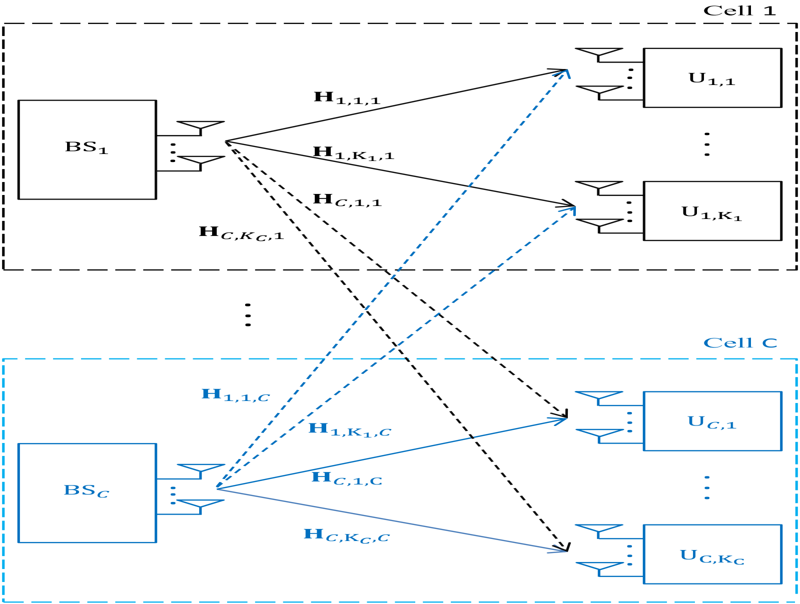
The Pathwise MIMO Interfering Broadcast Channel

Wassim Tabikh, Dirk Slock & Yi-Yuan-Wu[#]

Mobile Communications Department, EURECOM
[#]Orange Labs, Issy-les-Moulineaux, France

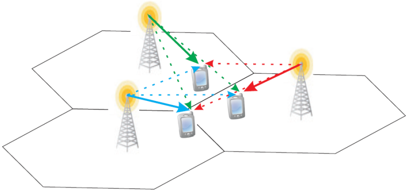
Information Theory and Applications (ITA) Workshop
San Diego, CA, Feb. 2, 2015

MIMO Interfering Broadcast Channel (IBC)

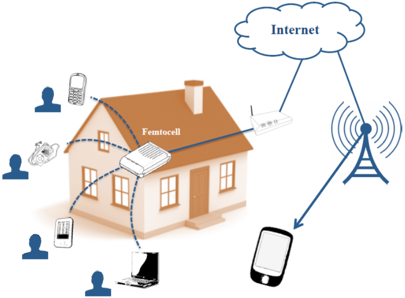


Possible Application Scenarios

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



- HetNets: Coexistence of macrocells and small cells, especially when small cells are considered part of the cellular solution.



Weighted Sum Rate w Partial CSIT: Motivation

- Interference Alignment (IA): showed that interference is not a fatality, essentially half of the interference-free rate can be attained
- requires perfect global CSIT, focuses on high SNR
- various IA flavors:
  asymptotic IA (infinite symbol extension & symbol extension),
  ergodic IA [NazerGastparJafarVishwanath:IT1012] explains the factor 2 loss in DoF of SISO IA w.r.t. an interference-free Tx scenario by transmitting the same signal twice at two paired channel uses in which all cross channel links cancel out each other: group channel realizations H_1, H_2 s.t. $\text{offdiag}(H_2) = -\text{offdiag}(H_1)$.
  Linear IA or Signal Space IA or MIMO IA: use MIMO dimensions: no latency, but DoF factor $\dfrac{2}{K+1}$: considered here.
- MIMO IA: joint Tx/Rx ZF
- at finite SNR: Tx & Rx not ZF but MMSE-style
- CSIT in practice: never perfect, partial due to feedback (estimation, quantization, noisy FB) or worse: e.g. covariance CSIT, location based, etc. (esp. for intercell CSIT)

Introduction

- Linear IA for MIMO (I)(B)C:design of any BS Tx filter depends on all Rx filters whereas in turn each Rx filter depends on all Tx filters [Negro:ita11]. As a result, all Tx/Rx filters are globally coupled and their design requires global CSIT.
- Massive MIMO from a DoF point of view it may seem like a suboptimal use of antennas. However [Wagner:IT0712, section V, Fig. 6], K^{opt} decreases as SNR decreases (also: Net DoF considerations and CSI acquisition).
- [Caire:IT1013],[Caire:mmWave]: MISO single cell. Statistical CSIT between user groups, instantaneous CSIT within user groups. Hypothesis: 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.
- [Lau:Hierarchical:SP0413] MISO hierarchical approach : Intercell statistical CSIT ZF BF, treating interfering links in a binary fashion (either ZF or ignore). Intracell BF is based on instantaneous CSIT and performs Regularized-ZF.
- "Massive MIMO" = "MU Massive MISO". Here: actual MU MC Massive MIMO.

Introduction (2)

- Whereas path CSIT by itself may allow ZF, which is of interest at high SNR, we are particularly concerned here with maximum Weighted Sum Rate (WSR) designs accounting for finite SNR.
- Massive MIMO makes the pathwise approach viable: 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 ZF requires much less coordination).

Key Points

- state of the art:
  - MIMO IA requires global MIMO channel CSIT
  - recent works focus on intercell exchange of only scalar quantities, at fast fading rate
  - 2-stage approaches with intercell interference zero-forced
- Massive MISO $\rightarrow$ Massive MIMO
- Pathwise: allows updating at slow fading rate
- Pathwise: may allow IA at reduced DoF
- don't waste resources for ZF but do max weighted sum rate
- Key contributions:
  - pathwise beamformer via uplink-downlink LMMSE duality
  - not only for data Tx but also for channel estimation

Max WSR Tx BF Design Outline

- max WSR Tx BF design with perfect CSIT
  - using WSR - WSMSE relation
  - from difference of concave to linearized concave
  - MIMO BC: local optima, deterministic annealing
- Gaussian partial CSIT
- max EWSR Tx design w partial CSIT
- Line of Sight (LoS) based partial CSIT
- max EWSR Tx design with LoS based CSIT

MIMO IBC with Linear Tx/Rx, single stream

- IBC with C cells with a total of K users. System-wide user numbering: the $N_k \times 1$ Rx signal at user k in cell b_k is

$$y_k = \underbrace{\mathbf{H}_{k,b_k}\,\mathbf{g}_k\,x_k}_{\text{signal}} + \underbrace{\sum_{\substack{i \neq k \\ b_i = b_k}} \mathbf{H}_{k,b_k}\,\mathbf{g}_i\,x_i}_{\text{intracell interf.}} + \underbrace{\sum_{j \neq b_k} \sum_{i:b_i=j} \mathbf{H}_{k,j}\,\mathbf{g}_i\,x_i}_{\text{intercell interf.}} + v_k$$

  where x_k = intended (white, unit variance) scalar signal stream, $\mathbf{H}_{k,b_k} = N_k \times M_{b_k}$ channel from BS b_k to user k. BS b_k serves $K_{b_k} = \sum_{i:b_i=b_k} 1$ users. Noise whitened signal representation $\Rightarrow$ $v_k \sim \mathcal{CN}(0, I_{N_k})$.
- The $M_{b_k} \times 1$ spatial Tx filter or 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 $\widehat{x}_k = \mathbf{f}_k^H y_k$

$$\begin{aligned}
\widehat{x}_k &= \mathbf{f}_k^H \mathbf{H}_{k,b_k}\,\mathbf{g}_k\,x_k + \sum_{i=1,\neq k}^{K} \mathbf{f}_k^H \mathbf{H}_{k,b_i}\,\mathbf{g}_i\,x_i + \mathbf{f}_k^H v_k \\
&= \mathbf{f}_k^H \mathbf{h}_{k,k}\,x_k + \sum_{i \neq k} \mathbf{f}_k^H \mathbf{h}_{k,i}\,x_i + \mathbf{f}_k^H v_k
\end{aligned}$$

  where $\mathbf{h}_{k,i} = \mathbf{H}_{k,b_i}\,\mathbf{g}_i$ is the channel-Tx cascade vector.

Max Weighted Sum Rate (WSR)

- Weighted sum rate (WSR)

$$WSR = WSR(\mathbf{g}) = \sum_{k=1}^{K} u_k \ \ln \frac{1}{e_k}$$

where $\mathbf{g} = \{\mathbf{g}_k\}$, the u_k are rate weights

- MMSEs $e_k = e_k(\mathbf{g})$

$$\frac{1}{e_k} = 1 + \mathbf{g}_k^H \mathbf{H}_k^H R_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k = (1 - \mathbf{g}_k^H \mathbf{H}_k^H R_k^{-1} \mathbf{H}_k \mathbf{g}_k)^{-1}$$
$$R_k = R_{\overline{k}} + \mathbf{H}_k \mathbf{g}_k \mathbf{g}_k^H \mathbf{H}_k^H \ , \ R_{\overline{k}} = \sum_{i \neq k} \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H + I_{N_k} \ ,$$

R_k, $R_{\overline{k}}$ = total, interference plus noise Rx cov. matrices resp.

- MMSE e_k obtained at the output $\widehat{x}_k = \mathbf{f}_k^H y_k$ of the optimal (MMSE) linear Rx

$$\mathbf{f}_k = R_k^{-1} \mathbf{H}_k \mathbf{g}_k \ .$$

From max WSR to min WSMSE

- For a general Rx filter $\mathbf{f}_k$ we have the MSE $e_k(\mathbf{f}_k, \mathbf{g})$
$= (1 - \mathbf{f}_k^H \mathbf{H}_k \mathbf{g}_k)(1 - \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{f}_k) + \sum_{i \neq k} \mathbf{f}_k^H \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H \mathbf{f}_k + ||\mathbf{f}_k||^2$
$= 1 - \mathbf{f}_k^H \mathbf{H}_k \mathbf{g}_k - \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{f}_k + \sum_i \mathbf{f}_k^H \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H \mathbf{f}_k + ||\mathbf{f}_k||^2.$
- The $WSR(\mathbf{g})$ is a non-convex and complicated function of $\mathbf{g}$. Inspired by [Christensen:TW1208], we introduced [Negro:ita10],[Negro:ita11] an augmented cost function, the Weighted Sum MSE, $WSMSE(\mathbf{g}, \mathbf{f}, w)$

$$= \sum_{k=1}^{K} u_k (w_k \; e_k(\mathbf{f}_k, \mathbf{g}) - \ln w_k) + \lambda (\sum_{k=1}^{K} ||\mathbf{g}_k||^2 - P)$$

where $\lambda =$ Lagrange multiplier and $P =$ Tx power constraint.
- After optimizing over the aggregate auxiliary Rx filters $\mathbf{f}$ and weights w, we get the WSR back:

$$\min_{\mathbf{f}, w} WSMSE(\mathbf{g}, \mathbf{f}, w) = -WSR(\mathbf{g}) + \overbrace{\sum_{k=1}^{K} u_k}^{\text{constant}}$$

From max WSR to min WSMSE (2)

- Advantage augmented cost function: alternating optimization
  $\Rightarrow$ solving simple quadratic or convex functions

$$\min_{w_k} WSMSE \;\Rightarrow\; w_k = 1/e_k$$

$$\min_{\mathbf{f}_k} WSMSE \;\Rightarrow\; \mathbf{f}_k = (\sum_i \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H + I_{N_k})^{-1} \mathbf{H}_k \mathbf{g}_k$$

$$\min_{\mathbf{g}_k} WSMSE \;\Rightarrow$$

$$\mathbf{g}_k = (\textstyle\sum_i u_i w_i \mathbf{H}_i^H \mathbf{f}_i \mathbf{f}_i^H \mathbf{H}_i + \lambda I_M)^{-1} \mathbf{H}_k^H \mathbf{f}_k u_k w_k$$

- UL/DL duality: optimal Tx filter g_k of the form of a MMSE linear Rx for the dual UL in which λ plays the role of Rx noise variance and $u_k w_k$ plays the role of stream variance.

Optimal Lagrange Multiplier λ

- (bisection) line search on $\sum_{k=1}^{K} ||\mathbf{g}_k||^2 - P = 0$ [Luo:SP0911].
- Or updated analytically as in [Negro:ita10],[Negro:ita11] by exploiting $\sum_k \mathbf{g}_k^H \frac{\partial WSMSE}{\partial \mathbf{g}_k^*} = 0$.
- This leads to the same result as in [Hassibi:TWC0906]: λ avoided by reparameterizing the BF to satisfy the power constraint: $\mathbf{g}_k = \sqrt{\frac{P}{\sum_{i=1}^{K} ||\mathbf{g}_i^{'}||^2}}\, \mathbf{g}_k^{'}$ with $\mathbf{g}_k^{'}$ now unconstrained

$$\text{SINR}_k = \frac{|\mathbf{f}_k \mathbf{H}_k \mathbf{g}_k^{'}|^2}{\sum_{i=1,\neq k}^{K} |\mathbf{f}_k \mathbf{H}_k \mathbf{g}_i^{'}|^2 + \frac{1}{P} ||\mathbf{f}_k||^2 \sum_{i=1}^{K} ||\mathbf{g}_i^{'}||^2} \; .$$

- This leads to the same Lagrange multiplier expression obtained in [Christensen:TW1208] on the basis of a heuristic that was introduced in [Joham:isssta02] as was pointed out in [Negro:ita10].

From WSR to WSSINR

- The WSR can be rewritten as

$$WSR = WSR(\mathbf{g}) = \sum_{k=1}^{K} u_k \ \ln(1 + \mathrm{SINR}_k)$$

where $1 + \mathrm{SINR}_k = 1/e_k$ or for general $\mathbf{f}_k$:

$$\mathrm{SINR}_k = \frac{|\mathbf{f}_k \mathbf{H}_k \mathbf{g}_k|^2}{\sum_{i=1, \neq k}^{K} |\mathbf{f}_k \mathbf{H}_k \mathbf{g}_i|^2 + ||\mathbf{f}_k||^2} \ .$$

- WSR variation

$$\partial WSR = \sum_{k=1}^{K} \frac{u_k}{1 + \mathrm{SINR}_k} \ \partial \mathrm{SINR}_k$$

interpretation: variation of a weighted sum SINR (WSSINR)

- The BFs obtained: same as for WSR or WSMSE criteria. But this interpretation shows: WSR = optimal approach to the SLNR or SJNR heuristics. WSSINR approach = [KimGiannakis:IT0511] below.

[KimGiannakis:IT0511] Difference of Convex fns (DC)

- Let $Q_k = \mathbf{g}_k \mathbf{g}_k^H$ be the transmit covariance for stream k $\Rightarrow$

$$WSR = \sum_{k=1}^{K} u_k [\ln\det(R_k) - \ln\det(R_{\overline{k}})]$$

w $R_k = \mathbf{H}_k (\sum_i Q_i) \mathbf{H}_k^H + I_{N_k}$, $R_{\overline{k}} = \mathbf{H}_k (\sum_{i \neq k} Q_i) \mathbf{H}_k^H + I_{N_k}$.

- Consider the dependence of WSR on Q_k alone:

$$WSR = u_k \ln\det(R_{\overline{k}}^{-1} R_k) + WSR_{\overline{k}}, \; WSR_{\overline{k}} = \sum_{i=1,\neq k}^{K} u_i \ln\det(R_{\overline{i}}^{-1} R_i)$$

where $\ln\det(R_{\overline{k}}^{-1} R_k)$ is concave in Q_k and $WSR_{\overline{k}}$ is convex in Q_k. Since a linear function is simultaneously convex and concave, consider the first order Taylor series expansion in Q_k around $\widehat{Q}$ (i.e. all $\widehat{Q}_i$) with e.g. $\widehat{R}_i = R_i(\widehat{Q})$, then

$$WSR_{\overline{k}}(Q_k, \widehat{Q}) \approx WSR_{\overline{k}}(\widehat{Q}_k, \widehat{Q}) - \text{tr}\{(Q_k - \widehat{Q}_k)\widehat{\mathbf{A}}_k\}$$

$$\widehat{\mathbf{A}}_k = -\left.\frac{\partial WSR_{\overline{k}}(Q_k, \widehat{Q})}{\partial Q_k}\right|_{\widehat{Q}_k, \widehat{Q}} = \sum_{i=1,\neq k}^{K} u_i \mathbf{H}_i^H (\widehat{R}_{\overline{i}}^{-1} - \widehat{R}_i^{-1}) \mathbf{H}_i$$

[KimGiannakis:IT0511] DC (2)

- Note that the linearized (tangent) expression for $WSR_{\overline{k}}$ constitutes a lower bound for it.
- Now, dropping constant terms, reparameterizing $Q_k = \mathbf{g}_k \mathbf{g}_k^H$ and performing this linearization for all users,

$$WSR(\mathbf{g}, \widehat{\mathbf{g}}) = \sum_{k=1}^{K} u_k \ln(1 + \mathbf{g}_k^H \mathbf{H}_k^H \widehat{R}_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k) - \mathbf{g}_k^H (\widehat{\mathbf{A}}_k + \lambda I) \mathbf{g}_k + \lambda P \,.$$

The gradient of this concave WSR lower bound is actually still the same as that of the original WSR or of the WSMSE criteria! Allows generalized eigenvector interpretation:

$$\mathbf{H}_k^H \widehat{R}_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k = \frac{1 + \mathbf{g}_k^H \mathbf{H}_k^H \widehat{R}_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k}{u_k} (\widehat{\mathbf{A}}_k + \lambda I) \mathbf{g}_k$$

or hence $\mathbf{g}_k^{'} = V_{max}(\mathbf{H}_k^H \widehat{R}_{\overline{k}}^{-1} \mathbf{H}_k, \widehat{\mathbf{A}}_k + \lambda I)$
which is proportional to the "LMMSE" $\mathbf{g}_k$,
with max eigenvalue $\sigma_k = \sigma_{max}(\mathbf{H}_k^H \widehat{R}_{\overline{k}}^{-1} \mathbf{H}_k, \widehat{\mathbf{A}}_k + \lambda I)$.

[KimGiannakis:IT0511] DC = Optimally Weighted SLNR

- Again, [KimGiannakis:IT0511] BF:

$$\mathbf{g}_k^{'} = V_{max}(\mathbf{H}_k^H \widehat{R}_{\overline{k}}^{-1} \mathbf{H}_k, \sum_{i=1,\neq k}^{K} u_i \mathbf{H}_i^H (\widehat{R}_{\overline{i}}^{-1} - \widehat{R}_i^{-1}) \mathbf{H}_i + \lambda I)$$

- This can be viewed as an optimally weighted version of SLNR (Signal-to-Leakage-plus-Noise-Ratio) [Sayed:SP0507]

$$SLNR_k = \frac{||\mathbf{H}_k \mathbf{g}_k||^2}{\sum_{i\neq k} ||\mathbf{H}_i \mathbf{g}_k||^2 + \sum_i ||\mathbf{g}_i||^2 / P} \text{ vs}$$

$$SINR_k = \frac{||\mathbf{H}_k \mathbf{g}_k||^2}{\sum_{i\neq k} ||\mathbf{H}_k \mathbf{g}_i||^2 + \sum_i ||\mathbf{g}_i||^2 / P}$$

- SLNR takes as Tx filter

$$\mathbf{g}_k^{'} = V_{max}(\mathbf{H}_k^H \mathbf{H}_k, \sum_{i\neq k} \mathbf{H}_i^H \mathbf{H}_i + I)$$

- Let $\sigma_k^{(1)} = \mathbf{g}_k^{'H} \mathbf{H}_k^H \widehat{R}_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k^{'}$ and $\sigma_k^{(2)} = \mathbf{g}_k^{'H} \widehat{\mathbf{A}}_k \mathbf{g}_k^{'}$.
- The advantage of this formulation is that it allows straightforward power adaptation: substituting $\mathbf{g}_k = \sqrt{p_k}\,\mathbf{g}_k^{'}$ yields

$$WSR = \lambda P + \sum_{k=1}^{K} \{u_k \ln(1 + p_k \sigma_k^{(1)}) - p_k(\sigma_k^{(2)} + \lambda)\}$$

which leads to the following interference leakage aware water filling

$$p_k = \left(\frac{u_k}{\sigma_k^{(2)} + \lambda} - \frac{1}{\sigma_k^{(1)}} \right)^{+}.$$

- For a given λ, $\mathbf{g}$ needs to be iterated till convergence.
- And λ can be found by duality (line search):

$$\min_{\lambda \geq 0} \max_{\mathbf{g}} \lambda P + \sum_{k} \{u_k \ln\det(R_{\overline{k}}^{-1} R_k) - \lambda p_k\} = \min_{\lambda \geq 0} WSR(\lambda).$$

High/Low SNR Behavior

- At high SNR, max WSR BF converges to ZF solutions with uniform power

$$\mathbf{g}_k^H = \mathbf{f}_k \mathbf{H}_k P^{\perp}_{(\mathbf{fH})^H_{\overline{k}}} / || \mathbf{f}_k \mathbf{H}_k P^{\perp}_{(\mathbf{fH})^H_{\overline{k}}} ||$$

  where $P^{\perp}_{\mathbf{X}} = I - P_{\mathbf{X}}$ and $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$ projection matrices
  $(\mathbf{fH})_{\overline{k}}$ denotes the (up-down) stacking of $\mathbf{f}_i \mathbf{H}_i$ for users $i = 1, \ldots, K$, $i \neq k$.
- At low SNR, matched filter for user with largest $||\mathbf{H}_k||_2$ (max singular value)

Simulated Annealing

- At high SNR: max WSR solutions are ZF. When ZF is possible (IA feasible), multiple ZF solutions typically exist.
- These different ZF solutions are the possible local optima for max WSR at infinite SNR. By homotopy, this remains the number of max WSR local optima as the SNR decreases from infinity. As the SNR decreases further, a stream for some user may get turned off until only a single stream remains at low SNR. Hence, the number of local optima reduces as streams disappear at finite SNR.
- At intermediate SNR, the number of streams may also be larger than the DoF though.
- Homotopy for finding global optimum: at low SNR, noise dominates interference $\Rightarrow$ optimal: one stream per power constraint, matched filter Tx/Rx. Gradually increasing SNR allows lower SNR solution to be in region of attraction of global optimum at next higher SNR. Phase transitions: add a stream.
- As a corollary, in the MISO case, the max WSR optimum is unique, since there is only one way to perform ZF BF.

Outline

- multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - max (weighted) sum rate (WSR)
  - UL/DL duality, WSR vs WSMSE
  - WSR as optimal SLNR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - problem of EWSMSE
  - Massive MIMO limit
  - large MIMO asymptotics
- distributed designs
  - per cell initialization
  - intercell update

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. The separable (or Kronecker) correlation model (for the channel itself, as opposed to its estimation error or knowledge) below is acceptable when the number of propagation paths N_p becomes large ($N_p \gg MN$) as possibly in indoor propagation.
- Given only mean and covariance information, the fitting maximum entropy distribution is Gaussian.

Mean and Covariance Gaussian CSIT (2)

- Hence consider

$$\text{vec}(\mathbf{H}) \sim \mathcal{CN}(\text{vec}(\overline{\mathbf{H}}), C_t^T \otimes C_r) \text{ or } \mathbf{H} = \overline{\mathbf{H}} + C_r^{1/2} \,\widetilde{\mathbf{H}}\, C_t^{1/2}$$

where $C_r^{1/2}$, $C_t^{1/2}$ are Hermitian square-roots of the Rx and Tx side covariance matrices

$$\begin{aligned} \text{E}\,(\mathbf{H}-\overline{\mathbf{H}})(\mathbf{H}-\overline{\mathbf{H}})^H &= \text{tr}\{C_t\}\; C_r \\ \text{E}\,(\mathbf{H}-\overline{\mathbf{H}})^H(\mathbf{H}-\overline{\mathbf{H}}) &= \text{tr}\{C_r\}\; C_t \end{aligned}$$

and the elements of $\widetilde{\mathbf{H}}$ are i.i.d. $\sim \mathcal{CN}(0,1)$. A scale factor needs to be fixed in the product $\text{tr}\{C_r\}\text{tr}\{C_t\}$ for unicity.

- In what follows, it will also be of interest to consider the total Tx side correlation matrix

$$R_t = \text{E}\;\mathbf{H}^H\mathbf{H} = \overline{\mathbf{H}}^H\overline{\mathbf{H}} + \text{tr}\{C_r\}\; C_t\;.$$

- Gaussian CSIT model could be considered an instance of Ricean fading in which the ratio $\text{tr}\{\overline{\mathbf{H}}^H\overline{\mathbf{H}}\}/(\text{tr}\{C_r\}\text{tr}\{C_t\}) =$ Ricean factor.

Max Expected WSR (EWSR)

- scenario of interest: perfect CSIR, partial (LoS) CSIT
- Imperfect CSIT $\Rightarrow$ various possible optimization criteria: outage capacity,.... Here: expected weighted sum rate $\mathrm{E}_{\mathbf{H}} WSR(\mathbf{g}, \mathbf{H}) =$

$$EWSR(\mathbf{g}) = \mathrm{E}_{\mathbf{H}} \sum_k u_k \ln(1 + \mathbf{g}_k^H \mathbf{H}_k^H R_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k)$$

perfect CSIR: optimal Rx filters $\mathbf{f}_k$ (fn of aggregate $\mathbf{H}$) have been substituted: $WSR(\mathbf{g}, \mathbf{H}) = \max_{\mathbf{f}} \sum_k u_k(-\ln(e_k(\mathbf{f}_k, \mathbf{g})))$.

EWSR Lower Bound: EWSMSE

- $EWSR(\mathbf{g})$: difficult to compute and to maximize directly. [Negro:iswcs12] much more attractive to consider $\mathrm{E}_{\mathbf{H}} e_k(\mathbf{f}_k, \mathbf{g}, \mathbf{H})$ since $e_k(\mathbf{f}_k, \mathbf{g}, \mathbf{H})$ is quadratic in $\mathbf{H}$. Hence optimizing $\mathrm{E}_{\mathbf{H}} WSMSE(\mathbf{g}, \mathbf{f}, w, \mathbf{H})$.

$$\begin{aligned} & \min_{\mathbf{f},w} \mathrm{E}_{\mathbf{H}} WSMSE(\mathbf{g}, \mathbf{f}, w, \mathbf{H}) \\ & \geq \mathrm{E}_{\mathbf{H}} \min_{\mathbf{f},w} WSMSE(\mathbf{g}, \mathbf{f}, w, \mathbf{H}) = -EWSR(\mathbf{g}) \end{aligned}$$

  or hence $EWSR(\mathbf{g}) \geq -\min_{\mathbf{f},w} \mathrm{E}_{\mathbf{H}} WSMSE(\mathbf{g}, \mathbf{f}, w, \mathbf{H})$.

- So now only a lower bound to the EWSR gets maximized, which corresponds in fact to the CSIR being equally partial as the CSIT.

$$\begin{aligned} & \mathrm{E}_{\mathbf{H}} e_k = 1 - 2\Re\{\mathbf{f}_k^H \overline{\mathbf{H}}_k \mathbf{g}_k\} + \sum_{i=1}^K \mathbf{f}_k^H \overline{\mathbf{H}}_k \mathbf{g}_i \mathbf{g}_i^H \overline{\mathbf{H}}_k^H \mathbf{f}_k \\ & + \mathbf{f}_k^H R_{r,k} \mathbf{f}_k \sum_{i=1}^K \mathbf{g}_i^H R_{t,k} \mathbf{g}_i + ||\mathbf{f}_k||^2. \end{aligned}$$

  $\Rightarrow$ signal term disappears if $\overline{\mathbf{H}}_k = 0$! Hence the EWSMSE lower bound is (very) loose unless the Rice factor is high, and is useless in the absence of mean CSIT.

Massive MIMO Limit

- We get a convergence for any term of the form

$$\mathbf{H}Q\mathbf{H}^H \overset{M\to\infty}{\longrightarrow} \mathrm{E}_{\,\mathbf{H}}\,\mathbf{H}Q\mathbf{H}^H = \overline{\mathbf{H}}Q\overline{\mathbf{H}}^H + \mathrm{tr}\{QC_t\}\,C_r\,.$$

Go one step further in separable channel correlation model: $C_{r,k,b_i} = C_{r,k}, \forall b_i$. This leads us to introduce

$$\mathbf{H}_k = [\mathbf{H}_{k,1}\cdots\mathbf{H}_{k,C}] = \overline{\mathbf{H}}_k + C_{r,k}^{1/2}\widetilde{\mathbf{H}}_k C_{t,k}^{1/2}$$

$$Q = \begin{bmatrix} \sum_{i:b_i=1} Q_i & & \\ & \ddots & \\ & & \sum_{i:b_i=C} Q_i \end{bmatrix} = \sum_{j=1}^{C}\sum_{i:b_i=j} I_j Q_i I_j^H$$

$$Q_{\overline{k}} = Q - I_{b_i} Q_i I_{b_i}^H$$

where $C_{t,k} = \mathrm{blockdiag}\{C_{t,k,1},\ldots,C_{t,k,C}\}$, and I_j is an all zero block vector except for an identity matrix in block j. Then we get for the *WSR* (= *EWSR*),

$$WSR = \sum_{k=1}^{K} u_k \ln\det(\breve{R}_{\overline{k}}^{-1}\breve{R}_k)$$

where

$$\breve{R}_k = I_{N_k} + \overline{\mathbf{H}}_k Q\overline{\mathbf{H}}_k^H + \mathrm{tr}\{QC_{t,k}\}\,C_{r,k}$$
$$\breve{R}_{\overline{k}} = I_{N_k} + \overline{\mathbf{H}}_k Q_{\overline{k}}\overline{\mathbf{H}}_k^H + \mathrm{tr}\{Q_{\overline{k}}C_{t,k}\}\,C_{r,k}$$

Massive MIMO Limit (2)

- This leads to

$$WSR = u_k \ln\det(I+\breve{R}_{\overline{k}}^{-1}\left(\overline{\mathbf{H}}_{k,b_k}\mathbf{g}_k\mathbf{g}_k^H\overline{\mathbf{H}}_{k,b_k}^H + \mathrm{tr}\{\mathbf{g}_k\mathbf{g}_k^H C_{t,k,b_k}\} C_{r,k}\right)) + WSR_{\overline{k}} \,.$$

- Consider simplified case: "Ricean factor" $\mu \sim$ SNR, for the direct links $\mathbf{H}_{k,b_k}$ (only) (properly organized (intracell) channel estimation and feedback) $\Rightarrow$ approximation

$$WSR = u_k \ln\det(I + \mathbf{g}_k^H \breve{\mathbf{B}}_k \mathbf{g}_k) + WSR_{\overline{k}} \quad \text{with}$$

$$\breve{\mathbf{B}}_k = \overline{\mathbf{H}}_{k,b_k}^H \breve{R}_{\overline{k}}^{-1} \overline{\mathbf{H}}_{k,b_k} + \mathrm{tr}\{C_{r,k}\breve{R}_{\overline{k}}^{-1}\} C_{t,k,b_k}$$

The linearization of $WSR_{\overline{k}}$ w.r.t. Q_k now involves

$$\breve{\mathbf{A}}_k = \sum_{i \neq k}^{K} u_i \left[\overline{\mathbf{H}}_{i,b_k}^H \left(\breve{R}_{\overline{i}}^{-1} - \breve{R}_i^{-1}\right)\overline{\mathbf{H}}_{i,b_k} + \mathrm{tr}\{\left(\breve{R}_{\overline{i}}^{-1} - \breve{R}_i^{-1}\right) C_{r,i}\} C_{t,i,b_k}\right] \,.$$

The rest of the development is now completely analogous to the case of perfect CSIT.

WSMSE-DC BF relation

- min WSMSE iteration $(i+1)$

$$
\begin{aligned}
\mathbf{A}_k^{(i)} &= \sum_j u_j w_j^{(i)} \mathbf{H}_i^H \mathbf{f}_i^{(i)} \mathbf{f}_j^{(i)H} \mathbf{H}_j + \lambda^{(i)} I_M \\
\mathbf{g}_k^{(i+1)} &= (\mathbf{A}_k^{(i)})^{-1} \mathbf{H}_k^H \mathbf{f}_k^{(i)} u_k w_k^{(i)} \\
&= (\mathbf{A}_k^{(i)})^{-1} \mathbf{B}_k^{(i)}\, \mathbf{g}_k^{(i)}\; u_k w_k^{(i)} \\
\mathbf{B}_k^{(i)} &= \mathbf{H}_k^H R_k^{-(i)} \mathbf{H}_k
\end{aligned}
$$

WSMSE does one power iteration of DC !!

$$
\mathbf{g}_k^{(i+1)} = V_{max}\{(\mathbf{A}_k^{(i)})^{-1} \mathbf{B}_k^{(i)}\}
$$

- partial CSIT (or MaMIMO) case: modified WSMSE:

$$
\mathbf{g}_k^{(i+1)} = (\mathrm{E}_{\,H} \mathbf{A}_k^{(i)})^{-1} (\mathrm{E}_{\,H} \mathbf{B}_k^{(i)})\, \mathbf{g}_k^{(i)}\; u_k w_k^{(i)}
$$

Outline

- multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - max (weighted) sum rate (WSR)
  - UL/DL duality, WSR vs WSMSE
  - WSR as optimal SLNR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - problem of EWSMSE
  - Massive MIMO limit
  - large MIMO asymptotics
- distributed designs
  - per cell initialization
  - intercell update

Reduced CSIT and Decoupled Tx/Rx Design

- for IA to apply to cellular: overall Tx/Rx design has to decompose so that the CSIT required is no longer global and remains bounded regardless of the network size.
- simplest case : local CSIT : a BS only needs to know the channels from itself to all terminals. In the TDD case : reciprocity. The local CSIT case arises when all ZF work needs to be done by the Tx: $d_{c,k} = N_{c,k}, \forall c, k$. The most straightforward such case is of course the MISO case: $d_{c,k} = N_{c,k} = 1$. It extends to cases of $N_{c,k} > d_{c,k}$ if less than optimal DoF are accepted. One of these cases is that of reduced rank MIMO channels.
- reduced CSIT [Lau:SP0913]: variety of approaches w reduced CSIT FB in exchange for DoF reductions.
- incomplete CSIT [deKerretGesbert:TWC13]: min some MIMO IC optimal DoF can be attained with less than global CSIT. Only occurs when M and/or N vary substantially so that subnetworks of a subgroup of BS and another subgroup of terminals arise in which the numbers of antennas available are just enough to handle the interference within the subnetwork.
- Massive MIMO leads to exploiting covariance CSIT, which will tend to have reduced rank and allows decoupled approaches.

Distributed IBC Design

- Global intracluster CSIT can also be gathered but it takes an overhead that evolves with C^2 [Negro:isccsp12]. Hence, high quality (high Ricean factor) intracell CSIT and Tx covariance only intercell CSIT may be a more appropriate setting. For what follows we shall assume the LoS Tx intercell CSIT. We shall focus on a MaMIMO setting.
- The approach considered here is non-iterative, or could be taken as initialization for further iterations.

Distributed IBC Design: Initialization

- Start with a per cell design.
- To simplify design, assume Rx antennas are used to handle intracell interference. Hence all intercell interference needs to be handled by Tx (BS) antennas.
- In that case, the crosslinks (cascades of channel and Rx) can be considered as independent from the intracell channels.
- In a MaMIMO setting, the ZF by BS j towards $K-K_j$ crosslink channels (or LoS components in fact) will tend to have a deterministic effect of reducing the effective number of Tx antennas by this amount and hence of reducing the Tx power by a factor $\frac{M_j}{M_j-(K-K_j)}$. Hence a per BS design can be carried out with (partial) intracell CSIT, with BS Tx power P_j replaced by $\frac{M_j}{M_j-(K-K_j)}P_j$, and with all intercell links $\mathbf{H}_{k,b_i}=0$, $b_i \neq b_k$.
- This first step (which is itself an iterative design for the scenario considered with reduced Tx power and no intercell links) leads to BFs $\mathbf{g}^{(0)}$ which lead to

$$\breve{R}_k = (1+\mathrm{tr}\{Q^{(0)}_{b_k} C_{t,k,b_k}\}) I_{N_k} + \overline{\mathbf{H}}_{k,b_k} Q^{(0)}_{b_k} \overline{\mathbf{H}}^H_{k,b_k}\,,$$

$$\text{where} \quad Q^{(0)}_{b_k} = \sum_{i:b_i=b_k} \mathbf{g}^{(0)}_i \mathbf{g}^{(0)H}_i$$

and similarly for $\breve{R}_{\overline{k}}$.

Distributed IBC Design: Iteration 1

- Do one iteration in order to adjust the Tx filters for the intercell interference.
- With the initial BFs $\mathbf{g}^{(0)}$, the local intercell CSIT C_{t,i,b_k} also, the correct power constraints, and $\breve{R}_k$, $\breve{R}_{\overline{k}}$ as above, we get $\breve{\mathbf{B}}_k$ as before, and $\breve{\mathbf{A}}_k$ becomes

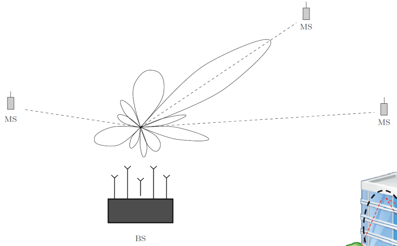
$$\breve{\mathbf{A}}_k = \sum_{i \neq k : b_i = b_k} u_i \left[\overline{\mathbf{H}}_{i,b_k}^H \left(\breve{R}_{\overline{i}}^{-1} - \breve{R}_i^{-1} \right) \overline{\mathbf{H}}_{i,b_k} + \mathrm{tr}\{ \left(\breve{R}_{\overline{i}}^{-1} - \breve{R}_i^{-1} \right) C_{r,i} \} C_{t,i,b_k} \right] + \sum_{i : b_i \neq b_k} \underbrace{u_i \mathrm{tr}\{ \breve{R}_{\overline{i}}^{-1} - \breve{R}_i^{-1} \}}_{= \mu_i} C_{t,i,b_k} \, .$$

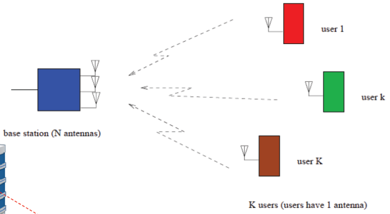
Hence the only information that needs to be fed back from user i in another cell is the positive scalar μ_i. This is related to the interference pricing in game theory [XuWang:JSAC1012].

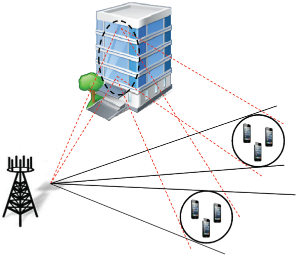
- The normalized BFs are then computed as $\mathbf{g}_k^{'} = V_{max}(\breve{\mathbf{B}}_k, \breve{\mathbf{A}}_k + \lambda_{b_k} I)$ where the λ_{b_k} are taken from the previous iteration.
- The stream powers are obtained from the interference-aware WF.

Massive MIMO: from spatial to spatiotemporal and back

- spatial: to null a user, need to null all paths of that user
  spatiotemporal: # antennas > # users
  spatial: # antennas > # paths ≫ # users
- but: paths are slowly fading, user channels are fast fading







- Keysight ieee comsoc M-MIMO tutorial, mmWave

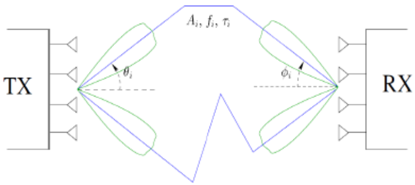


Figure 1 : MIMO transmission with M transmit and N receive antennas.

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 ψ_i (due to Doppler) and possibly the variation of the A_i correspond to the fast fading. All the other parameters vary on a slower time scale and correspond to slow fading.

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
- θ_i: direction of departure (AoD)
- ϕ_i: direction of arrival (AoA)
- $\mathbf{h}_t(.)$, $\mathbf{h}_r(.)$: $M/N \times 1$ Tx/Rx antenna array response

and

$$\mathbf{B} = [\mathbf{h}_r(\phi_1) \, \mathbf{h}_r(\phi_1) \cdots] \begin{bmatrix} e^{j\psi_1} & & \\ & e^{j\psi_2} & \\ & & \ddots \end{bmatrix}$$

$$\mathbf{A}^H = \begin{bmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{h}_t^T(\theta_1) \\ \mathbf{h}_t^T(\theta_2) \\ \vdots \end{bmatrix} \quad (2)$$

Mean and Covariance Gaussian CSIT
Dominant Paths Partial CSIT Channel Model

Given only mean and (separable) covariance information, the fitting maximum entropy distribution is Gaussian. Hence consider $\text{vec}(\mathbf{H}) \sim \mathcal{CN}(\text{vec}(\overline{\mathbf{H}}), \mathbf{C}_t^T \otimes \mathbf{C}_r)$ which can be rewritten as

$$\mathbf{H} = \overline{\mathbf{H}} + \mathbf{C}_r^{1/2}\,\widetilde{\mathbf{H}}\,\mathbf{C}_t^{1/2} \tag{3}$$

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

$$\begin{aligned} \mathrm{E}(\mathbf{H}-\overline{\mathbf{H}})(\mathbf{H}-\overline{\mathbf{H}})^H &= \mathrm{tr}\{\mathbf{C}_t\}\ \mathbf{C}_r \\ \mathrm{E}(\mathbf{H}-\overline{\mathbf{H}})^H(\mathbf{H}-\overline{\mathbf{H}}) &= \mathrm{tr}\{\mathbf{C}_r\}\ \mathbf{C}_t \end{aligned} \tag{4}$$

and the elements of $\widetilde{\mathbf{H}}$ 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

$$\mathbf{R}_t = \mathrm{E}\,\mathbf{H}^H\mathbf{H} = \overline{\mathbf{H}}^H\overline{\mathbf{H}} + \mathrm{tr}\{\mathbf{C}_r\}\,\mathbf{C}_t\ . \tag{5}$$

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

This Taylor series modeling of clusters is in contrast to the uniform DoA profile used in [Caire:mmWave], [Gesbert:arxiv1013].

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

$$\mathbf{H} = \mathbf{B}\,\mathbf{A}^H(\theta) + \sqrt{\beta}\,\widetilde{\mathbf{H}}' \tag{6}$$

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

$$\mathbf{C}_t = \mathbf{A}\mathbf{A}^H + N\beta\,I_M \tag{7}$$

since due to the normalization of the antenna array responses, $\mathrm{E}\,\mathbf{B}^H\mathbf{B} = \mathbf{I}$. Note that $\mu = \mathrm{tr}\{\mathbf{A}\mathbf{A}^H\}/\beta NM$ 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 θ is the nominal AoD and $\Delta\theta_i$ is small $\Rightarrow$ $\mathbf{h}_t(\theta_i) \approx \mathbf{h}_t(\theta) + \Delta\theta_i\,\dot{\mathbf{h}}_t(\theta)$: rank 2 contribution to $\mathbf{A}\mathbf{A}^H$.

Specular Wireless MIMO Channel Model

We get for the matrix impulse response of a time-varying MIMO channel $\mathbf{H}(t,\tau)$

$$\mathbf{H}(t,\tau) = \sum_{i=1}^{N_p} A_i(t)\, e^{j2\pi\, f_i\, t}\, \mathbf{h}_r(\phi_i)\, \mathbf{h}_t^T(\theta_i)\, p(\tau-\tau_i) \ \ .$$

The channel impulse response $\mathbf{H}$ has per path a rank 1 contribution in 4 dimensions (Tx and Rx spatial multi-antenna dimensions, delay spread and Doppler spread); there are N_p (specular) pathwise contributions where

- A_i: complex attenuation
- f_i: Doppler shift
- θ_i: direction of departure (AoD)
- ϕ_i: direction of arrival (AoA)
- τ_i: path delay (ToA)
- $\mathbf{h}_t(.)$, $\mathbf{h}_r(.)$: $M/N \times 1$ Tx/Rx antenna array response
- $p(.)$: pulse shape (Tx filter)

Specular Wireless MIMO Channel Model (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 phase in $e^{j2\pi\, f_i\, t}$ and possibly the variation of the A_i correspond to the fast fading. All the other parameters (including the Doppler frequency) vary on a slower time scale and correspond to slow fading.
- OFDM transmission

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

(not the same $A_i \geq 0$, path amplitude)

IA Feasibility Reduced Rank MIMO IBC

- The ZF from BS j to MT (i, k) requires

$$F_{i,k}^H \mathbf{H}_{i,k,j} G_{j,n} = F_{i,k}^H \mathbf{B}_{i,k,j} \, \mathbf{A}_{i,k,j}^H G_{j,n} = 0$$

which involves $\min(d_{i,k}d_{j,n}, d_{i,k}r_{i,k,j}, r_{i,k,}d_{j,n})$ constraints to be satisfied by the $(N_{i,k} - d_{i,k})d_{i,k}/(M_j - d_{j,n})d_{j,n}$ variables parameterizing the column subspaces of $F_{i,k}/G_{j,n}$.

- **IA feasibility singular MIMO IC with Tx/Rx decoupling**

$$F_{i,k}^H \mathbf{B}_{i,k,j} = 0 \text{ or } \mathbf{A}_{i,k,j}^H G_{j,n} = 0 \; .$$

This leads to a possibly increased number of ZF constraints $r_{i,k,j} \min(d_{i,k}, d_{j,n})$ and hence to possibly reduced IA feasibility. ZF of every cross link now needs to be partitioned between all Txs and Rxs, taking into account the limited number of variables each Tx or Rx has. The main goal of this approach however is that it leads to Tx/Rx decoupling and local CSI.

Covariance CSIT

Averaging over the (uniform) path phases ψ_i leads to

$$C_{\mathbf{hh}} = \sum_{i=1}^{N_p} A_i^2 \, \mathbf{h}_i \mathbf{h}_i^H = \sum_{i=1}^{N_p} A_i^2 \, (\mathbf{h}_r(\phi_i)\mathbf{h}_r^H(\phi_i)) \otimes (\mathbf{h}_t(\theta_i)\mathbf{h}_t^H(\theta_i))$$

where $C_{\mathbf{hh}} = \mathrm{E}\ \mathbf{hh}^H$, $\mathbf{h} = \mathrm{vec}(\mathbf{H})$ and $\mathbf{h}_i = \mathbf{h}_t(\theta_i) \otimes \mathbf{h}_r(\phi_i)$.
Note that the rank of $C_{\mathbf{hh}}$ 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 θ is the nominal AoD and $\Delta\theta_i$ is small. Hence

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

Such a cluster of paths only adds a rank 2 contribution to $C_{\mathbf{hh}}$.
Not of Kronecker form.

Tx side Covariance CSIT

Tx side covariance matrix C^t, which only explores the channel correlations as they can be seen from the BS side

$$C^t = \mathrm{E}\ \mathbf{H}^H\mathbf{H}$$

We can factor the channel response as

$$\mathbf{H} = \mathbf{B}\ \mathbf{A}^H\ , \quad \mathbf{B} = [\mathbf{h}_r(\phi_1)\,\mathbf{h}_r(\phi_1)\cdots] \begin{bmatrix} e^{j\psi_1} & & \\ & e^{j\psi_2} & \\ & & \ddots \end{bmatrix}\ ,$$

$$\mathbf{A}^H = \begin{bmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{h}_t^T(\theta_1) \\ \mathbf{h}_t^T(\theta_2) \\ \vdots \end{bmatrix}$$

Averaging of the path phases ψ_i, we get for the Tx side covariance matrix

$$C^t = \mathbf{A}\mathbf{A}^H$$

since due to the normalization of the antenna array responses, $\mathrm{E}\ \mathbf{B}^H\mathbf{B} = \mathrm{diag}\{[\mathbf{h}_r(\phi_1)\,\mathbf{h}_r(\phi_1)\cdots]^H\,[\mathbf{h}_r(\phi_1)\,\mathbf{h}_r(\phi_1)\cdots]\} = I$.

IA Feasibility Reduced Rank MIMO IBC

- The ZF from BS j to MT (i,k) requires

$$F_{i,k}^H \mathbf{H}_{i,k,j} G_{j,n} = F_{i,k}^H \mathbf{B}_{i,k,j} \, \mathbf{A}_{i,k,j}^H G_{j,n} = 0$$

which involves $\min(d_{i,k}d_{j,n}, d_{i,k}r_{i,k,j}, r_{i,k,}d_{j,n})$ constraints to be satisfied by the $(N_{i,k}-d_{i,k})d_{i,k}/(M_j-d_{j,n})d_{j,n}$ variables parameterizing the column subspaces of $F_{i,k}/G_{j,n}$.

- **IA feasibility singular MIMO IC with Tx/Rx decoupling**

$$F_{i,k}^H \mathbf{B}_{i,k,j} = 0 \text{ or } \mathbf{A}_{i,k,j}^H G_{j,n} = 0 \ .$$

This leads to a possibly increased number of ZF constraints $r_{i,k,j} \min(d_{i,k}, d_{j,n})$ and hence to possibly reduced IA feasibility. ZF of every cross link now needs to be partitioned between all Txs and Rxs, taking into account the limited number of variables each Tx or Rx has. The main goal of this approach however is that it leads to Tx/Rx decoupling.

Massive MIMO & Covariance CSIT

In massive MIMO, the Tx side channel covariance matrix is very likely to be (very) singular even though the channel response $\mathbf{H}$ may not be singular:

$$\mathrm{rank}(C^t_{i,k,j} = \mathbf{A}_{i,k,j}\mathbf{A}^H_{i,k,j}) = r_{i,k,j}\,,\ \mathbf{A}_{i,k,j} : M_j \times r_{i,k,j}$$

Let $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^H\mathbf{X})^{\#}\mathbf{X}^H$ and $P^{\perp}_{\mathbf{X}}$ be the projection matrices on the column space of $\mathbf{X}$ and its orthogonal complement resp. Consider now a massive MIMO IBC with C cells containing K_i users each to be served by a single stream. The following result states when this will be possible.

Theorem

Sufficiency of Covariance CSIT for Massive MIMO IBC *In the MIMO IBC with (local) covariance CSIT, all BS will be able to perform ZF BF if the following holds*

$$||P^{\perp}_{\mathbf{A}_{\overline{i,k},j}}\mathbf{A}_{i,k,j}|| > 0\ ,\ \forall i,k,j$$

where $\mathbf{A}_{\overline{i,k},j} = \{\mathbf{A}_{n,m,j},\, (n,m) \neq (i,k)\}$.

Massive MIMO & Covariance CSIT (2)

These conditions will be satisfied w.p. 1 if $\sum_{i=1}^{C}\sum_{k=1}^{K_i} r_{i,k,j} \leq M_j\,,\; j=1,\ldots,C$. In that case all the column spaces of the $\mathbf{A}_{i,k,j}$ will tend to be non-overlapping. However, the conditions could very well be satisfied even if these column spaces are overlapping, in contrast to what [Gesbert:arxiv1013],[Caire:arxiv0912] appear to require. In Theorem 1, we assume that all ZF work is done by the BS. However, if the MT have multiple antennas, they can help to a certain extent.

Theorem

Role of Receive Antennas in Massive MIMO IBC *If MT (i,k) disposes of $N_{i,k}$ antennas to receive a stream, it can perform rank reduction of a total amount of $N_{i,k}-1$ to be distributed over $\{r_{i,k,j}\,,\; j=1,\ldots,C\}$.*

Such rank reduction (by ZF of certain path contributions) facilitates the satisfaction of the conditions in Theorem 1.

FIR IA for Asynchronous FIR Frequency-Selective IBC

FIR frequency-selective channels : OFDM : assumes that the same OFDM is used by synchronized BS. In HetNets, this may not be the case. Then FIR Tx/Rx filters may be considered. We get in the z-domain:

$$F_{i,k}(z)\mathbf{H}_{i,k,j}(z)G_{j,n}(z) = 0 \;,\; (i,k) \neq (j,n) \;,$$

If we denote by L_F, L_H, L_G the length of the 3 types of filters, then in a symmetric configuration, the proper conditions become

$$\begin{aligned} & KC\left[d(ML_G - d) + d(NL_F - d)\right] \geq \\ & \qquad KC(KC-1)d^2(L_H + L_G + L_F - 2) \\ & \Rightarrow d \leq \frac{ML_G + NL_F}{(KC-1)(L_H + L_G + L_F - 2) + 2} \leq \frac{\max\{M,N\}}{KC-1} \end{aligned}$$

where the last inequality can be attained by letting L_G or L_F tend to infinity. Unless $M \gg N$, this represents reduced DoF compared to the frequency-flat case ($d \leq (M+N)/(KC+1)$).

Alternatively, the double convolution by both Tx and Rx filters can be avoided by considering most of the decoupled approaches above, leading to more traditional equalization configurations, with equal DoF possibilities for frequency-selective as for frequency-flat cases.

Comparison IA Feasbility Pathwise vs FIR

-

-

BF as Dual UL LMMSE

- DL Rx signal at user k in cell b_k

$$y_k = \underbrace{\mathbf{H}_{k,b_k}\,\mathbf{g}_k\,x_k}_{\text{signal}} + \underbrace{\sum_{\substack{i\neq k\\ b_i=b_k}} \mathbf{H}_{k,b_k}\,\mathbf{g}_i\,x_i}_{\text{intracell interf.}} + \underbrace{\sum_{j\neq b_k}\sum_{i:b_i=j} \mathbf{H}_{k,j}\,\mathbf{g}_i\,x_i}_{\text{intercell interf.}} + v_k$$

at output of Rx:

$$\widehat{x}_k = \mathbf{f}_k^H \mathbf{H}_{k,b_k}\,\mathbf{g}_k\,x_k + \sum_{i=1,\neq k}^{K} \mathbf{f}_k^H \mathbf{H}_{k,b_i}\,\mathbf{g}_i\,x_i + \mathbf{f}_k^H v_k$$

- Dual UL at BS k $\qquad \sigma^2_{\widetilde{x}_i} = u_i\,w_i,\; R_{\widetilde{v}_k\widetilde{v}_k} = \lambda_k\,I_{M_k}$

$$\widetilde{y}_k = \underbrace{\sum_{i:b_i=k} \mathbf{H}_{i,k}^H\,\mathbf{f}_i\,\widetilde{x}_i}_{\text{intracell users}} + \underbrace{\sum_{i:b_i\neq k} \mathbf{H}_{i,k}^H\,\mathbf{f}_i\,\widetilde{x}_i}_{\text{intercell users}} + \widetilde{v}_k$$

- pathwise dual UL at BS k

$$\widetilde{y}_k = \sum_{i:b_i=k} \mathbf{A}_{i,k} \underbrace{\mathbf{B}_{i,k}^H \mathbf{f}_i\,\widetilde{x}_i}_{\mathbf{s}_i \text{ intracell paths}} + \sum_{i:b_i\neq k} \mathbf{A}_{i,k} \underbrace{\mathbf{B}_{i,k}^H \mathbf{f}_i\,\widetilde{x}_i}_{\mathbf{s}_i \text{ intercell paths}} + \widetilde{v}_k$$

BF as Dual UL LMMSE (2)

- Dual UL at BS k

$$\widetilde{y}_k = \underbrace{\mathbf{H}_k^H F_k \widetilde{x}_k}_{\text{intracell users}} + \underbrace{\mathbf{H}_{\overline{k}}^H F_{\overline{k}} \widetilde{x}_{\overline{k}}}_{\text{intercell users}} + \widetilde{v}_k$$

- DL BF as UL LMMSE Rx: (for all intracell users)

$$G_k^H = R_{\widetilde{x}_k \widetilde{y}_k} R_{\widetilde{y}_k \widetilde{y}_k}^{-1} = (\mathrm{E}_{\,\widetilde{x},\widetilde{v}} \widetilde{x}_k \widetilde{y}_k)(\mathrm{E}_{\,\widetilde{x},\widetilde{v}} \widetilde{y}_k \widetilde{y}_k)^{-1}$$

- pathwise dual UL at BS k $\quad \widetilde{\mathbf{s}}_k = \mathbf{B}_k^H F_k \widetilde{x}_k, \; \widetilde{\mathbf{s}}_{\overline{k}} = \mathbf{B}_{\overline{k}}^H F_{\overline{k}} \widetilde{x}_{\overline{k}}$

$$\widetilde{y}_k = \mathbf{A}_k \underbrace{\widetilde{\mathbf{s}}_k}_{\text{intracell paths}} + \mathbf{A}_{\overline{k}} \underbrace{\widetilde{\mathbf{s}}_{\overline{k}}}_{\text{intercell paths}} + \widetilde{v}_k$$

- Pathwise DL BF as UL LMMSE Rx: (for all intracell paths)

$$\begin{aligned} \widetilde{G}_k^H &= R_{\widetilde{\mathbf{s}}_k \widetilde{y}_k} R_{\widetilde{y}_k \widetilde{y}_k}^{-1} = (\mathrm{E}_{\,\widetilde{x},\widetilde{v},\mathbf{B}}\, \widetilde{\mathbf{s}}_k \widetilde{y}_k)(\mathrm{E}_{\,\widetilde{x},\widetilde{v},\mathbf{B}}\, \widetilde{y}_k \widetilde{y}_k)^{-1} \\ &= R_{\widetilde{\mathbf{s}}_k \widetilde{\mathbf{s}}_k} \mathbf{A}_k^H (\mathbf{A}_k^H R_{\widetilde{\mathbf{s}}_k \widetilde{\mathbf{s}}_k} \mathbf{A}_k + \mathbf{A}_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}} \widetilde{\mathbf{s}}_{\overline{k}}} \mathbf{A}_{\overline{k}}^H + \lambda_k I)^{-1} \end{aligned}$$

2-stage BF: Pathwise + Userwise Intracell

- $$\begin{aligned}\widetilde{G}_k^H &= R_{\widetilde{\mathbf{s}}_k\widetilde{\mathbf{s}}_k}\mathbf{A}_k^H(\mathbf{A}_k^H R_{\widetilde{\mathbf{s}}_k\widetilde{\mathbf{s}}_k}\mathbf{A}_k + \mathbf{A}_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}}\widetilde{\mathbf{s}}_{\overline{k}}}\mathbf{A}_{\overline{k}}^H + \lambda_k I)^{-1} \\ &= \underbrace{(\mathbf{A}_k^H R^{-1}\mathbf{A}_k + I)^{-1}}_{\text{stage 2}}\underbrace{\mathbf{A}_k^H R^{-1}}_{\text{stage 1}}\,, \quad R = \mathbf{A}_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}}\widetilde{\mathbf{s}}_{\overline{k}}}\mathbf{A}_{\overline{k}}^H + \lambda_k I\end{aligned}$$
- stage 1: intercell path suppression

$$\mathbf{A}_k^H(\mathbf{A}_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}}\widetilde{\mathbf{s}}_{\overline{k}}}\mathbf{A}_{\overline{k}}^H + \lambda_k I)^{-1}$$

  allows pilot transmission without intercell path interference, but with intracell interference
- stage 1': intracell and intercell path suppression

$$\mathbf{A}_k^H(\mathbf{A}_k^H R_{\widetilde{\mathbf{s}}_k\widetilde{\mathbf{s}}_k}\mathbf{A}_k + \mathbf{A}_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}}\widetilde{\mathbf{s}}_{\overline{k}}}\mathbf{A}_{\overline{k}}^H + \lambda_k I)^{-1}$$

  allows pilot transmission on one user's paths without any interference from paths of any other user
  $\Rightarrow$ training lenth = max # paths of a user

2-stage BF (2)

- stage 1' mimimizes weighted sum MSE at path outputs

$$\widetilde{u}_i \widetilde{w}_i \widetilde{f}_i \widetilde{f}_i^H = (R_{\widetilde{\mathbf{s}}\widetilde{\mathbf{s}}})_{i,i}, \widetilde{\mathbf{g}}_i \Rightarrow \widetilde{f}_i, \widetilde{w}_i = 1/\widetilde{e}_i = 1/(1 - \widetilde{f}_i \mathbf{A}_i \widetilde{\mathbf{g}}_i)$$

not clear if this weighting is optimal for channel estimation

- stage 2: from paths to user signals (intracell)

$$R_{\widetilde{x}_k \widetilde{y}_k} = \underbrace{R_{\widetilde{x}_k \widetilde{\mathbf{s}}_k} R^{-1}_{\widetilde{\mathbf{s}}_k \widetilde{\mathbf{s}}_k}}_{\text{LMMSE: paths} \rightarrow \text{users}} R_{\widetilde{\mathbf{s}}_k \widetilde{y}_k}$$

$$R_{\widetilde{x}_k \widetilde{\mathbf{s}}_k} = R_{\widetilde{x}_k \widetilde{x}_k} F_k^H \mathbf{B}_k$$