Device-Centric Cooperation in Wireless Networks

Presented by David Gesbert and Paul de Kerret
gesbert@eurecom.fr, dekerret@eurecom.fr

July 4th, 2016

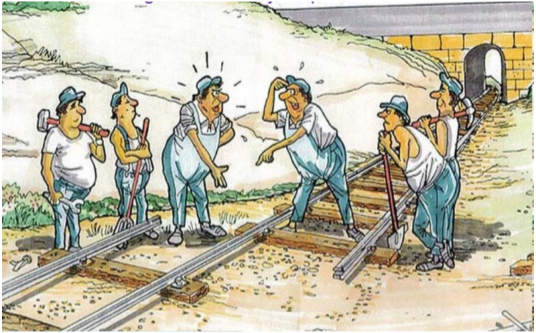
EURECOM
Sophia Antipolis

erc
European Research Council
Established by the European Commission

PERFUME

Coordination vs. cooperation

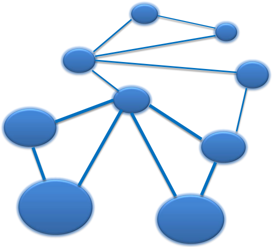
Coordination is a way to resolve complex problems among distributed agents
Can come with a notion of conflict: **coordination** → **cooperation**



Network coordination

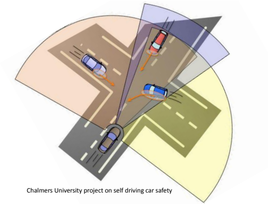
Coordination and cooperation have emerged as central concepts in many types of networks

- Autonomous robots networks
- Transportaton networks
- Sensor networks
- Processor networks
- Energy (Smart Grids) networks
- Wireless networks



Team-playing robots

- Driver-less vehicles
- Autonomous robot patrols
- Plant probes (nuclear sites,..)
- Military drones (ground, air)
- "Smart Factory" robots
- Robot sport teams "Robo-Cup"


Chalmers University project on self driving car safety

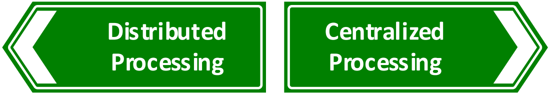
Network coordination is often, by essence, "myopic"

Outline

Outline

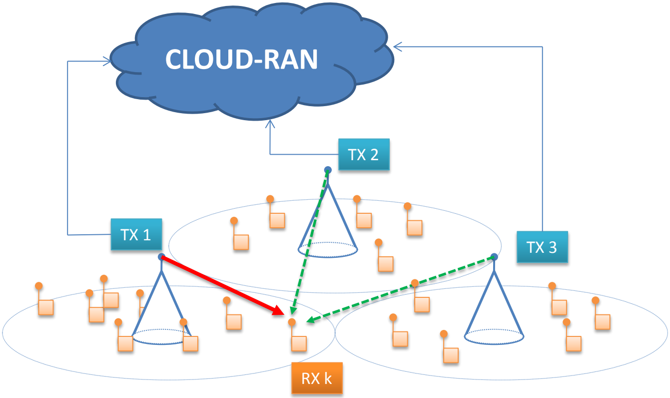
Network/Device cooperation beyond 5G

Where are we going?



Why beyond 5G may be "centralized"

- Cloud RAN is very popular, pushes for more centralization
- Centralized decision making is conceptually simple and efficient
- Coordination, coperation is easy
- Mobile service providers love it

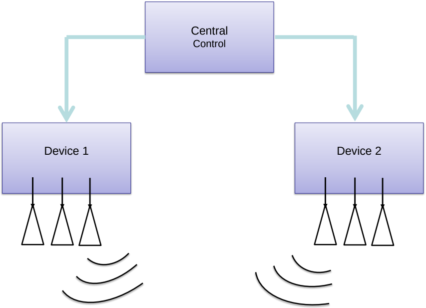


Enhancing spectral efficiency via coordination

Recent spectrum efficiency gains (or promise) from

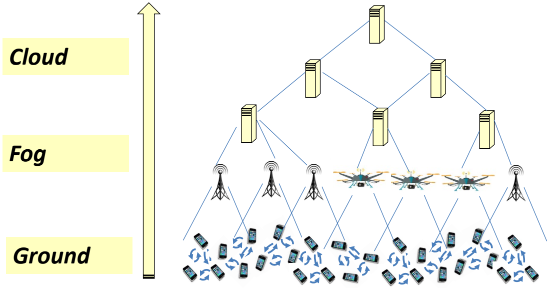
- MU-MIMO, Network MIMO (CoMP), Massive MIMO
- Dynamic cell clustering
- Beamforming
- Power control
- Channel aware scheduling
- Spectrum sharing

All made easy in centralized settings



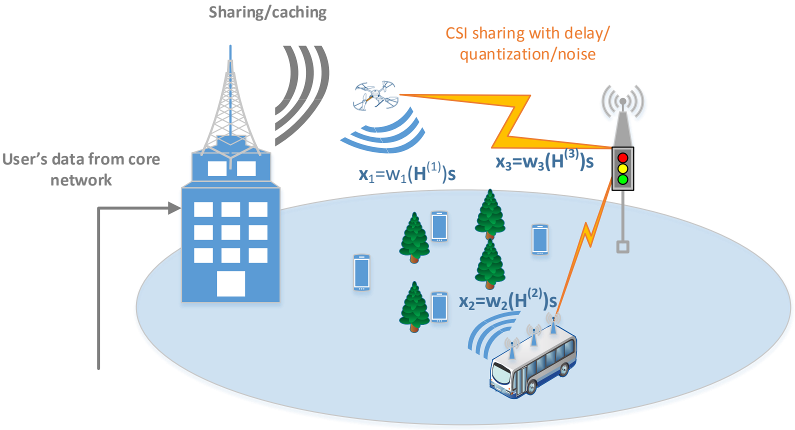
Why beyond 5G may be partly "decentralized"

- Centralization leads to expensive architectures
- Curse of dimension (IoT: billions of devices)
- Centralized processing increases latency, killer for the tactile internet.
- Wireless backhaul architectures are often heterogeneous

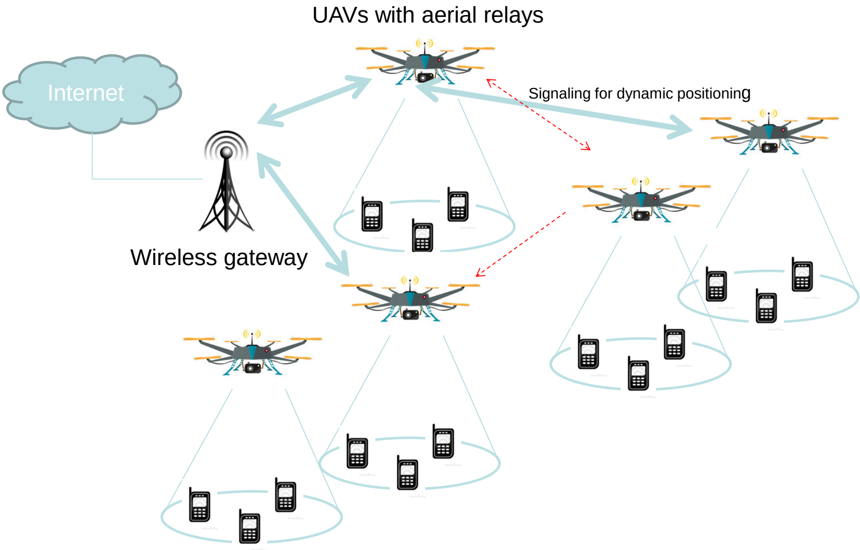


Centralization ↑ Backhaul ↑ latency ↑ timeliness ↓

Cooperation in heterogenous Wireless networks



Ultra flexible Wireless networks



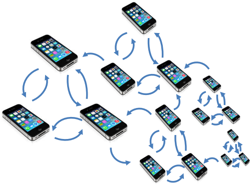
Device-centric Cooperation

Potential:

1. Many devices with substantial sensing/computing capabilities (phones, tablets, vehicles, drones, pico-BS..)
2. Huge collective intelligence
3. Local processing makes time-sensitive measurements more relevant

Challenges

1. How to model distributed information settings?
2. Is there a price of distributedness?
3. Are there robust approaches?



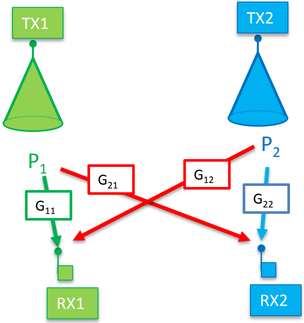
Ex 1: Power control over interference channels

- Two interfering devices, with interference channels $G_{i,j}, i = 1, 2, j = 1, 2$
- Transmit with binary power control is sum-rate optimal [Gjendemsjo et al., 2008, TWC]

$$(p_1^\star, p_2^\star) = \underset{(p_1,p_2)\in\mathcal{P}}{\operatorname{argmax}} \left[R(p_1(\{G_{i,j}\}), p_2(\{G_{i,j}\}))\right]$$

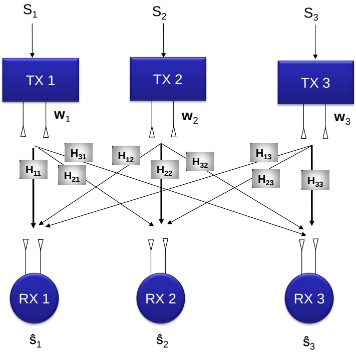
where

$$\mathcal{P} \triangleq \{(p_1, p_2) | p_j : \mathbb{R}^4 \to \{0, P_j^{\max}\}, j = 1, 2\}.$$



Hence the coordinated choice of "full power" or "stay silent" for each device requires full centralized CSI. What if not the case?

Ex 2: Interference Alignment



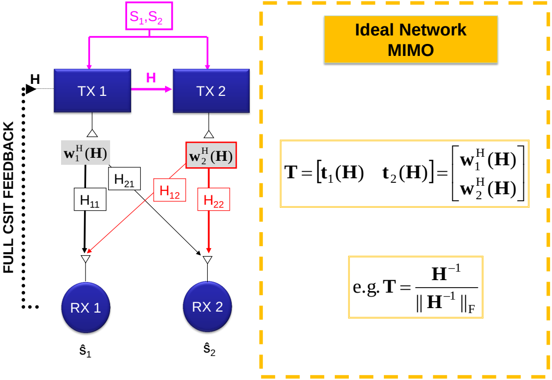
Interference Alignment Conditions with N_r=2:

$$\mathbf{H}_{12}\mathbf{w}_2 \propto \mathbf{H}_{13}\mathbf{w}_3$$
$$\mathbf{H}_{21}\mathbf{w}_1 \propto \mathbf{H}_{23}\mathbf{w}_3$$
$$\mathbf{H}_{31}\mathbf{w}_1 \propto \mathbf{H}_{32}\mathbf{w}_2$$

Alignment can be carried out in space, frequency, time domains. [Maddah-Ali et al., 2008, TIT][Cadambe and Jafar, 2008, TIT]

Realization of alignment conditions requires knowledge of all matrices $\mathbf{H}_{i,j}$ at all transmitters. What if not the case?

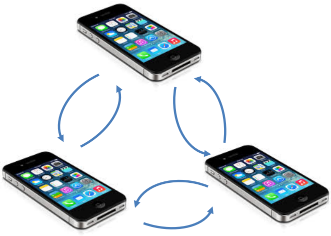
Ex 3: Network MIMO



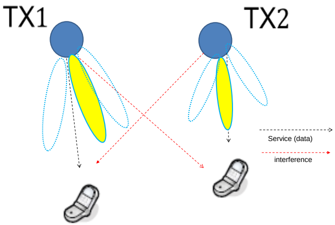
Network MIMO requires full knowledge of global **H** matrix (and data symbols) at all transmitters. What if not the case?

Ex 4: Distributed caching

- D2D can be leveraged for content sharing and caching among terminals [Golrezaei et al., 2014, TIT]
- Popular files can be cached in device memory for later use. Each device can store K files.
- N ideally close-by devices can coordinate to cache non overlapping subsets of K files, hence making the NK most popular files available in their vicinity.
- This requires full information exchange. What if not the case?



Ex 5: Coordinated beamforming/scheduling



- Each transmitter should design a beamforming vector $\mathbf{w}_i, i = 1, 2$
- The best beamformer choice strikes a optimal trade-off between matched filter (egoistic) solution and interference zero-forcing (altruistic) solution [Jorswieck et al., 2008, TSP]
- Optimal design based on knowledge of all direct and interference channel gains. what if not the case?

Ex 6: Cell coloring/clustering

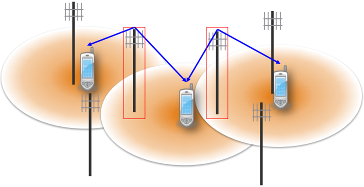


Figure: From Park, Lee, Heath, "Cooperative Base Station Coloring for Pair-wise Multi-Cell Coordination", arXiv March 2015

- Given a limited cooperation cluster size, cells can coordinate with other to design optimal clusters
- Clustering algorithms are usually centralized. But what if cells should attach to a cluster based on local CSI? (i.e. local user gains, local interference gains)
- Decentralized (heuristic) algorithm proposed in [Park et al., 2015]

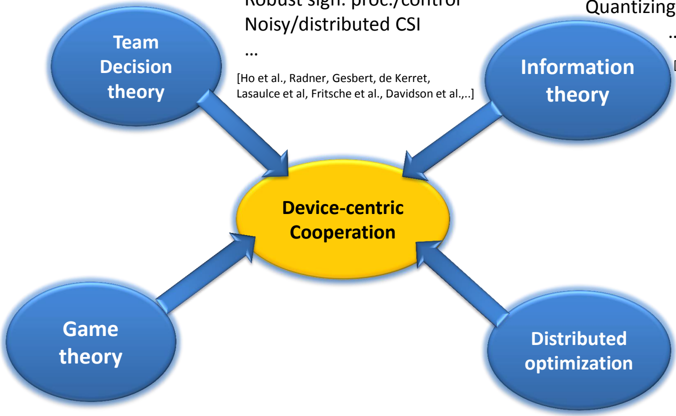
Device coordination: The many perspectives

One-shot decision
Robust sign. proc./control
Noisy/distributed CSI
...

[Ho et al., Radner, Gesbert, de Kerret,
Lasaulce et al, Fritsche et al., Davidson et al.,..]

Capacity/DoF analysis
Coordination theory
Quantizing with side info.
...

[Larousse et al,Cuff et al,
Li et al, Grover, ...]



Study of equilibria
Selfish behavior
Convergence studies
...

[Saad et al, Han et al, McKenzie et al, Lasaulce et al,
Poor et al., Rose et al., Jorswieck et al.,..]

Complexity/Convergence studies
Consensus algorithms
Delay tolerant applications
...

[Boyd et al, Inalhan et al, Colorni et al, Rabbat et al, Chen et al., Johansson et al.,
Palomar et al., Scaglione et al., Scutari et al.]

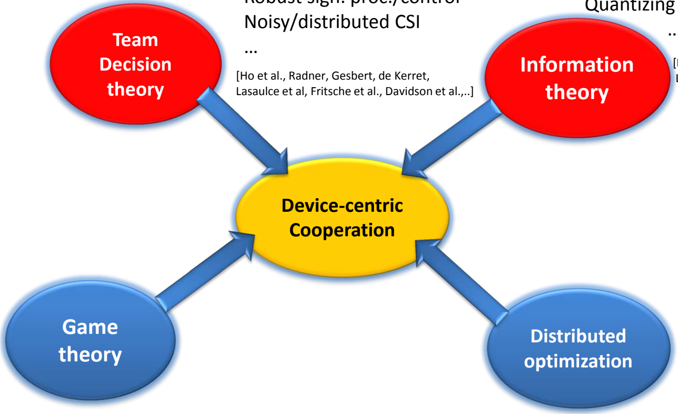
Device coordination: The many perspectives

One-shot decision
Robust sign. proc./control
Noisy/distributed CSI
...

[Ho et al., Radner, Gesbert, de Kerret, Lasaulce et al, Fritsche et al., Davidson et al.,..]

Capacity/DoF analysis
Coordination theory
Quantizing with side info.
...

[Larousse et al,Cuff et al, Li et al, Grover, ...]



Study of equilibria
Selfish behavior
Convergence studies
...

[Saad et al, Han et al, McKenzie et al, Lasaulce et al, Poor et al., Rose et al., Jorswieck et al.,..]

Complexity/Convergence studies
Consensus algorithms
Delay tolerant applications
...

[Boyd et al, Inalhan et al, Colorni et al, Rabbat et al, Chen et al., Johansson et al., Palomar et al., Scaglione et al., Scutari et al.]

Outline

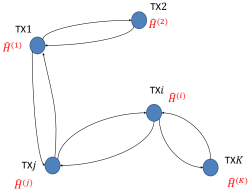
Distributed Information Models

Wireless Channel State Information (CSI) is by nature noisy and distributed

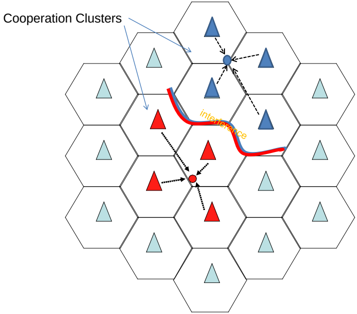
- Limited sensing and feedback
- Mobility
- Devices tend to be "myopic": They know better what is close
- CSI exchange is not free
- Devices do not need to know CSI for entire network

CSI is often transmitter dependent

→ "*Information Structure*"



Information stucture: Clustering



Approaches:

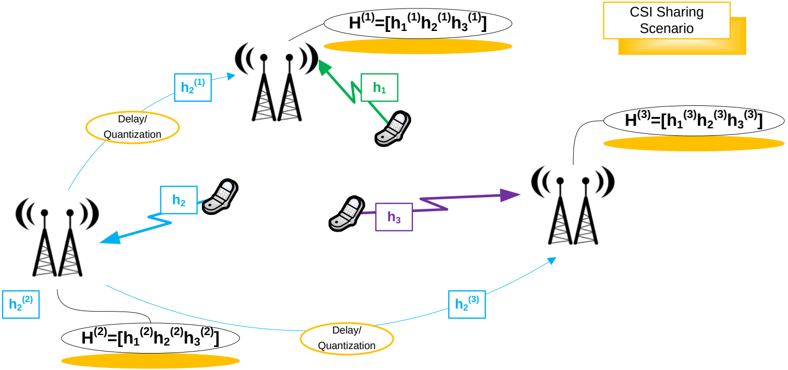
- Network-centric clustering
- User-centric clustering [Papadogiannis et al., 2008, ICC]

Limitations:

- Cluster too big: feedback sharing overhead heavy [Lozano et al., 2013, TIT]
- Cluster too small: edge-effects (inter-cluster interference) predominant

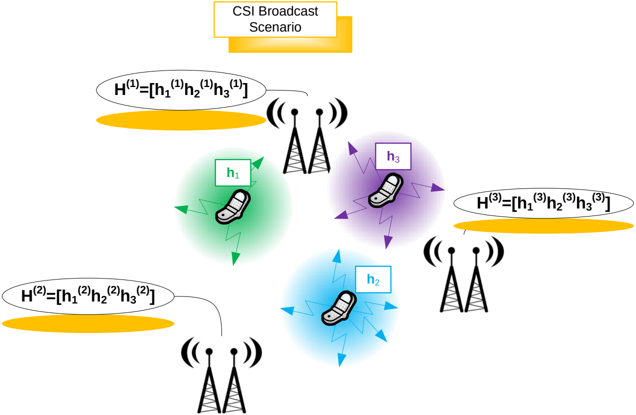
CSI information structure: LTE with limited backhaul

- Backhaul signaling introduces delays and possible quantization noise
- LTE compliant feedback: User feeds back to its home eNB only



CSI information structure: Feedback Broadcast

- CSIT can be shared directly over-the-air without backhaul links



Classical noisy CSI model (centralized)

- Every transmitter shares the same noisy channel estimate
- Imperfect (quantized, noisy, delayed,..) CSIT at TX modeled as [Wagner et al., 2012, TIT]

$$\{\hat{\mathbf{H}}\}_{i,k} = \sqrt{1-\sigma_{i,k}^2}\{\tilde{\mathbf{H}}\}_{i,k} + \sigma_{i,k}\{\mathbf{\Delta}\}_{i,k}, \qquad \forall i,k$$

where $\{\mathbf{\Delta}\}_{ik} \sim \mathcal{CN}(0,1)$

- With digital quantization $\sigma_{i,k}^2 = 2^{-B_{i,k}}$ (good approximation in the high resolution regime)
- CSIT allocation matrix $\mathbf{B}$ defined as

$$\{\mathbf{B}\}_{i,k} = B_{i,k}, \qquad \forall i,k$$

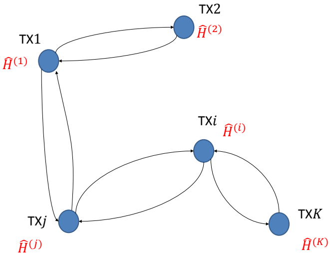
Distributed CSI Model

- CSIT is transmitter-dependent
- LOCAL CSIT at TX j modeled as

$$\{\hat{\mathbf{H}}^{(j)}\}_{i,k} = \sqrt{1-(\sigma_{i,k}^{(j)})^2}\{\mathbf{H}\}_{i,k} + \sigma_{i,k}^{(j)}\{\boldsymbol{\Delta}\}_{i,k}^{(j)}, \qquad \forall i,k$$

where $\{\boldsymbol{\Delta}\}_{i,k}^{(j)} \sim \mathcal{CN}(0,1)$

- $\sigma_{i,k}^{(j)}$ indicates quality of CSIT for channel element (i,k) at TX j



Distributed CSI structure models

Some useful particular cases:

- A CSI structure is *perfect* if $\hat{\mathbf{H}}^{(i)} = \mathbf{H}$, $\forall i$.
- A CSI structure is *centralized* if $\hat{\mathbf{H}}^{(i)} = \hat{\mathbf{H}}^{(j)}$, $\forall i, j$.
- A CSI structure is *distributed* if there exist i and j such that $\hat{\mathbf{H}}^{(i)} \neq \hat{\mathbf{H}}^{(j)}$.

Distributed CSI structure models (cont'd)

Some more particular cases:

- **Incomplete CSIT**: A CSI structure is *incomplete* if $\hat{\mathbf{H}}^{(i)}$ takes the form $\forall i\ \hat{\mathbf{H}}^{(i)} = \{\mathbf{H}_{k,l}, k \in \mathcal{S}_{\mathrm{TX}}, l \in \mathcal{S}_{\mathrm{RX}}\}$, where $\mathcal{S}_{\mathrm{TX}}$ (resp. $\mathcal{S}_{\mathrm{RX}}$) are subsets of the transmitter set (resp. receiver set).
- **Hierarchical CSIT**: A CSI structure is *hierarchical* if there exists an order of transmitter indices i_1, i_2, i_3.. such that $\hat{\mathbf{H}}^{(i_1)} \subset \hat{\mathbf{H}}^{(i_2)} \subset \hat{\mathbf{H}}^{(i_3)} \subset \ldots$
- **Master Slave**: Hierarchical where $\hat{\mathbf{H}}^{(i_1)} = [\ ]$, and $\hat{\mathbf{H}}^{(i_2)} = \mathbf{H}$ (can be extended to $K > 2$.)

Typical (practical) CSI structures

Consider the K transmitter (N antennas each) K user (single antenna) channel. Let $\boldsymbol{h}_{i,j}^{\mathrm{H}}$ be the $1 \times N$ vector channel between the jth transmitter and the ith user.

- Local CSIT with TDD reciprocity

$$(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}} = \begin{bmatrix} \mathbf{0} & \boldsymbol{h}_{1,j}^{\mathrm{H}} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \boldsymbol{h}_{K,j}^{\mathrm{H}} & \mathbf{0} \end{bmatrix}$$

- Local CSIT with LTE feedback mode

$$(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{h}_{j,1}^{\mathrm{H}} & \dots & \boldsymbol{h}_{j,K}^{\mathrm{H}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

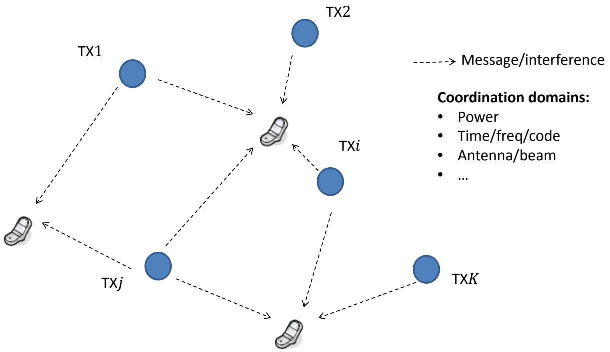
- Fully local CSIT

$$(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{h}_{j,j}^{\mathrm{H}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Outline

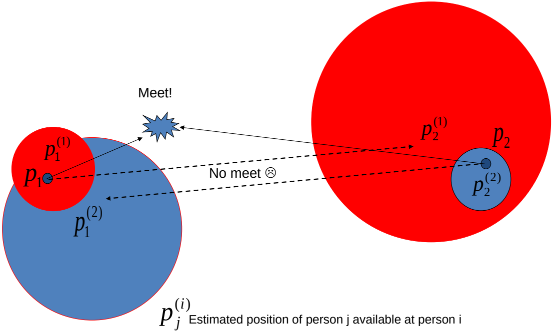
Device coordination Problem

- K nodes in a network seek to cooperate towards the maximization of a common utility
- Each node i must make best decision based on:
    - local measurement or feedback
    - finite rate signaling with neighbor nodes

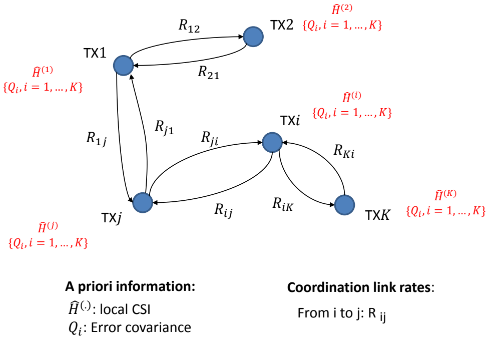


TD example: The Distributed Rendez-vous Problem

- Two visitors arrive independently in Edinburgh 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*



Coordination over finite communication graphs: The big picture



- No constraint over number bits exchanged: Distributed optimization $\rightarrow$ convergence speed?
- Constraint over number of bits exchanged: What to measure? What is the most relevant information to communicate among devices?
- Decision stage (after limited communication took place): What are robust coordinated decision techniques?
- Joint communication-decision framework (challenging)

Signaling for Coordination

What is most relevant to communicate of the signaling link?

- Many interesting heuristics (precoding decisions, measurements, etc.)
- Optimal signaling strategy coupled with optimum decision making W_i
- Heuristic strategies:
  1. Local decision W_i based on $\hat{\mathbf{H}}^{(i)}$ and $\mathbf{Q}_i, i = 1, .., K$, exchange quantized decisions over R_{ij} bits
     - But poorly informed nodes make bad decisions !
  2. Exchange quantized CSI $\hat{\mathbf{H}}^{(i)}$ over R_{ij} bits
     - But this ignores $\mathbf{Q}_i$!
- Optimal strategy (source coding with side-information): Create locally optimal codebooks, that are function of local CSI and neighbor CSI qualities [Li et al., 2014]

Distributed coordination

Team Decision theoretic problems:

- 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...
- Introduced first in economics and control [Witsenhausen68] [Ho, 1980, IEEE], recently in wireless [Zakhour and Gesbert, 2010, ITA]
- Fundamental limits rooted in Coordination Theory

Coordination Theory

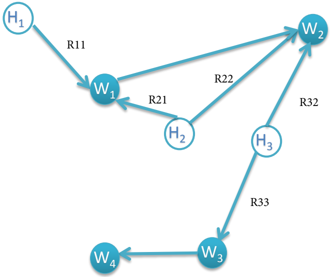


Figure: Coordination Framework[Cuff et al., 2010, TIT]

- $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$, arbitrary components of global system state, distributed according to $p_0(\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3)$
- W_1, W_2 and W_3 are actions selected by the nodes.
- What joint distribution $p_0(\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3)p(W_1, W_2, W_3|\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3)$ can be achieved?
- Answer: it depends on graph topology (capacity of each edge)

Example (One Isolated Node)

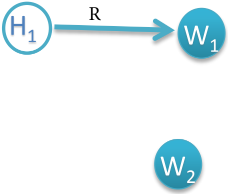


Figure: One isolated node scenario [Cuff et al., 2010, TIT]

Theorem

$$\mathcal{C}_{p_0} = \left\{ (\overbrace{R}^{rate}, \overbrace{p(W_1, W_2 | \mathbf{H}_1)}^{distribution}) \middle| R \geq I(\mathbf{H}_1; W_1 \middle| W_2) \right\}$$

Example (One Isolated Node)

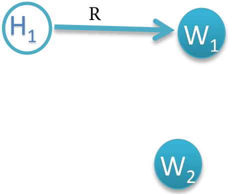


Figure: One isolated node scenario [Cuff et al., 2010, TIT]

- With Gaussian RVs, the condition becomes

$$(1 - 2^{-2R})^{-1}\rho^2_{H_1,W_1} + \rho^2_{W_1,W_2} \leq 1$$

- $R \to \infty$, $\rho^2_{H_1,W_1} + \rho^2_{W_1,W_2} \leq 1$
- $R \to 0$, $\rho^2_{H_1,W_1} = 0$ and $\rho^2_{W_1,W_2} \leq 1$

Further Results

- Results in more advanced topologies [Cuff et al., 2010, TIT]
- Polar codes used for coordination in [Chou et al., 2015, ISIT]
- Implicit coordination: Observation of action of one node by another is a non dedicated cooperation link
  - ➥ Coordination at low/no cost [Larrousse and Lasaulce, 2013, ISIT]
- Aim of this approach
  - Guidelines for network design
  - Insights for new cooperation methods

Team Decision (TD) Problems: A general formalism

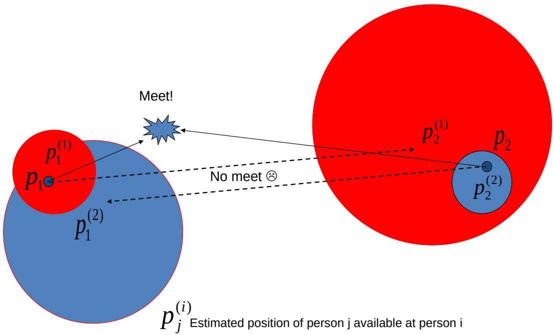
$$(\boldsymbol{s}_1^\star, \dots, \boldsymbol{s}_K^\star) = \underset{\boldsymbol{s}_1,\dots,\boldsymbol{s}_K}{\operatorname{argmax}} \, \mathbb{E}_{\boldsymbol{x},\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(K)}} \left[f\left(\boldsymbol{x}, \boldsymbol{s}_1(\boldsymbol{x}^{(1)}), \dots, \boldsymbol{s}_K(\boldsymbol{x}^{(K)})\right)\right]$$

where

- K: Number of Decision Makers (DMs)
- $\boldsymbol{x} \in \mathbb{C}^m$: State of the world
- $\boldsymbol{x}^{(j)} \in \mathbb{C}^m$: Estimate of the state of the world $\boldsymbol{x}$ at DM j
- $\boldsymbol{s}_j : \mathbb{C}^m \to \mathcal{A}_j \subset \mathbb{C}^{d_j}$: Strategy of the j-th DM
- $\boldsymbol{s}_j(\boldsymbol{x}^{(j)}) \in \mathcal{A}_j \subset \mathbb{C}^{d_j}$: Decision at DM j for the given realization $\boldsymbol{x}^{(j)}$
- $f : \mathbb{C}^m \times \Pi_{j=1}^K \mathbb{C}^{d_j} \to \mathbb{R}$: Joint objective of the K DMs
- $p_{\boldsymbol{x},\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(K)}}$: Joint probability distribution of the channel and the estimates

TD example: The Distributed Rendez-vous Problem

- Two visitors arrive independently in Edinburgh 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*



A robust solution: "Meet you at the City Hall!"

Can Team Problems be Solved with Games?

Key idea: Let autonomous transmitting devices interact to solve their interference conflicts
Players $\rightarrow$ transmitters
Actions $\rightarrow$ transmit decision (power, frequency, beam, ..)
Strategy $\rightarrow$ Utility maximization (max rate, min power, min delay,..)
Timing $\rightarrow$ simultaneous, sequential,..
Equilibrium $\rightarrow$ Nash, Stackelberg, Nash Bargaining,..

From Selfish Games to "Team Playing"

Why interference coordination can be different from a typical "game":

- Team agents (network nodes) are not conflicting players (different from players in a cooperative game)
- Agents seek maximization of the same network utility
- It is the lack of shared information which hinders cooperation, not the selfish of their interests
- Agents are not required to improve over the performance of the Nash equilibrium
- Connections to Bayesian games (see work by 1994 Nobel Prize winner John Harsanyi [Harsanyi, 1967])

A Fundamental Approach: Best Response

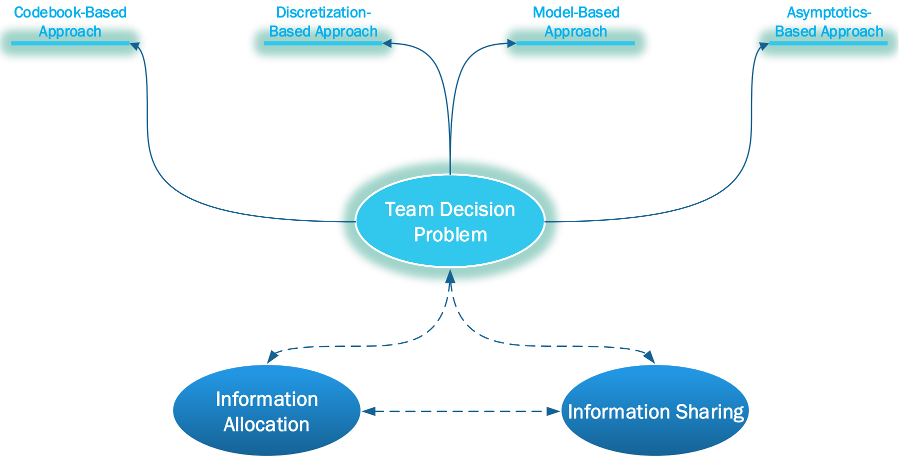
Best Response

A *Best-Response (BR)* strategy $\boldsymbol{s}_1^{\mathrm{BR}}, \dots, \boldsymbol{s}_K^{\mathrm{BR}}$ for the TD problem is a strategy such that

$$\boldsymbol{s}_j^{\mathrm{BR}} = \operatorname*{argmax}_{\boldsymbol{s}_j \in \mathcal{A}_j} \mathbb{E}_{\boldsymbol{x},\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(K)}|\boldsymbol{x}^{(j)}} \left[f\left(\boldsymbol{x}, \dots, \boldsymbol{s}_{j-1}^{\mathrm{BR}}(\boldsymbol{x}^{(j-1)}), \boldsymbol{s}_j(\boldsymbol{x}^{(j)}), \boldsymbol{s}_{j+1}^{\mathrm{BR}}(\boldsymbol{x}^{(j+1)}), \dots \right) \right], \quad \forall j$$

- Practical approach usually considered in the TD literature
- Still very challenging:
  - Functional optimization
  - Stochastic optimization
  - Channel space of large dimension (in most of the cases)
- In fact, Bayesian Cooperative Game with Incomplete Information [Harsanyi, 1967, Management Science]

Team Decision: Algorithm design



Model-Based Approach

Main idea: Restrict the space of possible strategies via a model

➥ Replace the strategy s_j by s_j^β with $\beta \in \mathbb{R}$ where s_j^β is a well chosen heuristic model

Example (Coordinated Beamforming [Jorswieck et al., 2008, TSP])

Beamformer in the MISO IC parameterized as

$$\boldsymbol{w}_k^\star(\lambda_k) = \frac{\lambda_k \boldsymbol{w}_k^{\mathsf{ZF}} + (1-\lambda_k)\boldsymbol{w}_k^{\mathrm{MF}}}{\|\lambda_k \boldsymbol{w}_k^{\mathsf{ZF}} + (1-\lambda_k)\boldsymbol{w}_k^{\mathrm{MF}}\|}$$

Model-based Team Decision Buying a Baguette or not?

- A french couple returns separately from work and wants baguette for dinner. their phone batteries are empty
- Personal cost for stopping at the baker is c_i .
- Each person knows its own cost c_i
- The c_i are uniformly distributed over $[0, 1]$.

Goal: maximize expectation of joint utility given by:

Person 2\Person 1	***Buy bread***	***Go home***
Buy bread	*a-c_1-c_2*	*1-c_1*
Go home	*1-c_2*	*0*

When should each person buy bread?

Optimal decision $\gamma_i^*(c_i)$ *of threshold form*

$$\boxed{\gamma_i^*(c_i) = \begin{cases} \text{Buy bread if } c_i \leq c_i^{th} \\ \text{Go home if } c_i > c_i^{th} \end{cases}}$$

Codebook-Based Approach

Main idea: Restrict the space of possible strategies to a codebook

➥ Choose s_j inside a codebook of function $\{s_j^1, \dots, s_j^m\}$

Example (Coordinated Beamforming)

- Restrict possible beamforming choices to $\mathcal{C} = \{\text{Matched Filter}, \text{Zero Forcing}\}$

Discretization-Based (1): Dimensionality Reduction

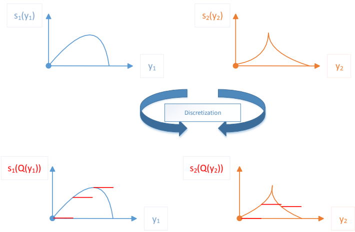
Main idea: Quantize the channel state space to reduce the dimension

➥ Replace the strategy s_j by $s_j(Q^{\mathrm{cb}})$ where

$$\begin{array}{rccc} Q^{\mathrm{cb}}: & \mathbb{C}^m & \to & \mathcal{C}^{\mathrm{cb}} \triangleq \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n^{\mathrm{cb}}}\} \\ & \boldsymbol{x}^{(j)} & \mapsto & Q^{\mathrm{cb}}(\boldsymbol{x}^{(j)}) = \operatorname{argmin}_{\boldsymbol{x} \in \mathcal{C}^{\mathrm{cb}}} \|\boldsymbol{x} - \boldsymbol{x}^{(j)}\|^2 \end{array}$$

- Optimization subspace reduced to a space of dimension n^{cb}:

$$\begin{array}{rccc} s_j: & \mathcal{C}^{\mathrm{cb}} & \to & \mathcal{A}_j \\ & \boldsymbol{x}_i & \mapsto & s_j(\boldsymbol{x}_i) \end{array}$$



Discretization-Based (2): Monte-Carlo Approximation

- Best-response optimization at DM 1: $\forall i_1 \in \{1, \dots, n^{\mathrm{cb}}\}$,

$$s_1^{\mathsf{BR}}(\boldsymbol{x}_{i_1}) = \underset{S_1 \in \mathcal{A}_1 \subset \mathbb{C}^{d_1}}{\operatorname{argmax}} \; \mathbb{E}\left[f(\boldsymbol{x}, S_1, s_2^{\mathsf{BR}}(\mathcal{Q}^{\mathrm{cb}}(\boldsymbol{x}^{(2)})), \dots, s_K^{\mathsf{BR}}(\mathcal{Q}^{\mathrm{cb}}(\boldsymbol{x}^{(K)}))) \middle| \boldsymbol{x}^{(1)} = \boldsymbol{x}_{i_1} \right]$$

- For a given $\boldsymbol{x}_i$ and given $s_2^{\mathsf{BR}}, \dots, s_K^{\mathsf{BR}}$: Standard stochastic optimization problem [Shapiro et al., 2014]
- Use Monte-Carlo approximation: $\forall i \in \{1, \dots, n^{\mathrm{cb}}\}$,

$$s_1^{\mathsf{BR}}(\boldsymbol{x}_i) = \underset{\mathbf{s}_1 \in \mathcal{A}_1}{\operatorname{argmax}} \frac{1}{n^{\mathrm{MC}}} \sum_{\ell=1}^{n^{\mathrm{MC}}} \mathrm{f}\left(\boldsymbol{x}_\ell, \mathbf{S}_1, \boldsymbol{s}_2(\mathcal{Q}(\boldsymbol{x}_\ell^{(2)})), \dots, \boldsymbol{s}_K(\mathcal{Q}(\boldsymbol{x}_\ell^{(K)}))\right)$$

where $\left(\boldsymbol{x}_\ell, \boldsymbol{x}_\ell^{(2)}, \dots, \boldsymbol{x}_\ell^{(K)}\right) \sim p_{\boldsymbol{x}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(K)} \,|\, \boldsymbol{x}^{(1)} = \boldsymbol{x}_i}$

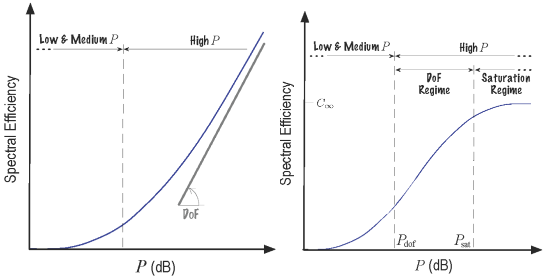
Asymptotics-Based Approach

Main idea: use asymptotic analysis to make the problem deterministic

➥ Possible to obtain new insights and transmission strategies

Example (DoF Analysis)

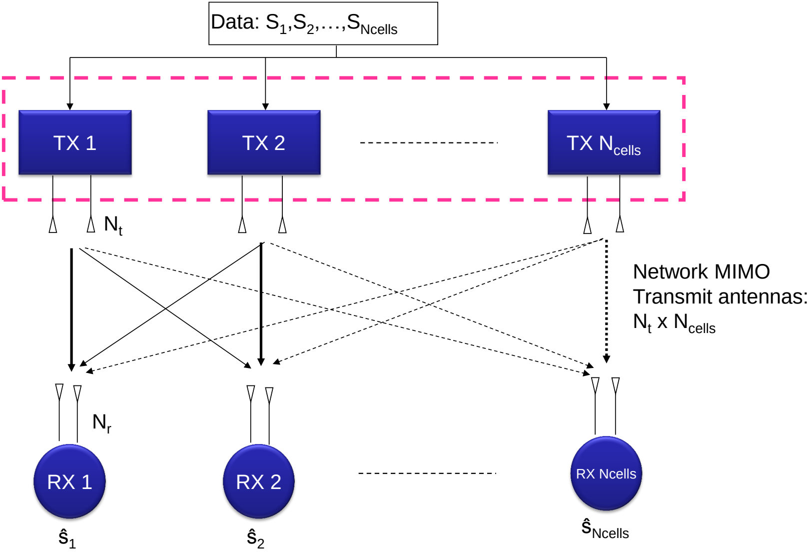
- Let the transmit SNR goes to infinity



* A. Lozano et al, "Fundamental limits of cooperation", IEEE Trans. On Information Theory, Sept. 2013.

Outline

Joint Precoding over Network MIMO



Team Decision Problem

$$(\boldsymbol{w}_1^\star, \ldots, \boldsymbol{w}_K^\star) = \underset{(p_1,\ldots,p_K)\in\mathcal{P}}{\operatorname{argmax}} \mathbb{E}[\mathrm{R}(\mathbf{H}, \boldsymbol{w}_1(\hat{\mathbf{H}}^{(1)}), \ldots, \boldsymbol{w}_K(\hat{\mathbf{H}}^{(K)}))]$$

where

$$\mathrm{R}(\mathbf{H}, \boldsymbol{w}_1(\hat{\mathbf{H}}^{(1)}), \ldots, \boldsymbol{w}_K(\hat{\mathbf{H}}^{(K)})) = \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_{d_k} + \mathbf{T}_k^{\mathrm{H}} \mathbf{H}_k^{\mathrm{H}} \left(\mathbf{R}_k + \sum_{i\neq k} \mathbf{H}_i \mathbf{T}_i \mathbf{T}_i^{\mathrm{H}} \mathbf{T}_i^{\mathrm{H}} \right)^{-1} \mathbf{H}_k \mathbf{T}_k \right|$$

with

- $\mathbf{H} \in \mathbb{C}^{N_{\text{tot}} \times M_{\text{tot}}}$ the multi-user channel
- $\boldsymbol{w}_j$ the precoding function:

$$\begin{aligned} \boldsymbol{w}_j : \quad \mathbb{C}^{N_{\text{tot}} \times M_{\text{tot}}} &\rightarrow \mathbb{C}^{M_j \times d_{\text{tot}}} \\ \hat{\mathbf{H}}^{(j)} &\mapsto \boldsymbol{w}_j(\hat{\mathbf{H}}^{(j)}) \end{aligned}$$

- $\mathbf{T} \in \mathbb{C}^{M_{\text{tot}} \times d_{\text{tot}}}$ the multi-user precoder

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & \ldots & \mathbf{T}_K \end{bmatrix} = \begin{bmatrix} \boldsymbol{w}_1(\hat{\mathbf{H}}^{(1)}) \\ \vdots \\ \boldsymbol{w}_K(\hat{\mathbf{H}}^{(K)}) \end{bmatrix}$$

A Key Example

- Particularly interesting because:
  - Continuous optimization with large channel state dimension
  - Strong dependency (state & TXs): see DoF results
- ➥ Many difficulties

Table: Team Decision Modeling for Joint Precoding

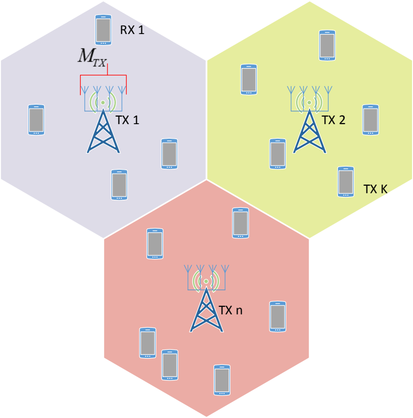
Notations for the Team Decision Problems		
State-of-the-world	$\boldsymbol{x}$	$\mathbf{H}$
Estimate at DM j	$\boldsymbol{x}^{(j)}$	$\hat{\mathbf{H}}^{(j)}$
Strategy at DM j	$\boldsymbol{s}_j$	$\boldsymbol{w}_j$
Decision space at DM j	$\mathcal{A}_j$	$\mathbb{C}^{M_j \times d_{\text{tot}}}$
Objective	f	R

Outline

A Result Based on Random Matrix Theory (RMT)

- n cooperating TXs
- Each TX has M_{TX} antennas
- K and M_{TX} grow large at the same rate

$$\beta \triangleq \lim_{M,K\to\infty} \frac{M}{K} \triangleq \lim_{M_{TX},K\to\infty} \frac{nM_{\mathrm{TX}}}{K} \geq 1$$



Distributed CSI with Correlated Noise

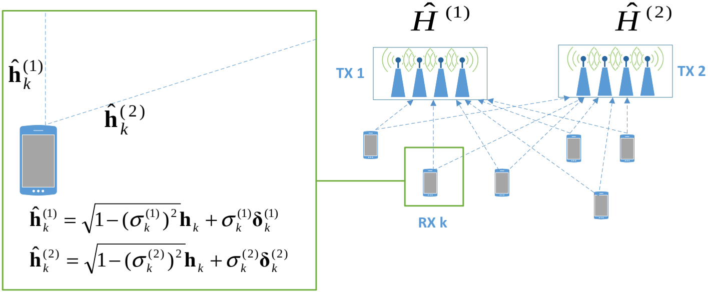
- Extend to spatial correlation in the CSI noise

$$\hat{\boldsymbol{h}}_k^{(j)} \triangleq \sqrt{1-(\sigma_k^{(j)})^2}\boldsymbol{h}_k + \sigma_k^{(j)}\boldsymbol{\delta}_k^{(j)}$$

with

$$\mathbb{E}\left[\boldsymbol{\delta}_k^{(j)}(\boldsymbol{\delta}_k^{(j')})^{\mathrm{H}}\right] = (\rho_k^{(j,j')})^2\mathbf{I}_M$$

- Extremely general model: Bridges the Gap from distributed CSIT to centralized CSIT: Can model partially centralized settings



A Practical Example

Example

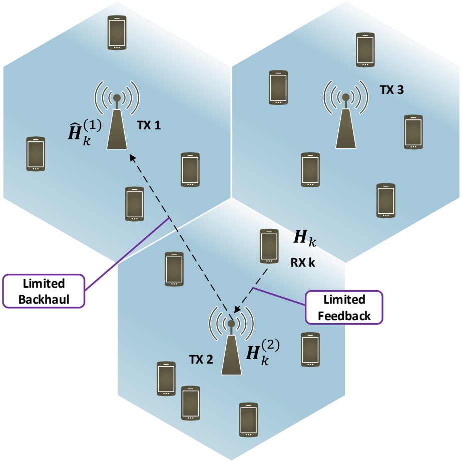
- Imperfect feedback

$$\hat{\boldsymbol{h}}_j^{(j)} = \sqrt{1-\sigma_{FB}{}^2}\boldsymbol{h}_j + \sigma_{FB}\boldsymbol{\delta}_j^{(j)}$$

- Imperfect backhaul

$$\hat{\boldsymbol{h}}_k^{(j')} = \sqrt{1-\sigma_{BH}{}^2}\hat{\boldsymbol{h}}_k^{(j)} + \sigma_{BH}\boldsymbol{\epsilon}_k^{(j,j')}$$

➥ CSI estimates error at different TXs are correlated



Model-Based Approach: Regularized ZF

- Modelization of the precoding decisions using Regularized ZF with sum power constraint P

$$\mathbf{T}_{\text{rZF}}^{(j)}\left(\gamma^{(j)}\right) \triangleq \left((\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\hat{\mathbf{H}}^{(j)} + M\gamma^{(j)}\mathbf{I}_M\right)^{-1}(\hat{\mathbf{H}}^{(j)})^{\mathrm{H}}\frac{\sqrt{P}}{\sqrt{\Psi^{(j)}}}$$

  with $\Psi^{(j)}$ the power normalization at TX j, and

$$\boldsymbol{w}_j(\hat{\mathbf{H}}^{(j)}) = \mathbf{E}_j^{\mathrm{H}}\mathbf{T}_{\text{rZF}}^{(j)}\left(\gamma^{(j)}\right)$$

  Where $\mathbf{E}_j$ is a row selection matrix

- Effective precoder is

$$\mathbf{T}^{\text{DCSI}} \triangleq \begin{bmatrix} \boldsymbol{w}_1(\hat{\mathbf{H}}^{(1)}) \\ \boldsymbol{w}_2(\hat{\mathbf{H}}^{(2)}) \\ \vdots \\ \boldsymbol{w}_n(\hat{\mathbf{H}}^{(n)}) \end{bmatrix}$$

Optimization of the Regularization Parameter

- Naive regularization

$$\gamma^{(j),\mathrm{naive}} = \mathrm{argmax}_{\gamma\in\mathbb{R}}\mathbb{E}[\mathrm{R}(\hat{\mathbf{H}}^{(j)},\ldots,\hat{\mathbf{H}}^{(j)})]$$

- Robust regularization

$$(\gamma^{(1),\star},\ldots,\gamma^{(n),\star}) = \underset{(\gamma^{(1)},\ldots,\gamma^{(n)})}{\mathrm{argmax}}\ \mathbb{E}[\mathrm{R}(\hat{\mathbf{H}}^{(1)},\ldots,\hat{\mathbf{H}}^{(n)})].$$

- Low complexity robust regularization with equal γ at all TXs

$$(\gamma^{\star},\ldots,\gamma^{\star}) = \underset{(\gamma,\ldots,\gamma)}{\mathrm{argmax}}\,\mathbb{E}[\mathrm{R}(\hat{\mathbf{H}}^{(1)},\ldots,\hat{\mathbf{H}}^{(n)})].$$

Main Result (1)

Theorem ([Li et al., 2015, Allerton])

In Joint Processing CoMP with Distributed CSI,

$$\mathrm{SINR}_k - \mathrm{SINR}_k^o \xrightarrow[K, M_{\mathrm{TX}} \to \infty]{a.s.} 0$$

with

$$\mathrm{SINR}_k^o \triangleq \frac{P\left(\frac{1}{n}\sum_{j=1}^{n}\sqrt{\frac{c_{0,k}^{(j)}}{\Gamma_{j,j}^o}}\frac{\delta^{(j)}}{1+\delta^{(j)}}\right)^2}{1+I_k^o}$$

with

$$c_{0,k}^{(j)} \triangleq 1-(\sigma_k^{(j)})^2, \quad c_{1,k}^{(j)} \triangleq (\sigma_k^{(j)})^2, \quad c_{2,k}^{(j)} \triangleq \sigma_k^{(j)}\sqrt{1-(\sigma_k^{(j)})^2}.$$

Main Result (2)

Theorem (continued)

$$I_k^o \triangleq P - P \sum_{j=1}^{n} \sum_{j'=1}^{n} \frac{\Gamma_{j,j'}^o}{\sqrt{\Gamma_{j,j}^o \Gamma_{j',j'}^o}} \left[\frac{2c_{0,k}^{(j)}}{n^2} \frac{\delta^{(j)}}{1+\delta^{(j)}} \frac{\left((\rho_k^{(j,j')})^2 c_{2,k}^{(j)} c_{2,k}^{(j')} + c_{0,k}^{(j)} c_{0,k}^{(j')} \right) \delta^{(j)} \delta^{(j')}}{n^2 \left(1+\delta^{(j)}\right) \left(1+\delta^{(j')}\right)} \right]$$

with

$$\Gamma_{j,j'}^o \triangleq \frac{\frac{1}{M} \sum_{\ell=1}^{K} \sqrt{c_{0,\ell}^{(j)} c_{0,\ell}^{(j')}} + \sqrt{c_{1,\ell}^{(j)} c_{1,\ell}^{(j')}} (\rho_\ell^{(j,j')})^2}{\frac{1+\delta^{(j)}}{\delta^{(j)}} \frac{1+\delta^{(j')}}{\delta^{(j')}} - \frac{1}{M} \sum_{\ell=1}^{K} \left(\sqrt{c_{0,\ell}^{(j)} c_{0,\ell}^{(j')}} + \sqrt{c_{1,\ell}^{(j)} c_{1,\ell}^{(j')}} (\rho_\ell^{(j,j')})^2 \right)^2}.$$

and

$$\delta^{(j)} \triangleq \frac{\beta - 1 - \gamma^{(j)}\beta + \sqrt{(\gamma^{(j)}\beta - \beta + 1)^2 + 4\gamma^{(j)}\beta^2}}{2\gamma^{(j)}\beta}.$$

Sanity Checks (1)

- Imperfect centralized CSIT:

$$\sigma_k^{(j)} = \sigma_k^{(j')} = \sigma_k, \qquad \text{(equal CSIT accuracy)}$$
$$\rho_k^{(j,j')} = 1, \qquad \text{(Full correlation)}$$
$$\gamma^{(j)} = \gamma^{(j')} = \gamma, \qquad \text{(Equal regularization)}$$

➥ Matches with [Wagner et al., 2012, TIT], [Couillet and Debbah, 2011, Theorem 14.1]

$$\mathrm{SINR}_k^{ID-DCSI,o} = \frac{(1-\sigma_k^2)\delta^2}{\Gamma^o\left(1-\sigma_k^2+(1+\delta)^2\sigma_k^2+\frac{(1+\delta)^2}{P}\right)}$$

- Also obtained with $n = 1$

Sanity Checks (2)

- Uncorrelated distributed CSIT with uniform accuracy and equal regularization:

$$\sigma_k^{(j)} = \sigma^{(j)}, \qquad \text{(Uniform CSIT)}$$
$$\rho_k^{(j,j')} = 0, \qquad \text{(Uncorrelated)}$$
$$\gamma^{(j)} = \gamma^{(j')} = \gamma, \qquad \text{(Equal regularization)}$$

➥ Matches with [de Kerret et al., 2015, ISIT]

$$\mathrm{SINR}_k^{EQ-DCSI,o} = \frac{\frac{P}{\Gamma^o}\left(\frac{1}{n}\sum_{j=1}^{n}\sqrt{c_{0,k}^{(j)}}\right)^2 \frac{\delta^2}{(1+\delta)^2}}{I_k^{EQ-DCSI,o}+1}$$

with

$$I_k^{EQ-DCSI} = P - P\sum_{j=1}^{n}\sum_{j'=1}^{n}\frac{\delta\Gamma_{j,j'}^o}{n^2(1+\delta)^2\Gamma^o}\cdot\left[2c_{0,k}^{(j)} + \delta\left(2c_{0,k}^{(j)} - c_{0,k}^{(j)}c_{0,k}^{(j')}\right)\right]$$

Cost of Distributedness

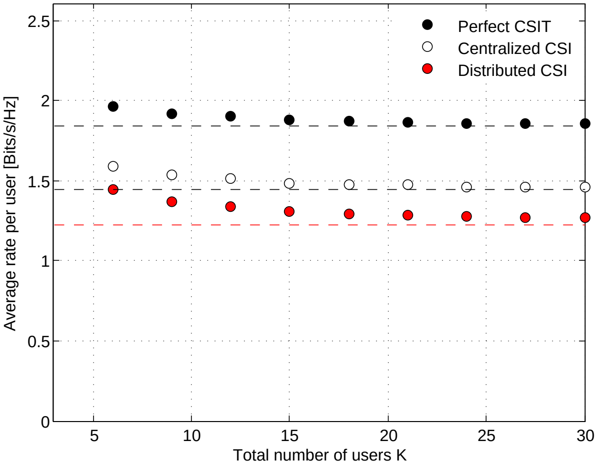


Figure: Average rate per user as a function of the number of users K with $(\sigma^{(j)})^2 = 0.1, \forall j$.

Optimization of the Regularization Parameter

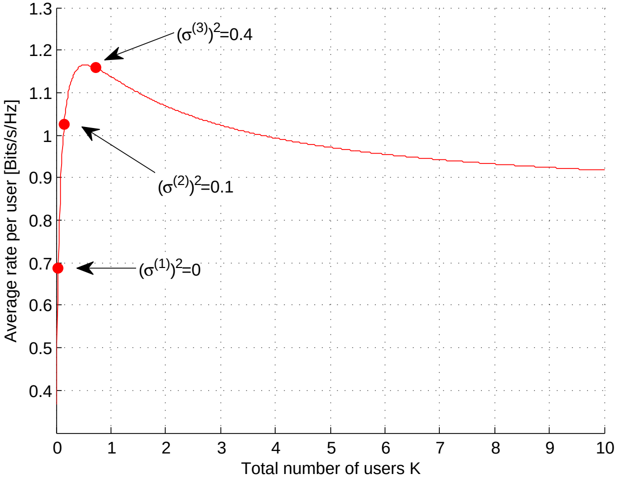
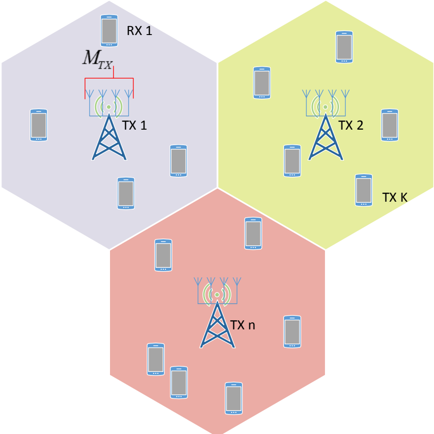


Figure: Average rate per user as a function of γ for $(\sigma^{(1)})^2 = 0, (\sigma^{(2)})^2 = 0.1, (\sigma^{(3)})^2 = 0.4$.

Simulation Settings

n	3	K	30
M	30	β	M/K=1
$(\sigma_k^{(1)})^2$	0.01	$(\sigma_k^{(2)})^2$	0.16
$(\sigma_k^{(3)})^2$	0.49	$\rho_k^{(j,j')}$	0.1
$\boldsymbol{h}_k$	$\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_M)$	$\boldsymbol{\delta}_k^{(j)}$	$\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_M)$



Simulations: Optimize γ

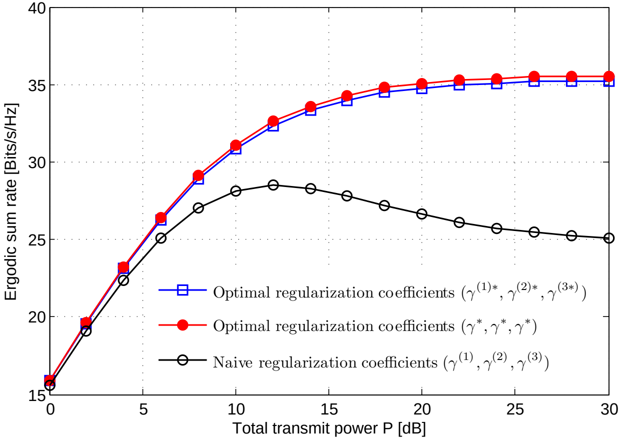


Figure: RZF, ergodic sum rate vs total transmit power P

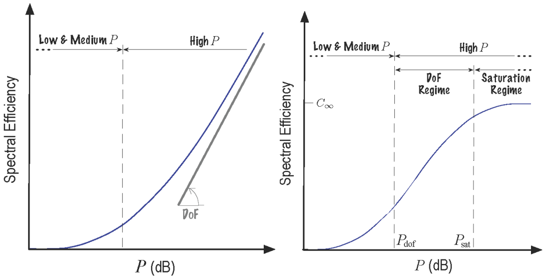
Outline

Asymtotical Analysis: DoF Approach

- First order approximation in the SNR

$$R^{\star} \approx \text{DoF} \log_2(SNR)$$

- Problem becomes deterministic: Possible to obtain analytical results to our complex TD problem

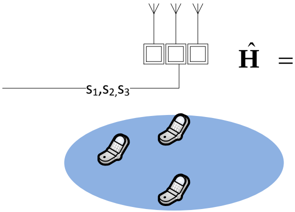


* A. Lozano et al, "Fundamental limits of cooperation", IEEE Trans. On Information Theory, Sept. 2013.

- Very successfull to obtain new innovative insights, discover new behaviours (MIMO, IA, delayed CSIT,...)

What is Known: Sum DoF with Centralized Noisy CSIT

- DoF in the K-users MIMO BC with imperfect CSIT recently confirmed [Davoodi and Jafar, 2014]

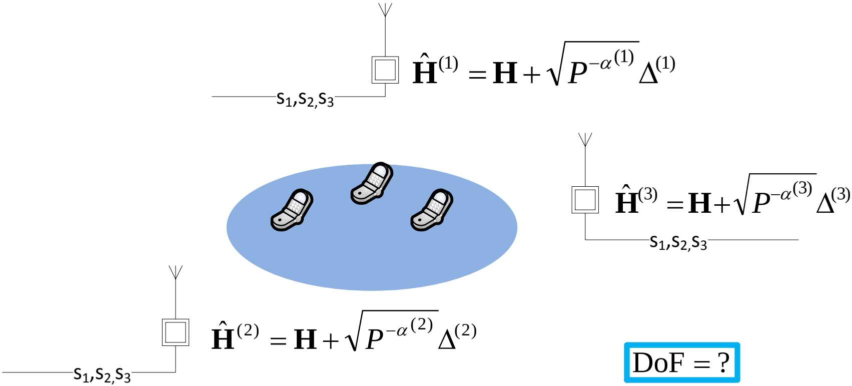


$$\hat{\mathbf{H}} = \mathbf{H} + \sqrt{P^{-\alpha}}\Delta$$

$$\boxed{\mathrm{DoF} = 1 + (K-1)\alpha}$$

- Achieved using simple ZF precoding + rate splitting

Distributed CSIT Configuration: $\alpha^{(1)}, \alpha^{(2)},..,\alpha^{(K)}$



$\boxed{\text{DoF} = ?}$

DoF under Distributed CSIT: Conventional (ZF) precoding

- DoF of Joint Precoding across K distributed TX under D-CSIT, K single-antenna users
- ZF shown to be very inefficient [de Kerret and Gesbert, 2012, TIT]:

$$\mathrm{DoF}^{\mathrm{ZF}} = 1 + (K-1)\min_{j} \alpha^{(j)}$$

- Can we do better?

Principles of New Scheme (Example for $K = 3$)

Key principles:

- Layered precoding
- Layer 1: Transmit with approximate precoder
- Layer 2: Best informed TX regenerates and quantizes interference created by layer 1
- Superpose (multicast) Layer 2 on top of layer 1
- Decode and suppress interference at each user.

We distinguish:

- Arbitrary CSIT regime ($\alpha_i \in [0, 1]\,, \forall i$)
- Weak CSIT regime (α_1,α_2,..α_K) are "small"

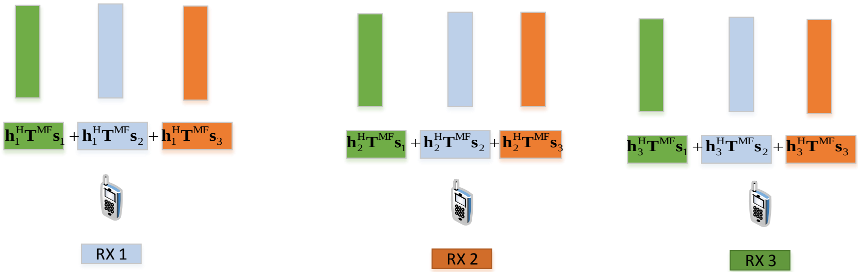
Case $K = 3$ Users: A First Simple Scheme

- Without loss of generality: TX 1 is best informed TX

$$\alpha^{(1)} \geq \alpha^{(2)} \geq \alpha^{(3)}$$

- We transmit 3 symbols per user using e.g. a distributed Matched Precoder with power $P^{(\alpha_1)}/9$

$$\mathbf{T}^{\mathrm{MF}(j)} \triangleq \frac{\hat{\mathbf{H}}^{(j)}}{\|\hat{\mathbf{H}}^{(j)}\|_{\mathrm{F}}}\sqrt{P}$$



Reconstructing the Approximate Interference

- TX 1 uses its CSIT to reconstruct the interference term:

$$\begin{aligned}(\hat{\boldsymbol{h}}_1^{(1)})^{\mathrm{H}}\mathbf{T}^{\mathrm{MF}}\boldsymbol{s}_2 &= (\boldsymbol{h}_1 + P^{-\alpha^{(1)}}\boldsymbol{\delta}_1^{(1)})^{\mathrm{H}}\mathbf{T}^{\mathrm{MF}}\boldsymbol{s}_2\\ &= \boldsymbol{h}_1^{\mathrm{H}}\mathbf{T}^{\mathrm{MF}}\boldsymbol{s}_2 + \underbrace{P^{-\alpha^{(1)}}(\boldsymbol{\delta}_1^{(1)})^{\mathrm{H}}\mathbf{T}^{\mathrm{MF}}\boldsymbol{s}_2}_{\sim P^0}\end{aligned}$$

➥ TX 1 can compute DoF-perfect estimates of the interference terms!

- Quantize the interference using $\alpha^{(1)}\log_2(P)$ bits per term if interference term scales in $P^{\alpha^{(1)}}$
- Superpose a multicast message of $(1-\alpha^{(1)})\log_2(P)$ bits, which will include quantized interference

DoF Analysis: The weak CSIT case ($\alpha^{(1)} < \frac{1}{1+K(K-1)} = 1/7$)

- The 6 quantized interference terms can be broadcast by TX 1 if

$$\underbrace{6\alpha^{(1)}\log_2(P)}_{\text{number of bits to quantize all interference terms}} \leq \underbrace{(1-\alpha^{(1)})\log_2(P)}_{\text{rate of the broadcast data symbol}} \Leftrightarrow \alpha^{(1)} \leq \frac{1}{7}$$

If the inequality is strict, we complete with fresh information bits

- DoF achieved is then

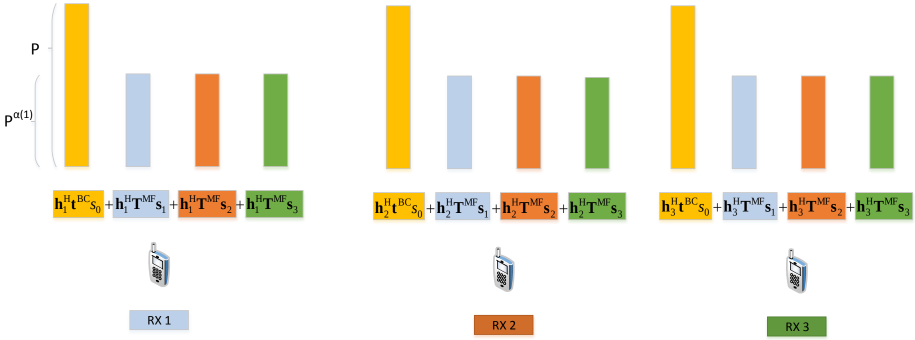
$$\text{DoF} = \underbrace{9\alpha^{(1)}}_{\text{information transmitted initially}} + \underbrace{\left(1-7\alpha^{(1)}\right)}_{\text{fresh information bits to complete the broadcast}} = 1+2\alpha^{(1)}$$

$$\text{DoF} = 1 + (K-1)\max_i \ \alpha^{(i)}$$

(instead of $\text{DoF} = 1 + (K-1)\min_i \ \alpha^{(i)}$!!)

A First Transmission Scheme in One Slide

- TX 1, 2, 3 jointly transmit K symbols to each user using a distributed Matched Precoder $\mathbf{T}^{\mathrm{MF}(j)} \in \mathbb{C}^{K\times K}$ with power $P^{\alpha^{(1)}}/K$
- TX 1 transmits the estimated quantized interference using the power $P - P^{\alpha^{(1)}}$ (equivalently from all TXs using the beamformer $\boldsymbol{t}^{\mathrm{BC}} \triangleq [1,0,0]^{\mathrm{T}}$)



Weak CSIT Regime: Improved results

- Improved scheme: TXs perform Active-Passive Zero-Forcing precoding
  - TX 2,..., TX K perform arbitrary precoding (passive)
  - TX 1 compensates with ZF precoding (active)

Theorem ([de Kerret and Gesbert, 2016, ISIT])

In the weak CSIT regime, defined by

$$\max_{j\in\{1,\dots,K\}} \alpha^{(j)} \leq \frac{1}{1+K(K-2)}$$

We have that:

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \geq 1+(K-1)\max_{j\in\{1,\dots,K\}} \alpha^{(j)}$$

Outer bound

Theorem

The Centralized Outerbound In the K-user Network MIMO channel with distributed CSIT:

$$\begin{aligned} \mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) &\leq \mathrm{DoF}^{\mathrm{CCSI}}(\max_{j\in\{1,\ldots,K\}} \alpha^{(j)}) \\ &= 1 + (K-1) \max_{j\in\{1,\ldots,K\}} \alpha^{(j)} \end{aligned}$$

Key ideas:

- DoF is upperbounded by DoF achieved by full CSIT exchange
- Having multiple CSIT with α_1,α_2,..,α_K doesnt help over having just best CSIT (α_1)

$\Rightarrow$ matches the achieved DoF for weak CSIT!

Arbitrary CSIT Regime with $K = 3$

Theorem

In the 3-user Network MIMO with distributed CSIT and $\alpha^{(1)} \geq \alpha^{(2)} \geq \alpha^{(3)}$, it holds that

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \geq \begin{cases} 1 + 2\alpha^{(1)} & \text{if } \alpha^{(1)} \leq \frac{1}{4} \\ 3\frac{2\alpha^{(1)} - \alpha^{(2)} + 2\alpha^{(1)}\alpha^{(2)}}{4\alpha^{(1)} - \alpha^{(2)}} & \text{if } \alpha^{(1)} \geq \frac{1}{4}. \end{cases}$$

Optimal DoF for $K = 3$ users:

- In the weak CSIT regime
- In any CSIT regime with $\alpha^{(1)} = \alpha^{(2)}$ (regardless of what user 3 knows)

DoF for $K = 3$ Users

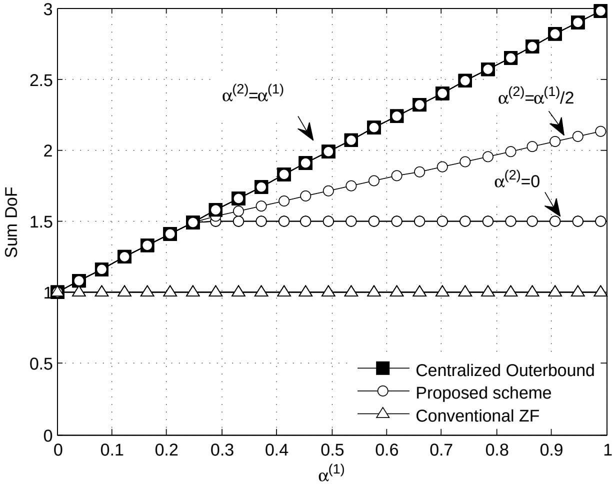


Figure: Sum DoF as a function of $\alpha^{(1)}$. User 3 has no CSIT ($\alpha^{(3)} = 0$)

Outline

Binary Power Control over Interference Channels

$$(p_1^\star, p_2^\star) = \underset{(p_1,p_2)\in\mathcal{P}}{\operatorname{argmax}} \left[\mathrm{R}(p_1(\mathbf{G}^{(1)}), p_2(\mathbf{G}^{(2)}))\right]$$

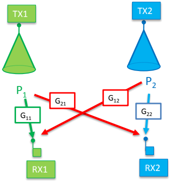
where

$$\mathrm{R}(P_1,P_2) = \log_2\left(1+\frac{G_{11}P_1}{1+G_{12}P_2}\right) + \log_2\left(1+\frac{G_{22}P_2}{1+G_{21}P_1}\right).$$

and

$$\begin{array}{cccc} p_j : & \mathbb{R}_+^4 & \rightarrow & \{P_j^{\min}, P_j^{\max}\} \\ & \mathbf{G}^{(j)} & \mapsto & p_j(\mathbf{G}^{(j)}) \end{array}$$

- Key Example because:
  - Binary optimization with (relatively) low dimensional channel state
  - Weaker depency with the channel state
  - ➥ Less difficulties



Identification of the Parameters

Table: Team Decision Modeling for Power Control

Notations for the Team Decision Problems		
State-of-the-world	$\boldsymbol{x}$	$\mathbf{G}$
Estimate at DM j	$\boldsymbol{x}^{(j)}$	$\mathbf{G}^{(j)}$
Strategy at DM j	$\boldsymbol{s}_j$	p_j
Decision space at DM j	$\mathcal{A}_j$	$\{P_j^{\min}, P_j^{\max}\}$
Objective	f	R

Outline

Discretization for Power Control –Refresher (1)–

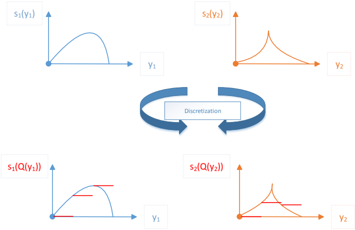
Main Idea: Quantize the channel state space to reduce the dimension

- Replace the strategy p_j by $p_j(Q^{\mathrm{cb}})$ where

$$\begin{array}{rccc} Q^{\mathrm{cb}}: & \mathbb{C}^{2\times 2} & \to & \mathcal{C}^{\mathrm{cb}} \triangleq \{\mathbf{G}_1^{\mathrm{cb}}, \ldots, \mathbf{G}_{n^{\mathrm{cb}}}^{\mathrm{cb}}\} \\ & \hat{\mathbf{G}}^{(j)} & \mapsto & \mathrm{argmin}_{\hat{\mathbf{G}}\in\mathcal{C}^{\mathrm{cb}}} \|\hat{\mathbf{G}} - \hat{\mathbf{G}}^{(j)}\|^2 \end{array}$$

- Optimization subspace reduced to a space of dimension n^{cb}

$$\begin{array}{rccc} p_j(Q^{\mathrm{cb}}): & \mathcal{C}^{\mathrm{cb}} & \to & \{P^{\mathrm{min}}, P^{\mathrm{max}}\} \\ & \mathbf{G}_i & \mapsto & p_j(\mathbf{G}_i) \end{array}$$



Discretization for Power Control –Refresher (2)–

- Best response power allocation strategy: Solve iteratively
  - At TX 1, $\forall i \in \{1, \ldots, n^{\mathrm{cb}}\}$,

$$\boldsymbol{p}_1^{\mathrm{BR}}(\mathbf{G}_i^{\mathrm{cb}}) = \operatorname*{argmax}_{P_1 \in \{P_1^{\mathrm{min}}, P_1^{\mathrm{max}}\}} \mathbb{E}\left[\mathrm{R}\left(\mathbf{G}, P_1, p_2^{\mathrm{BR}}(\mathcal{Q}(\mathbf{G}^{(2)}))\middle|\mathbf{G}^{(1)} = \mathbf{G}_i^{\mathrm{cb}}\right)\right]$$

  - At TX 2, $\forall i \in \{1, \ldots, n^{\mathrm{cb}}\}$,

$$\boldsymbol{p}_2^{\mathrm{BR}}(\mathbf{G}_i^{\mathrm{cb}}) = \operatorname*{argmax}_{P_2 \in \{P_2^{\mathrm{min}}, P_2^{\mathrm{max}}\}} \mathbb{E}\left[\mathrm{R}\left(\mathbf{G}, p_1^{\mathrm{BR}}(\mathcal{Q}(\mathbf{G}^{(1)}), P_2)\middle|\mathbf{G}^{(2)} = \mathbf{G}_i^{\mathrm{cb}}\right)\right]$$

Reach optimal strategy given the strategy of the other TX

Discretization for Power Control –Refresher (3)–

- Approximation of the expectation using Monte-Carlo runs with n^{MC} samples
- At TX 1, $\forall i \in \{1, \dots, n^{\mathrm{cb}}\}$,

$$p_1^{\mathsf{BR}}(\mathbf{G}_i^{\mathrm{cb}}) = \operatorname*{argmax}_{P_1 \in \{P_1^{\min}, P_1^{\max}\}} \frac{1}{n^{\mathrm{MC}}} \sum_{i=1}^{n^{\mathrm{MC}}} \mathrm{R}\left(\mathbf{G}_i, P_1, p_2^{\mathsf{BR}}(\mathcal{Q}^{\mathrm{cb}}(\mathbf{G}_i^{(2)}))\right), \quad \forall i_1 \in \{1, \dots, n^{\mathrm{cb}}\}$$

where $(\mathbf{G}_i, \mathbf{G}_i^{(2)}) \sim f_{\mathbf{G}, \mathbf{G}^{(2)} \big| \mathbf{G}^{(1)} = \mathbf{G}_i^{\mathrm{cb}}}$.

Simulations Parameters

- Rayleigh fading with uniform pathloss
- CSIT Model

$$\mathbf{H}^{(j)} \triangleq \sqrt{1-\sigma_j^2}\mathbf{H} + \sigma_j \boldsymbol{\Delta}$$

  where $\boldsymbol{\Delta} \sim \mathcal{N}_{\mathbb{C}}(0,1)$ and $\mathbf{H} \sim \mathcal{N}_{\mathbb{C}}(0,1)$, and

$$\{\mathbf{G}^{(j)}\}_{i,k} \triangleq \left|\{\mathbf{H}^{(j)}\}_{i,k}\right|^2, \qquad \forall i,k \in \{1,\dots,K\}.$$

- Codebook:
  - Product of scalar codebooks using 10 codewords from Lloyd algorithm for each scalar.
  - Hence: $n^{\mathrm{codebook}} = 10^4 = 10000$
  - Stochastic approximation using $n^{\mathrm{MC}} = 500$

Simulation Results (1)

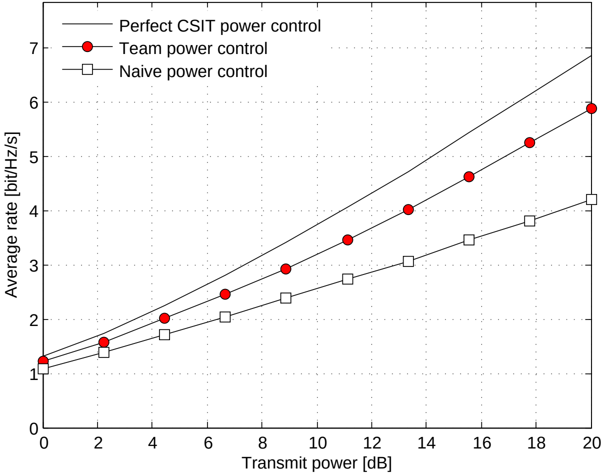


Figure: Average sum rate for $\sigma_1^2 = 1$ and $\sigma_2^2 = 0$.

Simulation Results (2)

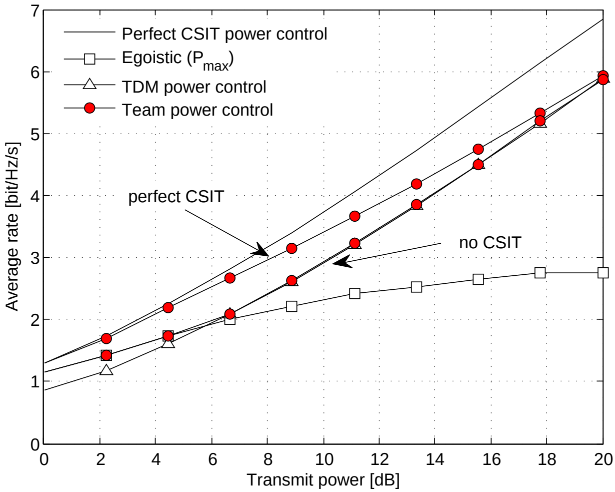
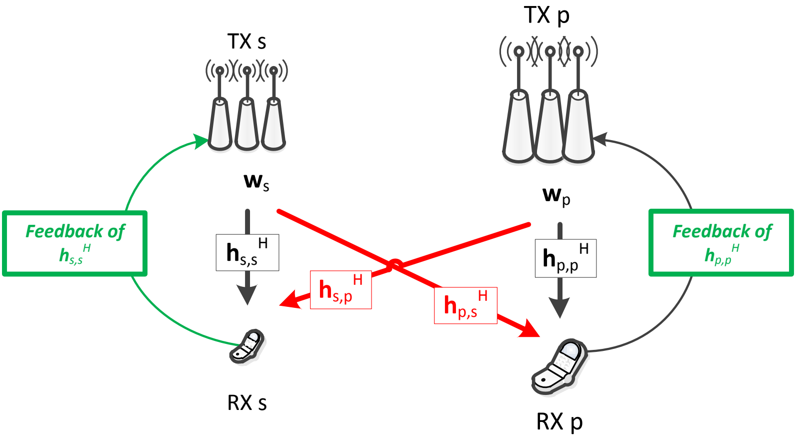


Figure: Comparison with schemes from the literature.

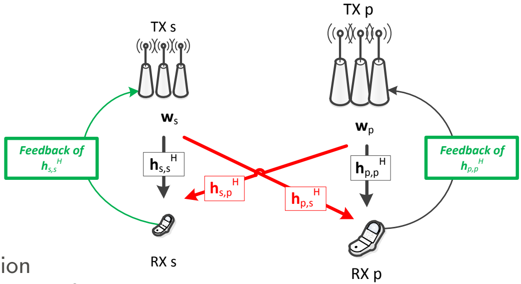
Outline

Cognitive Radio Beamforming

- Rate maximization of a secondary TX while preserving a rate constraint for the primary TX: Underlay Cognitive Radio [Haykin, 2005, JSAC]

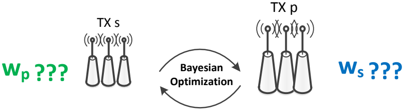


Team Decision Cognitive Radio Beamforming



- CSI configuration
  - TX s only knows $\boldsymbol{h}_{s,s}$ and multi-user statistics $\mathbf{R}_{i,j}$
  - TX p only knows $\boldsymbol{h}_{p,p}$ and multi-user statistics $\mathbf{R}_{i,j}$

➥ Coordination using only the statistics



Functional Optimization Problem

Optimization Problem

$$
\begin{aligned}
(\boldsymbol{w}_p^\star, \boldsymbol{w}_s^\star) = \underset{(\boldsymbol{w}_p, \boldsymbol{w}_s)}{\operatorname{argmax}} \, & \mathbb{E}\left[\mathrm{R}_s(\mathbf{w}_p(\mathbf{h}_{p,p}), \mathbf{w}_s(\mathbf{h}_{s,s}))\right] \\
\text{s. to} \quad & \mathbb{E}\left[\mathrm{R}_p(\mathbf{w}_p(\mathbf{h}_{p,p}), \mathbf{w}_s(\mathbf{h}_{s,s}))\right] \geq \tau > 0, \\
& 0 \leq \left\|\mathbf{w}_p(\mathbf{h}_{p,p})\right\|^2 \leq P_p^{\max} \\
& 0 \leq \left\|\mathbf{w}_s(\mathbf{h}_{s,s})\right\|^2 \leq P_s^{\max}
\end{aligned} \tag{R}
$$

where for $j \in \{s, p\}$,

$$
\begin{aligned}
\boldsymbol{w}_j : \quad \mathbb{C}^{M_j} & \rightarrow \mathbb{C}^{M_j} \\
\boldsymbol{h}_{j,j} & \mapsto \boldsymbol{w}_j(\boldsymbol{h}_{j,j})
\end{aligned}
$$

Identification of the Parameters

Table: Team Decision Modeling for Cognitive Radio

Notations for the Team Decision Problems		
State-of-the-world	$\boldsymbol{x}$	$\{\boldsymbol{h}_{s,s}, \boldsymbol{h}_{s,p}, \boldsymbol{h}_{p,s}, \boldsymbol{h}_{p,p}\}$
Estimate at DM j	$\boldsymbol{x}^{(j)}$	$\boldsymbol{h}_{j,j}$
Strategy at DM j	$\boldsymbol{s}_j$	$\boldsymbol{w}_j$
Decision space at DM j	$\mathcal{A}_j$	$\mathbb{C}^{M_j}$
Objective	f	R_s s.t. (R)

Outline

Key Idea: Codebook of Functions [Filippou et al., 2016, TWC]

Functional optimization difficult $\Rightarrow$ Parametrization of the decision space using a codebook of functions

- Here: 2 functions (strategies) labelled s and p
- Choose these strategies from efficient heuristics

Optimization Problem

$$
\begin{aligned}
(\boldsymbol{w}_p^\star, \boldsymbol{w}_s^\star) = \underset{(\boldsymbol{w}_p, \boldsymbol{w}_s)}{\operatorname{argmax}}\, & \mathbb{E}\left[\mathrm{R}_s(\mathbf{w}_p(\mathbf{h}_{p,p}), \mathbf{w}_s(\mathbf{h}_{s,s}))\right] \\
\text{s. to } \; & \mathbb{E}\left[\mathrm{R}_p(\mathbf{w}_p(\mathbf{h}_{p,p}), \mathbf{w}_s(\mathbf{h}_{s,s}))\right] \geq \tau > 0, \\
& 0 \leq \left\|\mathbf{w}_p(\mathbf{h}_{p,p})\right\|^2 \leq P_p^{\max} \\
& 0 \leq \left\|\mathbf{w}_s(\mathbf{h}_{s,s})\right\|^2 \leq P_s^{\max} \\
& (\boldsymbol{w}_p^\star, \boldsymbol{w}_s^\star) \in \left\{ (\boldsymbol{w}_p^{\mathrm{cb},1}, \boldsymbol{w}_s^{\mathrm{cb},1}), \ldots, (\boldsymbol{w}_p^{\mathrm{cb},n^{\mathrm{cb}}}, \boldsymbol{w}_s^{\mathrm{cb},n^{\mathrm{cb}}}) \right\}
\end{aligned} \tag{R}
$$

where for $j \in \{s, p\}$,

$$
\begin{aligned}
\boldsymbol{w}_j^{\mathrm{cb},i} : \quad \mathbb{C}^{M_j} & \rightarrow \mathbb{C}^{M_j} \\
\boldsymbol{h}_{j,j} & \mapsto \boldsymbol{w}_j^{\mathrm{cb}}(\boldsymbol{h}_{j,j})
\end{aligned}
$$

Strategy p

- TX p
  - TX p uses matched precoding
$$\boldsymbol{u}_p^{\mathrm{MF}} \triangleq \frac{\boldsymbol{h}_{p,p}}{\|\boldsymbol{h}_{p,p}\|}$$
  - TX p transmits with full power $\bar{P}_p = P_p^{\max}$
- TX s:
  - TX s transmits using the statistical ZF precoder
$$\boldsymbol{u}_s^{\mathrm{sZF}} \triangleq \operatorname*{argmin}_{\boldsymbol{u}} \boldsymbol{u}^{\mathrm{H}} \mathbf{R}_{p,s} \boldsymbol{u}$$
  - TX s controls its average transmit power $\bar{P}_s$ to fulfill the rate constraint (R)

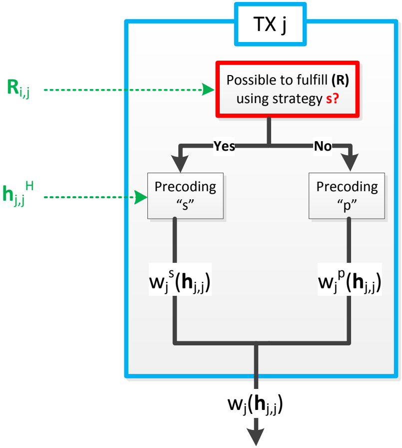
Strategy s

- TX s
  - TX s uses matched precoding
$$\boldsymbol{u}_s^{\mathrm{MF}} \triangleq \frac{\boldsymbol{h}_{s,s}}{\|\boldsymbol{h}_{s,s}\|}$$
  - TX s transmits with full power $\bar{P}_s = P_s^{\max}$
- TX p:
  - TX p transmits using the statistical ZF precoder
$$\boldsymbol{u}_p^{\mathrm{sZF}} \triangleq \operatorname*{argmin}_{\boldsymbol{u}} \boldsymbol{u}^{\mathrm{H}} \mathbf{R}_{s,p} \boldsymbol{u}$$
  - TX p controls its average transmit power $\bar{P}_p$ to fulfill the rate constraint (R)

Some Intuition

- TX p can reduce its power only if TX s can anticipate it: Coordination required to guarantee (R)
- Strategy s
  - Large objective
  - Rate constraint might be unfeasible
- Strategy p
  - Low objective
  - Rate constraint guaranteed

Statistical Coordination Algorithm [Filippou et al., 2016, TWC]



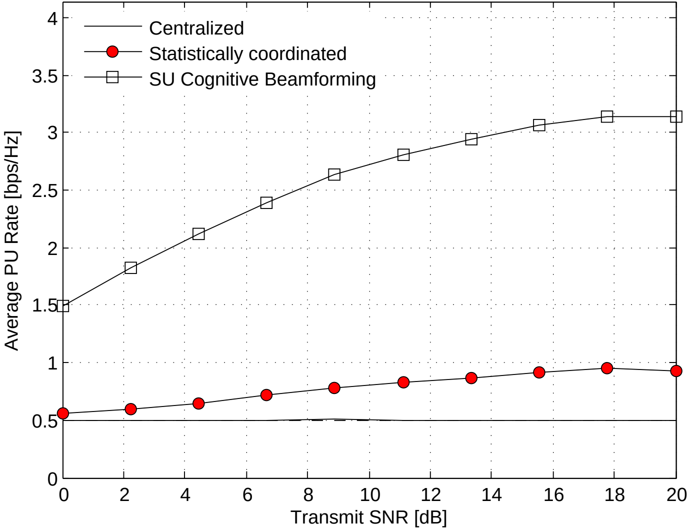
Simulations Parameters

- $M_s = M_p = 3$ antennas per-TX
- Correlation matrices

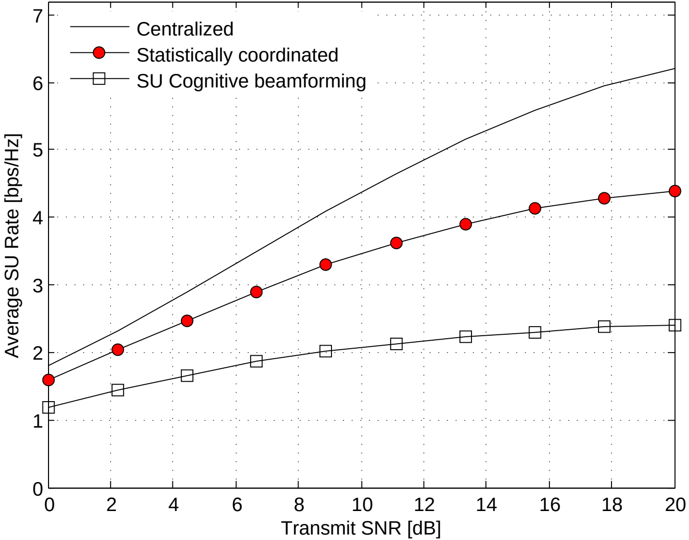
$$\mathbf{R}_{p,p} = \mathbf{R}_{s,s} = \mathbf{I}_3, \quad \mathbf{R}_{p,s} = \mathbf{R}_{s,p} = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

- Use in the following $\rho = 0.5$ and $\tau = 0.5$bps/Hz

Ergodic rate of the PU



Ergodic rate of the SU

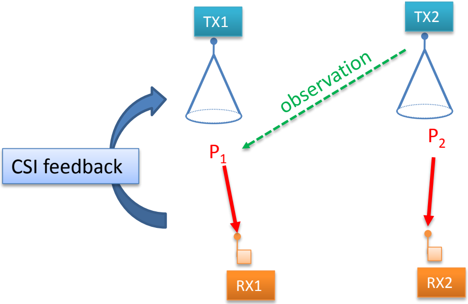


Outline

Coded Power Control

- TX 1 with non-causal CSI knowledge
- TX 2 observes transmit power P_1:

➥ implicit coordination



Modelization

- random state $H = \{G_{1,1}, G_{2,1}, G_{1,2}, G_{2,2}\}$ with fixed law $\rho(H)$ over $\mathcal{H}$
- TX i chooses its power levels P_i in $\mathcal{P}_i$
- TX 2 observes Z with fixed law $\Gamma(z|P_1)$.
- Strategy of Agent 1: $(w_{1,i})_{1\leq i\leq T}$ with:

$$w_{1,i} : \underbrace{\mathcal{H}^T}_{\text{CSI}} \times \underbrace{\mathcal{P}_1^{i-1}}_{\text{past actions}} \times \emptyset \rightarrow \mathcal{P}_1$$

- Strategy of Agent 2: $(w_{2,i})_{1\leq i\leq T}$ with:

$$w_{2,i} : \underbrace{\mathcal{H}^{i-1} \times \mathcal{Z}^{i-1} \times \mathcal{P}_2^{i-1}}_{\text{past}} \rightarrow \mathcal{P}_2$$

Auxiliary notion

Definition (Implementability)

$P_{H_i,P_{1,i},P_{2,i},Z_i}$: joint distribution induced by $(w_{1,i}, w_{2,i})_{i\geq 1}$ at stage i.
The distribution $\mathcal{Q}(h, p_1, p_2)$ is implementable if there exists a pair of strategies $(w_{1,i}, w_{2,i})_{i\geq 1}$ such that for all (h, p_1, p_2),

$$\frac{1}{T}\sum_{i=1}^{T}\sum_{y} P_{H_i,P_{1,i},P_{2,i},Y_i}(h, p_1, p_2, z) \to \mathcal{Q}(h, p_1, p_2)$$

as $T \to \infty$.

Feasible utilities

A certain utility value $\underline{f}$ is reachable if and only if there exists an implementable distribution $\mathcal{Q}$ such that $\underline{f} = \mathbb{E}_{\mathcal{Q}}[f]$.

Theorem ([Larrousse and Lasaulce, 2013])

Let $\overline{Q} \in \Delta(\mathcal{H} \times \mathcal{P}_1 \times \mathcal{P}_2)$ *with* $\sum_{(p_1,p_2)} \overline{Q}(h, p_1, p_2) = \rho(h)$. *The distribution* $\overline{Q}$ *is implementable if there exists* $Q \in \Delta(\mathcal{H} \times \mathcal{P}_1 \times \mathcal{P}_2 \times \mathcal{Z})$ *which verifies:*

$$I_Q(H; P_2) \leq I_Q(P_1; Z|H, P_2)$$

where the arguments of the mutual information $I_Q(.)$ *are defined from* Q *and* $Q(h, p_1, p_2, y) = \overline{Q}(h, p_1, p_2)\Gamma(z|p_1)$.

Remark: This theorem also characterizes expected payoff.

Convex Optimization Problem

$$
\begin{aligned}
\text{maximize } & \mathbb{E}_{\underline{q}}[f] = \sum_{\ell=1}^{L} q_\ell f_\ell \\
\text{subject to } & I_{\underline{q}}(H; P_2) \leq I_{\underline{q}}(P_1; Z|H, P_2) \\
& q_\ell \geq 0 \\
& \sum_{\ell=1}^{L} q_\ell = 1 \\
& \sum_{\ell \in \mathcal{L}_H(h)} q_\ell = \rho(h), \qquad \forall\, h, \\
& \frac{\sum_{\ell \in \mathcal{L}_{P_1,Z}(p_1,z)} q_\ell}{\sum_{\ell \in \mathcal{L}_{P_1}(p_1)} q_\ell} = \Gamma(z|p_1), \qquad \forall\, (p_1, z)
\end{aligned}
$$

Application: MAC Power Control

- 2-user MAC with binary power control
- 2 possible states: $\mathcal{H} = \{\{g_{\min}, g_{\max}\}, \{g_{\max}, g_{\min}\}\}$

	$P_{\min}$	$P_{\max}$
$P_{\min}$	0	20
$P_{\max}$	1	$\simeq 10$

$x_0 = (g_{\min}, g_{\max})$

	$P_{\min}$	$P_{\max}$
$P_{\min}$	0	1
$P_{\max}$	20	$\simeq 10$

$x_0 = (g_{\max}, g_{\min})$

Figure: Payoff table [Larrousse and Lasaulce, 2013]

Source Coding [Larrousse and Lasaulce, 2013]

- Source coding:

$$
\begin{array}{rccc}
f_{\mathrm{S}}: & \mathcal{H} & \rightarrow & \{m_0, m1\} \\
 & h & \mapsto & i
\end{array}
$$

TABLE V

PROPOSED SOURCE CODING AND DECODING FOR $p = \frac{1}{2}$.

x_0^3	Index $i = f_{\mathrm{S}}(x_0^3)$	$g_{\mathrm{S}}(i)$
(0,0,0)	m_0	(1,1,1)
(0,0,1)	m_0	(1,1,1)
(0,1,0)	m_0	(1,1,1)
(0,1,1)	m_0	(1,1,1)
(1,0,0)	m_0	(1,1,1)
(1,0,1)	m_0	(1,1,1)
(1,1,0)	m_1	(0,0,1)
(1,1,1)	m_1	(0,0,1)

Channel Coding

- Channel coding:

$$\begin{array}{cccc} f_{\mathrm{C}}: & \{m_0, m_1\} & \to & \mathcal{P}_1 \\ & i & \mapsto & \boldsymbol{p}_1 = [p_1(1), p_1(2), p_1(3)] \end{array}$$

- at block b, find optimal $\boldsymbol{p}_1^{\text{opt}} = [p_1(1)^{\text{opt}}, p_1(2)^{\text{opt}}, p_1(3)^{\text{opt}}]$ and second optimal $\boldsymbol{p}_1^{\text{opt}''} = [p_1(1)^{\text{opt}''}, p_1(2)^{\text{opt}''}, p_1(3)^{\text{opt}''}]$
  - If $i = m_0$, send with $\boldsymbol{p}_1^{\text{opt}}$
  - If $i = m_1$, send with $\boldsymbol{p}_1^{\text{opt}''}$

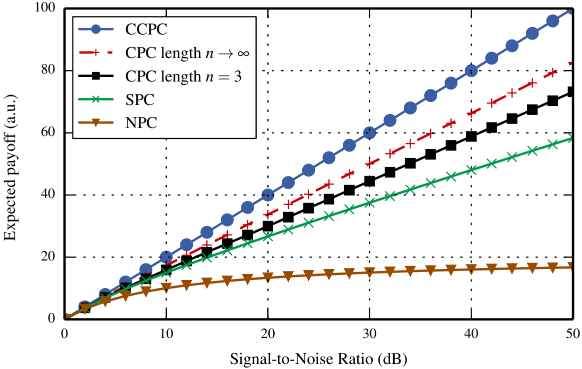
Channel Coding [Larrousse and Lasaulce, 2013]

TABLE VI

PROPOSED CHANNEL CODING FOR $p = \frac{1}{2}$.

$x_0^3(b)$	$x_2^3(b)$	i_{b+1}	$x_1^3(b)$
(0,0,0)	(1,1,1)	m_0	(0,0,0)
		m_1	(0,0,1)
(0,0,1)	(1,1,1)	m_0	(0,0,1)
		m_1	(0,0,0)
(0,1,0)	(1,1,1)	m_0	(0,1,0)
		m_1	(0,0,0)
(0,1,1)	(1,1,1)	m_0	(0,1,1)
		m_1	(0,0,1)
(1,0,0)	(1,1,1)	m_0	(1,0,0)
		m_1	(0,0,0)
(1,0,1)	(1,1,1)	m_0	(1,0,1)
		m_1	(0,0,1)
(1,1,0)	(0,0,1)	m_0	(1,1,0)
		m_1	(1,1,1)
(1,1,1)	(0,0,1)	m_0	(1,1,1)
		m_1	(1,1,0)

Simulations [Larrousse and Lasaulce, 2013]



Outline

Device centric coordination

- Relying on local communications and decentralized computations
- Decentralized cooperation can aim at the good of the network
- Challenge: Develop robust one-shot schemes that cope with arbitrary information structures
- Heuristics can be obtained by decoupling the communication from decision problems
- Open problems
  - Joint optimization of communications and decision is very challenging
  - Low complexity methods?
  - Information theoretic aspects (capacity under decentralized information settings..) ?
  - Coordination-aware feedback designs (hierarchical,...)

Other impact

- Coordination theory leads to new insights: Impact over network design?
- Implicit coordination: Coordination for free?
- Bridge the gap from implicit coordination to distributed optimization?
- Interactions with distributed optimization

Big thanks to

Artist4G

iJOIN

MITSUBISHI ELECTRIC
Changes for the Better

orange™

HUAWEI

erc
European Research Council
Established by the European Commission


thanks

References I

V. R. Cadambe and S. A. Jafar. Interference alignment and degrees of freedom of the K-user interference channel. *IEEE Trans. Inf. Theory*, 54(8): 3425–3441, Aug. 2008.

R. A. Chou, M. R. Bloch, and J. Kliewer. Polar coding for empirical and strong coordination via distribution approximation. In *2015 IEEE International Symposium on Information Theory (ISIT)*, 2015.

Romain Couillet and Merouane Debbah. *Random matrix methods for wireless Communications*. Cambridge University Press, 2011.

P. W. Cuff, H. H. Permuter, and T. M. Cover. Coordination Capacity. *IEEE Trans. Inf. Theo.*, 56(9):4181–4206, 2010.

A.G. Davoodi and S.A. Jafar. Settling conjectures on the collapse of degrees of freedom under finite precision csit. In *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2014.

P. de Kerret and D. Gesbert. Degrees of freedom of the network MIMO channel with distributed CSI. *IEEE Trans. Inf. Theory*, 58(11):6806–6824, Nov. 2012.

P. de Kerret and D. Gesbert. Network MIMO: Transmitters with no CSI Can Still be Very Useful. to be presented in ISIT, 2016.

P. de Kerret, D. Gesbert, and U. Salim. Large system analysis of joint regularized Zero Forcing precoding with distributed CSIT. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2015.

Miltiades C. Filippou, Paul de Kerret, David Gesbert, Tharmalingam Ratnarajah, Adriano Pastore, and George A Ropokis. Statistically coordinated precoding for the MISO cognitive radio channel. Accepted for publication in IEEE Trans. on Wireless Commun., 2016.

A. Gjendemsjo, D. Gesbert, G. E. Oien, and S. G. Kiani. Binary power control for sum rate maximization over multiple interfering links. *IEEE Trans. on Wireless Commun.*, 7(8):3164–3173, 2008.

N. Golrezaei, A. G. Dimakis, and A. F. Molisch. Scaling Behavior for Device-to-Device Communications With Distributed Caching. *IEEE Trans. Inf. Theory.*, 60(7):4286–4298, 2014.

John C. Harsanyi. Games with incomplete information played by Bayesian players, IIII. *Management Science*, 14(3):159–182, 1967.

S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE J. Sel. Areas Commun.*, 23(2):201–220, 2005.

Y. C. Ho. Team decision theory and information structures. *Proceedings of the IEEE*, 68(6):644–654, 1980.

E. A. Jorswieck, E. G. Larsson, and D. Danev. Complete characterization of the Pareto boundary for the MISO interference channel. *IEEE Trans. Signal Process.*, 56(10):5292–5296, Oct. 2008.

B. Larrousse and S. Lasaulce. Coded power control: Performance analysis. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2013.

Qianrui Li, David Gesbert, and Nicolas Gresset. Joint precoding over a master-slave coordination link. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.

References II

Qianrui Li, Paul de Kerret, David Gesbert, and Nicolas Gresset. Robust regularized zf in decentralized broadcast channel with correlated csi noise. In *Proc. Allerton Conference on Communication, Control, and Computing (Allerton)*, 2015.

A. Lozano, R. W. Heath, and J. G. Andrews. Fundamental limits of cooperation. *IEEE Trans. Inf. Theory*, 59(9):5213–5226, 2013.

M.A. Maddah-Ali, A.S. Motahari, and A.K. Khandani. Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis. *IEEE Trans. Inf. Theory*, 54(8):3457–3470, Aug. 2008.

A. Papadogiannis, D. Gesbert, and E. Hardouin. A dynamic clustering approach in wireless networks with multi-cell cooperative processing. In *Proc. IEEE International Conference on Communications (ICC)*, 2008.

Park, Lee, and Heath. Cooperative Base Station Coloring for Pair-wise Multi-Cell Coordination. In *arxiv*, 2015.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2014. ISBN 1611973422, 9781611973426.

S. Wagner, R. Couillet, M. Debbah, and D. Slock. Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback. *IEEE Trans. Inf. Theory*, 58(7):4509–4537, July 2012.

R. Zakhour and D. Gesbert. Team decision for the cooperative MIMO channel with imperfect CSIT sharing. In *Proc. Information Theory and Applications Workshop (ITA)*, 2010.