Improved channel estimation for massive MIMO systems using hybrid pilots with pilot anchoring

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

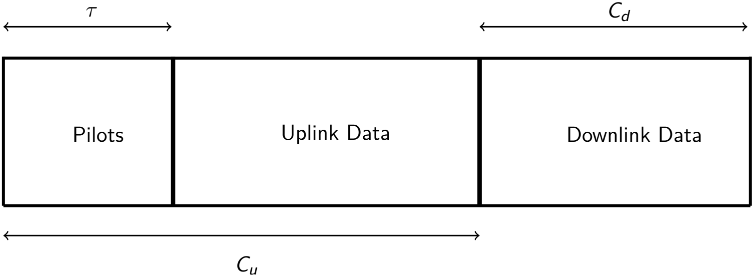
WSA 2017
16 Mar 2017
Berlin

Outline

- Channel estimation and pilot contamination in massive MIMO
- Review of superimposed pilots for massive MIMO
- Improving superimposed pilots with pilot anchoring
- Simulation Results

Time-multiplexed Pilots

- Uplink time slot is partitioned for pilots and data



- Orthogonal pilots are used to estimate the channel of K users
- The same pilot sequences are shared between L neighboring cells

Time-multiplexed Pilots: Uplink

- Received signal during training at desired BS ($\ell = 0$)

$$\mathbf{Y} = \sum_{\ell=0}^{L-1} \mathbf{H}_\ell \boldsymbol{\Phi}^T + \mathbf{W} \in \mathbb{C}^{M\times\tau}$$

$\mathbf{H}_\ell$: $M \times K$ channel matrix between desired BS and users in cell ℓ
$\boldsymbol{\Phi}$: $\tau \times K$ orthogonal pilot sequences ($\boldsymbol{\Phi}^T\boldsymbol{\Phi}^* = \tau \cdot \mathbf{I}$)

- The least squares estimate of the k'th user at desired cell

$$\widehat{\mathbf{h}}_{0,k}^{\mathrm{TP}} \triangleq \frac{1}{\tau}\mathbf{Y}\boldsymbol{\phi}_k^* = \mathbf{h}_{0,k} + \sum_{\ell=1}^{L-1} \mathbf{h}_{\ell,k} + \mathbf{n}_{0,k}$$

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

Superimposed Pilots (1/2)

- Pilots of length C_u are transmitted along the data
- Ideally $C_u \geq KL$ orthogonal pilots used; no overhead in time-domain, but causes interference between data and pilots
- Received signal at BS j when employing superimposed pilots

$$\mathbf{Y} = \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \mathbf{h}_{\ell,k} \left(\rho \mathbf{x}_{\ell,k} + \lambda \mathbf{p}_{\ell,k}\right)^T + \mathbf{W} \in \mathbb{C}^{M \times C_u}$$

ρ^2 and λ^2 : fractions of the transmit power reserved for data and pilots, respectively, such that $\rho^2 + \lambda^2 = 1$

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

$\mathbf{x}_{\ell,k}$: $C_u \times 1$ data vector transmitted by user k of cell ℓ

Superimposed Pilots (2/2)

- The least squares estimate of the channel when $C_u \geq KL$

$$\begin{aligned}\widehat{\mathbf{h}}_{0,k} &= \frac{\mathbf{Y}\mathbf{p}_{0,k}^{*}}{\lambda C_u} \\ &= \mathbf{h}_{0,k} + \frac{\rho}{\lambda C_u}\sum_{\ell=0}^{L-1}\sum_{m=0}^{K-1}\mathbf{h}_{\ell,m}\mathbf{x}_{\ell,m}^{T}\mathbf{p}_{0,k}^{*} + \mathbf{n}_{0,k}\end{aligned}$$

- Uplink data, estimated using a matched filter (MF)

$$\widehat{\mathbf{x}}_{0,k}^{T} = \frac{1}{\rho M \beta_{0,k}}\widehat{\mathbf{h}}_{0,k}^{H}\left(\mathbf{Y} - \lambda\widehat{\mathbf{h}}_{0,k}\mathbf{p}_{0,k}^{T}\right)$$

and the usual MF precoding with channel estimate in downlink

Performance Overview of Superimposed Pilots

- Values of ρ and λ that maximize SINR lower bound in UL

$$\text{Data power: } \rho^2_{\text{opt}} \approx \frac{1}{1+\sqrt{\frac{M+LK}{C_u}}}$$

$$\text{Pilot power: } \lambda^2_{\text{opt}} = 1 - \rho^2_{\text{opt}} \approx \frac{1}{1+\sqrt{\frac{C_u}{M+LK}}}$$

- The resulting DL SINR reads

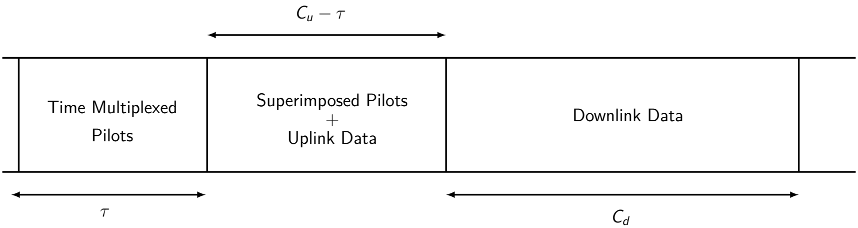
$$\left.\mathrm{SINR}_{0,k}^{\mathrm{SP-dl}}\right|_{\rho_{\text{opt}},\lambda_{\text{opt}}} \approx \frac{\sqrt{C_u\left(M+LK\right)}\beta_{0,k}^2}{\sum_{m=0}^{K-1}\sum_{\ell=0}^{L-1}\beta_{\ell,m}^2}$$

- Increases as $\sqrt{M}$ if the assumption $LK \leq C_u$ holds

Comments on Performance of Superimposed Pilots

- The UL performance of time-multiplexed pilots is limited by inter-cell interference.
- On the other hand, the UL performance of superimposed pilots is limited to a large extent by intra-cell interference.
- The UL performance of these pilot transmissions are complementary
- However, downlink performance of superimposed pilots is significantly better than time-multiplexed pilots
- Therefore, the objective of this work is to use superimposed pilots to improve the resilience of the channel estimate obtained from time-multiplexed pilots to pilot contamination

Proposed Pilot Structure



- The received signal in UL at BS $j = 0$ has two parts

$$\mathbf{Y} = [\mathbf{Y}_p, \mathbf{Y}_s]$$

where

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

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

Objective

- Reduce the inter-cell component of UL interference in superimposed pilots using time-multiplexed pilots
- This is accomplished by adding a shaping constraint on the LS objective function
- The pilot sequences are used as known data to implement this shaping constraint

Proposed Method (1/4)

- The optimization problem can be written as

$$\widehat{\mathbf{h}}_{0,k} = \arg\min_{\mathbf{h}} \|\mathbf{Y}_s - \mathbf{h}\lambda\mathbf{p}_{0,k}^T\|_F^2$$

subject to

$$\frac{1}{M}\mathbf{h}^H\mathbf{Y}_p = \phi_{0,k}^T$$

where

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

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

Proposed Method (2/4)

- This optimization problem has an equivalent form

$$\widehat{\mathbf{h}}_{0,k} = \arg\min_{\mathbf{h}} \|\mathbf{h} - \widehat{\mathbf{h}}^{\mathrm{SP}}_{0,k}\|_2^2$$

subject to

$$\frac{1}{M}\mathbf{h}^H \widehat{\mathbf{H}}^{\mathrm{TP}}_0 = \mathbf{e}_k^T$$

where $\mathbf{e}_k$ is the k'th column of the identity matrix.

$$\widehat{\mathbf{h}}^{\mathrm{SP}}_{0,k} \longleftarrow \mathbf{Y}_s = \sum_{\ell=0}^{L-1}\sum_{k=0}^{K-1} \mathbf{h}_{\ell,k}\left(\rho\mathbf{x}_{\ell,k} + \lambda\mathbf{p}_{\ell,k}\right)^T + \mathbf{W}_s$$

$$\widehat{\mathbf{H}}^{\mathrm{TP}}_0 \longleftarrow \mathbf{Y}_p = \sum_{\ell=0}^{L-1} \mathbf{H}_\ell \mathbf{\Phi}^T + \mathbf{W}_p$$

Proposed Method (3/4)

- Setting $\mathbf{x} = \mathbf{h} - \widehat{\mathbf{h}}_{0,k}^{\mathrm{SP}}$, the optimization problem can be simplified as

$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{x}\|_2^2 \\ &\text{subject to} \\ &\left(\widehat{\mathbf{H}}_{0,k}^{\mathrm{TP}}\right)^H \mathbf{x} = M\mathbf{e}_k - \left(\widehat{\mathbf{H}}_{0,k}^{\mathrm{TP}}\right)^H \widehat{\mathbf{h}}_{0,k}^{\mathrm{SP}} \end{aligned}$$

- This problem has the form of the generalized sidelobe canceller [†]

$$\min_{\mathbf{x}} \mathbf{x}^H \mathbf{R} \mathbf{x} \ \text{ subject to } \ \mathbf{C}^H \mathbf{x} = \mathbf{d}$$

which has the optimal solution $\mathbf{x}^\star = \mathbf{R}^{-1}\mathbf{C}\left(\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\right)\mathbf{d}$, where $\mathbf{R}$ is positive definite

---

[†] Sergiy A. Vorobyov, “Principles of minimum variance robust adaptive beamforming design”, Signal Processing, Volume 93, Issue 12, December 2013, Pages 3264-3277, ISSN 0165-1684

Proposed Method (4/4)

- Since $\mathbf{x} = \mathbf{h} - \widehat{\mathbf{h}}_{0,k}^{\mathrm{SP}}$, the output of the "channel combiner" is

$$\widehat{\mathbf{h}}_{0,k} = \widehat{\mathbf{H}}_{0,k}^{\mathrm{TP}} \left(\left(\widehat{\mathbf{H}}_{0,k}^{\mathrm{TP}} \right)^{H} \widehat{\mathbf{H}}_{0,k}^{\mathrm{TP}} \right)^{-1} \left(M \mathbf{e}_m - \left(\widehat{\mathbf{H}}_{0,k}^{\mathrm{TP}} \right)^{H} \widehat{\mathbf{h}}_{0,k}^{\mathrm{SP}} \right) + \widehat{\mathbf{h}}_{0,k}^{\mathrm{SP}}$$

- The UL data is then estimated as

$$\widehat{\mathbf{x}}_{0,k}^{T} = \frac{1}{M\rho} \widehat{\mathbf{h}}_{0,k}^{H} \mathbf{Y}_s - \lambda \mathbf{p}_{0,k}^{T}$$

Simulation Setup

- Number of cells in system (L) : 7
- Number of users per cell (K) : 5
- Uplink duration (C_u) : $40 > 7 \cdot 5$ symbols
- Cell Radius : 1km
- Path-loss exponent : 3
- No shadowing
- Constellation : QPSK
- Two channel scenarios are considered
  - *Scenario 1* : Users are uniformly distributed across the cell.
  - *Scenario 2* : Users in all cells are equally spaced on a circle.

Uplink BER Performance - Scenario 1 (Uniform)

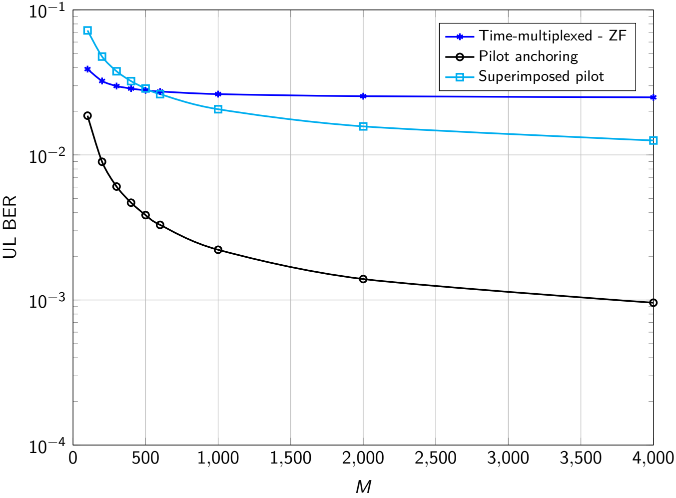


Figure: Uplink BER of users in reference cell vs M

Uplink SINR - Scenario 2 (Circle)

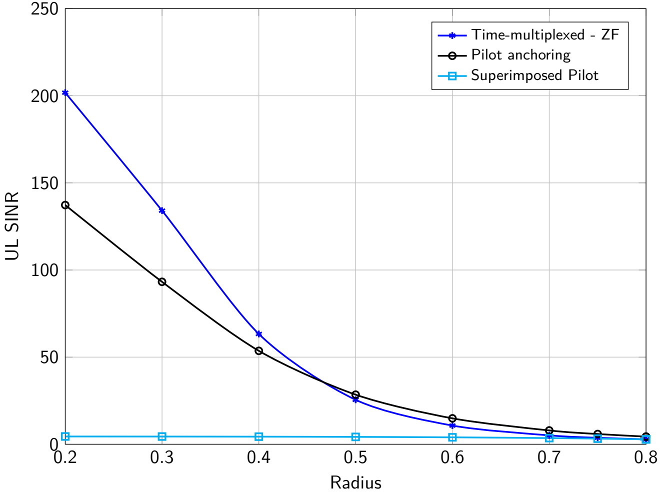


Figure: Uplink SINR of an arbitrary user vs user radius. $M = 300$.

Downlink BER - Scenario 1 (Uniform)

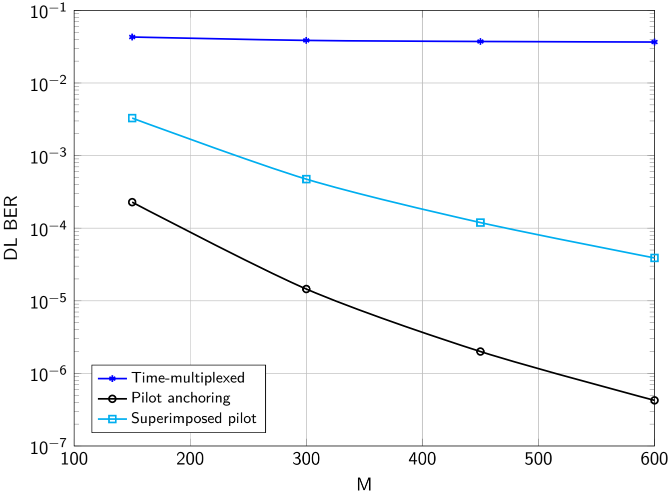


Figure: Downlink BER vs M

Downlink SINR - Scenario 2 (Circle)

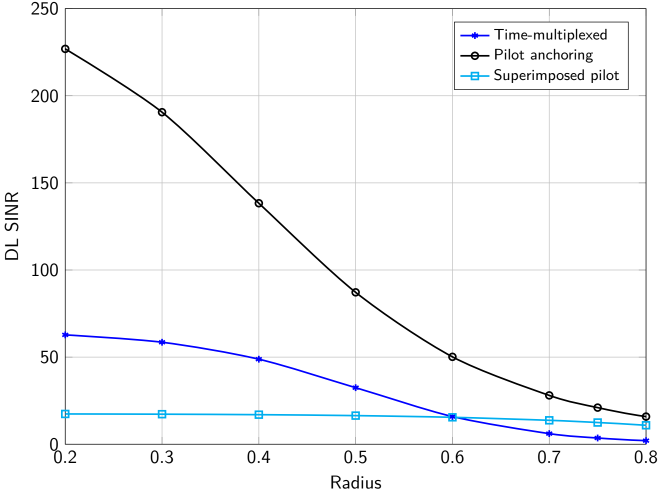


Figure: Downlink SINR of an arbitrary user vs user radius. $M = 300$.

Summary

- Superimposed pilots can augment a system with time-multiplexed pilots to improve receiver performance
- Instead of anchoring with known data/pilots, the proposed method can be easily modified to anchor with estimated data, and the equality constraint can be replaced with a norm-constraint
- The above can be implemented via an iterative method that tries to cancel both intra- and inter-cell interference

Thank You.

Any Questions?

Backup Slides

Alternative Norm Constraint for Estimated Data

- The optimization problem with the norm constraint can be written as

$$\widehat{\mathbf{h}}_{0,k} = \arg\min_{\mathbf{h}} \|\mathbf{Y}_s - \lambda \mathbf{h}\mathbf{p}_{0,k}^T\|_F^2$$

subject to

$$\left\| \frac{1}{M}\mathbf{h}^H \mathbf{Y}_d - \widehat{\mathbf{c}}_{0,k}^T \right\|^2 < \epsilon$$

  where ϵ is a design parameter, $\widehat{\mathbf{c}}_{0,k}$ are the N decoded symbols of user m in cell j, and $\mathbf{Y}_d$ is the corresponding matrix of received symbols.
- The above optimization problem is a quadratically constrained quadratic program (QCQP)

Performance Overview of Superimposed Pilots

- The asymptotic uplink SINR after MF

$$\mathrm{SINR}_{0,k}^{\mathrm{SP-ul}} = \frac{\beta_{0,k}^2}{\frac{1}{\lambda^2 C_u} \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \beta_{\ell,k}^2}$$

- The downlink SINR for large M when using MF precoding

$$\mathrm{SINR}_{0,k}^{\mathrm{SP-dl}} = \frac{\beta_{0,k}^2}{\frac{\rho^2}{\lambda^2 C_u} \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \beta_{\ell,k}^2}$$

Uplink BER Performance - Scenario 2

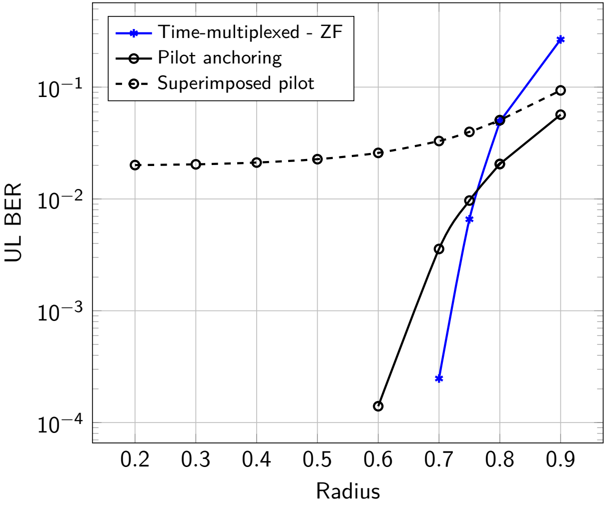


Figure: Uplink BER of users in reference cell vs user radius with $M = 300$

Downlink BER - Scenario 2

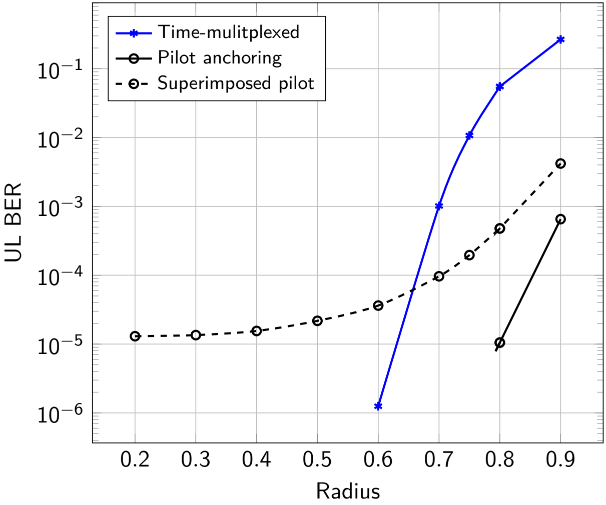


Figure: Downlink BER vs user radius