Understanding 5Gs Signal-Processing Demands: Device Centric Network Cooperation for 5G and Beyond: Theory and Algorithms

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

December 12th, 2016
Network Coordination

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

- Robots networks (autonomous drones, smart factory, plant probes, etc.)
- Transportation networks (driver less cars, truck trains, etc.)
- Sensor networks
- Processor networks
- Energy (Smart Grids) networks
- Wireless networks
Cooperation Scenarios in Wireless Network (5G and beyond)

Cooperation turns a resource usage conflict into a system gain, for instance in following scenarios:

1. Coordinated Multipoint (CoMP) transmission
2. Power control for interference reduction
3. Spectrum sublicensing (coordinated cognitive radios)
4. Beam alignment for massive MIMO in mmwave bands
5. Dynamic content caching
6. Inter-cell Interference Coordination (ICIC and eICIC)
7. Coordinated power transfer for battery life extension (IoT)
8. and more...
Joint Processing CoMP

- $n$ otherwise-interfering base stations jointly combine their $nM_{TX}$ antenna elements over ideal backhaul
- $K$ users are served simultaneously, free of interference (with $K$ up to $nM_{TX}$)

$$\text{Spatial Multiplexing Gain} = nM_{TX}$$

**Challenges:**
- All base stations must be synchronized and acquire knowledge of all served users’ channels
- All base stations must acquire knowledge of all served users’ channels
Power Control for Interference Reduction

- Neighboring base stations interfere and coordinate their power control policies
- Power control is subject to a maximum power constraint
- Optimum policy aims at maximizing the overall throughput, i.e. just the right amount of interference is generated
- **Challenges**: Coordination requires knowledge of all channel strengths $G_{ij}$
Spectrum Sub-licensing using Cognitive Radio Beamforming

- A primary operator (p) is sub-licensing its spectrum to a secondary operator (s)
- Both operator base stations control a beamforming vector

\[
\text{Maximize } \mathbb{E}[R_s] \quad \text{subject to } \mathbb{E}[R_p] \geq \tau > 0
\]

Secondary

Primary

Most common approach: Interference temperature constraint \( I_{\text{primary}} \leq \tau \)

Challenge: Full beamforming coordination requires centralized knowledge of primary and secondary channels.
Centralized VS Decentralized Signal Processing Architectures in 5G and Beyond

- Cloud RAN is popular, pushes for more centralization
- Centralized decision making is conceptually simple
- Coordination is easy
- Mobile service providers love it
Centralized vs. Decentralized Signal Processing Architectures in 5G and Beyond

- Centralization leads to **expensive** deployment (road digging, fiber,..)
- Backhaul architectures can be of **diverse nature**
- Curse of **dimension** (IoT: billions of devices)
- More centralization increases **latency**, decreases timeliness of **CSI**
Cooperation in LTE with Heterogeneous Backhaul: A Device Centric Perspective

Sharing/caching of user’s data symbols

Imperfect CSI sharing

\[ x_1 = w_1 (H^{(1)}) s \]
\[ x_2 = w_2 (H^{(2)}) s \]
\[ x_3 = w_3 (H^{(3)}) s \]
CSI for Device-Centric Cooperation

- CSI affected by mobility, limited training and feedback
- CSI exchange is **not** free
  - Devices are *myopic*: They know better what is close
  - Need for local (*device-centric*) decision-making
CSI for Device-Centric Cooperation

- CSI affected by mobility, limited training and feedback
- CSI exchange is not free
  - Devices are myopic: They know better what is close
  - Need for local (device-centric) decision-making

Today’s questions:
1. Distributed information models?
2. Price of myopia?
3. Myopia-robust approaches?
Outline

1 Distributed Information Models

2 Device-Centric Cooperation: Formulation and methods

3 Applications of Team Decision to Device-Centric Cooperation
Centralized VS Distributed Channel State Information

- Centralized (TX Independent)

Cloud RAN
\[ \hat{H} = H + \sigma N \]
Centralized VS Distributed Channel State Information

- **Centralized (TX Independent)**

  \[
  \hat{H} = H + \sigma N
  \]

- **Distributed (TX Dependent)**

  \[
  H^{(1)} = H + \sigma^{(1)} N^{(1)}
  
  H^{(2)} = H + \sigma^{(2)} N^{(2)}
  
  H^{(3)} = H + \sigma^{(3)} N^{(3)}
  \]
Outline

1 Distributed Information Models

2 Device-Centric Cooperation: Formulation and methods

3 Applications of Team Decision to Device-Centric Cooperation
Device-centric Coordination $\iff$ Team Decision

$K$ devices cooperate to maximize network performance $f$

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

where

- $x \in \mathbb{C}^m$: System State (for wireless: $x = H$)
- $x^{(j)} \in \mathbb{C}^m$: Observation of the state of the world $x$ at device $j$
- $s_j : \mathbb{C}^m \rightarrow A_j \subset \mathbb{C}^{d_j}$: Decision policy at device $j$
- $p_{x, x^{(1)}, \ldots, x^{(K)}}$: Joint probability distribution of the channel and the estimates
Device-Centric Cooperation: Formulation and methods

Team Decision theory

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

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

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

Distributed optimization

- 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.]
One-Shot Team Decision: Algorithm Design

- Quantizing the Policy Space
- Quantizing the Observation Space
- Model-Based Approach

Team Decision Methods

- Information Allocation
- Information Sharing
Outline

1. Distributed Information Models
2. Device-Centric Cooperation: Formulation and methods
3. Applications of Team Decision to Device-Centric Cooperation
First Application: Joint Processing CoMP

- Find precoders \( \{w_j\}_{j=1}^K \):

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

  - \( w_j \) being the precoder at TX \( j \)

  \[
  w_j : \mathbb{C}^{N_{\text{tot}} \times M_{\text{tot}}} \rightarrow \mathbb{C}^{M_j \times d_{\text{tot}}} 
  \]

  \[
  \hat{H}(j) \rightarrow w_j(\hat{H}(j)) 
  \]

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

- \( K \) and \( M_{\text{TX}} \) grow large at the same rate \( \beta \triangleq \lim_{M,K \to \infty} \frac{M}{K} \geq 1 \): Asymptotic analysis
Large Random Matrix Theory in Wireless Networks [A Short Digression]

Example

Let $A$ be the matrix of size $n \times n$ defined as

$$A \triangleq \begin{bmatrix}
0 & \pm 1 & \pm 1 \\
\pm 1 & 0 & \pm 1 \\
\pm 1 & \pm 1 & 0
\end{bmatrix}$$

Convergence of the eigenvalue distribution

Figure: from [Tulino and Verdu, 2004]

- Application to Wireless Networks for more than 10 years: Asymptotic expressions for SINR, rate, power, BER,...
- Made more relevant by Massive MIMO technology!
- Many books and lecture notes [Tulino and Verdu, 2004] [Couillet and Debbah, 2011]
Model-Based Optimization: Introduction of Regularized Zero Forcing

Modelization of the precoder using Regularized Zero Forcing:

\[ w_j(\hat{H}^{(j)}) = \begin{bmatrix} 0_j, 1, 0_{M-j} \end{bmatrix} \left( (\hat{H}^{(j)})^H \hat{H}^{(j)} + M \gamma^{(j)} I_M \right)^{-1} (\hat{H}^{(j)})^H \frac{\sqrt{P}}{\sqrt{\psi^{(j)}}} \]

- **Intuition**: Distributed (regularized) Channel Inversion
- **Tikhonov Regularization of channel inversion** [Golub et al., 2016], widely used [Shenouda and Davidson, 2006]

How to find optimization parameter \( \gamma^{(j)} \) at TX \( j \)?
Optimization of the Regularization Parameter $\gamma^{(j)}$

- **Myopic** regularization

  \[
  \gamma^{(j),\text{Myopic}} = \arg\max_{\gamma \in \mathbb{R}} \mathbb{E}[R(\hat{H}^{(j)}, \ldots, \hat{H}^{(j)})] 
  \]

- **Team-Based** regularization

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

- **Low Complexity Team-Based** regularization (Equal coefficient at all TXs)

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

- RMT allows to get rid of the expectation operator in the optimization
Performance of CoMP Transmission with Distributed CSIT

<table>
<thead>
<tr>
<th>Antenna Setting</th>
<th>n</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Channel Modeling</td>
<td>Fading</td>
<td>Rayleigh</td>
</tr>
<tr>
<td>Pathloss</td>
<td>Uniform</td>
<td></td>
</tr>
<tr>
<td>CSIT Configuration</td>
<td>$(\sigma_k^{(1)})^2$</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>$(\sigma_k^{(2)})^2$</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>$(\sigma_k^{(3)})^2$</td>
<td>0.49</td>
</tr>
<tr>
<td>$\rho_{j,j'}$</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>
II/ Second Application: On-Off Power Control

- Power control to reduce interference of two interfering wireless links:

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

where \( p_j \) is the power control function

\[
p_j : \mathbb{R}_+^4 \rightarrow \{ P_j^{\text{min}}, P_j^{\text{max}} \}
\]

\[
p_j(G^{(j)}) \rightarrow p_j(G^{(j)})
\]
Discretization of the Observation Space [de Kerret and Gesbert, 2016, SPAWC]

- Replace the strategy \( p_j(\hat{G}^{(j)}) \) by \( p_j(\text{Quant}(\hat{G}^{(j)})) \)

\[ \text{belongs to a codebook of size } n \]

- Optimizing a function over a **discrete set** is more easy than a **continuous one**
Best Response Optimization

- Solve iteratively
  - At TX 1, $\forall G_i \in \{G_1^{\text{Quant}}, \ldots, G_n^{\text{Quant}}\}$,
    \[ p_1^{\text{BR}} = \arg\max_{p_1} \mathbb{E}[R(p_1(G^{(1)}), p_2^{\text{BR}}(G^{(2)}))] \]
  - At TX 2, $\forall G_i \in \{G_1^{\text{Quant}}, \ldots, G_n^{\text{Quant}}\}$,
    \[ p_2^{\text{BR}} = \arg\max_{p_2} \mathbb{E}[R(p_1^{\text{BR}}(G^{(1)}), p_2(G^{(2)}))] \]

- Made possible by the discretization of the observation space
Simulations of On-Off Power Control with Local Feedback

<table>
<thead>
<tr>
<th>Channel Modeling</th>
<th>Fading</th>
<th>Rayleigh</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pathloss</td>
<td>Uniform</td>
</tr>
<tr>
<td>Algorithm parameters</td>
<td>Codebook size for quantization</td>
<td>$10^4$</td>
</tr>
<tr>
<td></td>
<td>Number of Monte-Carlo runs</td>
<td>500</td>
</tr>
<tr>
<td>CSIT Configuration</td>
<td>$\sigma^{(1)}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$\sigma^{(2)}$</td>
<td>0</td>
</tr>
</tbody>
</table>

![Graph showing average rate vs. transmit power for different power control methods.]
III/ Third Application: Cognitive Radio Beamforming with Local Feedback

Maximize $\mathbb{E}[R_s]$ subject to $\mathbb{E}[R_p] \geq \tau > 0$

- **CSI configuration**
  - Primary TX only knows $h_{p,p}$
  - Secondary TX only knows $h_{s,s}$

- **SOTA**: Primary user is oblivious of the secondary user

**Coordination scheme**

Primary TX adapts without any exchange of instantaneous information
Applications of Team Decision to Device-Centric Cooperation

Robust Distributed Optimization

Optimization Problem (P)

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

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

- \( w_p \) is the beamforming function at the primary TX

\[
w_p : \mathbb{C}^{M_p} \rightarrow \mathbb{C}^{M_p}
\]

\[
h_{p,p} \mapsto w_p(h_{p,p})
\]

- \( w_s \) is the beamforming function at the secondary TX

\[
w_s : \mathbb{C}^{M_s} \rightarrow \mathbb{C}^{M_s}
\]

\[
h_{s,s} \mapsto w_s(h_{s,s})
\]
Primary Friendly (PF) Strategy

- **Primary TX**: uses Matched Filtering with full power \( \bar{P}_p = P_p^{\text{max}} \)

\[
    u_p^{(PF)} \triangleq \frac{h_{p,p}}{\|h_{p,p}\|}
\]

- **Secondary TX**: uses the statistical Zero Forcing beamforming

\[
    u_s^{(PF)} \triangleq \arg\min_u u^H R_{p,s} u
\]

and average transmit power \( \bar{P}_s \) to fulfill the ergodic rate constraint
### Secondary Friendly (SF) Strategy

- **Secondary TX**: uses Matched Filtering with full power $\bar{P}_s = P_s^{\text{max}}$

  $$u_s^{(\text{SF})} \triangleq \frac{h_{s,s}}{\| h_{s,s} \|}$$

- **Primary TX**: uses the statistical Zero Forcing beamformer

  $$u_p^{(\text{SF})} \triangleq \arg\min_u u^H R_{s,p} u$$

  and average transmit power $\bar{P}_p$ to fulfill the ergodic rate constraint
Quantizing the Policy Space [Filippou et al., 2016, TWC]

- Restrict to 2 strategies labeled **Primary Friendly (PF) and Secondary Friendly (SF)**
- Need good heuristic choices

**Optimization Problem**

\[
(w_p^*, w_s^*) = \underset{(w_p, w_s) \in \mathcal{W}}{\text{argmax}} \mathbb{E}[R_s(w_p(h_{p,p}), w_s(h_{s,s}))]
\]

s. to \( \mathbb{E}[R_p(w_p(h_{p,p}), w_s(h_{s,s}))] \geq \tau > 0 \), \hspace{1cm} (P)

\[
\mathcal{W} = \left\{ (w_p^{(PF)}, w_s^{(PF)}), (w_p^{(SF)}, w_s^{(SF)}) \right\}
\]

- First Strategy
- Second Strategy
Cognitive Radio with Local Feedback: Rate of the Secondary User

Figure: Ergodic rate of the Primary User

- $M_S = M_p = 3$ antennas per-TX
- Correlation matrices
  
  $R_{p,p} = R_{s,s} = I_3,$
  
  $R_{p,s} = R_{s,p} = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$
  
  Use in the following $\rho = 0.5$ and $\tau = 0.5\text{bps/Hz}$

Figure: Ergodic rate of the Secondary User
Cognitive Radio with Local Feedback: Rate of the Primary User

- $M_s = M_p = 3$ antennas per-TX
- Correlation matrices
  \[
  R_{p,p} = R_{s,s} = I_3, \\
  R_{p,s} = R_{s,p} = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}
  \]
- Use in the following $\rho = 0.5$ and $\tau = 0.5$ bps/Hz

Figure: Ergodic rate of the Primary User

Figure: Ergodic rate of the Secondary User
Take home

- Device coordination is key to performance improvement in 5G and beyond
- Virtually all coordination schemes require extensive CSI acquisition and sharing among devices
- Coordination frameworks that are robust to CSI locality are desirable
- Several perspectives on the problem (i) control, (ii) signal processing, (iii) information theoretic

More applications (not covered here)
- Dynamic content caching at device side
- Coordinated beam alignment in millimeter wave Massive MIMO
- Coordinated power transfer for battery recharge in IoT networks
- More examples upon request
thanks
References I


