On the Noisy MIMO Interference Channel with Analog Feedback

Francesco Negro*, Irfan Ghauri†, Dirk Slock*

*Mobile Communication Department, EURECOM
†Intel Mobile Communications

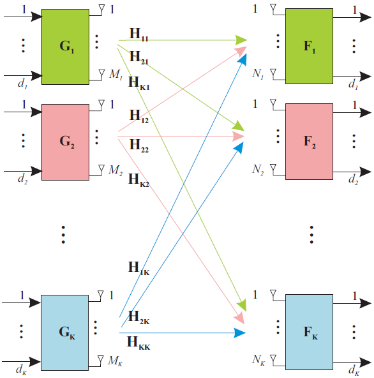
ITA, February 2012

Outline

- Introduction to the Noisy MIMO Interference Channel (IFC)
- Signal Space Interference Alignment (IA)
- Centralized CSIT Acquisition
- Distributed CSIT Acquisition
- Output Channel Feedback
- Concluding Remarks

MIMO IFC Introduction

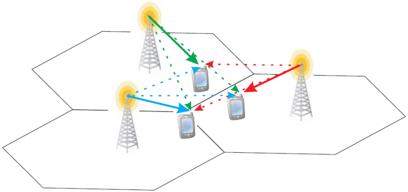
- Interference Alignment (IA) was introduced in [Cadambe,Jafar 2008]
- The objective of IA is to design the Tx beamforming matrices such that the interference at each non intended receiver lies in a common interference subspace
- If alignment is complete at the receiver simple Zero Forcing (ZF) can suppress interference and extract the desired signal
- In [SPAWC2010] we derive a set of interference alignment (IA) feasibility conditions for a K-link frequency-flat MIMO interference channel (IFC)
- $d = \sum_{k=1}^{K} d_k$



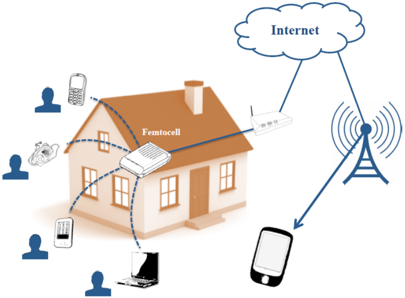
MIMO Interference Channel

Possible Application Scenarios

- Multi-cell cellular systems, modeling intercell interference.
  Difference from Network MIMO: no exchange of signals, "only" of channel impulse responses.



- Coexistence of cellular and femto-cells, especially when femtocells are considered part of the cellular solution.



Why IA?

- The number of streams (degrees of freedom (dof)) appearing in a feasible IA scenario correspond to prelogs of feasible multi-user rate tuples in the multi-user rate region.
  Max Weighted Sum Rate (WSR) becomes IA at high SNR.
- Noisy IFC: interfering signals are not decoded but treated as (Gaussian) noise.
  Apparently enough for dof.
- Lots of recent work more generally on rate prelog regions: involves time sharing, use of fractional power.

Noisy MIMO IFC: Some State of the Art

- IA: alternating ZF algorithm [Jafar etal: globecom08],[Heath etal: icassp09].
- IA feasibility: - $K = 2$ MIMO: [JafarFakhereddin:IT07]
  - [Yetis,Jafar:T10], [Slock etal:eusipco09,ita10, spawc10]
  - 3xNxN, 3xMxN: [BreslerTse:arxiv11]
- max WSR: single stream/link
  - approximately: max SINR [Jafar etal: globecom08]
  - eigenvector interpretation of WSR gradient w.r.t. BF: starting [Honig,Utschick:asilo09]
  - added DA-style approach in [Honig,Utschick:allerton10]
- max WSR: multiple streams/link
  - [Slock etal:ita10] application of [Christensen etal:TW08] from MIMO BC
  - further refined in [Negro etal:allerton10], independently suggested use of DA, developed in [Negro etal:ita11]

IA as a Constrained Compressed SVD

- $F_k^H : \; d_k \times N_k, \;\; H_{ki} : \; N_k \times M_i, \;\; G_i : \; M_i \times d_i \qquad F^H H G =$

$$\begin{bmatrix} F_1^H & 0 & \cdots & 0 \\ 0 & F_2^H & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1K} \\ H_{21} & H_{22} & \cdots & H_{2K} \\ \vdots & & \ddots & \vdots \\ H_{K1} & H_{K2} & \cdots & H_{KK} \end{bmatrix} \begin{bmatrix} G_1 & 0 & \cdots & 0 \\ 0 & G_2 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & G_K \end{bmatrix} = \begin{bmatrix} F_1^H H_{11} G_1 & 0 & \cdots & 0 \\ 0 & F_2^H H_{22} G_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H H_{KK} G_K \end{bmatrix}$$

  F^H, G can be chosen to be unitary for IA

- per user vs per stream approaches:

  IA: can absorb the $d_k \times d_k$ $F_k^H H_{kk} G_k$ in either F_k^H (per stream LMMSE Rx) or G_k or both.

  WSR: can absorb unitary factors of SVD of $F_k^H H_{kk} G_k$ in F_k^H, G_k without loss in rate $\Rightarrow F^H H G =$ diagonal.

Interference Alignment: Feasibility Conditions (1)

- To derive the existence conditions we consider the ZF conditions

$$\underbrace{\mathbf{F}_k^H}_{d_k \times N_k} \underbrace{\mathbf{H}_{kl}}_{N_k \times M_l} \underbrace{\mathbf{G}_l}_{M_l \times d_l} = \mathbf{0} \ , \quad \forall l \neq k$$

$$\text{rank}(\mathbf{F}_k^H \mathbf{H}_{kk} \mathbf{G}_k) = d_k \ , \quad \forall k \in \{1, 2, \dots, K\}$$

- rank requirement $\Rightarrow$ SU MIMO condition: $d_k \leq \min(M_k, N_k)$
- The total number of variables in $\mathbf{G}_k$ is
  $d_k M_k - d_k^2 = d_k(M_k - d_k)$
  Only the subspace of $\mathbf{G}_k$ counts, it is determined up to a $d_k \times d_k$ mixture matrix.
- The total number of variables in $\mathbf{F}_k^H$ is
  $d_k N_k - d_k^2 = d_k(N_k - d_k)$
  Only the subspace of $\mathbf{F}_k^H$ counts, it is determined up to a $d_k \times d_k$ mixture matrix.

Interference Alignment: Feasibility Conditions (2)

- A solution for the interference alignment problem can only exist if the total number of variables is greater than or equal to the total number of constraints i.e.,

$$\sum_{k=1}^{K} d_k(M_k - d_k) + \sum_{k=1}^{K} d_k(N_k - d_k) \geq \sum_{i \neq j=1}^{K} d_i \, d_j$$

$$\Rightarrow \sum_{k=1}^{K} d_k(M_k + N_k - 2d_k) \geq (\sum_{k=1}^{K} d_k)^2 - \sum_{k=1}^{K} d_k^2$$

$$\Rightarrow \sum_{k=1}^{K} d_k(M_k + N_k) \geq (\sum_{k=1}^{K} d_k)^2 + \sum_{k=1}^{K} d_k^2$$

- In the symmetric case: $d_k = d$, $M_k = M$, $N_k = N$:
  $d \leq \frac{M+N}{K+1}$
- For the $K = 3$ user case ($M = N$): $d = \frac{M}{2}$.
  With 3 parallel MIMO links, half of the (interference-free) resources are available!
  However $d \leq \frac{1}{(K+1)/2} M < \frac{1}{2} M$ for $K > 3$.

MWSR: Maximum Weighted Sum Rate (WSR)

The received signal at the $k-$th receiver is:

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

Introduce the interference plus noise covariance matrix at receiver k: $\mathbf{R}_{\bar{k}} = \mathbf{R}_{nn} + \sum_{l\neq k} \mathbf{H}_{kl}\mathbf{G}_l\mathbf{G}_l^H\mathbf{H}_{kl}^H$.
The WSR criterion is

$$\mathcal{R} = \sum_{k=1}^{K} \mathrm{u}_k \log\det(\mathbf{I} + \mathbf{G}_k^H\mathbf{H}_{kk}^H\mathbf{R}_{\bar{k}}^{-1}\mathbf{H}_{kk}\mathbf{G}_k) \tag{1}$$

$$\text{s.t.} \quad \mathrm{Tr}\{\mathbf{G}_k^H\mathbf{G}_k\} \leq P_k$$

This criterion is highly non convex in the Tx BFs $\mathbf{G}_k$.

WSR per stream

- Augmented WSR cost function: BFs $\mathbf{g}_{kn}$ plus Rx filters $\mathbf{f}_{kn}$ and weights w_{kn}:

$$\mathcal{O} = -\sum_{k=1}^{K} u_k \sum_{n=1}^{d_k} (-\ln(w_{kn}) + w_{kn}(1 - \mathbf{f}_{kn}^H \mathbf{H}_{kk} \mathbf{g}_{kn})(1 - \mathbf{f}_{kn}^H \mathbf{H}_{kk} \mathbf{g}_{kn})^H$$

$$+\mathbf{f}_{kn}^H \underbrace{(\mathbf{R}_{v_k} + \sum_{(im)\neq(kn)} \mathbf{H}_{ki} \mathbf{g}_{im} \mathbf{g}_{im}^H \mathbf{H}_{ki}^H)}_{\mathbf{R}_{\overline{kn}}} \mathbf{f}_{kn}) + \sum_{k=1}^{K} \lambda (P_k - \sum_{n=1}^{d_k} \mathbf{g}_{kn}^H \mathbf{g}_{kn}) \quad (2)$$

Alternating optimization $\Rightarrow$ quadratic or convex subproblems:

- Opt w_{kn} given $\mathbf{g}_{kn}$, $\mathbf{f}_{kn}$: $-\ln(w_{kn}) + w_{kn} e_{kn} \Rightarrow w_{kn} = e_{kn}^{-1}$
- Opt $\mathbf{f}_{kn}$ given w_{kn}, $\mathbf{g}_{kn}$: MMSE solution
- Opt $\mathbf{f}_{kn}$ given w_{kn}, $\mathbf{g}_{kn}$: MMSE-style solution (UL-DL duality)

State of the Art on MIMO IFC w Partial CSI

- MISO BC (MU-MISO DL) w CSIT acquisition: [KobayashiCaireJindal:IT10]
- TDD MISO BC w CSIT acquisition: [SalimSlock:JWCN11]
- Space-Time Coding for Analog Channel Feedback: [ChenSlock:isit08]
- [NegroShenoySlockGhauri: eusipco09]: TDD MIMO IFC IA iterative design via UL/DL duality and TDD reciprocity
- Interference Alignment with Analog CSI Feedback: [ElAyachHeath:Milcom10]
  Centralized approach: BS's are connected to a central unit gathering all CSI, performing BF computations and redistributing BF's.
- [Jafar:GLOBECOM10] Blind IA
- [VazeVaranasi:submIT] DoF region for MIMO IFC with feedback
- [SuhTse:IT11] GDoF for IFC with feedback

Key Points

- Distributed approach: no other connectivity assumed than the UL/DL IFC. FB over reversed IFC
- "distributed" = "duplicated" (decentralized)
- A distributed approach does not have to be iterative. It can be done with a finite overhead (finite prelog loss) and finite SNR loss compared to full CSI, even as SNR $\rightarrow \infty$. Hence of interest compared to non-coherent (no/outdated CSIT) IFC approaches.
- Distributed ($\mathcal{O}(K^2)$) requires more FB than Centralized ($\mathcal{O}(K)$).
- centralized/decentralized IFC CSIT estimation (only exchange of data at temporal coherence variation rate), vs NW-MIMO/CoMP (exchange of data at symbol/sample rate)
- Multiple Rx antennas $\Rightarrow$ Rx training also crucial!
- TDD vs FDD, depends on distributed/centralized.
- Channel FB vs Output feedback (OFB)
- "Practical" scheme far from unique

Signal Structure w Partial CSI

- Perfect CSI:
  Rx signal at the $k-$th receiver :

$$\mathbf{y}_k = \sum_{i=1}^{K} \sum_{m=1}^{d_i} \mathbf{H}_{ki} \mathbf{g}_{i,m} \, x_{i,m} + \mathbf{v}_k$$

  Estimate stream (k, n):

$$\widehat{x}_{k,n} = \mathbf{f}_{k,n} \mathbf{H}_{kk} \mathbf{g}_{k,n} \, x_{k,n} + \sum_{i=1}^{K} \sum_{m \neq n} \mathbf{f}_{k,n} \mathbf{H}_{ki} \mathbf{g}_{i,m} \, x_{i,m} \mathbf{f}_{k,n} \mathbf{v}_k$$

- Imperfect CSI:

$$\underbrace{\widehat{\widehat{\mathbf{f}}}_{k,n}}_{\text{est. at Rx } k} \qquad \underbrace{\mathbf{H}_{ki}}_{\text{true}} \qquad \underbrace{\widehat{\mathbf{g}}_{i,m}}_{\text{est. at Tx } i}$$

Rate Lower Bound

- signal of interest in direct link:

$$\widehat{\widehat{\mathbf{f}}}_{k,n}\mathbf{H}_{kk}\,\widehat{\mathbf{g}}_{k,n} = \underbrace{\widehat{\widehat{\mathbf{f}}}_{k,n}\widehat{\mathbf{H}_{kk}\,\widehat{\mathbf{g}}_{k,n}}}_{\text{known to Rx}} + \underbrace{\widehat{\widehat{\mathbf{f}}}_{k,n}\widetilde{\mathbf{H}_{kk}\,\widehat{\mathbf{g}}_{k,n}}}_{\text{put in interf.}}$$

1. Bound loss of partial CSI ergodic rate to full CSI ergodic rate.

$$\text{e.g.} \quad \mathcal{R}_k^{PCSI}(\rho) \leq (1 - \frac{T_{overhead}}{T})\, \mathcal{R}_k^{FCSI}(\rho/\alpha_k)$$

-
$$\begin{aligned} \widehat{\widehat{\mathbf{f}}}_{k,n}\, \mathbf{H}_{ki}\, \widehat{\mathbf{g}}_{i,m} &= (\mathbf{f}_{k,n} + \widetilde{\widetilde{\mathbf{f}}}_{k,n})\, \mathbf{H}_{ki}\, (\mathbf{g}_{i,m} + \widetilde{\mathbf{g}}_{i,m}) \\ &= \mathbf{f}_{k,n}\, \mathbf{H}_{k,i}\, \mathbf{g}_{i,m} + 3 \text{ error terms} \end{aligned}$$

2. Bound loss of partial CSI ergodic rate to full CSI ergodic rate for case of channel pdf = that of the estimated channel: provides closer bounds, but requires ergodic rate expressions with different channel statistics.

   - $$\begin{aligned}\widehat{\widehat{\mathbf{f}}}_{k,n}\,\mathbf{H}_{ki}\,\widehat{\mathbf{g}}_{i,m} &= (\widehat{\mathbf{f}}^{(i)}_{k,n}+\widetilde{\widehat{\mathbf{f}}}^{(i)}_{k,n})\,(\widehat{\mathbf{H}}^{(i)}_{ki}+\widetilde{\mathbf{H}}^{(i)}_{ki})\,\widehat{\mathbf{g}}_{i,m}\\ &= \widehat{\mathbf{f}}^{(i)}_{k,n}\widehat{\mathbf{H}}^{(i)}_{ki}\,\widehat{\mathbf{g}}_{i,m}+3\text{ error terms}\end{aligned}$$

3 Partial CSI Rate Analysis Approaches (3)

3. High SNR ρ rate asymptote: $\mathcal{R} = a\log(\rho) + b + \mathcal{O}(1/\rho)$
   a: multiplexing gain (prelog, dof), b: rate offset a, b independent of:
   - MMSE regularization (MMSE-ZF filters suffice)
   - optimized WF (uniform WF suffices)
   - LMMSE channel estimation (becomes deterministic estimation)
   -

$$
\begin{aligned}
\widehat{\widehat{\mathbf{f}}}_{k,n}\,\mathbf{H}_{ki}\,\widehat{\mathbf{g}}_{i,m} &= (\widehat{\mathbf{f}}^{(i)}_{k,n} + \widetilde{\widehat{\mathbf{f}}}^{(i)}_{k,n})\,(\widehat{\mathbf{H}}^{(i)}_{ki} + \widetilde{\mathbf{H}}^{(i)}_{ki})\,\widehat{\mathbf{g}}_{i,m} \\
&= \underbrace{\widehat{\mathbf{f}}^{(i)}_{k,n}\widehat{\mathbf{H}}^{(i)}_{ki}\widehat{\mathbf{g}}_{i,m}}_{=0} + \mathbf{f}_{k,n}\,\widetilde{\mathbf{H}}^{(i)}_{ki}\,\mathbf{g}_{i,m} + \widetilde{\widehat{\mathbf{f}}}^{(i)}_{k,n}\,\mathbf{H}_{ki}\,\mathbf{g}_{i,m}
\end{aligned}
$$

High SNR Rate Analysis

- Asymptote $\mathcal{R} = a\log(\rho) + b$ permits meaningful optimization for finite (but high) SNR, and may lead to more than minimal FB.
- At very high SNR ρ, only rate prelog a (dof) counts. Its maximization requires FB to be minimal (channel just identifiable).
- At moderate SNR, finding an optimal compromise between estimation overhead and channel quality will involve a properly adjusted overhead. However, the overhead issue is not the only reason for a possibly diminishing multiplexing gain a as SNR decreases, also reducing the number of streams $\{d_k\}$ may lead to a better compromise (as for full CSI).
- The rate offset b is already a non-trivial rate characteristic even in the full CSI case. b may increase as the number of streams decreases, due to reduced noise enhancement.

- Doppler Spectrum is bandlimited to $1/T$ ($1/D$ in figure)
- Nyquist's Theorem : downsampling possible with factor T
- Vectorize channel coefficients over T, matrix spectrum of rank 1, MIMO prediction error of rank 1.
- Hence channel coefficient evolution during current "coherence period" T is along a single basis vector, plus prediction from past.
- Block fading: basis vector = rectangular window and prediction from the past = 0

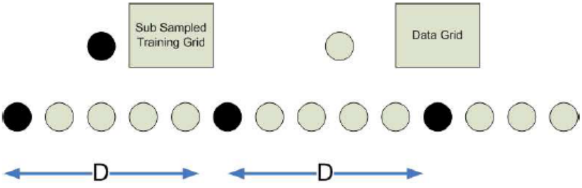
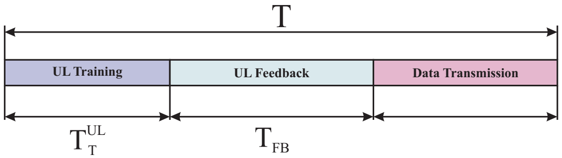


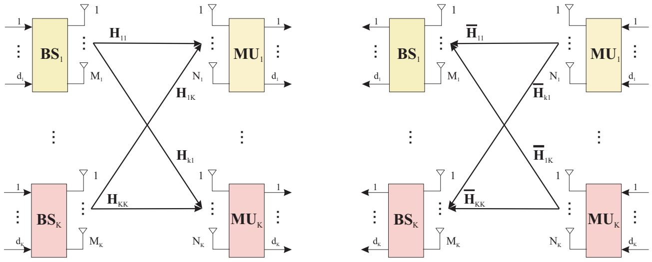
Figure 1: Subsampling Grid.

Centralized Approach

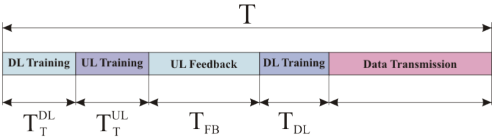


- Proposed by Heath [Milcom10,arxiv]
- The authors extrapolate the single antenna case, where only the estimate of the overall ch-BF gain and associated SINR is required
- In the MIMO IFC Rx not only needs to estimate the ch-BF cascade but also the I+N covariance matrix
- Not trivial. Training length similar as for the BF determination (order K) is required.
- Rate analysis of type 1.

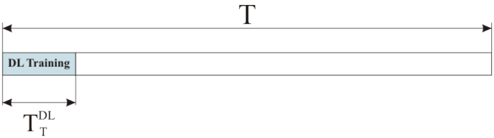
FDD Communication



- We Assume FDD transmission scheme
- Downlink channel matrix $\mathbf{H}_{ki}$ from BS_i to MU_k
- Uplink channel matrix $\overline{\mathbf{H}}_{ik}$ from MU_k to BS_i
- Analyze both centralized and distributed approaches.



- We consider a block fading channel model with Coherence time interval T
- The general channel matrix $\mathbf{H}_{ik} \sim \mathcal{N}(0, \mathbf{I})$
- To acquire the necessary CSI at BS and MU side several training and feedback phases are necessary
- Hence a total overhead of T_{ovrhd} channel usage is dedicated to BS-MU signaling
- Only part of the time $T_{data} = T - T_{ovrhd}$ is dedicated to real data transmission



Downlink Training Phase

- Each BS_i Tx ($\perp$) pilot sequences with power P_T^{DL}
- MU_k estimates all DL channels connected to it:
  $\mathbf{H}_k = [\mathbf{H}_{k1}, \ldots, \mathbf{H}_{kK}]$
- The duration of the DL training phase is

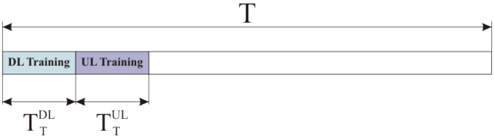
$$T_T^{DL} \geq \sum_{k=1}^{K} M_k$$

- Using MMSE estimation we get $\mathbf{H}_k = \widehat{\mathbf{H}}_k + \widetilde{\mathbf{H}}_k$

$$\widehat{\mathbf{H}}_k \sim \mathcal{N}(0, \frac{P_T^{DL}}{\sigma^2 + P_T^{DL}}\mathbf{I}), \qquad \widetilde{\mathbf{H}}_k \sim \mathcal{N}(0, \frac{\sigma^2}{\sigma^2 + P_T^{DL}}\mathbf{I})$$

we call $\sigma^2_{\widetilde{H}}$ and $\sigma^2_{\widehat{H}}$ the variance of the error and estimate

Uplink Training Phase



Uplink Training Phase (dual of DL Training Phase)

- Each MU_i sends a set of pilots symbols with power P_T^{UL}
- BS_k estimate the compound UL channel matrix: $\overline{\mathbf{H}}_k = [\overline{\mathbf{H}}_{k1}, \ldots, \overline{\mathbf{H}}_{kK}]$
- The duration of the UL training phase is

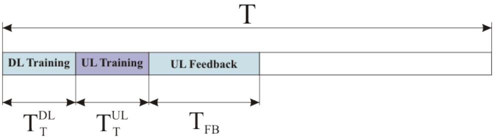
$$T_T^{UL} \geq \sum_{k=1}^{K} N_k$$

- Using MMSE estimation we get $\overline{\mathbf{H}}_k = \widehat{\overline{\mathbf{H}}}_k + \widetilde{\overline{\mathbf{H}}}_k$

$$\widehat{\overline{\mathbf{H}}}_k \sim \mathcal{N}(0, \frac{P_T^{UL}}{\sigma^2 + P_T^{UL}}\mathbf{I}), \qquad \widetilde{\overline{\mathbf{H}}}_k \sim \mathcal{N}(0, \frac{\sigma^2}{\sigma^2 + P_T^{UL}}\mathbf{I})$$

we call $\sigma^2_{\tilde{\overline{H}}}$ and $\sigma^2_{\hat{\overline{H}}}$ the variance of the error and estimate

Uplink Feedback Phase



- After the UL and DL training phases each device knows all channels directly connected to it
- To compute the Tx beamformers, complete IFC channel knowledge is required
- Each MU feeds back its channel knowledge (CFB) using *Analog Feedback*
- Two different approaches are possible:
  (a) Centralized Processing
  (b) Distributed Computation

(a) A Central Controller acquires complete CSI and computes all the BF, and disseminates this information.

(b) Each BS acquires complete CSI to compute all the BF, then uses only it own BF.

Uplink Feedback Phase: Centralized Processing

- The received symbol vector received at each BS is sent to the Central Controller for the estimation of DL channels. Staking all the received symbols together we get:

$$\overline{\mathbf{Y}} = \sqrt{P_{FB}} \underbrace{\begin{bmatrix} \overline{\mathbf{H}}_{11} & \cdots & \overline{\mathbf{H}}_{1K} \\ \vdots & \ddots & \vdots \\ \overline{\mathbf{H}}_{K1} & \cdots & \overline{\mathbf{H}}_{KK} \end{bmatrix}}_{M \times N} \underbrace{\begin{bmatrix} \widehat{\mathbf{H}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{H}}_2 & \cdots & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \cdots & \widehat{\mathbf{H}}_K \end{bmatrix}}_{N \times KM} \underbrace{\begin{bmatrix} \mathbf{\Phi}_1 \\ \vdots \\ \mathbf{\Phi}_K \end{bmatrix}}_{KM \times T_{FB}} + \underbrace{\begin{bmatrix} \mathbf{V}_1 \\ \vdots \\ \mathbf{V}_K \end{bmatrix}}_{\overline{\mathbf{V}}}$$

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

- To satisfy the identifiability condition the minimum CFB length is

$$T_{FB} \geq \frac{N \times M}{\sum_i \min\{N_i, M_i\}} \propto K$$

- To extract the i-th feedback contribution we use LS estimate based on the UL channel estimate $\widehat{\overline{\mathbf{H}}}_{ik}$

Uplink Feedback Phase: Centralized Processing

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

- Using the UL channel estimate the LS estimator is: $\overline{\mathbf{H}}_i^{LS} = P_{FB}^{-\frac{1}{2}} (\widehat{\overline{\mathbf{H}}}_i^H \widehat{\overline{\mathbf{H}}}_i)^{-1} \widehat{\overline{\mathbf{H}}}_i^H$

$$\widehat{\widehat{\mathbf{H}}}_i = \widehat{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}} \overline{\mathbf{H}}_i^{LS} \widetilde{\overline{\mathbf{H}}}_i \widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_i^{LS} \overline{\mathbf{V}}\mathbf{\Phi}_i = \mathbf{H}_i - \underbrace{\widetilde{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}} \overline{\mathbf{H}}_i^{LS} \widetilde{\overline{\mathbf{H}}}_i \widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_i^{LS} \overline{\mathbf{V}}\mathbf{\Phi}_i}_{\widetilde{\widehat{\mathbf{H}}}_i}$$

- The estimate of the CFB can be written in function of the true DL channel $\mathbf{H}_i$ plus the estimation error $\widetilde{\widehat{\mathbf{H}}}_i$

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

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

Uplink Feedback Phase: Distributed Processing

- The received symbols at BS_k can be described as follows

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

- To satisfy the identifiability condition the minimum CFB length is

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

- To extract the i-th feedback contribution at BS_k we use LS estimate based on the UL channel estimate $\widehat{\overline{\mathbf{H}}}_{ki}$

Uplink Feedback Phase: Distributed Processing

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

- Using the UL channel estimate the LS estimator is: $\overline{\mathbf{H}}_{ki}^{LS} = P_{FB}^{-\frac{1}{2}}(\widehat{\overline{\mathbf{H}}}_{ki}^H\widehat{\overline{\mathbf{H}}}_{ki})^{-1}\widehat{\overline{\mathbf{H}}}_{ki}^H$

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

- The estimate of the CFB can be written in function of the true DL channel $\mathbf{H}_i$ plus the estimation error $\widetilde{\widehat{\mathbf{H}}}_i$

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

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

Uplink Feedback Phase: Distributed Processing

- Simplest precoding: time multiplexing and repetition coding. To allow finer granularity of FB overhead: uses constant amplitude unitary spreading matrices (eg DFT).
- (linear) Rx strategies for analog FB: many variants possible
  - MMSE (Bayesian $\mathbf{H}$), MMSE-ZF (deterministic $\mathbf{H}$)
  - assuming $\widehat{\mathbf{H}}$ correct, accounting for $\widetilde{\mathbf{H}}$
- analog STC: more spatial multiplexing leads to less overhead but less noise enhancement (especially crucial in distributed approach)

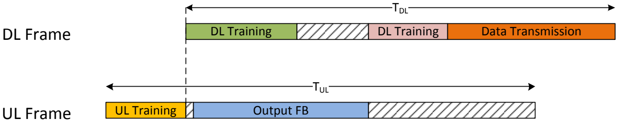
BF Computation Phase

- Once DL channel estimates available, need to perform BF design, e.g. according to Maximum Weighted Sum Rate.
- Can ignore channel estimation errors (full CSIT type design) or acknowledge them (partial CSIT type design).
- Similar considerations for centralized or distributed approaches.

WSR maximization with Partial CSI

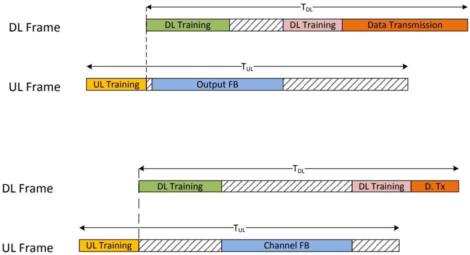
- Model the information of the channel at the transmit side in terms of a Gaussian prior representing mean and covariance information
$$\mathbf{H}_{ij} = \hat{\mathbf{H}}_{ij} + (\mathbf{R}^t_{ij})^{\frac{1}{2}} \tilde{\mathbf{H}}_{ij} (\mathbf{R}^r_{ij})^{\frac{H}{2}} \quad (3)$$
- $\mathbf{R}^t_{ij} = \mathbf{I}$ is the Tx side covariance matrix, $\mathbf{R}^r_{ij} = \tilde{\sigma}^2\mathbf{I}$ is the covariance matrix at the Rx side
- $\tilde{\mathbf{H}}_{ij}$ is a matrix with iid Gaussian, zero mean and unit variance, entries
- The relation between the WSR and the weighted mean squared error (WMSE) to approximate the maximization of the expected WSR with the expectation of the WMSE can be exploited
- This leads to an approximate solution but easy to be handled

Output Feedback



- Each MU feeds back to all BS the noiseless version of its received signal using un-quantized feedback: Output FB (OFB).
- In FDD systems UL and DL transmission can take place at the same time
- T_{UL} represents the UL coherence Time
- T_{DL} represents the DL coherence Time
- OFB phase can start one time instant after the beginning of the DL training phase

Output Feedback



- Output FB allows us to reduce the overhead due to CSI exchange
- In channel FB each MU has to wait the end of the DL training phase before being able to FB DL channel estimates
- For easy of exposition we consider $M_i = N_t \; \forall i$, $N_i = N_r \; \forall i$ where $N_t \geq N_r$

- The Rx signal at MU_k at time [t] during the DL training phase is

$$\mathbf{y}_k[t] = \sum_{i=1}^{K} \mathbf{H}_{ki}\mathbf{s}_i[t] + \mathbf{n}_k[t]$$

- At [t+1] MU_k feeds back to all BSs the noiseless version of the RX signal at time instant [t] (OFB).
- So BS_l receives:

$$\overline{\mathbf{y}}_l[t+1] = \sum_{j=1}^{K} \overline{\mathbf{H}}_{lj}\overline{\mathbf{x}}_j[t+1] + \overline{\mathbf{n}}_k[t+1] = \sum_{j=1}^{K} \overline{\mathbf{H}}_{lj}\alpha_j \sum_{i=1}^{K} \mathbf{H}_{ji}\mathbf{s}_i[t] + \overline{\mathbf{n}}_k[t+1]$$

α_j takes into account the TX power constraint at MU_j

Output Feedback

- In a distributed approach we use time multiplexing to allow all BSs to estimate all the DL channels
- Each BS has to estimate $\mathbf{H}_i = [\mathbf{H}_{i1}, \ldots, \mathbf{H}_{iK}]^{N_r \times KN_t}$, then each BS needs $\tau^o_{FB} = KN_r$ samples.
- The total length of the output feedback phase is:

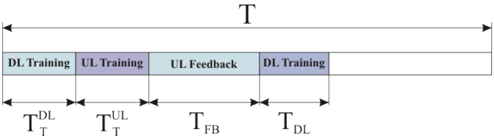
$$T^o_{FB} \geq K^2 N_r$$

- OFB length is the same as channel FB length
- OFB does not reduce FB duration but reduces overhead due to partial elimination of silent periods

Importance of CSIR

- CSIR is usually neglected
- Some schemes for arbitrary time-varying channels assume that Rxs know all channel matrices at all time: impossible to realize in practice
- An additional DL training phase is required to build the Rx filters

Downlink Training Phase



- Once the BFs have been calculated (Centralized/Distributed) they are used in the DL transmission
- Each MU applies a ZF RX filter to suppress interference
- To design Rx filter 2 approaches are possible
  - (a) DL Training
  - (b) Analog Transmission of Rx's

(a) Each BS sends BF'd pilots to estimate ch-BF cascade or Rx

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

(b) The entire Rx filter $\mathbf{F}_k$ is transmitted to the MU_k using analog signaling

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

In the end: Sum Rate (high SNR)

- SR

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

$$T_i \geq T_{i,min}$$

Assume $b_{kni} = b_i$ for what follows.

- Fixing $\sum_i T_i = T_{ovrhd}$, optimal $T_i = T_{ovrhd} \sqrt{b_i}/(\sum_i \sqrt{b_i})$.
- Optimizing over T_{ovrhd} now

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

TDD vs FDD

- Usually TDD transmission scheme is used to simplify the DL CSI acquisition at the BS side
- BS_k learns the DL channel $\mathbf{H}_{ki}$, $\forall i$ through reciprocity
- MU_i do not need to feedback $\mathbf{H}_{ki}$ to BS_k but this channel is required at $BS_{j \neq k}$
- In Distributed Processing reciprocity does NOT help in reducing channel feedback overhead $\Longrightarrow$ TDD equivalent FDD
- In Centralized Processing reciprocity makes channel feedback NOT required
- In what follow we concentrate on FDD transmission strategy

Further Optimizing DoF

- data Tx stage (as good as perfect CSI):
  - Can FB increase DoF with perfect CSIT?
    According to [HuangJafar:IT09] and [VazeVaranasi:ITsubm11]
    NO for $K = 2$ MIMO IFC; $K > 2$ is OPEN.
  - If not in general, then use of OFB is mainly (only) for CSI acquisition, not for augmenting DoF in presence of CSIT
- CSI acquisition stages:
  - Optimize number of streams/number of active antennas for small T: if less channel to learn then more time to Tx data, even if on reduced number of streams
  - Instead of going from $K = 1$ to full K immediately, could gradually increase number of interfering links (and their CSI acquisition) from 1 to K.
  - When T gets too short: delayed CSIT approaches.
  - A single (the largest) MIMO link can start transmitting right away w/o CSIT (possibly w/o CSIR also).

From Practical to More Optimal at Finite SNR

- When (analog) channel FB is of extended (non-minimal) duration, BF's can get computed and some DL transmission could start while FB is still going on. No need to wait until all CSI is gathered before transmission can get started.
- rate constants: partial CSI Tx/Rx design, diversity issues (optimized IA)
- optimization of training duration/power

Perspectives

- can OFB increase dof w perfect CSIT for $K \geq 3$?
- need to handle CSIR also in delayed CSIT approaches
- users with difference coherence time
- full duplex operation (2-way communications)
- minimum reciprocity:
  coherence times equal on UL and DL, feasible dof same on UL and DL
- real IFC system: doubly selective