Sparse Activity Detection in Multi-Cell Massive MIMO Exploiting Channel Large-Scale Fading

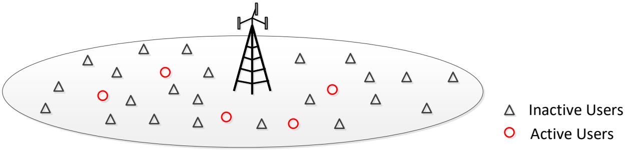
Wei Yu

Joint Work with Zhilin Chen, Foad Sohrabi

University of Toronto

November 2021

Massive Random Access for Internet-of-Things (IoT)



- A large number of devices with sporadic activity
- Low latency random access scheme for massive users is required
- Non-orthogonal signature sequences need to be used
- User activity detection (user identification) performed at base station (BS)

This talk: User activity detection problem for a multi-cell system.

System Model

- BS equipped with M antennas
- N single-antenna devices, K of which are active at a time
- Each device is associated with a length-L unique signature sequence $\mathbf{s}_n$
- Channel $\mathbf{h}_n$ of user n includes both (i.i.d.) Rayleigh and large-scale fading
- For single-cell system, received signal $\mathbf{Y} \in \mathbb{C}^{L \times M}$ at the BS is

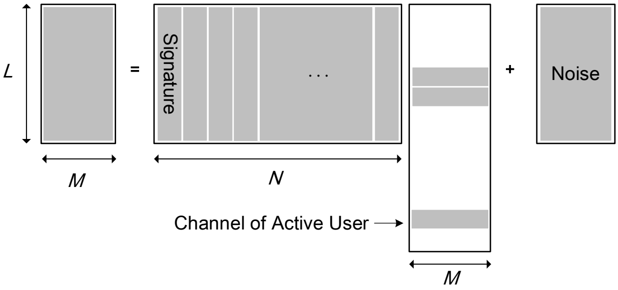
$$\mathbf{Y} = \sum_{n=1}^{N} \alpha_n \mathbf{s}_n \mathbf{h}_n^T + \mathbf{Z} = \mathbf{S}\mathbf{X} + \mathbf{Z}, \tag{1}$$

where

- $\alpha_n \in \{1, 0\}$ activity indicator; $\mathbf{Z} \in \mathbb{C}^{L \times M}$ Gaussian noise with variance σ^2
- $\mathbf{S} \triangleq [\mathbf{s}_1, \ldots, \mathbf{s}_N] \in \mathbb{C}^{L \times N}$; $\mathbf{X} \triangleq [\alpha_1 \mathbf{h}_1, \cdots, \alpha_N \mathbf{h}_N]^T \in \mathbb{C}^{N \times M}$

Joint Sparse Activity Detection and Channel Estimation

Aim to identify the K non-zero rows of $\mathbf{X}$ from $\mathbf{Y} = \mathbf{S}\mathbf{X} + \mathbf{Z}$.



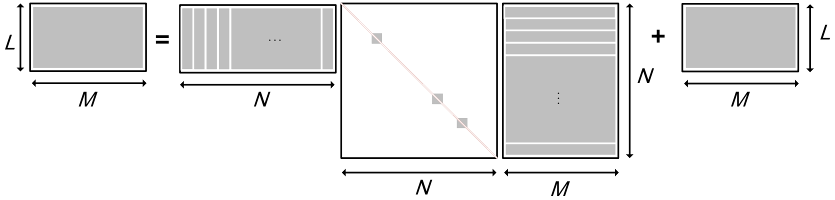
- Multiple measurement vector (MMV) problem in compressed sensing
  - Columns of $\mathbf{X}$ share the same sparsity pattern, i.e., row sparsity
- Efficiently solved by the approximate message passing (AMP) algorithm
- Detecting KM variables based on LM observations.

Joint Activity Detection and Large-Scale Fading Estimation

Key Assumption: We only need activity α_n and do not need $\mathbf{h}_n$.

Reformulate sparse activity detection as a large-scale-fading estimation problem:

$$\mathbf{Y} = \sum_{n=1}^{N} \alpha_n \mathbf{s}_n \mathbf{h}_n^T + \mathbf{Z} \triangleq \mathbf{S}\mathbf{\Gamma}^{\frac{1}{2}}\tilde{\mathbf{H}} + \mathbf{Z} \tag{2}$$



- $\mathbf{S} \triangleq [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_N] \in \mathbb{C}^{L\times N}$, signature matrix
- $\mathbf{\Gamma} \triangleq \text{diag}\{\alpha_1\beta_1, \alpha_2\beta_2, \cdots, \alpha_N\beta_N\} \in \mathbb{R}^{N\times N}$, where β_n is large-scale fading
- $\tilde{\mathbf{H}} \triangleq \left[\mathbf{h}_1/\sqrt{\beta_1}, \mathbf{h}_2/\sqrt{\beta_2}, \cdots, \mathbf{h}_N/\sqrt{\beta_N}\right]^T \in \mathbb{C}^{N\times M}$, normalized channel matrix

Maximum Likelihood Estimate of Activities and LS Fading

The maximum likelihood estimation of $\mathbf{\Gamma}$ can be expressed as

$$\min_{\mathbf{\Gamma}\geq\mathbf{0}} -\log p(\mathbf{Y}|\mathbf{\Gamma}) \propto \min_{\mathbf{\Gamma}\geq\mathbf{0}} \; -\frac{1}{M}\sum_{m=1}^{M}\left(\log\frac{1}{|\pi\mathbf{\Sigma}|}\exp\left(-\mathbf{y}_m^H\mathbf{\Sigma}^{-1}\mathbf{y}_m\right)\right) \tag{3}$$

$$= \min_{\mathbf{\Gamma}\geq\mathbf{0}} \; \log|\mathbf{\Sigma}| + \mathrm{tr}\left(\mathbf{\Sigma}^{-1}\hat{\mathbf{\Sigma}}\right) + const.$$

$$= \min_{\mathbf{\Gamma}\geq\mathbf{0}} \; \log|\mathbf{S}\mathbf{\Gamma}\mathbf{S}^H + \sigma^2\mathbf{I}| + \mathrm{tr}\left((\mathbf{S}\mathbf{\Gamma}\mathbf{S}^H + \sigma^2\mathbf{I})^{-1}\hat{\mathbf{\Sigma}}\right) + const. \tag{4}$$

Sample covariance matrix over the antennas is a sufficient statistics:

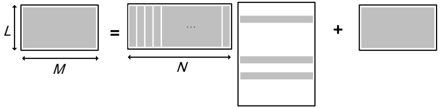
$$\hat{\mathbf{\Sigma}} \triangleq \frac{1}{M}\sum_{m=1}^{M}\mathbf{y}_m\mathbf{y}_m^H = \frac{1}{M}\mathbf{Y}\mathbf{Y}^H. \tag{5}$$

The above problem can be solved efficiently using coordinate descent.

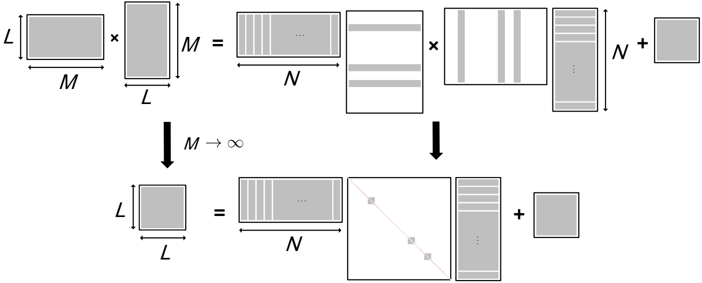
A. Fengler, S. Haghighatshoar, P. Jung, and G. Caire: "Non-Bayesian Activity Detection, Large-Scale Fading Coefficient Estimation, and Unsourced Random Access with a Massive MIMO Receiver", *T-IT*, May 2021.

Covariance Based Sparse Activity Detection

Instead of jointly estimating the channel, i.e., the non-zero rows in $\mathbf{X}$ based on $\mathbf{Y}$:



We now estimate large-scale fading $\mathbf{\Gamma}$ based on $\hat{\mathbf{\Sigma}} = \frac{1}{M}\mathbf{Y}\mathbf{Y}^H$:



Crucial Advantage: We now detect K variables based on L^2 observations!

Analysis of the Maximum Likelihood Estimate of $\mathbf{\Gamma}$

Theorem

Let $\mathcal{I}$ be an index set corresponding to zero entries of $\boldsymbol{\gamma}^0$, i.e., $\mathcal{I} \triangleq \{i \mid \gamma_i^0 = 0\}$. We define two sets $\mathcal{N}, \mathcal{C}$ in the space $\mathbb{R}^N$, respectively, as follows

$$\mathcal{N} \triangleq \{\mathbf{x} \mid \mathbf{x}^T \mathbf{J}(\boldsymbol{\gamma}^0)\mathbf{x} = 0, \mathbf{x} \in \mathbb{R}^N\}, \tag{6}$$

$$\mathcal{C} \triangleq \{\mathbf{x} \mid x_i \geq 0, i \in \mathcal{I}, \mathbf{x} \in \mathbb{R}^N\}, \tag{7}$$

where x_i is the i-th entry of $\mathbf{x}$. Then a necessary and sufficient condition for the consistency of $\hat{\boldsymbol{\gamma}}^{ML}$, i.e., $\hat{\boldsymbol{\gamma}}^{ML} \to \boldsymbol{\gamma}^0$ as $M \to \infty$, is $\mathcal{N} \cap \mathcal{C} = \{\mathbf{0}\}$.

$\mathcal{N}$ is the "null space" of the Fisher Information Matrix $\mathbf{J}(\boldsymbol{\gamma}^0)$;
$\mathcal{C}$ is a cone with non-negative entries indexed by the inactive users $\mathcal{I}$.

This condition leads to a phase analysis for the covariance based method, i.e., set of (N, L, K) outside of which $\hat{\boldsymbol{\gamma}}^{ML}$ cannot approach $\boldsymbol{\gamma}^0$ even in the large M limit.

Phase Transition of Covariance Approach

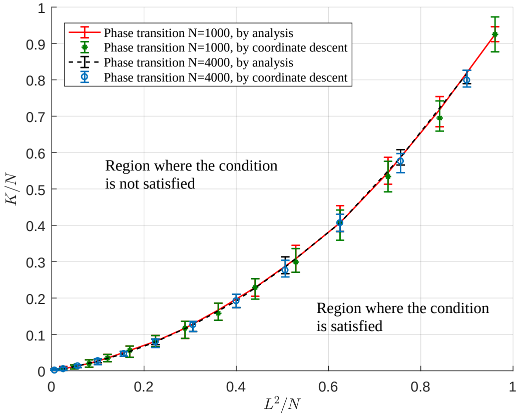


Figure: Phase transition in the space of N, L, K. All users are located at the cell-edge (1000m) with transmit power 23dBm. Path-loss is $128.1 + 37.6 \log(d[\text{km}])$. Generated by 100 Monte Carlo simulations. Error bars indicate the range below which all 100 realizations satisfy the condition and above which none satisfies the condition.

Activity Detection with Known Large-Scale Fading

- Suppose that large-scale fading is known, activity detection now becomes over the binary indicators α_n:

$$\min_{\{\alpha_n\}} \quad \log\left|\mathbf{S}\boldsymbol{\Gamma}\mathbf{S}^H + \sigma^2\mathbf{I}\right| + \operatorname{tr}\left((\mathbf{S}\boldsymbol{\Gamma}\mathbf{S}^H + \sigma^2\mathbf{I})^{-1}\hat{\boldsymbol{\Sigma}}\right) \tag{8a}$$

$$\text{s.t.} \quad \alpha_n \in \{0,1\},\ n = 1,2,\ldots,N \tag{8b}$$

- Binary α_n is challenging to deal with. We relax the constraint such that

$$\alpha_n \in [0,1], \qquad n = 1,2,\ldots,N \tag{9}$$

- The relaxed problem can be solved by coordinate descent:

$$d = \min\left\{\max\left\{\frac{\mathbf{s}_k^H\tilde{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{s}_k - \mathbf{s}_k^H\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{s}_k}{\beta_k(\mathbf{s}_k^H\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{s}_k)^2}, -\hat{\alpha}_k\right\}, 1-\hat{\alpha}_k\right\}. \tag{10}$$

- With unknown large-scale fading β_n, we estimate $\gamma_n = \alpha_n\beta_n$ in $[0,\infty]$. With known large-scale fading β_n, we estimate α_n in $[0,1]$.

Value of Knowing Large-Scale Fading

The covariance method directly estimates the activity indicator α_n in $[0,1]$ instead of $\gamma_n = \alpha_n \beta_n$ in $[0,\infty)$. Let $\boldsymbol{\alpha} \triangleq [\alpha_1, \ldots, \alpha_N]^T$ and let the true value of $\boldsymbol{\alpha}$ be $\boldsymbol{\alpha}^0$.

Theorem

Let $\mathcal{I}$ be an index set corresponding to zero entries of $\boldsymbol{\alpha}^0$, i.e., $\mathcal{I} \triangleq \{i \mid \alpha_i^0 = 0\}$. We define two sets $\mathcal{N}', \mathcal{C}'$ in the space $\mathbb{R}^N$, respectively, as follows

$$\mathcal{N}' \triangleq \{\mathbf{x} \mid \mathbf{x}^T \mathbf{J}(\boldsymbol{\gamma}^0)\mathbf{x} = 0, \mathbf{x} \in \mathbb{R}^N\}, \tag{11}$$

$$\mathcal{C}' \triangleq \{\mathbf{x} \mid x_i \geq 0, i \in \mathcal{I}, x_i \leq 0, i \notin \mathcal{I}, \mathbf{x} \in \mathbb{R}^N\}, \tag{12}$$

where x_i is the i-th entry of $\mathbf{x}$. Then a necessary and sufficient condition for the consistency of $\hat{\boldsymbol{\alpha}}^{ML}$, i.e., $\hat{\boldsymbol{\alpha}}^{ML} \to \boldsymbol{\alpha}^0$ as $M \to \infty$, is $\mathcal{N}' \cap \mathcal{C}' = \{\mathbf{0}\}$.

The extra constraint in the definition of $\mathcal{C}$ is due to that α_n is upper bounded.

Value of Knowing Large-Scale Fading

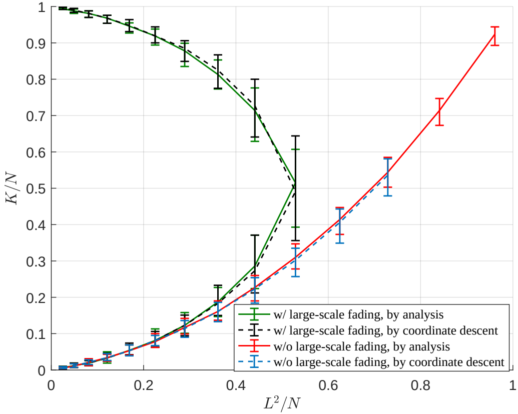
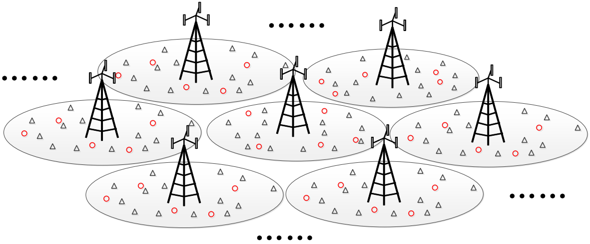


Figure: Phase transition comparison of the cases with and without knowing large-scale fading. $N = 1000$. With known large-scale fading, α_n is both lower and upper bounded.

When $\frac{K}{N} \approx 1$, then inactive users are sparse!

User Activity Detection in Multicell Systems

- What is the impact of the inter-cell interference?



- How to overcome the inter-cell interference?

Activity Detection in Multicell Systems

- Multi-cell system with B BSs each equipped with M antennas;
- N single-antenna devices per cell, K of which are active;
- Device n in cell b is assigned a length-L unique signature sequence $\mathbf{s}_{bn}$;
- Received signal $\mathbf{Y}_b \in \mathbb{C}^{L\times M}$ at BS b is

$$\begin{aligned}
\mathbf{Y}_b &= \sum_{n=1}^{N} \alpha_{bn}\mathbf{s}_{bn}\mathbf{h}_{bbn}^T + \sum_{j=1,j\neq b}^{j=B} \sum_{n=1}^{N} \alpha_{jn}\mathbf{s}_{jn}\mathbf{h}_{bjn}^T + \mathbf{Z}_b \\
&= \mathbf{S}_b\mathbf{A}_b\mathbf{G}_{bb}^{\frac{1}{2}}\tilde{\mathbf{H}}_{bb} + \sum_{j=1,j\neq b}^{j=B} \mathbf{S}_j\mathbf{A}_j\mathbf{G}_{bj}^{\frac{1}{2}}\tilde{\mathbf{H}}_{bj} + \mathbf{Z}_b \\
&= \mathbf{S}_b\mathbf{\Gamma}_{bb}^{\frac{1}{2}}\tilde{\mathbf{H}}_{bb} + \sum_{j=1,j\neq b}^{j=B} \mathbf{S}_j\mathbf{\Gamma}_{bj}^{\frac{1}{2}}\tilde{\mathbf{H}}_{bj} + \mathbf{Z}_b
\end{aligned} \tag{13}$$

where

- $\alpha_{bn} \in \{1,0\}$ activity indicator; $\mathbf{A}_j \triangleq \text{diag}\{\alpha_{j1}, \alpha_{j2}, \cdots, \alpha_{jN}\} \in \{0,1\}^{N\times N}$;
- $\mathbf{h}_{bjn} \in \mathbb{C}^{M\times 1}$ is the channel from user n in cell j to BS b
- $\tilde{\mathbf{H}}_{bj} \triangleq \left[\mathbf{h}_{bj1}/\sqrt{\beta_{bj1}}, \cdots, \mathbf{h}_{bjN}/\sqrt{\beta_{bjN}}\right]^T \in \mathbb{C}^{N\times M}$, normalized channel
- $\mathbf{S}_j \triangleq [\mathbf{s}_{j1}, \mathbf{s}_{j2}, \cdots, \mathbf{s}_{jN}] \in \mathbb{C}^{L\times N}$; $\mathbf{G}_{bj} \triangleq \text{diag}\{\beta_{bj1}, \beta_{bj2}, \cdots, \beta_{bjN}\} \in \mathbb{R}^{N\times N}$

Cooperative Activity Detection via Covariance Approach

- Assume that each BS is equipped with a large-scale antenna array.
- **Cooperative detection**: To alleviate the impact of inter-cell interference, we further consider BS cooperation by assuming all BSs are connected to a CU.
- Depending on whether the large-scale fading matrices $\mathbf{G}_{bj}, \forall b, j$ are known, the device activity detection problem can be formulated differently.
- When $\mathbf{G}_{bj}$ are not known, we need to estimate $\mathbf{\Gamma}_{bj} = \mathbf{A}_j \mathbf{G}_{bj}$, $\forall b, j$, which has

$$\boxed{B^2 N \text{ unknown parameters}}$$

- When $\mathbf{G}_{bj}$ are known, we only need to estimate $\mathbf{A}_b, \forall b$, which contains

$$\boxed{BN \text{ unknown parameters}}$$

Device activity detection is much easier if large-scale fading is known!

Cooperative Detection with Unknown Large-scale Fading

We aim to estimate $\mathbf{\Gamma}_{bj} = \mathbf{A}_j \mathbf{G}_{bj}, \forall b, j$ from the received signals $\mathbf{Y}_b, \forall b$.
The likelihood function of $\mathbf{Y}_b$'s given $\mathbf{\Gamma}_{bj}$'s can be expressed as

$$\begin{aligned} p(\mathbf{Y}_1, \ldots, \mathbf{Y}_B | \mathbf{\Gamma}_{11}, \mathbf{\Gamma}_{12}, \ldots, \mathbf{\Gamma}_{BB}) &= \prod_{b=1}^{B} p(\mathbf{Y}_b | \mathbf{\Gamma}_{11}, \mathbf{\Gamma}_{12}, \ldots, \mathbf{\Gamma}_{BB}) \\ &= \prod_{b=1}^{B} \frac{1}{|\pi \mathbf{\Sigma}_b|^M} \exp\left(-\operatorname{tr}\left(M \mathbf{\Sigma}_b^{-1} \hat{\mathbf{\Sigma}}_b\right)\right). \quad (14) \end{aligned}$$

The MLE problem can be cast as minimization of negative log-likelihood:

$$\min_{\{\mathbf{\Gamma}_{bj}\}} \quad \sum_{b=1}^{B} \left(\log|\mathbf{\Sigma}_b| + \operatorname{tr}\left(\mathbf{\Sigma}_b^{-1} \hat{\mathbf{\Sigma}}_b\right)\right) \quad (15a)$$

$$\text{s.t.} \quad \gamma_{bjn} \in [0, \infty), \forall b, j, n \quad (15b)$$

This is a challenging problem to solve.

Cooperative Detection with Known Large-scale Fading

Assuming that all large-scale fading matrices $\mathbf{G}_{bj}$'s, we directly estimate the device activity $\mathbf{A}_b$'s using the MLE. The likelihood function of $\mathbf{Y}_b$'s can be expressed as

$$
\begin{aligned}
p(\mathbf{Y}_1,\ldots,\mathbf{Y}_B|\mathbf{A}_1,\ldots,\mathbf{A}_B) &= \prod_{b=1}^{B} p(\mathbf{Y}_b|\mathbf{A}_1,\ldots,\mathbf{A}_B) \\
&= \prod_{b=1}^{B} \frac{1}{|\pi\boldsymbol{\Sigma}_b|^M} \exp\left(-\operatorname{tr}\left(M\boldsymbol{\Sigma}_b^{-1}\hat{\boldsymbol{\Sigma}}_b\right)\right). \qquad (16)
\end{aligned}
$$

Since the activity α_{bn} is binary, the maximization of likelihood can be cast as

$$
\min_{\{\mathbf{A}_b\}} \quad \sum_{b=1}^{B} \left(\log|\boldsymbol{\Sigma}_b| + \operatorname{tr}\left(\boldsymbol{\Sigma}_b^{-1}\hat{\boldsymbol{\Sigma}}_b\right)\right) \qquad (17a)
$$

$$
\text{s. t.} \quad \alpha_{bn} \in \{0,1\}, \forall b, n \qquad (17b)
$$

Known large-scale fading: Single-cell and multicell have same phase transition

- Multicell problem: Find a BN-dim sparse vector in $(BN - BL^2)$-dim subspace.
- Single-cell problem: Find a N-dim sparse vector in $(N - L^2)$-dim subspace.

Multicell Activity Detection: Known Large-Scale Fading

Let $\underline{\boldsymbol{\alpha}} \triangleq [\alpha_1, \ldots, \alpha_{BN}]^T$ be the activity indicators, and let true value of $\underline{\boldsymbol{\alpha}}$ be $\underline{\boldsymbol{\alpha}}^0$.

Theorem

Let $\underline{\mathcal{I}}$ be an index set corresponding to zero entries of $\underline{\boldsymbol{\alpha}}^0$, i.e., $\underline{\mathcal{I}} \triangleq \{i \mid \underline{\alpha}_i^0 = 0\}$. We define two sets $\mathcal{N}'', \mathcal{C}''$ in the space $\mathbb{R}^{BN}$, respectively, as follows

$$\mathcal{N}'' \triangleq \{\mathbf{x} \mid \mathbf{x}^T \mathbf{J}(\gamma^0)\mathbf{x} = 0, \mathbf{x} \in \mathbb{R}^{BN}\}, \tag{18}$$

$$\mathcal{C}'' \triangleq \{\mathbf{x} \mid x_i \geq 0, i \in \underline{\mathcal{I}}, x_i \leq 0, i \notin \underline{\mathcal{I}}, \mathbf{x} \in \mathbb{R}^{BN}\}, \tag{19}$$

where x_i is the i-th entry of $\mathbf{x}$. Then a necessary and sufficient condition for the consistency of $\hat{\underline{\boldsymbol{\alpha}}}^{ML}$, i.e., $\hat{\underline{\boldsymbol{\alpha}}}^{ML} \to \underline{\boldsymbol{\alpha}}^0$ as $M \to \infty$, is $\mathcal{N}'' \cap \mathcal{C}'' = \{\mathbf{0}\}$.

All the dimensions scale by B, so we expect the same phase transition curve.

Phase Transition of Multicell vs. Single-Cell

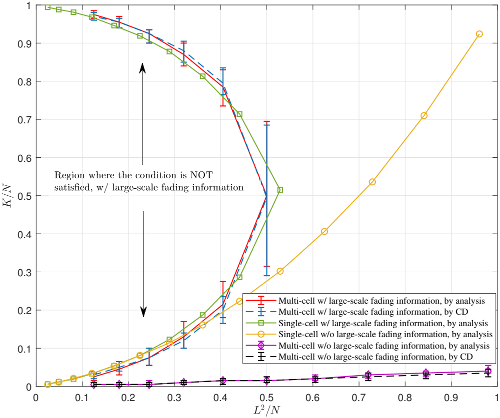


Figure: Phase transition comparison for multicell system with $B = 7$, $N = 200$.

Performance of Covariance Based Detection for Multi-cell

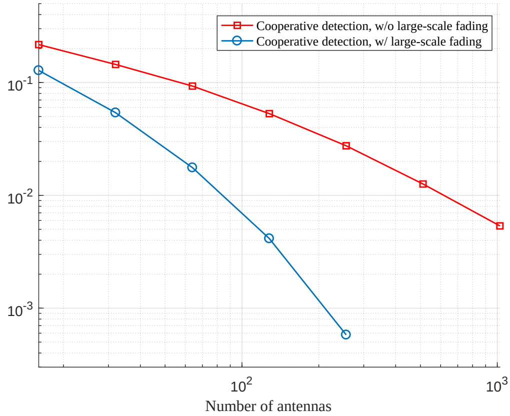


Figure: Performance comparison of the multicell covariance approach with and without knowing large-scale fading. $B = 7$, $N = 200$, $K = 20$, and $L = 20$. We observe that knowing the large-scale fading brings substantial improvement.

Performance of Covariance Based Detection for Multi-cell

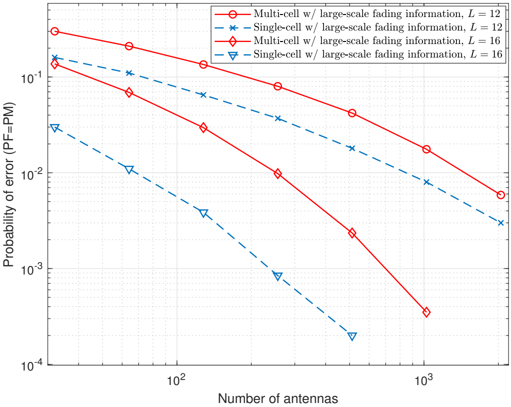


Figure: Performance comparison of the multicell covariance approach with single-cell system, with knowing large-scale fading. $B = 7$, $N = 200$, $K = 20$.

Conclusions

- Device activity detection for massive random access in machine-type and IoT communications is a sparse recovery problem.
- Covariance based MLE can detect activities without estimating the channels.
- Analysis and algorithm can be extended from single-cell to multicell systems.
- Multicell cooperative detection has a similar phase transition as single-cell system, *but only if the large-scale fading is known*.
- In practice, there is a performance degradation due to intercell interference.

Further Information

Zhilin Chen, Foad Sohrabi, Ya-Feng Liu, Wei Yu,

"Phase Transition Analysis for Covariance Based Massive Random Access with Massive MIMO",

[Online] available: https://arxiv.org/abs/2003.04175
Submitted to IEEE Transactions on Information Theory, March 2020.

Zhilin Chen, Foad Sohrabi, and Wei Yu,

"Sparse Activity Detection in Multi-Cell Massive MIMO Exploiting Channel Large-Scale Fading",

IEEE Transactions on Signal Processing, June 2021.