Feedback and Cooperation in Interference-Limited Networks

Newcom# Summer School, Sophia Antipolis, France

David Gesbert, EURECOM, gesbert@eurecom.fr

May 28, 2013

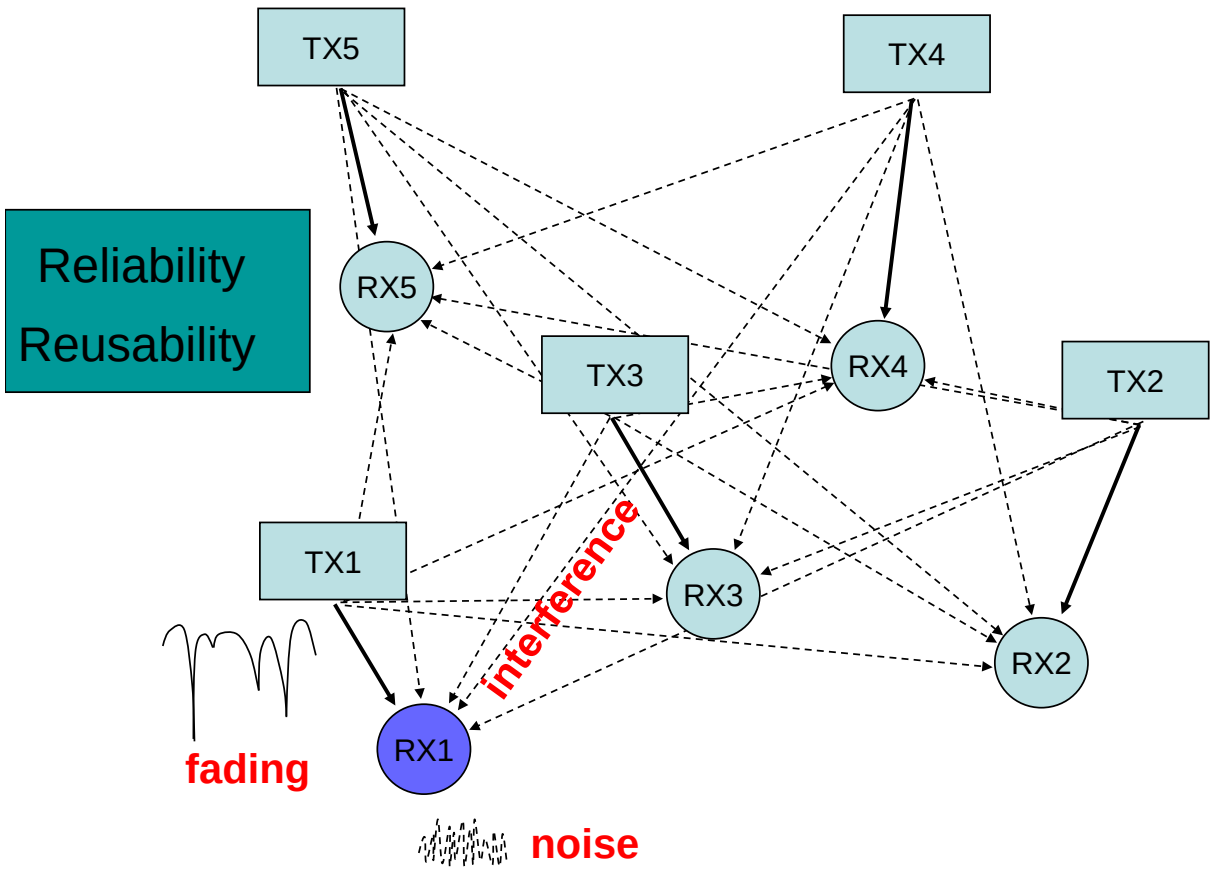
EURECOM
Sophia Antipolis

Acknowledgements

Special thanks to my students and collaborators!

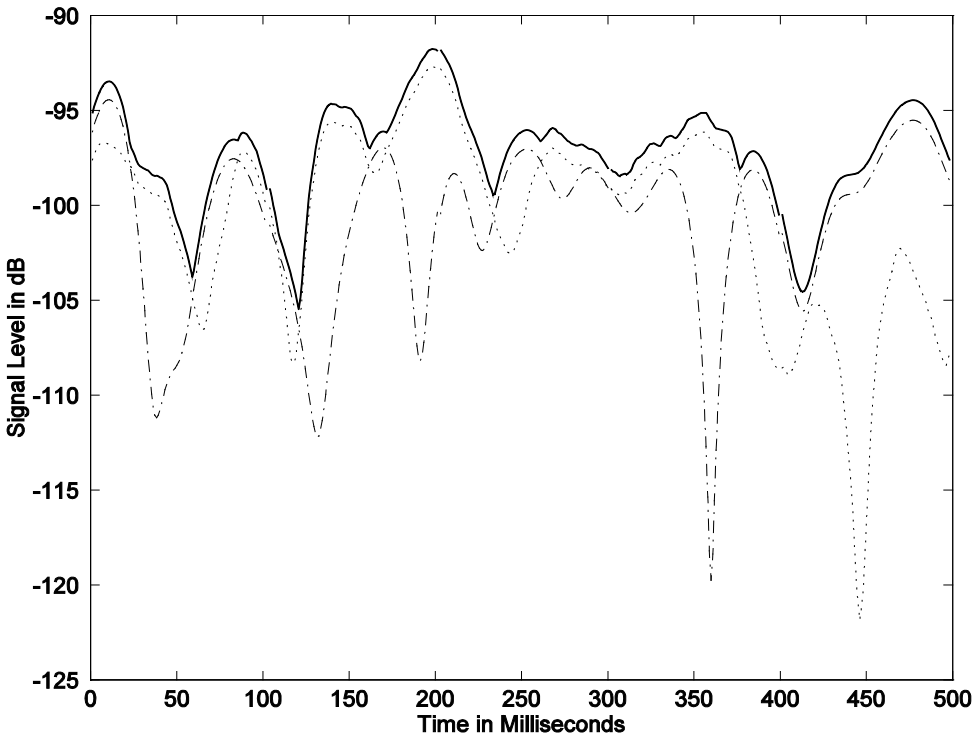
- Paul de Kerret
- Xinping Yi
- Haifan Yin
- Miltiades Filippou
- Randa Zakhour
- Marios Kountouris
- Mari Kobayashi
- Sheng Yang

WHY AREN'T WIRELESS COMMUNICATIONS BETTER ?



CERTAIN PROBLEMS CAN BE TURNED AROUND...

Exploiting fading using multi-user diversity scheduling



What about interference?

Goals of the tutorial

- Highlight fundamental limitation posed by interference, regardless of wireless scenario (cellular, ad-hoc, cognitive)
- Emphasize special role played by spatial processing
- Present unique features as well as commonalities behind methods
- Bring out connections between information theory and system viewpoint
- Raise awareness for crucial role of feedback and information exchange
- Give a sense of where standards are going
- Point out recent results and open exciting research topics!

Do interrupt me!

Shameless (but useful) self-references

- Multi-cell MIMO cooperative networks : A new look at interference Gesbert, David; Hanly, Stephen; Huang, Howard; Shamai, Shlomo; Simeone, Osvaldo; Yu, Wei IEEE Journal on Selected Areas in Communications, December 2010
- CSI Sharing Strategies for transmitter cooperation in wireless networks, P. de Kerret, D. Gesbert, in IEEE Wireless Communications Magazine, Feb. 2013.

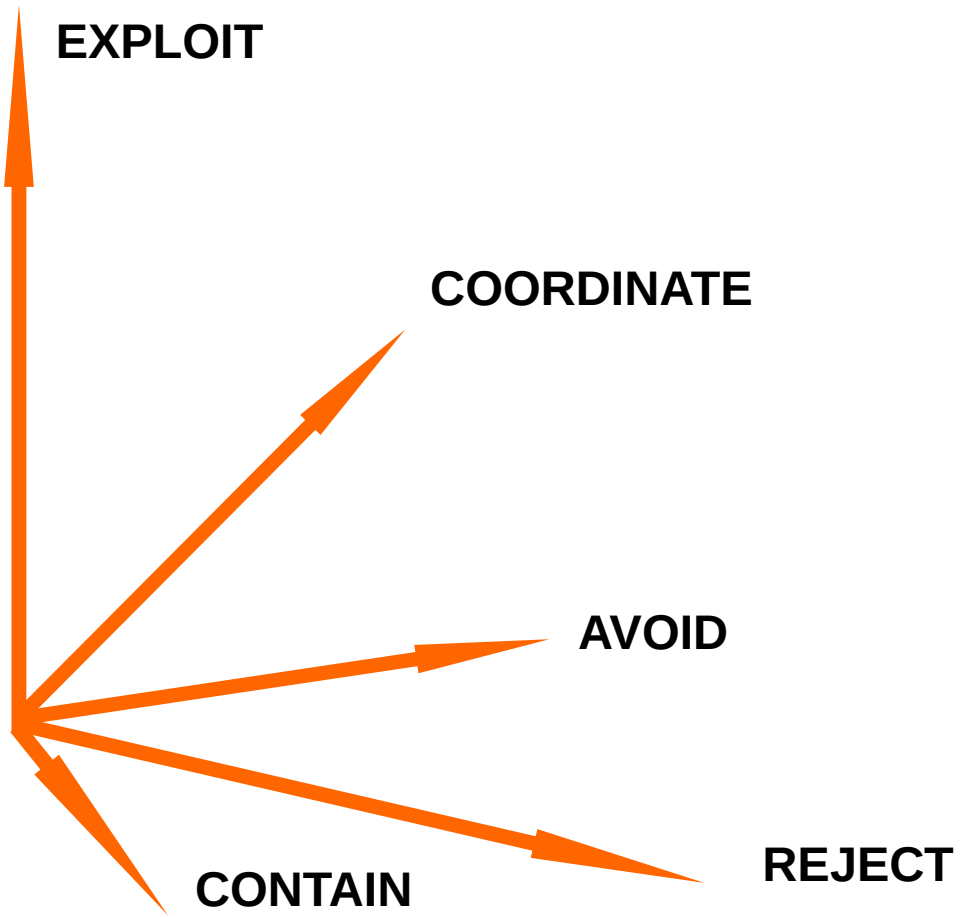
OUTLINE

- Part 1: Key approaches to interference control
- Part 2: Dealing with delayed feedback
- Part 3: Transmitter cooperation with limited information sharing
- Part 4: Distributed cooperation: Using the high-dimensional case

THE ESSENCE OF THIS TALK

Part 1: Key approaches to interference control

THE DIMENSIONS OF INTERFERENCE MANAGEMENT



Interference Containment

- When RX can sustain a given level of interference without impact on QoS (e.g. voice communications, underlay cognitive radios)
- Just enough interference is leaked from transmitter
- Excess interference is "avoided".

Interference Rejection

- RX cancels out interference
- Two approaches:
  - Direct rejection based on beamforming
  - Interference decoding followed by subtraction "SIC" (requires knowledge of codebook and modulations of interferer)

Interference avoidance

- When RX has little or no interference rejection capability
- Interference is avoided from TX side
- Key concept: *Orthogonalizing* transmit resources

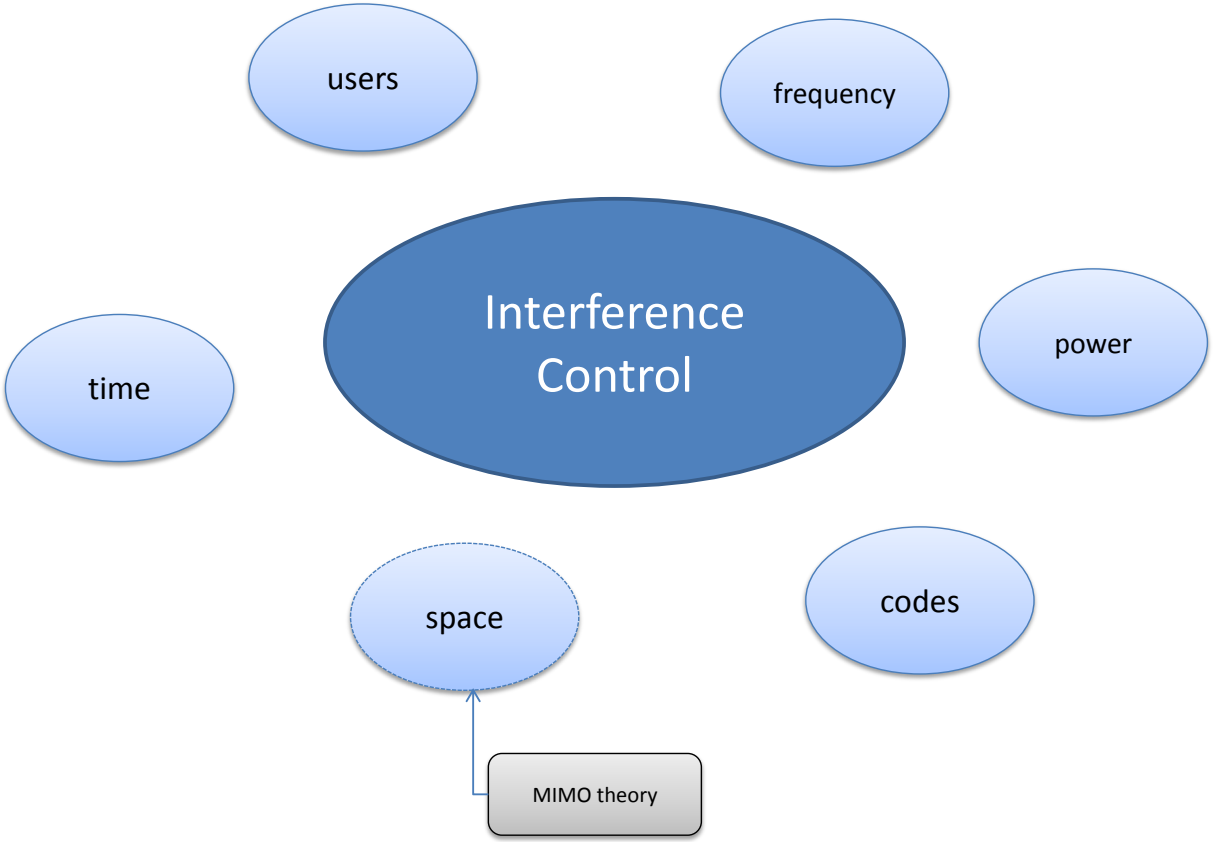
Interference coordination (shaping)

- Shaping from transmitter side (power control, time or frequency assignement, beamforming)
- TX does not avoid interference but facilitates receiver based rejection
- Optimizing the interference distribution to minimize impact.
- Example: Interference Alignement

Interference exploitation

- Interfering transmitter receives prior information about useful data messages
- Interfering transmitter is exploited to contribute to useful transmission (similar to relay)
- Requires tight CSIT and synchronization control
- Example: Network MIMO

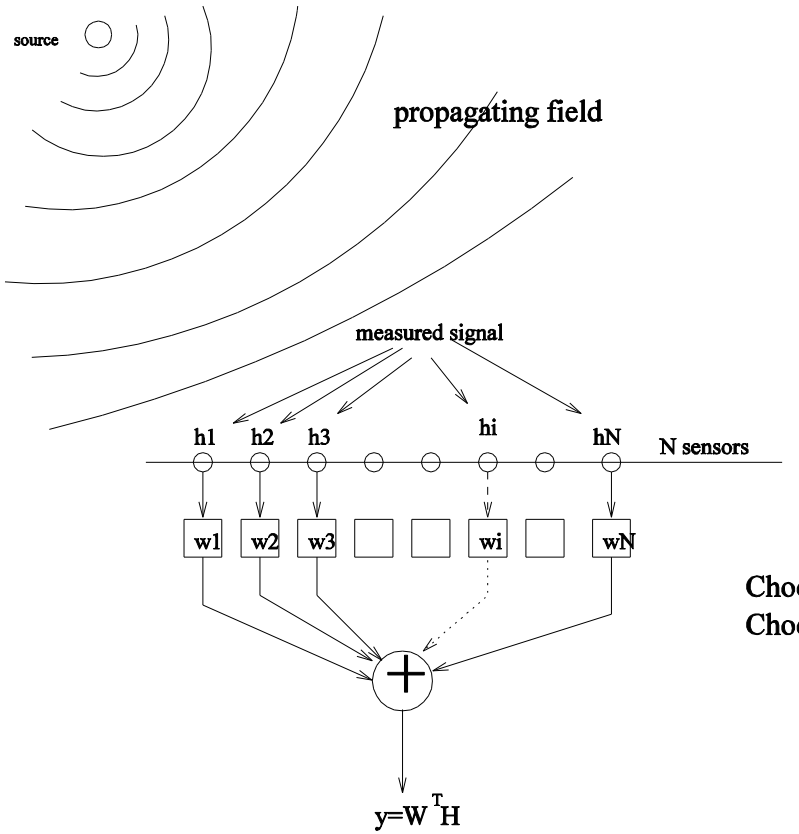
Interference control domains



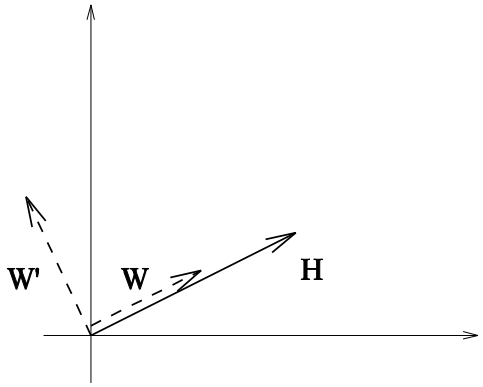
The special role of the spatial dimension

- Many leading concepts for interference management rely on spatial dimension
- Main ideas for multicell case are captured by the multi-user MIMO single cell setting.
- Let us briefly review key principles.

BEAMFORMING AND INTERFERENCE CANCELING



Vector space analogy (for two sensors)

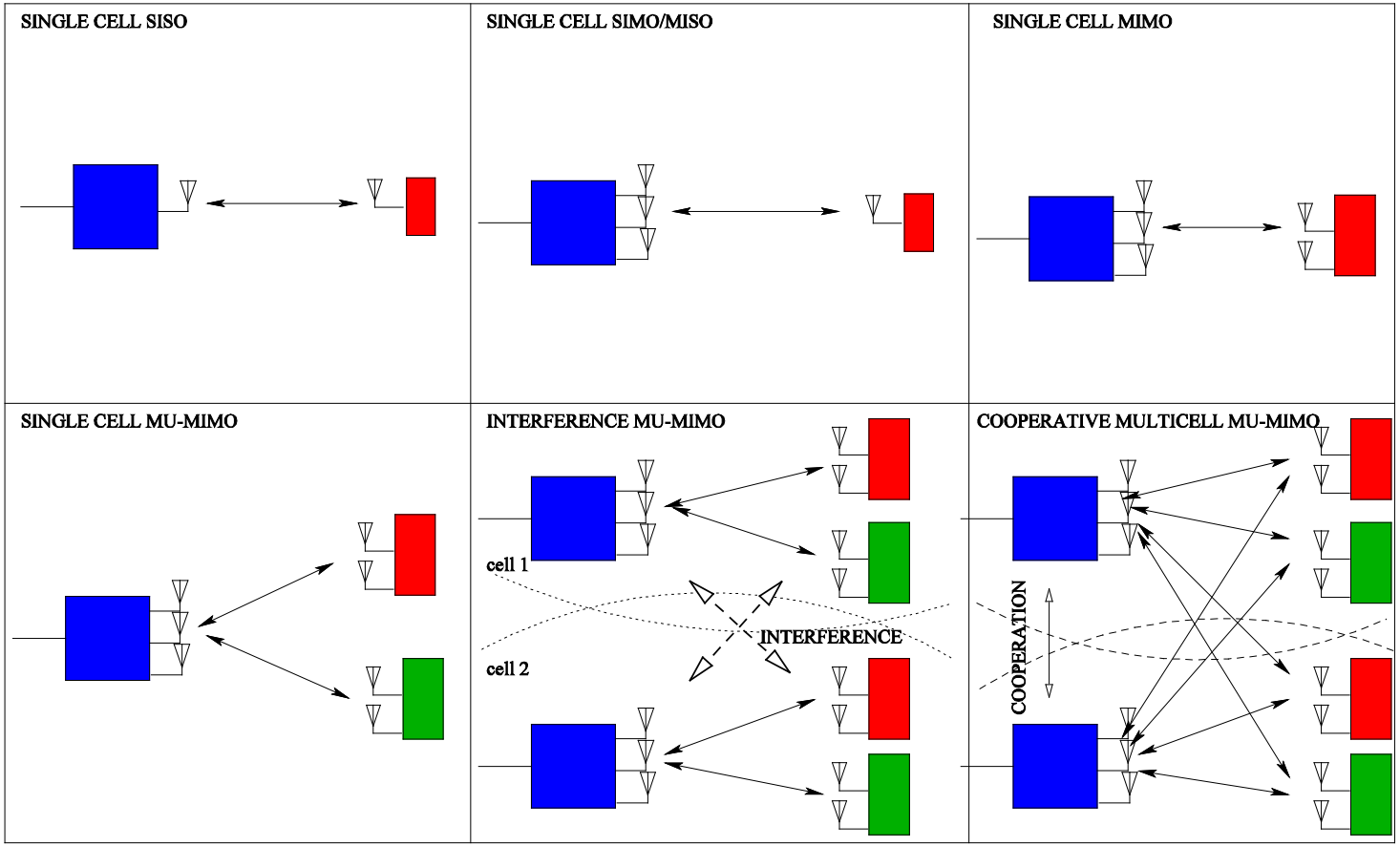


Choosing W enhances the source (beamforming)
Choosing W' nulls the source out (interference nulling)

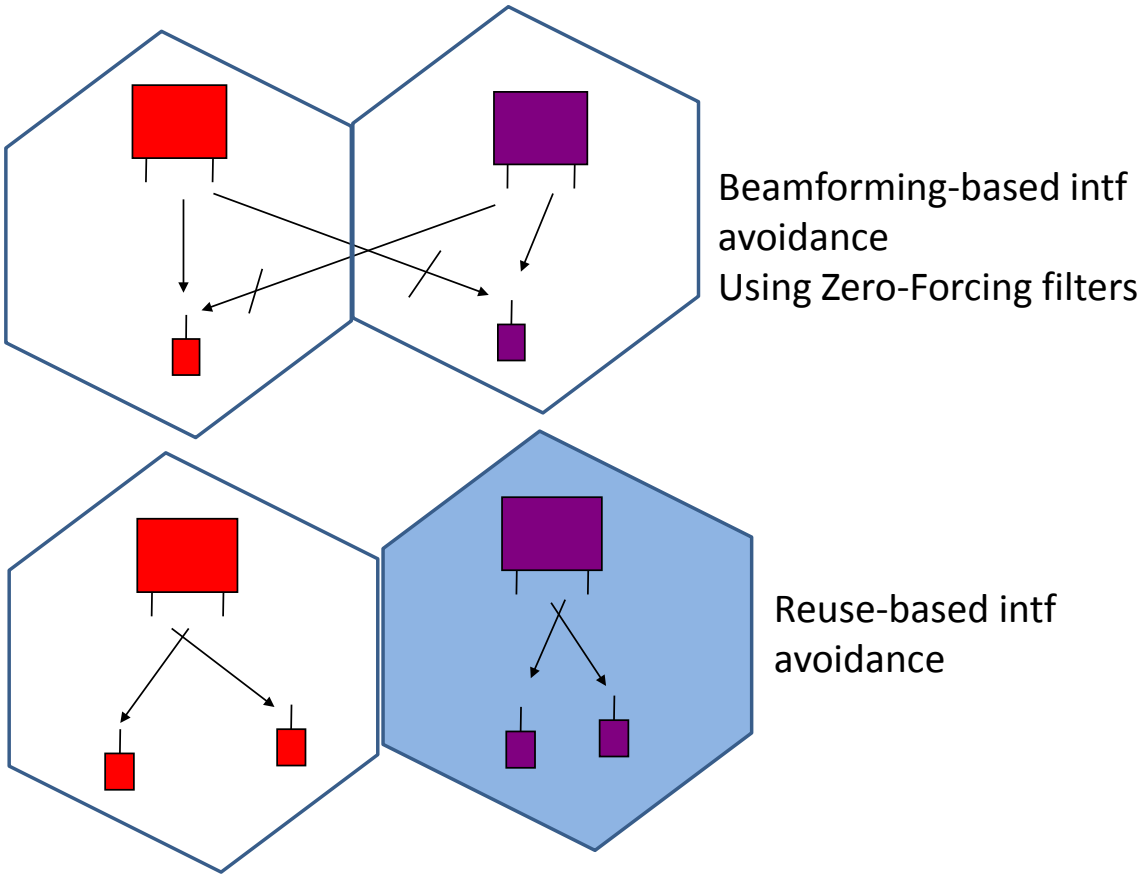
KEY LIMITS OF BEAMFORMING

- A N-antenna beamformer can amplify one source by factor N in average SNR.
- A N-antenna beamformer can extract one source and cancel out $N-1$ interferers simultaneously.
- N sources can be simultaneously extracted (assuming the other $N-1$ are viewed as interferers) by beamforming superposition.
- TX beamforming realizes same benefits as RX beamforming assuming CSIT.
- It does not matter that TXers are co-located or not (assuming they can "talk")

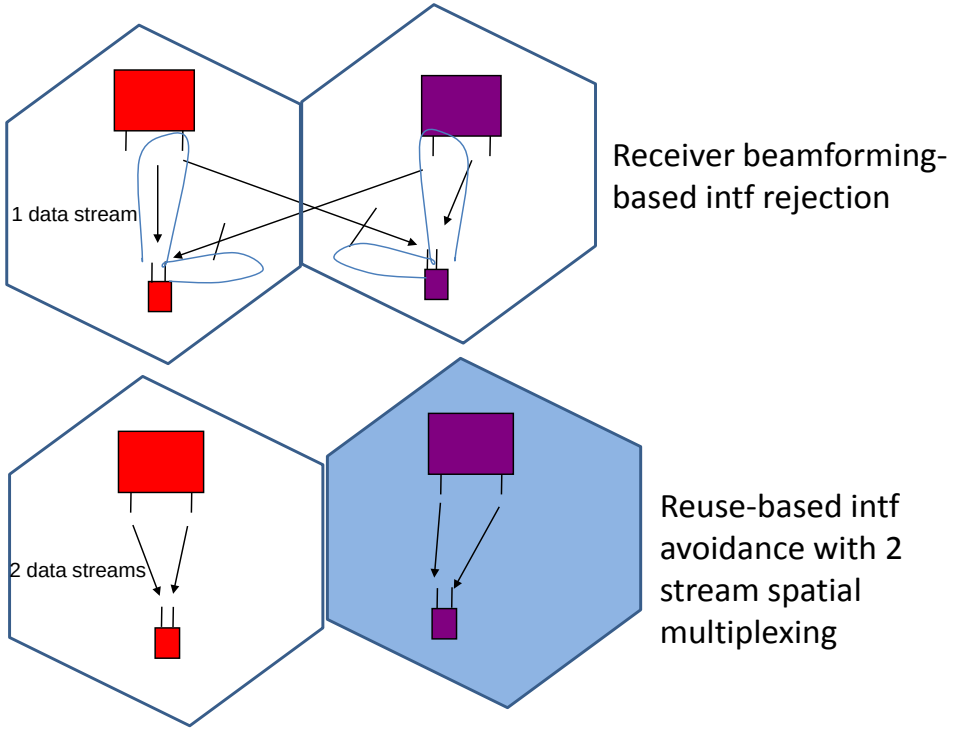
MIMO Configurations



Exploiting spatial DoF wisely: reuse vs. beamforming



Exploiting spatial DoF wisely: The inefficiency of rejection

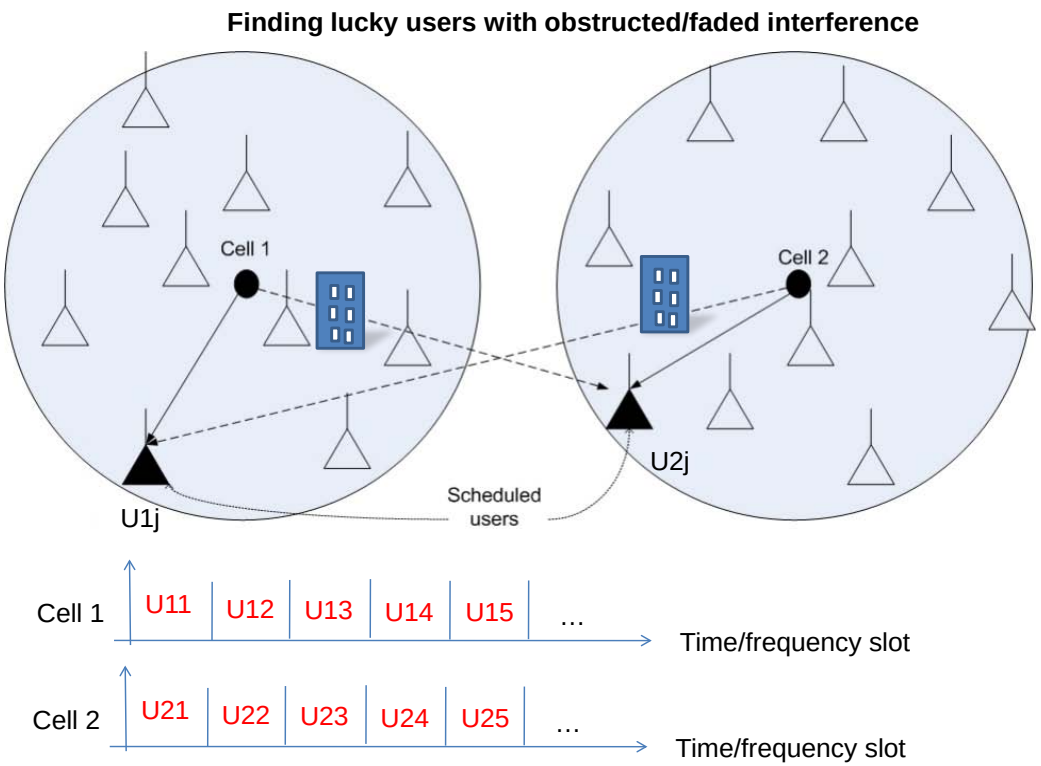


- Avoidance and rejection are not free: consume receiver's degrees of freedom
- Are there better ways to handle interference?

Interference coordination

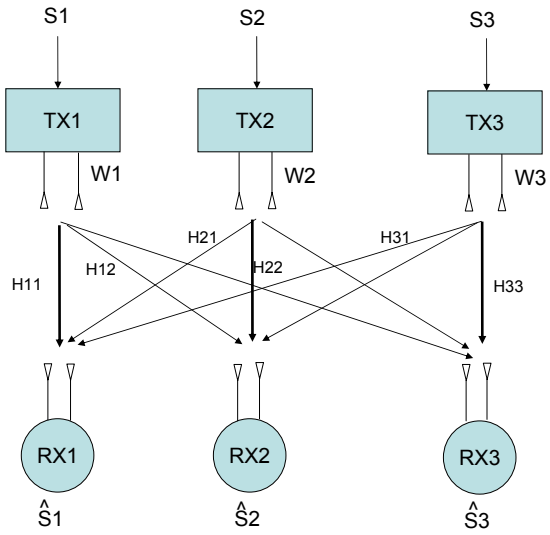
- Softer approach to reducing interference
- Non zero interference is leaked from the transmitter
- Choice of transmission parameter(s) is coordinated across devices
  - transmit power
  - time slot
  - subcarrier
  - beam
  - user
  - ...

Coordination using resource allocation



- Exploits the variability (fading) of interference
- Power control/beamforming couples the decisions at all cells

Coordination using multiple antenna: Alignement



Interference Alignement Conditions with Nr=2:

$$H_{21}W_2 \propto H_{31}W_3$$
$$H_{12}W_1 \propto H_{32}W_3$$
$$H_{13}W_1 \propto H_{23}W_2$$

Alignement can be carried out in space, frequency, time domains. A optimal MG of $1/2$ can be achieved (everyone gets half the cake) *[Maddah-Ali, Motahari, Khandani, Trans IT 2008] [Cadambe, Jafar, Trans IT 2008]*

Interference Alignement: Algorithm design

Exploiting uplink downlink duality of alignement [Gomadam et al., 08]

1. Let U_i be the receiver beamforming vector at user i.
2. Let I_i be the total noise summed at RX i, with covariance Q_i.
3. Take U_i as minimum eigenvector of Q_i, $\forall i$.
4. Use U_i as transmit beamforming vector from user i.
5. Take W_i as RX vector at base i, on reciprocal channel.
6. Find W_i as minimum eigenvector of noise covariance matrix at base i.
7. Back to step 2 and iterate.

Interference Alignement: Finite SNR

- IA is optimal at infinite SNR case only
- At finite SNR, key is to balance desired signal enhancement with interference canceling*
  - max SINR
  - MMSE
  - max sum rate
  - Game theoretic approach (Altruism vs. Egoism)

* [Gomadam et al., 08] [Tresh et al. 09] [Peters et al. 09][Ho et al. 10] (many more)

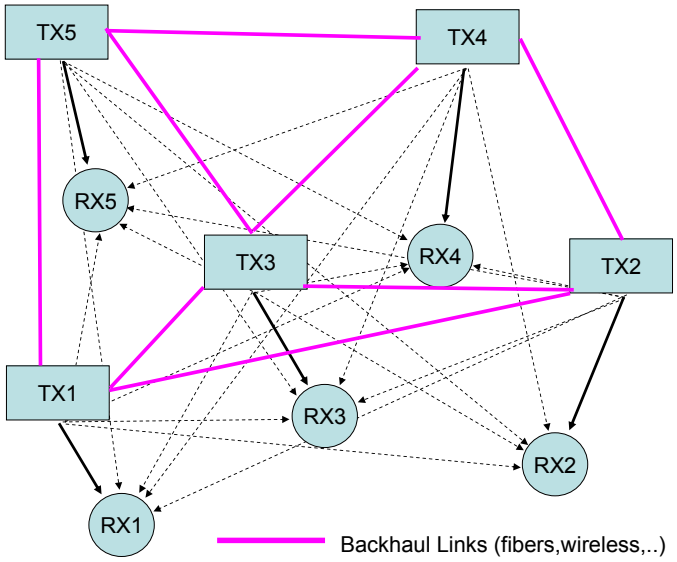
Interference exploitation: Learning from scaling laws

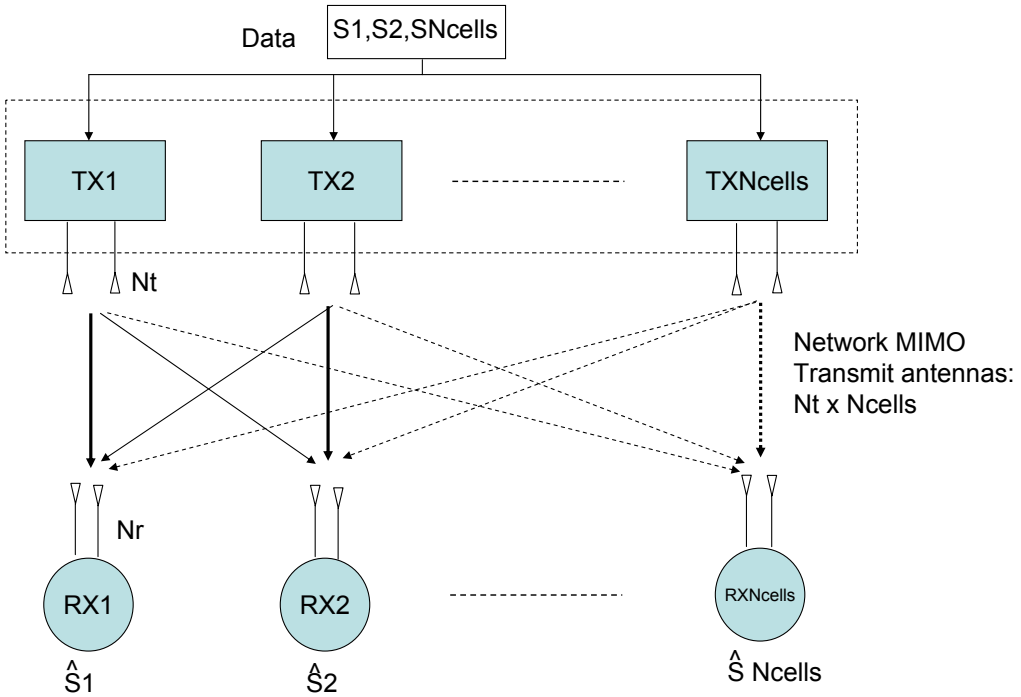
Let R_{DPC} be the sum rate obtained with optimum CSIT-based downlink precoding for K users, M antennas at user side, N total antennas across all base stations:
It is found that [Hassibi05]:

$$\lim_{K\to\infty} \frac{E(R_{DPC})}{N \log\log(MK)} = 1 \tag{1}$$

Interpretation: With large K, the base stations can select and spatially multiplex the N best users out of K with negligible interference loss.
How to realize this in practice?

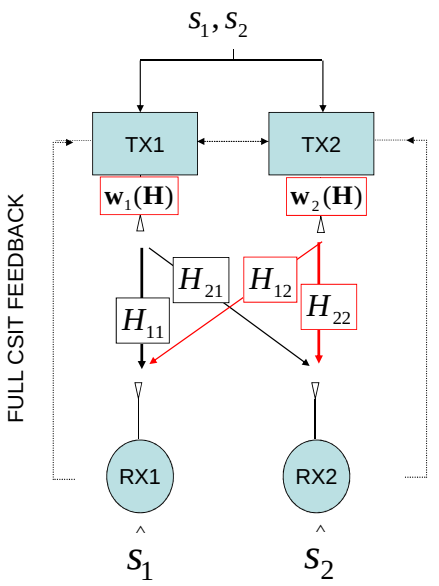
MULTI-CELL (NETWORK) MIMO





[Hanly et al 1993, Shamai et al. 2001, ...]

How does it work?



Ideal Network MIMO

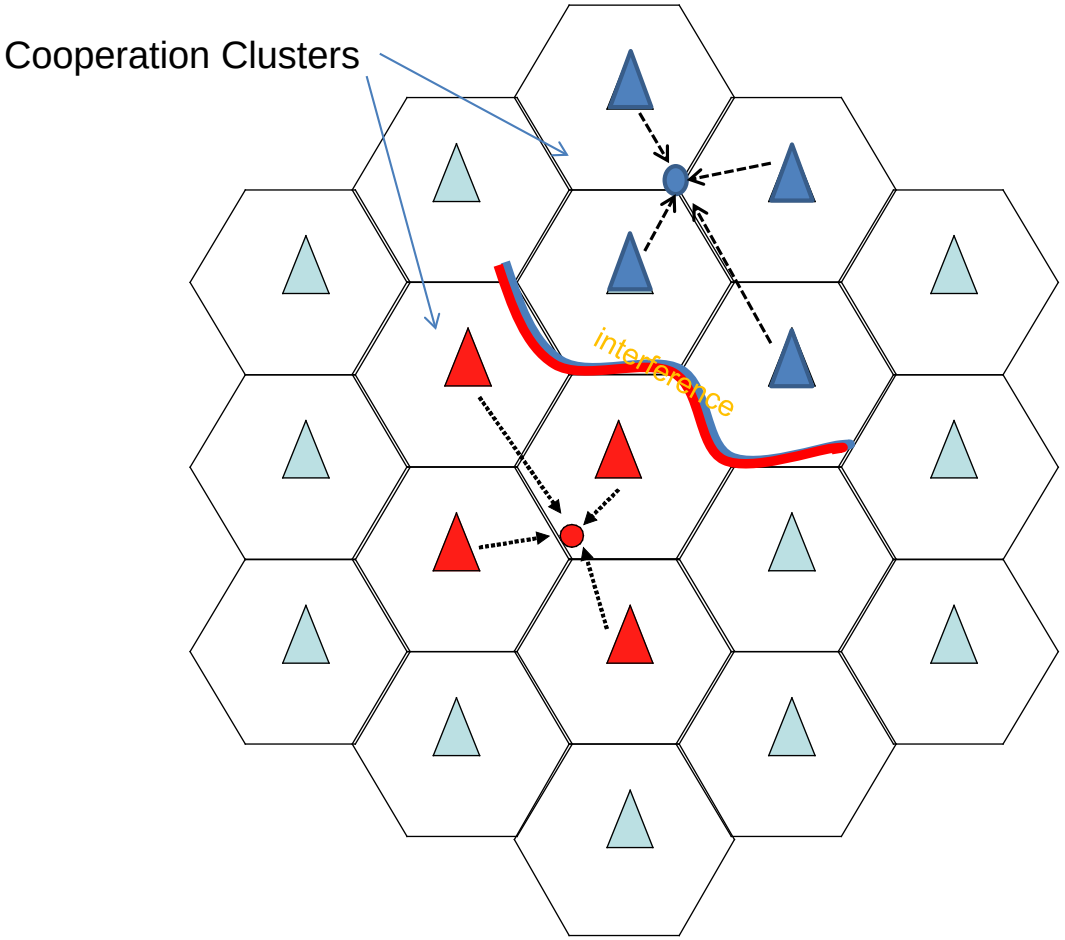
Transmit Matrix:

$$\mathbf{T} = \begin{bmatrix} \mathbf{w}_1(\mathbf{H}) \\ \mathbf{w}_2(\mathbf{H}) \end{bmatrix}$$

$$\text{e.g.}: \mathbf{T} = \frac{\mathbf{H}^{-1}}{\| \mathbf{H}^{-1} \|_F}$$

Modify standard MU-MIMO schemes to reflect per base power constraint (ZF, MMSE, non-linear precoding: Dirty Paper Coding, vector perturbation, ..)

MULTI-CELL MIMO WITH CLUSTERING



Channel feedback in spatial interference control

- No feedback
- Quantized feedback
- Noisy analog feedback
- Delayed feedback

Quantized feedback in MU-MIMO and Network MIMO

Assume a total of M antennas across all cooperating BS, single antenna users

No feedback

no CSIT, then multiplexing gain (MG) $\rightarrow$ 1

Quantized feedback

Theorem [Jindal2006] *Assume Random Vector Quantization with B bits used to encode the channel of one user. Then* $B \geq \alpha(M-1)\log(\mathrm{SNR})$ *is necessary to achieve MG of* $M\alpha$.

Crucial assumption: The quantized feedback is ideally shared across all transmit antennas.

CSIT in Interference Alignement

Current IA schemes are based on

- CSIT feedback followed exchange across links
- Pilot based CSI estimation + TX-RX iterations

Quantized feedback in Interference Alignement

Similar results as broadcast channels apply (feedback bits must grow with $MN \log SNR$, to achieve maximum MG) [Thukral et al. 09]

Possible to improve over this feedback rate by exploiting rotational invariance [Mohsen and Guillaud ITW 2012]

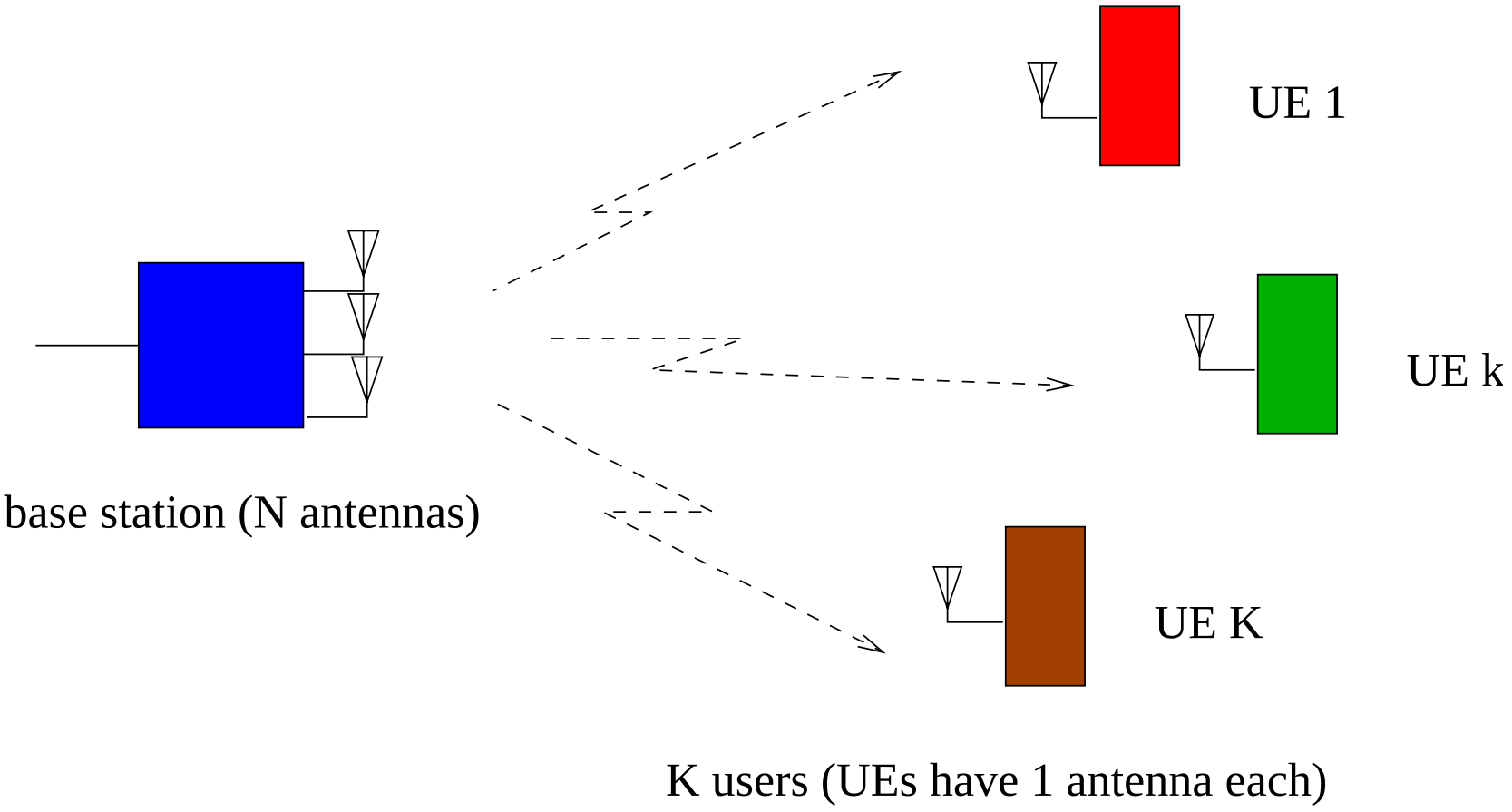
Part 2: Dealing with delayed feedback

Delayed feedback in MU-MIMO

- Feeddback is (maybe) perfect but arrives with delay
- Can one exploits delayed CSI when delay $>$ channel coherence time ? $\rightarrow$ [Maddah Ali et al. 2011]
- Can one exploit delayed CSI when delay $<$ channel coherence time ? $\rightarrow$ introduced in [1]

[1] *M. Kobayashi, S. Yang, D. Gesbert, X. Yi, "On the Degrees of Freedom of time correlated MISO broadcast channel with delayed CSIT" in Proc. IEEE Intern. Symposium on Information Theory, 2012*

THE MULTIPLE ANTENNA (MISO) BROADCAST CHANNEL



Degrees of Freedom (Dof)$= \lim_{P\to\infty} R/\log P$
$=$ nbr of interference-free streams at high SNR

Consider a 2-user BC

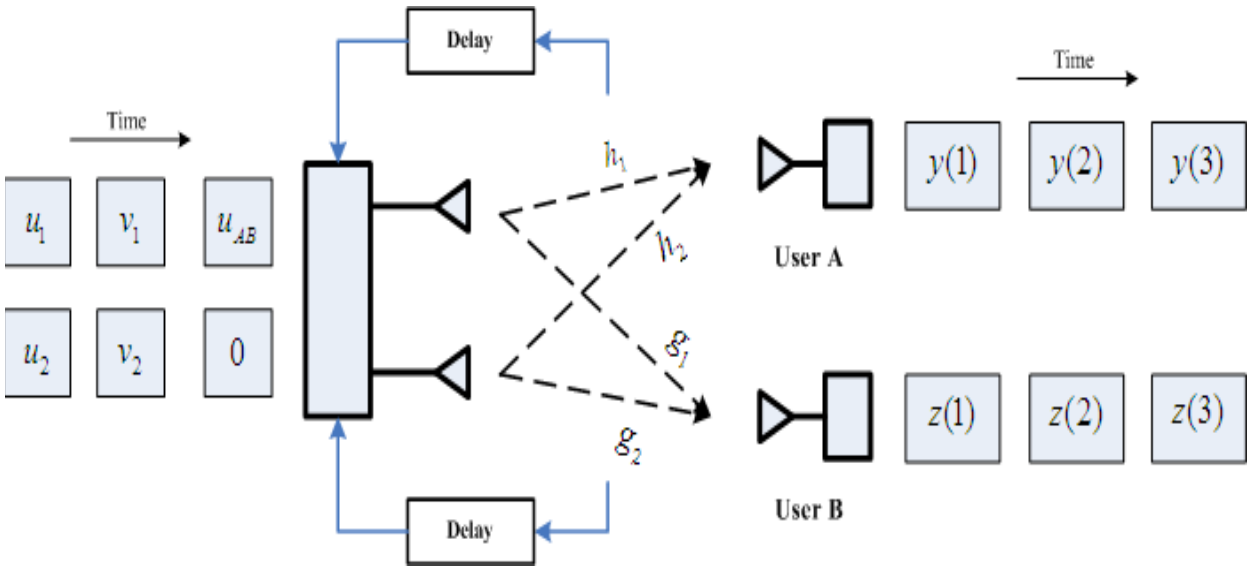
Current CSIT available:

- DoF per user of 1 with perfect CSIT.
- for a CSIT error in $\sigma^2 \sim P^{-1}$, the full DoF is achievable with simple zero-forcing [Caire et al. 2010].

Current CSIT NOT available:

- DoF of $\frac{1}{2}$ with no CSIT whatsoever
- Assume delayed (perfect) CSIT is available at base and other user
  - A surprising result (Maddah-ALi Tse 2010): DoF is $\frac{2}{3} > \frac{1}{2}$
  - Result applies no matter how outdated CSIT is!

MU-MISO PRECODING WITH DELAYED CSIT



At time t, transmitter gets $\mathbf{h}(t-1)$ How to precode when coherence time > 1?

Completely Stale Channel Feedback is still Useful [Maddah Ali et al (MAT) 2010]

Tx

- Slot-1: $\mathbf{x}(1) = \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- Slot-2: $\mathbf{x}(2) = \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- Slot-3: $\mathbf{x}(3) = \begin{bmatrix} u_{AB} \\ 0 \end{bmatrix}$

where
$u_{AB} = \boldsymbol{g}^T(1)\boldsymbol{u} + \boldsymbol{h}^T(2)\boldsymbol{v}$

Rxs

$$y(1) = \boldsymbol{h}^T(1)\boldsymbol{u} + e(1)$$
$$z(1) = \boldsymbol{g}^T(1)\boldsymbol{u} + b(1)$$

$$y(2) = \boldsymbol{h}^T(2)\boldsymbol{v} + e(2)$$
$$z(2) = \boldsymbol{g}^T(2)\boldsymbol{v} + b(2)$$

$$y(3) = h_1(3)u_{AB} + e(3)$$
$$z(3) = g_1(3)u_{AB} + b(3)$$

Completely Stale Channel Feedback is still Useful [Maddah Ali et al (MAT) 2010]

Tx

- Slot-1: $\mathbf{x}(1) = \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- Slot-2: $\mathbf{x}(2) = \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- Slot-3: $\mathbf{x}(3) = \begin{bmatrix} u_{AB} \\ 0 \end{bmatrix}$

  where
  $u_{AB} = \boldsymbol{g}^T(1)\boldsymbol{u} + \boldsymbol{h}^T(2)\boldsymbol{v}$

Rxs

$$y(1) = \boldsymbol{h}^T(1)\boldsymbol{u} + e(1)$$
$$z(1) = \boldsymbol{g}^T(1)\boldsymbol{u} + b(1)$$

$$y(2) = \boldsymbol{h}^T(2)\boldsymbol{v} + e(2)$$
$$z(2) = \boldsymbol{g}^T(2)\boldsymbol{v} + b(2)$$

$$y(3) = h_1(3)u_{AB} + e(3)$$
$$z(3) = g_1(3)u_{AB} + b(3)$$

Completely Stale Channel Feedback is still Useful [Maddah Ali et al (MAT) 2010]

Tx

- Slot-1: $\mathbf{x}(1) = \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- Slot-2: $\mathbf{x}(2) = \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- Slot-3: $\mathbf{x}(3) = \begin{bmatrix} u_{AB} \\ 0 \end{bmatrix}$

  where
  $u_{AB} = \boldsymbol{g}^T(1)\boldsymbol{u} + \boldsymbol{h}^T(2)\boldsymbol{v}$

Rxs

$$y(1) = \boldsymbol{h}^T(1)\boldsymbol{u} + e(1)$$
$$z(1) = \boldsymbol{g}^T(1)\boldsymbol{u} + b(1)$$

$$y(2) = \boldsymbol{h}^T(2)\boldsymbol{v} + e(2)$$
$$z(2) = \boldsymbol{g}^T(2)\boldsymbol{v} + b(2)$$

$$y(3) = h_1(3)u_{AB} + e(3)$$
$$z(3) = g_1(3)u_{AB} + b(3)$$

Completely Stale Channel Feedback is still Useful [Maddah Ali et al (MAT) 2010]

Tx

- Slot-1: $\mathbf{x}(1) = \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- Slot-2: $\mathbf{x}(2) = \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- Slot-3: $\mathbf{x}(3) = \begin{bmatrix} u_{AB} \\ 0 \end{bmatrix}$

  where
  $u_{AB} = \boldsymbol{g}^T(1)\boldsymbol{u} + \boldsymbol{h}^T(2)\boldsymbol{v}$

Rxs

$$y(1) = \boldsymbol{h}^{\mathrm{T}}(1)\boldsymbol{u} + e(1)$$
$$z(1) = \boldsymbol{g}^{\mathrm{T}}(1)\boldsymbol{u} + b(1)$$

$$y(2) = \boldsymbol{h}^{\mathrm{T}}(2)\boldsymbol{v} + e(2)$$
$$z(2) = \boldsymbol{g}^{\mathrm{T}}(2)\boldsymbol{v} + b(2)$$

$$y(3) = h_1(3)u_{AB} + e(3)$$
$$z(3) = g_1(3)u_{AB} + b(3)$$

Completely Stale Channel Feedback is still Useful [Maddah Ali et al (MAT) 2010]

Tx

- Slot-1: $\mathbf{x}(1) = \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- Slot-2: $\mathbf{x}(2) = \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- Slot-3: $\mathbf{x}(3) = \begin{bmatrix} u_{AB} \\ 0 \end{bmatrix}$

  where
  $u_{AB} = \boldsymbol{g}^T(1)\boldsymbol{u} + \boldsymbol{h}^T(2)\boldsymbol{v}$

Rxs

$$y(1) = \boldsymbol{h}^{\mathsf{T}}(1)\boldsymbol{u} + e(1)$$
$$z(1) = \boldsymbol{g}^{\mathsf{T}}(1)\boldsymbol{u} + b(1)$$

$$y(2) = \boldsymbol{h}^{\mathsf{T}}(2)\boldsymbol{v} + e(2)$$
$$z(2) = \boldsymbol{g}^{\mathsf{T}}(2)\boldsymbol{v} + b(2)$$

$$y(3) = h_1(3)u_{AB} + e(3)$$
$$z(3) = g_1(3)u_{AB} + b(3)$$

Space-time Interference Alignment (MAT Alignment)

$$\begin{bmatrix} y(1) \\ y(2) \\ y(3) \end{bmatrix} = \underbrace{\begin{bmatrix} h_1(1) & h_2(1) \\ 0 & 0 \\ h_1(3)g_1(1) & h_1(3)g_2(1) \end{bmatrix}}_{rank=2} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 \\ h_1(2) & h_2(2) \\ h_1(3)h_1(2) & h_1(3)h_2(2) \end{bmatrix}}_{rank=1} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Important unanswered questions

MAT scheme is both optimistic and pessimistic

- MAT assumes infinite SNR.
  - What happens at finite SNR?
- MAT scheme based on pessimistic delay assumption
  - What happens for CSIT delay $<$ coherence time ?

Can we do better?

A finite SNR approach: Generalized MAT

- TS1: $\mathbf{x}_1 = \mathbf{u}$
- TS2: $\mathbf{x}_2 = \mathbf{v}$
- TS3: $\mathbf{x}_3 = \begin{bmatrix} l_{uv} \\ 0 \end{bmatrix}$ where l_{uv} intoduces past channel-aware precoders

$$l_{uv} = \mathbf{w}_1^T \mathbf{u} + \mathbf{w}_2^T \mathbf{v}$$

Signal model (at user 1)

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}_u \mathbf{u} + \bar{\mathbf{H}}_v \mathbf{v} + \mathbf{m},$$

where

$$\bar{\mathbf{H}}_u = \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{0} \\ h_{1,3}\mathbf{w}_1^T \end{bmatrix}, \ \bar{\mathbf{H}}_v = \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_2^T \\ h_{1,3}\mathbf{w}_2^T \end{bmatrix},$$

Particularly for MAT algorithm

$$\mathbf{w}_1 = \mathbf{g}_1, \quad \mathbf{w}_2 = \mathbf{h}_2.$$

GMAT Precoder Design

Due to $\mathbf{h}_i(3)$ being unknown, consider virtual signals:

$$\begin{aligned}\mathbf{y} &= \mathbf{H}_u\mathbf{u} + \mathbf{H}_v\mathbf{v} + \mathbf{m}\\ \mathbf{z} &= \mathbf{G}_u\mathbf{u} + \mathbf{G}_v\mathbf{v} + \mathbf{m}\end{aligned} \tag{2}$$

where

$$\mathbf{H}_u = \begin{bmatrix}\mathbf{h}_1^T\\ \mathbf{0}\\ \mathbf{w}_1^T\end{bmatrix}, \ \mathbf{H}_v = \begin{bmatrix}\mathbf{0}\\ \mathbf{h}_2^T\\ \mathbf{w}_2^T\end{bmatrix}$$

$$\mathbf{G}_u = \begin{bmatrix}\mathbf{g}_1^T\\ \mathbf{0}\\ \mathbf{w}_1^T\end{bmatrix}, \ \mathbf{G}_v = \begin{bmatrix}\mathbf{0}\\ \mathbf{g}_2^T\\ \mathbf{w}_2^T\end{bmatrix}$$

Goal:
$\mathbf{w}_i = f_i(\mathbf{h}_1, \mathbf{g}_1, \mathbf{h}_2, \mathbf{g}_2)$ trades-off alignment for signal orthogonality

GMAT Precoder Design — Optimization

- Sum mutual information in high SNR region

$$I(\mathbf{u};\mathbf{y}) + I(\mathbf{v};\mathbf{z}) \approx \log\left(\frac{\mathbf{w}_1^H\mathbf{R}_1\mathbf{w}_1}{\mathbf{w}_2^H\mathbf{R}_2\mathbf{w}_2}\frac{\mathbf{w}_2^H\mathbf{Q}_2\mathbf{w}_2}{\mathbf{w}_1^H\mathbf{Q}_1\mathbf{w}_1}\right) + \log C$$

where

$$\begin{aligned}
\mathbf{R}_1 &= C_1\left(\mathbf{I} + \rho\mathbf{h}_1^{\perp}\mathbf{h}_1^{\perp H}\right)\\
\mathbf{R}_2 &= C_2\left(\gamma_1\mathbf{I} + \rho\mathbf{h}_2^{\perp}\mathbf{h}_2^{\perp H}\right)\\
\mathbf{Q}_1 &= C_3\left(\gamma_2\mathbf{I} + \rho\mathbf{g}_1^{\perp}\mathbf{g}_1^{\perp H}\right)\\
\mathbf{Q}_2 &= C_4\left(\mathbf{I} + \rho\mathbf{g}_2^{\perp}\mathbf{g}_2^{\perp H}\right)
\end{aligned}$$

GMAT Precoder Design — Optimization

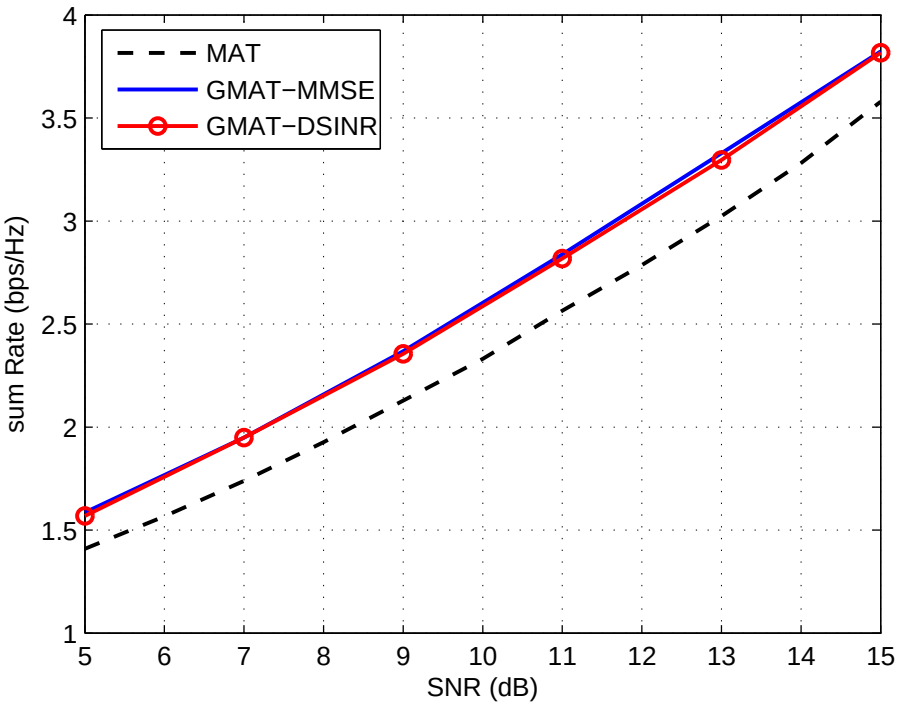
The two precoders are found separately from generalized eigenvector problems

$$\max_{\|\mathbf{w}_1\|^2=1} \frac{\mathbf{w}_1^H \mathbf{R}_1 \mathbf{w}_1}{\mathbf{w}_1^H \mathbf{Q}_1 \mathbf{w}_1}$$

$$\max_{\|\mathbf{w}_2\|^2=1} \frac{\mathbf{w}_2^H \mathbf{Q}_2 \mathbf{w}_2}{\mathbf{w}_2^H \mathbf{R}_2 \mathbf{w}_2}$$

X. Yi, D. Gesbert "Precoding Methods for MISO Broadcast Channel with Delayed CSIT", to appear in IEEE Trans. Wireless 2013.

Simulation Results



Exploiting time-correlation of delayed feedback

Channel prediction at Txs according to the delayed feedback from the Rxs

$$\boldsymbol{h}_t = \hat{\boldsymbol{h}}_t + \tilde{\boldsymbol{h}}_t$$

where $\sigma^2 \triangleq \mathrm{E}(\|\tilde{\boldsymbol{h}}_t\|^2) \sim P^{-(1-2F)}$ and $2F$ is the normalized bandwidth.
Define

$$\alpha \triangleq -\lim_{P\to\infty} \frac{\log \sigma^2}{\log P} = 1 - 2F$$

reflecting the quality of imperfect current CSIT, i.e.,

- $\alpha = 0$: the channel is not predictable at all, and no current CSIT is available
- $0 < \alpha < 1$: the channel is predictable to a certain extent
- $\alpha = 1$: the channel is totally predictable and perfect in the sense of DoF

Beamforming using imperfect current CSIT

$$\mathbf{x}(1) = \begin{bmatrix} \hat{\boldsymbol{g}}(1) & \hat{\boldsymbol{g}}^{\perp}(1) \end{bmatrix} \boldsymbol{u} + \begin{bmatrix} \hat{\boldsymbol{h}}(1) & \hat{\boldsymbol{h}}^{\perp}(1) \end{bmatrix} \boldsymbol{v}$$

where the overheard interferences

$$\eta_1 = \boldsymbol{h}^{\mathsf{H}}(1)\hat{\boldsymbol{h}}(1)v_1 + \boldsymbol{h}^{\mathsf{H}}(1)\hat{\boldsymbol{h}}^{\perp}(1)v_2$$
$$\eta_2 = \boldsymbol{g}^{\mathsf{H}}(1)\hat{\boldsymbol{g}}(1)u_1 + \boldsymbol{g}^{\mathsf{H}}(1)\hat{\boldsymbol{g}}^{\perp}(1)u_2$$

With suitable power allocation across u_1 and u_2 (v_1 and v_2):

$$\mathsf{E}(|\eta_1|^2) = \mathsf{E}(|\eta_2|^2) \sim P^{1-\alpha}$$

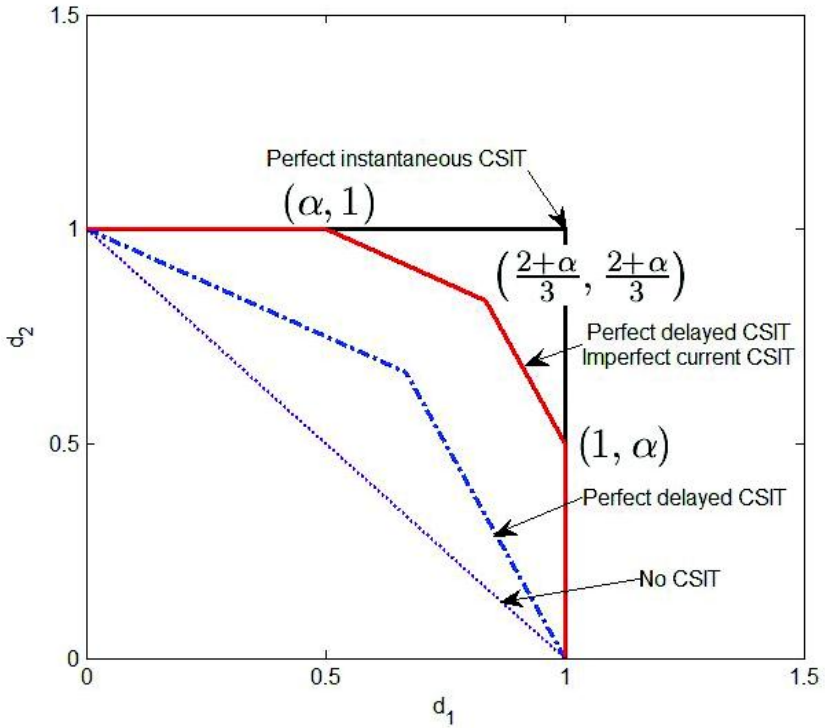
due to $\mathsf{E}(|\boldsymbol{h}^{\mathsf{H}}(t)\hat{\boldsymbol{h}}^{\perp}(t)|^2) = \sigma^2 \sim P^{-\alpha}$.
Key idea: Quantize and forward residual interference over just $(1-\alpha)\log P$ bits

$$\eta_1 \Rightarrow \hat{\eta}_1 \qquad\qquad \eta_2 \Rightarrow \hat{\eta}_2$$

$$\mathbf{x}(2) = \begin{bmatrix} \hat{\eta}_2 \\ 0 \end{bmatrix} \qquad\qquad \mathbf{x}(3) = \begin{bmatrix} \hat{\eta}_1 \\ 0 \end{bmatrix}$$

Time slots 2 and 3 consume $(1-\alpha)$ channel uses only

Optimal DoF region for Time-correlated MISO Broadcast Channel (2 antennas, 2 users)



Key: Add new symbols in slots 2 and 3 to get optimal DoF

S. Yang, M. Kobayashi, D. Gesbert, X. Yi, "On the Degree of Freedom Region of Time Correlated MISO Broadcast Channels with Delayed CSIT", to appear in IEEE Transactions on Information Theory, 2013.

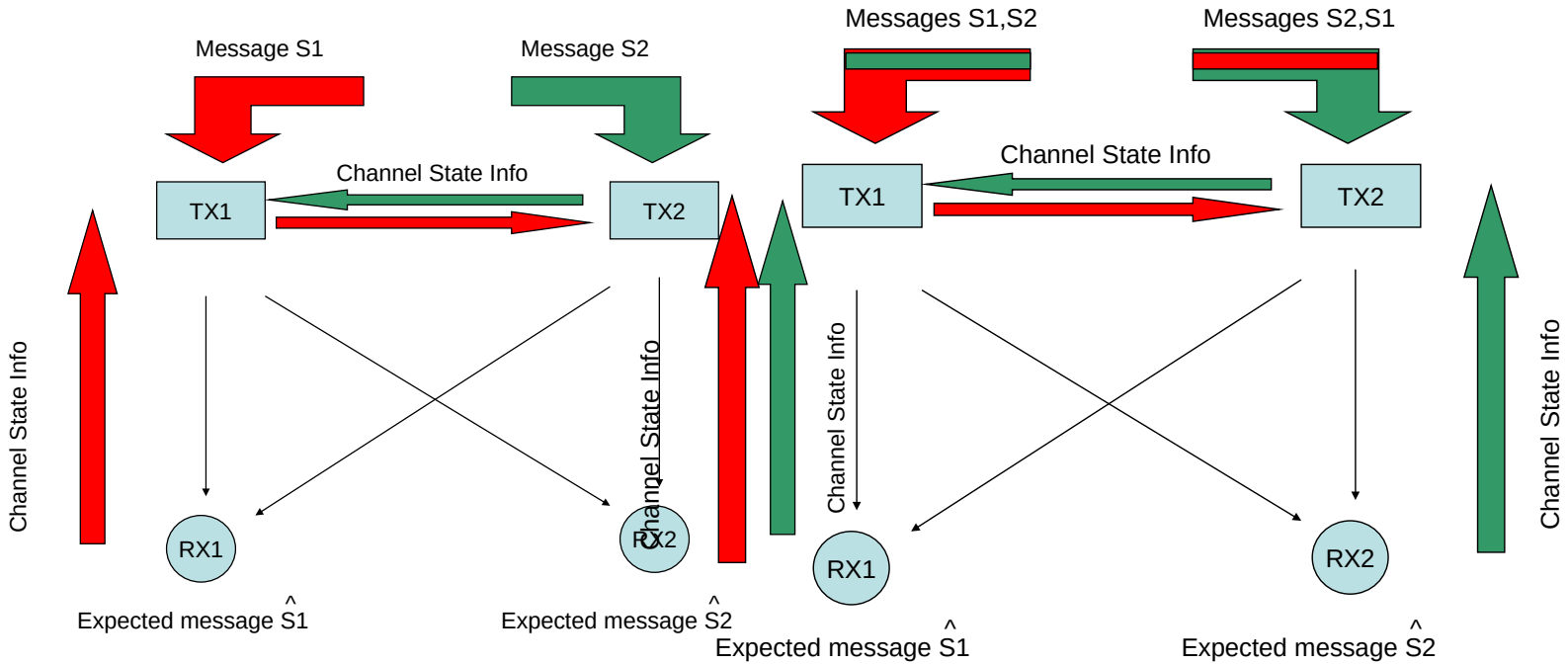
Extension to MIMO and IC cases:

X. Yi, S. Yang, D. Gesbert, M. Kobayashi "The Degree of Freedom Region of Temporally Correlated MIMO Networks with Delayed CSIT", to appear in IEEE Trans. on Information Theory, March. 2013

Part 3: Transmitter cooperation with limited information sharing

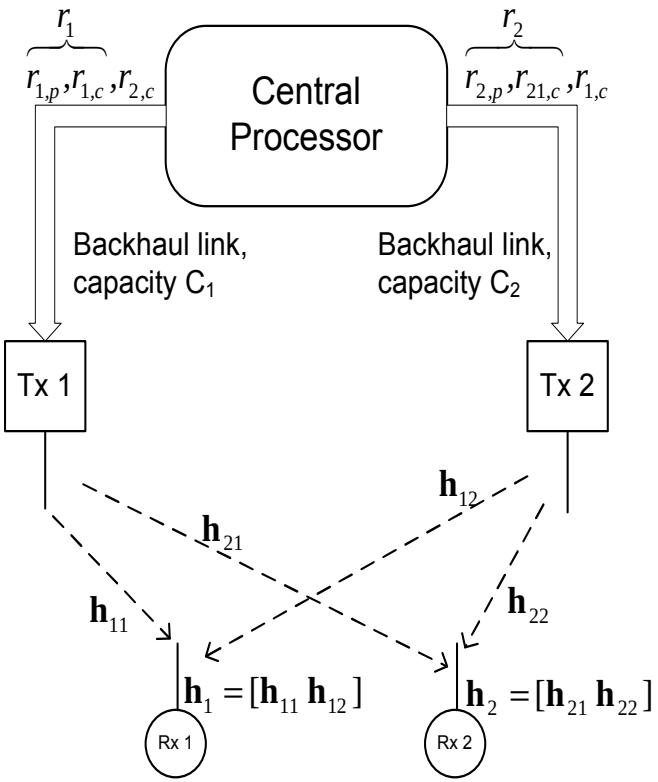
Transmitter cooperation with limited information sharing

- Limited user data sharing
- Limited channel state information (CSI) sharing



Limited user data sharing: bridging the IC and BC channels

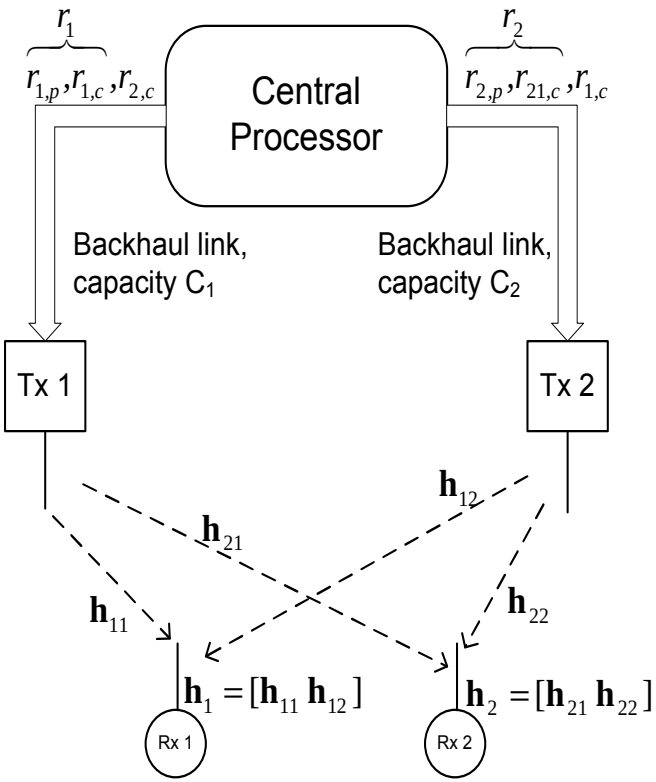
Partial user data sharing with fully shared CSI



- Rate of user i, r_i, is split into two parts:
  - Private message, from Tx i alone, $r_{i,p}$
  - Shared message, from both Txs, $r_{i,c}$.
- Extreme cases:
  - $r_{i,c} = 0 \Rightarrow r_i = r_{i,p}$: IC.
  - $r_{i,p} = 0 \Rightarrow r_i = r_{i,c}$: BC (network MIMO).

Limited user data sharing: bridging the IC and BC channels

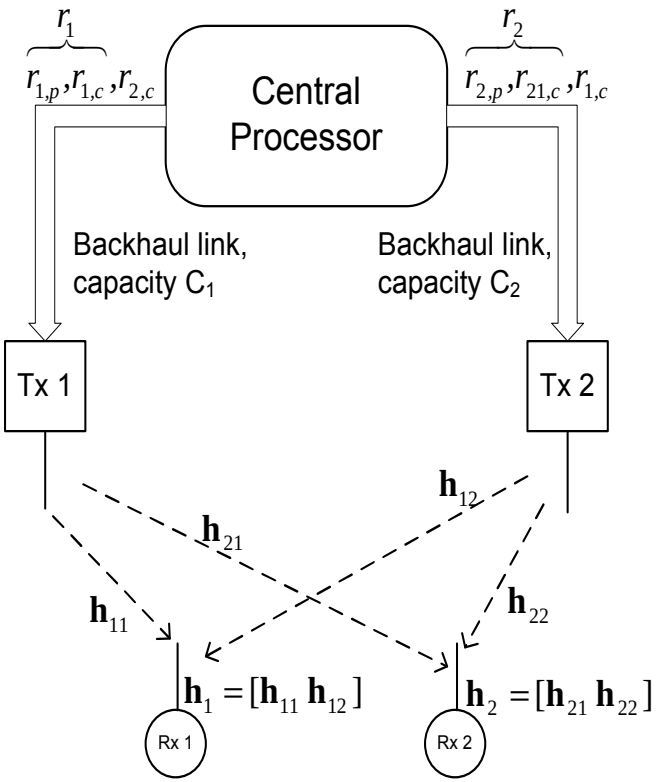
Partial user data sharing with fully shared CSI



- Rate of user i, r_i, is split into two parts:
  - Private message, from Tx i alone, $r_{i,p}$
  - Shared message, from both Txs, $r_{i,c}$.
- Extreme cases:
  - $r_{i,c} = 0 \Rightarrow r_i = r_{i,p}$: IC.
  - $r_{i,p} = 0 \Rightarrow r_i = r_{i,c}$: BC (network MIMO).

Limited user data sharing: bridging the IC and BC channels

Partial user data sharing with fully shared CSI



- Rate of user i, r_i, is split into two parts:
  - Private message, from Tx i alone, $r_{i,p}$
  - Shared message, from both Txs, $r_{i,c}$.
- Extreme cases:
  - $r_{i,c} = 0 \Rightarrow r_i = r_{i,p}$: IC.
  - $r_{i,p} = 0 \Rightarrow r_i = r_{i,c}$: BC (network MIMO).

SIGNAL MODEL

Linear beamforming on private and common messages:

$$\mathbf{x} = \begin{bmatrix} \mathbf{w}_{1,c} & \mathbf{w}_{2,c} \end{bmatrix} \begin{bmatrix} s_{1,c} \\ s_{2,c} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_{1,p} \\ \mathbf{0} \end{bmatrix} s_{1,p} + \begin{bmatrix} \mathbf{0} \\ \mathbf{w}_{2,p} \end{bmatrix} s_{2,p}, \quad (3)$$

Subject to power constraints:

$$\|\mathbf{D}_i \mathbf{w}_{1,c}\|^2 + \|\mathbf{D}_i \mathbf{w}_{2,c}\|^2 + \|\mathbf{w}_{i,p}\|^2 \leq P_i, \quad i = 1, 2. \quad (4)$$

Theorem: The following rates are achievable

$$r_{i,p} \leq \log_2 \left(1 + \frac{\left|\mathbf{h}_{ii}\mathbf{w}_{i,p}\right|^2}{\sigma_i^2} \right),$$

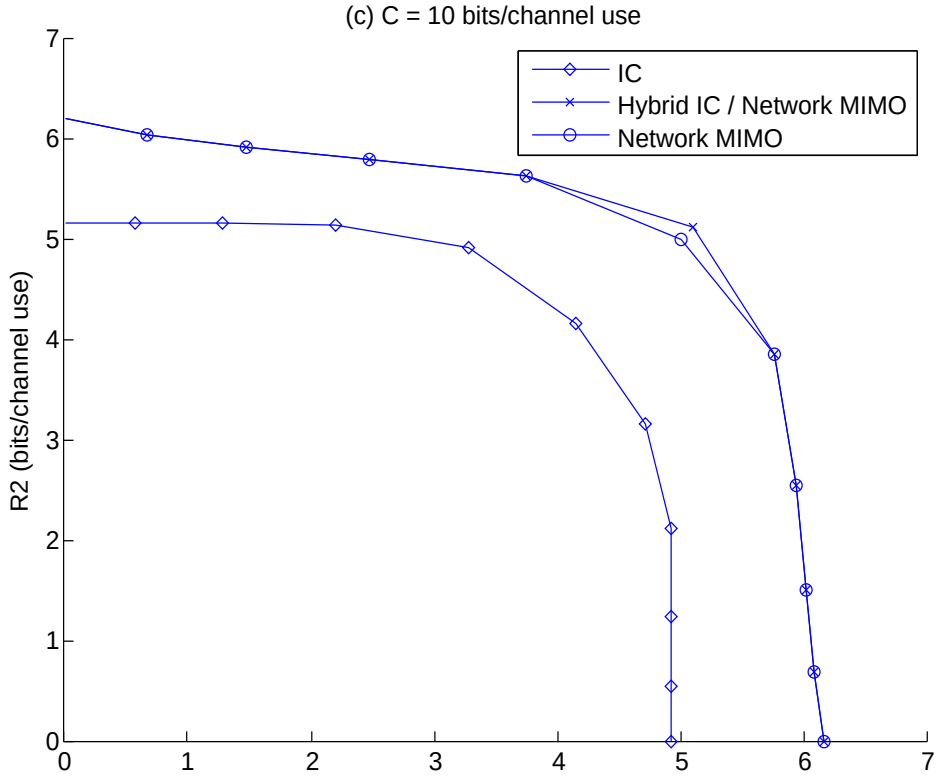
$$r_i = r_{i,p} + r_{i,c} \leq \log_2 \left(1 + \frac{\left|\mathbf{h}_{ii}\mathbf{w}_{i,p}\right|^2 + \left|\mathbf{h}_i\mathbf{w}_{i,c}\right|^2}{\sigma_i^2} \right) \tag{5}$$

where

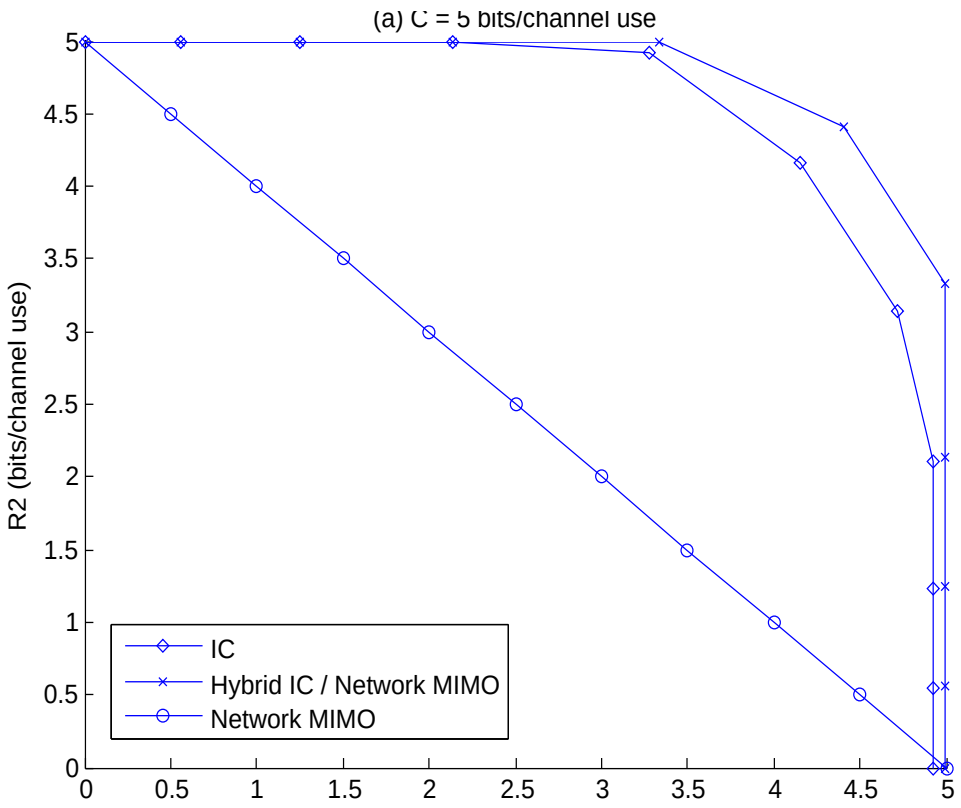
$$\sigma_i^2 = \sigma^2 + \left|\mathbf{h}_{i\bar{i}}\mathbf{w}_{\bar{i},p}\right|^2 + \left|\mathbf{h}_i\mathbf{w}_{\bar{i},c}\right|^2 \tag{6}$$

R. Zakhour, D. Gesbert, "Optimized data sharing in multicell MIMO with finite backhaul capacity", in IEEE Transactions on Signal Processing, Dec. 2011.

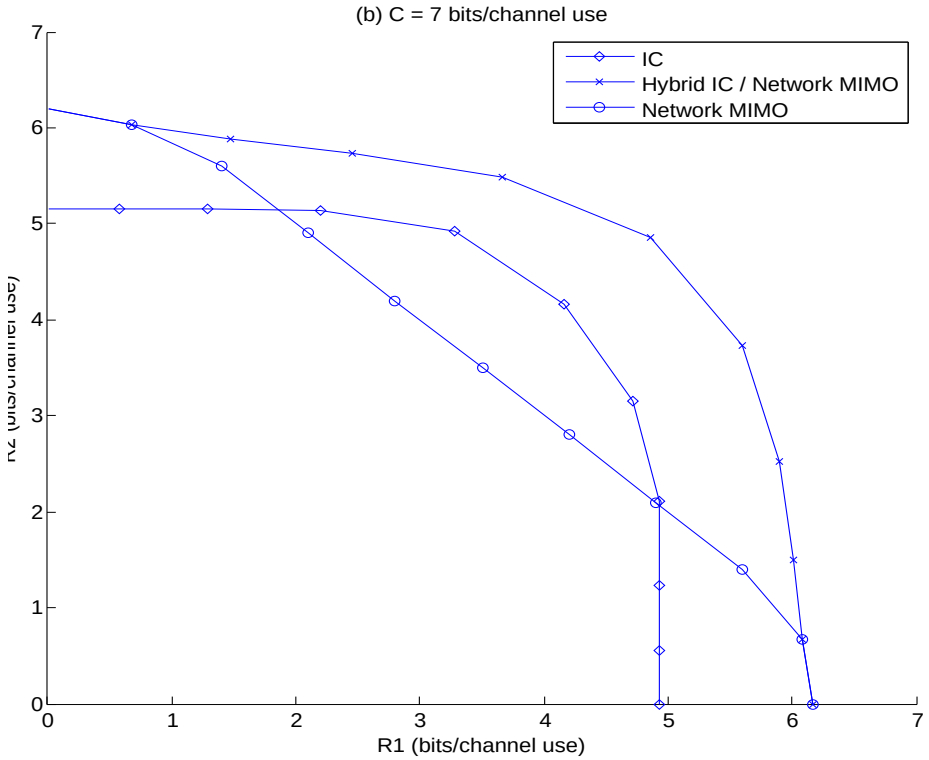
Generous backhaul (C=10 Bits/Sec/Hz)



Restricted backhaul (C=5 Bits/Sec/Hz)



Intermediate backhaul (C=7 Bits/Sec/Hz)



Limited channel state information (CSI) sharing

Key assumptions:

- Mutually interfering transmitters are willing to cooperate
- Each transmitter obtained limited CSI feedback from terminals
- Each transmitter willing to or capable of sharing a limited amount of CSI with others

Learning from team decision theory

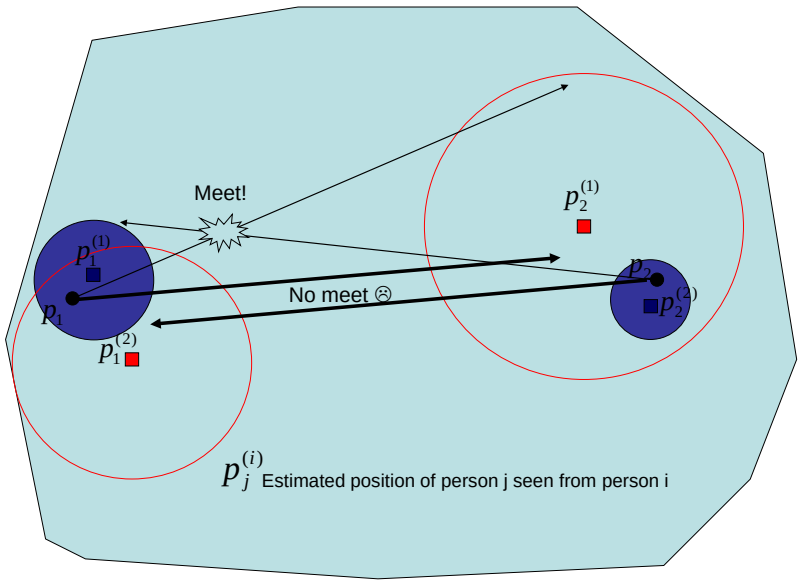
A Team decision theoretic problem[1]:

- Several network agents wish to cooperate towards maximization of a common utility
- Each agent has its own limited view over the system state
- All need to come up with consistent actions
- Classical "robust" design does not work...

[1] Yu-Chi Ho, "Team decision theory and information structures" Proceedings of IEEE, June 1980
R. Zakhour and D. Gesbert, "Team decision for the cooperative MIMO channel with imperfect CSIT sharing", The Information Theory and Applications (ITA) Workshop, San Diego CA., February 2010

THE DISTRIBUTED RENDEZ-VOUS PROBLEM

- Two visitors arrive independently in Nice and seek to meet as quickly as possible.
- They have different and imprecise information about their own and each other's position.
- **Problem**: *Pick a direction to walk into*



Team decisional transmitter cooperation

Let us consider the following K-transmitter framework:

1. Define global system state $\mathbf{H}$ (e.g. multi-user channel matrix)
2. A distributed information structure: each transmitter i has knowledge of $\hat{\mathbf{H}}^{(i)}$, which exhibits *some* arbitrary correlation with $\mathbf{H}$.
3. A decision space for each transmitter i. Example: $\mathbf{w}_i(\hat{\mathbf{H}}^{(i)})$ where $\mathbf{w}_i$ is a complex matrix. (e.g. $\mathbf{w}_i \in \mathbb{C}^{N\times K}$ for K-user network MIMO)
4. A common network utility $u = \sum_{i=1}^{K} u_i(\mathbf{w}_1(\hat{\mathbf{H}}^{(1)}), .., \mathbf{w}_K(\hat{\mathbf{H}}^{(K)}), \mathbf{H})$

CHARACTERIZING THE INFORMATION STRUCTURE

- **Perfect CSIT**: A CSI structure is *perfect* if $\hat{\mathbf{H}}^{(i)} = \mathbf{H}$, $\forall i$.
- **Centralized CSIT**: A CSI structure is *centralized* if $\hat{\mathbf{H}}^{(i)} = \hat{\mathbf{H}}^{(j)}$, $\forall i, j$.
- **Distributed CSIT**: A CSI structure is *distributed* if there exist i and j such that $\hat{\mathbf{H}}^{(i)} \neq \hat{\mathbf{H}}^{(j)}$.
- **Incomplete CSIT**: A CSI structure is *incomplete* if $\hat{\mathbf{H}}^{(i)}$ takes the form $\forall i$ $\hat{\mathbf{H}}^{(i)} = \{\mathbf{h}_{kl}, k \in \mathcal{S}_{tx},\ l \in \mathcal{S}_{rx}\}$, where $\mathcal{S}_{tx}$ (resp. $\mathcal{S}_{rx}$) are subsets of the transmitter set (resp. receiver set).

CHARACTERIZING THE INFORMATION STRUCTURE

The (lack of) correlation among the information structure (IS) elements $\hat{\mathbf{H}}^{(i)}$ determine the degree of distributed-ness of the cooperation. We have the following:

- Limited feedback $\not\Rightarrow$ distributed but distributed $\Rightarrow$ limited feedback somewhere
- (strictly) incomplete $\Rightarrow$ ditributed

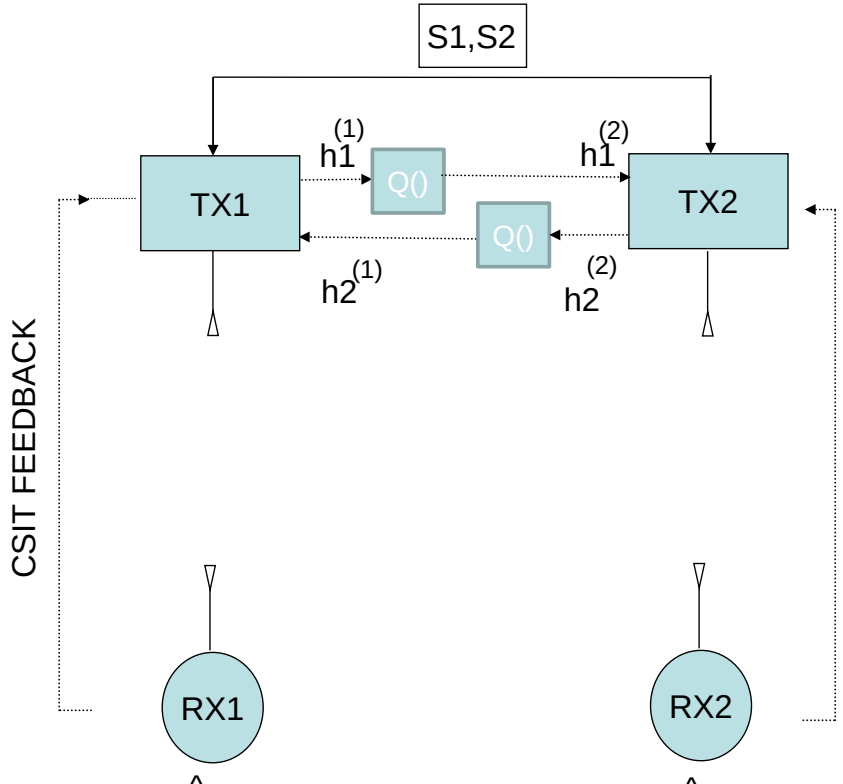
IDEAL DISTRIBUTED INFORMATION STRUCTURES

Consider the K transmitter (N antennas each) K user (single antenna) channel. Let $\mathbf{h}_{ji}$ be the $1 \times N$ vector channel between the i-th transmitter and the j-th user.

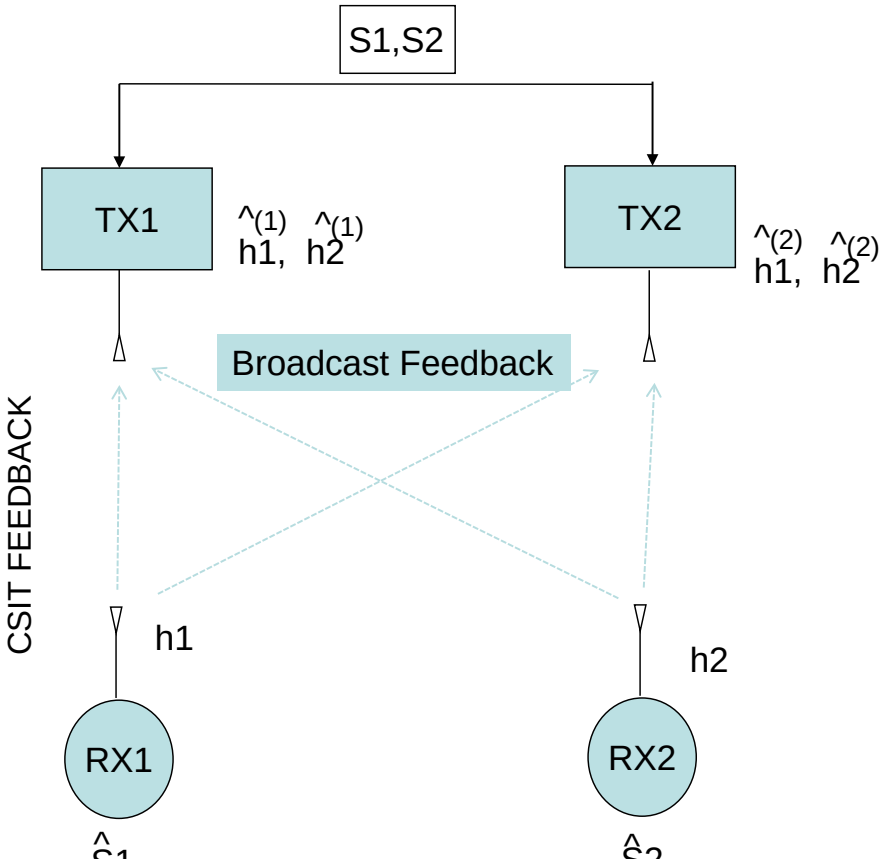
- Local CSIT with TDD reciprocity $\rightarrow \hat{\mathbf{H}}^{(i)} = [\mathbf{h}_{1i}^T, .., \mathbf{h}_{Ki}^T]^T$
- Local CSIT with LTE feedback mode $\rightarrow \hat{\mathbf{H}}^{(i)} = [\mathbf{h}_{i1}, .., \mathbf{h}_{iK}]$
- Fully local (*Fully distributed*) CSIT $\rightarrow \hat{\mathbf{H}}^{(i)} = \mathbf{h}_{ii}$
- Others (imagination is the limit!)

- Over the air feedback introduces thermal noise, quantization noise (unless analog feedback) and delays
- Backhaul signaling introduces delays and possible quantization noise

CSI sharing over quantized backhaul



Broadcast feedback



TEAM DECISION MAKING

Distributed coordination = team decision making = A difficult problem in general! (functional optimization).

$$\max_{\mathbf{w}_i(\hat{\mathbf{H}}^{(i)}), i=1..K} E\left\{\sum_i u_i(\mathbf{w}_1(\hat{\mathbf{H}}^{(1)}), .., \mathbf{w}_K(\hat{\mathbf{H}}^{(K)}), \mathbf{H})\right\} \tag{7}$$

The model-based approach:

- Replace $\mathbf{w}_i(\hat{\mathbf{H}}^{(i)})$ by $\mathbf{f}(\mathbf{a}_i, \hat{\mathbf{H}}^{(i)})$ where $\mathbf{f}(.,.)$ is a functional model and $\mathbf{a}_i$ a vector of deterministic parameters to be determined at transmitter i.
- Solve for (still hard ;-))

$$\max_{\mathbf{a}_i, i=1..K} E\left\{\sum_i u_i(\mathbf{f}(\mathbf{a}_1, \hat{\mathbf{H}}^{(1)}), .., \mathbf{f}(\mathbf{a}_K, \hat{\mathbf{H}}^{(K)}), \mathbf{H})\right\} \tag{8}$$

A DoF perspective on distributed coordination

- Finite-SNR rate maximisation is difficult
- Can we learn from high SNR regime?
  - $\rightarrow$ hint: Use DoF as figure of merit
- Examine different problems at local and global scale

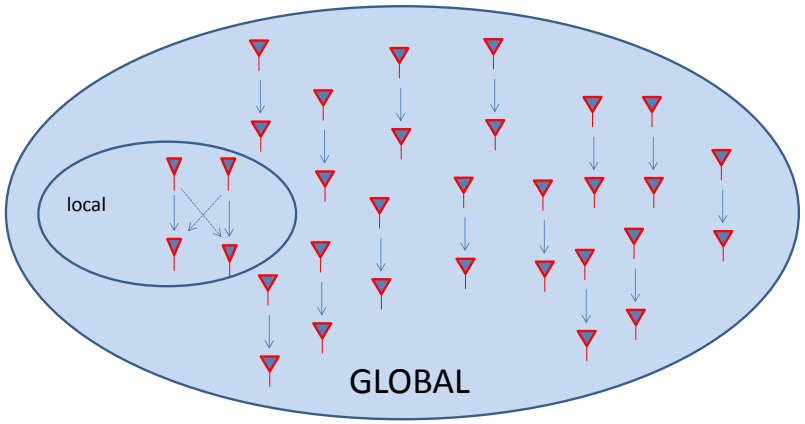
Local vs. global

At the global scale:

- Do all transmitters need to share the same CSIT across the network?
- Can "better" transmitters go away with less CSIT?
- What is an optimal spatial CSIT allocation policy across the network?

At the local scale:

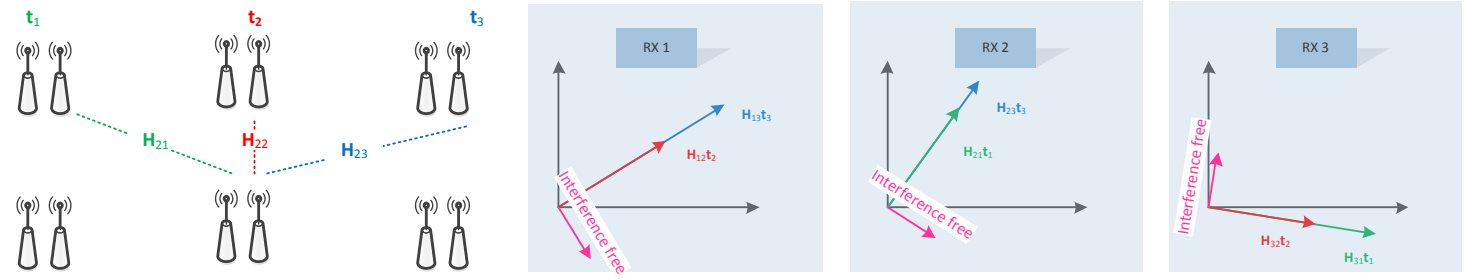
- How bad is coordinated transmission based on different (inconsistent) CSIT?
- Can robust solutions be developped? (e.g. for precoding)

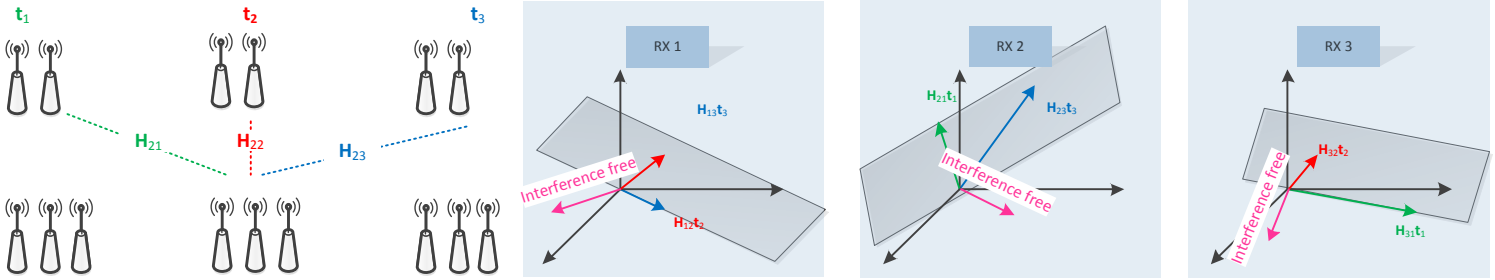


Two examples of results:

1. The interference channel: DoF of interference alignement with *incomplete* CSIT
2. The Network MIMO channel: DoF-preserving reduced spatial CSIT allocation policies

Interference alignment case: Building intution





- Traditional IA feasibility studies ignores CSIT setting (in fact assumes complete CSIT)
- Trade-off between extra antennas and CSIT requirements?
- How incomplete can CSIT be while preserving alignment?

Tightly-Feasible and Super-Feasible Settings

IA feasibility based on counting equations and variables (single stream transmission):

Definition

Tightly-feasible IC $\Leftrightarrow$ feasible IC and $\mathcal{N}_{\text{var}}(\mathcal{K},\mathcal{K}) = \mathcal{N}_{\text{Equality}}(\mathcal{K},\mathcal{K})$

Definition

Super-feasible IC $\Leftrightarrow$ feasible IC and $\mathcal{N}_{\text{var}}(\mathcal{K},\mathcal{K}) > \mathcal{N}_{\text{Equality}}(\mathcal{K},\mathcal{K})$

Remark: $\mathcal{N}_{\text{var}}(\mathcal{K},\mathcal{K}) = \sum_{i=1}^{K} N_i + M_i$ *and* $\mathcal{N}_{\text{Equality}}(\mathcal{K},\mathcal{K}) = K(K+1)$

Tightly-Feasible and Super-Feasible Settings

IA feasibility based on counting equations and variables (single stream transmission):

Definition

Tightly-feasible IC $\Leftrightarrow$ feasible IC and $\mathcal{N}_{\text{var}}(\mathcal{K},\mathcal{K}) = \mathcal{N}_{\text{Equality}}(\mathcal{K},\mathcal{K})$

Definition

Super-feasible IC $\Leftrightarrow$ feasible IC and $\mathcal{N}_{\text{var}}(\mathcal{K},\mathcal{K}) > \mathcal{N}_{\text{Equality}}(\mathcal{K},\mathcal{K})$

Remark: $\mathcal{N}_{\text{var}}(\mathcal{K},\mathcal{K}) = \sum_{i=1}^{K} N_i + M_i$ *and* $\mathcal{N}_{Equality}(\mathcal{K},\mathcal{K}) = K(K+1)$

IA feasibility seen under prism of CSIT

Looking for the smallest CSIT Allocation $\mathcal{A}$

- Counting the size of feedback for a complete CSIT allocation $\mathcal{A}$: $\mathsf{s}\left(\mathcal{A}_{\text{comp}}\right) = K(\sum_{i=1}^{K} N_i)(\sum_{i=1}^{K} M_i)$

Optimization Problem

Find the most incomplete CSIT allocation with IA feasible:

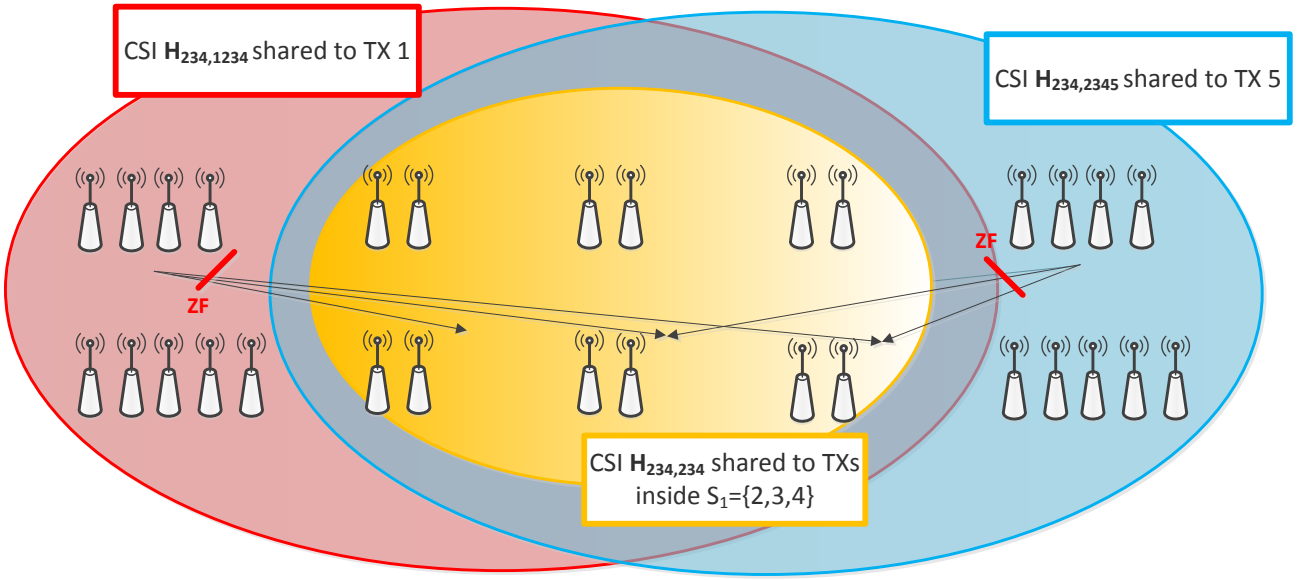
$$\mathcal{A}_{\text{min}} = \underset{\mathcal{A} \in \mathbb{A}_{\text{Feas}}}{\operatorname{argmin}}\, \mathsf{s}(\mathcal{A})$$

where $\mathsf{s}(\mathcal{A})$ is total number of scalar feedbacks.

A RESULT FOR TIGHTLY FEASIBLE SETTINGS

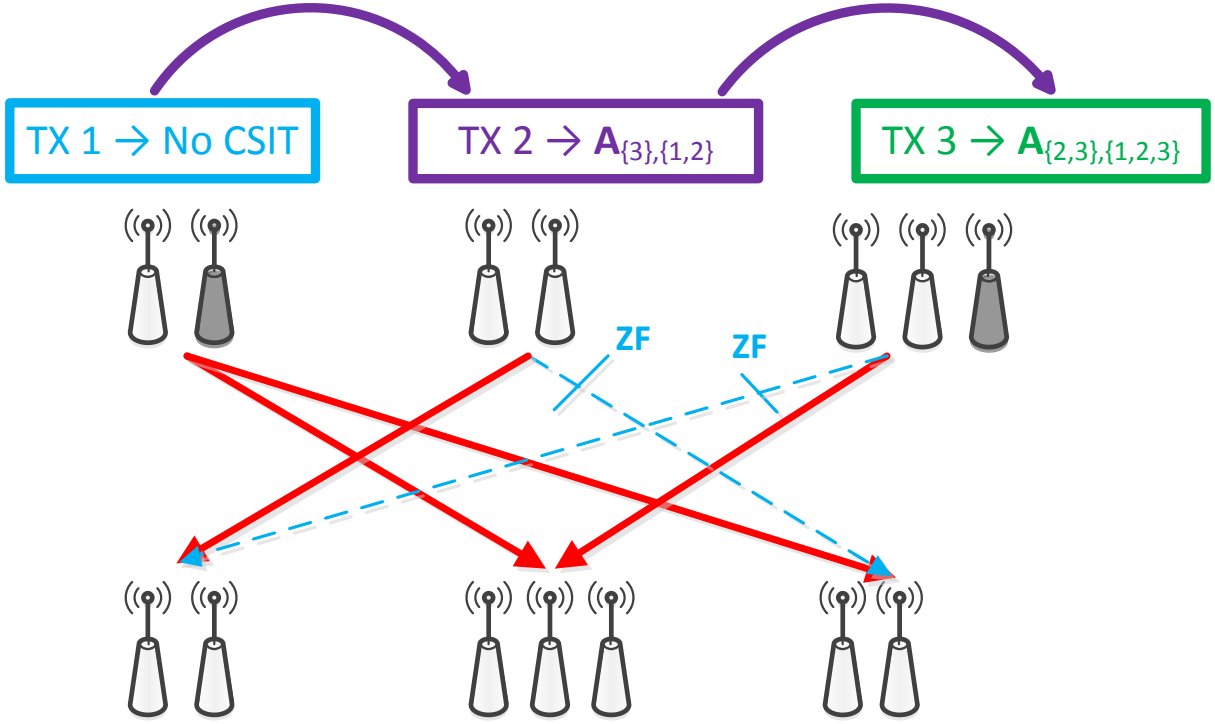
THEOREM (1)

In a tightly-feasible IC, there exists a strictly incomplete CSIT allocation preserving IA feasibility if there exists a tightly-feasible IC strictly included in the full IC.



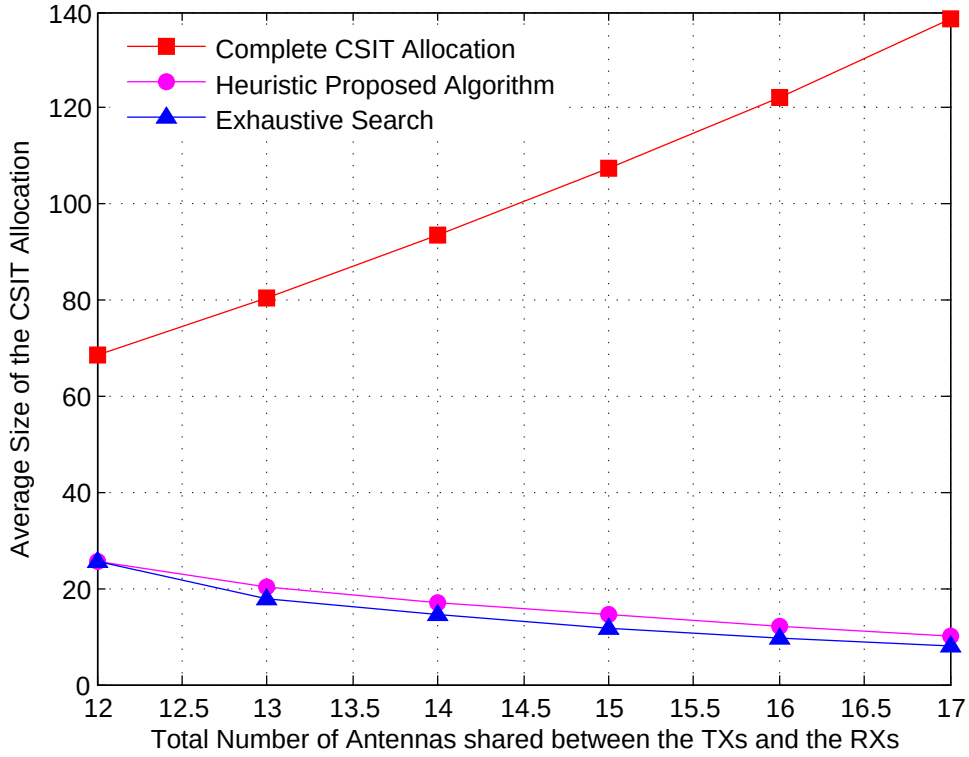
[1] P. de Kerret and D. Gesbert, *"CSI Sharing Strategies in Wireless Networks"*, IEEE Wireless Communication Magazine, Feb. 2013.

Extension to super feasible settings



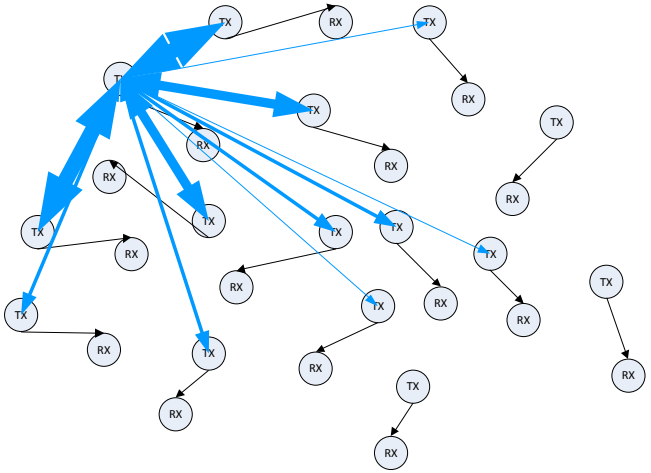
CSI allocation algorithm for general IC

- Combinatorial algorithm to find where to exploit the additional antennas
- Developed a heuristic algorithm exploiting the analysis for the Tightly-feasible setting



The Network MIMO channel: DoF-preserving reduced spatial CSIT allocation policies

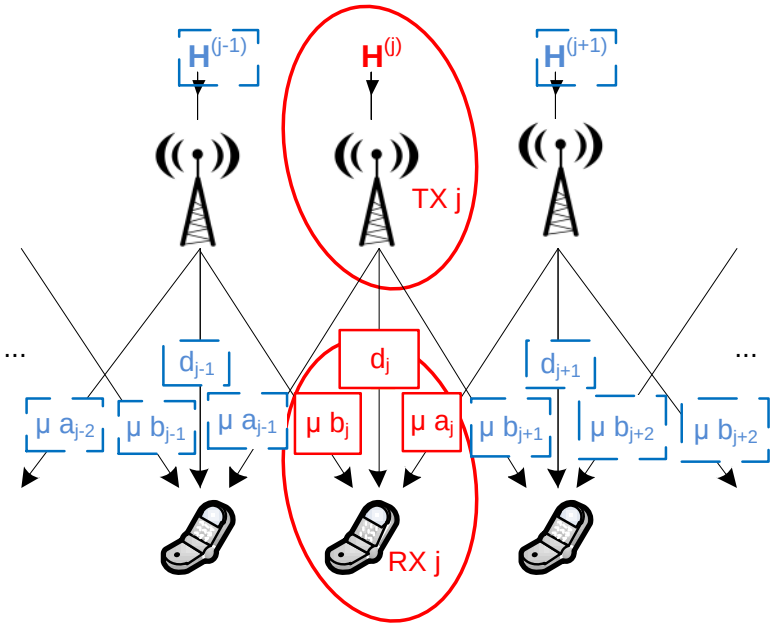
- Large literature on MIMO-BC with imperfect CSIT but CSIT always perfectly shared between the TX antennas
- Is it necessary for every TX to know every coefficient perfectly?
- Intuitively cooperation should decrease with the distance



Our question: With which accuracy should which user channel be known at which TX?

THE SIMPLE WYNER MODEL CAN TEACH A LOT

- Interference only from direct neighboring TXs with an inter-cell attenuation factor $\mu \in (0, 1)$



Remark: Results not dependent on this pathloss structure. Extension to linear exponentially decaying channels in [2]

[2] P. de Kerret and D. Gesbert, *"CSI feedback allocation in multicell MIMO channels"*, Proc. ICC 2012.

Generalized DoFs

- Following R. Etkin, D. Tse, and H. Wang IT08, define the interference level α at a RX

$$\forall i, \alpha \triangleq \frac{\log(\text{INR})}{\log(\text{SNR})} = \frac{\log(\sum_{\ell \neq i} |\{\mathbf{H}\}_{i\ell}|^2 P)}{\log(|\{\mathbf{H}\}_{ii}|^2 P)}$$

by $\mu^2 \sim P^{\alpha-1}$

- Generalized DoFs:

$$\forall \alpha \in (0,1)\, \text{DoF}_i(\alpha, \mathbf{B}) \triangleq \lim_{\text{SNR},\text{INR}\to\infty, \frac{\log(\text{INR})}{\log(\text{SNR})}=\alpha} \frac{R_i(\mathbf{B})}{\log(\text{SNR})} \qquad (9)$$

with **B** the CSIT allocation matrix defined in the following

remark: Generalized DoF necessary to represent the impact of the pathloss over the DoFs

CSIT Allocation

- CSIT allocation matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$ such that i-th row of $\mathbf{H}^{(j)}$ obtained using $\{\mathbf{B}\}_{ij}$ bits
- In [4] shown necessary to have $B_i^{(j)} \propto \log_2(P)$ for DoF > 0
- Size $s(\bullet)$ of a CSIT allocation at TX j

$$s(\mathbf{B}^{(j)}) = \lim_{P \to \infty} \frac{\sum_{i=1}^{K} B_i^{(j)}}{\log_2(P)} \tag{10}$$

- Set of DoF-optimal CSIT allocations $\mathbb{B}_{\text{DoF}}$

$$\mathbb{B}_{\text{DoF}} \triangleq \{\mathbf{B} | \forall i, \text{DoF}_i(\alpha, \mathbf{B}) = 1\} \tag{11}$$

[3] N. Jindal, *"MIMO Broadcast Channels with Finite Rate Feedback"*, Trans. IT, 2006.

CSIT Allocation

- CSIT allocation matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$ such that i-th row of $\mathbf{H}^{(j)}$ obtained using $\{\mathbf{B}\}_{ij}$ bits
- In [4] shown necessary to have $B_i^{(j)} \propto \log_2(P)$ for DoF > 0
- Size $\mathsf{s}(\bullet)$ of a CSIT allocation at TX j

$$\mathsf{s}(\mathbf{B}^{(j)}) = \lim_{P \to \infty} \frac{\sum_{i=1}^{K} B_i^{(j)}}{\log_2(P)} \tag{10}$$

- Set of DoF-optimal CSIT allocations $\mathbb{B}_{\mathsf{DoF}}$

$$\mathbb{B}_{\mathsf{DoF}} \triangleq \{\mathbf{B} | \forall i, \mathsf{DoF}_i(\alpha, \mathbf{B}) = 1\} \tag{11}$$

[3] N. Jindal, *"MIMO Broadcast Channels with Finite Rate Feedback"*, Trans. IT, 2006.

CSIT Allocation

- CSIT allocation matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$ such that i-th row of $\mathbf{H}^{(j)}$ obtained using $\{\mathbf{B}\}_{ij}$ bits
- In [4] shown necessary to have $B_i^{(j)} \propto \log_2(P)$ for DoF > 0
- Size s(•) of a CSIT allocation at TX j

$$\mathrm{s}(\mathbf{B}^{(j)}) = \lim_{P \to \infty} \frac{\sum_{i=1}^{K} B_i^{(j)}}{\log_2(P)} \tag{10}$$

- Set of DoF-optimal CSIT allocations $\mathbb{B}_{\mathrm{DoF}}$

$$\mathbb{B}_{\mathrm{DoF}} \triangleq \{\mathbf{B} | \forall i, \mathrm{DoF}_i(\alpha, \mathbf{B}) = 1\} \tag{11}$$

[3] N. Jindal, *"MIMO Broadcast Channels with Finite Rate Feedback"*, Trans. IT, 2006.

CSIT Allocation

- CSIT allocation matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$ such that i-th row of $\mathbf{H}^{(j)}$ obtained using $\{\mathbf{B}\}_{ij}$ bits
- In [4] shown necessary to have $B_i^{(j)} \propto \log_2(P)$ for DoF > 0
- Size s(•) of a CSIT allocation at TX j

$$\mathrm{s}(\mathbf{B}^{(j)}) = \lim_{P \to \infty} \frac{\sum_{i=1}^{K} B_i^{(j)}}{\log_2(P)} \tag{10}$$

- Set of DoF-optimal CSIT allocations $\mathbb{B}_{\mathrm{DoF}}$

$$\mathbb{B}_{\mathrm{DoF}} \triangleq \{\mathbf{B} | \forall i, \mathrm{DoF}_i(\alpha, \mathbf{B}) = 1\} \tag{11}$$

$$\mathrm{minimize}_{\mathbf{B}}\ \mathrm{s}(\mathbf{B}),\ \text{subject to } \mathbf{B} \in \mathbb{B}_{\mathrm{DoF}} \tag{12}$$

[3] N. Jindal, *"MIMO Broadcast Channels with Finite Rate Feedback"*, Trans. IT, 2006.

CSIT Allocation

- CSIT allocation matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$ such that i-th row of $\mathbf{H}^{(j)}$ obtained using $\{\mathbf{B}\}_{ij}$ bits
- In [4] shown necessary to have $B_i^{(j)} \propto \log_2(P)$ for DoF > 0
- Size s(•) of a CSIT allocation at TX j

$$\mathrm{s}(\mathbf{B}^{(j)}) = \lim_{P \to \infty} \frac{\sum_{i=1}^{K} B_i^{(j)}}{\log_2(P)} \tag{10}$$

- Set of DoF-optimal CSIT allocations $\mathbb{B}_{\mathrm{DoF}}$

$$\mathbb{B}_{\mathrm{DoF}} \triangleq \{\mathbf{B} | \forall i, \mathrm{DoF}_i(\alpha, \mathbf{B}) = 1\} \tag{11}$$

$$\mathrm{minimize}_{\mathbf{B}}\ \mathrm{s}(\mathbf{B}),\ \text{subject to } \mathbf{B} \in \mathbb{B}_{\mathrm{DoF}} \tag{12}$$

[3] N. Jindal, *"MIMO Broadcast Channels with Finite Rate Feedback"*, Trans. IT, 2006.

CSIT Allocation

- CSIT allocation matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$ such that i-th row of $\mathbf{H}^{(j)}$ obtained using $\{\mathbf{B}\}_{ij}$ bits
- In [4] shown necessary to have $B_i^{(j)} \propto \log_2(P)$ for DoF > 0
- Size s(•) of a CSIT allocation at TX j

$$\mathrm{s}(\mathbf{B}^{(j)}) = \lim_{P \to \infty} \frac{\sum_{i=1}^{K} B_i^{(j)}}{\log_2(P)} \tag{10}$$

- Set of DoF-optimal CSIT allocations $\mathbb{B}_{\mathrm{DoF}}$

$$\mathbb{B}_{\mathrm{DoF}} \triangleq \{\mathbf{B} | \forall i, \mathrm{DoF}_i(\alpha, \mathbf{B}) = 1\} \tag{11}$$

$$\text{minimize}_{\mathbf{B}}\ \mathrm{s}(\mathbf{B}), \text{ subject to } \mathbf{B} \in \mathbb{B}_{\mathrm{DoF}} \tag{12}$$

[3] N. Jindal, *"MIMO Broadcast Channels with Finite Rate Feedback"*, Trans. IT, 2006.

CSIT Allocation

- CSIT allocation matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$ such that i-th row of $\mathbf{H}^{(j)}$ obtained using $\{\mathbf{B}\}_{ij}$ quantization bits
- It is known to be necessary to have $B_i^{(j)} \propto \log_2(P)$ to preserve rate scaling with SNR.

Theorem

Let's define $\mathbf{B}^{Dist}$ *as*

$$\forall i, j,\ \{\mathbf{B}^{Dist}\}_{ij} = \lceil [1+(\gamma-1)|i-j|]^+ + 2[\gamma+(\gamma-1)|i-j|]^+ \log_2(P) \rceil. \tag{13}$$

then $\mathbf{B}^{Dist} \in \mathbb{B}_{\text{DoF}}$

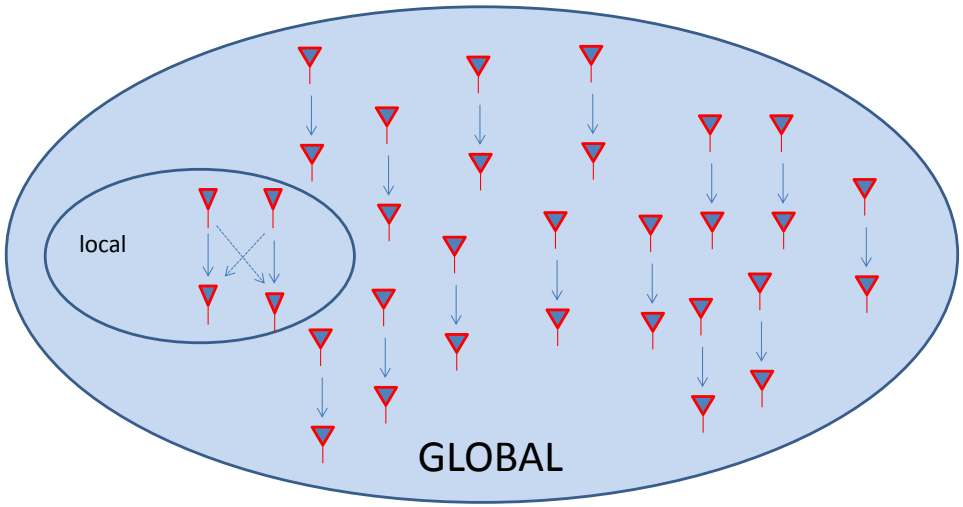
Lesson: It is only necessary to share CSIT and symbol s_i at TX j if distance not too large

$$1 + (\alpha - 1)(2|i - j|) > 0 \tag{14}$$

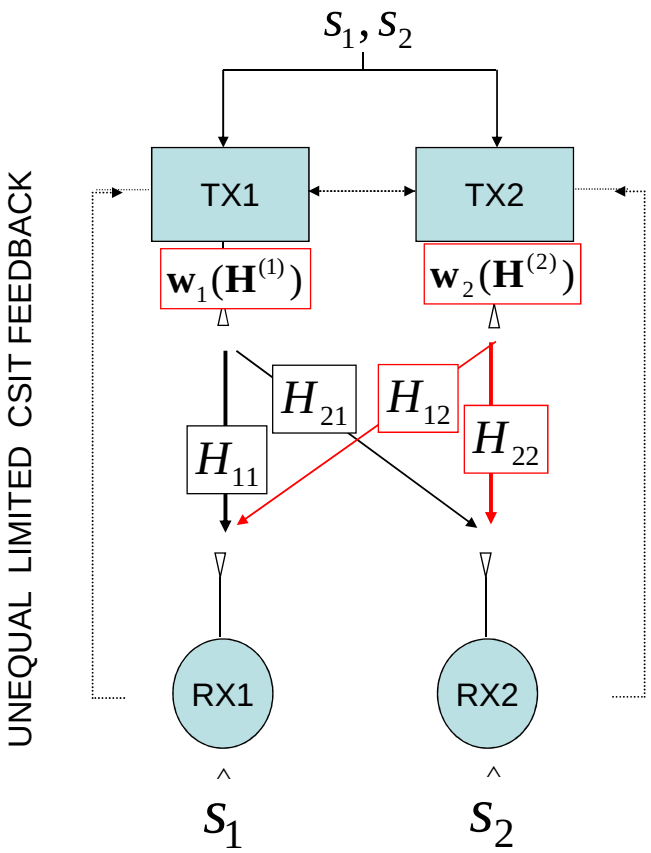
DISTRIBUTED COORDINATION: THE LOCAL SCALE

Let us consider the two cell network MIMO setting:

- What is the DoF of conventional precoding with distributed CSIT?
- What the optimal DoF with distributed CSIT?



MIMO PRECODING WITH ARBITRARY CSI SHARING



Distributed Transmit Matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{w}_1(\mathbf{H}^{(1)}) \\ \hline \mathbf{w}_2(\mathbf{H}^{(2)}) \end{bmatrix}$$

$$\mathbf{H}^{(1)} \neq \mathbf{H}^{(2)}$$

Network MIMO with Distributed CSI

- Received signal y_1 at RX 1 and y_2 at RX 2 written as:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1^{\mathrm{H}} \\ \boldsymbol{h}_2^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

where:

- s_i is the symbol transmitted to RX i
- $\boldsymbol{t}_i$ is the beamformer carrying symbol s_i
- $\boldsymbol{h}_i^{\mathrm{H}}$ is the channel from the TXs to RX i
- $\boldsymbol{\eta}$ is the additive white Gaussian noise

- Main figure of merit is the Multiplexing Gain (MG) defined as: $\forall i \in \{1, 2\}$,

$$\mathrm{M_{G}}_i \triangleq \lim_{P \to \infty} \frac{R_i(P)}{\log_2(P)} \tag{15}$$

- Study the average sum rate:

$$\forall i \in \{1, 2\}, R_i(P) \triangleq \mathrm{E}_{\mathbf{H}, \mathcal{W}} \left[\log_2 \left(1 + \frac{|\boldsymbol{h}_i^{\mathrm{H}} \boldsymbol{t}_i|^2}{1 + |\boldsymbol{h}_i^{\mathrm{H}} \boldsymbol{t}_{\bar{i}}|^2} \right) \right] \tag{16}$$

Distributed Precoding

- $\boldsymbol{h}_i^{(j)\mathrm{H}}$ estimate at TX j of the normalized channel $\tilde{\boldsymbol{h}}_i^{\mathrm{H}}$ to RX i
- $B_i^{(j)}$ number of bits quantizing $\boldsymbol{h}_i^{(j)\mathrm{H}}$
- At TX j, computation of

$$\boldsymbol{T}^{(j)} = \begin{bmatrix} \boldsymbol{t}_1^{(j)} & \boldsymbol{t}_2^{(j)} \end{bmatrix} = \begin{bmatrix} T_{11}^{(j)} & T_{12}^{(j)} \\ T_{21}^{(j)} & T_{22}^{(j)} \end{bmatrix} \tag{17}$$

- Only the j-th row of $\boldsymbol{T}^{(j)}$ implemented:

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 \end{bmatrix} = \begin{bmatrix} T_{11}^{(1)} & T_{12}^{(1)} \\ T_{21}^{(2)} & T_{22}^{(2)} \end{bmatrix} \tag{18}$$

DISTRIBUTED CSI MODEL

Inspired from Jindal's result in broadcast case, we define the Feedback scaling matrix $\boldsymbol{\alpha} \in \mathbb{R}^{2\times 2}$ as

$$\forall i, j \in \{1, 2\}, \quad \{\boldsymbol{\alpha}\}_{ij} \triangleq \alpha_i^{(j)} \triangleq \lim_{P\to\infty} \frac{B_i^{(j)}}{\log_2(P)} \tag{19}$$

Conventional Zero Forcing (two single ant. users)

$$\boldsymbol{t}_i^{\mathrm{C-ZF}(j)} \triangleq \begin{bmatrix} t_{1i}^{\mathrm{C-ZF}(j)} \\ t_{2i}^{\mathrm{C-ZF}(j)} \end{bmatrix} \triangleq \sqrt{\frac{P}{2}} \frac{\Pi^{\perp}_{\tilde{\boldsymbol{h}}_{\bar{i}}^{(j)}}\left(\tilde{\boldsymbol{h}}_i^{(j)}\right)}{\|\Pi^{\perp}_{\tilde{\boldsymbol{h}}_{\bar{i}}^{(j)}}\left(\tilde{\boldsymbol{h}}_i^{(j)}\right)\|}, \quad j \in \{1,2\} \tag{20}$$

Theorem

The MG achieved with C-ZF is equal to

$$M_G^{\mathrm{ZF}} = 2 \min_{i,j \in \{1,2\}} \alpha_i^{(j)}. \tag{21}$$

Can we do better?

Beacon-ZF Beamformer defined as:

$$\boldsymbol{t}_i^{\mathrm{B-ZF}(j)} \triangleq \begin{bmatrix} t_{1i}^{\mathrm{B-ZF}(j)} \\ t_{2i}^{\mathrm{B-ZF}(j)} \end{bmatrix} \triangleq \sqrt{\frac{P}{2}} \frac{\Pi^{\perp}_{\tilde{\boldsymbol{h}}_i^{(j)}}(\boldsymbol{c}_i)}{\|\Pi^{\perp}_{\tilde{\boldsymbol{h}}_i^{(j)}}(\boldsymbol{c}_i)\|} \tag{22}$$

where $\boldsymbol{c}_i$ is any vector known to both TXs beforehand.

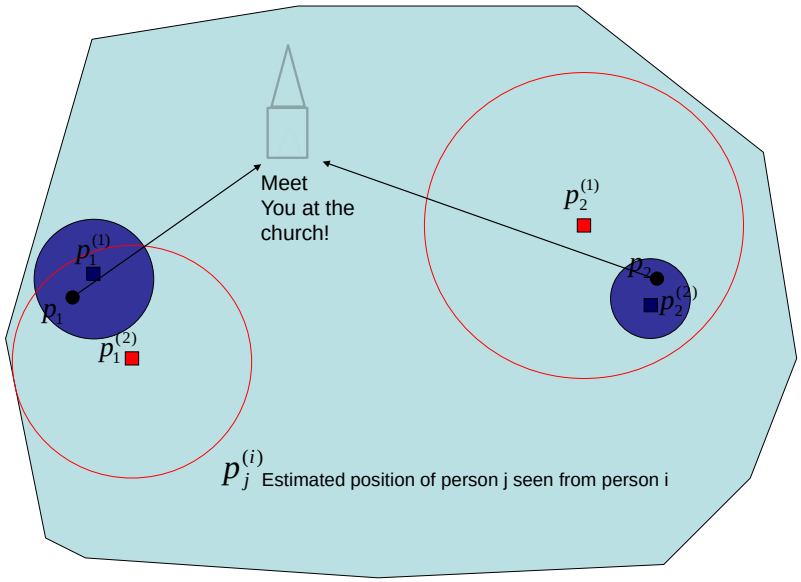
THEOREM

The MG achieved with B-ZF is

$$M_G^{\mathrm{B-ZF}} = \min_{j\in\{1,2\}} \alpha_1^{(j)} + \min_{j\in\{1,2\}} \alpha_2^{(j)} \tag{23}$$

THE DISTRIBUTED RENDEZ-VOUS PROBLEM

- "'Beacon" = pre-agreed meeting location.
- E.g. the two lovers agree to meet at main city church and walk into this direction



Can we do even better?

Active-Passive Zero Forcing

- Assume we wish to transmit message i, $(i = 1, 2)$
- Assume w.l.o.g. that $\alpha_{\tilde{i}}^{(2)} \geq \alpha_{\tilde{i}}^{(1)}$, then

$$\boldsymbol{t}_i^{\mathrm{AP-ZF}} \triangleq \sqrt{\frac{P}{2\log_2(P)}} \begin{bmatrix} 1 \\ -\frac{H_{\bar{i}1}^{(2)}}{H_{\bar{i}2}^{(2)}} \end{bmatrix} \tag{24}$$

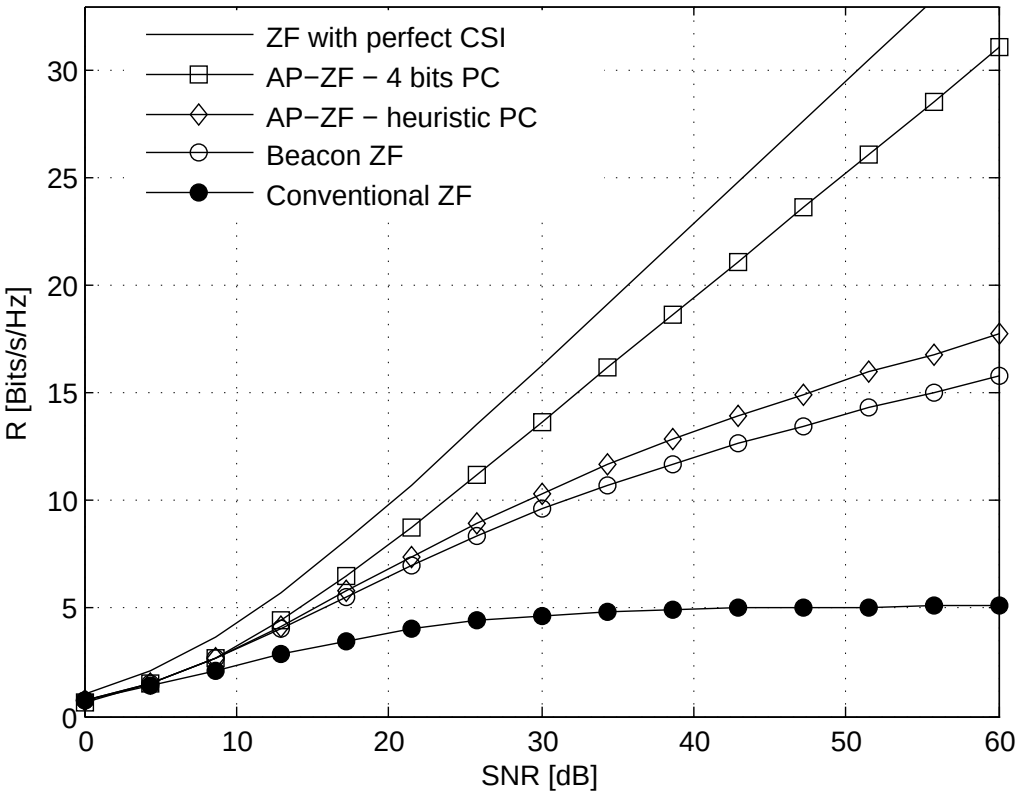
Theorem

AP-ZF is MG maximizing across all distributed precoders and

$$M_G^{\mathrm{AP-ZF}} = \max_{j\in[1,2]} \alpha_1^{(j)} + \max_{j\in[1,2]} \alpha_2^{(j)} \tag{25}$$

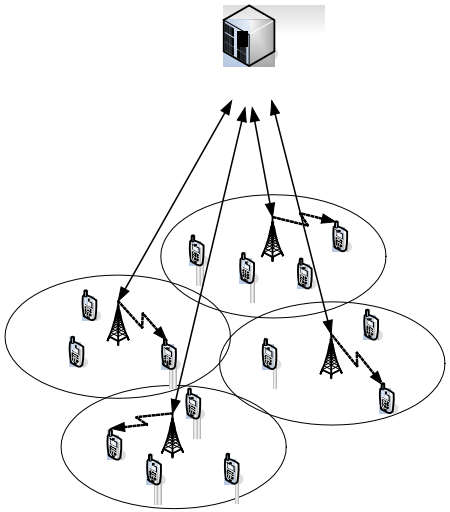
P. de Kerret, D. Gesbert "Degrees of freedom of the network MIMO channel with distributed CSI", November 2012

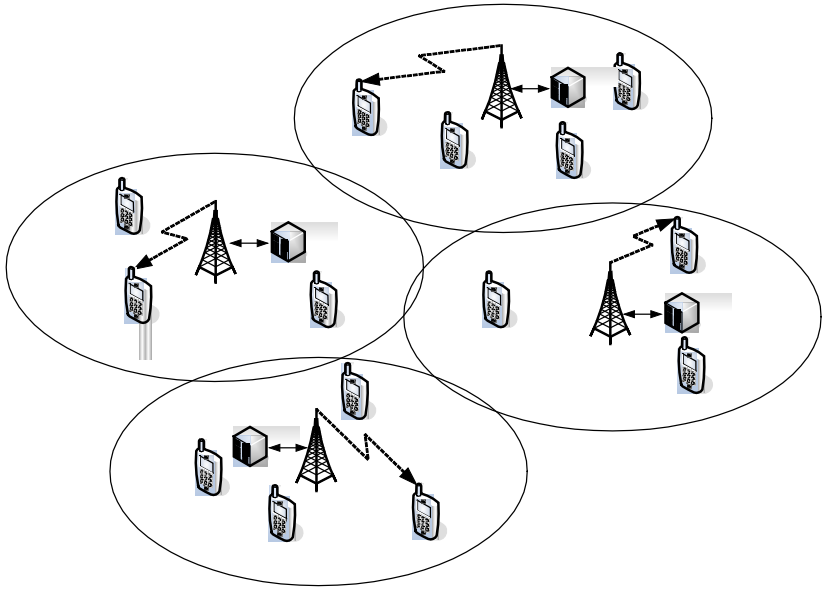
Sum rate vs SNR with $[\alpha_1^{(1)}, \alpha_1^{(2)}] = [1, 0.5]$, $[\alpha_2^{(1)}, \alpha_2^{(2)}] = [0, 0.7]$



Part 4: Distributed coordination: Using large dimensions

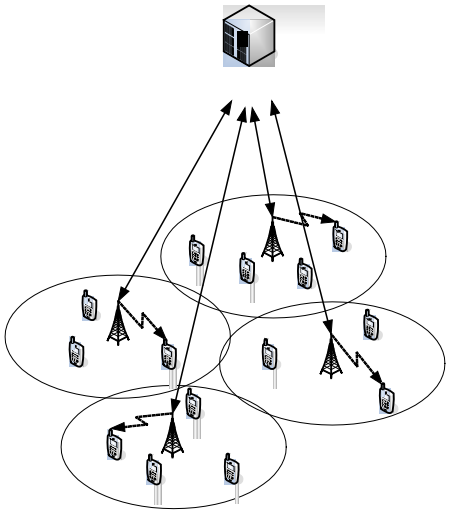
- Fully distributed scheduling (in the many user regime)
- Fully distributed beamforming (in the many antenna regime)

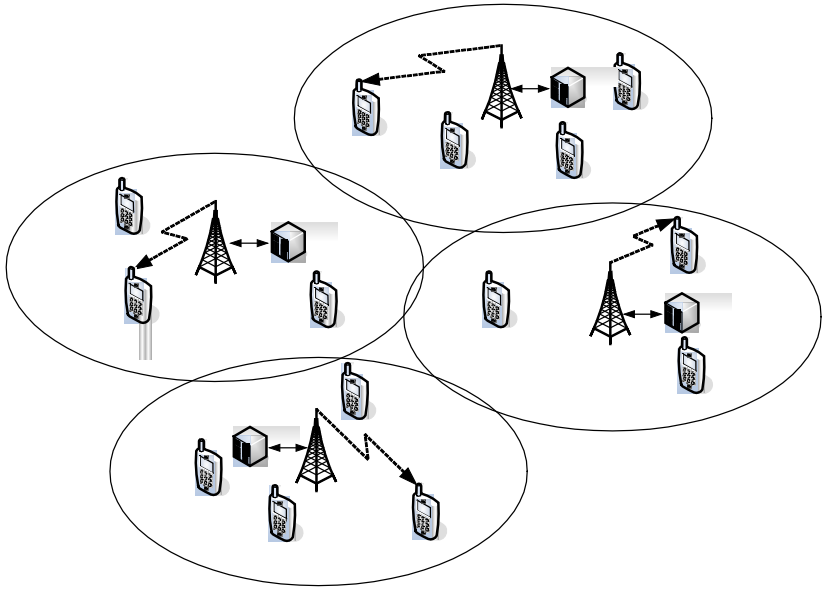




Part 4: Distributed coordination: Using large dimensions

- Fully distributed scheduling (in the many user regime)
- Fully distributed beamforming (in the many antenna regime)





Coordinated scheduling over N cells with $U >> 1$ users

Definition : A ***scheduling vector*** $\boldsymbol{U}$ for a given resource slot contains the set of users simultaneously scheduled across all cells:

$$\boldsymbol{U} = [u_1 \; u_2 \; \cdots \; u_n \; \cdots \; u_N] \; 1 \leq u_j \leq U$$

Definition : A ***transmit power vector*** $\boldsymbol{P}$ contains the transmit power values used by each transmitter towards its respective user:

$$\boldsymbol{P} = [P_{u_1} \; P_{u_2} \; \cdots \; P_{u_n} \; \cdots \; P_{u_N}]$$

where $[\boldsymbol{P}]_n = P_{u_n} = \mathbb{E}|X_{u_n}|^2 \leq P_{\max}$.

The SINR for the user selected in cell n is

$$\Gamma([\boldsymbol{U}]_n, \boldsymbol{P}) = \frac{G_{u_n,n} P_{u_n}}{\sigma^2 + \sum_{i \neq n}^{N} G_{u_n,i} P_{u_i}}, \tag{26}$$

The system capacity under single user decoding is

$$\mathcal{C}(\boldsymbol{U}, \boldsymbol{P}) \triangleq \frac{1}{N} \sum_{n=1}^{N} \log\Big(1 + \Gamma([\boldsymbol{U}]_n, \boldsymbol{P})\Big). \tag{27}$$

$$\text{Problem:} \quad (\boldsymbol{U}^*, \boldsymbol{P}^*) = \arg\max_{\substack{\boldsymbol{U} \in \Upsilon \\ \boldsymbol{P} \in \Omega}} \mathcal{C}(\boldsymbol{U}, \boldsymbol{P}), \tag{28}$$

UPPER BOUND AND LOWER BOUND ON CAPACITY

Assuming no interference and path loss+rayleigh based channel gains $G_{u_n,i}$:

$$\mathcal{C}(\boldsymbol{U}^*, \boldsymbol{P}^*) \leq \mathcal{C}^{ub} = \frac{1}{N}\sum_{n=1}^{N} \log\left(1 + \max_{u_n=1..U}\{G_{u_n,n}\}P_{max}/\sigma^2\right). \quad (29)$$

Assuming full-powered interference:

$$\mathcal{C}(\boldsymbol{U}^*, \boldsymbol{P}^*) \geq \mathcal{C}^{lb} = \frac{1}{N}\sum_{n=1}^{N} \log\left(1 + \max_{u_n=1..U}\{\frac{\{G_{u_n,n}\}P_{max}}{\sigma^2 + \sum_{i\neq n}^{N} G_{u_n,i}P_{max}}\}\right). \quad (30)$$

CAPACITY SCALING FOR LARGE NUMBER OF USERS

Theorem 1: The upper bound on capacity behaves like:

$$E(\mathcal{C}^{ub}) \approx \frac{\epsilon}{2} \log U \text{ for large } U \tag{31}$$

Theorem 2: The lower bound on capacity behaves like:

$$E(\mathcal{C}^{lb}) \approx \frac{\epsilon}{2} \log U \text{ for large } U \tag{32}$$

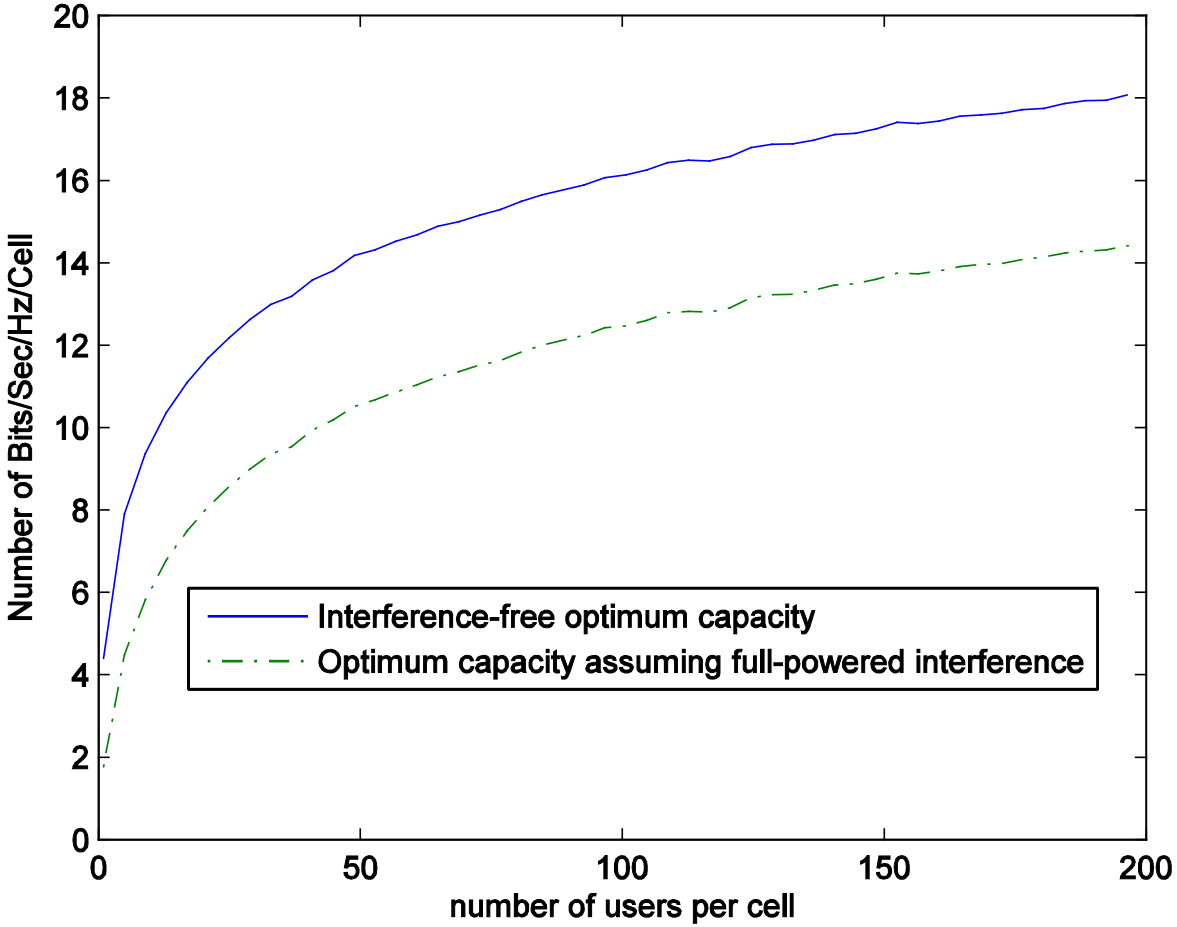
Corollary :

$$E(\mathcal{C}(\boldsymbol{U}^*, \boldsymbol{P}^*)) \approx \frac{\epsilon}{2} \log U \tag{33}$$

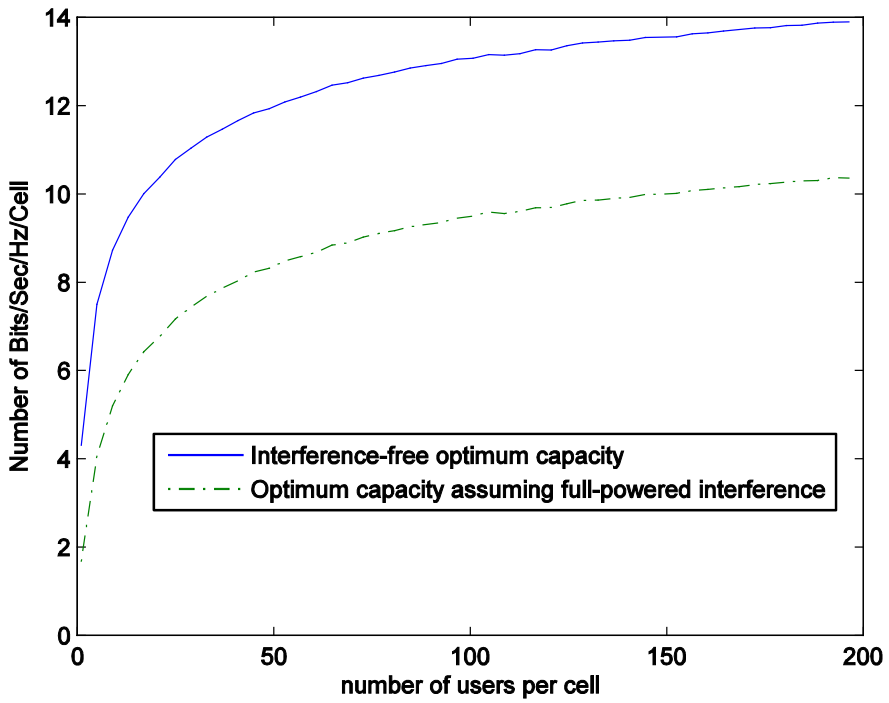
- Interference loss vanish for large number of users
- No big price paid for looking for small intf users!

D. Gesbert, M. Kountouris, "Rate Scaling Laws in Multicell Networks under Distributed Power Control and User Scheduling", in IEEE Trans. On Information Theory, Jan. 2011

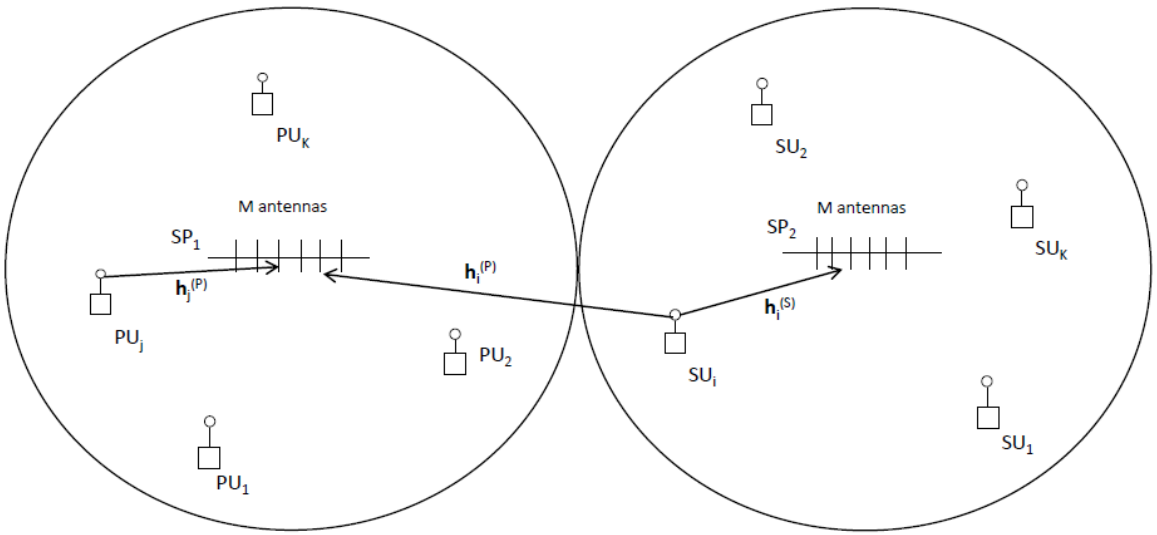
Capacity scaling for 4 cell network



Capacity scaling with disk of exclusion around base

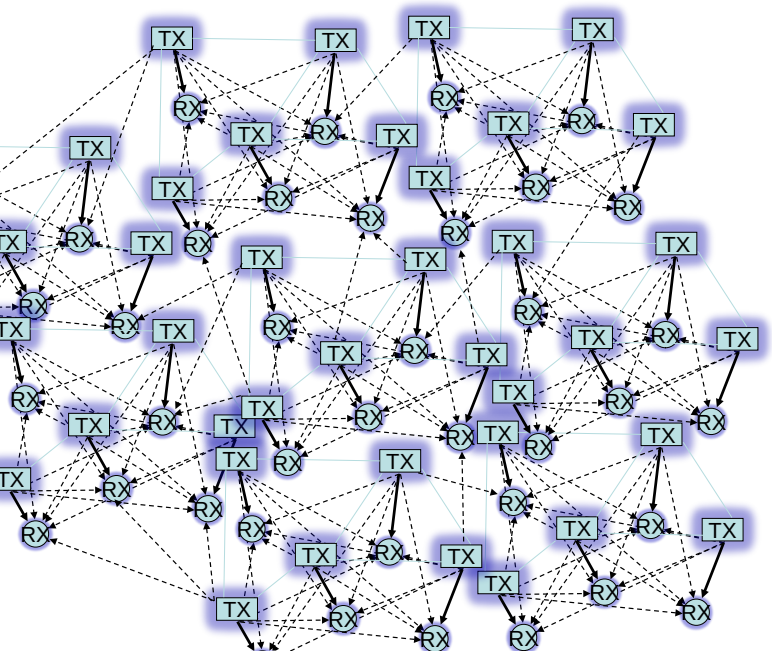


Achieving distributed multi-cell beamforming using Massive MIMO



THE "DENSE VS. MASSIVE" DEBATE

Dense cooperation (single antenna base station)



Massive MIMO base station (no cooperation)

Massive MIMO principles

- Let M antennas be used at BS 1 and BS 2.
- As $M \to \infty$ (normalized) useful and interference channel vector become quasi orthogonal
- Matched filter maximizes SNR *and* cancels interference *simultaneously* [Marzetta 2010]
- Performance analysis as finite number of antennas (random matrices) [Hoydis et al 2011]
- Matched filter solution is fully distributed!

But there is a problem...

PILOT CONTAMINATION

- Non orthogonal pilots -> pilot contamination (PC)
- PC destroys Massive MIMO theoretical benefits

Pilot sequence in l-th cell: $\mathbf{s}_l = [\ s_{l1} \quad s_{l2} \quad \cdots \quad s_{l\tau}\ ]^T$
the $M \times \tau$ signal at the target base station (with noise $\mathbf{N}$) is

$$\mathbf{Y} = \sum_{l=1}^{L} \mathbf{h}_l \mathbf{s}_l^T + \mathbf{N} \tag{34}$$

LS estimator for: $\widehat{\mathbf{h}}_1^{\mathrm{LS}} = \mathbf{Y}\mathbf{s}^*(\mathbf{s}^T\mathbf{s}^*)^{-1}$
With full pilot reuse:

$$\widehat{\mathbf{h}}_1^{\mathrm{LS}} = \mathbf{h}_1 + \sum_{l \neq 1}^{L} \mathbf{h}_l + \mathbf{N}\mathbf{s}^*/\tau \tag{35}$$

A MORE POWERFUL ESTIMATOR

$$p(\mathbf{h}|\mathbf{y}) = \frac{\exp\left(-\left(\mathbf{h}^H\mathbf{R}^{-1}\mathbf{h} + (\mathbf{y}-\mathbf{S}\mathbf{h})^H(\mathbf{y}-\mathbf{S}\mathbf{h})/\sigma_n^2\right)\right)}{AB}$$

where

$$\mathbf{R} \triangleq \mathrm{diag}(\mathbf{R}_1, \cdots, \mathbf{R}_L) \tag{36}$$

$A \triangleq (\pi\sigma_n^2)^{M\tau}$ and

$$B \triangleq \pi^{LM}(\det\mathbf{R})^M \tag{37}$$

Develop covariance-based (Bayesian) estimator

$$\widehat{\mathbf{h}}_1 = \mathbf{R}_1\left(\sigma_n^2\mathbf{I}_M + \tau\sum_{l=1}^{L}\mathbf{R}_l\right)^{-1}\mathbf{S}^H\mathbf{y} \tag{38}$$

$\mathbf{R}_l$ is covariance matrix of l-th interference channel.

E.g specular channel model: $\mathbf{h}_i = \frac{1}{\sqrt{P}} \sum_{p=1}^{P} \mathbf{a}(\theta_{ip})\alpha_{ip}$

where P is number of paths and

$$\mathbf{a}(\theta) \triangleq \begin{bmatrix} 1 \\ e^{-j2\pi \frac{D}{\lambda} \cos(\theta)} \\ \vdots \\ e^{-j2\pi \frac{(M-1)D}{\lambda} \cos(\theta)} \end{bmatrix} \tag{39}$$

Density function of random variable θ contains useful information, captured by correlation matrix.

$$\mathbf{R_i} = \frac{\delta_i^2}{P} \sum_{p=1}^{P} \mathbb{E}\{\mathbf{a}(\theta_{ip})\mathbf{a}(\theta_{ip})^H\} = \delta_i^2 \mathbb{E}\{\mathbf{a}(\theta_i)\mathbf{a}(\theta_i)^H\}$$

Theorem

Assume multipath AOA θ for user j (at target BS 1) has density $p_j(\theta)$ with bounded support, i.e. $p_j(\theta) = 0$ for $\theta \notin [\theta_j^{\min}, \theta_j^{\max}]$ for some fixed $\theta_j^{\min} \leqslant \theta_j^{\max} \in [0, \pi]$. If the $L-1$ intervals $[\theta_i^{\min}, \theta_i^{\max}]$, $i = 2, \ldots, L$ are strictly non-overlapping with $[\theta_1^{\min}, \theta_1^{\max}]$, we have

$$\lim_{M \to \infty} \widehat{\mathbf{h}}_1 = \widehat{\mathbf{h}}_1^{\text{no int}} \tag{40}$$

If desired and interference multipath ranges do not overlap, pilot contamination vanish asymptotically.

H. Yin. D. Gesbert, M. Filippou, Y. Liu "A Coordinated Approach to Channel Estimation in Large-scale Multiple-antenna Systems", in IEEE Journal on Selected Areas in Communications, Special Issue on Large Scale Antenna Systems. Feb. 2013.

SKETCH OF PROOF

Proof relies on three lemmas:

Lemma 1:

Define $\boldsymbol{\alpha}(\mathrm{x}) \triangleq [\ 1 \quad e^{-j\pi x} \quad \cdots \quad e^{-j\pi(M-1)x}\ ]^T$ and $\mathcal{A} \triangleq \mathrm{span}\{\boldsymbol{\alpha}(\mathrm{x}) | \mathrm{x} \in [-1, 1]\}$. Given $b_1, b_2 \in [-1, 1]$ and $b_1 < b_2$, define $\mathcal{B} \triangleq \mathrm{span}\{\boldsymbol{\alpha}(\mathrm{x}) | \mathrm{x} \in [b_1, b_2]\}$, then

- $\dim\{\mathcal{A}\} = M$
- $\dim\{\mathcal{B}\} \sim (b_2 - b_1)M/2$ when M grows large.

lemma 2 *When M grows large,*

$$\text{rank}(\mathbf{R}_i) \leqslant d_i M$$

where

$$d_i \triangleq \left(\cos(\theta_i^{min}) - \cos(\theta_i^{max})\right) \frac{D}{\lambda}$$

Lemma 1 indicates that for large M, there exists a null space $\text{null}(\mathbf{R}_i)$ of dimension $(1 - d_i)M$.

lemma 3 *When M is large, the null space* $\text{null}(\mathbf{R}_i)$ *includes the following set of unit norm vectors:*

$$\text{null}(\mathbf{R}_i) \supset \text{span}\left\{\frac{\mathbf{a}(\Phi)}{\sqrt{M}}, \forall \Phi \notin [\theta_i^{\text{min}}, \theta_i^{\text{max}}]\right\}$$

Coordinated Pilot Assignement (CPA):

- Estimate and exchange covariance information between cells (slow varying)
- Apply a coordinated pilot assignement based on covariance information to fulfill (almost) non-overlap condition between signal subspaces
  A given pilot sequence is assigned to a user set $\mathcal{U}$ over L cells, minimizing a utility function

$$\mathsf{F}(\mathcal{U}) \triangleq \sum_{j=1}^{|\mathcal{U}|} \frac{\mathcal{M}_j(\mathcal{U})}{\mathrm{tr}\left\{\mathbf{R}_{jj}(\mathcal{U})\right\}} \tag{41}$$

  where $\mathcal{M}_j(\mathcal{U})$ is the MSE for the desired channel at the j-th base station
- Use a greedy approach to avoid exhaustive search

Angle spread 10 degrees

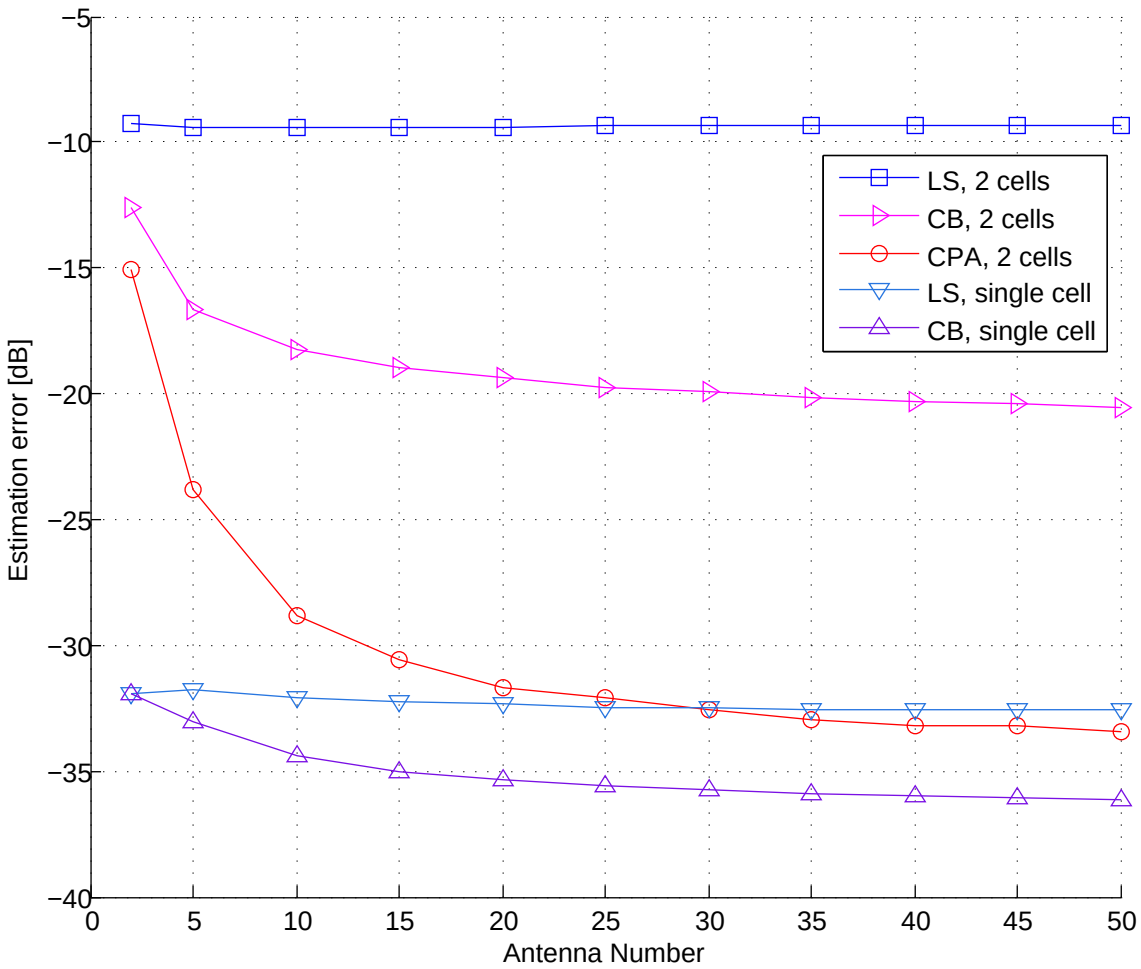


FIGURE: Estimation MSE vs. antenna number, Gaussian distributed AOAs with $\sigma = 10$ degrees.

10 Antennas

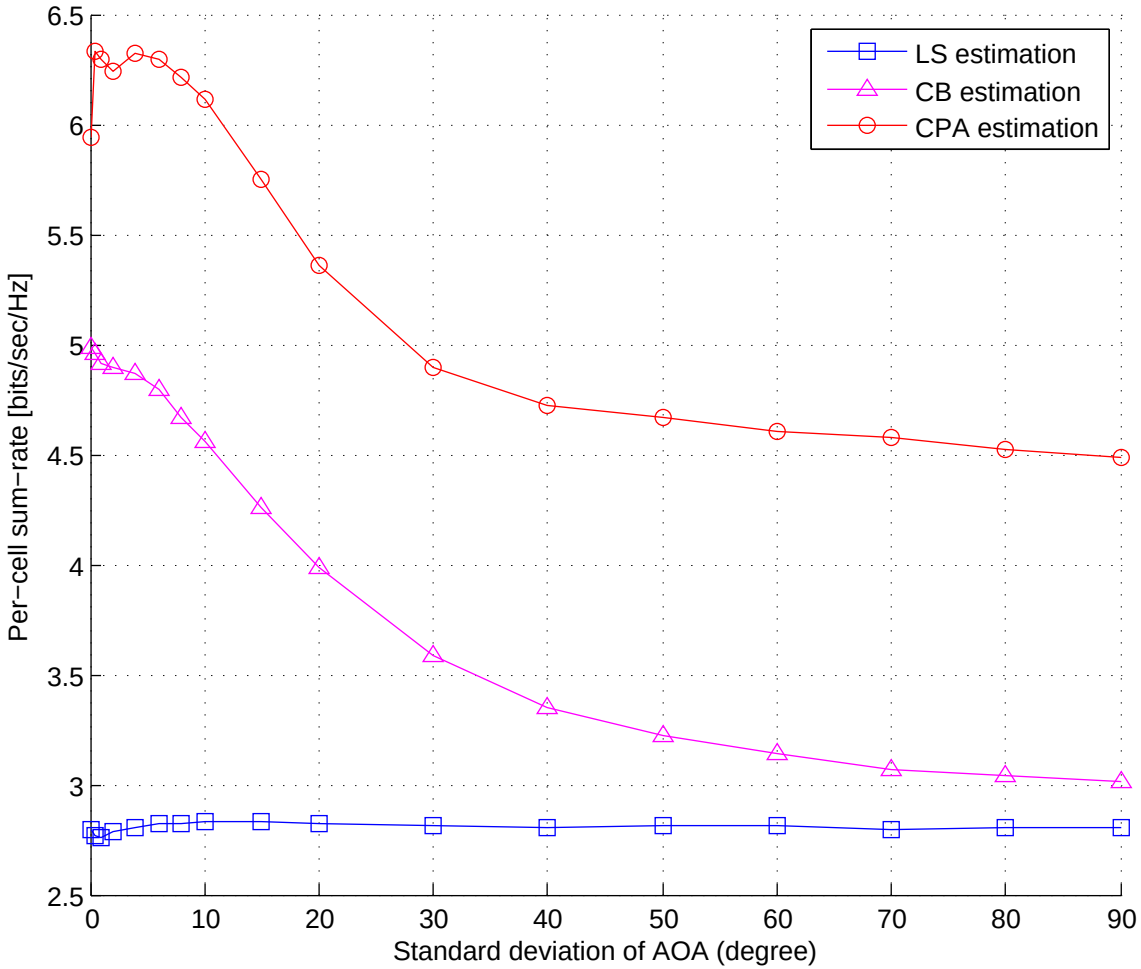


FIGURE: Per-cell sum-rate vs. standard deviation of AOA (Gaussian distribution) with $M = 10$, 7-cell network.

PERSPECTIVES

- Cooperation and coordination a powerful weapon against interference
- Cooperation performance is traded-off for data exchange overhead
- Asymptotic regimes (users, antennas) can lead to simple fully distributed schemes
- Some interesting open problems yet to solve:
  - Just how much information of any given channel/user data is needed at each cooperating device?
  - Robust precoding designs needed to cope with partial information sharing
  - Architecture design: How to best disseminate limited information exchange? (feedback topologies)

THANK YOU!


Let's solve our conflict!
Okay, tell me your point of view
Sorry, but I don't have time to talk!