf{g}_i x_i + \mathbf{f}_k^H \mathbf{v}_k \\ &= \mathbf{f}_k^H \mathbf{h}_{k,k} x_k + \sum_{i \neq k} \mathbf{f}_k^H \mathbf{h}_{k,i} x_i + \mathbf{f}_k^H \mathbf{v}_k \end{aligned} \quad (10)$$

where  $\mathbf{h}_{k,i} = \mathbf{H}_{k,b_i} \mathbf{g}_i$  is the channel-Tx cascade vector. ZF feasibility, in particular also when only based on Tx side covariance CSIT, has been discussed in [15], where also the role of Rx antennas is highlighted.

### IV. PARAMETERIZING MMSE FILTERS IN TERMS OF ZF

The linear Rx maximizing the rate of a user is the MMSE Rx. Let us focus on the Rx for user  $k$ . In this section, assume that  $N \geq K$ . Let  $\mathbf{h}_{\bar{k}} = \mathbf{h}_{\bar{k},1:K-1} = [\mathbf{h}_{k,i}, i = 1, \dots, K, \neq k]$  and let in this section  $\mathbf{h}_k = \mathbf{h}_{k,k}$  and  $\mathbf{H} = [\mathbf{h}_k \mathbf{h}_{\bar{k}}]$  (we shall omit some subscripts to simplify notation). The MMSE Rx can be written as

$$\mathbf{f} = (\mathbf{H}\mathbf{H}^H + I)^{-1} \mathbf{h}_k = (\mathbf{H}\mathbf{H}^H + I)^{-1} \mathbf{H}\mathbf{e} = \mathbf{H}(\mathbf{H}^H \mathbf{H} + I)^{-1} \mathbf{e} \quad (11)$$

where  $\mathbf{e} = [1 \ 0 \ 0 \ \dots]^H$ . Let us introduce the projection matrices  $P_{\mathbf{H}} = \mathbf{H}(\mathbf{H}^H \mathbf{H})^\# \mathbf{H}^H$ ,  $P_{\mathbf{H}}^\perp = I - P_{\mathbf{H}}$  onto the column space of  $\mathbf{H}$  or its orthogonal complement (and  $(\cdot)^\#$  denotes Moore-Penrose pseudo-inverse). In [16], it was shown for the case  $K = 2$  that the MMSE Rx filter (or optimal Tx filter actually, see further) can be written as a linear combination of the matched filter and the ZF filter. This result was generalized in [17] to general  $K$  by showing that the MMSE filter is a linear combination of the matched filter and  $K-1$  ZF filters, obtained by ZF to one of the  $K-1$  interferers. While this result is correct, this generalization is of limited interest. We present here a more interesting interpretation which will allow us to approximate the MMSE filter by one ZF filter, at any given SNR.

**Theorem 1: Parameterization of a MU MMSE filter in terms of ZF filters.** A MMSE filter can be parameterized as a linear combination of all ZF filters of increasing order:

$$\mathbf{f} = \sum_{i=0}^{K-1} \alpha_i P_{\mathbf{h}_{\bar{k},1:i}}^\perp \mathbf{h}_k \quad (12)$$

where  $P_{\mathbf{h}_{\bar{k},1:0}}^\perp \mathbf{h}_k = \mathbf{h}_k$  is the matched filter and  $\boldsymbol{\alpha} = [\alpha_0 \ \dots \ \alpha_{K-1}]^T$  is the vector of combination coefficients.  $\square$

*Proof:* To link MMSE and ZF designs, consider the usual introduction of extended channel vectors  $\mathbf{H}'$ . Then we get from (11) and the inverse of partitioned matrices

$$\begin{aligned} \mathbf{H}' &= \begin{bmatrix} \mathbf{H} \\ I \end{bmatrix} \Rightarrow \mathbf{f} = \mathbf{H}(\mathbf{H}'^H \mathbf{H}')^{-1} \mathbf{e} \\ &= [\mathbf{h}_k \ \mathbf{h}_{\bar{k}}] \begin{bmatrix} 1 \\ -(\mathbf{h}_{\bar{k}}^H \mathbf{h}_{\bar{k}} + I)^{-1} \mathbf{h}_{\bar{k}}^H \mathbf{h}_k \end{bmatrix} \|P_{\mathbf{h}_{\bar{k}}}^\perp \mathbf{h}_k\|^{-2} \end{aligned} \quad (13)$$

The ZF Rx is obtained by omitting the “ $+I$ ” regularization in the denominator(s):  $f^{ZF} = \|P_{\mathbf{h}_{\bar{k}}}^\perp \mathbf{h}_k\|^{-2} P_{\mathbf{h}_{\bar{k}}}^\perp \mathbf{h}_k$ .

Introduce the LDU factorization  $\mathbf{h}_{\bar{k}}^H \mathbf{h}_{\bar{k}} = \mathbf{L}\mathbf{D}\mathbf{L}^H$  where  $\mathbf{D}$  is diagonal with positive real diagonal elements and  $\mathbf{L}$  is lower triangular with unit diagonal, and consider the result  $\tilde{\mathbf{h}}_{\bar{k}}$  of applying Gram-Schmidt orthogonalization to the consecutive columns of  $\mathbf{h}_{\bar{k}}$ :

$$\tilde{\mathbf{h}}_{\bar{k}} = [P_{\mathbf{h}_{\bar{k},1:0}}^\perp \mathbf{h}_{\bar{k},1} \ P_{\mathbf{h}_{\bar{k},1:1}}^\perp \mathbf{h}_{\bar{k},2} \ \dots \ P_{\mathbf{h}_{\bar{k},1:K-2}}^\perp \mathbf{h}_{\bar{k},K-1}]. \quad (14)$$

Then  $\mathbf{h}_{\bar{k}} = \tilde{\mathbf{h}}_{\bar{k}} \mathbf{L}^H$  is the QR factorization of  $\mathbf{h}_{\bar{k}}$  and  $\mathbf{D} = \tilde{\mathbf{h}}_{\bar{k}}^H \tilde{\mathbf{h}}_{\bar{k}}$ . We can write equivalently

$$P_{\mathbf{h}_{\bar{k}}}^\perp \mathbf{h}_k = [\mathbf{h}_k \ \tilde{\mathbf{h}}_{\bar{k}}] \begin{bmatrix} 1 \\ \mathbf{r} \end{bmatrix}, \quad \mathbf{r} = -\mathbf{D}^{-1} \tilde{\mathbf{h}}_{\bar{k}}^H \mathbf{h}_k = -\tilde{\mathbf{h}}_{\bar{k}}^\# \mathbf{h}_k. \quad (15)$$

This allows to write all consecutive ZF solutions as

$$P_{\mathbf{h}_{\bar{k},1:i}}^\perp \mathbf{h}_k = \mathbf{h}_k + \sum_{n=1}^i P_{\mathbf{h}_{\bar{k},n}}^\perp \mathbf{h}_k = \mathbf{h}_k + \sum_{n=1}^i \tilde{\mathbf{h}}_{\bar{k},n} r_n \quad (16)$$

with  $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_{K-1}]^T$ . Now we are able to write the transformation

$$[\mathbf{h}_k \ \tilde{\mathbf{h}}_{\bar{k}}] \underbrace{\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & r_1 & r_1 & \dots & r_1 \\ 0 & 0 & r_2 & \dots & r_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & r_{K-1} \end{bmatrix}}_R = [P_{\mathbf{h}_{\bar{k},1:0}}^\perp \mathbf{h}_k \ \dots \ P_{\mathbf{h}_{\bar{k},1:K-1}}^\perp \mathbf{h}_k] \quad (17)$$

where the upper triangular transformation matrix  $R$  is nonsingular since all  $r_i \neq 0$  (assuming  $\mathbf{H}$  to have full column rank). Note that for the ZF case,  $R^{-1} \begin{bmatrix} 1 \\ \mathbf{r} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$ . (12) can now be obtained with

$$\boldsymbol{\alpha} = R^{-1} \begin{bmatrix} 1 \\ (I + \mathbf{D}^{-1} \mathbf{L}^{-1} \mathbf{L}^{-H})^{-1} \mathbf{r} \end{bmatrix} \|P_{\mathbf{h}_k}^\perp \mathbf{h}'_k\|^{-2}. \quad (18)$$

Theorem 1 does not require the columns of  $\mathbf{h}_k$  to be ordered in any particular way but it will be advantageous to order the interferers in the order of decreasing power. In the case in which the powers of the various interferers are all quite different, this analysis suggests an approximation of the MMSE filter by a single ZF filter, or in other words to approximate the vector of combination coefficients  $\boldsymbol{\alpha}$  by a sparse vector with a single non-zero entry. Indeed, assume there is an index  $i$  such that

$$\|\mathbf{h}_{k,j}\|^2 \begin{cases} \gg 1 & , j \leq i, \\ \ll 1 & , j > i. \end{cases} \quad (19)$$

Then this leads for the MMSE filter to the ZF approximation

$$\begin{aligned} (\mathbf{h}_k^H \mathbf{h}_k + I)^{-1} \mathbf{h}_k^H \mathbf{h}_k &\approx \begin{bmatrix} \mathbf{h}_{k,1:i}^H \mathbf{h}_{k,1:i} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{h}_{k,1:i}^H \mathbf{h}_k \\ \mathbf{0} \end{bmatrix} \\ \Rightarrow \mathbf{f} &\approx \|P_{\mathbf{h}_{k,1:i}}^\perp \mathbf{h}_k\|^{-2} P_{\mathbf{h}_{k,1:i}}^\perp \mathbf{h}_k. \end{aligned} \quad (20)$$

## V. INTRACELL CHANNEL AND INTERCELL PATH CSIT

The CSI in the received signal (9) is quite different between Rxs and TxS. From the point of view of the TxS, introducing the channel model (6) for the intercell links leads to

$$\begin{aligned} \mathbf{y}_k &= \mathbf{H}_{k,b_k} \mathbf{g}_k x_k + \sum_{i:b_i=b_k} \mathbf{H}_{k,b_k} \mathbf{g}_i x_i + \tilde{\mathbf{v}}_k \\ \mathbf{C}_{\tilde{\mathbf{v}}_k} \tilde{\mathbf{v}}_k &= (1 + \sum_{i:b_i \neq b_k} (\|\mathbf{A}_{k,b_i}^H \mathbf{g}_i\|^2 + \beta_{k,b_i} \|\mathbf{g}_i\|^2)) I \end{aligned} \quad (21)$$

where we model the intercell interference plus noise as a Gaussian noise  $\tilde{\mathbf{v}}_k$ , leading to a mutual information lower bound. In contrast to the intercell CSIT, BS  $b_k$  is assumed to possess (pretty good) partial CSIT for  $\mathbf{H}_{k,b_k}$ . On the basis of the model (21), it is possible to formulate a BF design based on maximizing an overall Expected Weighted Sum Rate (EWSR) as in [18]. However, in spite of the slightly simplified pathwise model in (21), such designs still require global CSIT concerning the intercell CSIT  $\|\mathbf{A}_{k,b_i}^H \mathbf{g}_i\|$ ,  $\beta_{k,b_i}$  and  $\|\mathbf{g}_i\|$  plus the usual intracell CSIT. It is then not surprising that simplified approaches have been proposed e.g. based on interference pricing [19], [20].

We shall consider a hierarchical BF of the form

$$\mathbf{g}_k = \mathbf{G}_{b_k} \mathbf{c}_k \quad (22)$$

where  $\mathbf{G}_{b_k}$  is the outer BF or intercell BF for BS  $b_k$  and  $\mathbf{c}_k$  is the inner or intracell BF for user  $k$ . Again, whereas the two levels of BF could be designed jointly based on EWSR, we shall here consider here a more decoupled approach.

## A. Intercell Path BF $\mathbf{G}_{b_k}$ Design

It is well known through uplink-downlink duality that the (optimal) WSR maximizing BF (Tx) has the form of a MMSE Rx. We consider the developments of Section IV to approximate the optimal BF by a possibly reduced order ZF (RO-ZF) BF. Using the pathwise channel model (6) for the intercell links, in the downlink interpreted as dual uplink, the unknown and rapidly varying complex path gains (phases) decouple the pathwise contributions and make each path appear as an individual source with its associated Tx side antenna arrays response. For the RO-ZF BF, the BS needs to determine which paths can be ignored in the ZF operation. The intercell BF will then be determined from

$$P_{\mathbf{A}_{b_k, b_k}^\perp}^\perp = \mathbf{G}_{b_k} \mathbf{G}_{b_k}^H \quad (23)$$

which assumes a (column) unitary  $\mathbf{G}_{b_k}$ . So,  $\mathbf{A}_{b_k, b_k}^\perp = [\mathbf{A}_{i, b_k}, b_i \neq b_k]$  and the column space of  $\mathbf{G}_{b_k}$  spans the orthogonal complement of that of  $\mathbf{A}_{b_k, b_k}^\perp$ . Now, the actual  $\mathbf{A}_{b_k, b_k}^\perp$  considered here is of reduced order, counting only the significant paths as determined below.

1) *Role of Rx Antennas:* Whereas in the MU-MIMO case a multitude of ZF configurations are possible, the precise distribution of the roles of interference reduction over Tx and Rx affects the resulting diversity for each user. In a Massive MIMO setting, with limited antennas on the Rx side but an excess of antennas at the Tx side, it is preferable in order to maximize diversity for Rx antennas to cancel intercell and not intracell interference, and actually to only do a minimal amount of ZF.

2) *WSR Inspired Interference Pricing for ZF Order Determination:* In the RO-ZF simplification of an optimal BF, the effect of adding one more path to be ZF'd is relatively straightforward to analyze: the additional ZF leads to a reduction in signal power for all intracell users, and the disappearance of one interference term. Consider the WSR

$$WSR = - \sum_k u_k \ln e_k \quad (24)$$

where the  $u_k$  is the weight for the rate of user  $k$  and  $e_k$  is the MMSE attained at the (optimal) Rx for user  $k$ . The variation of the WSR due to a variation in  $\mathbf{G}_{b_k}$  by possibly adding ZF to an additional path from BS  $b_k$  to user  $i$  is of the form

$$\Delta WSR = - \sum_{j:b_j=b_k} \frac{u_j}{e_j} \Delta e_j - \frac{u_i}{e_i} \Delta e_i \quad (25)$$

where  $\Delta e_j > 0$  represent an increase and  $\Delta e_i < 0$  represents a decrease in MSE. The ZF order should be increased as long as  $\Delta WSR > 0$ . The variation  $\Delta e_j$  can be determined from (16).

We get an expression of the form  $\frac{\Delta e_j}{e_j} = \frac{|r_n|^2 \|\tilde{\mathbf{h}}_n\|^2}{\|P_{\mathbf{h}_{j,n}}^\perp \mathbf{h}_j\|^2}$  which involves information within cell  $k$  (apart from the intercell path array responses). For  $\Delta e_i$ , which corresponds to the disappearance of an interference path from BS  $b_k$  to MT  $i$ , we continue the modeling in (21). The power of e.g. all paths corresponds to  $\|\mathbf{A}_{i, b_k}^H \mathbf{g}_k\|^2 = \|\mathbf{A}_{i, b_k}^H \mathbf{G}_{b_k} \mathbf{c}_k\|^2$ . Assuming omnidirectional  $\mathbf{c}_k$  (or random  $\mathbf{c}_k$  which is independent of the intercell paths), this becomes  $\frac{P_{b_k}}{M} \text{tr}\{\mathbf{A}_{i, b_k}^H P_{b_k}^\perp \mathbf{A}_{i, b_k}\}$  where

$p_{b_k}$  represents the BS Tx power, and  $P_{b_k}^\perp$  projects onto the orthogonal complement of the paths already accounted for. To evaluate the impact of interference reduction at intercell user  $i$ , BS  $b_k$  needs to know also  $\frac{u_i}{e_i}$  where for  $e_i$  some averaged MSE could be used. The RO-ZF should progressively consider the intercell paths to all affected users  $i$ .

3) *MT Initiated ZF Order Determination*: Here we assume that for any intercell link, the BS and MT are aware of the various (dominant) paths involved, ordered according to decreasing strength. In this approach, MT  $i$  simply informs BS  $b_k$  of how many paths in their cross link should be canceled by the BS, on the basis of a desired residual interference plus noise level. The MT can possibly consider to cancel some remaining paths (of intermediate strength) itself.

4) *CSI Acquisition*: For intercell interference management purposes, there should be isotropic DL training so that MTs can estimate the strengths of the various intercell paths. This operation does not require full estimation of the DL channels by the MTs, the path induced low rank nature (of  $C_i$ ) can be exploited to reduce the training overhead (similar ideas appear in [21]).

Concerning FDD (assumed so far) vs. TDD, TDD may lead to more instantaneous local CSIT, but the problem of global coupling in e.g. WSR optimization remains.

### B. Intracell Partial Channel CSIT BF $c_k$ Design

Once the intercell BF has been designed, the design of the intracell BF becomes a classical single cell (MU DL or BC) design with partial CSIT. The intracell design requires DL training of the cascades of intercell channels and intercell BFs, to allow the MTs to estimate the (assumed white and small) residual intercell interference plus noise level, and of course the intracell channels. We refer to [18] for a number of attractive BF designs maximizing Expected WSR (EWSR) in a BC based on Gaussian CSIT. Note that whereas the outer precoder  $G_{b_k}$ , which only depends on the path structure, can be slowly time-varying, whereas the inner precoder  $c_k$  will (need to be) be fastly varying if it depends on instantaneous (mean) CSIT.

## VI. ALTERNATIVE CSIT FORMULATIONS

1) *Receiver Feedback*: Rx feedback could be considered as in [22]. There are two dimensions to be accounted for here. One is the desire to obtain a decoupled, non-iterative BF design. This implies that the Rx design should not depend on the BF, but only on local CSIR. The other dimension is that the Rx depends in any case on instantaneous CSIR and hence a design of BF based on Rxs becomes fastly time-varying, requiring fast feedback also.

2) *Per User (Channel) CSIT and Inter-User Path CSIT*: The split between the pathwise and userwise approaches can be put at a different level, for instance by considering userwise only for direct links and pathwise for (intercell and intracell) cross links.

3) *Only Path CSIT*: The pathwise approach can be adopted also for the direct links. In this case the exploitation of diversity becomes highly desirable. The Rx antennas (with good CSIR) can provide one source of diversity. Diversity can be augmented by space-time (or -frequency) coding over the various direct link paths.

### Acknowledgements

EURECOM's research is partially supported by its industrial members: ORANGE, BMW Group, Swisscom, SFR, ST Microelectronics, Symantec, SAP, Monaco Telecom, iABG, and also by the EU FP7 projects ADEL and NEWCOM#. The research of Orange Labs is partially supported by the EU FP7 project METIS.

### REFERENCES

- [1] F. Negro, I. Ghauri, and D. T. M. Slock, "Deterministic Annealing Design and Analysis of the Noisy MIMO Interference Channel," in *Proc. IEEE Information Theory and Applications workshop (ITA)*, San Diego, CA, USA, Feb. 2011.
- [2] F. Negro, D. Slock, and I. Ghari, "On the Noisy MIMO Interference Channel with CSI through Analog Feedback," in *Int'l Symp. Communications, Control and Sig. Proc. (ISCCSP)*, Rome, Italy, May 2012.
- [3] Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Foresighted Delayed CSIT Feedback for Finite Rate of Innovation Channel Models and Attainable NetDoFs of the MIMO Interference Channel," in *Proc. Wireless Days*, Valencia, Spain, Nov. 2013.
- [4] A. Lozano, R. Heath, and J. Andrews, "Fundamental Limits of Cooperation," *IEEE Trans. Info. Theory*, Sep. 2013.
- [5] S. Jafar, "Topological Interference Management Through Index Coding," *IEEE Trans. Info. Theory*, Jan. 2014.
- [6] Y. Lejosne, M. Bashar, D. Slock, and Y. Yuan-Wu, "MIMO Interfering Broadcast Channels based on Local CSIT," in *Proc. European Wireless (EW)*, Barcelona, Spain, May 2014.
- [7] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for Next Generation Wireless Systems," *IEEE Comm's Mag.*, Feb. 2014.
- [8] S. Wagner, R. Couillet, M. Debbah, and D. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels Under Limited Feedback," *IEEE Trans. Info. Theory*, Jul. 2012.
- [9] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing: the Large-Scale Array Regime," *IEEE Trans. Info. Theory*, Oct. 2013.
- [10] A. Adhikary, E. E. Sadafi, M. Samimi, R. Wang, G. Caire, T. Rappaport, and A. Molisch, "Joint Spatial Division and Multiplexing for mm-Wave Channels," [Online]. Available: <http://arxiv.org/abs/1312.2045>.
- [11] A. Liu and V. Lau, "Hierarchical Interference Mitigation for Large MIMO Cellular Networks," *IEEE Trans. Signal Processing*, Apr. 2013.
- [12] J. Brady and A. Sayeed, "Beamspace MU-MIMO for High-Density Gigabit Smallm Cell Access at Millimeter-Wave Frequencies," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Toronto, Canada, Jun. 2014.
- [13] A. Sayeed, "Deconstructing Multiantenna Fading Channels," *IEEE Trans. Sig. Proc.*, Oct. 2002.
- [14] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with Interference in Distributed Large-scale MIMO Systems: A Statistical Approach," [Online]. Available: <http://arxiv.org/abs/1310.6674>.
- [15] Y. Lejosne, M. Bashar, D. Slock, and Y. Yuan-Wu, "Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels," in *Proc. IEEE Int'l Symp. on Communications, Control and Sig. Proc. (ISCCSP)*, Athens, Greece, May 2014.
- [16] E. Jorswieck, E. Larsson, and D. Danev, "Complete Characterization of the Pareto Boundary for the MISO Interference Channel," *IEEE Trans. Sig. Proc.*, Oct. 2008.
- [17] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative Multicell Precoding: Rate Region Characterization and Distributed Strategies With Instantaneous and Statistical CSI," *IEEE Trans. Sig. Proc.*, Aug. 2010.
- [18] M. Bashar and D. Slock, "Cognitive Multi-User MIMO Downlink with Mixed Feedback/Location based Gaussian CSIT," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Toronto, Canada, June 2014.
- [19] D. Schmidt, C. Shi, R. Berry, M. Honig, and W. Utschick, "Comparison of Distributed Beamforming Algorithms for MIMO Interference Networks," *IEEE Trans. Sig. Proc.*, July 2013.
- [20] R. Berry and D. Tse, "Shannon meets Nash on the Interference Channel," *IEEE Trans. Information Theory*, Apr. 2011.
- [21] X. Rao, V. Lau, and X. Kong, "CSIT Estimation and Feedback for FDD Multi-user Massive MIMO Systems," in *Proc. IEEE Int'l Conf. Acoustics Speech and Sig. Proc. (ICASSP)*, Florence, Italy, May 2014.
- [22] Z. Ho and D. Gesbert, "Balancing Egoism and Altruism on Interference Channel: The MIMO Case," in *Proc. IEEE Int'l Conf. Communications (ICC)*, Cape Town, South Africa, May 2010.