The Noisy MIMO Interference Channel with Distributed CSI Acquisition and Filter Computation

Francesco Negro*, Umer Salim†, Irfan Ghauri†, Dirk T.M. Slock*
†Intel Mobile Communications, GAIA, 2600 Route des Crêtes, 06560 Sophia Antipolis Cedex, France
Email: {umer.salim, irfan.ghauri}@intel.com
*Mobile Communications Department, EURECOM, 2229 Route des Crêtes, BP 193, 06904 Sophia Antipolis Cedex, France
Email: {francesco.negro, dirk.slock}@eurecom.fr

***Abstract*— We study the frequency-flat noisy MIMO interference channel (IFC) without symbol extension and with initial assumption of no channel state information (CSI) neither at the base stations (BS) nor at the user equipments (UE). In the noisy IFC, interference is treated as noise and hence linear transmit (Tx) and receive (Rx) filtering is considered. A transmission strategy is proposed through which the BS and the UE get the necessary CSI for CSI based transmit and receive processing, by channel training and analog channel feedback, leading to transmission overhead. Both FDD and TDD scenarios are considered (the details for which are not too different in this distributed approach). In the limited overhead approach considered, discrepancies arise between the local estimates of the global CSI established at the various Tx/Rx units, on the basis of which the local Tx/Rx filters are computed. These filters are computed by optimizing a lower bound (and close approximation) of the weighted sum rate, accounting for partial CSI. The approach allows for simple approximate sum rate expressions that can easily be optimized for any set of system parameters to unveil the trade-off between the cost and the gains associated to CSI acquisition overhead. A centralized approach is briefly discussed also.**

I. Introduction

The main bottleneck of system capacity in modern wireless cellular systems is interference. This comes from the aggressive frequency reuse factor 1 that has been recently introduced by the standardization bodies to increase the overall system performances. A systematic study of the performance of cellular communication systems severely affected by intercell interference can be model using the K-user Interference Channel (IFC). The application of the IFC to new challenging practical problems has increased the research effort in studying the interference channel from an information theoretic point of view. While the interference channel has been the focus of intense research over the past few decades, its capacity in general remains an open problem and is not well understood even for simple cases [1]. In the seminal work [2] the authors have shown that the conventional approach of orthogonalizing the signal dimensions can be overcome by the use of a new signaling technique called Interference Alignment (IA). They have proved that for(time or frequency) varying SISO channels a total of $\frac{K}{2}$ interference-free streams can be received with IA instead of the 1 obtained through orthogonalization. This significant increase of degrees of freedom (DoF) can be achieved by asymptotic signal-space expansion in time or frequency called *symbol extension*. The sum degrees of freedom for a general MIMO IFC is still an open problem, the only known result is given in [3] for a $K = 2$ user MIMO IFC. For the case $K > 2$ some bounds have been provided in [4]. IA requires perfect and global channel state information (CSI) at all Tx/Rx. This assumption does not come for free in practical time-varying channels. For this reason different studies have been conducted for more practical situations. In [5] the authors consider the SISO IFC with frequency selective channels. Using quantized channels over a Grassmann manifold fed back using an error-free feedback channel they show that full multiplexing gain can be achieved if the feedback bit rate scales sufficiently fast with the SNR. This result is extended in [6] to the MISO and MIMO IFC. In [7] the author shows for different selected multiuser communication scenarios that it is possible to align the interference when the transmitters do not know the channel coefficients but they only have information about the channel autocorrelation structure of different users. In [7] a staggered block fading channel model is the only assumption required to achieve complete IA. The resulting multiplexing gain is much lower however than for the case of full CSI. The authors of [8] propose to use analog feedback for the acquisition of full CSIT. The channel coefficients are directly fed back to the base station (BS) without any quantization process. This has the advantage, in contrast to digital feedback, that the complexity does not increase with SNR. They show that using IA with the acquisition of CSIT using analog feedback incurs no loss of multiplexing gain if the feedback power scales with the SNR.

In this paper, as in [8], we assume no other connectivity than the Uplink/Downlink (UL/DL) wireless IFC. We want to show that if the channel is varying not too fast in time all the CSI acquisition, necessary for distributed IA beamformer design, can be done with a finite overhead that causes only a finite SNR loss compared to the full CSI case. This goes in contrast to the distributed solutions where iterations between Transmitters and Receiver, e.g. [9], are required in order to converge to a proper IA solution. Those solutions require a

The research of Intel Mobile Communications France is supported in part by the EU FP7 projects CROWN, SACRA.
EURECOM's research is partially supported by its industrial members: BMW Group, Swisscom, Cisco, ORANGE, SFR, ST Ericsson, Thales, Symantec, SAP, Monaco Telecom and by the EU FP7 projects CROWN, SACRA and WHERE2.

huge number of iterations that decrease a lot the gain of an IA transmission. We note also that in the MIMO case, the presence of multiple receive antennas implies that the design of the rceivers, in a final training phase, is not negliglible (in contrast to what was assumed in [8]). Another key point in this paper is that an approach with even distributed CSI acquisition as proposed here, followed by CSI based IA (or optimized Tx/Rx), leads to significantly higher multiplexing gains (rate prelogs) than so-called blind IA or non-coherent IA or other IA approaches based on delayed CSIT. Finally, a practical scheme for the training based approach considered here is far from unique because many different training and feedback solutions and (e.g. partial CSI based) Tx/Rx designs are possible. We just consider one simple approach to make the point.

II. Signal Model

A. K-Link MIMO Interference Channel

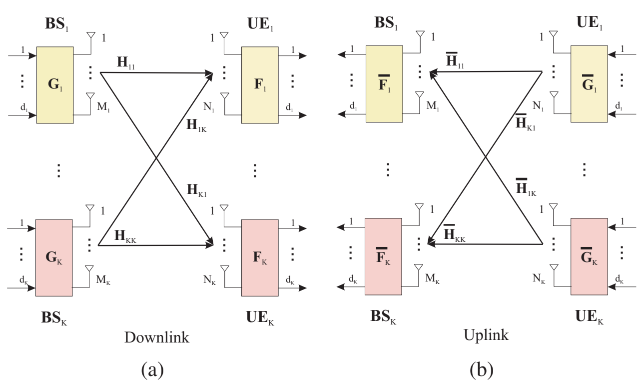

Fig. 1: MIMO DL/UL Interference Channel

Fig. 1(a) depicts a K-link MIMO interference channel with K transmitter-receiver pairs. In the following we assume a Time Division Duplexing (TDD) transmission scheme. To differentiate the two transmitting and receiving devices we assume that each of the K pairs is composed of a Base station (BS) and a Mobile user (MU). This is only for notational purposes. The k-th BS and its corresponding MU are equipped with M_k and N_k antennas respectively. The k-th transmitter generates interference at all $l \neq k$ receivers. The received signal in the Downlink (DL) phase $\mathbf{y}_k$ at the k-th MU, can be represented as

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{x}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{H}_{kl}\mathbf{x}_l + \mathbf{n}_k \qquad (1)$$

where $\mathbf{H}_{kl} \in \mathbb{C}^{N_k \times M_l}$ represents the channel matrix between the l-th BS and k-th MU, $\mathbf{x}_k$ is the $\mathbb{C}^{M_k \times 1}$ transmit signal vector of the k-th BS and the $\mathbb{C}^{N_k \times 1}$ vector $\mathbf{n}_k$ represents (temporally white) AWGN with zero mean and covariance matrix $\mathbf{R}_{n_k n_k}$. The channel is assumed to follow a block-fading model having a coherence time of T symbol intervals without channel variation. Each entry of the channel matrix is a complex random variable drawn from a continuous distribution. It is assumed that each transmitter has complete knowledge of all channel matrices corresponding to its direct link and all the other cross-links in addition to the transmitter power constraints and the receiver noise covariances.

We denote by $\mathbf{G}_k$, the $\mathbb{C}^{M_k \times d_k}$ precoding matrix of the k-th transmitter. Thus $\mathbf{x}_k = \mathbf{G}_k\mathbf{s}_k$, where $\mathbf{s}_k$ is a $d_k \times 1$ vector representing the d_k independent symbol streams for the k-th user pair. We assume $\mathbf{s}_k$ to have a spatio-temporally white Gaussian distribution with zero mean and unit variance, $\mathbf{s}_k \sim \mathcal{N}(0, \mathbf{I}_{d_k})$. The k-th receiver applies $\mathbf{F}_k \in \mathbb{C}^{d_k \times N_k}$ to suppress interference and retrieve its d_k desired streams. The output of such a receive filter is then given by

$$\mathbf{r}_k = \mathbf{F}_k\mathbf{H}_{kk}\mathbf{G}_k\mathbf{s}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{F}_k\mathbf{H}_{kl}\mathbf{G}_l\mathbf{s}_l + \mathbf{F}_k\mathbf{n}_k$$

In the reverse transmission link, Fig. 1(b) Uplink (UL) phase, the received signal at the k-th BS is given by:

$$\bar{\mathbf{r}}_k = \overline{\mathbf{F}}_k\overline{\mathbf{H}}_{kk}\overline{\mathbf{G}}_k\bar{\mathbf{s}}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \overline{\mathbf{F}}_k\overline{\mathbf{H}}_{kl}\overline{\mathbf{G}}_l\bar{\mathbf{s}}_l + \overline{\mathbf{F}}_k\bar{\mathbf{n}}_k$$

where $\overline{\mathbf{F}}_k$ and $\overline{\mathbf{G}}_l$ denote respectively the $d_k \times M_k$ Rx filter at BS number k and the $N_l \times d_l$ BF matrix applied at Tx l. The UL channel form the l-th MU and the k-th BS is denoted as $\overline{\mathbf{H}}_{kl}$.

B. Interference Alignment (IA)

The objective in IA is to design spatial filters to be applied at the transmitters such that, the interference caused by all transmitters at each non-intended RX lies in a common *interference subspace*. In this paper we do not consider the use of symbol extension [2]. Moreover, the interference subspace and the *desired signal subspace* of each RX should be non-overlapping (linearly independent). If alignment is complete, simple ZF can be applied to suppress the interference and extract the desired signal in the high-SNR regime. Since IA is a condition for joint transmit-receive linear ZF, we need to satisfy the following conditions:

$$\mathbf{F}_k\mathbf{H}_{kl}\mathbf{G}_l = \mathbf{0} \quad \forall l \neq k \qquad (2)$$

$$\mathrm{rank}(\mathbf{F}_k\mathbf{H}_{kk}\mathbf{G}_k) = d_k \quad \forall k \in \{1, 2, \ldots, K\} \qquad (3)$$

This last rank condition leads to the traditional single user MIMO constraint $d_k \leq \min(M_k, N_k)$ for d_k streams to be able to pass over the k-th link. A closed form expression for the BF and the Rx filters is not known in general, it is derived only for a few simple MIMO interference channel configurations. To determine the IA solution of a general K-User MIMO IFC the only possible alternative is the computation using an iterative algorithm. In the literature several of such algorithms have been proposed based on different optimization criteria, for example [9].

III. Transmission Phases

We assume a block fading model, in which the channel is assumed to be constant over T channel uses. This time period T will need to be shared between the different training T_{ovrhd} and data transmission phases $T_{data} = T - T_{ovrhd}$ of the overall transmission scheme. In this section we describe all the necessary transmission phases required to set up a communication using IA beamformers and receiver filters. In particular study in detail the FDD transmission schemes, this approach will be justified later on in the paper.

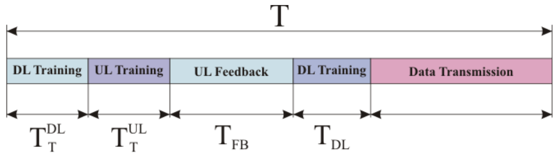

Fig. 2: MIMO Uplink Interference Channel

A. Downlink Training Phase

During this phase each BS_k sends orthogonal pilot sequences that can be received by all the MU for a total duration of T_T^{DL}. In this way UE_i can easily estimate the DL channels $\mathbf{H}_i = [\mathbf{H}_{i1}, \ldots, \mathbf{H}_{iK}]$ directly connected to it. Because the compound channel matrix $\mathbf{H}_i$ has dimensions $N_i \times \sum_k M_k$ the minimum total duration of this training phase is

$$T_T^{DL} \geq \sum_{k=1}^{K} M_k.$$

Each BS independently transmits an orthogonal matrix $\mathbf{\Psi}_k$ of dimension $M_k \times T_T^{DL}$ with power P_T^{DL} hence the total received $N_i \times T_T^{DL}$ matrix at Rx i is:

$$\mathbf{Y}_i = \sum_{k=1} \sqrt{P_T^{DL}} \mathbf{H}_{ik} \mathbf{\Psi}_k + \mathbf{V} \tag{4}$$

where $\mathbf{V}$ represents the zero mean additive white Gaussian noise with variance σ_v^2. The DL Tx power can be related to the time duration of the corresponding Tx phase as

$$P_T^{DL} = \frac{T_T^{DL}}{\sum M_k} \overline{P}_T^{DL}. \tag{5}$$

where $\overline{P}_T^{DL}$ represents the DL power constraint. Using an MMSE estimate on $\mathbf{Y}_i \mathbf{\Psi}_l$ each DL channel can be written as $\mathbf{H}_i = \widehat{\mathbf{H}}_i + \widetilde{\mathbf{H}}_i$ where:

$$\widehat{\mathbf{H}}_i \sim \mathcal{N}\left(0, \frac{P_T^{DL}}{\sigma_v^2 + P_T^{DL}} \mathbf{I}\right), \quad \widetilde{\mathbf{H}}_i \sim \mathcal{N}\left(0, \frac{\sigma_v^2}{\sigma_v^2 + P_T^{DL}} \mathbf{I}\right) \tag{6}$$

we call $\sigma^2_{\widehat{\mathbf{H}}_i}$ and $\sigma^2_{\widetilde{\mathbf{H}}_i}$ the variance of the channel estimate and error respectively.

B. Uplink Training Phase

This phase can be seen as the dual of the DL training where now all MU send orthogonal pilots to each BS for the estimation of the UL channel matrices. The time duration of this phase is:

$$T_T^{UL} \geq \sum_{k=1}^{K} N_k.$$

Then BS_k can estimate the compound channel matrix $\overline{\mathbf{H}}_i = [\overline{\mathbf{H}}_{i1}, \ldots, \overline{\mathbf{H}}_{iK}]$ using an MMSE estimator as described for the DL training phase. Each Ul channel can be represented in terms of channel estimate and channel estimation error with variance respectively $\sigma^2_{\widehat{\overline{\mathbf{H}}}}$ and $\sigma^2_{\widetilde{\overline{\mathbf{H}}}}$:

$$\widehat{\overline{\mathbf{H}}}_i \sim \mathcal{N}\left(0, \frac{P_T^{UL}}{\sigma_v^2 + P_T^{UL}} \mathbf{I}\right), \quad \widetilde{\overline{\mathbf{H}}}_i \sim \mathcal{N}\left(0, \frac{\sigma_v^2}{\sigma_v^2 + P_T^{UL}} \mathbf{I}\right). \tag{7}$$

The DL training power is now defined as:

$$P_T^{UL} = \frac{T_T^{UL}}{\sum N_k} \overline{P}_T^{UL}. \tag{8}$$

where $\overline{P}_T^{UL}$ represents the UL power constraint. We are describing all the transmission phases for the FDD transmission scheme, hence different frequency bands are used for UL and DL communications. This separation implies that transmission and reception can take place at the same time. If we take advantage of this possibility the two training phases, UL and DL, can collapse in only one training slot that have duration $T_T = \max\{T_T^{DL}, T_T^{UL}\}$. Accounting for this new training phase implies a reduction of the total overhead T_{ovrhd}.

C. Uplink Feedback Phase

Once the UL and DL training phases are completed each terminal knows the channel directly connected to it in the UL and DL respectively. In order to compute the IA BF matrices full DL CSI is required. In FDD case, the one under investigation, each MU has to feedback the DL channel estimate $\widehat{\mathbf{H}}_i$ to all BS, this task can be done using Analog Feedback (AFB). This particular transmission phase should be designed according to the particular type of processing used for the computation of the BF matrices. We can describe two approaches: centralized and distributed. In the former a central controller acquire the necessary CSI, computes the BFs and then it disseminates this information among the K BSs. In the latter approach each BS should have full CSI to compute the IA BF, using e.g. the approach of [10]. This solution can be also called *Duplicated* because each BS essentially solves the same problem and find the complete solution, all the IA BF, and then it will use only its own BF.

Centralized Processing
The Rx signal vector at each BS is sent to the centralized controller that retrieves the useful channel information and computes the BF matrices. If we stack all the received vector from the K BSs in $\overline{\mathbf{Y}}$ we get:

$$\underbrace{\begin{bmatrix} \overline{\mathbf{Y}}_1 \\ \vdots \\ \overline{\mathbf{Y}}_K \end{bmatrix}}_{\overline{\mathbf{Y}}} = \sqrt{P_{FB}} \underbrace{\begin{bmatrix} \overline{\mathbf{H}}_{11} & \ldots & \overline{\mathbf{H}}_{1K} \\ \vdots & \ddots & \vdots \\ \overline{\mathbf{H}}_{K1} & \ldots & \overline{\mathbf{H}}_{KK} \end{bmatrix}}_{M \times N} \underbrace{\begin{bmatrix} \widehat{\mathbf{H}}_1 \\ \vdots \\ \widehat{\mathbf{H}}_K \end{bmatrix}}_{N \times M} \underbrace{\begin{bmatrix} \mathbf{\Phi}_1 \\ \vdots \\ \mathbf{\Phi}_K \end{bmatrix}}_{M \times T_{FB}} + \underbrace{\begin{bmatrix} \mathbf{V}_1 \\ \vdots \\ \mathbf{V}_K \end{bmatrix}}_{\overline{\mathbf{V}}}$$

where $N = \sum_i N_i$ and $M = \sum_i M_i$ and

$$P_{FB} = \overline{P}_{FB} \frac{T_{FB}}{N_i M} \tag{9}$$

with $\overline{P}_{FB}$ is the feedback power constraint. Using a centralized controller to gather all Rx data the entire system can be interpreted as a unique single user MIMO link with a BS that is equipped with M total antennas and a MU with N antennas. With this interpretation we can calculate the total amount of time necessary to satisfy the identifiability conditions. In particular we get:

$$T_{FB} \geq \frac{N \times M}{\min\{N, M\}} = \max\{N, M\} \propto K. \tag{10}$$

To extract the $i-$th AFB contribution we pre-multiply the received matrix $\overline{\mathbf{Y}}$ by the i-th orthonormal matrix $\mathbf{\Phi}_i$:

$$\overline{\mathbf{Y}} \mathbf{\Phi}_i = \sqrt{P_{FB}} \underbrace{\begin{bmatrix} \overline{\mathbf{H}}_{i1} \\ \vdots \\ \overline{\mathbf{H}}_{iK} \end{bmatrix}}_{\overline{\mathbf{H}}_i} \widehat{\mathbf{H}}_i + \overline{\mathbf{V}} \mathbf{\Phi}_i$$

then we perform a least square (LS) estimate based on the UL channel estimates $\widehat{\overline{\mathbf{H}}}_{ik}$: $\overline{\mathbf{H}}_i^{LS} = P_{FB}^{-\frac{1}{2}}(\widehat{\overline{\mathbf{H}}}_i^H\widehat{\overline{\mathbf{H}}}_i)^{-1}\widehat{\overline{\mathbf{H}}}_i^H$. Using this estimator we obtain the following AFB estimates:

$$\begin{aligned}\widehat{\widehat{\mathbf{H}}}_i &= \overline{\mathbf{H}}_i^{LS}\overline{\mathbf{Y}}\mathbf{\Phi}_i = \widehat{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}}\overline{\mathbf{H}}_i^{LS}\widetilde{\overline{\mathbf{H}}}_i\widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_i^{LS}\overline{\mathbf{V}}\mathbf{\Phi}_i \\ &= \mathbf{H}_i - \widetilde{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}}\overline{\mathbf{H}}_i^{LS}\widetilde{\overline{\mathbf{H}}}_i\widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_i^{LS}\overline{\mathbf{V}}\mathbf{\Phi}_i = \mathbf{H}_i - \widetilde{\widetilde{\mathbf{H}}}_i\end{aligned}$$

The AFB estimate can be written in function of the true DL channel and the AFB estimation error: $\widehat{\widehat{\mathbf{H}}}_i = \mathbf{H}_i - \widetilde{\widetilde{\mathbf{H}}}_i$. The error contribution is due to the DL and UL channel estimation errors $(\widetilde{\mathbf{H}}_i, \widetilde{\overline{\mathbf{H}}}_i)$ in the DL and UL training phases respectively. The AFB estimation error $\widetilde{\widetilde{\mathbf{H}}}_i$ is distributed as $\mathcal{N}(0, \sigma^2_{\widetilde{\widetilde{\mathbf{H}}}_i})$ where

$$\mathrm{Cov}(\widetilde{\widetilde{\mathbf{H}}}_i|\widehat{\overline{\mathbf{H}}}_i) = \sigma^2_{\widetilde{\mathbf{H}}_i}\mathbf{I} + [(\sigma^2_{\widehat{\mathbf{H}}_i}\sigma^2_{\widetilde{\overline{\mathbf{H}}}_i}) + \frac{\sigma^2}{P_{FB}}](\widehat{\overline{\mathbf{H}}}_i^H\widehat{\overline{\mathbf{H}}}_i)^{-1}.$$

Assuming that $\mathbb{E}\{(\widehat{\overline{\mathbf{H}}}_i^H\widehat{\overline{\mathbf{H}}}_i)^{-1}\} \propto \frac{1}{M-N_i}$ we get:

$$\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}_i} = \sigma^2_{\widetilde{\mathbf{H}}_i} + \frac{1}{M-N_i}[(\sigma^2_{\widehat{\mathbf{H}}_i}\sigma^2_{\widetilde{\overline{\mathbf{H}}}_i}) + \frac{\sigma^2}{P_{FB}}]$$

Distributed Processing

In this case the AFB transmission is organized in such a way that each BS can gather full channel knowledge from all MU. The Rx matrix at BS_k can be written as:

$$\overline{\mathbf{Y}}_k = \sqrt{P_{FB}}\underbrace{\left[\overline{\mathbf{H}}_{k1} \;\dots\; \overline{\mathbf{H}}_{kK}\right]}_{M_k\times N}\underbrace{\begin{bmatrix}\widehat{\mathbf{H}}_1 & \mathbf{0} & \dots & \mathbf{0}\\ \mathbf{0} & \widehat{\mathbf{H}}_2 & \dots & \mathbf{0}\\ \vdots & & \ddots & \mathbf{0}\\ \mathbf{0} & \dots & \dots & \widehat{\mathbf{H}}_K\end{bmatrix}}_{N\times KM}\underbrace{\begin{bmatrix}\mathbf{\Phi}_1\\ \vdots\\ \mathbf{\Phi}_K\end{bmatrix}}_{KM\times T_{FB}} + \mathbf{V}_k$$

where

$$P_{FB} = \overline{P}_{FB}\frac{T_{FB}}{N_iM} \tag{11}$$

with $\overline{P}_{FB}$ is the feedback power constraint. In the distributed approach to satisfy the identifiability conditions the AFB length should be:

$$T_{FB} \geq \frac{N\times M}{\min_i\{M_i, N_i\}} \propto K^2 \tag{12}$$

To extract the $i-$th AFB contribution we pre-multiply the received matrix $\overline{\mathbf{Y}}_k$ by the i-th orthonormal matrix $\mathbf{\Phi}_i$:

$$\overline{\mathbf{Y}}_k\mathbf{\Phi}_i = \sqrt{P_{FB}}\overline{\mathbf{H}}_{ki}\widehat{\mathbf{H}}_i + \mathbf{V}_k\mathbf{\Phi}_i.$$

Also in this case we use a LS estimator, based on the UL channel estimate $\widehat{\overline{\mathbf{H}}}_{ki}$, $\overline{\mathbf{H}}_i^{LS} = P_{FB}^{-\frac{1}{2}}(\widehat{\overline{\mathbf{H}}}_{ki}^H\widehat{\overline{\mathbf{H}}}_{ki})^{-1}\widehat{\overline{\mathbf{H}}}_{ki}^H$. The AFB estimate can be written in function of the true DL channel and the AFB estimation error: $\widehat{\widehat{\mathbf{H}}}_i = \mathbf{H}_i - \widetilde{\widetilde{\mathbf{H}}}_i$. The error contribution is due to the DL and UL channel estimation errors $(\widetilde{\mathbf{H}}_i, \widetilde{\overline{\mathbf{H}}}_{ki})$ in the DL and UL training phases respectively:

$$\begin{aligned}\widehat{\widehat{\mathbf{H}}}_i &= \overline{\mathbf{H}}_{ki}^{LS}\overline{\mathbf{Y}}_k\mathbf{\Phi}_i = \widehat{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}}\overline{\mathbf{H}}_{ki}^{LS}\widetilde{\overline{\mathbf{H}}}_{ki}\widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_{ki}^{LS}\mathbf{V}_k\mathbf{\Phi}_i \\ &= \mathbf{H}_i - \widetilde{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}}\overline{\mathbf{H}}_{ki}^{LS}\widetilde{\overline{\mathbf{H}}}_{ki}\widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_{ki}^{LS}\mathbf{V}_k\mathbf{\Phi}_i = \mathbf{H}_i - \widetilde{\widetilde{\mathbf{H}}}_i\end{aligned}$$

where the estimation error is then distributed as $\mathcal{N}(0, \sigma^2_{\widetilde{\widetilde{\mathbf{H}}}_i})$, with

$$\mathrm{Cov}(\widetilde{\widetilde{\mathbf{H}}}_i|\widehat{\overline{\mathbf{H}}}_{ki}) = \sigma^2_{\widetilde{\mathbf{H}}_i}\mathbf{I} + [(\sigma^2_{\widehat{\mathbf{H}}_i}\sigma^2_{\widetilde{\overline{\mathbf{H}}}_{ki}}) + \frac{\sigma^2}{P_{FB}}](\widehat{\overline{\mathbf{H}}}_{ki}^H\widehat{\overline{\mathbf{H}}}_{ki})^{-1}$$

Assuming that $\mathbb{E}\{(\widehat{\overline{\mathbf{H}}}_{ki}^H\widehat{\overline{\mathbf{H}}}_{ki})^{-1}\} \propto \frac{1}{M_k-N_i}$ we get:

$$\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}_i} = \sigma^2_{\widetilde{\mathbf{H}}_i} + \frac{1}{M_k-N_i}[(\sigma^2_{\widehat{\mathbf{H}}_i}\sigma^2_{\widetilde{\overline{\mathbf{H}}}_{ki}}) + \frac{\sigma^2}{P_{FB}}]$$

Another possible strategy to receive the analog feedback is to use linear MMSE estimate instead of the least square approach described in this section. The two solutions will be identical at high SNR but in different SNR regimes LMMSE should give better performances.

The analog feedback transmission described here is based on the assumption that the number of Tx and Rx antennas satisfy the relation that $\min\{M_i\} \geq N_j, \forall j$. If this condition is not satisfied then a different transmission scheme should be applied. In particular the precoding matrix should be such that the identifiability conditions should be satisfied at all BS, this require a more careful precoding design. This solution can be also used to introduced more redundancy in the transmission that can increase the performances of the feedback reception.

D. Downlink Training Phase

One the beamformers have been computed, using a centralized or distributed approach, they can be used for the DL communications. According to IA each MU should apply a ZF receiver, in order to compute the Rx filters each MU requires some additional information on the DL communication. On this purpose two approaches are possible: DL training or analog transmission of the entire Rx filters. In the former case BS_k sends a set of beamformed pilots that allow UE_i to estimate the cascade $\mathbf{H}_{ik}\mathbf{G}_k$. This phases lasts

$$T_{DL} \geq \sum_k d_k.$$

Then each MU can estimate the interference subspace and the signal subspace for the Rx filter design. The other possibility consists in the transmission to the i-th MU of the entire Rx filter matrix $\mathbf{F}_i$ using analog transmission. This solution requires a duration

$$T_{DL} \geq \sum_k \frac{N_kd_k}{\min\{N_k, M_k\}}$$

The two solution proposed here are not equivalent. Training is shorter but the estimation error will have a bigger impact in the calculation of the Rx filter compare to the one in the analog transmission. Which solution should be preferred depends also on the operating SNR point. For example in high SNR, where we are interested more in maximizing the total degrees of freedom the duration of this phase has a bigger impact compare to the estimation error then DL training is the preferable solution.

In the following we consider the approach based on training. Using a sequence of orthogonal pilots ϕ_{km} for stream (k,m) of length $1\times T_{DL}$, the Rx signal at MU k we get:

$$\mathbf{Y}_{km} = \sqrt{P_T}\mathbf{H}_{kk}\widehat{\mathbf{g}}_{km}\phi_{km} + \sum_{(in)\neq(km)}\sqrt{P_T}\mathbf{H}_{ki}\widehat{\mathbf{g}}_{in}\phi_{in} + \mathbf{V}_{km}$$

where the two $N_k \times T_{DL}$ matrices are $\mathbf{Y}_{km} = [\mathbf{y}_{km}[1], \dots, \mathbf{y}_{km}[T_{DL}]]$ and $\mathbf{V}_{km} = [\mathbf{v}_{km}[1], \dots, \mathbf{v}_{km}[T_{DL}]]$. The least square estimate of the cascade channel and BF is given as:

$$\widehat{\mathbf{H}_{kl}\widehat{\mathbf{g}}_{lt}} = \frac{1}{T_{DL}\sqrt{P_T}}\mathbf{Y}_{km}\phi_{lt}^H = \mathbf{H}_{kl}\widehat{\mathbf{g}}_{lt} + \underbrace{\frac{1}{T_{DL}\sqrt{P_T}}\mathbf{V}_{km}\phi_{lt}^H}_{\widetilde{\mathbf{H}_{kl}\widehat{\mathbf{g}}_{lt}}} \tag{13}$$

the elements of the estimation error matrix are distributed according to $\mathcal{N}(0, \sigma^2_{\widetilde{\mathbf{H}_{kl}\widehat{\mathbf{g}}_{lt}}}\mathbf{I})$, where $\sigma^2_{\widetilde{\mathbf{H}_{kl}\widehat{\mathbf{g}}_{lt}}} = \frac{\sigma^2}{T_{DL}P_T}$. Using channel estimate (13) we can build the MMSE Rx filter as:

$$\widehat{\widehat{\mathbf{f}}}_{km} = P_T\left(\widehat{\mathbf{H}_{kk}\widehat{\mathbf{g}}_{km}}\right)^H\left[\sum_{in}\widehat{\mathbf{H}_{ki}\widehat{\mathbf{g}}_{in}}P_T\left(\widehat{\mathbf{H}_{ki}\widehat{\mathbf{g}}_{in}}\right)^H + \sigma^2\mathbf{I}\right]^{-1} \tag{14}$$

We can further develop (13) in order to express in term of the DL channel estimate at the BS_l obtained using AFB in section III-C:

$$\widehat{\mathbf{H}_{kl}\widehat{\mathbf{g}}_{lt}} = \widehat{\widehat{\mathbf{H}}}^{(l)}_{kl}\widehat{\mathbf{g}}_{lt} + \widetilde{\widehat{\mathbf{H}}}^{(l)}_{kl}\widehat{\mathbf{g}}_{lt} + \widetilde{\mathbf{H}_{kl}\widehat{\mathbf{g}}_{lt}}. \quad (15)$$

$\widehat{\widehat{\mathbf{H}}}^{(l)}_{kl}$ represent the DL channel estimate calculated at BS_l used for the calculation of the BF vector $\widehat{\mathbf{g}}_{lt}$. With the expression (15) and the first order approximation: $(\mathbf{A}+\mathbf{\Delta A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{\Delta A}\mathbf{A}^{-1}$ we can decompose the Rx filter (14) as:

$$\widehat{\widehat{\mathbf{f}}}^{(l)}_{km} = P_T\widehat{\widehat{\mathbf{H}}}^{(l)}_{kk}\widehat{\mathbf{g}}_{km}\widehat{\mathbf{R}}^{(l)-1}_{\mathbf{yy}} + \widetilde{\widehat{\mathbf{f}}}^{(l)}_{km} = \widehat{\mathbf{f}}^{(l)}_{km} + \widetilde{\widehat{\mathbf{f}}}^{(l)}_{km} \quad (16)$$

$\widehat{\mathbf{f}}^{(l)}_{km}$ corresponds to the MMSE Rx filter calculated using the DL channel estimated at BS_l. It is the same MMSE Rx filter that would have been calculated at BS_l as a sub-product of the iterative algorithm used for calculating the IA BF. Then $\widehat{\mathbf{f}}^{(l)}_{km}\widehat{\widehat{\mathbf{H}}}^{(l)}_{kl}\widehat{\mathbf{g}}_{li} = 0$ at high SNR. $\widetilde{\widehat{\mathbf{f}}}^{(l)}_{km}$ contains all the error contributions of (14) up to first order.

IV. TDD Vs FDD transmission strategy

Usually the TDD transmission scheme is used arguing that thanks to reciprocity the amount of feedback required to acquire CSIT is (significantly) reduced. In this communication strategy UE_i does not need to feedback channel $\mathbf{H}_{ki}$ to BS_k because this information can be acquired using the corresponding UL channel $\overline{\mathbf{H}}_{ki}$. On the other hand this information is needed at the other base stations $BS_{j\neq k}$ for the design of their own BF matrix. From this observation we realized that for distributed BF process the organization of feedback is very complicated if we consider the possibility of reducing feedback using reciprocity, and hence we can conclude that TDD does not help in reducing the feedback overhead compare to FDD transmission scheme. On the contrary if we consider a centralized BF calculation then TDD makes feedback not required because the reduced set of CSI available at each BS using reciprocity is shared and hence the computation center can collect the total required information on the DL channels based on the UL channel estimates available at each BS. For the reasons described above we described all the transmission phases only for the FDD solution.

V. Final Goal: Sum Rate Optimization

In the end we get an expression for the sum rate (SR) at high SNR of the form $\mathcal{R}^{PCSI} =$

$$\sum_{k,n} \underbrace{(1-\frac{\sum T_i}{T})}_{\text{reduced data channel uses}} \ln(|\mathbf{f}_{kn}\mathbf{H}_{kk}\mathbf{g}_{kn}|^2 \underbrace{\rho/(1+\sum_i \frac{b_{kni}}{T_i}))}_{\text{SNR loss}},$$

assuming the various $T_i \geq T_{i,min}$. Now assume $b_{kni} = b_i$ for what follows. Then fixing $\sum_i T_i = T_{ovrhd}$, we get for the optimal $T_i = T_{ovrhd}\ \sqrt{b_i}/(\sum_i \sqrt{b_i})$. Next we can optimize over T_{ovrhd} which yields (at high SNR):

$$T_{ovrhd} = \frac{\sqrt{T}\ (\sum_i \sqrt{b_i})}{\sqrt{\mathcal{R}^{FCSI}}}\ .$$

Hence, if the coherence time T is large, the rate loss due to CSI acquisition is small.

VI. From Practical to more Optimal Solutions

All the different transmission phases described in section III are done one after the other but other solutions are possible to optimize the overhead. In a possible alternative approach, one does not need to wait to gather all CSIT before starting transmission. For example one user can start to transmit directly after the DL training phase as a single user MIMO link without any CSIT. Or also, it is possible to start with blind/noncoherent IA first. Then, instead of going from $K = 1$ to full K immediately another possible strategy is to build intermediate IA solutions adding gradually interfering links as soon as the corresponding transmitters get the required CSI to design the IA beamformers for the given interfering subsystem. Another consideration is that when the (analog) channel feedback duration is non-minimal, beamformers can be computed right after the minimal feedback has been obtained and DL transmission can start. Then the beamformers can get further updated during the continuing feedback. This is one advantage of analog feedback (similar to repetition coding), that ”decoding” can be done before the full ”codeword” has been received. In any case, there is a myriad of possibilities for alternative solutions, to increase the system performances using a more optimized transmission strategy.

References

[1] R. H. Etkin, D. N. C. Tse, and Hua Wang, “Gaussian interference channel capacity to within one bit,” *IEEE Trans. on Inform. Theory*, vol. 54, no. 12, pp. 5534–5562, 2008.

[2] V.R. Cadambe and S.A. Jafar, “Interference alignment and degrees of freedom of the K-user interference channel,” *IEEE Trans. on Inform. Theory*, vol. 54, no. 8, pp. 3425 –3441, Aug. 2008.

[3] S.A. Jafar and M.J. Fakhereddin, “Degrees of freedom for the MIMO interference channel,” *Information Theory, IEEE Transactions on*, vol. 53, no. 7, pp. 2637–2642, July 2007.

[4] Tiangao Gou and Syed Ali Jafar, “Degrees of freedom of the K user MxN MIMO interference channel,” *CoRR*, vol. abs/0809.0099, 2008, http://arxiv.org/abs/0809.0099.

[5] H. Bolcskei and I.J. Thukral, “Interference alignment with limited feedback,” in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, 28 2009-july 3 2009, pp. 1759 –1763.

[6] R.T. Krishnamachari and M.K. Varanasi, “Interference alignment under limited feedback for mimo interference channels,” in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, june 2010, pp. 619 –623.

[7] S.A. Jafar, “Exploiting channel correlations - simple interference alignment schemes with no CSIT,” in *GLOBECOM 2010, 2010 IEEE Global Telecommunications Conference*, dec. 2010, pp. 1 –5.

[8] Omar El Ayach and Robert W. Heath Jr., “Interference alignment with analog channel state feedback,” *CoRR*, vol. abs/1010.2787, 2010, http://arxiv.org/abs/1010.2787.

[9] K. Gomadam, V.R. Cadambe, and S.A. Jafar, “Approaching the capacity of wireless networks through distributed interference alignment,” in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, Dec 2008.

[10] F. Negro, I. Ghauri, and D.T.M. Slock, “Deterministic Annealing Design and Analysis of the Noisy MIMO Interference Channel,” in *IEEE Information Theory and Applications Workshop (ITA)*, UCSD, CA, USA, Feb. 2011.