Blind Pilot Decontamination

Ralf R. Müller

Professor for Digital Communications
Friedrich-Alexander University Erlangen-Nuremberg

Adjunct Professor for Wireless Networks
Norwegian University of Science and Technology

joint work with

Laura Cottatellucci — Institute Eurecom, France

Mikko Vehkaperä — Aalto University, Finland

25-Jun-2013

This work was supported in part by the FP7 project
HARP

Massive MIMO

Massive MIMO mimics the idea of spread spectrum.

- Spread spectrum:
  - Massive use of bandwidth

Massive MIMO

Massive MIMO mimics the idea of spread spectrum.

- Spread spectrum:
  - Massive use of bandwidth
  - Large processing gain

Massive MIMO

Massive MIMO mimics the idea of spread spectrum.

- Spread spectrum:
  - Massive use of bandwidth
  - Large processing gain
- Massive MIMO:
  - Massive use of antenna elements

Massive MIMO

Massive MIMO mimics the idea of spread spectrum.

- Spread spectrum:
  - Massive use of bandwidth
  - Large processing gain
- Massive MIMO:
  - Massive use of antenna elements
  - Large array gain

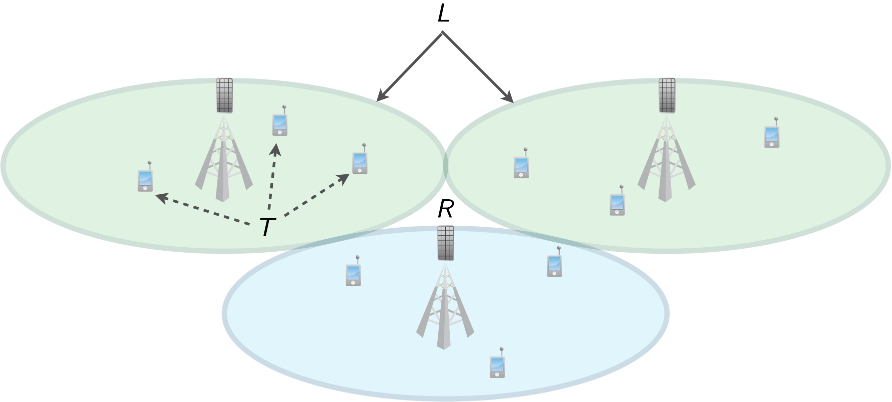
Massive MIMO

Massive MIMO mimics the idea of spread spectrum.

- Spread spectrum:
  - Massive use of bandwidth
  - Large processing gain
- Massive MIMO:
  - Massive use of antenna elements
  - Large array gain

Both systems can operate in arbitrarily strong noise and interference.

Uplink (Reverse Link) System Model



$R \gg T$

$L \sim T$

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z}$$

Pilot Contamination

For T transmit antennas and R receive antennas, even for a static channel, RT channel coefficients must be estimated.

- Linear channel estimation:
  - The array gain, can be utilized for data detection, but **not** for channel estimation.

Pilot Contamination

For T transmit antennas and R receive antennas, even for a static channel, RT channel coefficients must be estimated.

- Linear channel estimation:
  - The array gain, can be utilized for data detection, but **not** for channel estimation.
  - Channel estimation ultimately limits performance.

Pilot Contamination

For T transmit antennas and R receive antennas, even for a static channel, RT channel coefficients must be estimated.

- Linear channel estimation:
  - The array gain, can be utilized for data detection, but **not** for channel estimation.
  - Channel estimation ultimately limits performance.
- General channel estimation:
  - Can the array gain can be utilized for both channel estimation and data detection?

Pilot Contamination

For T transmit antennas and R receive antennas, even for a static channel, RT channel coefficients must be estimated.

- Linear channel estimation:
  - The array gain, can be utilized for data detection, but **not** for channel estimation.
  - Channel estimation ultimately limits performance.
- General channel estimation:
  - Can the array gain can be utilized for both channel estimation and data detection?
  - Is the performance limited by channel estimation?

Pilot Contamination

For T transmit antennas and R receive antennas, even for a static channel, RT channel coefficients must be estimated.

- Linear channel estimation:
  - The array gain, can be utilized for data detection, but **not** for channel estimation.
  - Channel estimation ultimately limits performance.
- General channel estimation:
  - Can the array gain can be utilized for both channel estimation and data detection?
  - Is the performance limited by channel estimation?

How to estimate a massive MIMO channel appropriately?

Blind Interference Rejection

This topic was well studied in the '90s in context of spread-spectrum, *see e.g. U. Madhow: "Blind adaptive interference suppression for direct sequence CDMA," Proceedings of the IEEE, Oct. 1998.*

Blind Interference Rejection

This topic was well studied in the '90s in context of spread-spectrum, *see e.g. U. Madhow: "Blind adaptive interference suppression for direct sequence CDMA," Proceedings of the IEEE, Oct. 1998.*

- Idea:
  - The signal of interest and the interference are almost orthogonal.

Blind Interference Rejection

This topic was well studied in the '90s in context of spread-spectrum, *see e.g. U. Madhow: "Blind adaptive interference suppression for direct sequence CDMA," Proceedings of the IEEE, Oct. 1998.*

- Idea:
  - The signal of interest and the interference are almost orthogonal.
  - We need not know the channel coefficients of the interference, but only the subspace the interference occupies.

Blind Interference Rejection

This topic was well studied in the '90s in context of spread-spectrum, *see e.g. U. Madhow: "Blind adaptive interference suppression for direct sequence CDMA," Proceedings of the IEEE, Oct. 1998.*

- Idea:
  - The signal of interest and the interference are almost orthogonal.
  - We need not know the channel coefficients of the interference, but only the subspace the interference occupies.
- Implementation:
  - Project onto the orthogonal complement of the interference subspace.

Blind Interference Rejection

This topic was well studied in the '90s in context of spread-spectrum, *see e.g. U. Madhow: "Blind adaptive interference suppression for direct sequence CDMA," Proceedings of the IEEE, Oct. 1998.*

- Idea:
  - The signal of interest and the interference are almost orthogonal.
  - We need not know the channel coefficients of the interference, but only the subspace the interference occupies.
- Implementation:
  - Project onto the orthogonal complement of the interference subspace.

How to find the interference subspace or its orthogonal complement?

Matched Filter Projection

Let us start the considerations with a SIMO system and white noise only.

Matched Filter Projection

Let us start the considerations with a SIMO system and white noise only.

- Let $\mathbf{y}_c$ be the column vector received at the receive array at time c and $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_C]$ with C denoting the coherence time.

Matched Filter Projection

Let us start the considerations with a SIMO system and white noise only.

- Let $\mathbf{y}_c$ be the column vector received at the receive array at time c and $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_C]$ with C denoting the coherence time.
- We would like to find a linear filter $\mathbf{m}$, such that $\mathbf{m}^\dagger \mathbf{Y}$ has high SNR.

Matched Filter Projection

Let us start the considerations with a SIMO system and white noise only.

- Let $\mathbf{y}_c$ be the column vector received at the receive array at time c and $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_C]$ with C denoting the coherence time.
- We would like to find a linear filter $\mathbf{m}$, such that $\mathbf{m}^\dagger \mathbf{Y}$ has high SNR.
- We get

$$\mathbf{m} = \operatorname*{argmax}_{\mathbf{m}_0} \frac{||\mathbf{m}_0^\dagger \mathbf{Y}||^2}{||\mathbf{m}_0||^2} = \operatorname*{argmax}_{\mathbf{m}_0} \frac{\mathbf{m}_0^\dagger \mathbf{Y}\mathbf{Y}^\dagger \mathbf{m}_0}{\mathbf{m}_0^\dagger \mathbf{m}_0}$$

is that eigenvector of $\mathbf{Y}\mathbf{Y}^\dagger$ that corresponds to the largest eigenvalue.

Matched Filter Projection II

Next, consider a MIMO system with $T > 1$ transmit antennas and white noise.

Matched Filter Projection II

Next, consider a MIMO system with $T > 1$ transmit antennas and white noise.

- Now, we look for a basis $\mathbf{M}$ of the T-dimensional subspace containing the signal of interest.

Matched Filter Projection II

Next, consider a MIMO system with $T > 1$ transmit antennas and white noise.

- Now, we look for a basis $\mathbf{M}$ of the T-dimensional subspace containing the signal of interest.
- We find it by an eigenvalue decomposition of $\mathbf{YY}^{\dagger}$ picking those eigenvectors which correspond to the T largest eigenvalues.

Matched Filter Projection II

Next, consider a MIMO system with $T > 1$ transmit antennas and white noise.

- Now, we look for a basis $\mathbf{M}$ of the T-dimensional subspace containing the signal of interest.
- We find it by an eigenvalue decomposition of $\mathbf{Y}\mathbf{Y}^\dagger$ picking those eigenvectors which correspond to the T largest eigenvalues.
- We now project the received signal onto that subspace

$$\mathbf{Y}' = \mathbf{M}^\dagger \mathbf{Y}$$

and dismiss all noise components outside that subspace.

Matched Filter Projection II

Next, consider a MIMO system with $T > 1$ transmit antennas and white noise.

- Now, we look for a basis $\mathbf{M}$ of the T-dimensional subspace containing the signal of interest.
- We find it by an eigenvalue decomposition of $\mathbf{Y}\mathbf{Y}^\dagger$ picking those eigenvectors which correspond to the T largest eigenvalues.
- We now project the received signal onto that subspace

$$\mathbf{Y}' = \mathbf{M}^\dagger \mathbf{Y}$$

  and dismiss all noise components outside that subspace.
- By the massive MIMO philosophy, i.e. $T \ll R$, this subspace is much smaller than the full space.

Matched Filter Projection II

Next, consider a MIMO system with $T > 1$ transmit antennas and white noise.

- Now, we look for a basis $\mathbf{M}$ of the T-dimensional subspace containing the signal of interest.
- We find it by an eigenvalue decomposition of $\mathbf{Y}\mathbf{Y}^\dagger$ picking those eigenvectors which correspond to the T largest eigenvalues.
- We now project the received signal onto that subspace

$$\mathbf{Y}' = \mathbf{M}^\dagger \mathbf{Y}$$

  and dismiss all noise components outside that subspace.
- By the massive MIMO philosophy, i.e. $T \ll R$, this subspace is much smaller than the full space.

We have utilized the array gain without estimating the channel.

Matched Filter Projection III

Consider now the general case (noise, interference and a MIMO system with $T > 1$ transmit antennas and $R \gg T$ receive antennas).

Matched Filter Projection III

Consider now the general case (noise, interference and a MIMO system with $T > 1$ transmit antennas and $R \gg T$ receive antennas).

- While white noise is small in all components if

$$\mathrm{SNR} \gg \frac{T}{R}$$

  the interference typically concentrates in few signal dimensions where it is strong.

Matched Filter Projection III

Consider now the general case (noise, interference and a MIMO system with $T > 1$ transmit antennas and $R \gg T$ receive antennas).

- While white noise is small in all components if

$$\mathrm{SNR} \gg \frac{T}{R} \ll 1,$$

the interference typically concentrates in few signal dimensions where it is strong.

Matched Filter Projection III

Consider now the general case (noise, interference and a MIMO system with $T > 1$ transmit antennas and $R \gg T$ receive antennas).

- While white noise is small in all components if

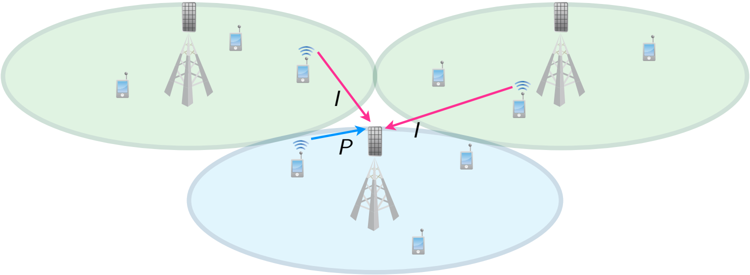
$$\mathrm{SNR} \gg \frac{T}{R} \ll 1,$$

the interference typically concentrates in few signal dimensions where it is strong.

How to distinguish the signal of interest from interference?

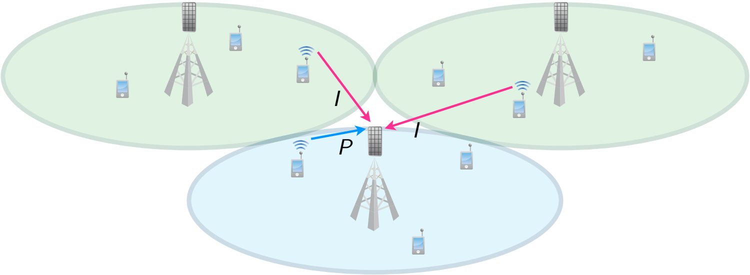
Power Controlled Hand-Off

Consider power-controlled hand-off and perfect received power control.



Power Controlled Hand-Off

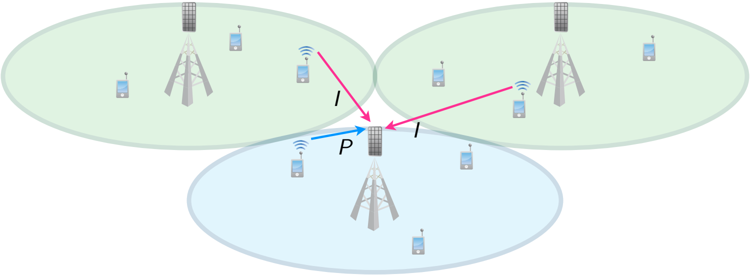
Consider power-controlled hand-off and perfect received power control.



- Interfering signals cannot be stronger than signals of interest, i.e. $P \geq I$.

Power Controlled Hand-Off

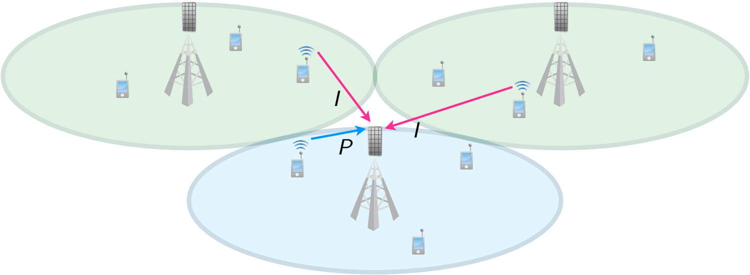
Consider power-controlled hand-off and perfect received power control.



- Interfering signals cannot be stronger than signals of interest, i.e. $P \geq I$.
- Most interfering signals are noticeably weaker than the signals of interest.

Power Controlled Hand-Off

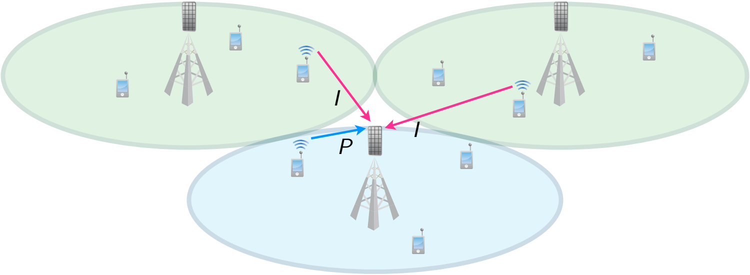
Consider power-controlled hand-off and perfect received power control.



- Interfering signals cannot be stronger than signals of interest, i.e. $P \geq I$.
- Most interfering signals are noticeably weaker than the signals of interest.
- For vanishing load $\alpha = T/R \to 0$, the signals of interest can be separated from the interference.

Power Controlled Hand-Off

Consider power-controlled hand-off and perfect received power control.



- Interfering signals cannot be stronger than signals of interest, i.e. $P \geq I$.
- Most interfering signals are noticeably weaker than the signals of interest.
- For vanishing load $\alpha = T/R \to 0$, the signals of interest can be separated from the interference.

What if the load is small, but not vanishing?

Asymptotic Eigenvalue Distribution

The exact *asymptotic* eigenvalue distribution can be given implicitly in terms of its Stieltjes transform

$$\mathrm{G}(s) = \int \frac{\mathrm{dP}(x)}{x-s}.$$

For an iid. channel, we find

$$\begin{aligned} s\mathrm{G}(s)+1 = &- \frac{PTC\alpha\big(s\mathrm{G}(s)+1-\kappa\big)\mathrm{G}(s)}{\alpha\kappa - PTC\big(s\mathrm{G}(s)+1-\kappa\big)\mathrm{G}(s)} \\ &- \int \frac{xLTC\alpha\big(s\mathrm{G}(s)+1-\kappa\big)\mathrm{G}(s)\,\mathrm{dP_I}(x)}{\alpha\kappa\,\mathrm{p_I}(x) - xTC\big(s\mathrm{G}(s)+1-\kappa\big)\mathrm{G}(s)} \\ &- \frac{WC\alpha\big(s\mathrm{G}(s)+1-\kappa\big)\mathrm{G}(s)}{\kappa} \end{aligned}$$

with W denoting the noise power, $\kappa = \frac{C}{R}$, and $\mathrm{P_I}(x)$ denoting the power distribution of the interference.

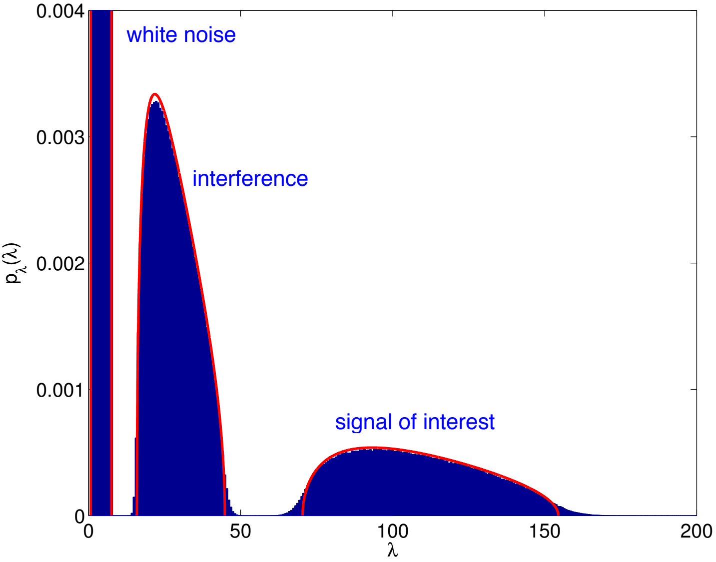
Asymptotic Eigenvalue Distribution

Assuming that all LT interferers have power I, i.e. $\mathrm{p_I}(x) = \delta(x - I)$, the fixed-point equation for the Stieltjes transform simplifies to

$$
\begin{aligned}
s\mathrm{G}(s) + 1 = &- \frac{PTC\alpha\big(s\mathrm{G}(s) + 1 - \kappa\big)\mathrm{G}(s)}{\alpha\kappa - PTC\big(s\mathrm{G}(s) + 1 - \kappa\big)\mathrm{G}(s)} \\
&- \frac{ILTC\alpha\big(s\mathrm{G}(s) + 1 - \kappa\big)\mathrm{G}(s)}{\alpha\kappa - ITC\big(s\mathrm{G}(s) + 1 - \kappa\big)\mathrm{G}(s)} \\
&- \frac{WC\alpha\big(s\mathrm{G}(s) + 1 - \kappa\big)\mathrm{G}(s)}{\kappa}
\end{aligned}
$$

with W denoting the noise power and $\kappa = \frac{C}{R}$.

Empirical Eigenvalue Distribution



$$R = 300$$
$$T = 10$$
$$C = 1000$$
$$L = 2$$
$$W = 1000$$
$$P = 100$$
$$I = 25$$

Uplink vs. Downlink

We have detected the uplink data without estimating the full channel.

For energy concentration on the downlink (forward link), we need a good estimate of the full channel matrix $\mathbf{H}$.

Uplink vs. Downlink

We have detected the uplink data without estimating the full channel.

For energy concentration on the downlink (forward link), we need a good estimate of the full channel matrix $\mathbf{H}$.

1. We use time-division duplex.

Uplink vs. Downlink

We have detected the uplink data without estimating the full channel.

For energy concentration on the downlink (forward link), we need a good estimate of the full channel matrix $\mathbf{H}$.

1. We use time-division duplex.
2. We project the received signal $\mathbf{Y}$ onto the orthogonal complement of the interference.

Uplink vs. Downlink

We have detected the uplink data without estimating the full channel.

For energy concentration on the downlink (forward link), we need a good estimate of the full channel matrix $\mathbf{H}$.

1. We use time-division duplex.
2. We project the received signal $\mathbf{Y}$ onto the orthogonal complement of the interference.
3. We use all uplink data to estimate the downlink (forward link) channel to high accuracy.

Eigenvalue Spread

Assume an i.i.d. channel matrix and $R \gg T \to \infty$.

The eigenvalues of the signal of interest are confined in an interval centered at the received power P with width

$$4P\sqrt{\frac{T}{R} + \frac{T}{C}}.$$

Eigenvalue Spread

Assume an i.i.d. channel matrix and $R \gg T \to \infty$.

The eigenvalues of the signal of interest are confined in an interval centered at the received power P with width

$$4P\sqrt{\frac{T}{R}+\frac{T}{C}}.$$

The eigenvalues of the interference are confined in an interval centered at the interference power I with width

$$4I\sqrt{\frac{LT}{R}+\frac{LT}{C}}$$

where L denotes the number of interfering cells.

For massive MIMO, the two widths are quite small.

Eigenvalue Separation

The two intervals do not overlap if

$$\frac{P}{I} > \frac{1 + 2\sqrt{\frac{LT}{R} + \frac{LT}{C}}}{1 - 2\sqrt{\frac{T}{R} + \frac{T}{C}}}.$$

Eigenvalue Separation

The two intervals do not overlap if

$$\frac{P}{I} > \frac{1 + 2\sqrt{\frac{LT}{R} + \frac{LT}{C}}}{1 - 2\sqrt{\frac{T}{R} + \frac{T}{C}}}.$$

If the two intervals do not overlap, we can totally reject the interference by means of eigenvalue decomposition.

Eigenvalue Separation

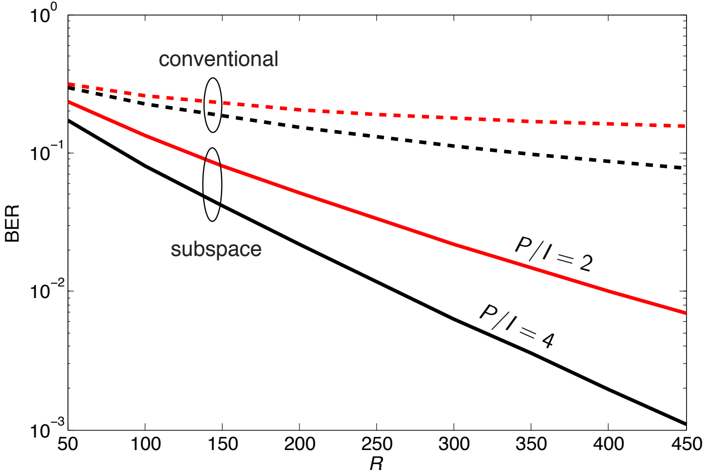
The two intervals do not overlap if

$$\frac{P}{I} > \frac{1 + 2\sqrt{\frac{LT}{R} + \frac{LT}{C}}}{1 - 2\sqrt{\frac{T}{R} + \frac{T}{C}}}.$$

If the two intervals do not overlap, we can totally reject the interference by means of eigenvalue decomposition.

For finite number of receive antennas, the interval boundaries are not sharp, but have exponentially decaying tails.

BER vs. Array Size



$T = 3$

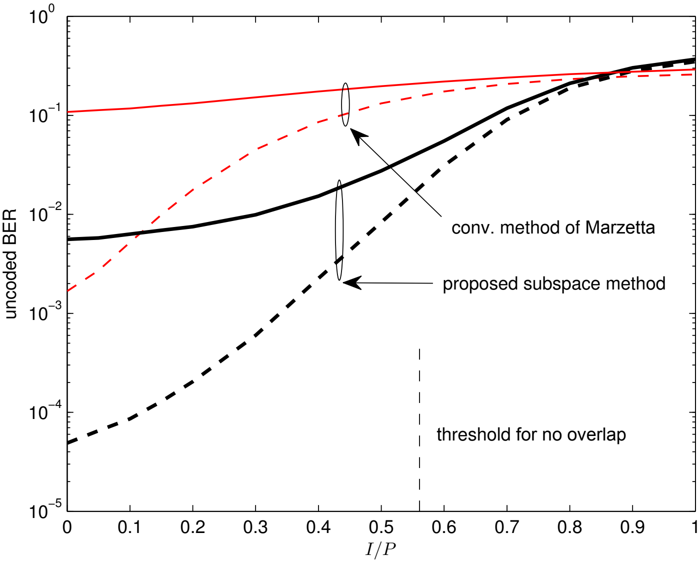
$C = 1000$

$L = 2$

$\mathrm{SNR} = -10\mathrm{dB}$

1 pilot symbol per transmit antenna and cell

BER vs. Power Margin



$$
\begin{aligned}
R &= 200 \\
T &= 2 \\
C &= 400 \\
L &= 2 \\
W &= 1 \\
P &= 0.1
\end{aligned}
$$

1 (–) or 10 (- -) pilot symbols per transmit antenna and cell

Power Margin

How to guarantee a sufficient power margin between the signal of interest and the interference?

Power Margin

How to guarantee a sufficient power margin between the signal of interest and the interference?

Two antennas per user.

Power Margin

How to guarantee a sufficient power margin between the signal of interest and the interference?

Two antennas per user.

If a user experiences equally good channel conditions to several base stations/access points, the user forms a beam that favors one of the base stations/access points over the others.

Power Margin

How to guarantee a sufficient power margin between the signal of interest and the interference?

Two antennas per user.

If a user experiences equally good channel conditions to several base stations/access points, the user forms a beam that favors one of the base stations/access points over the others.

If the power margin is sufficient without beam forming, the user can use the two antennas for spatial multiplexing.

Power Margin

How to guarantee a sufficient power margin between the signal of interest and the interference?

Two antennas per user.

If a user experiences equally good channel conditions to several base stations/access points, the user forms a beam that favors one of the base stations/access points over the others.

If the power margin is sufficient without beam forming, the user can use the two antennas for spatial multiplexing.

Pro: A sufficient power margin can be established (with high probability).
Con: Users at cell boundaries may suffer from reduced data rate.

Conclusions

- Pilot contamination is not a fundamental effect, but an artefact of linear channel estimation.

Conclusions

- Pilot contamination is not a fundamental effect, but an artefact of linear channel estimation.
- Pilot decontamination based on power control works well under the simulated conditions.

Conclusions

- Pilot contamination is not a fundamental effect, but an artefact of linear channel estimation.
- Pilot decontamination based on power control works well under the simulated conditions.
- The algorithm requires real-time eigenvalue or singular value decompositions.

Literature

- U. Madhow,”Blind adaptive interference suppression for direct sequence CDMA,“ Proc. of the IEEE, vol. 86, no. 10, pp. 2049–2069, Oct. 1998.
- H. Q. Ngo and E. G. Larsson,”EVD-based channel estimation in multicell multiuser MIMO system with very large antenna arrays,“ *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012.
- H. Yin, D. Gesbert, M. Filippou, and Y. Liu, ”A coordinated approach to channel estimation in large-scale multiple-antenna systems,“ *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- R. R. Müller, M. Vehkaperä, and L. Cottatellucci,”Blind pilot decontamination,“*Proc. of 17th International ITG Workshop on Smart Antennas (WSA 2013)*, Stuttgart, Germany, Mar. 2013.
- L. Cottatellucci, R. R. Müller, M. Vehaperä, ”Analysis of pilot decontamination based on power control“, *Proc. of IEEE Vehicular Technology Conference (VTC)*, Dresden, Germany, Jun. 2013.