Device-Centric Cooperation in Wireless Networks

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

July 4th, 2016
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
- Transportation 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”

Network coordination is often, by essence, ”myopic”
Outline

1. Wireless Device Cooperation
2. Distributed Information Models
3. Coordination and Team decision: Problem formulation
4. Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
   - A Different Point of View : Implicit Coordination
5. Key Aspects and Open Problems
Outline

1. Wireless Device Cooperation
2. Distributed Information Models
3. Coordination and Team decision: Problem formulation
4. Applications of Team Decision to Device-Centric Cooperation
5. Key Aspects and Open Problems
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, cooperation 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
Cooperation in heterogenous Wireless networks

User’s data from core network

Sharing/caching

CSI sharing with delay/quantization/noise

\[ x_1 = w_1(H^{(1)})s \]
\[ x_2 = w_2(H^{(2)})s \]
\[ x_3 = w_3(H^{(3)})s \]
Ultra flexible Wireless networks

UAVs with aerial relays

Internet

Wireless gateway

Signaling for dynamic positioning
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]

$$(p_1^*, p_2^*) = \arg \max_{(p_1, p_2) \in P} [R(p_1(\{G_{i,j}\}), p_2(\{G_{i,j}\}))]$$

where

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

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

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 $H_{i,j}$ at all transmitters. **What if not the case?**
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 \( w_i, i = 1, 2 \)
- 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

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

- **Team Decision theory**
  - 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.,...]

- **Game theory**
  - 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.,...]

- **Information theory**
  - Complexity/Convergence studies
  - Consensus algorithms
  - Delay tolerant applications
  - ... [Boyd et al, Inalhan et al, Colorni et al, Rabber et al, Chen et al., Johansson et al., Palomar et al., Scaglione et al., Scutari et al.]

- **Distributed optimization**
  - Capacity/DoF analysis
  - Coordination theory
  - Quantizing with side info.
  - ... [Larrousse et al, Cuff et al, Li et al, Grover, ...]
Device coordination: The many perspectives

**Team Decision theory**
- 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.,...]

**Information theory**
- Capacity/DoF analysis
- Coordination theory
- Quantizing with side info.
- ...
  - [Larousse et al, Cuff et al, Li et al, Grover, ...]

**Game theory**
- 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, ...]

**Distributed optimization**
- Complexity/Convergence studies
- Consensus algorithms
- Delay tolerant applications
- ...
Outline

1. Wireless Device Cooperation
2. Distributed Information Models
3. Coordination and Team decision: Problem formulation
4. Applications of Team Decision to Device-Centric Cooperation
5. Key Aspects and Open Problems
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 structure: 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

\[
H^{(1)} = [h_1^{(1)} h_2^{(1)} h_3^{(1)}] \\
H^{(2)} = [h_1^{(2)} h_2^{(2)} h_3^{(2)}] \\
H^{(3)} = [h_1^{(3)} h_2^{(3)} h_3^{(3)}]
\]
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]

\[
\{\hat{H}\}_{i,k} = \sqrt{1 - \sigma^2_{i,k}} \{\tilde{H}\}_{i,k} + \sigma_{i,k} \{\Delta\}_{i,k}, \quad \forall i, k
\]

where \(\{\Delta\}_{ik} \sim CN(0, 1)\)
- With digital quantization \(\sigma^2_{i,k} = 2^{-B_{i,k}}\) (good approximation in the high resolution regime)
- CSIT allocation matrix \(B\) defined as

\[
\{B\}_{i,k} = B_{i,k}, \quad \forall i, k
\]
Distributed CSI Model

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

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

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

- $\sigma_{i,k}^{(j)}$ 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 $\hat{H}^{(i)} = H$, $\forall i$.
- A CSI structure is *centralized* if $\hat{H}^{(i)} = \hat{H}^{(j)}$, $\forall i, j$.
- A CSI structure is *distributed* if there exist $i$ and $j$ such that $\hat{H}^{(i)} \neq \hat{H}^{(j)}$. 
Distributed CSI structure models (cont’d)

Some more particular cases:

- **Incomplete CSIT**: A CSI structure is *incomplete* if \( \hat{H}^{(i)} \) takes the form \( \forall i \hat{H}^{(i)} = \{H_k, l, k \in S_{TX}, l \in S_{RX}\} \), where \( S_{TX} \) (resp. \( S_{RX} \)) 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 \( \hat{H}^{(i_1)} \subset \hat{H}^{(i_2)} \subset \hat{H}^{(i_3)} \subset ... \)

- **Master Slave**: Hierarchical where \( \hat{H}^{(i_1)} = [ ] \), and \( \hat{H}^{(i_2)} = H \) (can be extended to \( K > 2 \)).
Typical (practical) CSI structures

Consider the $K$ transmitter ($N$ antennas each) $K$ user (single antenna) channel. Let $h_{i,j}^H$ be the $1 \times N$ vector channel between the $j$th transmitter and the $i$th user.

- **Local CSIT with TDD reciprocity**

$$ (\hat{H}(j))^H = \begin{bmatrix} 0 & h_{1,j}^H & 0 \\ \vdots & \vdots & \vdots \\ 0 & h_{K,j}^H & 0 \end{bmatrix} $$

- **Local CSIT with LTE feedback mode**

$$ (\hat{H}(j))^H = \begin{bmatrix} 0 & 0 & 0 \\ h_{j,1}^H & \cdots & h_{j,K}^H \\ 0 & 0 & 0 \end{bmatrix} $$

- **Fully local CSIT**

$$ (\hat{H}(j))^H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & h_{j,j}^H & 0 \\ 0 & 0 & 0 \end{bmatrix} $$
Outline

1. Wireless Device Cooperation
2. Distributed Information Models
3. Coordination and Team decision: Problem formulation
4. Applications of Team Decision to Device-Centric Cooperation
5. Key Aspects and Open Problems
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

Coordination domains:
- Power
- Time/freq/code
- Antenna/beam
- …
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*

- Estimated position of person $j$ available at person $i$
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 $\hat{H}^{(i)}$ and $Q_i$, $i = 1, \ldots, K$, exchange quantized decisions over $R_{ij}$ bits
     - But poorly informed nodes make bad decisions!
  2. Exchange quantized CSI $\hat{H}^{(i)}$ over $R_{ij}$ bits
     - But this ignores $Q_i$!

- 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** [Zakhor and Gesbert, 2010, ITA]
- Fundamental limits rooted in **Coordination Theory**
**Coordination Theory**

- $H_1$, $H_2$ and $H_3$, arbitrary components of global system state, distributed according to $p_0(H_1, H_2, H_3)$
- $W_1$, $W_2$ and $W_3$ are actions selected by the nodes.
- What joint distribution $p_0(H_1, H_2, H_3)p(W_1, W_2, W_3|H_1, H_2, H_3)$ can be achieved?
- Answer: it depends on graph topology (capacity of each edge)

**Figure:** Coordination Framework [Cuff et al., 2010, TIT]
Example (One Isolated Node)

\[ C_{p_0} = \left\{ (R, p(W_1, W_2|H_1)) \left| R \geq I(H_1; W_1|W_2) \right. \right\} \]

Figure: One isolated node scenario [Cuff et al., 2010, TIT]
Example (One Isolated Node)

With Gaussian RVs, the condition becomes

\[(1 - 2^{-2R})^{-1} \rho_{H_1, W_1} + \rho_{W_1, W_2}^2 \leq 1\]

- \(R \to \infty\), \(\rho_{H_1, W_1}^2 + \rho_{W_1, W_2}^2 \leq 1\)
- \(R \to 0\), \(\rho_{H_1, W_1}^2 = 0\) and \(\rho_{W_1, W_2}^2 \leq 1\)

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

\[(s_1^*, \ldots, s_K^*) = \arg\max_{s_1, \ldots, s_K} \mathbb{E}_{x, x^{(1)}, \ldots, x^{(K)}} \left[ f \left( x, s_1(x^{(1)}), \ldots, s_K(x^{(K)}) \right) \right] \]

where

- \( K \): Number of Decision Makers (DMs)
- \( x \in \mathbb{C}^m \): State of the world
- \( x^{(j)} \in \mathbb{C}^m \): Estimate of the state of the world \( x \) at DM \( j \)
- \( s_j : \mathbb{C}^m \rightarrow A_j \subset \mathbb{C}^{d_j} \): Strategy of the \( j \)-th DM
- \( s_j(x^{(j)}) \in A_j \subset \mathbb{C}^{d_j} \): Decision at DM \( j \) for the given realization \( x^{(j)} \)
- \( f : \mathbb{C}^m \times \prod_{j=1}^K \mathbb{C}^{d_j} \rightarrow \mathbb{R} \): Joint objective of the \( K \) DMs
- \( p_{x,x^{(1)},\ldots,x^{(K)}} \): 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 \textit{Best-Response (BR)} strategy $s_1^{BR}, \ldots, s_K^{BR}$ for the TD problem is a strategy such that

$$s_j^{BR} = \arg\max_{s_j \in A_j} \mathbb{E}_{x, x^{(1)}, \ldots, x^{(K)} | x^{(j)}} \left[ f \left( x, \ldots, s_{j-1}^{BR}(x^{(j-1)}), s_j(x^{(j)}), s_{j+1}^{BR}(x^{(j+1)}), \ldots \right) \right], \quad \forall j$$

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

- Codebook-Based Approach
- Discretization-Based Approach
- Model-Based Approach
- Asymptotics-Based Approach

Team Decision Problem

Information Allocation

Information Sharing
Model-Based Approach

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

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

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

Beamformer in the MISO IC parameterized as

$$w_k^*(\lambda_k) = \frac{\lambda_k w^\text{ZF}_k + (1 - \lambda_k) w^\text{MF}_k}{\|\lambda_k w^\text{ZF}_k + (1 - \lambda_k) w^\text{MF}_k\|}$$
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:

<table>
<thead>
<tr>
<th>Person 2</th>
<th>Person 1</th>
<th>Buy bread</th>
<th>Go home</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buy bread</td>
<td>$a - c_1 - c_2$</td>
<td>$1 - c_1$</td>
<td></td>
</tr>
<tr>
<td>Go home</td>
<td>$1 - c_2$</td>
<td></td>
<td>$0$</td>
</tr>
</tbody>
</table>

When should each person buy bread?

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

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


**Codebook-Based Approach**

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

Choose $s_j$ inside a codebook of function \( \{s_j^1, \ldots, s_j^m\} \)

**Example (Coordinated Beamforming)**

- Restrict possible beamforming choices to \( C = \{\text{Matched Filter, Zero Forcing}\} \)
Discretization-Based (1): Dimensionality Reduction

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

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

$$Q^{cb}: \mathbb{C}^m \rightarrow \mathbb{C}^{cb} \triangleq \{x_1, \ldots, x_{n^{cb}}\}$$

$$x^{(j)} \mapsto Q^{cb}(x^{(j)}) = \arg\min_{x \in \mathbb{C}^{cb}} \|x - x^{(j)}\|^2$$

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

$$s_j: \mathbb{C}^{cb} \rightarrow \mathcal{A}_j$$

$$x_i \mapsto s_j(x_i)$$
Discretization-Based (2): Monte-Carlo Approximation

- Best-response optimization at DM 1: \( \forall i_1 \in \{1, \ldots, n^{cb}\} \),

\[
s^\text{BR}_1(x_{i_1}) = \arg\max_{s_1 \in A_1 \subseteq C_{d1}} \mathbb{E} \left[ f(x, S_1, s^\text{BR}_2(Q^{cb}(x^{(2)})), \ldots, s^\text{BR}_K(Q^{cb}(x^{(K)}))) \middle| x^{(1)} = x_{i_1} \right]
\]

- For a given \( x_i \) and given \( s^\text{BR}_2, \ldots, s^\text{BR}_K \): Standard stochastic optimization problem [Shapiro et al., 2014]

- Use Monte-Carlo approximation: \( \forall i \in \{1, \ldots, n^{cb}\} \),

\[
s^\text{BR}_1(x_i) = \arg\max_{s_1 \in A_1} \frac{1}{n^{MC}} \sum_{\ell=1}^{n^{MC}} f \left( x_\ell, S_1, s_2(Q(x_\ell^{(2)})), \ldots, s_K(Q(x_\ell^{(K)})) \right)
\]

where \( \left( x_\ell, x_\ell^{(2)}, \ldots, x_\ell^{(K)} \right) \sim p_{x,x^{(2)},\ldots,x^{(K)}} \middle| x^{(1)} = x_i \)
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

---

Outline

1 Wireless Device Cooperation

2 Distributed Information Models

3 Coordination and Team decision: Problem formulation

4 Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
   - A Different Point of View: Implicit Coordination

5 Key Aspects and Open Problems
Joint Precoding over Network MIMO

Data: $S_1, S_2, \ldots, S_{N_{\text{cells}}}$

Network MIMO
Transmit antennas: $N_t \times N_{\text{cells}}$

Data: $S_1, S_2, \ldots, S_{N_{\text{cells}}}$
**Team Decision Problem**

\[
(w_1^*, \ldots, w_K^*) = \arg\max_{(p_1, \ldots, p_K) \in \mathcal{P}} \mathbb{E}[R(H, w_1(\hat{H}^{(1)}), \ldots, w_K(\hat{H}^{(K)}))] \]

where

\[
R(H, w_1(\hat{H}^{(1)}), \ldots, w_K(\hat{H}^{(K)})) = \sum_{k=1}^{K} \log_2 \left| I_{d_k} + T_k^H H_k^H \left( R_k + \sum_{i \neq k} H_i T_i T_i^H T_i^H \right)^{-1} H_k T_k \right|
\]

with

- \( H \in \mathbb{C}^{N_{\text{tot}} \times M_{\text{tot}}} \) the multi-user channel
- \( w_j \) the precoding function:
  \[
  w_j : \mathbb{C}^{N_{\text{tot}} \times M_{\text{tot}}} \rightarrow \mathbb{C}^{M_j \times d_{\text{tot}}}
  \]
  \[
  \hat{H}^{(j)} \mapsto w_j(\hat{H}^{(j)})
  \]
- \( T \in \mathbb{C}^{M_{\text{tot}} \times d_{\text{tot}}} \) the multi-user precoder

\[
T = \begin{bmatrix} T_1 & \ldots & T_K \end{bmatrix} = \begin{bmatrix} w_1(\hat{H}^{(1)}) \\ \vdots \\ w_K(\hat{H}^{(K)}) \end{bmatrix}
\]
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

<table>
<thead>
<tr>
<th>Notations for the Team Decision Problems</th>
<th>( x )</th>
<th>( \hat{H} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>State-of-the-world ( x )</td>
<td>( x^{(j)} )</td>
<td>( \hat{H}^{(j)} )</td>
</tr>
<tr>
<td>Estimate at DM ( j ) ( \hat{x} )</td>
<td>( \hat{x}^{(j)} )</td>
<td></td>
</tr>
<tr>
<td>Strategy at DM ( j ) ( s_j )</td>
<td>( s_j )</td>
<td>( w_j )</td>
</tr>
<tr>
<td>Decision space at DM ( j ) ( A_j )</td>
<td>( A_j )</td>
<td>( \mathbb{C}^{M_j \times d_{tot}} )</td>
</tr>
<tr>
<td>Objective ( f )</td>
<td>( f )</td>
<td>( R )</td>
</tr>
</tbody>
</table>
Outline

1. Wireless Device Cooperation
2. Distributed Information Models
3. Coordination and Team decision: Problem formulation
4. Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
     - A Different Point of View: Implicit Coordination
5. Key Aspects and Open Problems
Applications of Team Decision to Device-Centric Cooperation

Application to Network MIMO Precoding

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

\[ \beta \triangleq \lim_{M,K \to \infty} \frac{M}{K} \triangleq \lim_{M_{TX},K \to \infty} \frac{nM_{TX}}{K} \geq 1 \]
Distributed CSI with Correlated Noise

- Extend to spatial correlation in the CSI noise

\[ \hat{h}_k^{(j)} \triangleq \sqrt{1 - \left(\sigma_k^{(j)}\right)^2} h_k + \sigma_k^{(j)} \delta_k^{(j)} \]

with

\[ \mathbb{E} \left[ \delta_k^{(j)} (\delta_k^{(j')})^\text{H} \right] = (\rho_k^{(j,j')})^2 I_M \]

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

\[ \hat{h}_k^{(1)} = \sqrt{1 - \left(\sigma_k^{(1)}\right)^2} h_k + \sigma_k^{(1)} \delta_k^{(1)} \]

\[ \hat{h}_k^{(2)} = \sqrt{1 - \left(\sigma_k^{(2)}\right)^2} h_k + \sigma_k^{(2)} \delta_k^{(2)} \]
A Practical Example

Example

- Imperfect feedback
  \[
  \hat{h}_j^{(j)} = \sqrt{1 - \sigma_{FB}^2} h_j + \sigma_{FB} \delta_j^{(j)}
  \]

- Imperfect backhaul
  \[
  \hat{h}_k^{(j')} = \sqrt{1 - \sigma_{BH}^2} \hat{h}_k^{(j)} + \sigma_{BH} \epsilon_k^{(j,j')}
  \]

- 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 $P$

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

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

$$ w_j(\hat{H}^{(j)}) = E_j^H T^{(j)}_{rZF} \left( \gamma^{(j)} \right) $$

Where $E_j$ is a row selection matrix

- Effective precoder is

$$ T^{DCSI} \triangleq \begin{bmatrix} w_1(\hat{H}^{(1)}) \\ w_2(\hat{H}^{(2)}) \\ \vdots \\ w_n(\hat{H}^{(n)}) \end{bmatrix} $$
Optimization of the Regularization Parameter

- **Naive regularization**
  \[ \gamma^{(j),\text{naive}} = \arg\max_{\gamma \in \mathbb{R}} \mathbb{E}[R(\hat{H}^{(j)}, \ldots, \hat{H}^{(j)})] \]

- **Robust regularization**
  \[ (\gamma^{(1),*}, \ldots, \gamma^{(n),*}) = \arg\max_{(\gamma^{(1)}, \ldots, \gamma^{(n)})} \mathbb{E}[R(\hat{H}^{(1)}, \ldots, \hat{H}^{(n)})]. \]

- **Low complexity robust** regularization with equal \(\gamma\) at all TXs
  \[ (\gamma^*, \ldots, \gamma^*) = \arg\max_{(\gamma, \ldots, \gamma)} \mathbb{E}[R(\hat{H}^{(1)}, \ldots, \hat{H}^{(n)})]. \]
Main Result (1)

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

In Joint Processing CoMP with Distributed CSI,

\[ \text{SINR}_k - \text{SINR}^o_k \xrightarrow{\text{a.s.}} \frac{K, M_{TX} \to \infty}{0} \]

with

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

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)

\[ I^o_k \triangleq P - P \sum_{j=1}^{n} \sum_{j'=1}^{n} \frac{\Gamma^o_{j,j'}}{\sqrt{\Gamma^o_{j,j} \Gamma^o_{j',j'}}} \min \left\{ \frac{2c_{0,k}^{(j)}}{n^2} \frac{\delta^{(j)}}{1+\delta^{(j)}} \frac{\left( \rho_k^{(j,j')} \right)^2 c_{2,k}^{(j)} c_{2,k}^{(j')} + c_{0,k}^{(j)} c_{0,k}^{(j')} \delta^{(j)} \delta^{(j')}}{n^2 \left( 1+\delta^{(j)} \right) \left( 1+\delta^{(j')} \right)} \right\} \]

with

\[ \Gamma^o_{j,j'} \triangleq \frac{1}{M} \sum_{\ell=1}^{K} \sqrt{c_{0,\ell}^{(j)} c_{0,\ell}^{(j')}} + \sqrt{c_{1,\ell}^{(j)} c_{1,\ell}^{(j')}} \left( \rho_{\ell}^{(j,j')} \right)^2 \]

\[ \sqrt{\frac{1+\delta^{(j)}}{\delta^{(j)}}} \frac{1+\delta^{(j')}}{\delta^{(j')}} - \frac{1}{M} \sum_{\ell=1}^{K} \left( \sqrt{c_{0,\ell}^{(j)} c_{0,\ell}^{(j')}} + \sqrt{c_{1,\ell}^{(j)} c_{1,\ell}^{(j')}} \left( \rho_{\ell}^{(j,j')} \right)^2 \right)^2 \]

and

\[ \delta^{(j)} \triangleq \beta - 1 - \gamma^{(j)} \beta + \sqrt{\left( \gamma^{(j)} \beta - \beta + 1 \right)^2 + 4 \gamma^{(j)} \beta^2} \]

\[ 2 \gamma^{(j)} \beta \]
Sanity Checks (1)

- Imperfect centralized CSIT:
  
  \[ \sigma_k^{(j)} = \sigma_k^{(j')} = \sigma_k, \]  
  \( \text{equal CSIT accuracy} \)

  \[ \rho_k^{(j,j')} = 1, \]  
  \( \text{Full correlation} \)

  \[ \gamma^{(j)} = \gamma^{(j')} = \gamma, \]  
  \( \text{Equal regularization} \)

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

  \[ \text{SINR}^{ID-DCSI, o}_k = \frac{(1 - \sigma_k^2)\delta^2}{\Gamma_o \left( 1 - \sigma_k^2 + (1 + \delta)^2 \sigma_k^2 + \frac{(1+\delta)^2}{P} \right)} \]

- Also obtained with \( n = 1 \)
Sanity Checks (2)

- Uncorrelated distributed CSIT with uniform accuracy and equal regularization:
  \[
  \sigma_k^{(j)} = \sigma^{(j)}, \quad (\text{Uniform CSIT})
  \]
  \[
  \rho_k^{(j,j')} = 0, \quad (\text{Uncorrelated})
  \]
  \[
  \gamma^{(j)} = \gamma^{(j')} = \gamma, \quad (\text{Equal regularization})
  \]

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

\[
{\text{SINR}}_{k}^{\text{EQ-DCSI},o} = \frac{P \left( \frac{1}{n} \sum_{j=1}^{n} \sqrt{c_{0,k}^{(j)}} \right)^2}{I_{k}^{\text{EQ-DCSI},o} + 1}
\]

with

\[
I_{k}^{\text{EQ-DCSI}} = P - P \sum_{j=1}^{n} \sum_{j'=1}^{n} \frac{\delta \Gamma_{j,j'}^{o}}{n^2(1 + \delta)^2 \Gamma_{j,j'}^{o}} \cdot \left[ 2c_{0,k}^{(j)} + \delta \left( 2c_{0,k}^{(j)} - c_{0,k}^{(j)}c_{0,k}^{(j')} \right) \right]
\]
Cost of Distributedness

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

68/127
Optimization of the Regularization Parameter

Figure: Average rate per user as a function of $\gamma$ for $\sigma^2(1) = 0$, $\sigma^2(2) = 0.1$, $\sigma^2(3) = 0.4$. 
### Simulation Settings

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>3</td>
<td>$K$</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>30</td>
<td>$\beta$</td>
<td>$M/K=1$</td>
<td></td>
</tr>
<tr>
<td>$(\sigma_k^{(1)})^2$</td>
<td>0.01</td>
<td>$(\sigma_k^{(2)})^2$</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>$(\sigma_k^{(3)})^2$</td>
<td>0.49</td>
<td>$\rho_k^{(j:j')}$</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>$h_k$</td>
<td>$\sim \mathcal{N}_{\mathbb{C}}(0, I_M)$</td>
<td>$\delta_k^{(j)}$</td>
<td>$\sim \mathcal{N}_{\mathbb{C}}(0, I_M)$</td>
<td></td>
</tr>
</tbody>
</table>

**Figure**: Simulation setting
Simulations: Optimize $\gamma$

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

1 Wireless Device Cooperation

2 Distributed Information Models

3 Coordination and Team decision: Problem formulation

4 Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
   - A Different Point of View: Implicit Coordination

5 Key Aspects and Open Problems
Asymptotical Analysis: DoF Approach

- First order approximation in the SNR

\[ R^* \approx \text{DoF} \log_2(SNR) \]

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


- Very successful 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 \cite{Davoodi and Jafar, 2014}

\[
\hat{H} = H + \sqrt{P^{-\alpha}} \Delta
\]

- Achieved using simple ZF precoding + rate splitting

\[
\text{DoF} = 1 + (K - 1)\alpha
\]
Distributed CSIT Configuration: $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(K)}$

\[ \hat{H}^{(1)} = H + \sqrt{P^{-\alpha^{(1)}}} \Delta^{(1)} \]

\[ \hat{H}^{(2)} = H + \sqrt{P^{-\alpha^{(2)}}} \Delta^{(2)} \]

\[ \hat{H}^{(3)} = H + \sqrt{P^{-\alpha^{(3)}}} \Delta^{(3)} \]

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]:
  \[
  \text{DoF}^{ZF} = 1 + (K - 1) \min_j \alpha^{(j)}
  \]
- 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** ($\alpha_1, \alpha_2, .. \alpha_K$) 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 $P^{(\alpha_1)}/9$
  \[ T^{MF(j)} \triangleq \frac{\hat{H}(j)}{\|\hat{H}(j)\|_F} \sqrt{P} \]
Reconstructing the Approximate Interference

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

\[
(\hat{h}_1^{(1)})^H T^{MF} s_2 = (h_1 + P^{-\alpha^{(1)}} \delta_1^{(1)})^H T^{MF} s_2 = h_1^H T^{MF} s_2 + P^{-\alpha^{(1)}} (\delta_1^{(1)})^H T^{MF} s_2 \sim P^0
\]

- 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

\[
\frac{6\alpha^{(1)} \log_2(P)}{\leq (1 - \alpha^{(1)}) \log_2(P)} \Leftrightarrow \alpha^{(1)} \leq \frac{1}{7}
\]

number of bits to quantize all interference terms rate of the broadcast data symbol

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

- DoF achieved is then

\[
\text{DoF} = \underbrace{9\alpha^{(1)}} + \underbrace{(1 - 7\alpha^{(1)})} = 1 + 2\alpha^{(1)}
\]

information transmitted initially fresh information bits to complete the broadcast

\[
\text{DoF} = 1 + (K - 1) \max \alpha^{(i)}
\]

(instead of DoF = 1 + (K - 1) \min \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 $T_{\text{MF}}(j) \in \mathbb{C}^{K \times K}$ 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 $t_{\text{BC}} \triangleq [1, 0, 0]^T$)

$$
\begin{align*}
P^{\alpha(1)} &= h_1^T t_{\text{BC}} s_0 + h_1^T T_{\text{MF}} s_1 + h_1^T T_{\text{MF}} s_2 + h_1^T T_{\text{MF}} s_3 \\
&= h_2^T t_{\text{BC}} s_0 + h_2^T T_{\text{MF}} s_1 + h_2^T T_{\text{MF}} s_2 + h_2^T T_{\text{MF}} s_3 \\
&= h_3^T t_{\text{BC}} s_0 + h_3^T T_{\text{MF}} s_1 + h_3^T T_{\text{MF}} s_2 + h_3^T T_{\text{MF}} s_3
\end{align*}
$$
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*

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

*We have that:*

$$
\text{DoF}^{\text{DCSI}}(\alpha) \geq 1 + (K - 1) \max_{j \in \{1,\ldots,K\}} \alpha(j)
$$
Outer bound

Theorem

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

$$\text{DoF}^{\text{DCSI}}(\alpha) \leq \text{DoF}^{\text{CCSI}}(\max_{j \in \{1, \ldots, K\}} \alpha^{(j)})$$

$$= 1 + (K - 1) \max_{j \in \{1, \ldots, K\}} \alpha^{(j)}$$

Key ideas:

- DoF is upperbounded by DoF achieved by full CSIT exchange
- Having multiple CSIT with $\alpha_1, \alpha_2, \ldots, \alpha_K$ doesn't help over having just best CSIT ($\alpha_1$)

$\Rightarrow$ 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

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

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

1 Wireless Device Cooperation

2 Distributed Information Models

3 Coordination and Team decision: Problem formulation

4 Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
   - A Different Point of View: Implicit Coordination

5 Key Aspects and Open Problems
Binary Power Control over Interference Channels

\[(p_1^*, p_2^*) = \arg\max_{(p_1, p_2) \in \mathcal{P}} [R(p_1(G^{(1)}), p_2(G^{(2)}))]\]

where

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

and

\[p_j : \mathbb{R}_+^4 \rightarrow \{P_j^\text{min}, P_j^\text{max}\} \quad G^{(j)} \mapsto p_j(G^{(j)})\]

- Key Example because:
  - Binary optimization with (relatively) low dimensional channel state
  - Weaker dependency with the channel state

\[\Rightarrow \text{Less difficulties}\]
Identification of the Parameters

Table: Team Decision Modeling for Power Control

<table>
<thead>
<tr>
<th>Notations for the Team Decision Problems</th>
<th>$x$</th>
<th>$G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>State-of-the-world</td>
<td>$x^{(i)}$</td>
<td>$G^{(i)}$</td>
</tr>
<tr>
<td>Estimate at DM $j$</td>
<td>$s_j$</td>
<td>$p_j$</td>
</tr>
<tr>
<td>Strategy at DM $j$</td>
<td>$A_j$</td>
<td>${P_j^{\text{min}}, P_j^{\text{max}}}$</td>
</tr>
<tr>
<td>Decision space at DM $j$</td>
<td>$f$</td>
<td>$R$</td>
</tr>
<tr>
<td>Objective</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Applications of Team Decision to Device-Centric Cooperation

Outline

1. Wireless Device Cooperation
2. Distributed Information Models
3. Coordination and Team decision: Problem formulation
4. Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
   - A Different Point of View: Implicit Coordination
5. Key Aspects and Open Problems
Discretization for Power Control –Refresher (1)–

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

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

\[
Q^{cb} : \mathbb{C}^{2 \times 2} \rightarrow \mathcal{C}^{cb} \triangleq \{G_{1cb}, \ldots, G_{p^{cb}}\} \\
\hat{G}(j) \mapsto \underset{\hat{G} \in \mathcal{C}^{cb}}{\text{argmin}} \| \hat{G} - \hat{G}(j) \|^2
\]

Optimization subspace reduced to a space of dimension \( n^{cb} \)

\[
p_j(Q^{cb}) : \mathcal{C}^{cb} \rightarrow \{P^{min}, P^{max}\} \\
G_i \mapsto p_j(G_i)
\]
Discretization for Power Control –Refresher (2)–

- **Best response power allocation strategy:** Solve iteratively
  
  - At TX 1, \( \forall i \in \{1, \ldots, n_{cb} \} \),
    
    \[
    p_{1}^{\text{BR}}(G_{i}^{cb}) = \arg\max_{P_1 \in \{P_{1}^{\text{min}}, P_{1}^{\text{max}} \}} \mathbb{E}\left[ R\left( G, P_1, p_2^{\text{BR}}(Q(G^{(2)})) | G^{(1)} = G_{i}^{cb} \right) \right]
    \]

  - At TX 2, \( \forall i \in \{1, \ldots, n_{cb} \} \),
    
    \[
    p_{2}^{\text{BR}}(G_{i}^{cb}) = \arg\max_{P_2 \in \{P_{2}^{\text{min}}, P_{2}^{\text{max}} \}} \mathbb{E}\left[ R\left( G, P_1^{\text{BR}}(Q(G^{(1)})), P_2 | G^{(2)} = G_{i}^{cb} \right) \right]
    \]

**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, $\forall i \in \{1, \ldots, n^{cb}\}$,

$$p_1^{BR}(G_i^{cb}) = \arg\max_{P_1 \in \{P_1^{min}, P_1^{max}\}} \frac{1}{n^{MC}} \sum_{i=1}^{n^{MC}} R \left( G_i, P_1, p_2^{BR}(Q^{cb}(G_i^{(2)})) \right), \quad \forall i \in \{1, \ldots, n^{cb}\}$$

where $(G_i, G_i^{(2)}) \sim f_{G,G^{(2)}}|_{G^{(1)}=G_i^{cb}}$. 
Simulations Parameters

- Rayleigh fading with uniform pathloss
- CSIT Model

\[ H^{(j)} \triangleq \sqrt{1 - \sigma_j^2H + \sigma_j\Delta} \]

where \( \Delta \sim \mathcal{N}_C(0,1) \) and \( H \sim \mathcal{N}_C(0,1) \), and

\[ \{G^{(j)}\}_{i,k} \triangleq \left|\{H^{(j)}\}_{i,k}\right|^2, \quad \forall i, k \in \{1, \ldots, K\}. \]

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

1 Wireless Device Cooperation

2 Distributed Information Models

3 Coordination and Team decision: Problem formulation

4 Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
     - A Different Point of View: Implicit Coordination

5 Key Aspects and Open Problems
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 $h_{s,s}$ and multi-user statistics $R_{i,j}$
  - TX $p$ only knows $h_{p,p}$ and multi-user statistics $R_{i,j}$

- Coordination using only the statistics

\[ W_p \quad ??? \quad W_s \quad ??? \]
Functional Optimization Problem

Optimization Problem

\[(w_p^*, w_s^*) = \arg\max_{(w_p, w_s)} \mathbb{E}[R_s(w_p(h_{p,p}), w_s(h_{s,s}))]\]

s. to \(\mathbb{E}[R_p(w_p(h_{p,p}), w_s(h_{s,s}))] \geq \tau > 0, \quad (R)\)

\[0 \leq \|w_p(h_{p,p})\|^2 \leq P_p^{\text{max}}\]

\[0 \leq \|w_s(h_{s,s})\|^2 \leq P_s^{\text{max}}\]

where for \(j \in \{s, p\},\)

\[w_j : \mathbb{C}^{M_j} \rightarrow \mathbb{C}^{M_j}\]

\[h_{j,j} \mapsto w_j(h_{j,j})\]
Identification of the Parameters

**Table:** Team Decision Modeling for Cognitive Radio

<table>
<thead>
<tr>
<th>Notations for the Team Decision Problems</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>State-of-the-world</td>
<td>$x$</td>
</tr>
<tr>
<td>Estimate at DM $j$</td>
<td>$x^{(j)}$</td>
</tr>
<tr>
<td>Strategy at DM $j$</td>
<td>$s_j$</td>
</tr>
<tr>
<td>Decision space at DM $j$</td>
<td>$\mathcal{A}_j$</td>
</tr>
<tr>
<td>Objective</td>
<td>$f$</td>
</tr>
<tr>
<td></td>
<td>$R_s$</td>
</tr>
<tr>
<td></td>
<td>s.t. $(R)$</td>
</tr>
</tbody>
</table>

$x = \{ h_{s,s}, h_{s,p}, h_{p,s}, h_{p,p} \}$
Outline

1 Wireless Device Cooperation
2 Distributed Information Models
3 Coordination and Team decision: Problem formulation
4 Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
     - A Different Point of View: Implicit Coordination
5 Key Aspects and Open Problems
Key Idea: Codebook of Functions [Filippou et al., 2016, TWC]

Functional optimization difficult ⇒ 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

\[
(w^*_p, w^*_s) = \arg\max_{(w_p, w_s)} \mathbb{E} [R_s(w_p(h_{p,p}), w_s(h_{s,s}))]
\]

s. to \[ \mathbb{E} [R_p(w_p(h_{p,p}), w_s(h_{s,s}))] \geq \tau > 0, \]
\[ 0 \leq \|w_p(h_{p,p})\|^2 \leq P_p^{\text{max}} \]
\[ 0 \leq \|w_s(h_{s,s})\|^2 \leq P_s^{\text{max}} \]
\[ (w^*_p, w^*_s) \in \{ (w_{cb,1}^{p}, w_{cb,1}^{s}), \ldots, (w_{cb,n_{cb}}^{p}, w_{cb,n_{cb}}^{s}) \} \]

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

\[
w_{j,1}^{cb,j} : \mathbb{C}^{M_j} \rightarrow \mathbb{C}^{M_j}
\]
\[ h_{j,j} \mapsto w_{j}^{cb}(h_{j,j}) \]
Strategy $p$

- TX $p$
  - TX $p$ uses matched precoding
    \[ u_p^{\text{MF}} \triangleq \frac{h_{p,p}}{\|h_{p,p}\|} \]
  - TX $p$ transmits with full power $\bar{P}_p = P_{p}^{\text{max}}$

- TX $s$:
  - TX $s$ transmits using the statistical ZF precoder
    \[ u_s^{s\text{ZF}} \triangleq \arg\min_u u^H R_{p,s} u \]
  - TX $s$ controls its average transmit power $\bar{P}_s$ to fulfill the rate constraint $\langle R \rangle$
**Strategy** $s$

- **TX $s$**
  - TX $s$ uses matched precoding
    \[
    u_s^{MF} \triangleq \frac{h_{s,s}}{\|h_{s,s}\|}
    \]
  - TX $s$ transmits with full power $\bar{P}_s = P_s^{\max}$

- **TX $p$:**
  - TX $p$ transmits using the statistical ZF precoder
    \[
    u_p^{ZF} \triangleq \arg\min_u u^H R_{s,p} u
    \]
  - 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
  
  $R_{p,p} = R_{s,s} = I_3$, \quad $R_{p,s} = R_{s,p} = \begin{bmatrix}
  1 & \rho & \rho^2 \\
  \rho & 1 & \rho \\
  \rho^2 & \rho & 1
\end{bmatrix}$

- Use in the following $\rho = 0.5$ and $\tau = 0.5 \text{bps/Hz}$
Ergodic rate of the PU

![Graph showing the ergodic rate of the PU with different transmit SNR values for centralized, statistically coordinated, and SU cognitive beamforming.]
Ergodic rate of the SU

![Graph showing the ergodic rate of the SU for different transmit SNR values.

- Centralized
- Statistically coordinated
- SU Cognitive beamforming

The graph plots the average SU rate in bps/Hz against transmit SNR in dB.]
Outline

1. Wireless Device Cooperation
2. Distributed Information Models
3. Coordination and Team decision: Problem formulation
4. Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
   - A Different Point of View : Implicit Coordination
5. Key Aspects and Open Problems
Coded Power Control

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

![Diagram showing TX1, RX1, TX2, RX2, CSI feedback, and observation arrows between TX1 and TX2, with implicit coordination arrow]
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:
  $w_{1,i} : \mathcal{H}^T \times \mathcal{P}_1^{i-1} \times \emptyset \rightarrow \mathcal{P}_1$
  \[\text{(CSI, past actions)}\]
- Strategy of Agent 2: $(w_{2,i})_{1\leq i \leq T}$ with:
  $w_{2,i} : \mathcal{H}^{i-1} \times \mathcal{Z}^{i-1} \times \mathcal{P}_2^{i-1} \rightarrow \mathcal{P}_2$
  \[\text{(past)}\]
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 \( Q(h, p_1, p_2) \) is implementable if there exists a pair of strategies \((w_{1, i}, w_{2, i})_{i \geq 1}\) such that for all \((h, p_1, p_2)\),

\[
\frac{1}{T} \sum_{i=1}^{T} \sum_{y} P_{H_i, P_1, i, P_2, i, Y_i}(h, p_1, p_2, z) \rightarrow Q(h, p_1, p_2)
\]
as \(T \rightarrow \infty\).

Feasible utilities

A certain utility value \(\underline{f}\) is reachable if and only if there exists an implementable distribution \(Q\) such that \(\underline{f} = \mathbb{E}_Q[f]\).
Theorem ([Larrousse and Lasaulce, 2013])

Let $Q \in \Delta(\mathcal{H} \times \mathcal{P}_1 \times \mathcal{P}_2)$ with $\sum_{(p_1, p_2)} Q(h, p_1, p_2) = \rho(h)$. The distribution $Q$ 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 $Q(h, p_1, p_2, y) = Q(h, p_1, p_2) \Gamma(z|p_1)$.

Remark: This theorem also characterizes expected payoff.
Convex Optimization Problem

\[
\begin{align*}
\text{maximize} \quad & \mathbb{E}_q[f] = \sum_{\ell=1}^{L} q_{\ell} f_{\ell} \\
\text{subject to} \quad & I_q(H; P_2) \leq I_q(P_1; Z|H, P_2) \\
& q_{\ell} \geq 0 \\
& \sum_{\ell=1}^{L} q_{\ell} = 1 \\
& \sum_{\ell \in \mathcal{L}_H(h)} q_{\ell} = \rho(h), \quad \forall \, h, \\
& \frac{\sum_{\ell \in \mathcal{L}_{P_1, Z}(p_1, z)} q_{\ell}}{\sum_{\ell \in \mathcal{L}_{P_1}(p_1)} q_{\ell}} = \Gamma(z|p_1), \quad \forall \, (p_1, z)
\end{align*}
\]
Application: MAC Power Control

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

\[
\begin{array}{cc}
P_{\text{min}} & P_{\text{max}} \\
P_{\text{min}} & 0 & 20 \\
P_{\text{max}} & 1 & \approx 10
\end{array}
\]

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

\[
\begin{array}{cc}
P_{\text{min}} & P_{\text{max}} \\
P_{\text{min}} & 0 & 1 \\
P_{\text{max}} & 20 & \approx 10
\end{array}
\]

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

**Figure:** Payoff table [Larrousse and Lasaulce, 2013]
Source Coding [Larrousse and Lasaulce, 2013]

Source coding:

\[ f_S : \mathcal{H} \rightarrow \{ m_0, m_1 \} \]
\[ h \mapsto i \]

TABLE IV
OPTIMAL MARGINAL AND JOINT DISTRIBUTIONS (EXPRESSED IN %) FOR THE SUM-RATE PAYOFF FUNCTION OF THE CPC POLICY, WITH SNR = 10 dB AND WITH FOUR POSSIBLE TRANSMIT POWER LEVELS \{0, 10/3, 20/3, 10\}.

TABLE V
PROPOSED SOURCE CODING AND DECODING FOR \( p = \frac{1}{2} \).

<table>
<thead>
<tr>
<th>( x_0^3 )</th>
<th>Index ( i = f_S(x_0^3) )</th>
<th>( g_S(i) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0,0)</td>
<td>( m_0 )</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>(0,0,1)</td>
<td>( m_0 )</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>(0,1,0)</td>
<td>( m_0 )</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>(0,1,1)</td>
<td>( m_0 )</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>(1,0,0)</td>
<td>( m_0 )</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>(1,0,1)</td>
<td>( m_0 )</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>(1,1,0)</td>
<td>( m_1 )</td>
<td>(0,0,1)</td>
</tr>
<tr>
<td>(1,1,1)</td>
<td>( m_1 )</td>
<td>(0,0,1)</td>
</tr>
</tbody>
</table>
Channel Coding

- Channel coding:

\[ f_C : \{m_0, m_1\} \rightarrow P_1 \]

\[ i \mapsto p_1 = [p_1(1), p_1(2), p_1(3)] \]

- at block \( b \), find optimal \( p_1^{\text{opt}} = [p_1(1)^{\text{opt}}, p_1(2)^{\text{opt}}, p_1(3)^{\text{opt}}] \) and second optimal \( p_1^{\text{opt}''} = [p_1(1)^{\text{opt}''}, p_1(2)^{\text{opt}''}, p_1(3)^{\text{opt}''}] \)

  - If \( i = m_0 \), send with \( p_1^{\text{opt}} \)
  - If \( i = m_1 \), send with \( p_1^{\text{opt}''} \)
**Channel Coding** [Larrousse and Lasaulce, 2013]

<table>
<thead>
<tr>
<th>$x_0^3(b)$</th>
<th>$x_2^3(b)$</th>
<th>$i_{b+1}$</th>
<th>$x_1^3(b)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0,0)</td>
<td>(1,1,1)</td>
<td>$m_0$</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td></td>
</tr>
<tr>
<td>(0,0,1)</td>
<td>(1,1,1)</td>
<td>$m_0$</td>
<td>(0,0,1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>(0,1,0)</td>
<td>(1,1,1)</td>
<td>$m_0$</td>
<td>(0,1,0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>(0,1,1)</td>
<td>(1,1,1)</td>
<td>$m_0$</td>
<td>(0,1,1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td>(0,0,1)</td>
</tr>
<tr>
<td>(1,0,0)</td>
<td>(1,1,1)</td>
<td>$m_0$</td>
<td>(1,0,0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>(1,0,1)</td>
<td>(1,1,1)</td>
<td>$m_0$</td>
<td>(1,0,1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td>(0,0,1)</td>
</tr>
<tr>
<td>(1,1,0)</td>
<td>(0,0,1)</td>
<td>$m_0$</td>
<td>(1,1,0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>(1,1,1)</td>
<td>(0,0,1)</td>
<td>$m_0$</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m_1$</td>
<td>(1,1,0)</td>
</tr>
</tbody>
</table>
Simulations [Larrousse and Lasaulce, 2013]
Outline

1 Wireless Device Cooperation

2 Distributed Information Models

3 Coordination and Team decision: Problem formulation

4 Applications of Team Decision to Device-Centric Cooperation
   - Application to Network MIMO Precoding
     - Model-Based Approach
     - DoF Approach
   - Application to Power Control
     - Functional Optimization by Discretization
   - Application to Cognitive Radio Beamforming
     - Codebook-Based Approach
   - A Different Point of View : Implicit Coordination

5 Key Aspects and Open Problems
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
thanks
References I


