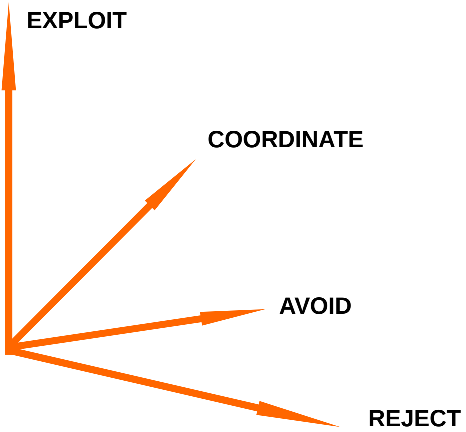
DISCRIMINATING AGAINST INTERFERENCE IN MASSIVE MIMO SYSTEMS: A STATISTICAL APPROACH

FUELING THE DENSE VS. MASSIVE DEBATE

David Gesbert,

EURECOM,
gesbert@eurecom.fr
Joint work with Haifan Yin

June 10, 2013


EXPLOIT
COORDINATE
AVOID
REJECT

Exploiting and coordinating interference

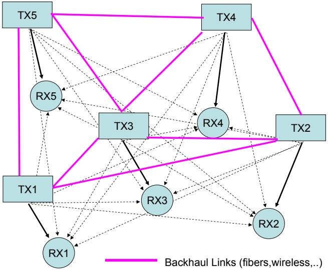
Transmitter cooperation involve substantial information sharing

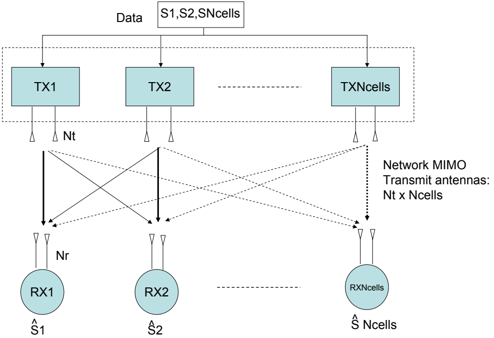
- **Coordination**: transmitters exchange CSIT
  - Coordinated beamforming (CoMP in LTE-A), interference alignement, coordinated scheduling, coordinated power control..
- **Exploitation**: transmitters exchange CSIT and user data
  - Network (multicell) MIMO, Joint Processing CoMP
- **Rejection**: Simple per-terminal per-cell processing, little info exchange

Some questions:

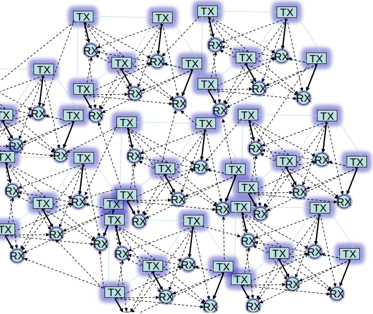
- Are such methods scalable?
- Do distributed implementation exist?

EXPLOITING INTERFERENCE VIA MULTICELL MIMO



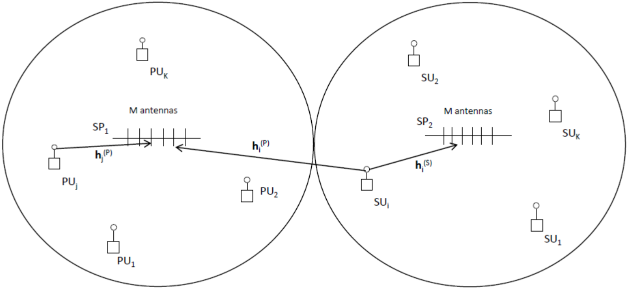


Dense cooperation (single antenna base station)



Massive MIMO base station (no cooperation)

WHAT CAN SIMPLE DISTRIBUTED BEAMFORMING ACHIEVE?



- Let M antennas be used at BS 1 and BS 2.
- As $M \to \infty$ (normalized) useful and interference channel vector become quasi orthogonal
- Matched filter maximizes SNR *and* cancels interference *simultaneously* [Marzetta 2010]
- Matched filter solution is fully distributed!

But there is a problem...

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

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

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

Least Squares (LS) estimator with full pilot reuse:

$$\widehat{\mathbf{h}}_1^{\mathrm{LS}} = \mathbf{h}_1 + \sum_{l \neq 1}^{L} \mathbf{h}_l + \mathbf{N}\mathbf{s}^* / \tau \tag{2}$$

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

where

$$\mathbf{R} \triangleq \text{diag}(\mathbf{R}_1, \cdots, \mathbf{R}_L) \tag{3}$$

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

$$B \triangleq \pi^{LM}(\det\mathbf{R})^M \tag{4}$$

Develop covariance-based (Bayesian or MMSE) estimator

$$\widehat{\mathbf{h}}_1 = \mathbf{R}_1\left(\sigma_n^2\mathbf{I}_M + \tau\sum_{l=1}^{L}\mathbf{R}_l\right)^{-1}\mathbf{S}^H\mathbf{y} \tag{5}$$

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

Theorem [Yin, Gesbert, Filippou, Liu JSAC 2013]
Assume multipath angle-of-arrival θ for user j (at target BS 1) has density $p_j(\theta)$ with bounded support, i.e. $p_j(\theta) = 0$ for $\theta \notin [\theta_j^{\text{min}}, \theta_j^{\text{max}}]$ for some fixed $\theta_j^{\text{min}} \leqslant \theta_j^{\text{max}} \in [0, \pi]$. If the $L-1$ intervals $[\theta_i^{\text{min}}, \theta_i^{\text{max}}]$, $i = 2, \ldots, L$ are strictly non-overlappipng with $[\theta_1^{\text{min}}, \theta_1^{\text{max}}]$, we have

$$\lim_{M \to \infty} \widehat{\mathbf{h}}_1 = \widehat{\mathbf{h}}_1^{\text{no int}} \tag{6}$$

If desired and interference multipath ranges do no overlap, pilot contamination vanish asymptotically.

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

where P is number of paths and

$$\mathbf{a}(\theta) \triangleq \begin{bmatrix} 1 \\ e^{-j2\pi \frac{D}{\lambda}\cos(\theta)} \\ \vdots \\ e^{-j2\pi \frac{(M-1)D}{\lambda}\cos(\theta)} \end{bmatrix} \qquad (7)$$

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

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

Proof relies on three lemmas:

Lemma 1:

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

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

lemma 2 *When M grows large,*

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

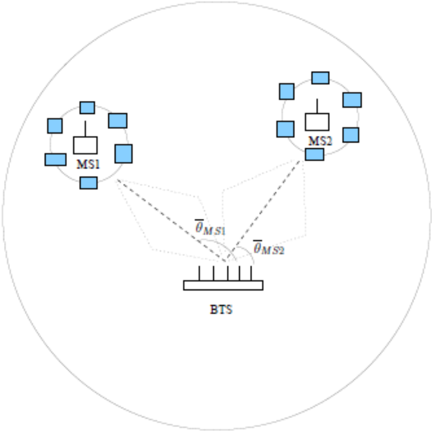
where

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

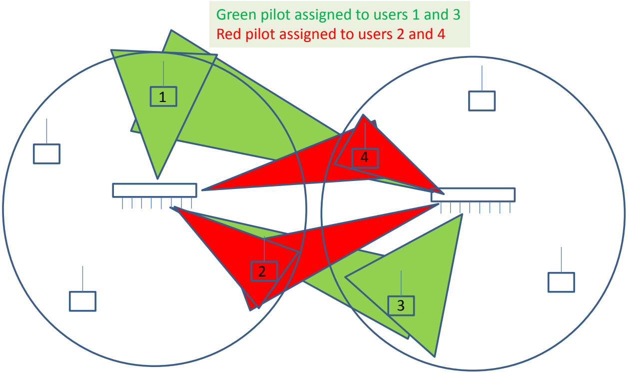
Lemma 1 indicates that for large M, there exists a null space $\text{null}(\mathbf{R}_i)$ of dimension $(1 - d_i)M$.

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

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


MS1
MS2
$\overline{\theta}_{MS1}$
$\overline{\theta}_{MS2}$
BTS

DECONTAMINATING PILOTS PUT TO PRACTICE



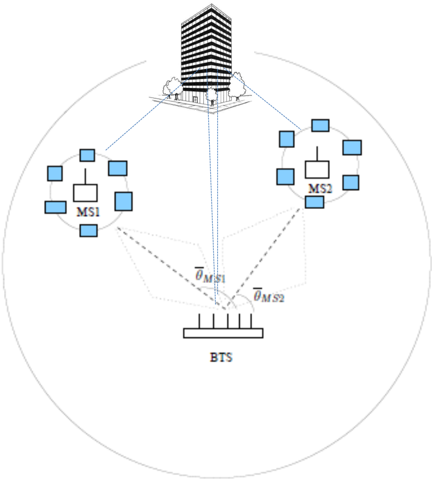
Coordinated Pilot Assignement (CPA):

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

$$\mathrm{F}(\mathcal{U}) \triangleq \sum_{j=1}^{|\mathcal{U}|} \frac{\mathcal{M}_j(\mathcal{U})}{\mathrm{tr}\left\{\mathbf{R}_{jj}(\mathcal{U})\right\}} \tag{8}$$

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

THE SKYSCRAPER EFFECT



Angle spread 10 degrees

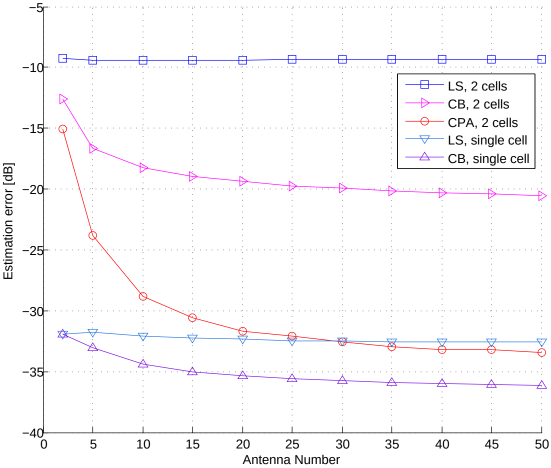


FIGURE: Estimation MSE vs. antenna number, Gaussian distributed AOAs with $\sigma = 10$ degrees.

10 Antennas

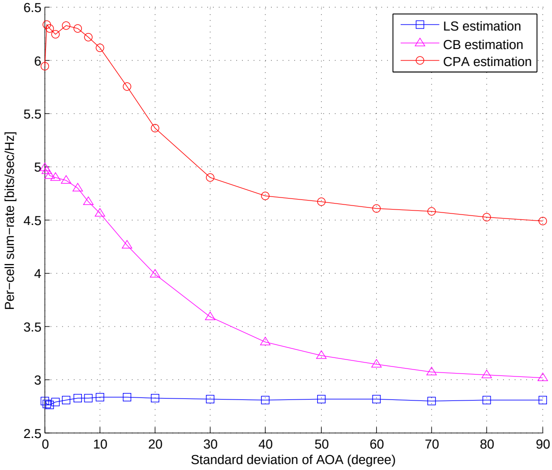


FIGURE: Per-cell sum-rate vs. standard deviation of AOA (Gaussian distribution) with $M = 10$, 7-cell network.

Uplink channel estimation with reused pilots:

$$\mathbf{Y} = \mathbf{h}_1\mathbf{s}^T + \mathbf{h}_2\mathbf{s}^T + \mathbf{N}, \quad (9)$$

Define the null space of $\mathbf{R}_2$:

$$\mathbf{R}_2 = \mathbf{U}\Sigma\mathbf{U}^H \quad \mathbf{W}_1 = [\mathbf{u}_{m+1}|\mathbf{u}_{m+2}|\ldots|\mathbf{u}_M]^H \quad (10)$$

Assume $\mathbf{h}_1 \in$ null space of $\mathbf{R}_2$, then $\mathbf{h}_1 = \mathbf{W}_1^H\underline{\mathbf{h}}_1$ where $\underline{\mathbf{h}}_1$ is the effective channel. The subspace-based channel estimate is

$$\widehat{\mathbf{h}}_1 = \mathbf{W}_1^H\widehat{\underline{\mathbf{h}}}_1 = \mathbf{W}_1^H\mathbf{W}_1\mathbf{Y}\mathbf{s}^*(\mathbf{s}^T\mathbf{s}^*)^{-1} \quad (11)$$

Note 1: One can also use the fact that $\mathbf{h}_1 \in$ signal subspace of $\mathbf{R}_2$.
Note 2: These properties can be exploited for feedback reduction in FDD context (Adhikary, Caire 2012).

Angle spread 30 degrees+random scheduler ⇒ very poor performance!

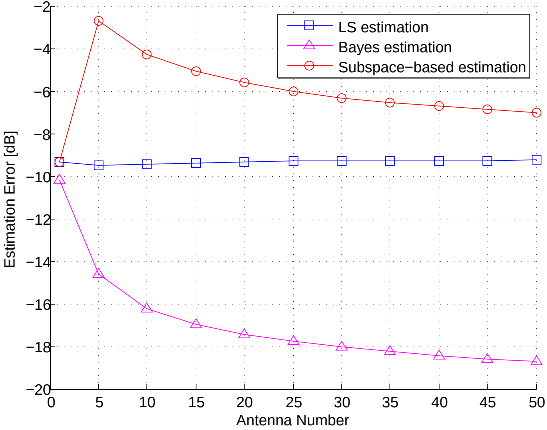


FIGURE: Estimation MSE vs. antenna number, uniformly distributed AOAs with $\theta_\Delta = 30$ degrees, 2-cell network.

Is it really necessary to have a good channel estimate? No!

Uplink received data:

$$\mathbf{y} = \mathbf{h}_1\mathbf{s}_1^T + \mathbf{h}_2\mathbf{s}_2^T + \mathbf{n}, \tag{12}$$

The subspace-based MRC beamformer is $\widehat{\underline{\mathbf{h}}}_1^H \mathbf{W}_1$

$$\widehat{\underline{\mathbf{h}}}_1^H \mathbf{W}_1 \mathbf{y} = \widehat{\underline{\mathbf{h}}}_1^H \underline{\mathbf{h}}_1 \mathbf{s}_1^T + \underbrace{\widehat{\underline{\mathbf{h}}}_1^H \mathbf{W}_1 \mathbf{h}_2 \mathbf{s}_2^T}_{\approx \mathbf{0}} + \widehat{\underline{\mathbf{h}}}_1^H \mathbf{W}_1 \mathbf{n} \tag{13}$$

Subspace-based massive-MIMO beamformer yields good interference reduction signal enhancement trade-off...

Angle spread 30 degrees. Worst channel estimate yields best data rate!

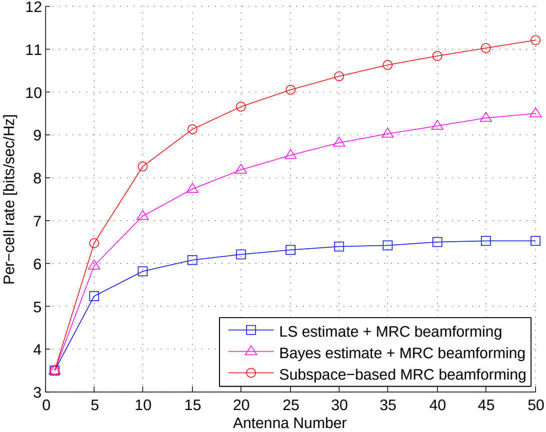


FIGURE: Per-cell rate vs. antenna number, uniformly distributed AOAs with $\theta_\Delta = 30$ degrees, 2-cell network.

PERSPECTIVES

- Massive MIMO leads to strongly subspace structured covariances
- subspace orthogonality can be exploited for pilot decontamination, beamforming design, feedback reduction
- Orthogonality can be boosted with the help of coordinated pilot assignement and user scheduling
- Open issues: estimation of covariance matrices, random antenna arrays, dealing with skyscraper effects ...