Coded Reference Signals for Massive MIMO Channel Estimation in FDD Systems

Rikke Apelfröjd[1,3], Wolfgang Zirwas[2], and Mikael Sternad[1]

[1]Uppsala University
[2]Nokia Bell Labs, Munich
[3]Now at Ericsson Research

April 2019

Mainly based on R. Apelfröjd, W. Zirwas and M. Sternad, "Low-overhead cyclic reference signals for channel estimation in FDD massive MIMO," *IEEE Trans. on Communications*, vol. 67, no. 5, May 2019.

Outline

In 5G downlinks, we need channel state information (CSI) at the transmitter for link adaptation, for scheduling, and for future downlink adaptive beamforming and coordinated transmission.

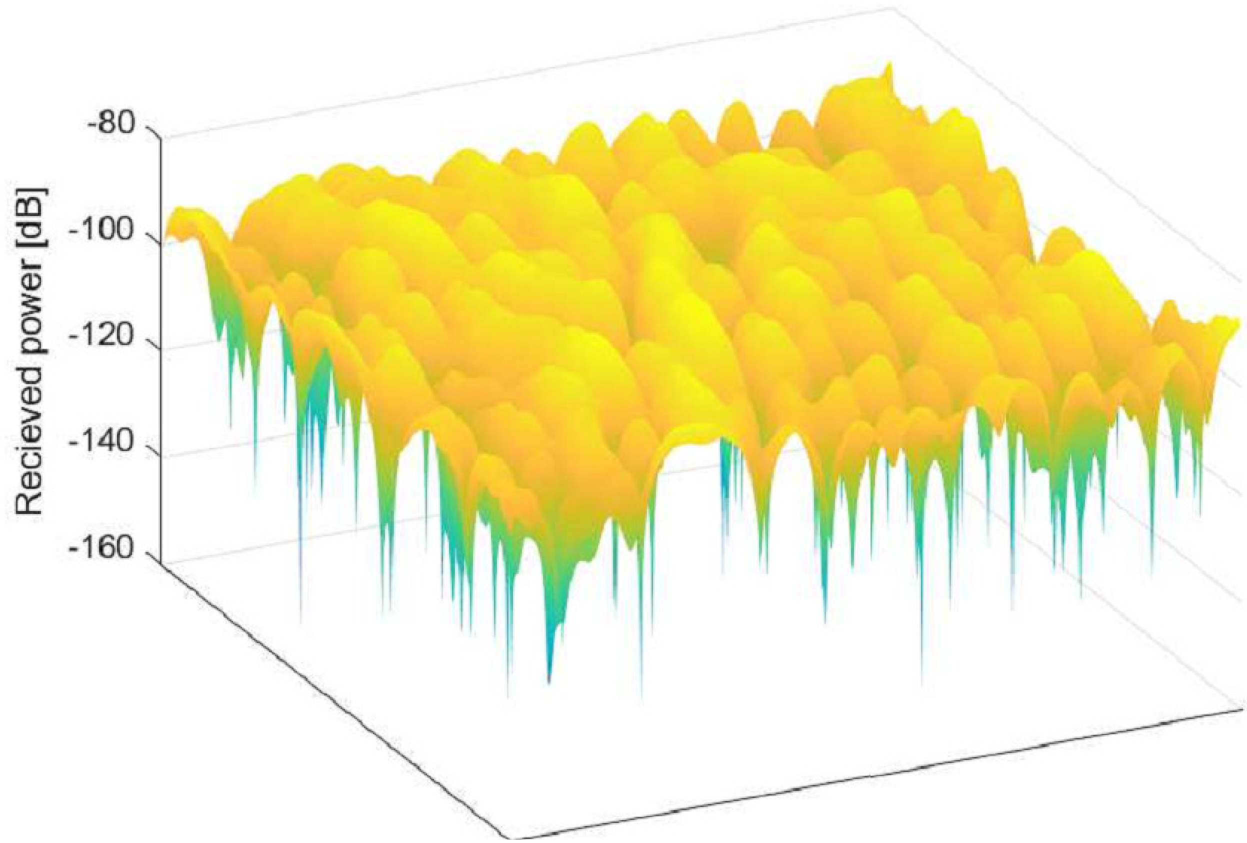
In frequency division duplex (FDD) downlinks, this requires ***many complex-valued channel gains to be estimated*** based on known downlink reference signals, ***with acceptable training (and feedback) overhead.***

This is one of the main remaining open problems in the research on massive Multiple Input Multiple Output (MIMO) systems.

We here present a solution.
For each user equipment (UE), the most important of hundreds of channels can be estimated, with low ref. signal overhead (4%-10%).

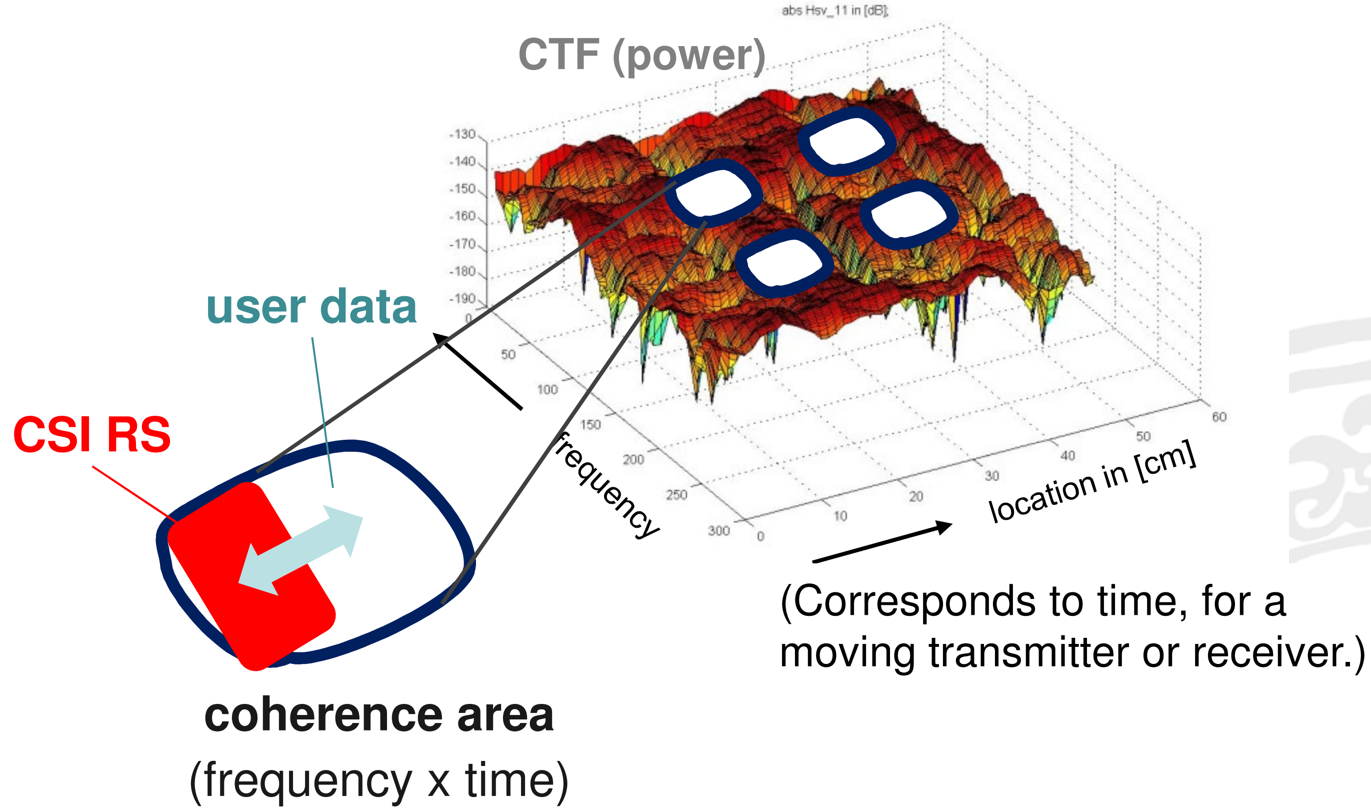
$\Rightarrow$ Massive MIMO and multi-cell cooperation is enabled for FDD.

Received RFpower in the horizontal plane, from one transmitter



Due to multipath propagation, the received power varies over space, on a distance scale of λ/2 between fading dips.

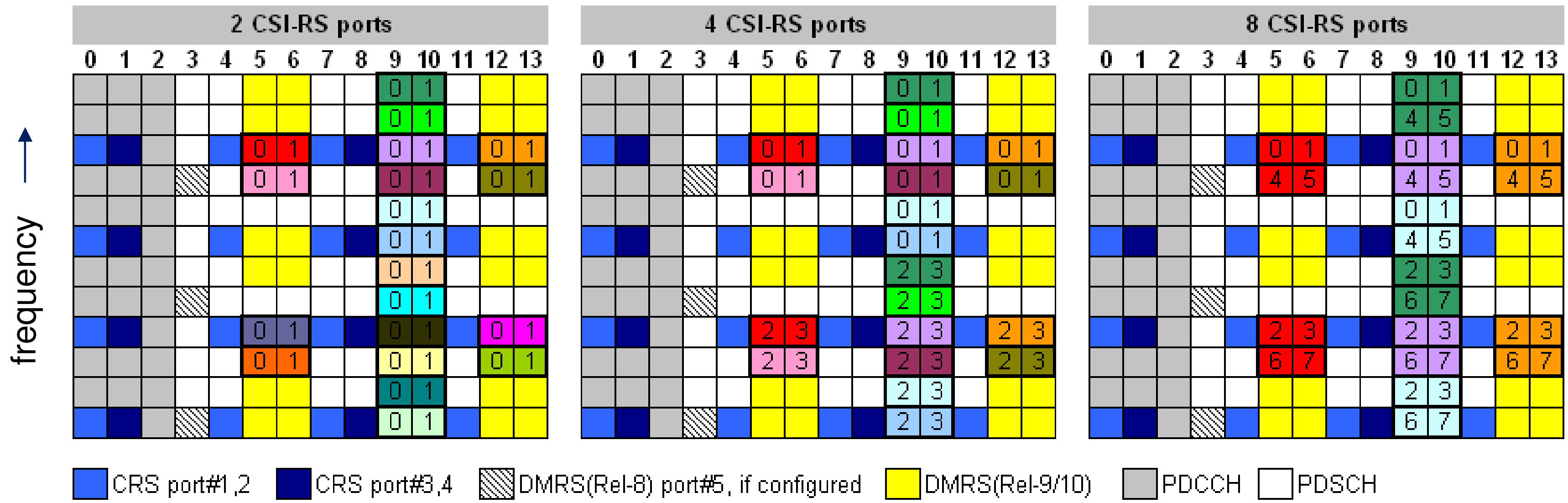
The channel is (almost) constant within a “coherence area”



To estimate the downlink channel, we must probe (estimate, predict) the complex radio channel gain within each coherence area.

Known ***reference signals (RS),*** also called ***pilot symbols***, are transmitted for this purpose. - But the fraction of RS should not be too large!

CSI Reference Signals in 3GPP LTE (4G) OFDM downlinks



Resource blocks of 12 subcarriers (**180 kHz**) x 14 OFDM symbols (**1 ms**).

CSI-RS: The numbered 40 symbols in each 180 kHz band, repeated at most every 5th ms, to limit the overhead. We will use such a resource efficiently.

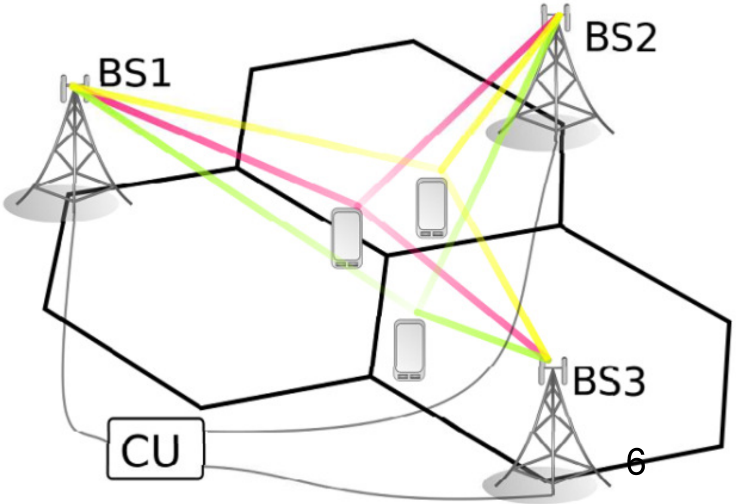
Channel estimation of a scalar complex gain ***h***, based on a received signal ***y*** (in a square in figure above) and a known transmitted reference symbol ***d:***

y = hd + interference from other transmitters + noise (+ ISI + ICI)

Many channels will need to be estimated simultaneously

- In 5G systems, **massive MIMO antennas** are being introduced, having N_{tx} = 32 -1024 antenna elements, with $N_B \leq N_{tx}$ "antenna ports".
- Radio channels from each of the N_B antenna ports will differ, in general.

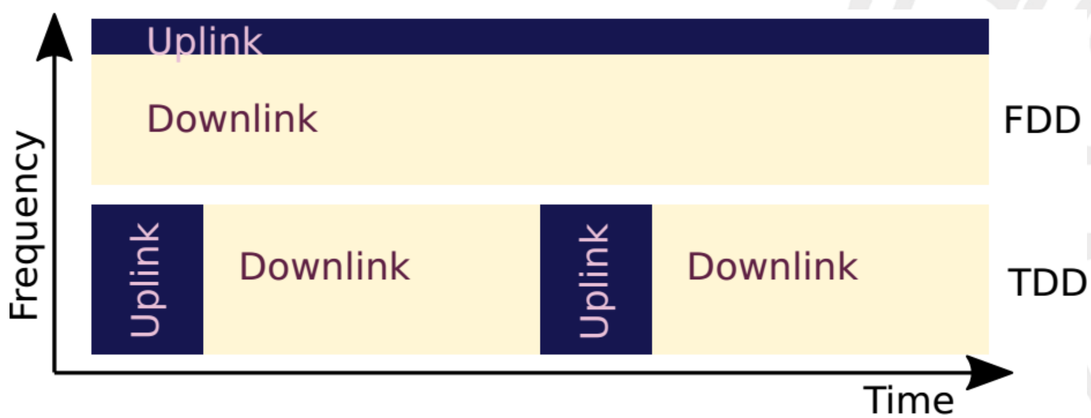
- **Coordinated transmission** from N_{BS} base stations increases performance.
- But to support it, we would have to estimate $N_{CC} = N_B N_{BS}$ channels for each UE antenna…



Channel estimations for Massive MIMO in FDD and TDD systems

- **FDD (Frequency Division Duplex)** systems use different frequency bands for downlinks (network to user) and uplinks.
- Therefore, the **downlink** channels have to be estimated based on known **downlink** reference signals in each coherence area, + uplink feedback.

- **TDD (Time Division Duplex)** systems use the same band for both uplink and downlink.
- Estimates of **uplink** channels can then be used as **downlink** estimates.

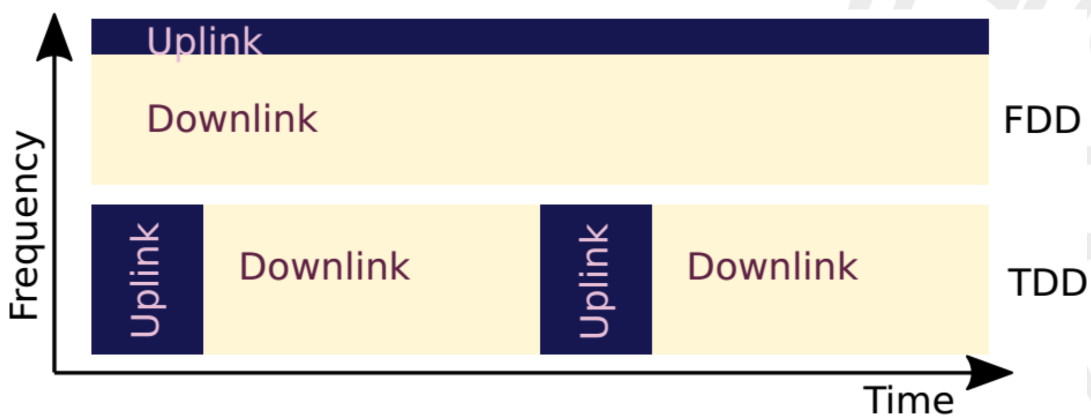


Channel estimations for Massive MIMO in FDD and TDD systems

- In **TDD**, uplink channel estimates can approximate downlink channels, based on **channel reciprocity** and calibration.
  - Then, only the uplink reference signals would be used.

  Estimating uplink channels from e.g. 10 users seems much simpler than directly estimating all downlinks from several hundred antennas.

- But, as most currently used spectrum is paired, and is used for FDD
  – we would like to be able to increase throughput also for these cases!

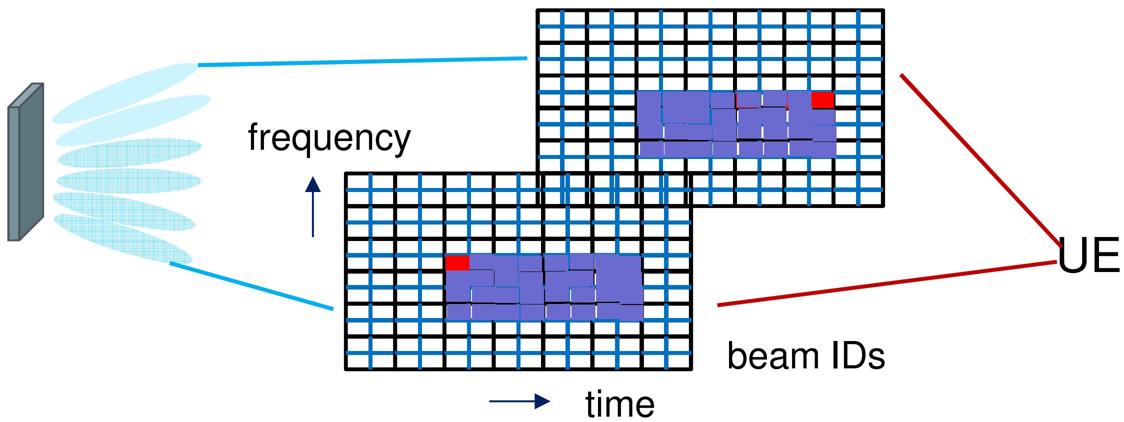


Channel estimations for Massive MIMO in FDD and TDD systems

- **Orthogonal** reference signals (RS) provide the best channel state information (CSI). E.g. where one antenna sends, all others are quiet.

  - ***but*** this requires N_{CC} resource units to estimate N_{CC} channels.

=> Large RS overhead in massive MIMO FDD downlinks.



Due to this, massive MIMO is often seen as something that can be used in TDD systems only.

Different approaches for FDD

- Just allow the large overhead and optimize the number of transmit antennas
- Well, we can probably do better
- Design *non-orthogonal* (superposed) reference signals and also utilize correlations to improve estimates
- Yes, how should these be designed though?
- User-specific *optimized* patterns of reference signals (cognitive sensing)
- Sure, but we would like to include many users – capacity is logarithmic.
- Iteratively design reference signals specific for a *group* of scheduled users.
- What about bursty traffic. Then what do we do? (Also, iterative ref. signal design will cause large delays.)
- Optimize reference signals over all potential users
- There is a large risk that we end up with a solution that is just equally bad for all potential groups or good for a few potential groups and bad for all others

Aims

- Limit the downlink ref. signal overhead and decouple it from the (large) total number of downlink channels from the transmit antennas.

- Instead, the RS overhead should ***scale with the number of channels that each user would actually need to estimate***.

  (Need to assure that this number of channels is not excessively large.)

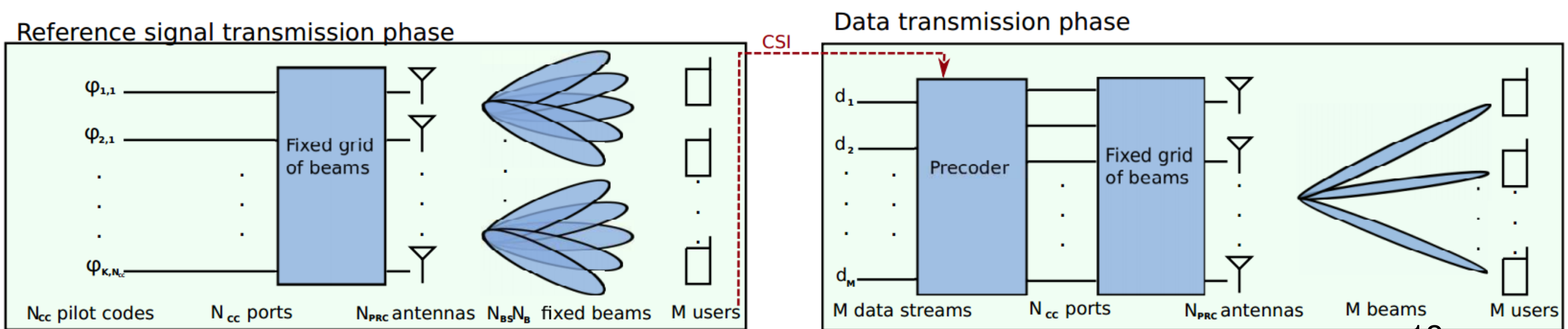
- Estimate the relevant (strongest, most important) channels to each user **accurately**, and with **short delay**, for many users.

- **User fairness**: The design should allow any user within the coverage area to gain good estimates of its relevant downlink channels.

  We will use **pre-determined non-orthogonal superposed (overlapping)** downlink reference signal vectors.
  (Optimization for particular users/channels is avoided.)

Step 1: Create spatial sparsity

- Interference from many other reference signals will be the critical problem. It is manageable if most channels to a user are weak.
- Signals from **different** (cooperating) base stations to a user often have **different** average received powers. Good for our purpose.
- But signals received from antennas at **one** base station have **similar** average received powers – making them hard to separate.
- We propose use of **fixed grids of beams** to break up this similarity:



Step 2: Use coded superposed reference signals

- Unique reference signal patterns are now sent over each beam.
- We use coded reference signals over *K* symbols from the *N* beams such that **any user can estimate up to *K* channels** – provided that the others are sufficiently weak.

Example, Block fading channels with *K=6*, and *N=9* beams

$$\bar{y}(\tau) = \underline{\Phi}(\tau)\underline{\bar{h}}(\tau) + \bar{v}(\tau)$$

Beam 1 (Beam 6 deactivated) Beam 9

$$\underline{\Phi}(\tau) = \begin{bmatrix} -1 & 0 & 1 & 1 & -1 & 0 & -1 & -1 & 1 \\ 0 & 0 & -1 & 0 & -1 & 0 & -1 & 0 & -1 \\ -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 & -1 & 0 & 1 & -1 & -1 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

Step 2: Use coded superposed reference signals

- Unique reference signal patterns are sent over each beam.
- We use coded reference signals over *K* symbols from the *N* beams such that **any user can estimate up to *K* channels** – provided that the others are sufficiently weak.

Example, Block fading channels with *K=6*, and *N=9* beams, 3 strong:

$$\bar{y}(\tau) = \underline{\Phi}(\tau)\underline{\bar{h}}(\tau) + \bar{v}(\tau)$$

$$\underline{\hat{h}}_{\mathrm{rel}}(\tau) = \underline{\Phi}^{\dagger}_{\mathrm{rel}}(\tau)\bar{y}(\tau)$$ Left pseudo-inverse of col. 3,5 and 9

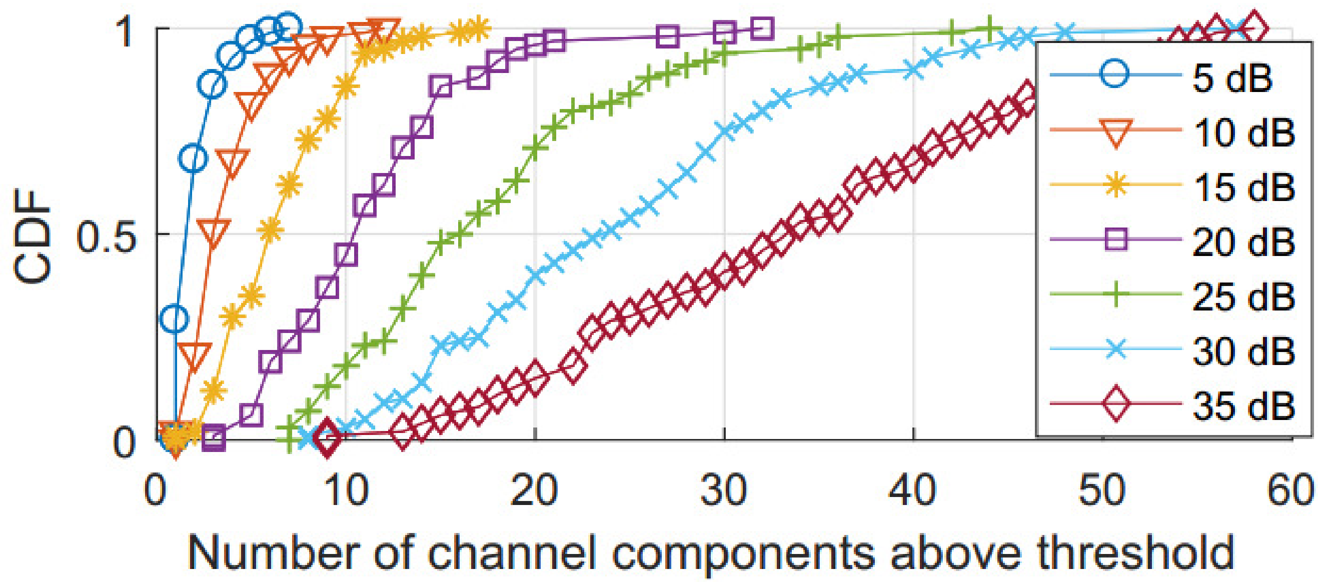
$$\underline{\Phi}(\tau) = \begin{bmatrix} -1 & 0 & 1 & 1 & -1 & 0 & -1 & -1 & 1 \\ 0 & 0 & -1 & 0 & -1 & 0 & -1 & 0 & -1 \\ -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 & -1 & 0 & 1 & -1 & -1 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

(This basic scheme was introduced in Zirwas et. al, WSA, Munich, 2016)

Reference signal structure

- **Send rarely (shadow fading timescale):**
  – Large blocks with orthogonal ref. signals from all beams, to identify the (on average) strongest beams for each user.

  (Could be substituted by estimates based on uplink sounding.)



Reference signal structure

- **Send often (fast fading timescale):**
  - Blocks with *K* coded reference signals to estimate the CSI.

”Code words” are *K*-dimensional complex vectors, with unique directions for each beam within a multi-cell cooperation area.

For *N* beams, we need *2* ≤ *K*, not *K* ≥ *N* for this !

Sets of ref. signal vectors can be constructed in many ways.

In the paper, we use constant-modulus complex numbers at time/frequency symbol *k* in beam *n* at block (time) *τ* defined by:

$$\varphi(k,n,\tau) = \exp(\theta(k,n,\tau)j),$$

with phase

$$\theta(k,n,\tau) = (k\Phi(\tau))^n \,,$$

where $\Phi(\tau)$ is a scalar parameter that defines the code.

Step 3: Use correlations in the estimates of relevant channels

The pseudoinverse-based estimate used in the example can be improved by using correlations (estimated over whole bandwidth).

Based on each RS code block of size *K*, estimate channels by

- **LMSE (Wiener) filtering** based on beam correlations and noise statistics, produces regularized estimates of up to *K* channels,
- or **Kalman filtering**, that also uses correlations over time and measurements from previous resource blocks. Can est. ≥ *K* chan.

LMSE acts as a fast start-up estimator. Kalman estimates can be used later, when autoregressive fading models become available.

Non-relevant channels are treated as noise in simplest case. (In high-complexity case, *all* >> *K* channels could be Kalman-estimated by using the temporal correlation.)

Step 4: Use time-varying reference signal codes that repeat over time

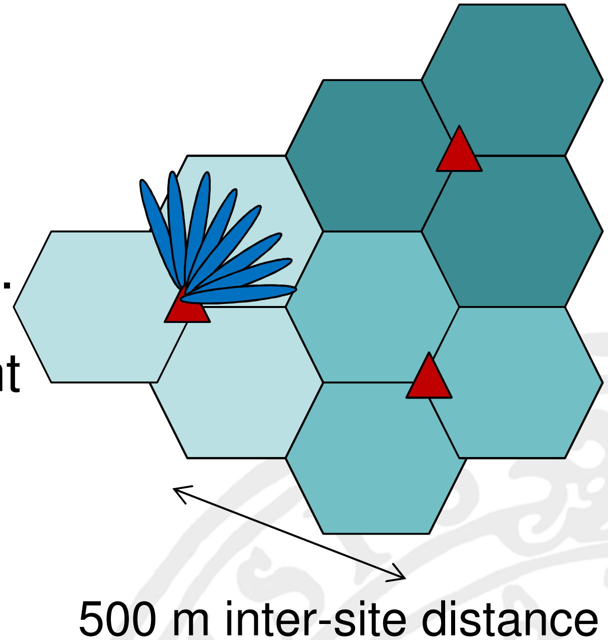
- Any fixed code structure might be bad (make the strongest channels hard to separate) in a few particular user positions.
- Different code structures are likely not bad at the same place.
- Therefore, we ***introduce diversity***: Use different codes (different scalars $\Phi(\tau)$) at different times τ, that repeat cyclically.
- The Kalman filter, which averages over time, will then produce better channel estimates.

  The worst cases become much better.

  (For cyclic reference signals, the Kalman Riccati difference equation converges to a cyclo-stationary solution.)

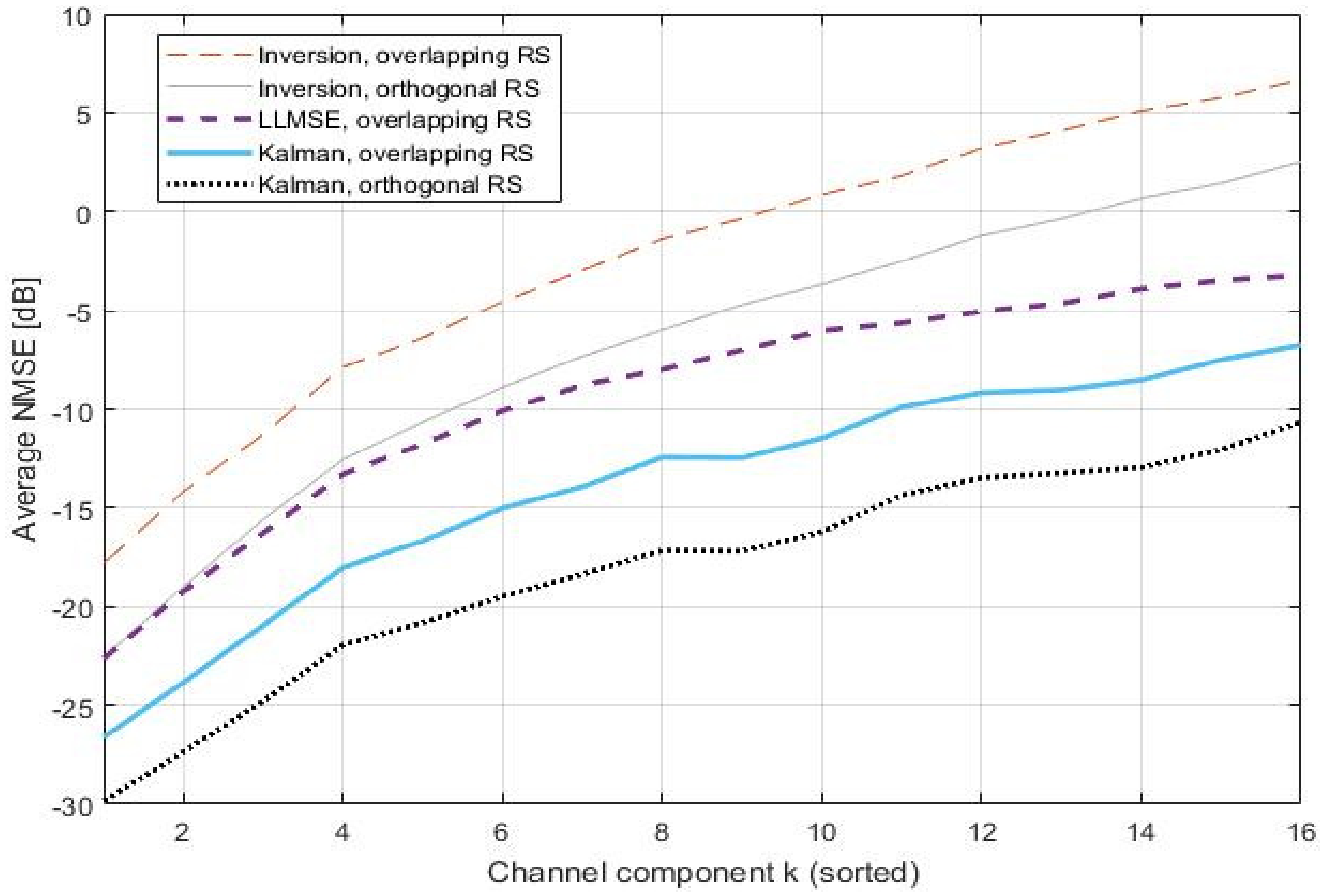
Simulations: 72 beams in 9 cells

- 3 sites x 3 sectors x 8 beams (based on 32 antennas) = 72 channels (beams).
- Quadriga channel simulator, NLOS scenario.
- *K=18* reference signal resources (6 adjacent subcarriers á 15kHz x 3 subsequent OFDM symbols). **Not** perfect flat fading: Channel correlation 0.9-0.95 within these resources.
- RS sampling time of 5 ms (4.3% overhead)
- 144 subcarriers, 2.1 MHz bandwidth used.
- 100 random user positions, with pedestrian moving users, in circle with radius 500 m centered in middle of cell cluster.



Channel estimation performance

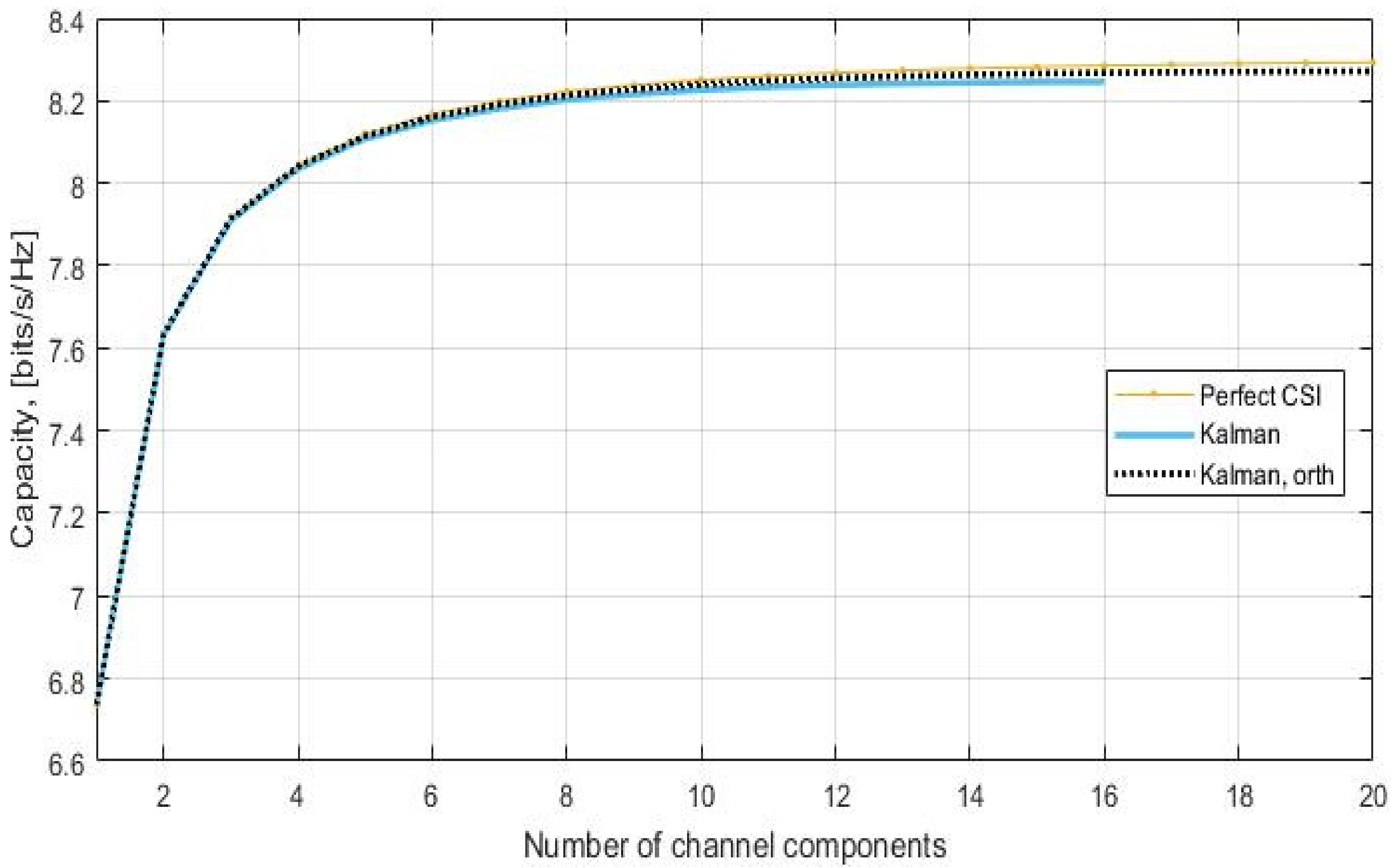
Pseudoinverse, **LMSE** and **Kalman** estimation of 16 strongest channels. Normalized mean square estimation error (NMSE) for strongest, next strongest, etc, channel, averaged over frequency and 100 user positions.



Comparisons to cases with orthogonal RS, (grey and black dotted) which would require unrealistic (100%) RS overhead.

Resulting beamforming capacity

Maximum ratio transmit beamforming to one user (combining fixed beams), as a function of how many fixed beams are combined. We here assume channel estimation accuracies = the average values from the previous slide.



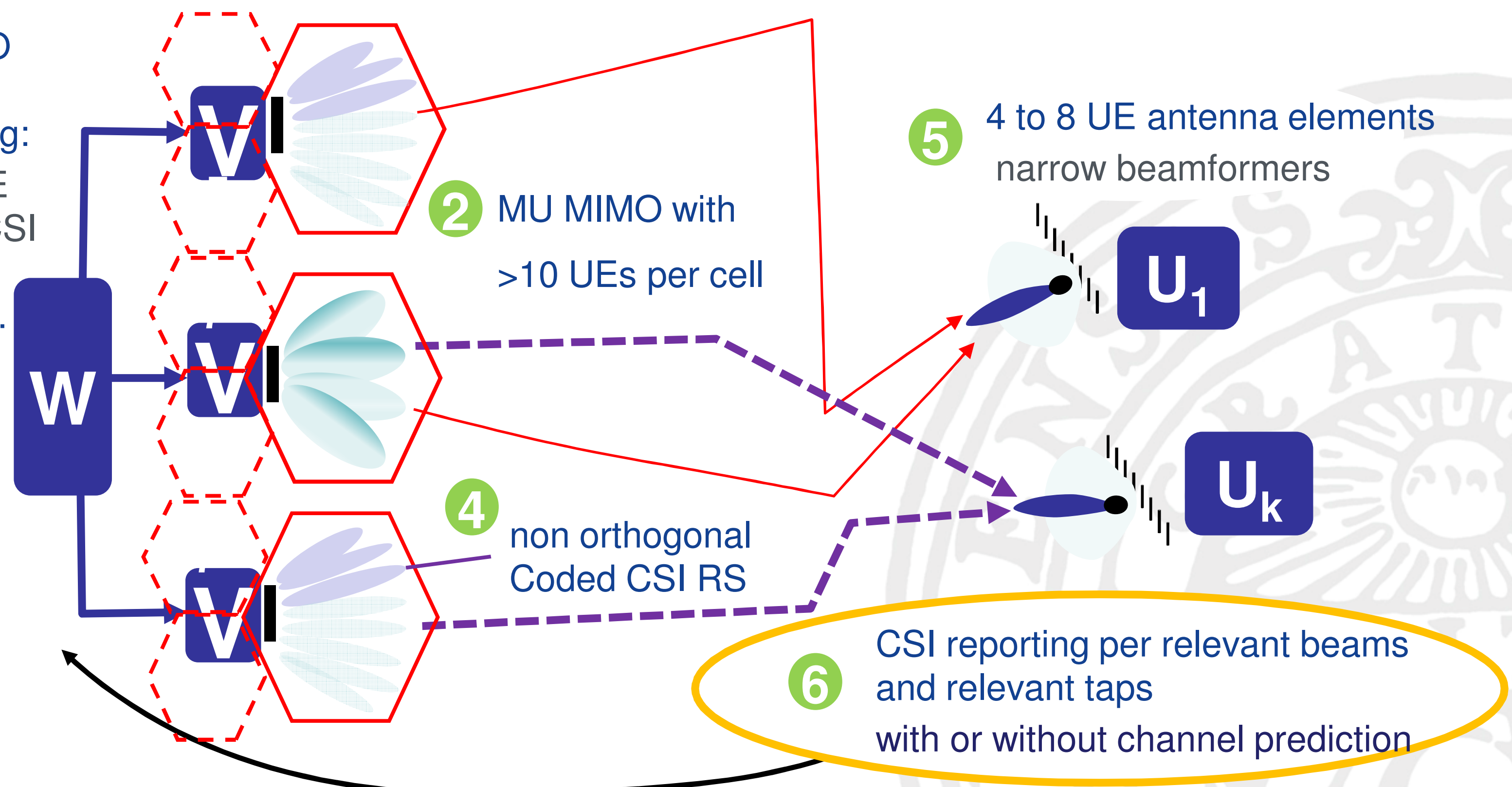
Kalman estimation with low (4.3%) RS overhead gives insignificant performance degradation as compared to the use of perfect CSI.

Flexible interference mitigation framework for 5G below 6 GHz

7-10 x spectral efficiency over LTE 4x4
5-10% overhead for CSI RS
2 to < 10 bit/subframe CSI feedback

1 ***massive MIMO***: 4 to 1028 antenna elements
fixed GoB → limited # of beams per cell

3 Network MIMO precoder with mode switching:
adaption to UE speed, load, CSI reporting parameters, …



7 3-site cooperation areas. Interference floor shaping btw cooperation areas.

System-level simulation – Zero forcing over 288 beams in 9 cooperating cells

Ideal CSI:

gross spectral eff. = **47.2** b/s/Hz/cell
avg_DR_per_UE = **4.16** b/s/Hz

Est. CSI:

gross spectral eff. = **40.2** b/s/Hz/cell
avg_DR_per_UE = **3.55** b/s/Hz

Key Solutions for massive MIMO in FDD, which synergistically produce these results:

- Massive MIMO base stations, using grid-of-beams (32 beams x 9 cells = 288 beams)
- Also, spatial UE processing over 8 UE antennas, by maximum ratio combining
- Beam deactivation per user group: Turn off under-used beams
- Use coded CSI reference signals (2x60 symb each 5 ms, Pseudo-inverse channel est.)
- Adaptive quantization of feedback => 4 bits/res. block (0.02b/s/Hz) uplink overhead
- User selection and regularized zero-forcing network MIMO precoding over all beams.

(For details, see: Zirwas, Sternad and Apelfröjd, IEEE PIMRC, Montreal, 2017.)

UPPSALA UNIVERSITET

Summary and discussion

- Using **fixed grid-of-beams** from massive MIMO antennas ensures that users have sparse channel vectors (few relevant beams).

- **Superposed reference signals** then causes a NMSE loss of ≈ 5dB.
  - Does not affect max. ratio single user beamforming performance
  - Gives some, but acceptable, loss for zero forcing precoding.

- For each user, the most important out of hundreds of channels can now be estimated, with low (4%-10%) reference signal overhead
- Also, reasonable uplink feedback overhead (see PIMRC 2017).

Could solve the problem using the present CSI-RS resources.

⇒ **Massive MIMO and multi-cell cooperation is enabled for FDD.**

Summary and discussion

TDD downlink estimation based on uplink estimates (channel reciprocity) has several challenges:

- tight calibration needed
- pilot contamination for (many) uplink sounding reference signals
- limited UE transmit power and battery lifetime
- hard to estimate interference at UEs by uplink channel estimation.
- limited support for channel prediction, requiring long term observations of the radio channel.

→ ***Explicit CSI feedback based on downlink Coded CSI RS could be considered also for TDD as add on.***

References

- V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfröjd and T. Svensson,
  The role of Small Cells, Coordinated Multipoint, and Massive MIMO in 5G.
  IEEE Communications Magazine, May 2014, pp. 44-51.
  http://www.signal.uu.se/Publications/abstracts/p1402.html

- W. Zirwas, M. Bilal Amin and M. Sternad,
  Coded CSI Reference Signals for 5G - Exploiting Sparsity of FDD Massive MIMO Radio Channels. *20th International ITG Workshop on Smart Antennas (WSA 2016)*, Munich, Germany, March 2016.
  http://www.signal.uu.se/Publications/abstracts/c1601.html

- W. Zirwas, M. Sternad and R. Apelfröjd,
  Key Solutions for a Massive MIMO FDD System.
  IEEE Personal Indoor and Mobile Radio Communications, PIMRC´17, Montreal, Canada, October 2017.
  http://www.signal.uu.se/Publications/abstracts/c1704.html

- R. Apelfröjd, W. Zirwas and M. Sternad,
  Low-Overhead Cyclic Reference Signals for Channel Estimation in FDD Massive MIMO. *IEEE Transactions on Communications, vol. 67, no. 5, May 2019.*
  http://www.signal.uu.se/Publications/abstracts/p1901.html

For more information, see: http://www.signal.uu.se/Research/4G5Gwireless.html
Ph.D. Thesis by R. Apelfröjd: http://www.signal.uu.se//Publications/abstracts/a181.html