Sum Utility Optimization in MIMO Multi-User Multi-Cell: Centralized and Distributed, Perfect and Partial CSIT, Fast and Slow CSIT

Dirk Slock

Mobile Communication Department, EURECOM

CWC, Oulu, 18/09/2015

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

Utility Functions

- single user (MIMO) in Gaussian noise: Gaussian signaling optimal (avg. power constr.)
- rate stream k : $R_k = \ln(1 + \text{SINR}_k)$
- SINR balancing: $\max_{BF} \min_k \text{SINR}_k/\gamma_k$ under Tx power P, fairness
- related: min Tx power under $\text{SINR}_k \geq \gamma_k$ GREEN
- max Weighted Sum Rate (WSR): $\max_{BF} \sum_k u_k R_k$, given P

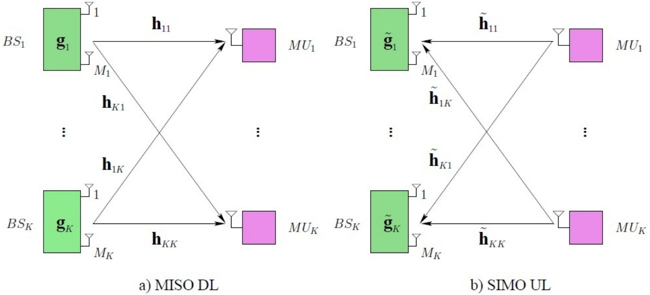
  weights u_k may reflect state of queues (to minimize queue overflow)

  weights also allow to vary orientation of normal to Pareto boundary of rate region and hence to explore whole Pareto boundary if rate region convex
  Pareto boundary: cannot increase an R_k without decreasing some R_i.

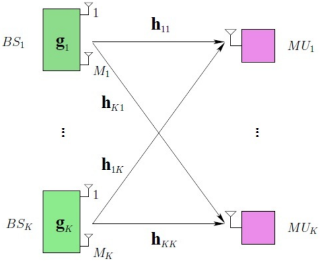
MISO Interference Channel

- K pairs of multiantenna Base Station (BS) and single antenna Mobile User (MU)
- BS number k is equipped with M_k antennas
- $\mathbf{g}_k$ ($\tilde{\mathbf{g}}_k$) is the beamformer (RX filter) applied at the k-th BS in DL (UL) transmission
- y_k is Rx signal at the k-th MU in the DL phase,
  $\tilde{r}_k$ is output of Rx filter at the k-th BS in the UL phase:

$$y_k = \mathbf{h}_{kk}\mathbf{g}_k s_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{h}_{kl}\mathbf{g}_l s_l + n_k \qquad \tilde{r}_k = \tilde{\mathbf{g}}_k\tilde{\mathbf{h}}_{kk}\tilde{s}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \tilde{\mathbf{g}}_k\tilde{\mathbf{h}}_{kl}\tilde{s}_l + \tilde{\mathbf{g}}_k\tilde{n}_k$$



- MISO DL IFC



- The SINR for the DL channel is:

$$SINR_k^{DL} = \frac{p_k \mathbf{g}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk} \mathbf{g}_k}{\sum_{l \neq k} p_l \mathbf{g}_l^H \mathbf{h}_{kl}^H \mathbf{h}_{kl} \mathbf{g}_l + \sigma^2}$$

p_k is the TX power at the k-th BS.

- Imposing a set of DL SINR constraints at each mobile station: $SINR_k^{DL} = \gamma_k$ we obtain in matrix notation:

$$\mathbf{\Phi p} + \boldsymbol{\sigma} = \mathbf{D}^{-1}\mathbf{p}$$

with:

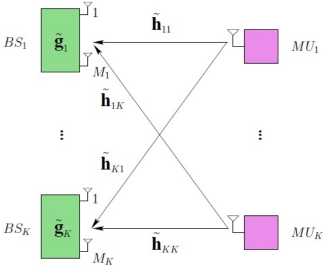
$$[\mathbf{\Phi}]_{ij} = \begin{cases} \mathbf{g}_j^H \mathbf{h}_{ij}^H \mathbf{h}_{ij} \mathbf{g}_j, & j \neq i \\ 0, & j = i \end{cases}$$

$$\mathbf{D} = \mathrm{diag}\{\frac{\gamma_1}{\mathbf{g}_1^H \mathbf{h}_{11}^H \mathbf{h}_{11} \mathbf{g}_1}, \ldots, \frac{\gamma_K}{\mathbf{g}_K^H \mathbf{h}_{KK}^H \mathbf{h}_{KK} \mathbf{g}_K}\}.$$

- We can determine the TX power solving w.r.t. $\mathbf{p}$ obtaining:

$$\mathbf{p} = (\mathbf{D}^{-1} - \mathbf{\Phi})^{-1}\boldsymbol{\sigma} \qquad (1)$$

- SIMO UL IFC



- Assuming that $\tilde{\mathbf{h}}_{ij} = \mathbf{h}_{ji}^H$ and $\tilde{\mathbf{g}}_i = \mathbf{g}_i^H$ the SINR for the UL channel can be written as:

$$SINR_k^{UL} = \frac{q_k \mathbf{g}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk} \mathbf{g}_k}{\mathbf{g}_k^H (\sum_{l \neq k} q_l \mathbf{h}_{lk}^H \mathbf{h}_{lk} + \sigma^2 \mathbf{I}) \mathbf{g}_k}$$

q_k represents the Tx power from the k-th MS.

- Imposing the same SINR constraints also in the UL: $SINR_k^{UL} = \gamma_k$ it is possible to rewrite that constraints as:

$$\tilde{\boldsymbol{\Phi}}\mathbf{q} + \boldsymbol{\sigma} = \mathbf{D}^{-1}\mathbf{q}$$

with:

$$[\tilde{\boldsymbol{\Phi}}]_{ij} = \begin{cases} \mathbf{g}_i^H \mathbf{h}_{ji}^H \mathbf{h}_{ji} \mathbf{g}_i, & j \neq i \\ 0, & j = i \end{cases}$$

$$\mathbf{D} = \mathrm{diag}\{\frac{\gamma_1}{\mathbf{g}_1^H \mathbf{h}_{11}^H \mathbf{h}_{11} \mathbf{g}_1}, \ldots, \frac{\gamma_K}{\mathbf{g}_K^H \mathbf{h}_{KK}^H \mathbf{h}_{KK} \mathbf{g}_K}\}.$$

- The power vector can be found as:

$$\mathbf{q} = (\mathbf{D}^{-1} - \tilde{\boldsymbol{\Phi}})^{-1}\boldsymbol{\sigma} \tag{2}$$

- Comparing the definition we can see that $\tilde{\mathbf{\Phi}} = \mathbf{\Phi}^T$. This implies that there exists a duality relationship between the DL MISO and UL SIMO IFCs.
- We can extend the results for UL-DL duality for MAC/BC [Schubert & Boche'04] to the MISO/SIMO IFC:

Targets $\gamma_1, \ldots, \gamma_K$ are jointly feasible in UL and DL if and only if the spectral radius ρ of the weighted coupling matrix satisfies $\rho(\mathbf{D\Phi}) < 1$.

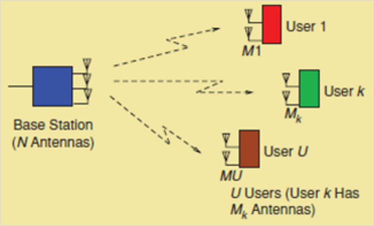
Both UL and DL have the same SINR feasible region under a sum-power constraint, i.e., target SINRs are feasible in the DL if and only if the same targets are feasible in the UL:

$$\sum_i q_i = \mathbf{1}^T\mathbf{q} = \sigma\mathbf{1}^T(\mathbf{D}^{-1} - \mathbf{\Phi}^T)^{-T} = \sigma\mathbf{1}^T(\mathbf{D}^{-1} - \mathbf{\Phi})^{-1} = \sum_i p_i \quad (3)$$

- Using this results it is possible to extend some BF design techniques used in the BC [Schubert & Boche'04] to the MISO IFC:
  - Max-Min SINR (SINR Balancing)
  - Power minimization under SINR constraints

MIMO BroadCast

- MIMO BC = Multi-User MIMO Downlink
- N_t transmission antennas.
- K users with N_k receiving antennas.
- Assume perfect CSI
- Possibly multiple streams/user d_k.
- Power constraint P
- Noise variance $\sigma^2 = 1$.
- $\mathbf{H}_k$ the MIMO channel for user k.



$$\begin{aligned}\mathbf{F}_k\mathbf{y}_k &= \mathbf{F}_k\mathbf{H}_k \textstyle\sum_{i=1}^{K} \mathbf{G}_i\mathbf{s}_i + \mathbf{F}_k\mathbf{z}_k \\ &= \underbrace{\mathbf{F}_k\mathbf{H}_k\mathbf{G}_k\mathbf{s}_k}_{\text{useful signal}} + \underbrace{\sum_{i=1, i\neq k}^{K} \mathbf{F}_k\mathbf{H}_k\mathbf{G}_i\mathbf{s}_i}_{\text{inter-user interference}} + \underbrace{\mathbf{F}_k\mathbf{z}_k}_{\text{noise}}\end{aligned}$$

System Model (2)

- Rx signal: $\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{z}_k = \mathbf{H}_k \sum_{i=1}^{K} \mathbf{G}_i \mathbf{s}_i + \mathbf{z}_k$
- $$\underbrace{\mathbf{F}_k}_{d_k \times N_k} \underbrace{\mathbf{y}_k}_{N_k \times 1} = \underbrace{\mathbf{F}_k}_{d_k \times N_k} \underbrace{\mathbf{H}_k}_{N_k \times N_t} \sum_{i=1}^{K} \underbrace{\mathbf{G}_i}_{N_t \times d_i} \underbrace{\mathbf{s}_i}_{d_i \times 1} + \underbrace{\mathbf{F}_k}_{d_k \times N_k} \underbrace{\mathbf{z}_k}_{N_k \times 1}$$
- [Christensen etal:T-WCdec08]: use of linear receivers in MIMO BC is not suboptimal (full CSIT, // SU MIMO): can prefilter $\mathbf{G}_k$ with a $d_k \times d_k$ unitary matrix to make interference plus noise prewhitened channel matrix - precoder cascade of user k orthogonal (columns)
- Optimal MIMO BC design requires DPC, which is significantly more complicated than BF.
- Multiple receive antennas cannot improve the sum rate prelog.
  So what benefit can they bring?
  Of course: cancellation of interference from other transmitters (spatially colored noise): not considered here.

Zero-Forcing (ZF)

- ZF-BF $\mathbf{F}_{1:i}\mathbf{H}_{1:i}\mathbf{G}_{1:i} =$

$$\begin{bmatrix} \mathbf{F}_1 & 0 \cdots 0 \\ 0 & \mathbf{F}_2 \ddots \vdots \\ \vdots & \ddots 0 \\ 0 \cdots & 0\, \mathbf{F}_i \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_i \end{bmatrix} [\mathbf{G}_1\ \mathbf{G}_2 \cdots \mathbf{G}_i] = \begin{bmatrix} \mathbf{F}_1\mathbf{H}_1\mathbf{G}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{F}_2\mathbf{H}_2\mathbf{G}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{F}_i\mathbf{H}_i\mathbf{G}_i \end{bmatrix}$$

- ZF-DPC (modulo reordering issues) $\mathbf{F}_{1:i}\mathbf{H}_{1:i}\mathbf{G}_{1:i} =$

$$\begin{bmatrix} \mathbf{F}_1 & 0 \cdots 0 \\ 0 & \mathbf{F}_2 \ddots \vdots \\ \vdots & \ddots 0 \\ 0 \cdots & 0\, \mathbf{F}_i \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_i \end{bmatrix} [\mathbf{G}_1\ \mathbf{G}_2 \cdots \mathbf{G}_i] = \begin{bmatrix} \mathbf{F}_1\mathbf{H}_1\mathbf{G}_1 & 0 & \cdots & 0 \\ * & \mathbf{F}_2\mathbf{H}_2\mathbf{G}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ * & \cdots & * & \mathbf{F}_i\mathbf{H}_i\mathbf{G}_i \end{bmatrix}$$

Stream Selection Criterion from Sum Rate

- At high SNR, both
  - optimized (MMSE style) filters vs. ZF filters
  - optimized vs. uniform power allocation

  only leads to $\frac{1}{\text{SNR}}$ terms in rates.
- At high SNR, the sum rate is of the form

$$\underbrace{N_t}_{\color{red}{\text{DoF}}} \log(\text{SNR}/N_t) + \underbrace{\sum_i \log\det(\mathbf{F}_i\mathbf{H}_i\mathbf{G}_i)}_{\color{blue}{\text{constant}}} + O(\frac{1}{\text{SNR}}) + \underbrace{O(\log\log(\text{SNR}))}_{\color{red}{\text{noncoherent Tx}}}$$

for properly normalized ZF Rx $\mathbf{F}_i$ and ZF Tx $\mathbf{G}_i$ (BF or DPC).

Role of Rx antennas?

- Different distributions of ZF between Tx and Rx give different ZF channel gains! If Rx ZF's k streams, hence Tx only has to ZF $M-1-k$ streams! So, number of possible solutions (assuming $d_k \equiv 1$):

$$\prod_{k=1}^{M}\left(\sum_{i=0}^{N_k-1}\frac{(M-1)!}{k!(M-1-k)!}\right)$$

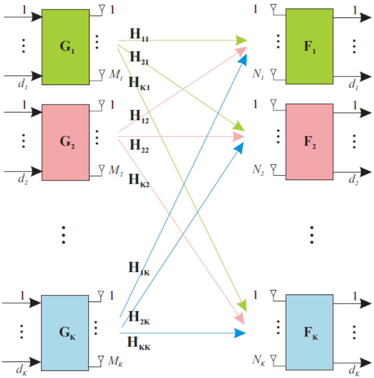
  for each user, Rx can ZF k between 0 and N_k-1 streams, to choose among $M-1$.
  Explains non-convexity of MIMO SR at high SNR.
- ZF by Rx can alternatively be interpreted as IA by Tx (Rx adapts Rx-channel cascades to lie in reduced dimension subspace).
- SESAM (and all existing MIMO stream selection algorithms): assumes that all ZF is done by Tx only. Hence, Rx can be a MF, matched to channel-BF cascade.

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

MIMO IFC Introduction

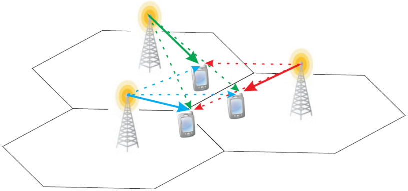
- Interference Alignment (IA) was introduced in [Cadambe,Jafar 2008]
- The objective of IA is to design the Tx beamforming matrices such that the interference at each non intended receiver lies in a common interference subspace
- If alignment is complete at the receiver simple Zero Forcing (ZF) can suppress 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)
- $d = \sum_{k=1}^{K} d_k$



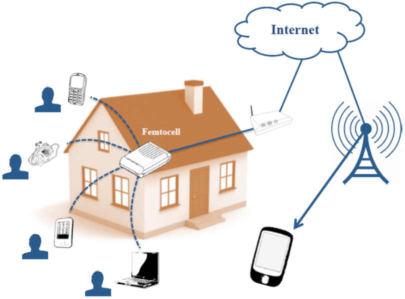
MIMO Interference Channel

Possible Application Scenarios

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



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



Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

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.

Various IA Flavors

- *linear* IA [GouJafar:IT1210], also called *signal space* IA, only uses the spatial dimensions introduced by multiple antennas.
- *asymptotic* IA [CadambeJafar:IT0808] uses symbol extension (in time and/or frequency), leading to (infinite) symbol extension involving diagonal channel matrices, requiring infinite channel diversity in those dimensions. This leads to infinite latency also. The (sum) DoF of asymptotic MIMO IA are determined by the *decomposition* bound [WangSunJafar:isit12].
- *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. $\text{offdiag}(H_2) = -\text{offdiag}(H_1)$. Ergodic IA also suffers from uncontrolled latency but provides the factor 2 rate loss at any SNR. The DoF of ergodic MIMO IA are also determined by the decomposition bound [LejosneSlockYuan:icassp14].
- *real* IA [MotahariGharanMaddah-AliKhandani:arxiv09], also called it signal scale IA, exploits discrete signal constellations and is based on the Diophantine equation. Although this approach appears still quite exploratory, some related work based on lattices appears promising.

IA as a Constrained Compressed SVD

- (compressed) SVD:

$$H = F\,D'\,G'^H = F\,[D\;0]\left[G\;G''\right]^H = F\,D\,G^H \;\Rightarrow\; F^H H G = D$$

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

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

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 $\Rightarrow$ $F^H HG =$ diagonal.

Interference Alignment: Feasibility Conditions (1)

- To derive the existence conditions we consider the ZF conditions

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

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

- rank requirement $\Rightarrow$ SU MIMO condition: $\quad d_k \leq \min(M_k, N_k)$
- 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)

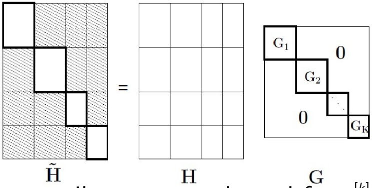
- A solution for the interference alignment problem can only exist if the total number of variables is greater than or equal to the total number of constraints i.e.,

$$\sum_{k=1}^{K} d_k(M_k - d_k) + \sum_{k=1}^{K} d_k(N_k - d_k) \geq \sum_{i\neq j=1}^{K} d_i\, d_j$$
$$\Rightarrow \; \sum_{k=1}^{K} d_k(M_k + N_k - 2d_k) \geq (\sum_{k=1}^{K} d_k)^2 - \sum_{k=1}^{K} d_k^2$$
$$\Rightarrow \; \sum_{k=1}^{K} d_k(M_k + N_k) \geq (\sum_{k=1}^{K} d_k)^2 + \sum_{k=1}^{K} d_k^2$$

- 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 ($M = N$): $\quad d \;=\; \frac{M}{2}$.
  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 $K > 3$.

Interference Alignment: Feasibility Conditions (3)

The main idea of IA is to convert the alignment requirements at each RX into a rank condition of an associated interference matrix $\mathbf{H}_I^{[k]}$=$[\mathbf{H}_{k1}\mathbf{G}_1, \ldots \mathbf{H}_{k(k-1)}\mathbf{G}_{(k-1)}, \mathbf{H}_{k(k+1)}\mathbf{G}_{(k+1)}, \ldots \mathbf{H}_{kk}\mathbf{G}_K]$, that spans the interference subspace at the k-th RX (the shaded blocks in each block row).Thus the dimension of the Interference subspace must satisfy $\text{rank}(\mathbf{H}_I^{[k]}) = r_I^{[k]} \leq N_k - d_k$



The equation above prescribes an upperbound for $r_I^{[k]}$ but the nature of the channel matrix (full rank) and the rank requirement of the BF specifies the following lower bound $r_I^{[k]} \geq \max_{l \neq k}(d_l - [M_l - N_k]_+)$. Imposing a rank $r_I^{[k]}$ on $\mathbf{H}_I^{[k]}$ implies imposing $(N_k - r_I^{[k]})(\sum_{\substack{l=1 \\ l \neq k}}^{K} d_l - r_I^{[k]})$ constraints at RX k. Enforcing the minimum number of constraints on the system implies to have maximum rank: $r_I^{[k]} \leq \min(d_{tot}, N_k) - d_k$

Interference Alignment: Feasibility Conditions (3)

- [BreslerTse:arxiv11]: counting equations and variables not the whole story!
- appears in very "rectangular" ($\neq$ square) MIMO systems
- example: $(M, N, d)^K = (4, 8, 3)^3$ MIMO IFC system
  comparing variables and ZF equations:
  $d = \frac{M+N}{K+1} = \frac{4+8}{3+1} = \frac{12}{4} = 3$ should be possible
- supportable interference subspace dim. $= N - d =$ 8-3 $= 5$
- however, the 2 interfering 8×4 cross channels generate 4-dimensional subspaces which in an 8-dimensional space do not intersect w.p. 1 !
- hence, the interfering 4×3 transmit filters cannot massage their 6-dimensional joint interference subspace into a 5-dimensional subspace!
- This issue is not captured by # variables vs # equations:
  $d = \frac{M+N}{K+1}$ only depends on $M + N$: $(5, 7, 3)^3$, $(6, 6, 3)^3$ work.

Feasibility Linear IA

- We shall focus here on linear IA, in which the spatial Tx filters align their various interference terms at a given user in a common subspace so that a Rx filter can zero force (ZF) it. Since linear IA only uses spatial filtering, it leads to low latency.
- The DoF of linear IA are upper bounded by the so-called **proper bound** [Negro:eusipco09], [Negro:spawc10], [YetisGouJafarKayran:SP10], which simply counts the number of filter variables vs. the number of ZF constraints.
- The proper bound is not always attained though because to make interference subspaces align, the channel subspaces in which they live have to sufficiently overlap to begin with, which is not always the case, as captured by the so-called **quantity bound** [Tingting:arxiv0913] and first elucidated in [BreslerCartwrightTse:allerton11], [BreslerCartwrightTse:itw11], [WangSunJafar:isit12].
- The transmitter coordination required for DL IA in a multi-cell setting corresponds to the Interfering Broadcast Channel (IBC). Depending on the number of interfering cells, the BS may run out of antennas to serve more than one user, which then leads to the Interference Channel (IC).

I and Q components: IA with Real Symbol Streams

- Using real signal constellations in place of complex constellations, transmission over a complex channel of any given dimension can be interpreted as transmission over a real channel of double the original dimensions (by treating the I and Q components as separate channels).
- This doubling of dimensions provides additional flexibility in achieving the total DoF available in the network.
- Split complex quantities in I and Q components:

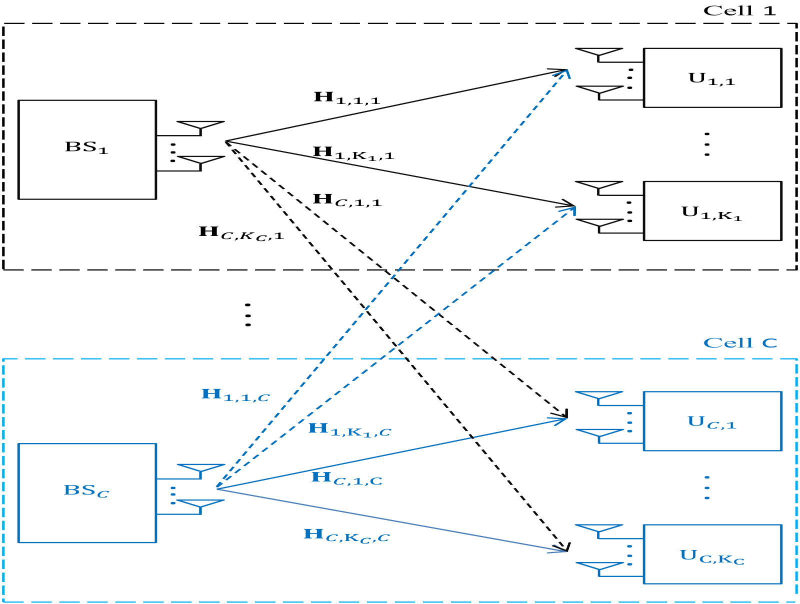
$$\mathbf{H}_{ij} = \begin{bmatrix} \mathrm{Re}\{\mathbf{H}_{ij}\} & -\mathrm{Im}\{\mathbf{H}_{ij}\} \\ \mathrm{Im}\{\mathbf{H}_{ij}\} & \mathrm{Re}\{\mathbf{H}_{ij}\} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \mathrm{Re}\{\mathbf{x}\} \\ \mathrm{Im}\{\mathbf{x}\} \end{bmatrix}$$

- Example: GMSK in GSM: was considered as wasting half of the resources, but in fact unknowingly anticipated interference treatment: 3 interfering GSM links can each support one GMSK signal without interference by proper joint Tx/Rx design! (SAIC: handles 1 interferer, requires only Rx design).

Outline

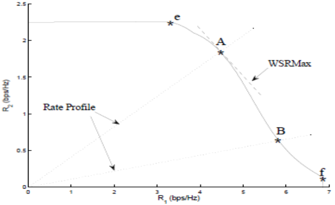
- interference single cell: Broadcast Channel (BC)
    - utility functions: SINR balancing, (weighted) sum rate (WSR)
    - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
    - Degrees of Freedom (DoF) and Interference Alignment (IA)
    - multi-cell multi-user: Interfering Broadcast Channel (IBC)
    - Weighted Sum Rate (WSR) maximization and UL/DL duality
    - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
    - CSIT: perfect, partial, LoS
    - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
    - distributed CSIT acquisition, netDoF
    - topology, rank reduced, decoupled Tx/Rx design, local CSIT
    - distributed designs
    - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

MIMO Interfering Broadcast Channel (IBC)



From IA to Optimized IFC's

- from Interference Alignment (=ZF) to max Sum Rate (SR) for the "Noisy IFC".
- to vary the point reached on the rate region boundary: SR $\rightarrow$ Weighted SR (WSR)
- problem: IFC rate region not convex $\Rightarrow$ multiple (local) optima for WSR (multiple boundary points with same tangent direction)
- solution of [CuiZhang:ita10]: WSR $\rightarrow$ max SR under rate profile constraint: $\frac{R_1}{\alpha_1} = \frac{R_2}{\alpha_2} = \cdots = \frac{R_K}{\alpha_K}$: $K-1$ constraints.
  Pro: explores systematically rate region boundary.
  Con: for a fixed rate profile, bad links drag down good links.
  $\Rightarrow$ stick to (W)SR
  (monitoring global opt issues).
  Note: multiple WSR solutions
  $\Leftrightarrow$ multiple IA solutions.



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

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

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 λ = 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

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

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

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

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

- 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

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

- 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 $\mathbf{f}_k$:

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

- 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

- 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

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

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

[KimGiannakis:IT0511] (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,

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

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:

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

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

[KimGiannakis:IT0511] = 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

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

[KimGiannakis:IT0511] Interference Aware WF

- Let $\sigma_k^{(1)} = \mathbf{g}_k^{'H}\mathbf{H}_k^H\widehat{R}_{\overline{k}}^{-1}\mathbf{H}_k\mathbf{g}_k^{'}$ and $\sigma_k^{(2)} = \mathbf{g}_k^{'H}\widehat{A}_k\mathbf{g}_k^{'}$.
- 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

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

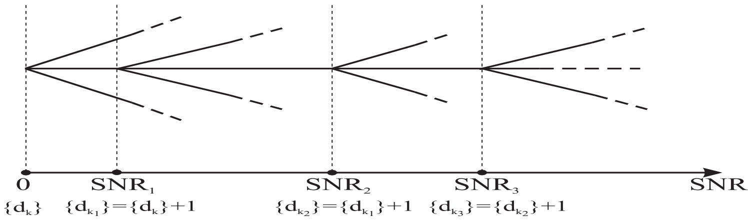
  where $P^{\perp}_{\mathbf{X}} = I - P_{\mathbf{X}}$ and $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$ projection matrices
  $(\mathbf{fH})_{\overline{k}}$ denotes the (up-down) stacking of $\mathbf{f}_i \mathbf{H}_i$ for users $i = 1, \ldots, K$, $i \neq k$.
- At low SNR, matched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

Deterministic Annealing

- At high SNR: max WSR solutions are ZF. When ZF is possible (IA feasible), multiple ZF solutions typically exist.
  Homotopy on the MIMO channel SVD:
  $H_{ji} = \sum_{k=1}^{d} \sigma_{jik} \mathbf{u}_{jik} \mathbf{v}_{jik}^H + t \sum_{k=d+1} \sigma_{jik} \mathbf{u}_{jik} \mathbf{v}_{jik}^H$
  The IA (ZF) condition for rank 1 link $i - j$ can be written as
  $$\sigma_{ji} \mathbf{f}_j^H \mathbf{u}_{ji} \mathbf{v}_{ji}^H \mathbf{g}_i = 0$$
- Two configurations are possible: $\mathbf{f}_j^H \mathbf{u}_{ji} = 0$ or $\mathbf{v}_{ji}^H \mathbf{g}_i = 0$
  Either the Tx or the Rx suppresses one particular interfering stream
- 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.

Deterministic Annealing (2)

- 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.



Difference of Convex Functions vs Majorization

- Difference of Convex functions: linearize convex part in terms of Tx covariance matrices Q_k to make it concave
- afterwards work with BF in $Q_k = \mathbf{g}_k \mathbf{g}_k^H$
- but the linearization in Q_k does not correspond to second-order Taylor series or any precise development in $\mathbf{g}_k$

- other interpretation: majorization: replace cost function to be maximized by one below it that touches the original one in one point [Stoica:SPmagJan04]
- specifically: matrix version of $x - 1 - \ln(x) \geq 0\,,\ x > 0$: Itakura-Saito distance in AR modeling ($x =$ ratio of true spectrum and AR model spectrum)
- majorized cost function can be optimized with any parameterization

MIMO Tx/Rx Design w Other Utility Functions

- in all cases (e.g. also SINR balancing),

  Rx filter = LMMSE
  Tx filter = LMMSE in dual uplink

- influence of precise utility function is in the design of the actual & dual stream powers and noise variances

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

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.

Location Aided Partial CSIT LoS Channel Model

- Assuming the Tx disposes of not much more than the LoS component information, model

$$\mathbf{H} = \mathbf{h}_r\,\mathbf{h}_t^H(\theta) + \widetilde{\mathbf{H}}'$$

where θ is the LoS AoD and the Tx side array response is normalized: $||\mathbf{h}_t(\theta)||^2 = 1$.

- Since the orientation of the MT is random, model the Rx side LoS array response $\mathbf{h}_r$ as vector of i.i.d. complex Gaussian

$$\begin{array}{ll}\mathbf{h}_r \;\; \text{i.i.d.} & \sim \mathcal{CN}(0, \frac{\mu}{\mu+1}) \quad \text{and} \\ \widetilde{\mathbf{H}}' \;\; \text{i.i.d.} & \sim \mathcal{CN}(0, \frac{1}{\mu+1}\frac{1}{M}) \text{ , independent of } \mathbf{h}_r,\end{array}$$

where the matrix $\widetilde{\mathbf{H}}$ represents the aggregate NLoS components.

Location Aided Partial CSIT LoS Channel Model (2)

- Note that

$$\mathrm{E}\ ||\mathbf{H}||_F^2 = \mathrm{E}\ \mathrm{tr}\{\mathbf{H}^H\mathbf{H}\} =$$
$$||\mathbf{h}_t(\theta)||^2\ \mathrm{E}\ ||\mathbf{h}_r||^2 + \mathrm{E}\ ||\widetilde{\mathbf{H}}'||_F^2 = \tfrac{\mu N}{\mu+1} + \tfrac{N}{\mu+1} = N\,,$$

$(\mathrm{E}\ ||\mathbf{h}_r\,\mathbf{h}_t^T(\theta)||_F^2)/(\mathrm{E}\ ||\widetilde{\mathbf{H}}'||_F^2) = \mu =$ a Rice factor.

- In fact the only parameter additional to the LoS AoD θ is μ.
- So, this is a case of zero mean CSIT and Tx side covariance CSIT

$$R_t = \mathrm{E}\ \mathbf{H}^H\mathbf{H} = \frac{\mu N}{\mu+1}\mathbf{h}_t(\theta)\mathbf{h}_t^H(\theta) + \frac{N}{\mu+1}\frac{1}{M}I_M\ .$$

Location Aided LoS ZF BF

- For ZF BF, the BS shall use for user k a spatial filter $\mathbf{g}_k = \sqrt{p_k}\,\mathbf{g}_k^{'}$ such that $\mathbf{g}_k^{'} = \mathbf{g}_k^{"}/||\mathbf{g}_k^{"}||$

$$\mathbf{g}_k^{"} = P^{\perp}_{\mathbf{h}_{t,\overline{k}}}\mathbf{h}_{t,k}$$

  where $\mathbf{h}_{t,\overline{k}} = [\mathbf{h}_{t,1} \cdots \mathbf{h}_{t,k-1}\,\mathbf{h}_{t,k+1} \cdots \mathbf{h}_{t,K}]$.
- And uniform power distribution $p_k = P/K,\ k = 1, \ldots, K$.
- The $\mathbf{g}_k^{"}$ can also be computed from

$$\mathbf{g}^{"} = [\mathbf{g}_1^{"} \cdots \mathbf{g}_K^{"}] = \mathbf{h}_t(\mathbf{h}_t^H\mathbf{h}_t)^{-1}\ ,\ \ \mathbf{h}_t = [\mathbf{h}_{t,1} \cdots \mathbf{h}_{t,K}]\,.$$

LoS based Partial CSIT: Beyond ZF

- Go beyond the asymptotics of high SNR and high Ricean factor: even if the Tx ignores the multipath and the Rx can handle it, it would be better to have a multipath aware Tx design. Note that the Ricean factor μ satisfies uplink/downlink (UL/DL) reciprocity, even in a FDD. Solution: previous partial CSIT design.

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: $WSR(\mathbf{g}, \mathbf{H}) = \max_{\mathbf{f}} \sum_k u_k(-\ln(e_k(\mathbf{f}_k, \mathbf{g})))$.

Max EWSR by Stochastic Approximation

- In [Luo:spawc13] a stochastic approximation approach for maximizing the EWSR was introduced: replace statistical average by sample average (samples of $\mathbf{H}$ get generated according to its Gaussian CSIT distribution in a Monte Carlo fashion), and one iteration of the min WSMSE approach gets executed per term added in the sample average.
- Some issues: in this case the number of iterations may get dictated by a sufficient size for the sample average rather than by a convergence requirement for the iterative approach.
- Another issue is that this approach converges to a local maximum of the EWSR. It is not immediately clear how to combine this stochastic approximation approach with deterministic annealing.
- Below: various deterministic approximations and bounds for the EWSR, which can then be optimized as in the full CSI case.

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.

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

  $\Rightarrow$ 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.

EWSR Upper Bound: WSESINR

- Using the concavity of ln(.), we get

$$EWSR(\mathbf{g}) \leq \sum_{k=1}^{K} u_k \ln(1 + \mathrm{E}_{\mathbf{H}_k} \mathrm{SINR}_k(\mathbf{g}, \mathbf{H}_k)) .$$

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

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

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

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

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

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

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

- SU MIMO asymptotics from [Loubaton:IT0310],[Taricco:IT0808] (in which both $M, N \to \infty$, which tends to give more precise approximations when M is not so large) for a term of the form $\ln\det(Q\mathbf{H}^H\mathbf{H}+I)$ correspond to replacing $\mathbf{H}_k^H\mathbf{H}_k$ in the $\widetilde{R}_k$ and $\widetilde{R}_{\overline{k}}$ with a kind of $R_{t,k}$ with a different weighting of the $\overline{\mathbf{H}}_k^H\overline{\mathbf{H}}_k$ and $C_{t,k}$ portions, of the form $R_{t,k}^{'} = a_k C_{t,k} + \overline{\mathbf{H}}_k^H\mathbf{B}_k\overline{\mathbf{H}}_k$ for some scalar a_k and matrix $\mathbf{B}_k$ that depends on $C_{r,k}$.
- For the general case of Gaussian CSIT with separable (Kronecker) covariance, get

$$\begin{aligned}&\mathrm{E}_{\mathbf{H}}\ln\det(I+\mathbf{H}Q\mathbf{H}^H)\\ &= \max_{z,w}\left\{\ln\det\begin{bmatrix} I+wC_r & \overline{\mathbf{H}} \\ -Q\overline{\mathbf{H}}^H & I+zQC_t\end{bmatrix} - zw\right\}.\end{aligned}$$

$\max_{z,w}$ interpretation is new.

Large MIMO Asymptotics Refinement (2)

- Simpler case: zero channel means $\overline{\mathbf{H}}_k = 0$ and no Rx side correlations $C_r = I$, and with per user Tx side correlations $C_t \leftarrow C_k$, the EWSR w large MIMO asymptotics:

$$
\begin{aligned}
&EWSR = \\
&\sum_{k=1}^{K} \Big\{ u_k \max_{z_k, w_k} \left[\ln\det(I + z_k G G^H C_k) + N_k \ln(1 + w_k) - z_k w_k \right] \\
&- u_k \max_{z_{\overline{k}}, w_{\overline{k}}} \left[\ln\det(I + z_{\overline{k}} G_{\overline{k}} G_{\overline{k}}^H C_k) + N_k \ln(1 + w_{\overline{k}}) - z_{\overline{k}} w_{\overline{k}} \right] \Big\}
\end{aligned}
$$

where $G = [\mathbf{g}_1 \cdots \mathbf{g}_K]$ and $G_{\overline{k}}$ is the same as G except for column $\mathbf{g}_k$. Can be maximized by alternating optimization.

WSMSE-DC BF relation

- min WSMSE iteration $(i+1)$

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

WSMSE does one power iteration of DC !!

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

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

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

MaMIMO: more interesting Large System Regime

- MaMIMO regime considered above: $M, N \to \infty$, $M/N \to$ constant, K finite
- MaMIMO regime of more interest: $M, K \to \infty$, $M/K \to$ constant, N finite
- considered in [1], for (R-)ZF BF, Gaussian channel vectors with arbitrary covariance matrices and CSIT errors
- optimal BF are considered in [2]

[1] S. Wagner, R. Couillet, M. Debbah, D. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels under Limited Feedback," IEEE Trans. Information Theory, July 2012.

[2] S. Wagner and D. Slock, "Weighted Sum Rate Maximization of Correlated MISO Broadcast Channels under Linear Precoding: A Large System Analysis," in *Proc. SPAWC*, San Francisco, CA, USA, June 2011.

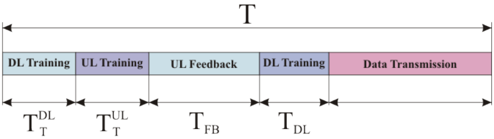
Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

IC CSI Acquisition Outline

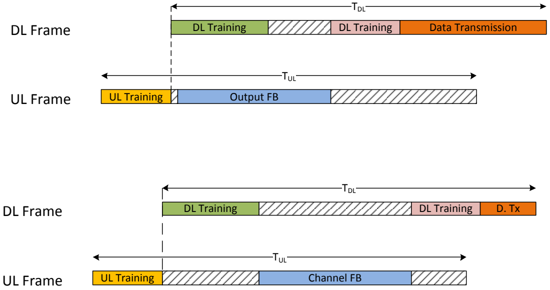
- Centralized CSIT Acquisition
- Distributed CSIT Acquisition
- Channel Feedback & Output Feedback
- DoF optimization as a function of coherence time

Transmission Phases



- 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

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 $N_t \geq N_r$

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

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 $\mathbf{H}_{ik}$, $\forall i$ through reciprocity
- MU_i do not need to feedback $\mathbf{H}_{ik}$ 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 almost equivalent to FDD
- In Centralized Processing reciprocity makes channel feedback NOT required

Concluding Remarks

- DoF in multi-user systems accounting for (channel) feedback are extremely sensitive to channel model.
- All this argues for shrinking the Feedback delay as much as possible: in FDD, feedback delay can be shrunk to roundtrip delay! **Immediate Feedback**.

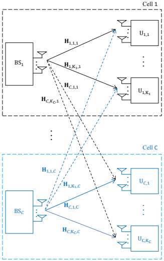
Outline

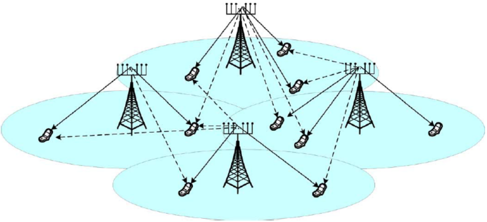
- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

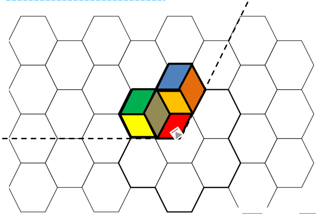
Reduced CSIT and Decoupled Tx/Rx Design

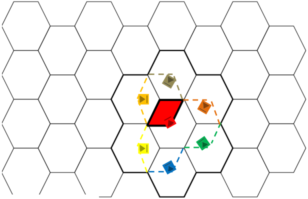
- for IA to apply to cellular: overall Tx/Rx design has to decompose so that the CSIT required is no longer global and remains bounded regardless of the network size.
- simplest case : local CSIT : a BS only needs to know the channels from itself to all terminals. In the TDD case : reciprocity. The local 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.

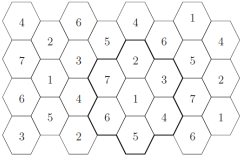
Massive MIMO: topological aspects











Clustered Topological MIMO IBC

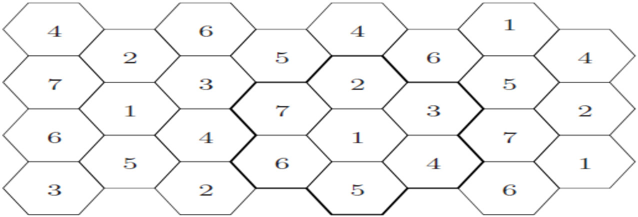


Figure : Hexagonal cellular system with cluster size $C = 7$.

We propose here an approach to an infinite IBC network by exploiting topology, enforcing CSI to be local to clusters, and reverse engineering the numbers of antennas required.

Consider partitioning an infinite IBC into finite IBC clusters. Within a finite IBC cluster, CSI acquistion can be performed in a distributed fashion. Then antennas get added to the BS in order to perform ZF of the finite inter-cluster links (due to topology, longer links can be neglected). Traditional cellular system, with interference limited to the first tier (6 cells). Hence we get a cluster size of $C = 7$.

As for GSM frequency reuse, the whole area can be covered by contiguous repetition of the cluster pattern. However, here cell numbering in a cluster has nothing to do with freq. reuse (=1).

Apart from these additional constraints for the cluster edge cell Txs, the Tx/Rx design within a cluster may seem like that of a standard MIMO IBC. However, the topology also affects within a cluster (alternatively, this could be not exploited). Hence if we consider the channel blocks between the 7 cells, we get an overall channel matrix of the form

$$\mathbf{H} = \begin{bmatrix} * & * & * & * & * & * & * \\ * & * & * & 0 & 0 & 0 & * \\ * & * & * & * & 0 & 0 & 0 \\ * & 0 & * & * & * & 0 & 0 \\ * & 0 & 0 & * & * & * & 0 \\ * & 0 & 0 & 0 & * & * & * \\ * & * & 0 & 0 & 0 & * & * \end{bmatrix}$$

where the "$*$" entries denote non-zero blocks. The $d_{c,k}$ streams for user (c,k) get extracted from its Rx signal $y_{c,k}$ by a $d_{c,k} \times N_{c,k}$ Rx filter $F_{c,k}$. To get the DoF, we need to count the number of streams that can pass through the Tx/Rx filters in parallel without suffering interference.

Sectored Cells

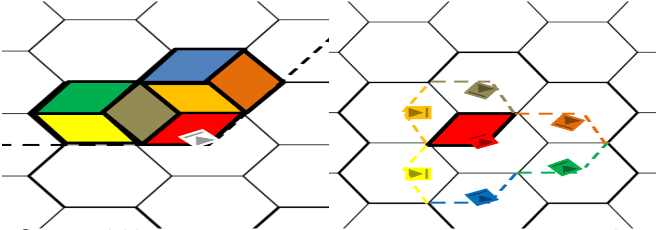


Figure : Sectored Hexagonal cellular system with 3 sectors and cluster size $C = 7$. The left figure indicates the sectors in which MTs receive a certain sector BS, as for data Tx or training. The right figure indicates the BS sectors that Rx FB from MTs in a certain sector.

The topological (distance) aspect introduces a certain "banded" character in the overall channel matrix $\mathbf{H}$: the number of non-zero blocks in any block row or block column remains finite (of cluster size C) regardless of the overall matrix size. Sectoring furthermore adds a certain spatial causality. Indeed a certain sector BS will only affect a portion $\frac{1}{3}$ of the MTs in the case of 3 sectors. This leads to a "triangular" $\mathbf{H}$. Nevertheless, as only the BS Tx/Rx are sectored, and not the MTs, with interference up to the first tier, again a cluster size of $C = 7$ (figure)

Interfering HetNets

- The design can be applied to the case of HetNets (heterogeneous networks), with multiple small cells per macro cell.
- In the topological case, we can e.g. assume that the small cell MTs Rx macro interference just like the macro MTs, but the small cell BS only interfere to the (all) MTs within the cell.
- In case of K_m macro cell MTs per cell and K_s small cell MTs, the design for the Tx filters at the macro BS remains unchanged, after replacing $K = K_m + K_s$. The small cell BS only needs Tx antennas to ZF to local users within the macro cell.

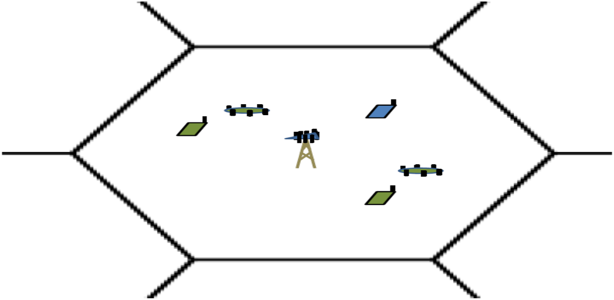


Figure : Zoom on a HetNet macro cell with two small cells in the macro cell and one MT for each of the three BS.

In the HetNet scenario, it may be of interest to adapt the Rx with interference that is only aligned for a subset of the interferers. For the remaining interferers, the Rx then appears as fixed and the ZF work has to be done by the corresponding Txs. As the concept of incomplete CSIT [deKerretGesbert:TWC13] shows, this may be not that suboptimal, depending on the antenna configurations. For the HetNet scenario, consider an IBC design per macro cell (as cluster)

Topological IA

- topological CSIT: only CSIT is knowledge about which Tx is connected to which Rx
- example: 3 single-antenna Tx code over 3 subcarriers. cascade precoders - channels :

$$\mathbf{H} = \begin{bmatrix} h_{11}c_{11} & h_{12}c_{12} & h_{13}c_{13} \\ h_{21}c_{21} & h_{22}c_{22} & h_{23}c_{23} \\ h_{31}c_{31} & h_{32}c_{32} & h_{33}c_{33} \end{bmatrix}$$

- for IA, need to align 3 interference columns in 2D subspace $\det(\mathbf{H}) = 0, \forall h_{ij}$
  $\Rightarrow$ all 6 monomials need to be zero, e.g. $c_{11}c_{22}c_{33} = 0 \Rightarrow$ one of them $= 0$. $\Rightarrow$ optimality of orthogonalization (frequency reuse).
- Graph from rows to columns of matrix $\mathbf{H}$: matching number of bipartite graph should be at most 2. For general n: matching number at most $n-1$.

Topological IA (2)

- Case of multiple Tx antennas (matrix encoder blocks): generalization of matching number: matroids.
- topic initiated by Jafar who exposed a link to index coding.
- topological CSIT is a very poor CSIT

A. El Gamal, N. Naderializadeh and A.S. Avestimehr, "When Does an Ensemble of Matrices with Randomly Scaled Rows Lose Rank?," in *Proc. ISIT*, Hong Kong, June 2015.

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

Distributed Sum Utility Optimization

- weighted sum rate (WSR) inspired interference pricing:

$$\text{WSR} = \sum_k u_k \; \ln(1+\text{SINR}_k) = \sum_k u_k \; \ln(1+\frac{\text{S}_k}{\text{I}_k+\text{N}_k})$$

  sensitivity rate user k to beamformer for user i

$$\frac{\partial}{\partial \mathbf{g}_i} \ln(1+\text{SINR}_k) = \underbrace{-\frac{\text{SINR}_k}{1+\text{SINR}_k}\frac{1}{\text{I}_k+\text{N}_k}}_{\text{interference price}} \frac{\partial}{\partial \mathbf{g}_i} \text{I}_k$$

  Feed back interference prices instead of CSIT.
  $\frac{\partial}{\partial \mathbf{g}_i}\text{I}_k$ involves $\mathbf{f}^H\mathbf{H}$, Rx-channel cascade.
- UE initiated BS zero-forcing: UE decides which cross links (paths) need to be zero-forced by the corresponding BS
- Cognitive Radio approaches for distributed design: interference temperatures for intercell or intercluster links
- distributed utility optimization may call for new paradigms

Distributed IBC Designs

- Distributed here: distribution over cells (BS), assuming BS in a cell has all intracell CSIT.

0. Distributed replication of centralized design.
1. Distributed Tx design with (iterative) fast fading exchange: take centralized design and exchange (between cells) whatever (fast fading) information is required.
2. MaMIMO motivated distributed Tx design with 2 stage fast fading exchange.
3. MaMIMO motivated distributed Tx design with slow fading exchange.
4. 2-directional training

0. Distributed replication of centralized design

- Global intracluster CSIT can be gathered but it takes an overhead that evolves with C^2 [Negro:isccsp12].
- Hence such an approch does not scale with the number of cells.
- Other issues: different partial CSIT at every BS $\Rightarrow$ if let converge to a local optimum, different BS may converge to different local optima!
  Even deterministic annealing may lead to different presumed global optima at different BS!

1. Distributed Tx design with (iterative) fast fading exchange

- Some of the centralized designs with alternating optimization are very sensitive to a strict execution of the alternating optimization strategy and do not support asynchronous updating at different BS. Was pointed out in
- [ScutariFacchineiSongPalomarPang:T-SPfeb14], where 2 mechanisms are introduced that can can guarantee convergence:
  (i) introduction of a stepsize, controlling the magnitude of the update from the old to the new BF solution, and
  (ii) the addition of a proximal term to the cost function, a quadratic deviation between new and old solution.
- [YangScutariPalomarPesavento:arxiv1410.5076] extension to stochastic approximation with guaranteed convergence, in which iterative process is used to average over channel realizations also: converges to the solution of max Expected WSR without requiring to learn the channel distributions!

2. Tx design with 2 stage fast fading exchange

- In this approach, the fast fading Rx's are designed in a first pure intracell design.
  Hence the role of the Rx is here constrained to handle intracell interference, whereas intercell interference is (actively) handled by Tx only.
- To account for the intercell management in the second stage, the transmit power constraint for the intracell design in the first stage is reduced, based on large system design guidelines.

- Variant considered here: 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.

2'. 2-stage 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 $\breve{R}_{\overline{k}}$.

2'. 2-stage 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 $\breve{A}_k$ becomes

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

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

3. Tx design with slow fading exchange (?)

- multi Rx antenna extension of [ZhangCui:T-SPoct10], [DahroujYu:T-WCmay10]
- In the large system regime (large number of BS antennas and users, finite UE antennas), not all finite dimensional quantities converge to a deterministic limit (e.g. Rxs). But (scalar) quantities of the nature of (co)variance, SINR, MSE, etc. do.
- [ChenYou:arXiv1309.4034] minimax Lagrangian duality

$$\max_{Q_k \geq 0} \min_{R_{\overline{k}}} \sum_{k=1}^{K} u_k [\ln\det(R_{\overline{k}} + \mathbf{H}_k Q_k \mathbf{H}_k^H) - \ln\det(R_{\overline{k}})]$$

$$\mathrm{tr}\{Q_k\} \leq P_k \,,\; R_{\overline{k}} \geq \mathbf{H}_k (\textstyle\sum_{i \neq k} Q_i) \mathbf{H}_k^H + I_{N_k}$$

- main quantity to be exchanged: $R_{\overline{k}}$. But it does not harden in MaMIMO.
- Related work by Antti Tölli & coworkers [icc14,eusipco14], in [MüllerCouilletBjörnsonWagnerDebbah:submT-SP14] and in [LagenVidal:T-WC15] where in a TDD MaMIMO setting only local fast CSIT would be required.

4. 2-directional training

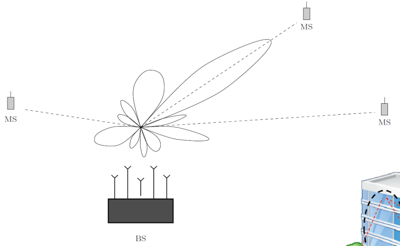
- Introduced in [ShiBerryHonig:T-SPfeb14], followed up by [JayasingheTolliKalevaLatva-aho:icc15].
- Exploits channel reciprocity in TDD and the fact that a MIMO IA solution in the DL is immediately also a MIMO IA solution in the UL and vice versa:
  $F\,\mathbf{H}\,G = \text{diag} \;\Rightarrow\; G^T\,\mathbf{H}^T\,F^T = \text{diag}.$
- In a MU setting, adaptation cannot be done completely blindly since the user signals need to be distinguished. Hence a streamwise differentiation needs to be introduced which can take many forms, for instance superimposed pilots. In superfast 2-way adaptation, each device manages a linear MMSE adaptive Rx filter that uses the pilot signal as desired response. This Rx filter is directly used as Tx filter also at all time instants.
- reciprocal channels but non-reciprocal environment & utilities: non-reciprocity of Tx powers, Rx noises, utility functions. In this case separate filters need to be adapted for Rx and Tx.

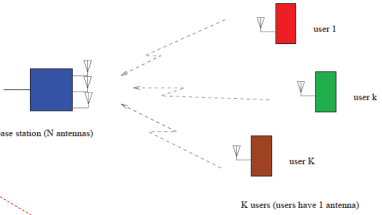
Outline

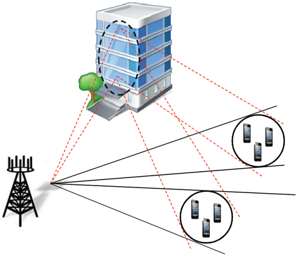
- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

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

Specular Wireless (Massive) MIMO Channel Model

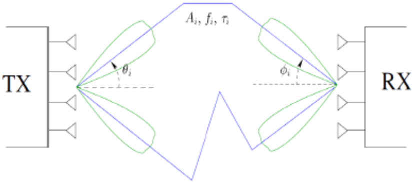


Figure 1 : MIMO transmission with M transmit and N receive antennas.

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

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

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

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

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

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

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

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

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

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

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

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

(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

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

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**

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

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

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

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

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

$$C^t = AA^H$$

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

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

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**

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

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:

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

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*

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

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

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 $A_{i,k,j}$ 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.

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - MIMO BC : DPC vs BF, role of Rx antennas (IA, local optima)
- interference multi-cell/HetNets: Interference Channel (IC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - multi-cell multi-user: Interfering Broadcast Channel (IBC)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
- Max WSR with Partial CSIT
  - CSIT: perfect, partial, LoS
  - EWSMSE, Massive MIMO limit, large MIMO asymptotics
- CSIT acquisition and distributed designs
  - distributed CSIT acquisition, netDoF
  - topology, rank reduced, decoupled Tx/Rx design, local CSIT
  - distributed designs
  - Massive MIMO, mmWave : covariance CSIT, pathwise CSIT

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:

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

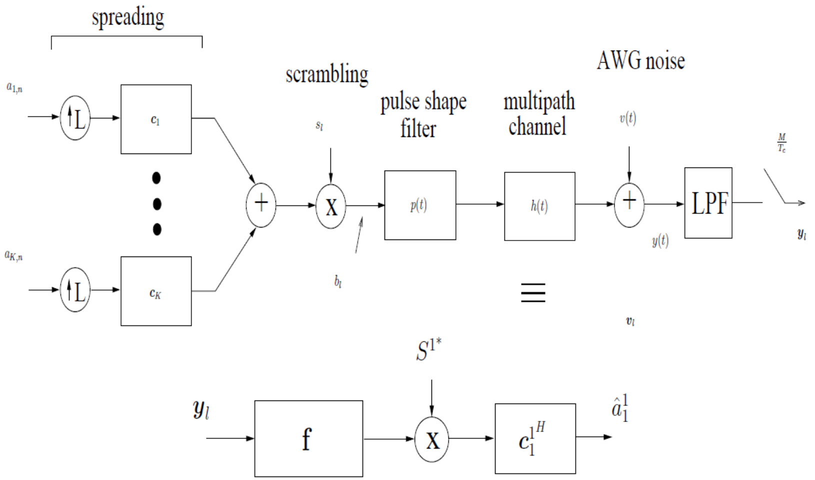
If we denote by L_F, L_H, L_G the length of the 3 types of filters, then in a symmetric configuration, the proper conditions become

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

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.

Reminder: UMTS DL Chip Equalization



- chip equalizer Rx structure: filter + descrambler + correlator
  filter = channel Matched Filter (MF): RAKE receiver
  filter = channel equalizer: chip equalizer

Reminder: UMTS DL Chip Equalization (2)

LMMSE chip equalizer vs true LMMSE receiver

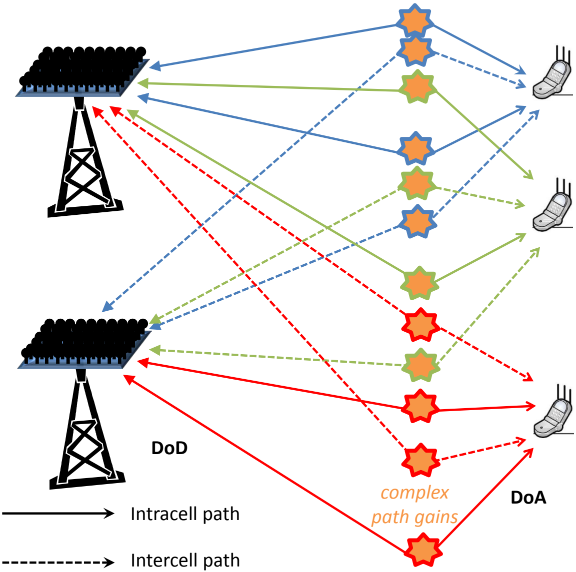
- SINR maximized by MMSE receiver (within class of linear RXs)
- True MMSE Rx:

$$\widehat{a} = F\,Y = R_{aY} R_{YY}^{-1}\,Y = (E_{A,V}\,a Y^H)\,(E_{A,V}\,Y\,Y^H)^{-1}\,Y$$

  F time-varying since scrambler considered deterministic here (known)
- Approximate MMSE Rx:

$$\begin{aligned}\widehat{a} &= (E_{A,V}\,a Y^H)\,(E_{A,V,S}\,Y\,Y^H)^{-1}\,Y \\ &= \sigma_a^2 \underbrace{c^H}_{\text{correlator}} \underbrace{S^H}_{\text{descrambler}} \underbrace{\underbrace{\mathcal{T}(h)^H}_{\text{MF}} (\underbrace{E_{A,V,S}\,Y\,Y^H}_{\text{(block) Toeplitz}})^{-1}}_{\text{LTI chip rate MMSE equalizer}} \; Y\end{aligned}$$

Pathwise Multi-User Multi-Cell



scrambler $\rightarrow$ path gains, DoAs ; DL $\rightarrow$ dual UL LMMSE

DL BF as Dual UL LMMSE Rx

- 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 $\qquad \sigma^2_{\widetilde{x}_i} = u_i\,w_i,\ R_{\widetilde{v}_k\widetilde{v}_k} = \lambda_k\,I_{M_k}$

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

where the ficitious dual UL Tx signals $\widetilde{x}_i$ are uncorrelated zero mean with variance $\sigma^2_{\widetilde{x}_i} = u_i\,w_i$ and the fictitious dual UL Rx noise has covariance matrix $R_{\widetilde{v}_k\widetilde{v}_k} = \lambda_k\,I_{M_k}$.

DL BF as Dual UL LMMSE Rx (2)

- The dual UL Rx signal at BS k can be rewritten as

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

$$\begin{aligned}
\mathbf{H}_k^H &= [\mathbf{H}_{k,m_k+1}^H \cdots \mathbf{H}_{k,m_k+K_k}^H],\\
\mathbf{H}_{\overline{k}}^H &= [\mathbf{H}_1^H \cdots \mathbf{H}_{k-1}^H \, \mathbf{H}_{k+1}^H \cdots \mathbf{H}_C^H],\\
F_k &= \text{blockdiag}\{\mathbf{f}_{m_k+1}, \ldots, \mathbf{f}_{m_k+K_k}\},\\
F_{\overline{k}} &= \text{blockdiag}\{F_1, \ldots, F_{k-1} \, F_{k+1}, \ldots, F_C\}
\end{aligned}$$

$m_k = \sum_{i=1}^{k-1} K_i$ and corresponding block structure for the super vectors $\widetilde{x}_k$, $\widetilde{x}_{\overline{k}}$.

- This leads to the DL BF as an UL LMMSE Rx: (for all intracell users jointly)

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

which can be seen to correspond to

$$\mathbf{g}_k^H = u_k w_k \mathbf{f}_k^H \mathbf{H}_{k,b_k} (\sum_i u_i w_i \mathbf{H}_{i,b_k}^H \mathbf{f}_i \mathbf{f}_i^H \mathbf{H}_{i,b_k} + \lambda_{b_k} I_M)^{-1}$$

Pathwise Dual UL

- motivation pathwise dual UL LMMSE: to what extent to cancel (intercell) interfering paths?
- Substituting the channel response matrices in terms of their pathwise factored form

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

$\widetilde{\mathbf{s}}_{i,k}$ = (vectors of) fictious pathwise UL Tx signals from user i to BS k.

- The factors $\mathbf{B}_{i,k}$ are now treated as unknown, modeled as independent with zero mean i.i.d. elements of variance $\frac{1}{N_i}$. As a result we get for the correlation matrices $R_{\widetilde{\mathbf{s}}_{i,k}\widetilde{\mathbf{s}}_{i,k}} = \frac{||\mathbf{f}_i||^2}{N_i}\sigma^2_{\widetilde{x}_i} I$.
- Similarly to the userwise, the pathwise dual UL Rx signal at BS k above can be rewritten as

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

where $\widetilde{\mathbf{s}}_k = \mathbf{B}_k^H F_k \widetilde{x}_k$, $\widetilde{\mathbf{s}}_{\overline{k}} = \mathbf{B}_{\overline{k}}^H F_{\overline{k}} \widetilde{x}_{\overline{k}}$ and A_k,$\mathbf{B}_k$ and $A_{\overline{k}}$,$\mathbf{B}_{\overline{k}}$ have similar block structure as $\mathbf{H}_k$ and $\mathbf{H}_{\overline{k}}$ resp. except for different block sizes.

Pathwise Dual UL (2)

- Hence we get the pathwise DL BF as an UL LMMSE Rx (for all intracell paths jointly)

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

where $R_{\widetilde{\mathbf{s}}_k \widetilde{\mathbf{s}}_k}$, $R_{\widetilde{\mathbf{s}}_{\overline{k}} \widetilde{\mathbf{s}}_{\overline{k}}}$ are diagonal.

2-stage BF: Pathwise + Userwise Intracell

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

  $$A_k^H R^{-1} = A_k^H (A_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}}\widetilde{\mathbf{s}}_{\overline{k}}} A_{\overline{k}}^H + \lambda_k I)^{-1}$$

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

  $$A_k^H R^{-1} \rightarrow A_{i,k}^H (A_{\overline{i},k} R_{\widetilde{\mathbf{s}}_{\overline{i},k}\widetilde{\mathbf{s}}_{\overline{i},k}} A_{\overline{i},k}^H + A_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}}\widetilde{\mathbf{s}}_{\overline{k}}} A_{\overline{k}}^H + \lambda_k I)^{-1}$$

  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' BFs minimize dual weighted sum MSE at path outputs

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

  It is not clear if this weighting is optimal for channel estimation also (but it goes in the right direction).

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

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

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

Relation Pathwise - Cognitive Radio Design

- The pathwise intercell design can be interpreted as a cell-wise intracell design with intercell interference constraints of the form (for cell i)

$$\sum_{k:b_k=i} u_k \ \ln \frac{1}{e_k} + \sum_{k:b_k=i} \sum_{n:b_n\neq i} \mu_{k,n}(|\mathbf{g}_k^H \mathbf{H}_{k,b_n}^H \mathbf{f}_n|^2 - Q_{k,n})$$

where the $\mu_{k,n} = \sigma^2_{\tilde{x}_n}$ are Lagrange multipliers and the $Q_{k,n}$ are linkwise interference power constraints.

- The pathwise approach is obtained by replacing the second term by its expected value w.r.t. the $\mathbf{B}$ factors, leading for the quadratic terms to

$$\mathrm{E}_{\,\mathbf{B}} \sum_{k:b_k=i} \sum_{n:b_n\neq i} \mu_{k,n} |\mathbf{g}_k^H \mathbf{H}_{k,b_n}^H \mathbf{f}_n|^2 = \sum_{k:b_k=i} \mathbf{g}_k^H \, A_{\overline{k}} R_{\widetilde{\mathbf{s}}_{\overline{k}} \widetilde{\mathbf{s}}_{\overline{k}}} A_{\overline{k}}^H \, \mathbf{g}_k \, .$$

- The pathwise philosophy corresponds to no intercell exchange of fast fading information. Hence the intercell exchange involves long-term averages for $\sigma^2_{\tilde{x}_n} = u_n/e_n$ and for the noise variance (which includes residual intercell interference).

5G Perspectives

- multi-user multi-cell interference management: theoretical possibilities, but (global) CSIT required
  - FB delay $\Rightarrow$ channel prediction and channel Doppler models crucial
  - analog channel FB?
  - FDD: immediate channel FB
  - distributed : yes but watch for fast fading
  - full duplex radio: reciprocity since single frequency, immediate use of channel estimates (no FDD FB or TDD ping-pong delay)
- Massive MIMO simplifications:
  separating fast and slow fading channel components
  decoupling of cells?
- mmWave (beamforming, bandwidth), spectrum aggregation
- new waveforms: windowed OFDM ?
- beyond classical cellular:
  - HetNets (macro/small):
  - wireless/self backhauling
  - D2D, cloud, IoT (low rate), COM for control (low latency),

A Few References

- Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Ergodic Interference Alignment for MIMO Interference Networks," in *Proc. ICASSP*, Florence, Italy, May 2014.
- T. Liu and C. Yang, "Genie Chain and Degrees of Freedom of Symmetric MIMO Interference Broadcast Channels," Sep. 2013, [Online]. Available: http://arxiv.org/abs/1309.6727.
- F. Negro, I. Ghauri, and D. T. M. Slock, "Deterministic Annealing Design and Analysis of the Noisy MIMO Interference Channel," in *Proc. IEEE Information Theory and Applications workshop (ITA)*, San Diego, CA, USA, Feb. 2011.
- F. Negro, M. Cardone, I. Ghauri, and D. T. M. Slock, "SINR Balancing and Beamforming for the MISO Interference Channel," in *Proc. IEEE PIMRC*, Toronto, Canada, Sept. 2011.
- F. Negro, D. Slock, and I. Ghauri, "On the Noisy MIMO Interference Channel with CSI through Analog Feedback," in *Int'l Symp. Communications, Control and Sig. Proc. (ISCCSP)*, Rome, Italy, May 2013.
- Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Foresighted Delayed CSIT Feedback for Finite Rate of Innovation Channel Models and Attainable NetDoFs of the MIMO Interference Channel," in *Proc. Wireless Days*, Valencia, Spain, Nov. 2013.
- S. Jafar, "Topological Interference Management Through Index Coding," *IEEE Trans. Info. Theory*, Jan. 2014.
- L. Ruan, V. Lau, and M. Win, "The Feasibility Conditions for Interference Alignment in MIMO Networks," *IEEE Trans. Signal Processing*, Apr. 2013.
- M. Razaviyayn, L. Gennady, and Z. Luo, "On the Degrees of Freedom Achievable Through Interference Alignment in a MIMO Interference Channel," *IEEE Trans. Signal Processing*, Feb. 2012.
- X. Rao, L. Ruan, and V. Lau, "CSI Feedback Reduction for MIMO Interference Alignment," *IEEE Trans. Signal Processing*, Sep. 2013.
- P. de Kerret and D. Gesbert, "Interference Alignment with Incomplete CSIT Sharing," *IEEE Trans. Wireless Comm.*, 2013.
- Y. Lejosne, M. Bashar, D. Slock, and Y. Yuan-Wu, "Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels," in *Proc. ISCCSP*, Athens, Greece, May 2014.
- M. Bashar, Y. Lejosne, D. Slock, and Y. Yuan-Wu, "MIMO Broadcast Channels with Gaussian CSIT and Application to Location based CSIT," in *Proc. ITA*, San Diego, CA, USA, Feb. 2014.

A Few More References

F. Negro, I. Ghauri, and D.T.M. Slock,"Sum Rate Maximization in the Noisy MIMO Interfering Broadcast Channel with Partial CSIT via the Expected Weighted MSE," in *IEEE Int'l Symp. on Wireless Commun. Systems (ISWCS)*, Paris, France, Aug. 2012.

W. Tabikh, D. Slock, and Y. Yuan-Wu, "The Pathwise MIMO Interfering Broadcast Channel," in *Proc. IEEE Workshop on Signal Proc. Advances in Wireless Comm's (SPAWC)*, 2015.

Y. Lejosne, A. B. Nasser, D. Slock, and Y. Yuan-Wu, "Multi-Cell Multi-User MIMO Downlink with Partial CSIT and Decentralized Design," in *IEEE Broadband Wireless Access workshop (BWA) at GLOBECOM*, Austin, 2014.

P. Elia and D. Slock, "MIMO Broadcast and Interference Channels Towards 5G: Feedback, Performance and Topological Considerations," tutorial at *IEEE SPAWC*, 2015.