Superimposed Pilots : An Alternative Pilot Structure to Mitigate Pilot Contamination in Massive MIMO

Karthik Upadhya †, Sergiy A. Vorobyov†, Mikko Vehkapera‡

†Department of Signal Processing and Acoustics
Aalto University, School of Electrical Engineering
Espoo, Finland

‡Department of Electronic and Electrical Engineering
The University of Sheffield
Sheffield, UK

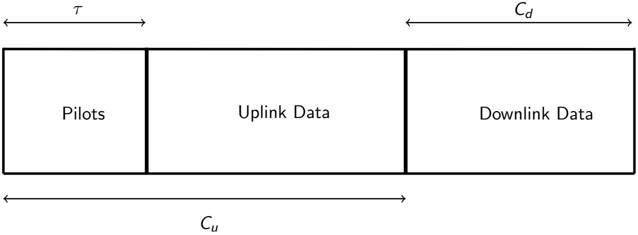
ICASSP 2016
03.22.2016
Shanghai, China

Outline

- Channel Estimation and Pilot Contamination in Massive MIMO
- Channel Estimation Using Superimposed Pilots
- Iterative Channel Estimation
- Simulation Results
- Some Comments on Performance

Channel Estimation in Massive MIMO

- Uplink time slot is partitioned for transmitting uplink pilots and data



- Orthogonal pilots are used for estimating the channel
- Pilots are shared between neighboring cells

Channel Estimation in Massive MIMO

- Received pilot sequence at the j^{th} BS (implicitly assumed in equations) with M antennas and K users per cell

$$\mathbf{Y} = \sum_{\ell=0}^{L-1} \mathbf{H}_\ell \mathbf{\Phi}^T + \mathbf{W}$$

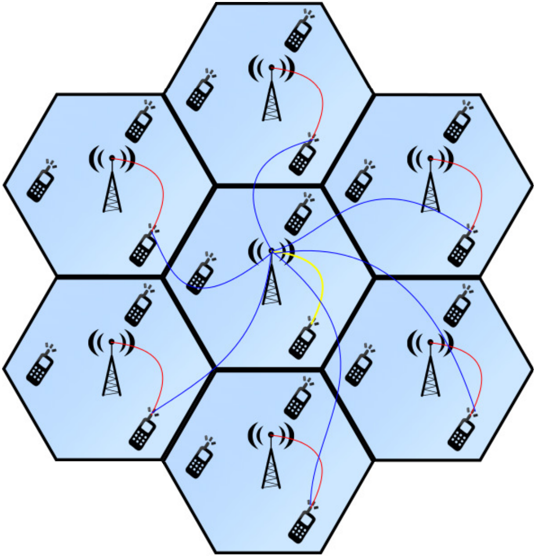
$\mathbf{H}_\ell$: Channel matrix between the j^{th} BS and users in the ℓ^{th} cell
$\mathbf{\Phi}$: $\tau \times K$ Matrix of orthogonal pilot sequences ($\mathbf{\Phi}^T \mathbf{\Phi}^* = \tau \mathbf{I}$)
$\mathbf{W}$: Additive white Gaussian noise at j^{th} BS

- LS channel estimate

$$\widehat{\mathbf{h}}_{j,m}^{\text{CP}} \triangleq \frac{1}{\tau}\mathbf{Y}\boldsymbol{\phi}_m^* = \mathbf{h}_{j,m} + \sum_{\substack{\ell=0 \\ \ell \neq j}}^{L-1} \mathbf{h}_{\ell,m} + \frac{1}{\tau}\mathbf{W}\boldsymbol{\phi}_m^*$$

$\mathbf{h}_{\ell,k}$: k^{th} column of $\mathbf{H}_\ell$, $\mathbf{h}_{\ell,k} \sim \mathcal{CN}(0, \beta_{\ell,k}\mathbf{I})$
$\boldsymbol{\phi}_m$: m^{th} column of $\mathbf{\Phi}$

Pilot Contamination



Pilot Contamination

- SINR in the uplink and downlink

$$\underset{M\to\infty}{\mathrm{SINR}_{j,m}^{\mathrm{CP-ul}}} = \underset{M\to\infty}{\mathrm{SINR}_{j,m}^{\mathrm{CP-dl}}} = \frac{\beta_{j,m}^2}{\sum_{\ell\neq j}\beta_{\ell,m}^2}$$

$\beta_{\ell,k}$: Large-scale path-loss coefficient between j^{th} BS and the k^{th} user in the ℓ^{th} cell

- Corresponding rates

$$\mathrm{R}_{j,m}^{\mathrm{CP-ul}} = \frac{(C_u-\tau)}{C}\log_2\left(1+\mathrm{SINR}_{j,m}^{\mathrm{CP-ul}}\right)$$
$$\mathrm{R}_{j,m}^{\mathrm{CP-dl}} = \frac{C_d}{C}\log_2\left(1+\mathrm{SINR}_{j,m}^{\mathrm{CP-dl}}\right)$$

C : Coherence time of the channel. $C = C_u + C_d$

Superimposed Pilots

- Pilots of length C_u are transmitted simultaneously along with data
- If $KL \leq C_u$, each user is assigned a unique pilot
- Received signal at j^{th} BS (implicitly assumed in equations) when employing superimposed pilots

$$\mathbf{Y} = \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \mathbf{h}_{\ell,k} \left(\rho_{\ell,k}\mathbf{x}_{\ell,k} + \lambda_{\ell,k}\mathbf{p}_{\ell,k}\right)^T + \mathbf{W}$$

$\rho^2_{\ell,k}$ and $\lambda^2_{\ell,k}$: fractions of the transmit power reserved for pilot and data, respectively

$\mathbf{p}_{\ell,k}$: $C_u \times 1$ orthogonal pilot transmitted by the k^{th} user of the ℓ^{th} cell

- The total transmitted power is $\mu_{\ell,k} = \lambda^2_{\ell,k} + \rho^2_{\ell,k}$

Superimposed Pilots

- The least squares estimate of the channel

$$\begin{aligned}\widehat{\mathbf{h}}_{j,\ell,k} &= \mathbf{Y}\left(\lambda_{\ell,k}^2\,\mathbf{p}_{\ell,k}^H\mathbf{p}_{\ell,k}\right)^{-1}\lambda_{\ell,k}\mathbf{p}_{\ell,k}^* \\ &= \mathbf{h}_{\ell,k} + \frac{1}{C_u\lambda_{\ell,k}}\sum_{n=0}^{L-1}\sum_{p=0}^{K-1}\rho_{n,p}\mathbf{h}_{n,p}\mathbf{x}_{n,p}^T\mathbf{p}_{\ell,k}^* + \frac{1}{C_u\lambda_{n,p}}\mathbf{W}\mathbf{p}_{\ell,k}^*\end{aligned}$$

- The data, estimated using a matched filter

$$\begin{aligned}\widetilde{\mathbf{x}}_{j,m}^T &= \frac{1}{M\rho_{j,m}\beta_{j,m}}\widehat{\mathbf{h}}_{j,m}^H\left(\mathbf{Y} - \lambda_{j,m}\widehat{\mathbf{h}}_{j,m}\mathbf{p}_{j,m}^T\right) \\ \widehat{\mathbf{x}}_{j,m} &= \eta\left(\widetilde{\mathbf{x}}_{j,m}\right)\end{aligned}$$

$\eta(\cdot)$: element-by-element hard-decision function that replaces each element of the input vector with the constellation point that is closest in Euclidean distance to that element

Performance

- The MSE of the superimposed pilot based channel estimate

$$\mathrm{MSE}^{\mathrm{SP}}_{j,m} = \frac{1}{C_u \lambda^2_{j,m}} \left(\sum_{k=0}^{K-1} \rho^2_{j,k}\beta_{j,k} + \sum_{\substack{\ell=0 \\ \ell \neq j}}^{L-1} \sum_{k=0}^{K-1} \rho^2_{\ell,k}\beta_{\ell,k} \right) + \frac{\sigma^2}{C_u \lambda^2_{j,m}}$$

- The SINR after MF

$$\mathrm{SINR}^{\mathrm{SP-ul}}_{\substack{j,m \\ M\to\infty}} = \frac{\beta^2_{j,m}}{\frac{1}{C_u \lambda^2_{j,m}\rho^2_{j,m}} \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \rho^2_{\ell,k}\mu_{\ell,k}\beta^2_{\ell,k}}$$

- For comparison, recall the SINR of a system using conventional pilots

$$\mathrm{SINR}^{\mathrm{CP-ul}}_{\substack{j,m \\ M\to\infty}} = \frac{\beta^2_{j,m}}{\sum_{\ell \neq j} \beta^2_{\ell,m}}$$

Implications

$$\underset{M\to\infty}{\mathrm{SINR}_{j,m}^{\mathrm{SP-ul}}} = \frac{\beta_{j,m}^2}{\frac{1}{C_u \lambda_{j,m}^2 \rho_{j,m}^2} \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \rho_{\ell,k}^2 \mu_{\ell,k} \beta_{\ell,k}^2} \qquad \underset{M\to\infty}{\mathrm{SINR}_{j,m}^{\mathrm{CP-ul}}} = \frac{\beta_{j,m}^2}{\sum_{\ell\neq j} \beta_{\ell,m}^2}$$

- Power control is essential for a system with superimposed pilots
- There is a coherence time beyond which a system with superimposed pilots performs better than a system with conventional pilots, i.e., if $\beta_{j,m} = 1,\ \forall m,\ \beta_{\ell,m} = \beta,\ \forall \ell \neq j, m$, and $\rho_{j,m}^2 = \lambda_{j,m}^2, \forall\, j, m$, then

$$\underset{M\to\infty}{\mathrm{SINR}_{j,m}^{\mathrm{SP-ul}}} > \underset{M\to\infty}{\mathrm{SINR}_{j,m}^{\mathrm{CP-ul}}} \quad \text{if} \quad C_u > 2K\left(1 + \frac{1}{(L-1)\beta^2}\right)$$

- This dependence on C_u can be reduced by using an iterative data aided scheme.

Iterative Channel Estimation

- Replacing k, ℓ with a single index n such that $0 \leq n \leq N-1$, where $N \triangleq KL$, received signal at the j^{th} BS when employing superimposed pilots

$$\mathbf{Y} = \sum_{n=0}^{N-1} \mathbf{h}_n \left(\rho_n \mathbf{x}_n + \lambda_n \mathbf{p}_n\right)^T + \mathbf{W}$$

- Users are arranged in decreasing order of their path-loss coefficients, i.e., $\beta_0 > \beta_1 > \ldots > \beta_{N-1}$
- $\mathcal{U}_m^{(i)}$ is the set of users whose channel and data vectors are used in feedback to estimate channel of m^{th} user in i^{th} iteration

Iterative Channel Estimation

- Set $\widehat{\mathbf{h}}_m^{(0)} = \mathbf{0}, \ \forall \ m$ and $\widehat{\mathbf{x}}_m^{(0)} = \mathbf{0}, \ \forall \ m = 0, \ldots, N-1$
- Channel estimation step

$$\widehat{\mathbf{h}}_m^{(i)} = \frac{1}{C_u \lambda_m} \left[\mathbf{Y} - \sum_{\substack{n=0 \\ n \in \mathcal{U}_m^{(i)}}}^{m-1} \rho_n \widehat{\mathbf{h}}_n^{(i)} \left(\widehat{\mathbf{x}}_n^{(i)} \right)^T - \sum_{\substack{n=m \\ n \in \mathcal{U}_m^{(i)}}}^{N-1} \rho_n \widehat{\mathbf{h}}_n^{(i-1)} \left(\widehat{\mathbf{x}}_n^{(i-1)} \right)^T \right]$$

- Data estimation step

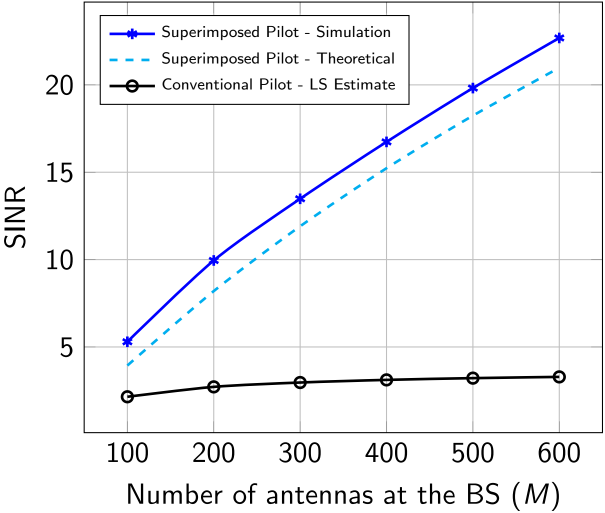
$$\left(\widetilde{\mathbf{x}}_m^{(i)} \right)^T = \frac{1}{M \rho_m \beta_m} \left(\widehat{\mathbf{h}}_m^{(i)} \right)^H \left(\mathbf{Y} - \lambda_m \widehat{\mathbf{h}}_m^{(i)} \mathbf{p}_m^T \right)$$

$$\widehat{\mathbf{x}}_m^{(i)} = \eta \left(\widetilde{\mathbf{x}}_m^{(i)} \right)$$

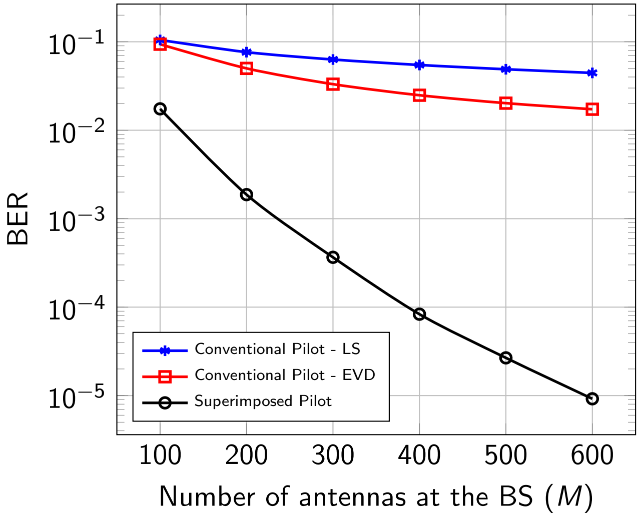
Simulation Setup

- $M = 300$ antennas
- $C_u = 100$ symbols
- 6 interfering cells ($L = 6$)
- $K = 5$ users per cell
- Cell Radius - 1km
- Users are equally spaced on a circle of radius 800m
- Path-loss exponent - 3

SINR Performance



BER Performance



Performance Benefits

- Superior uplink SINR performance in scenarios with high pilot contamination.
- As $M \to \infty$, optimal values of $\rho_{j,m} \to 0$ and $\lambda_{j,m} \to 1$, resulting in an interference-free downlink among the KL users
- Can augment systems with conventional pilots to increase the spectral efficiency

Caveats of Using Superimposed Pilots

- Number of users in all cells combined have to be less than uplink duration, i.e., $KL \leq C_u$
- Is superior to conventional pilots in uplink performance, in scenarios with high pilot contamination
- Uplink performance depends on $\frac{KL}{C}$
- $M \gg LK$ instead of $M \gg K$
- Power control is mandatory
- Coordination between cells is required for assigning pilots and for learning the path-loss coefficients $\beta_{j,\ell,k}$ of all users

References I

[1] T. L. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” in IEEE Trans. on Wireless Comm., vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin and R. Zhang, “An Overview of Massive MIMO: Benefits and Challenges,” in IEEE J. of Sel. Topics in Signal Proc., vol. 8, no. 5, pp. 742-758, Oct. 2014.

[3] R. R. Müller, L. Cottatellucci and M. Vehkaperä, “Blind Pilot Decontamination,” in IEEE J. of Sel. Topics in Signal Proc., vol. 8, no. 5, pp. 773-786, Oct. 2014.

[4] E. Björnson, E. G. Larsson and M. Debbah, “Massive MIMO for Maximal Spectral Efficiency: How Many Users and Pilots Should Be Allocated?,” in IEEE Trans. on Wireless Comm., vol. 15, no. 2, pp. 1293-1308, Feb. 2016.

References II

[5] K. Upadhya, S. Vorobyov, M. Vehkapera, "Superimposed Pilots are Superior for Mitigating Pilot Contamination in Massive MIMO - Part I: Theory and Channel Estimation", in IEEE. Trans. on Signal Proc., Submitted, Nov. 2015, http://arxiv.org/abs/1603.00648

[6] K. Upadhya, S. Vorobyov, M. Vehkapera, "Superimposed Pilots are Superior for Mitigating Pilot Contamination in Massive MIMO - Part II: Performance and Optimization", in IEEE. Trans. on Signal Proc., Submitted, Mar. 2016,

Thank You

Backup Slides

Choice of $\mathcal{U}_m^{(i)}$ - Assumptions

- $\mathbf{e}_m^{(i)} \triangleq \mathbf{x}_m - \widetilde{\mathbf{x}}_m^{(i)}$ is the error in the matched filtered data.
- $\Delta\mathbf{x}_m^{(i)} \triangleq \mathbf{x}_m - \widehat{\mathbf{x}}_m^{(i)}$ is the error vector after slicing.
- $\Delta\mathbf{h}_m^{(i)} \triangleq \mathbf{h}_m - \widehat{\mathbf{h}}_m^{(i)}$ is the error in the channel estimate.
- Assuming that the elements of $\mathbf{e}_m^{(i)}$ are i.i.d. circular complex-Gaussian random variables with zero mean and variance $\mathcal{I}_m^{(i)}$.
- In deriving $\mathcal{I}_m^{(i)}$, the following simplifications have been made to obtain a closed form expression:
  1. $\mathbf{e}_m^{(i)}$ is independent of $\mathbf{x}_k$ and $\mathbf{W}$, $\forall\ k, i$.
  2. $\Delta\mathbf{x}_m^{(i)}$ is independent of $\mathbf{x}_k$, $\mathbf{W}$, and $\mathbf{h}_k$, $\forall\ k, i$.
  3. $\Delta\mathbf{x}_m^{(i)}$ is independent of $\Delta\mathbf{x}_k^{(p)}$, $\forall p \neq i, m \neq k$ and the elements of $\Delta\mathbf{x}_m^{(i)}$ are i.i.d.
  4. $\Delta\mathbf{h}_m^{(i)}$ is independent of $\mathbf{x}_k$, $\mathbf{W}$, and $\Delta\mathbf{x}_k^{(p)}$, $\forall\ k, p$.

Choice of $\mathcal{U}_m^{(i)}$

- The user set $\mathcal{U}_m^{(i)}$ is selected as

$$\mathcal{U}_m^{(i)} = \arg \min_{\mathcal{U} \in \mathcal{P}(\mathcal{S})} \left\{ \mathcal{I}_m^{(i)} (\mathcal{U}) \right\}$$

$\mathcal{I}_m^{(i)}(\mathcal{U})$ is the interference power

$$\mathcal{I}_m^{(i)} (\mathcal{U}) \approx \frac{1}{\beta_m^2} \left(\frac{1}{M\rho_m^2} \sum_{\substack{n=0 \\ n \neq m}}^{N-1} \beta_n \beta_m + \frac{\sigma^2 \beta_m}{M\rho_m^2} + \frac{1}{M^2 \rho_m^2} \psi_m^{(i)}(\mathcal{U}) \right)$$

Choice of $\mathcal{U}_m^{(i)}$

- $\psi_m^{(0)} = 0$. For $i \geq 1$

$$\psi_m^{(i)} = \frac{M^2}{C_u \lambda_m^2} \left[\sum_{\substack{n=0 \\ n \notin \mathcal{U}_m^{(i)}}}^{N-1} \rho_n^2 \beta_n^2 + \sum_{\substack{n=0 \\ n \in \mathcal{U}_m^{(i)}}}^{m-1} \rho_n^2 \beta_n^2 \alpha_n^{(i)} + \sum_{\substack{n=m \\ n \in \mathcal{U}_m^{(i)}}}^{N-1} \rho_n^2 \beta_n^2 \alpha_n^{(i-1)} + o\left(\frac{N}{M}\right) \right]$$

$\alpha_k^{(i)} \triangleq \mathbb{E}\left\{ \left| \left[\Delta \mathbf{x}_k^{(i)} \right]_n \right|^2 \right\}$ is the variance of $\Delta \mathbf{x}_m^{(i)} \triangleq \mathbf{x}_m - \widehat{\mathbf{x}}_m^{(i)}$

- For a P-QAM constellation

$$\alpha_k^{(i)} = \begin{cases} \frac{24}{\sqrt{P}(\sqrt{P}+1)} Q\left(\sqrt{\frac{\frac{3}{(P-1)}}{\mathcal{I}_k^{(i)}}} \right), & i \geq 1 \\ 1, & i = 0 \end{cases}$$

Choice of $\mathcal{U}_m^{(i)}$

- The optimal set is given as

$$\mathcal{U}_m^{(i)} = \Big\{ k \in \mathbb{N} \,\Big|\, \alpha_k^{(i)} < \gamma_k^{(i)} \text{ when } k < m \\ \text{and } \alpha_k^{(i-1)} < \gamma_k^{(i-1)} \text{ when } k \geq m \Big\}$$

$$\gamma_k^{(i)} \triangleq \frac{\left\{ \beta_k^2 + \frac{1}{M} \sum_{n=0}^{N-1} \beta_k \beta_n - \frac{\psi_k^{(i)} \big|_{k \in \mathcal{U}_m^{(i)}}}{M^2} \right\}}{\left\{ \beta_k^2 + \frac{1}{M} \sum_{n=0}^{N-1} \beta_k \beta_n + \frac{\psi_k^{(i)} \big|_{k \in \mathcal{U}_m^{(i)}}}{M^2} \right\}}$$

- Alternatively, a fixed conservative threshold can be chosen as

$$\mathcal{U}_{\text{fixed}} = \Big\{ m \in \mathbb{N} \,\Big|\, \mathcal{I}_m^{(2)}(\{m\}) < \mathcal{I}_m^{(2)}(\varnothing) = \mathcal{I}_m^{(1)}(\varnothing) \Big\}$$