The Massive MIMO Paradigm – *Fundamentals and State-of-the-Art*

Dr. Emil Björnson

Department of Electrical Engineering (ISY)

Linköping University

Sweden

Outline

1. Massive MIMO: What and Why?
2. System Capacity with Massive MIMO
3. Resource Allocation and Access
4. Energy Efficiency
5. Impact of Hardware Impairments

Dr. Emil Björnson

- *PhD from KTH Royal Institute of Technology*
- *Associate professor at Linköping University*
- *10 year experience of MIMO research*
- *1 book and 5 best paper awards*
- *Several pending patent applications*

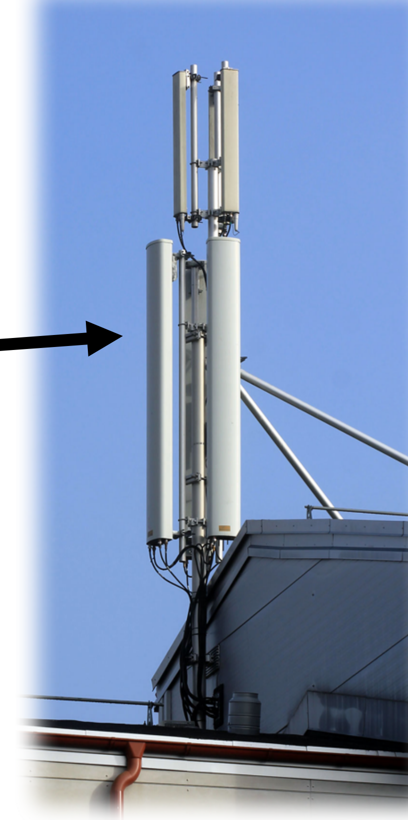
Part 1

Massive MIMO: What and Why?

LINKÖPING UNIVERSITY

INTRODUCTION

Wireless Communications



Martin Cooper's law

The number of voice/data connections has doubled every 2.5 years (+32% per year)

Early technology

1890s: Telegraph

1920s: Audio radio

1930s: Walkie-talkie

1930s: Television

→

Cellular

1991: 2G

2001: 3G

2012: 4G

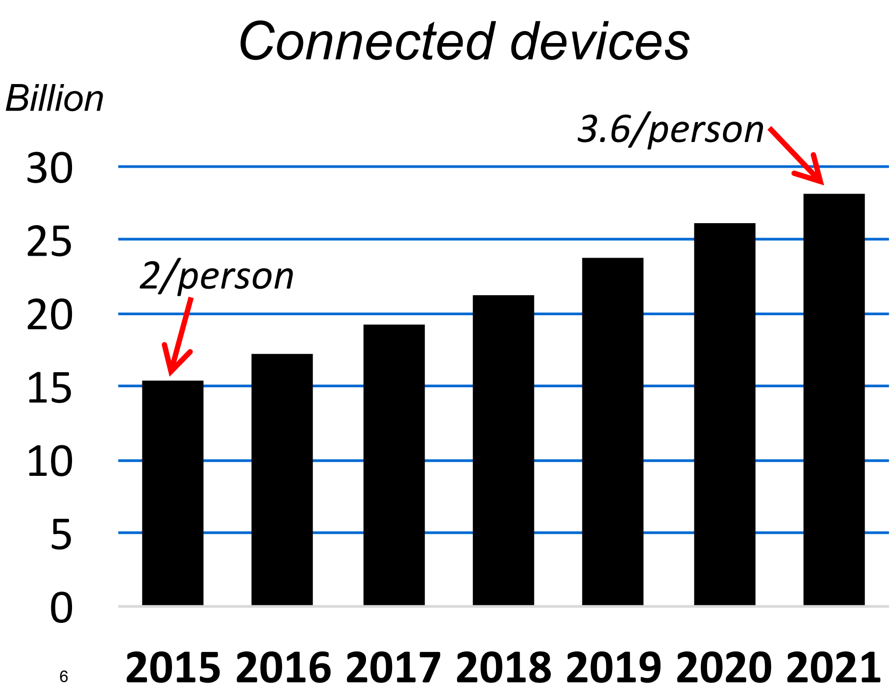
2020: 5G?

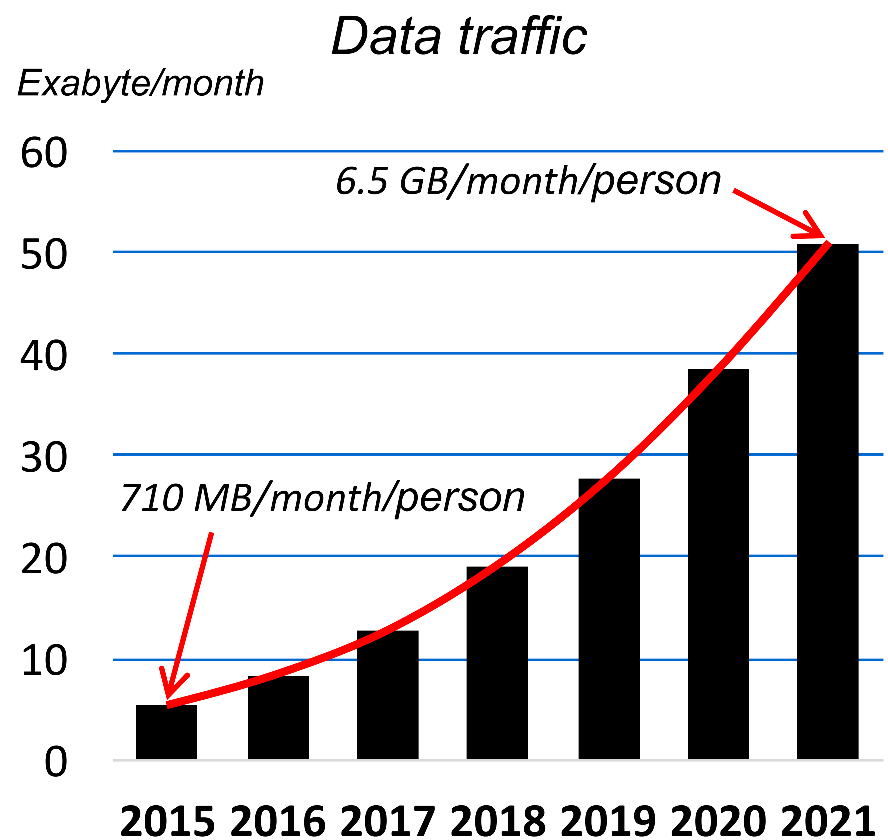
Image source: Wikipedia

Current Trends

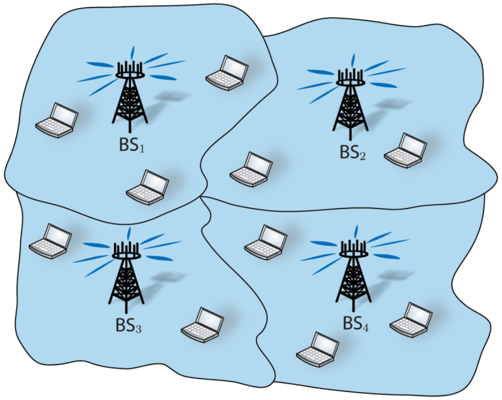
Data source:
Ericsson Mobility Report
(November 2015)

- Future Network Traffic Growth
  - 10% more connected devices per year
  - 45% annual data traffic growth – faster than in the past!





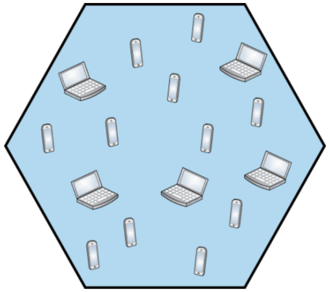
Evolving Networks for Higher Traffic



Cellular networks

Coverage area divided into cells
Users served by a base station

- Increase Network Throughput [bit/s/km^2]
  - Consider a given area



- Simple Formula for Network Throughput:

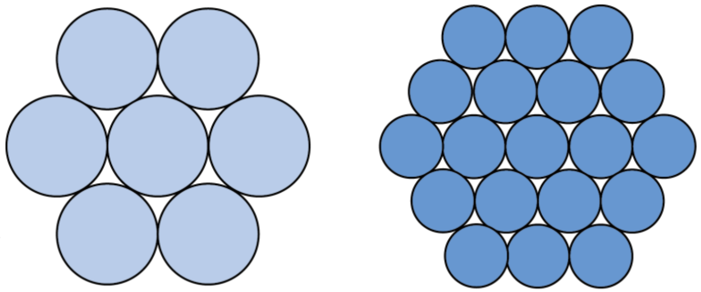
$$\underbrace{\text{Throughput}}_{\text{bit/s/km}^2} = \underbrace{\text{Cell density}}_{\text{Cell/km}^2} \cdot \underbrace{\text{Available spectrum}}_{\text{Hz}} \cdot \underbrace{\text{Spectral efficiency}}_{\text{bit/s/Hz/Cell}}$$

  - Ways to achieve 1000x improvement:

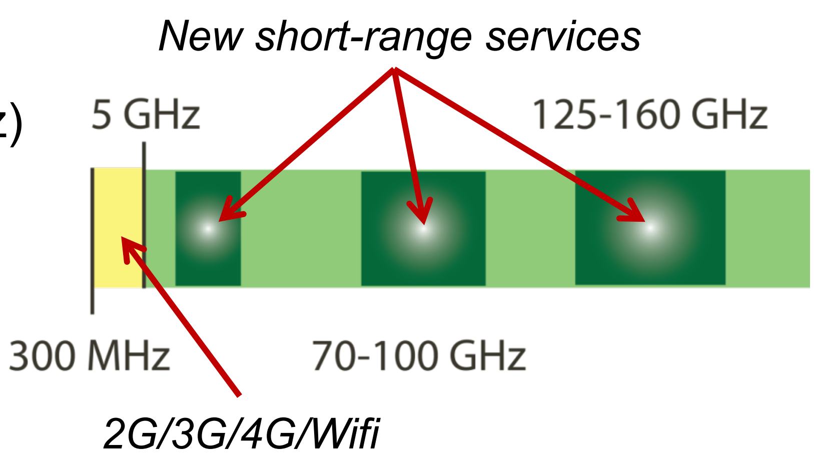
	Higher cell density	More spectrum	Higher spectral efficiency
Nokia (2011)	10x	10x	10x
SK Telecom (2012)	56x	3x	6x

Conventional Solutions

- Higher Cell Density
  - Traditional way to improve throughput
  - Cut cell radius by z → z^2 times more cells
  - Issues: High rent and deployment costs
    Interference is getting worse
    WiFi + Cellular is already dense: *Coverage is the issue*!



- More Spectrum
  - Range suitable for coverage: < 5 GHz
  - Already allocated for services!
    (Cellular: ~ 550 MHz, WiFi: ~ 550 MHz)
  - Far above 5 GHz: High propagation losses → Mainly short-range hotspots



Higher Spectral Efficiency

“Imagine that we decided to reward the first person who finds a way to make spectrum use below 5 GHz 50 or 100 times more efficient over the next decade. The reward could be something simple—say 10 megahertz of spectrum suitable for mobile broadband.”

FCC Commissioner Jessica Rosenworcel
Marconi Society Anniversary Symposium, Oct. 2, 2014.

Price of sub-5 GHz Spectrum

January 2015: FCC sold 65 MHz at 1.7-2.1 GHz for $45 billion

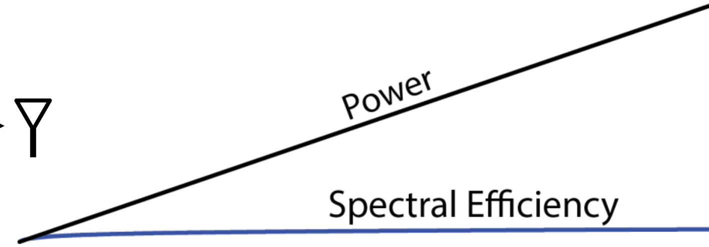
Can FCC’s $50\times$ goal be reached?

Principles to achieve

HIGHER SPECTRAL EFFICIENCY

Higher Spectral Efficiency

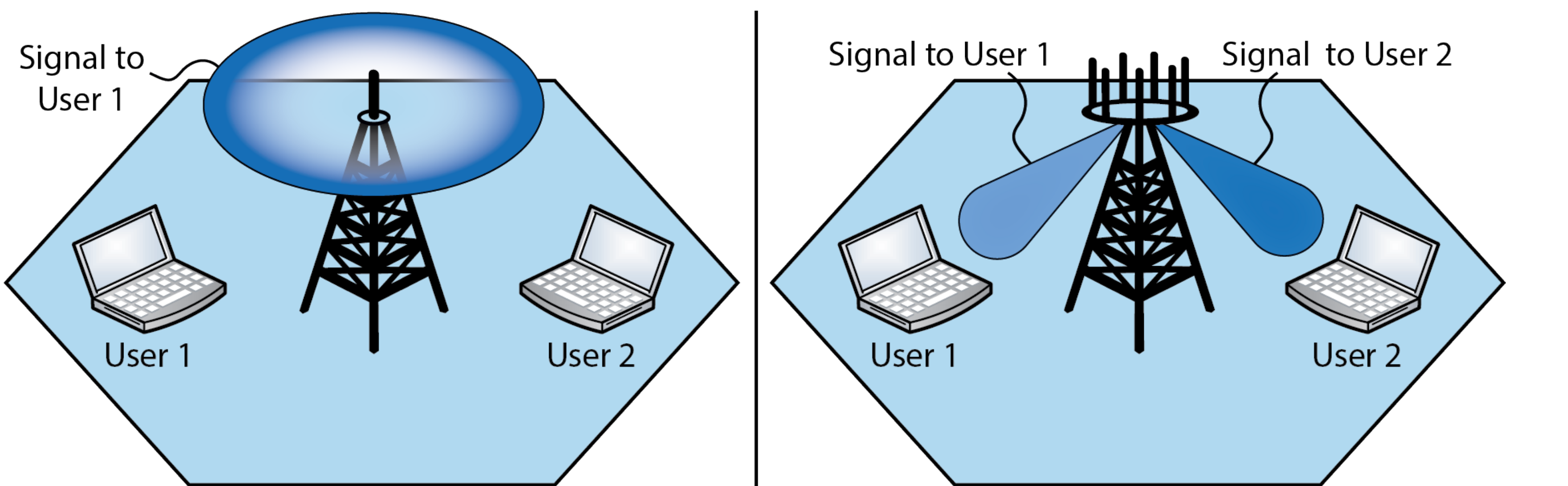
- Point-to-Point Spectral Efficiency: 
  - Governed by Shannon's capacity limit:



$$\log_2\left(1 + \frac{\text{Received Signal Power}}{\text{Interference Power} + \text{ Noise Power}}\right) \quad [\text{bit/s/Hz/user}]$$

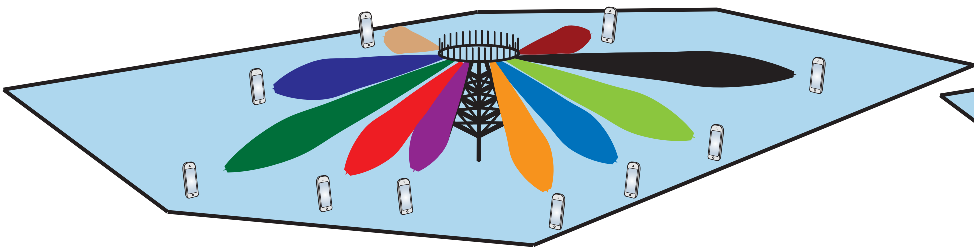
  - Issue: 4 bit/s/Hz → 8 bit/s/Hz requires $17\times$ more power!

- Many Parallel Transmissions: *Spatially focused to each desired user*

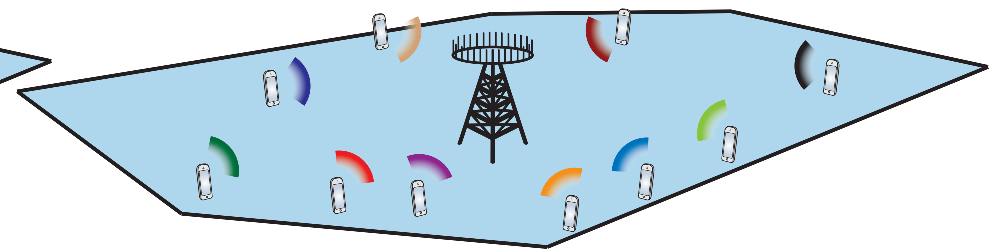


Single-Antenna Transmission | Multi-Antenna Transmission

Multi-User MIMO (Multiple-input Multiple-output)



Downlink: Multi-user precoding



Uplink: Multi-user detection

Implementation concepts

1987: J. Winters, "*Optimum combining for indoor radio systems with multiple users*," IEEE Trans. Commun.

1990: S. Swales, M. Beach, D. Edwards, J. McGeehan, "*The performance enhancement of multibeam adaptive base-station antennas for cellular land mobile radio systems,*" IEEE Trans. Veh. Technol.

1991: S. Anderson, M. Millnert, M. Viberg, B. Wahlberg, "*An adaptive array for mobile communication systems*," IEEE Trans. Veh. Technol.

Information-theoretic Limits

2003: G. Caire, S. Shamai, "*On the achievable throughput of a multiantenna Gaussian broadcast channel*," IEEE Trans. Inf. Theory.

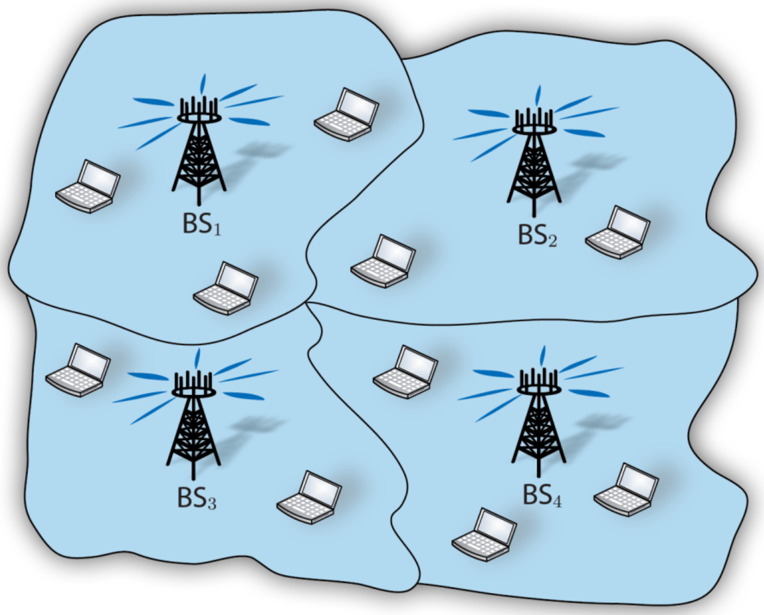
2003: A. Goldsmith, S. Jafar, N. Jindal, S. Vishwanath, "*Capacity limits of MIMO channels*," IEEE J. Sel. Areas Commun.

2003: P. Viswanath, D. Tse, "*Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality*," IEEE Trans. Inf. Theory.

Wireless standards: *2005: HC-SDMA, 2008: LTE, 2013: WiFi 802.11ac*

Multi-User MIMO: Theory and Practice

- Cellular Multi-User MIMO
  - Base stations (BSs) with M antennas
  - Parallel uplink/downlink for K users
  - Channel coherence block: τ_c symbols



- Theory: Hardware is Limiting
  - Spectral efficiency roughly prop. to

$$\min\left(M, K, \frac{\tau_c}{2}\right)$$

  - $2\times$ improvement $=$ $2\times$ antennas and users $(\tau_c \in [100{,}10000])$

- Practice: Co-User Interference is Limiting
  - Multi-user MIMO in LTE-A: Up to 8 antennas
  - Small gains: Hard to learn users' channels
    Hard to coordinate BSs

End of the MIMO road?

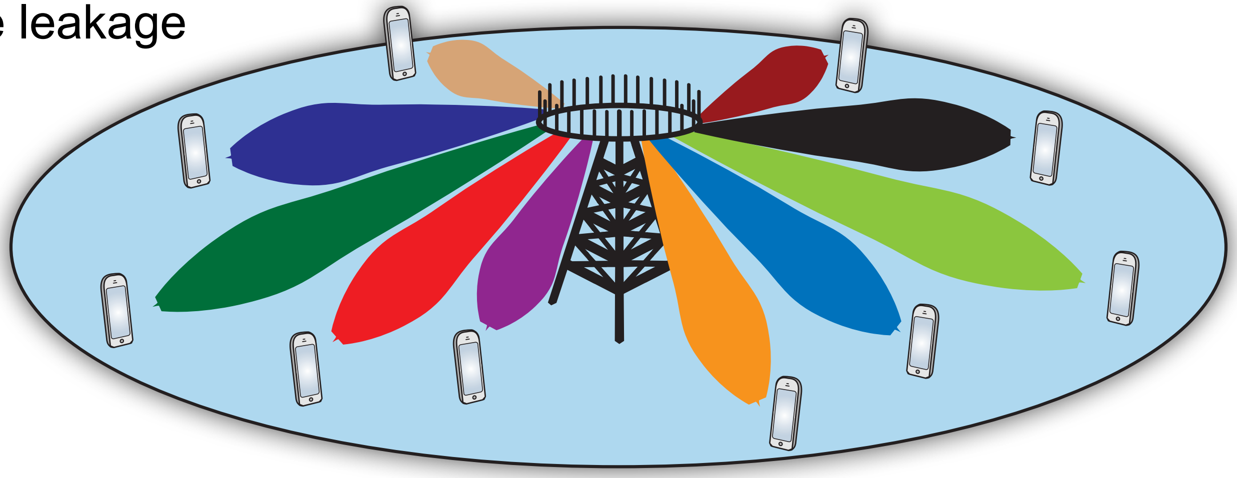
No reason to add more antennas/users?

Taking Multi-User MIMO to the Next Level

- Network Architecture: Massive MIMO
  - Many BS antennas; e.g., $M \approx 200$ antennas, $K \approx 40$ single-antenna users
  - Key: Many more antennas than users: $M \gg K$
  - Very directive signals
  - Little interference leakage

Spectral efficiency prop. to number of users!

$$\min\left(M, K, \frac{\tau_c}{2}\right) \approx K$$



- Seminal work:

  T. Marzetta, "*Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas*," IEEE Trans. Wireless Communications, 2010.
  - 2013 IEEE Guglielmo Marconi Prize Paper Award
  - 2015 IEEE W. R. G. Baker Award

What is the Key Difference from Today?

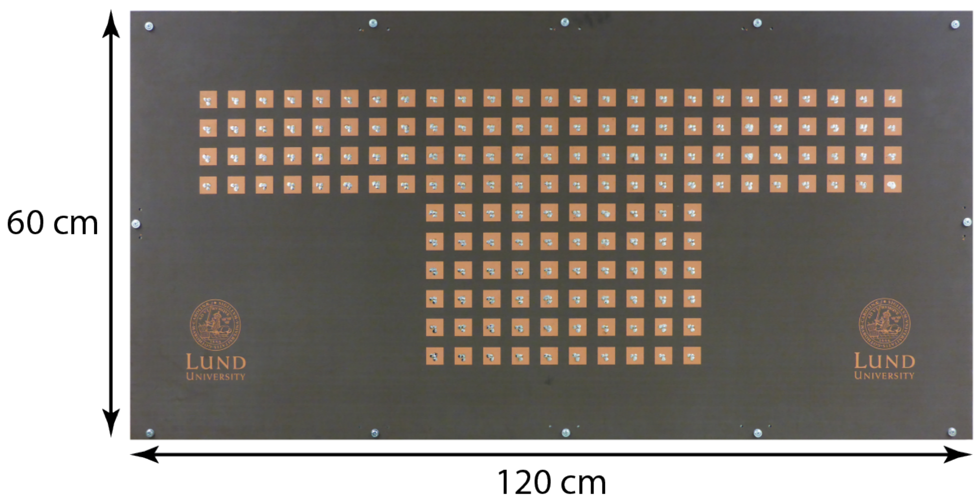
- Number of Antennas? **No, we already have many antennas!**
  - 3G/UMTS: 3 sectors x 20 element-arrays = 60 antennas
  - 4G/LTE-A: 8-MIMO x 30 = 240 antennas

Typical vertical array:
10 antennas x 2 polarizations
Only 2 antenna ports

Massive MIMO Characteristics

Many fully digital steerable antennas

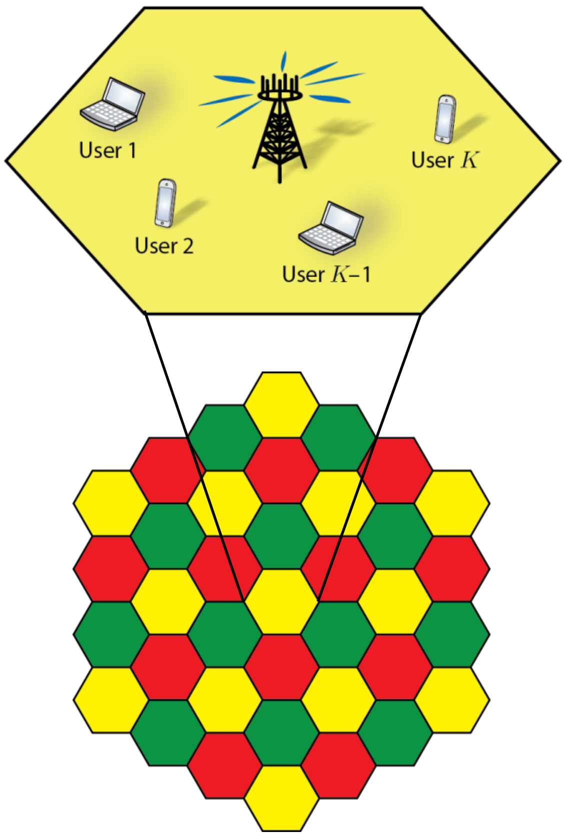
Massive in numbers – not massive in size



160 antenna elements, *LuMaMi testbed, Lund University*

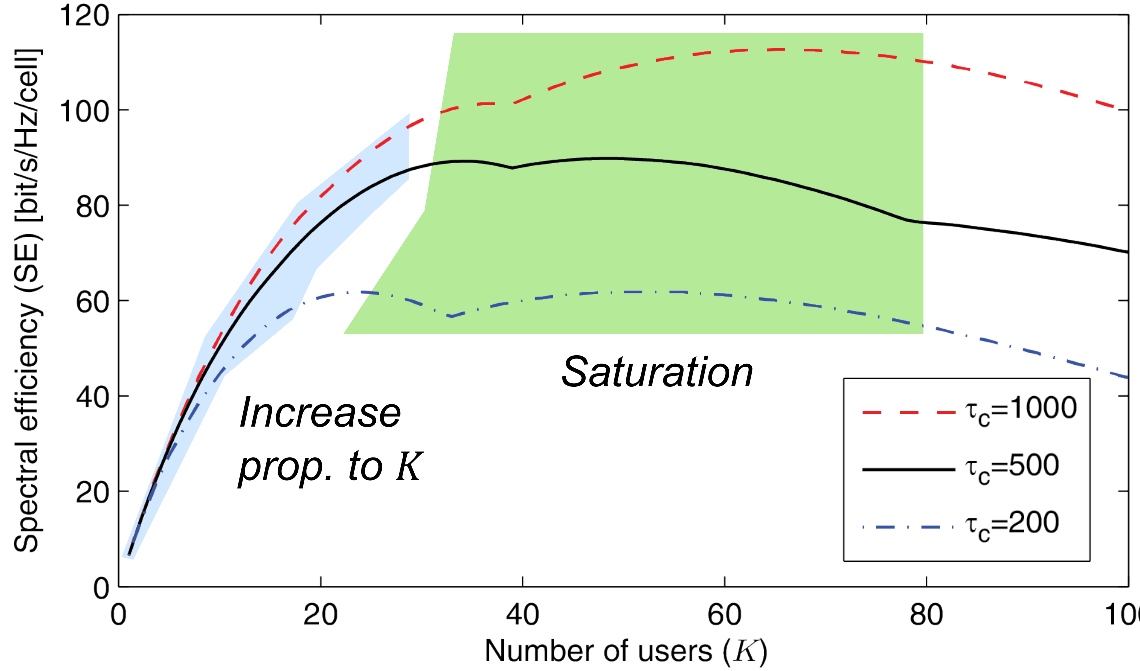
3 sectors, 4 vertical arrays per sector

How Much can Spectral Efficiency be Improved?



Uplink simulation

$M = 200$ antennas, SNR 0 dB, Rayleigh fading



- *Baseline:* 2.25 bit/s/Hz/cell (IMT-Advanced)
- *Massive MIMO:* $25\times - 50\times$ improvement
- Large coherence τ_c is key → Use lower frequencies

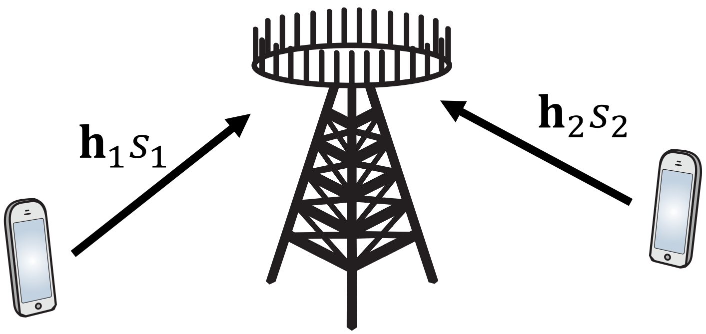
Yes, FCC's $50\times$ goal is within reach!

Key aspects of having

MASSIVE ANTENNA NUMBERS

Asymptotic Channel Orthogonality

- Example: Uplink with i.i.d. Rayleigh Fading
  - Two users, send signals s_k for $k = 1{,}2$
  - Channels: $\mathbf{h}_k = [h_{k1} \dots h_{kM}]^T \sim CN(\mathbf{0}, \mathbf{I}_M)$
  - Noise: $\mathbf{n} \sim CN(\mathbf{0}, \mathbf{I}_M)$
  - Received: $\mathbf{y} = \mathbf{h}_1 s_1 + \mathbf{h}_2 s_2 + \mathbf{n}$



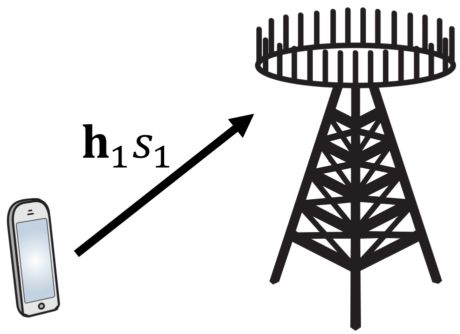
- Linear Detector $\mathbf{v}_1$ for User 1: $\tilde{y}_1 = \mathbf{v}_1^H \mathbf{y} = \boxed{\mathbf{v}_1^H \mathbf{h}_1} s_1 + \boxed{\mathbf{v}_1^H \mathbf{h}_2} s_2 + \boxed{\mathbf{v}_1^H \mathbf{n}}$
  - Maximum ratio filter: $\mathbf{v}_1 = \frac{1}{M} \mathbf{h}_1$
  - Signal remains: $\mathbf{v}_1^H \mathbf{h}_1 = \frac{1}{M} \|\mathbf{h}_1\|^2 \xrightarrow{M \to \infty} \mathrm{E}[|h_{11}|^2] = 1$
  - Interference vanishes: $\mathbf{v}_1^H \mathbf{h}_2 = \frac{1}{M} \mathbf{h}_1^H \mathbf{h}_2 \xrightarrow{M \to \infty} \mathrm{E}[h_{11}^H h_{21}] = 0$
  - Noise vanishes: $\mathbf{v}_1^H \mathbf{n} = \frac{1}{M} \mathbf{h}_1^H \mathbf{n} \xrightarrow{M \to \infty} \mathrm{E}[h_{11}^H n_1] = 0$

Asymptotically noise/interference-free communication: $\tilde{y}_1 \xrightarrow{M \to \infty} s_1$

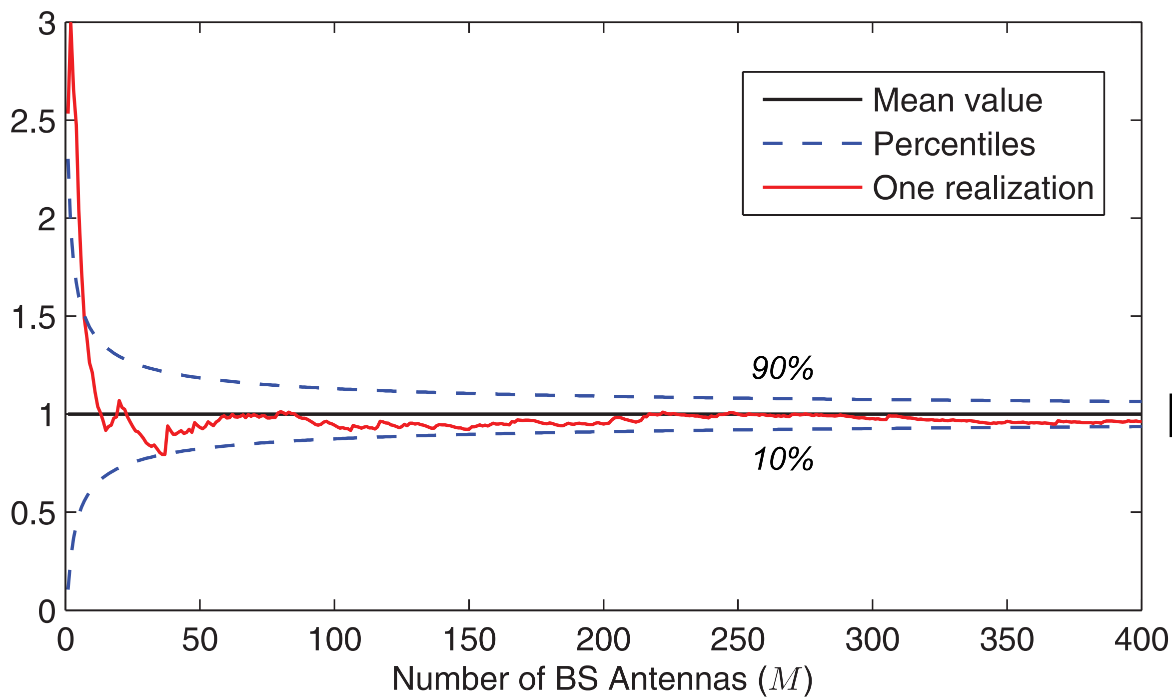
Asymptotic Channel Hardening

Variations of effective channel reduce with M:

$$\frac{1}{M}\|\mathbf{h}_1\|^2 \text{ has } \begin{cases} \text{Mean: } 1 \\ \text{Variance: } 1/M \end{cases}$$



$\mathbf{h}_1 \sim CN(\mathbf{0}, \mathbf{I}_M)$



$\|\mathbf{h}_1\|^2 \approx \mathrm{E}\{\|\mathbf{h}_1\|^2\}$

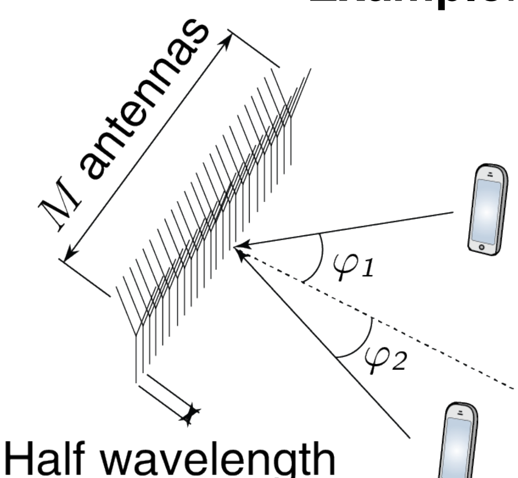
Double benefits: $\|\mathbf{h}_1\|^2$ *scales with* M*, variations reduces*

Orthogonality Only in Isotropic Fading?

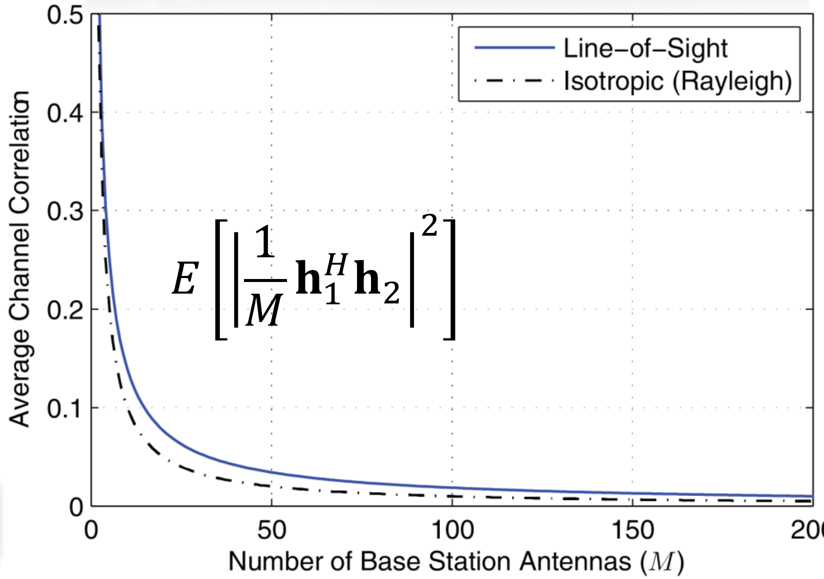
- Assumptions in i.i.d. Rayleigh Fading
  - No dominant directivity
  - Very many scattering objectives

Less true as $M \to \infty$

Example: Line-of-Sight Channels



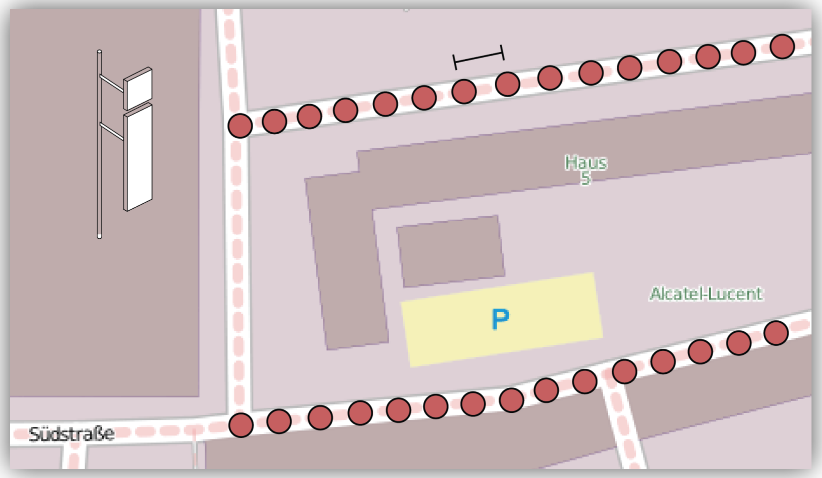
- Random user angles
- M observations:
  - Stronger signal
  - Suppressed noise
- What is $\mathbf{h}_1^H \mathbf{h}_2$?



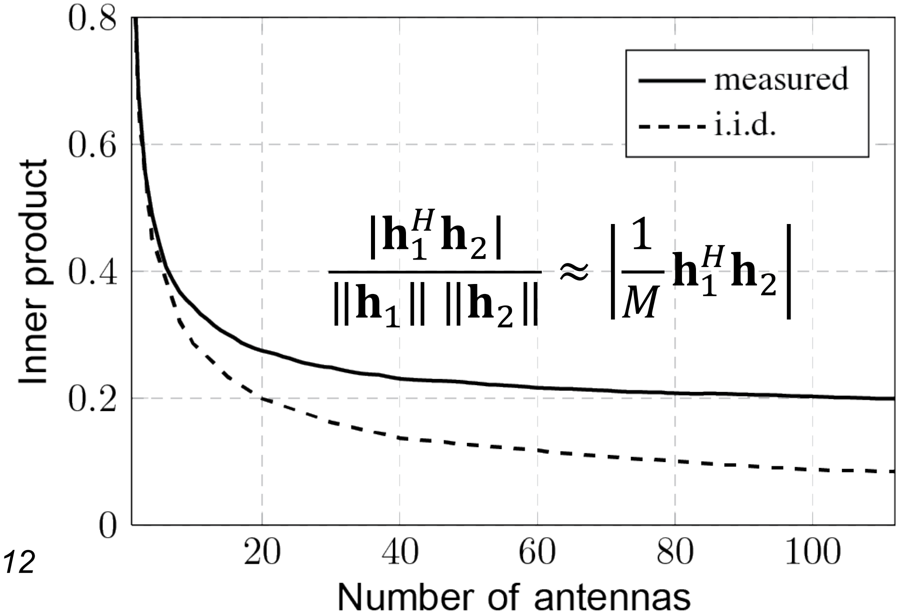
***Difference**: How quickly orthogonality appears*

How Do Practical Channels Behave?

- Measurements show similar results



Source: J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, "Channel Measurements for Large Antenna Arrays," ISWCS 2012

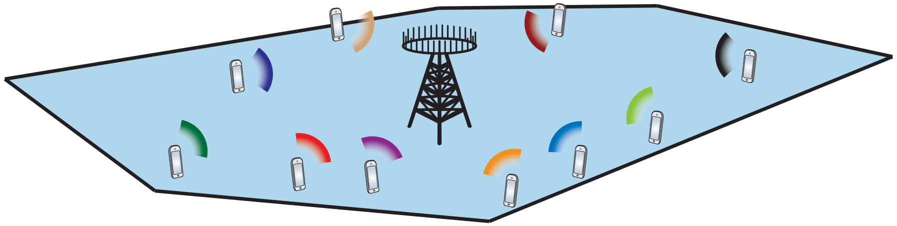


- Asymptotic Favorable Propagation: $\frac{1}{M}\mathbf{h}_1^H \mathbf{h}_2 \to 0$ as $M \to \infty$
  - Achieved in Rayleigh fading and line-of-sight – two extremes!
  - Same behavior expected and observed in practice

Achieving a scalable

MASSIVE MIMO PROTOCOL

Classical Multi-User MIMO vs. Massive MIMO

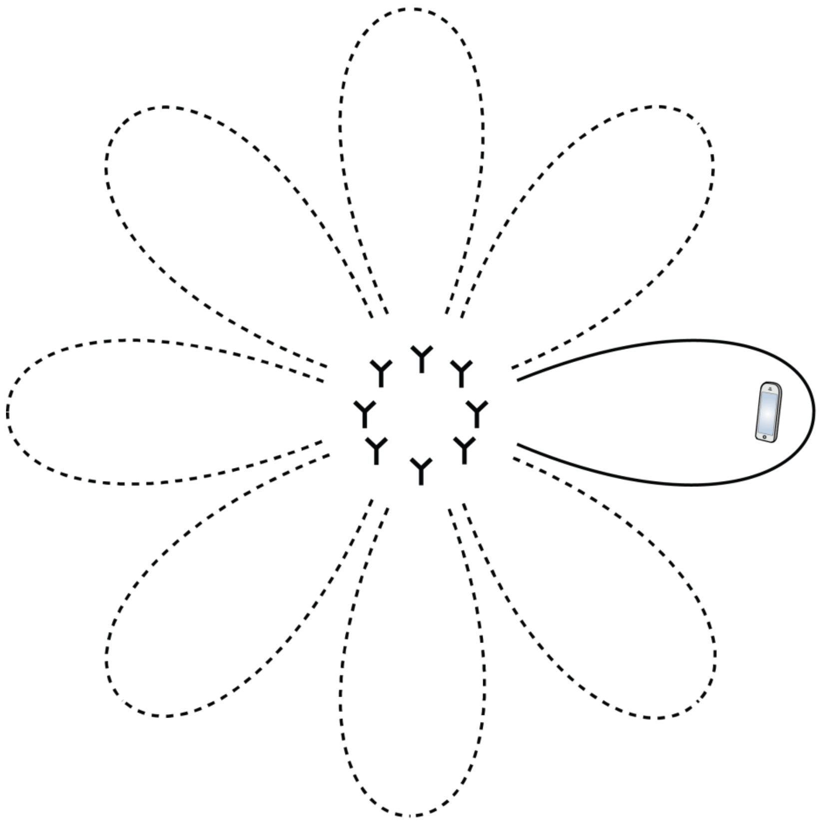


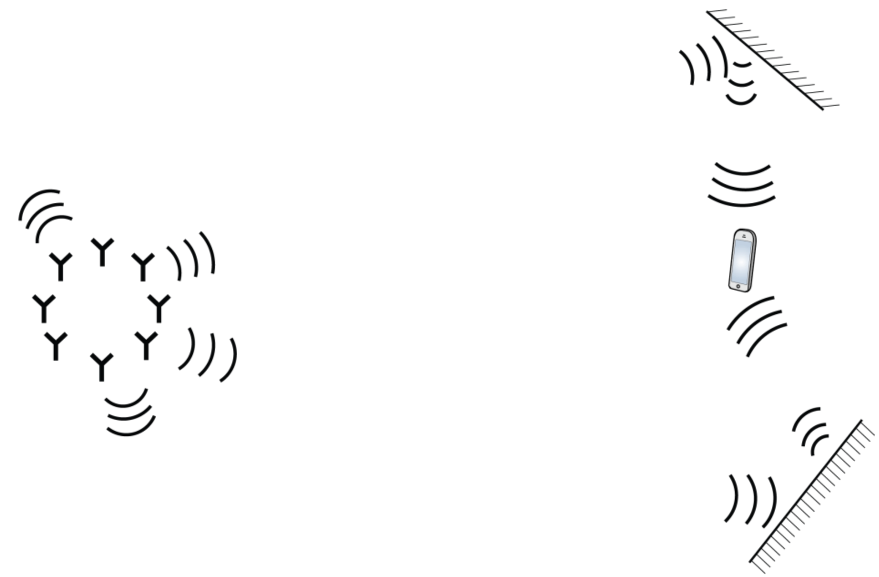
	Classic multi-user MIMO	Massive MIMO (Canonical)
Antennas M, users K	$M \approx K$	$M \gg K$
Signal processing	Non-linear is preferred	Linear is near optimal
Duplexing mode	Designed for TDD and FDD	Designed for TDD w. reciprocity
Instantaneous channel	Known at BS and user	Only needed at BS (hardening)
Channel quality	Affected by frequency-selective and fast fading	Almost no channel quality variations (hardening)
Variations in user load	Scheduling needed if $K > M$	Scheduling seldom needed
Resource allocation	Rapid due to fading	Only on a slow time scale
Cell-edge performance	Only good if BSs cooperate	Improved by array gain of M
BS cooperation	Highly beneficial if rapid	Only long-term coordination

FDD = Frequency-division duplex, *TDD = Time-division duplex*

Downlink MIMO Precoding

Same principle for MIMO detection





Line-of-Sight

Channels characterized by angles

1-2 parameters to estimate per user

Precoding = Angular beamforming

Non-Line-of-Sight

Rich multipath propagation

M parameters to estimate per user

Precoding $\neq$ Angular beamforming

Easy: Codebooks usable (phased array)

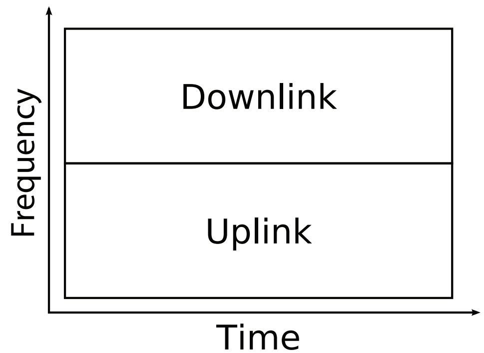
Hard: Requires pilot transmission!

How to Limit the Pilot Overhead?

Frequency-division duplex (FDD)

Downlink: M pilots + K feedback

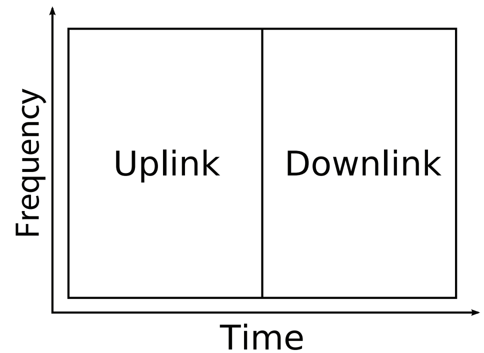
Uplink: K pilots + M feedback

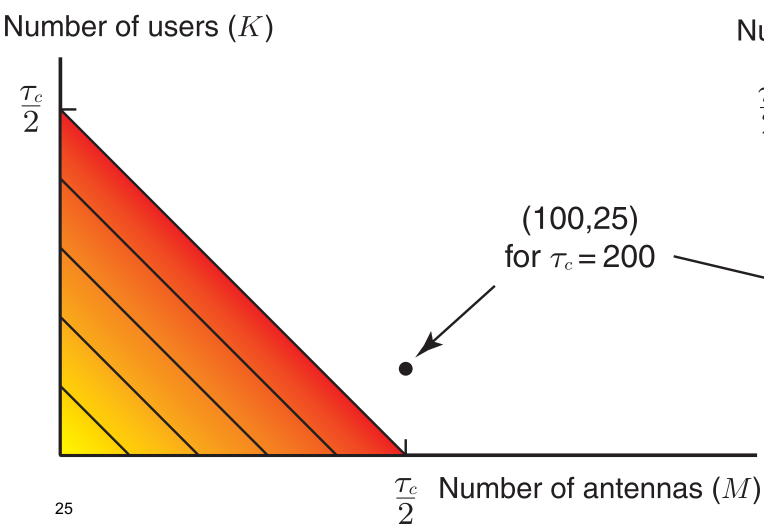


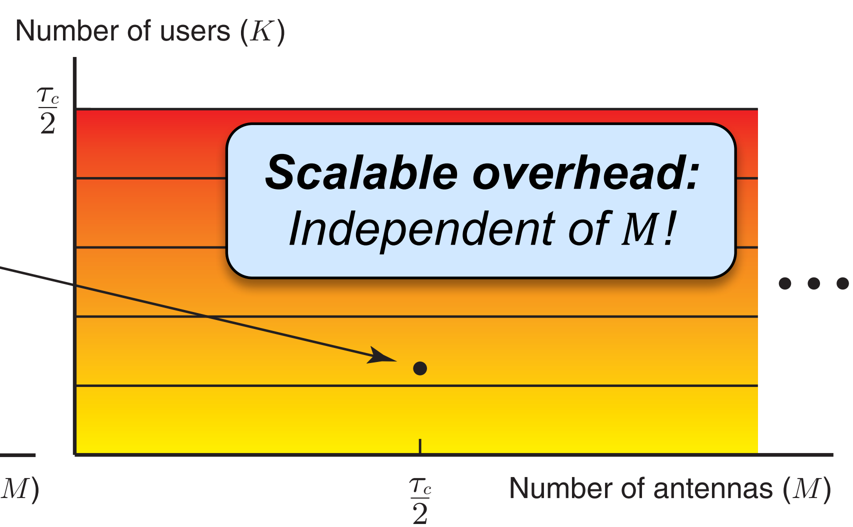
Time-division duplex (TDD)

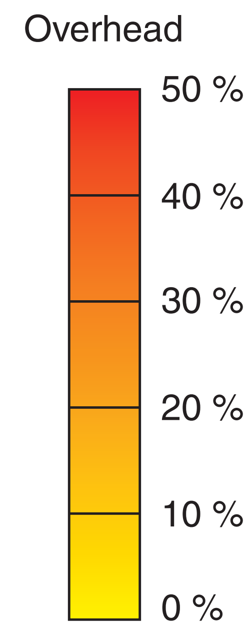
Uplink: K pilots, exploit channel reciprocity

Downlink: K precoded pilots (optional)



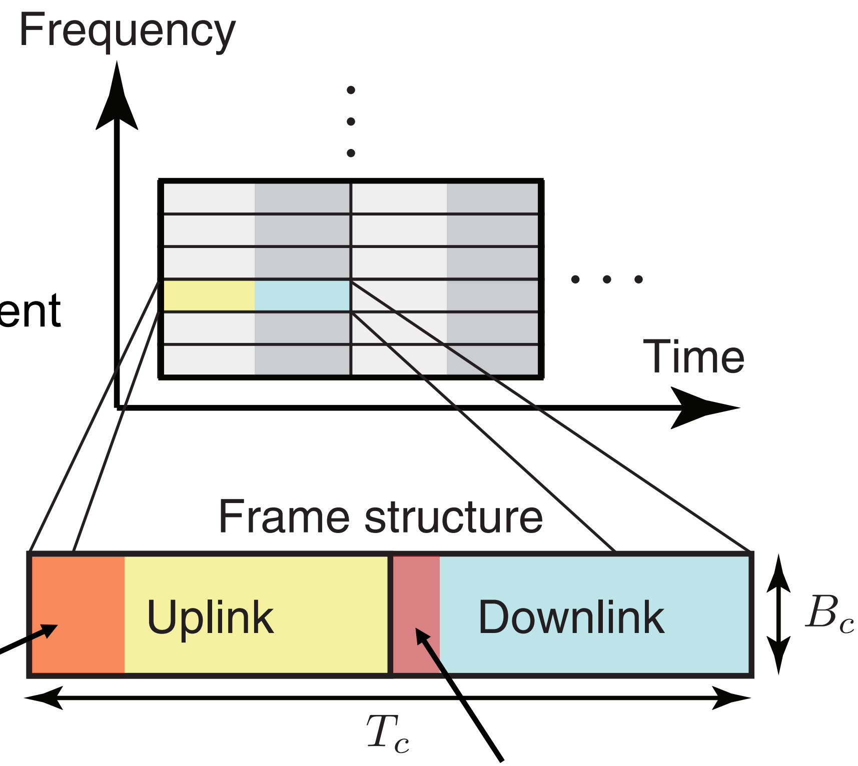






Massive MIMO TDD Protocol

- Coherence Blocks
  - Fixed channel responses
  - Coherence time: T_c s
  - Coherence bandwidth: B_c Hz
  - Depends on mobility and environment
  - Block length: $\tau_c = T_c B_c$ symbols
  - Typically: $\tau_c \in [100,10000]$



***Uplink pilots**:*
τ_p symbols/block
($\tau_p \geq K$)

Optional:
downlink pilots
(recall hardening)

Uplink and/or downlink payload data
$\tau_c - \tau_p$ symbols/block

Signal Processing Scheme

- Capacity-Achieving Non-linear Processing
  - Downlink: Dirty paper coding
  - Uplink: Successive interference cancellation

Why not used in Massive MIMO?

Linear Processing

Bad when $M \approx K$

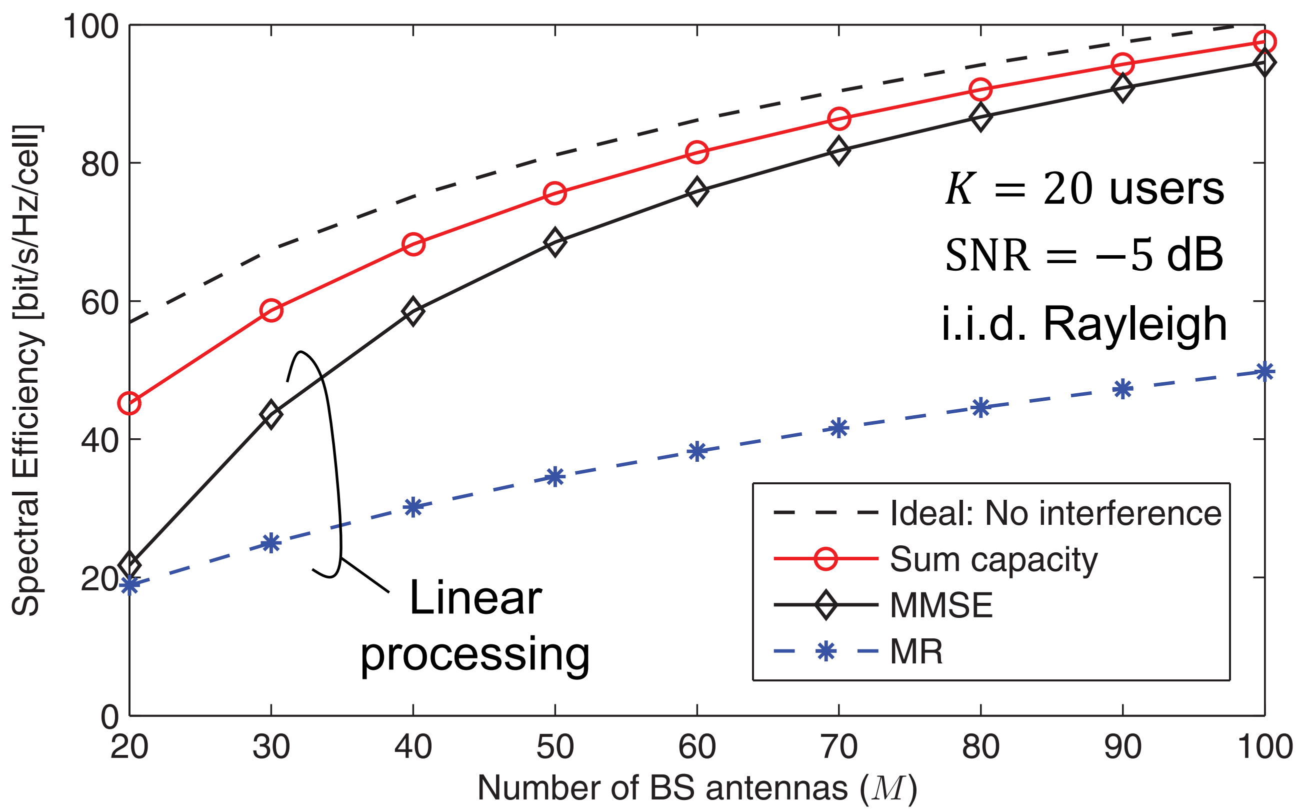
Good when $M/K > 2$

Easier to implement

Scalable Complexity in Massive MIMO

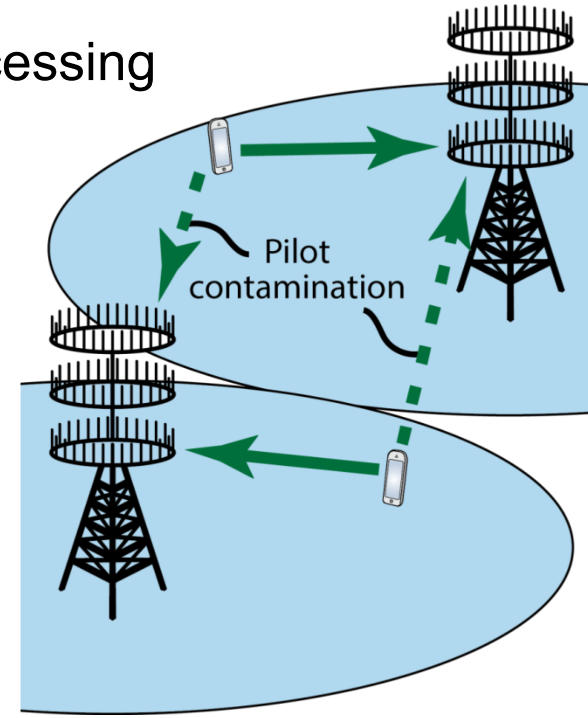
Minimum mean-square error (MMSE): $O(MK^2)$

Maximum ratio (MR): $O(MK)$



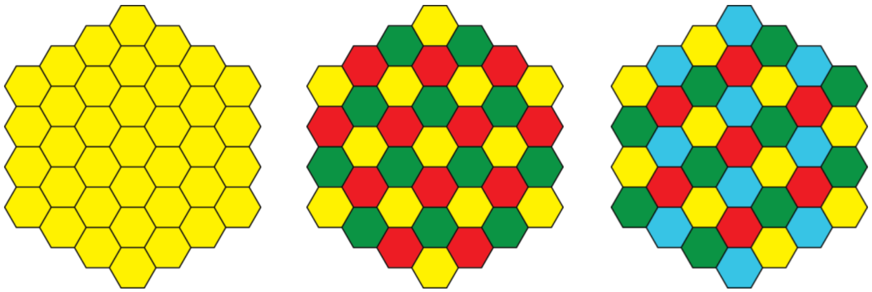
Channel Acquisition in Massive MIMO

- BS Needs Channel Responses for Linear Processing
  - Estimate using $\tau_p \leq \tau_c$ pilot symbols
  - Must reuse pilot sequences in different cells



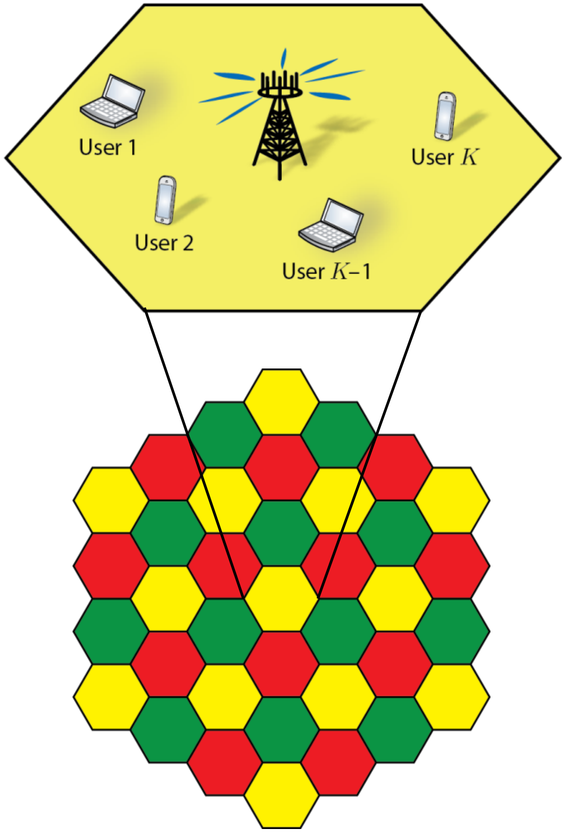
- Called: Pilot Contamination
  - BSs cannot tell some users apart
  - Recall: Noise and interference vanish as $M \to \infty$
  - Not interference between users with same pilot!

- Scalable Solution: Select how often pilots are reused
  - Pilot reuse factor $f \geq 1$
  - Users per cell: $K \leq \tau_p / f$
  - Higher f $\to$ Fewer users per cell,
    but interferers further away



Pilot reuse $f = 1$ ***Pilot reuse*** $f = 3$ ***Pilot reuse*** $f = 4$

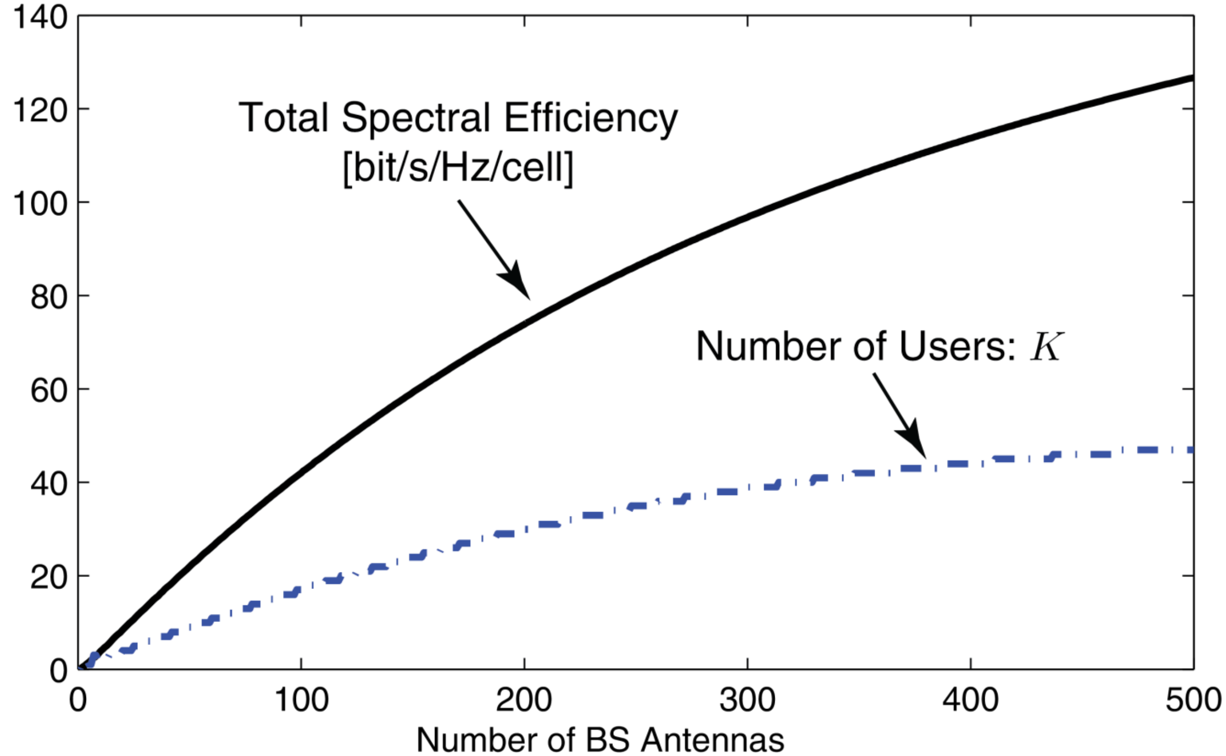
How Many Antennas Are Needed?



Uplink simulation

SNR 5 dB, Rayleigh fading,

ZF detection, $\tau_c = 500$, *pilot reuse* $f = 3$

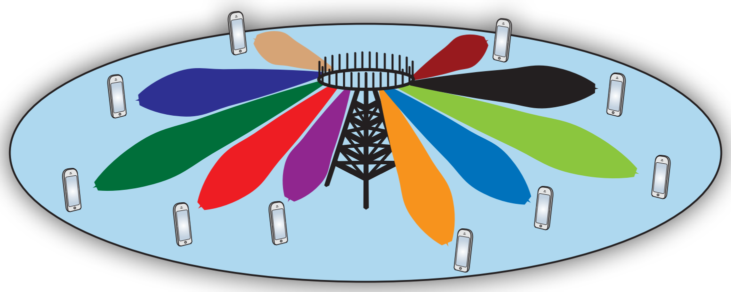


Massive MIMO is an incredibly scalable technology!

PART 1: SUMMARY

Part 1: Summary

- Massive MIMO: The way to increase spectral efficiency in 5G
  - >20x gain over IMT-Advanced are foreseen
  - BSs with many small antennas and transceiver chains
  - Many potential deployment strategies



- Facts to Remember
  - Massive MIMO ≠ Massive size: TV sized panels at cellular frequencies
  - Favorable propagation in most propagation environments
  - Signal processing is simplified, not complicated

- Further Reading
  - Emil Björnson, Erik G. Larsson, Thomas L. Marzetta, "*Massive MIMO: 10 Myths and One Critcal Question*," IEEE Commun. Magazine, Feb. 2016.

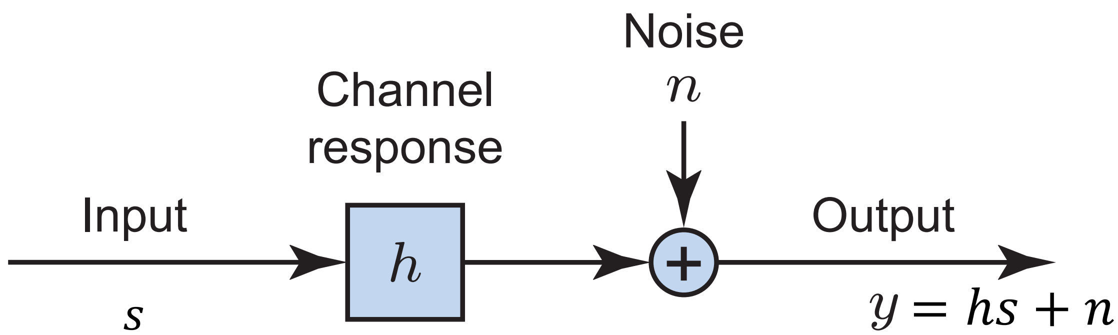
Part 2

System Capacity with Massive MIMO

LINKÖPING UNIVERSITY

BASIC COMMUNICATION THEORETIC RESULTS

Capacity of Memoryless AWGN Channel



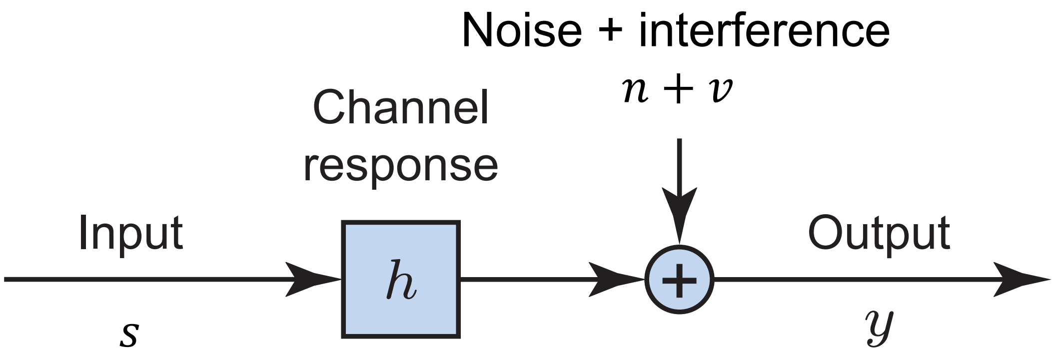
- Assumptions
  - Limited input power: $\mathbb{E}\{|s|^2\} \leq p$
  - Independent noise: $n \sim CN(0, \sigma_{\text{AWGN}}^2)$
  - Channel h is deterministic

- Capacity:

$$C = \log_2\left(1 + \frac{p|h|^2}{\sigma_{\text{AWGN}}^2}\right) \quad [\text{bit/s/Hz}]$$

Capacity achieved by $s \sim CN(0, p)$

Capacity Lower Bound with Interference



- Assumptions
  - Same as before: $\mathbb{E}\{|s|^2\} \leq p, \quad n \sim CN(0, \sigma^2_{\mathrm{AWGN}}), \; h$ is deterministic
  - Uncorrelated interference: $\mathbb{E}\{s^* v\} = 0$
  - Interference power: $\mathbb{E}\{|v|^2\} = p_v$

- Capacity lower bound:

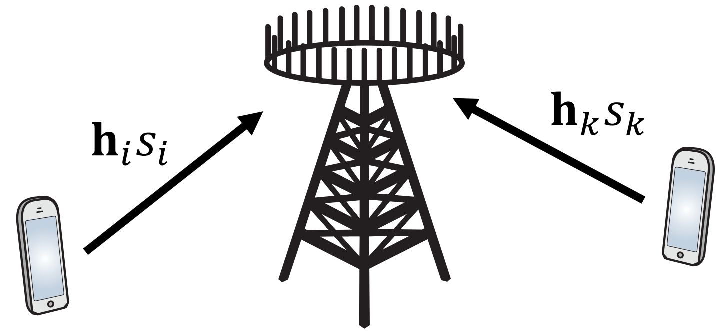
$$C \geq \log_2\left(1 + \frac{p|h|^2}{p_v + \sigma^2_{\mathrm{AWGN}}}\right) \quad [\mathrm{bit/s/Hz}]$$

Called achievable ***spectral efficiency***

Lower bound achieved by $s \sim CN(0, p)$ and treating interference as Gaussian noise

SINGLE-CELL MASSIVE MIMO PERFORMANCE

Uplink Spectral Efficiency (1)



- Single-Cell Uplink
  - K users: Signals s_k with $\mathbb{E}\{|s_k|^2\} = p_k$
  - Channels: $\mathbf{h}_k = [h_{k1} \dots h_{kM}]^T$
  - Noise: $\mathbf{n} \sim CN(\mathbf{0}, \sigma^2 \mathbf{I}_M)$
  - Received signal at BS: $\mathbf{y} = \sum_{i=1}^{K} \mathbf{h}_i \, s_i + \mathbf{n}$

- Select a detection vector $\mathbf{v}_k$ for User k: $(k = 1, \dots, K)$

$$\mathbf{v}_k^H \mathbf{y} = \sum_{i=1}^{K} \mathbf{v}_k^H \mathbf{h}_i \, s_i + \mathbf{v}_k^H \mathbf{n}$$

$$= \underbrace{\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}}_{h\text{: Constant gain}} s_k + \underbrace{\left(\mathbf{v}_k^H \mathbf{h}_k - \mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}\right) s_k}_{v_1\text{: Uncorrelated deviation}} + \underbrace{\sum_{i \neq k} \mathbf{v}_k^H \mathbf{h}_i s_i}_{v_2\text{: Interference}} + \underbrace{\mathbf{v}_k^H \mathbf{n}}_{n\text{: Noise}}$$

$$\boxed{C_k \geq \log_2\left(1 + \frac{p_k |h|^2}{p_{v_1} + p_{v_2} + \sigma^2_{\text{AWGN}}}\right) = \log_2\left(1 + \frac{p_k \left|\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}\right|^2}{\sum_{i=1}^{K} p_i \mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} - p_k \left|\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}\right|^2 + \sigma^2 \mathbb{E}\{\|\mathbf{v}_k\|^2\}}\right)}$$

Uplink Spectral Efficiency (2)

- Uplink Spectral Efficiency [bit/s/Hz] of User k:

$$\overline{\mathrm{SE}}_k = \log_2\left(1 + \frac{p_k \left|\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}\right|^2}{\sum_{i=1}^{K} p_i \mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} - p_k \left|\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}\right|^2 + \sigma^2 \mathbb{E}\{\|\mathbf{v}_k\|^2\}}\right)$$

  - Achieved by Gaussian codebook: $s_k \sim CN(0, p_k)$
  - Resembles an AWGN channel
  - Depends on channel distribution: $\mathbf{h}_1, \dots, \mathbf{h}_K$
  - Depends on the choice of $\mathbf{v}_k$ – Ideally based on the channel

We need to estimate $\mathbf{h}_1, \dots, \mathbf{h}_K$ at BS to select $\mathbf{v}_k$ properly

Uplink Channel Estimation (1)

- Recall: τ_p pilot symbols per coherence block ($\tau_p \geq K$)
  - Use K orthogonal τ_p-length pilot sequences: $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_K$
  - Properties: $\boldsymbol{\phi}_i^H \boldsymbol{\phi}_j = 0$ if $i \neq j$
    $\|\boldsymbol{\phi}_i\|^2 = \tau_p$

- Uplink pilot transmission
  - User k transmits pilot $\sqrt{p_k}\boldsymbol{\phi}_k^H$
  - Received $M \times \tau_p$ matrix:

Independent elements $CN(0, \sigma^2)$

$$\mathbf{Y}_{\text{pilot}} = \sum_{i=1}^{K} \sqrt{p_i}\mathbf{h}_i \boldsymbol{\phi}_i^H + \mathbf{N}$$

  - Extract signal from User k:

$$\mathbf{y}_{\text{pilot},k} = \mathbf{Y}_{\text{pilot}}\boldsymbol{\phi}_k / \|\boldsymbol{\phi}_k\| = \sqrt{p_k \tau_p}\mathbf{h}_k + \overbrace{\mathbf{N}\boldsymbol{\phi}_k / \|\boldsymbol{\phi}_k\|}^{\mathbf{n}_k \sim CN(\mathbf{0}, \sigma^2 \mathbf{I}_M)}$$

***Processing gain**: Power $p_k \tau_p$ instead of p_k*

Uplink Channel Estimation (2)

- Estimation of $\mathbf{h}_k$ based on:

$$\mathbf{y}_{\text{pilot},k} = \sqrt{p_k \tau_p}\mathbf{h}_k + \mathbf{n}_k$$

  - Assume $\mathbf{h}_k \sim CN(\mathbf{0}, \beta_k \mathbf{I}_M)$

MMSE estimator and error covariance:

$$\hat{\mathbf{h}}_k = \frac{\sqrt{p_k \tau_p}\beta_k}{p_k \tau_p \beta_k + \sigma^2} \mathbf{y}_{\text{pilot},k}$$

$$\mathbb{E}\left\{(\boldsymbol{h}_k - \hat{\mathbf{h}}_k)(\mathbf{h}_k - \hat{\mathbf{h}}_k)^H\right\} = \left(\beta_k - \underbrace{\frac{p_k \tau_p \beta_k^2}{p_k \tau_p \beta_k + \sigma^2}}_{= \gamma_k}\right)\mathbf{I}_M$$

- Independent estimate and error:
  - $\mathbf{e}_k = \mathbf{h}_k - \hat{\mathbf{h}}_k \sim CN(\mathbf{0}, (\beta_k - \gamma_k)\mathbf{I}_M)$
  - $\hat{\mathbf{h}}_k \sim CN(\mathbf{0}, \gamma_k \mathbf{I}_M)$

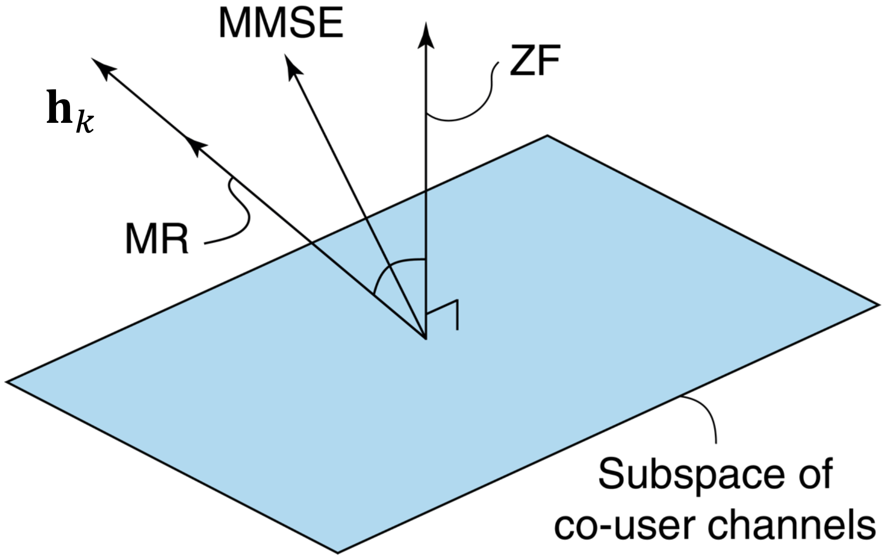
***Note**:*

$\gamma_k \to \beta_k$ *as* $p_k \tau_p \to \infty$

Asymptotically error free

Combining Schemes for Uplink Detection

- How to Choose Detection Vectors?
  - Based on $\widehat{\mathbf{H}} = [\hat{\mathbf{h}}_1 \; \ldots \; \hat{\mathbf{h}}_K]$



- Different Criteria
  - Maximize channel gain $\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}$:

$$\mathbf{v}_k = \hat{\mathbf{h}}_k \longleftarrow \textit{Maximum ratio (MR)}$$

  - Minimize interference terms $\mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_i\right|^2\right\}$:

$$[\mathbf{v}_1 \; \ldots \; \mathbf{v}_K] = \widehat{\mathbf{H}}\left(\widehat{\mathbf{H}}^H \widehat{\mathbf{H}}\right)^{-1} \longleftarrow \textit{Zero-forcing (ZF), if } M \geq K$$

  - Balance channel gain against interference:

$$[\mathbf{v}_1 \; \ldots \; \mathbf{v}_K] = \widehat{\mathbf{H}}\left(\widehat{\mathbf{H}}^H \widehat{\mathbf{H}} + \left(\sum_i (\beta_i - \gamma_i) + \sigma^2\right)\mathbf{I}_M\right)^{-1} \longleftarrow \textit{MMSE}$$

Spectral Efficiency: Maximum Ratio Combining

- Uplink Spectral Efficiency of User k with $\mathbf{v}_k = \hat{\mathbf{h}}_k$:

$$\overline{\mathrm{SE}}_k = \log_2\left(1 + \frac{p_k \left|\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}\right|^2}{\sum_{i=1}^{K} p_i \mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} - p_k \left|\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\}\right|^2 + \sigma^2 \mathbb{E}\{\|\mathbf{v}_k\|^2\}}\right)$$

  - $\mathbb{E}\{\|\mathbf{v}_k\|^2\} = \mathbb{E}\{\|\hat{\mathbf{h}}_k\|^2\} = \mathrm{tr}(\gamma_k \mathbf{I}_M) = \gamma_k M$
  - $\mathbb{E}\{\mathbf{v}_k^H \mathbf{h}_k\} = \mathbb{E}\{\hat{\mathbf{h}}_k^H (\hat{\mathbf{h}}_k + \mathbf{e}_k)\} = \mathbb{E}\left\{\left\|\hat{\mathbf{h}}_k\right\|^2\right\} = \gamma_k M$
  - $\mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} = \mathbb{E}\left\{\left\|\hat{\mathbf{h}}_k\right\|^2\right\} \beta_i = \beta_i \gamma_k M$ if $i \neq k$
  - $\mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_k\right|^2\right\} = \mathbb{E}\left\{\left\|\hat{\mathbf{h}}_k\right\|^4\right\} + \mathbb{E}\left\{\left|\hat{\mathbf{h}}_k^H \mathbf{e}_k\right|^2\right\} = \gamma_k^2 (M + M^2) + \gamma_k (\beta_k - \gamma_k) M$
    $= \gamma_k^2 M^2 + \gamma_k \beta_k M$

$$\overline{\mathrm{SE}}_k^{\mathrm{MR}} = \log_2\left(1 + \frac{p_k \gamma_k^2 M^2}{\sum_{i=1}^{K} p_i \beta_i \gamma_k M + \sigma^2 \gamma_k M}\right) = \log_2\left(1 + \frac{p_k \gamma_k M}{\sum_{i=1}^{K} p_i \beta_i + \sigma^2}\right)$$

Spectral Efficiency: Zero-Forcing Combining

- Uplink Spectral Efficiency of User k with $[\mathbf{v}_1 \ \dots \ \mathbf{v}_K] = \widehat{\mathbf{H}}\left(\widehat{\mathbf{H}}^H\widehat{\mathbf{H}}\right)^{-1}$:

$$\overline{\mathrm{SE}}_k = \log_2\left(1 + \frac{p_k\left|\mathbb{E}\{\mathbf{v}_k^H\mathbf{h}_k\}\right|^2}{\sum_{i=1}^K p_i\mathbb{E}\left\{\left|\mathbf{v}_k^H\mathbf{h}_i\right|^2\right\} - p_k\left|\mathbb{E}\{\mathbf{v}_k^H\mathbf{h}_k\}\right|^2 + \sigma^2\mathbb{E}\{\|\mathbf{v}_k\|^2\}}\right)$$

  - $\mathbb{E}\{\|\mathbf{v}_k\|^2\} = \mathbb{E}\left\{\left[\left(\widehat{\mathbf{H}}^H\widehat{\mathbf{H}}\right)^{-1}\right]_{k,k}\right\} = [\text{Wishart matrix properties}] = \frac{1}{\gamma_k(M-K)}$
  - $\mathbb{E}\left\{\mathbf{v}_k^H\mathbf{h}_k\right\} = \mathbb{E}\{1\} + \mathbb{E}\left\{\mathbf{v}_k^H\right\}\mathbb{E}\{\mathbf{e}_k\} = 1$
  - $\mathbb{E}\left\{\left|\mathbf{v}_k^H\mathbf{h}_i\right|^2\right\} = \mathbb{E}\left\{\left|\mathbf{v}_k^H\mathbf{e}_i\right|^2\right\} = \mathbb{E}\{\|\mathbf{v}_k\|^2\}(\beta_i - \gamma_i) = \frac{\beta_i-\gamma_i}{\gamma_k(M-K)}$ if $i \neq k$
  - $\mathbb{E}\left\{\left|\mathbf{v}_k^H\mathbf{h}_k\right|^2\right\} = \mathbb{E}\{|1|^2\} + \mathbb{E}\left\{\left|\mathbf{v}_k^H\mathbf{e}_k\right|^2\right\} = 1 + \frac{\beta_k-\gamma_k}{\gamma_k(M-K)}$

$$\overline{\mathrm{SE}}_k^{\mathrm{ZF}} = [\text{Multiply terms by } \gamma_k(M-K)] = \log_2\left(1 + \frac{p_k\gamma_k(M-K)}{\sum_{i=1}^K p_i(\beta_i - \gamma_i) + \sigma^2}\right)$$

Uplink Spectral Efficiency: Interpretation

Uplink Spectral Efficiency of User k:

Array gain (coherent combining)

Estimation quality
$0 \leq \gamma_k \leq \beta_k$

$$\overline{\mathrm{SE}}_k = \log_2\left(1 + \frac{p_k G \gamma_k}{\underbrace{\sum_{i=1}^{K} p_i z_i}_{\text{Non-coherent interference}} + \sigma^2}\right)$$

Noise power

Non-coherent interference

where

$$G = \begin{cases} M & \text{with MR} \\ M - K & \text{with ZF} \end{cases} \qquad z_i = \begin{cases} \beta_i & \text{with MR} \\ \beta_i - \gamma_i & \text{with ZF} \end{cases}$$

***MR**:*

Maximal array gain

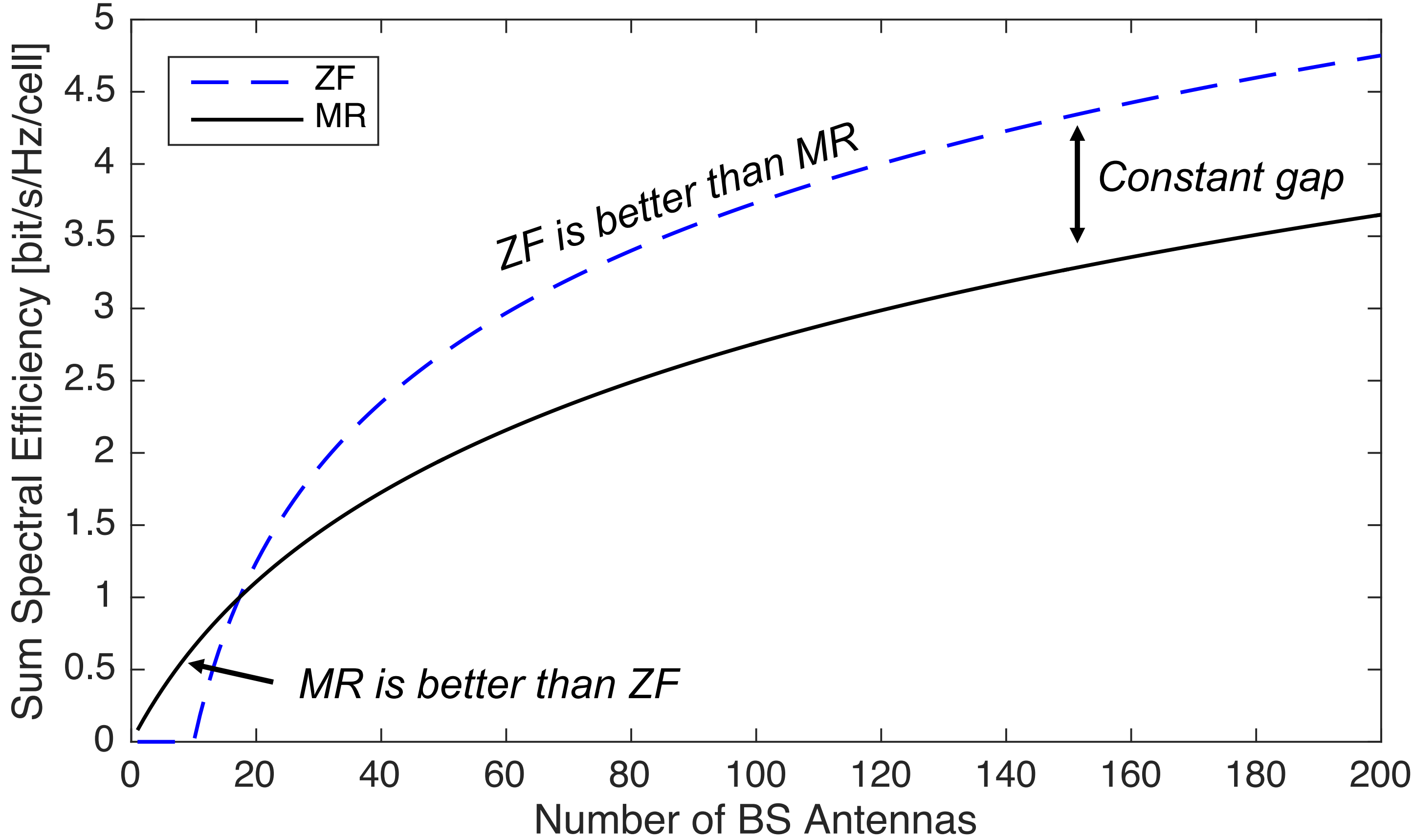
Full non-coherent interference

***ZF**:*

Reduced array gain

Reduced non-coherent interference

Simulation: Single-Cell Scenario



Parameters: $K = 10,\ \tau_p = 10,\ \mathrm{SNR} = \frac{p\beta}{\sigma^2} = -5$ dB

Downlink Spectral Efficiency (1)

- Channel reciprocity for User k
  - Uplink channel: $\mathbf{h}_k$
  - Downlink channel: $\mathbf{h}_k^H$ (with added conjugate for simplicity)

- Received signal at User k:

$$y_k = \sum_{i=1}^{K} \mathbf{h}_k^H \mathbf{w}_i \varsigma_i + n_k$$

  - Precoding vector $\mathbf{w}_i$ for User i (with $\mathbb{E}\{\|\mathbf{w}_i\|^2\} = 1$)
  - Signal $\varsigma_i \sim CN(0, q_i)$ for User i (power q_i)
  - Noise at User k: $n_k \sim CN(0, \sigma^2)$

Downlink Spectral Efficiency (2)

Channel State Information

No downlink pilots: Only statistical CSI

- Received signal at User k:

$$y_k = \sum_{i=1}^{K} \mathbf{h}_k^H \mathbf{w}_i \varsigma_i + n_k$$

$$= \underbrace{\mathbb{E}\{\mathbf{h}_k^H \mathbf{w}_k\}}_{h\text{: Constant gain}} \varsigma_k + \underbrace{\left(\mathbf{h}_k^H \mathbf{w}_k - \mathbb{E}\{\mathbf{h}_k^H \mathbf{w}_k\}\right) \varsigma_k}_{v_1\text{: Uncorrelated deviation}} + \underbrace{\sum_{i=1, i \neq k}^{K} \mathbf{h}_k^H \mathbf{w}_i \varsigma_i}_{v_2\text{: Interference}} + \underbrace{n_k}_{n\text{: Noise}}$$

- Downlink Spectral Efficiency [bit/s/Hz] of User k:

$$\underline{\mathrm{SE}}_k = \log_2\left(1 + \frac{q_k \left|\mathbb{E}\{\mathbf{h}_k^H \mathbf{w}_k\}\right|^2}{\sum_{i=1}^{K} q_i \mathbb{E}\left\{\left|\mathbf{h}_k^H \mathbf{w}_i\right|^2\right\} - q_k \left|\mathbb{E}\{\mathbf{h}_k^H \mathbf{w}_k\}\right|^2 + \sigma^2}\right)$$

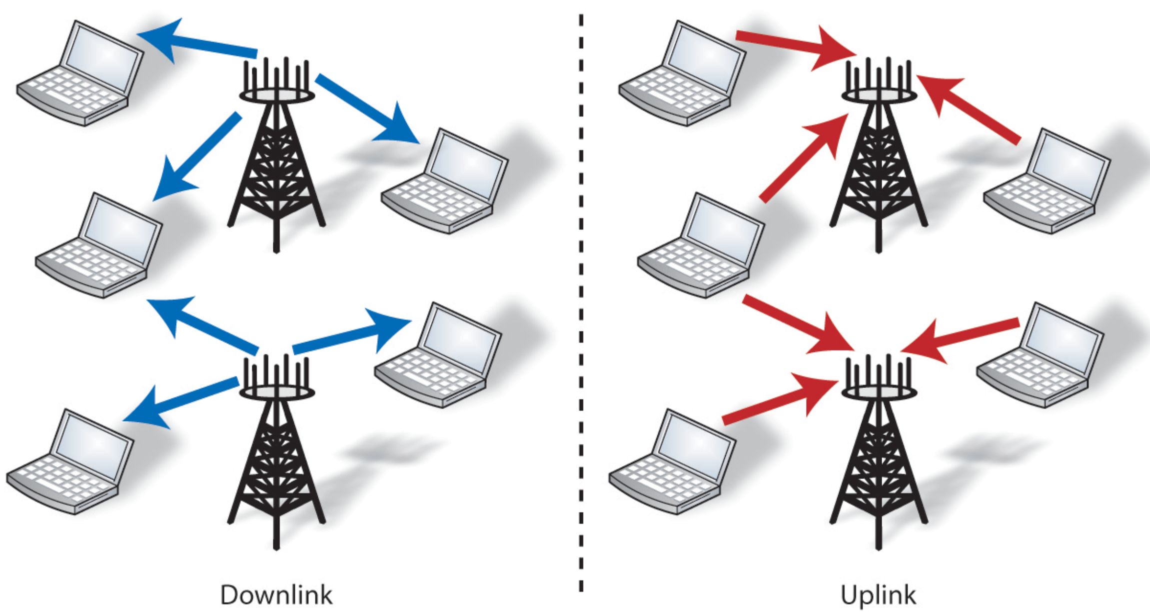
What are the Good Precoding Vectors?

Uplink-Downlink Duality Theorem

The uplink SEs are achievable in the downlink using same sum power.

Use $\mathbf{w}_k = \mathbf{v}_k / \sqrt{\mathbb{E}\{\|\mathbf{v}_k\|^2\}}$, *but different power allocation*

- Consequence
  - Use detection vectors as precoding vectors
  - Use whatever power allocation



Downlink Spectral Efficiency (3)

- Downlink Spectral Efficiency [bit/s/Hz] of User k:

$$\underline{\mathrm{SE}}_k = \log_2\left(1 + \frac{q_k \left|\mathbb{E}\{\mathbf{h}_k^H \mathbf{w}_k\}\right|^2}{\sum_{i=1}^{K} q_i \mathbb{E}\left\{\left|\mathbf{h}_k^H \mathbf{w}_i\right|^2\right\} - q_k \left|\mathbb{E}\{\mathbf{h}_k^H \mathbf{w}_k\}\right|^2 + \sigma^2}\right)$$

  - Set $\mathbf{w}_k = \mathbf{v}_k / \sqrt{\mathbb{E}\{\|\mathbf{v}_k\|^2\}}$ and get same expectations as in uplink

  - Result:

Array gain (coherent precoding)

Estimation quality
$0 \leq \gamma_k \leq \beta_k$

$$\underline{\mathrm{SE}}_k = \log_2\left(1 + \frac{q_k G \gamma_k}{\underbrace{\sum_{i=1}^{K} q_i z_k}_{\text{Non-coherent interference}} + \sigma^2}\right)$$

Noise power

where

$$G = \begin{cases} M & \text{with MR} \\ M - K & \text{with ZF} \end{cases} \qquad z_i = \begin{cases} \beta_i & \text{with MR} \\ \beta_i - \gamma_i & \text{with ZF} \end{cases}$$

Example of Uplink-Downlink Duality (1)

- Consider $K = 2$ users, MR or ZF processing
- Make $R = \overline{\mathrm{SE}}_k = \underline{\mathrm{SE}}_k$

User 1: $$R = \log_2\left(1 + \frac{p_1 G \gamma_1}{p_1 z_1 + p_2 z_2 + \sigma^2}\right) = \log_2\left(1 + \frac{q_1 G \gamma_1}{q_1 z_1 + q_2 z_1 + \sigma^2}\right)$$

User 2: $$R = \log_2\left(1 + \frac{p_2 G \gamma_2}{p_1 z_1 + p_2 z_2 + \sigma^2}\right) = \log_2\left(1 + \frac{q_2 G \gamma_2}{q_1 z_2 + q_2 z_2 + \sigma^2}\right)$$

- Similar equations, but one key difference:
  - Uplink: Interference comes through different user channels
  - Downlink: Interference comes from the BS

Writing $R = \overline{\mathrm{SE}}_1$ as a linear equation

$$2^R - 1 = \frac{p_1 G \gamma_1}{p_1 z_1 + p_2 z_2 + \sigma^2} \quad \leftrightarrow \quad \frac{p_1 G \gamma_1}{2^R - 1} = p_1 z_1 + p_2 z_2 + \sigma^2$$

Example of Uplink-Downlink Duality (2)

- Uplink equations:

$$\begin{bmatrix} \frac{G\gamma_1}{2^{\bar{\mathrm{R}}}-1} - z_1 & -z_2 \\ -z_1 & \frac{G\gamma_2}{2^{\bar{\mathrm{R}}}-1} - z_2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- Downlink equations:

$$\begin{bmatrix} \frac{G\gamma_1}{2^{\bar{\mathrm{R}}}-1} - z_1 & -{\color{red}z_1} \\ -{\color{red}z_2} & \frac{G\gamma_2}{2^{\bar{\mathrm{R}}}-1} - z_2 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

  - Just a transpose is different between the matrices
  - If one equation is solvable → Both equations are solvable

Example of Uplink-Downlink Duality (3)

- Define

$$\boldsymbol{B} = \begin{bmatrix} \frac{G\gamma_1}{2^{\bar{\mathrm{R}}}-1} - z_1 & -z_2 \\ -z_1 & \frac{G\gamma_2}{2^{\bar{\mathrm{R}}}-1} - z_2 \end{bmatrix}$$

- Transmit powers:

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \sigma^2 \boldsymbol{B}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \sigma^2 \boldsymbol{B}^{-T} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \boldsymbol{B}^{-T}\boldsymbol{B} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

- Sum transmit power:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \boldsymbol{B}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Same sum power, but distributed differently over users

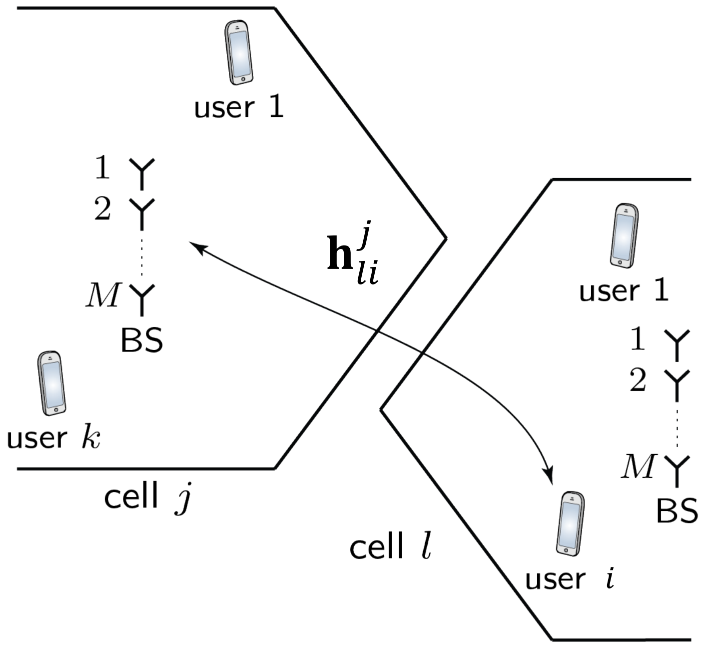
MULTI-CELL MASSIVE MIMO PERFORMANCE

Multi-Cell Channel Model

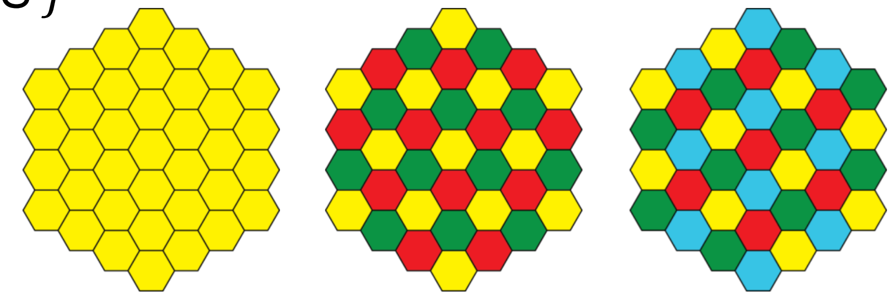
- System Model
  - L cells, K users per cell
  - Channel from BS j to user i in cell l

$$\mathbf{h}_{li}^{j} \sim \mathcal{CN}\left(\mathbf{0}, \beta_{li}^{j} \mathbf{I}_M\right)$$

  - Uplink transmit power: p_{li}
  - Downlink transmit power: q_{li}



- Recall: Pilot reuse factor $f \geq 1$
  - Users per cell: $K \leq \tau_p / f$
  - $\mathcal{P}_j(f)$: Cells with same pilots as BS j
  - Higher f $\rightarrow$ Fewer users per cell, but fewer interferers in $\mathcal{P}_j$



Pilot reuse $f = 1$ ***Pilot reuse*** $f = 3$ ***Pilot reuse*** $f = 4$

Multi-Cell Channel Estimation

- Estimation of $\mathbf{h}_{li}^{l}$ based on:

Noise: $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$

$$\mathbf{y}_{\text{pilot},li} = \sqrt{p_{li}\tau_p}\mathbf{h}_{li}^{l} + \underbrace{\sum_{j \in \mathcal{P}_l(f) \setminus \{l\}} \sqrt{p_{ji}\tau_p}\mathbf{h}_{ji}^{l}}_{\text{Interference from pilot reuse}} + \mathbf{n}_{li}$$

Desired channel

Interference from pilot reuse

- MMSE estimate: $\hat{\mathbf{h}}_{li}^{l} = \frac{\sqrt{p_{li}\tau_p}\beta_{li}^{l}}{\sum_{l' \in \mathcal{P}_l(f)} p_{l'i}\tau_p\beta_{l'i}^{l} + \sigma^2}\mathbf{y}_{\text{pilot},li} \sim \mathcal{CN}\left(\mathbf{0}, \gamma_{li}^{l}\mathbf{I}_M\right)$
- Estimation error: $\mathbf{e}_{li}^{l} \sim \mathcal{CN}\left(\mathbf{0}, \left(\beta_{li}^{l} - \gamma_{li}^{l}\right)\mathbf{I}_M\right)$
- Estimation quality: $\gamma_{li}^{j} = \frac{p_{li}\tau_p\left(\beta_{li}^{j}\right)^2}{\sum_{l' \in \mathcal{P}_l(f)} p_{l'i}\tau_p\beta_{l'i}^{l} + \sigma^2}$ (Note: $0 \leq \gamma_{li}^{j} \leq \beta_{li}^{j}$)

Pilot Contamination *(Parallel channel estimates)*

$$\hat{\mathbf{h}}_{ji}^{l} = \frac{\sqrt{p_{ji}}\beta_{ji}^{l}}{\sqrt{p_{li}}\beta_{li}^{l}}\hat{\mathbf{h}}_{li}^{l} \quad \text{for } j \in \mathcal{P}_l(f)$$

Uplink Multi-Cell Spectral Efficiency (1)

- Received signal at BS l:

$$\mathbf{y}_l = \underbrace{\sum_{i=1}^{K} \mathbf{h}_{li}^{l} s_{li}}_{\text{Intra-cell signals}} + \overbrace{\sum_{j=1,\, j\neq l}^{L} \sum_{i=1}^{K} \mathbf{h}_{ji}^{l} s_{ji}}^{\text{Inter-cell interference}} + \mathbf{n}_l$$

Noise: $CN(\mathbf{0}, \sigma^2 \mathbf{I}_M)$

- Select a detection vector $\mathbf{v}_{lk}$ for User k in cell l:

$$\mathbf{v}_{lk}^{H}\mathbf{y}_l = \sum_{i=1}^{K} \mathbf{v}_{lk}^{H}\mathbf{h}_{li}^{l} s_{li} + \sum_{j=1,\, j\neq l}^{L} \sum_{i=1}^{K} \mathbf{v}_{lk}^{H}\mathbf{h}_{ji}^{l} s_{ji} + \mathbf{v}_{lk}^{H}\mathbf{n}_l$$

$$= \underbrace{\mathbb{E}\{\mathbf{v}_{lk}^{H}\mathbf{h}_{lk}^{l}\}s_{lk}}_{\text{Constant gain}} + \underbrace{\left(\mathbf{v}_{lk}^{H}\mathbf{h}_{lk}^{l} - \mathbb{E}\{\mathbf{v}_{lk}^{H}\mathbf{h}_{lk}^{l}\}\right)s_{lk}}_{\text{Uncorrelated deviation}} + \underbrace{\sum_{i\neq k} \mathbf{v}_{lk}^{H}\mathbf{h}_{li}^{l} s_{li}}_{\text{Intra-cell interf.}} + \underbrace{\sum_{j=1,\, j\neq l}^{L} \sum_{i=1}^{K} \mathbf{v}_{lk}^{H}\mathbf{h}_{ji}^{l} s_{ji}}_{\text{Inter-cell interf.}} + \underbrace{\mathbf{v}_{lk}^{H}\mathbf{n}_l}_{\text{Noise}}$$

$$\boxed{C_{lk} \geq \log_2\left(1 + \frac{p_{lk}\left|\mathbb{E}\{\mathbf{v}_{lk}^{H}\mathbf{h}_{lk}^{l}\}\right|^2}{\sum_{j=1}^{L}\sum_{i=1}^{K} p_{ji}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{H}\mathbf{h}_{ji}^{l}\right|^2\right\} - p_{lk}\left|\mathbb{E}\{\mathbf{v}_{lk}^{H}\mathbf{h}_{lk}^{l}\}\right|^2 + \sigma^2\mathbb{E}\{\|\mathbf{v}_{lk}\|^2\}}\right)}$$

Uplink Multi-Cell Spectral Efficiency (2)

- Spectral Efficiency of User k in cell l:

$$\overline{\mathrm{SE}}_{lk} = \log_2\left(1 + \frac{p_{lk}\left|\mathbb{E}\{\mathbf{v}_{lk}^H\mathbf{h}_{lk}^l\}\right|^2}{\sum_{j=1}^{L}\sum_{i=1}^{K} p_{ji}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^H\mathbf{h}_{ji}^l\right|^2\right\} - p_{lk}\left|\mathbb{E}\{\mathbf{v}_{lk}^H\mathbf{h}_{lk}^l\}\right|^2 + \sigma^2\mathbb{E}\{\|\mathbf{v}_{lk}\|^2\}}\right)$$

  - Closed-form expression with MR or ZF combining:

Array gain (coherent)

$$\overline{\mathrm{SE}}_{lk} = \log_2\left(1 + \frac{p_{lk}G\gamma_{lk}^l}{\underbrace{\sum_{j=1}^{L}\sum_{i=1}^{K} p_{ji}z_{ji}^l}_{\text{Non-coherent interf.}} + \underbrace{\sum_{j\in\mathcal{P}_l(f)\setminus\{l\}} p_{jk}\,G\gamma_{jk}^l}_{\text{Coherent interf.}} + \sigma^2}\right)$$

Noise power

where

$$G = \begin{cases} M & \text{with MR} \\ M-K & \text{with ZF} \end{cases} \qquad z_{ji}^l = \begin{cases} \beta_{ji}^l & \text{with MR} \\ \beta_{ji}^l - \gamma_{ji}^l & \text{with ZF} \end{cases}$$

Downlink Multi-Cell Spectral Efficiency

Indices are switched just as in single-cell

- Spectral Efficiency of User k in cell l:

$$\underline{\mathrm{SE}}_{lk} = \log_2\left(1 + \frac{q_{lk}\left|\mathbb{E}\{\mathbf{w}_{lk}^H \mathbf{h}_{lk}^l\}\right|^2}{\sum_{j=1}^{L}\sum_{i=1}^{K} q_{ji}\mathbb{E}\left\{\left|\mathbf{w}_{ji}^H \mathbf{h}_{lk}^j\right|^2\right\} - q_{lk}\left|\mathbb{E}\{\mathbf{w}_{lk}^H \mathbf{h}_{lk}^l\}\right|^2 + \sigma^2}\right)$$

  - Closed-form expression with MR or ZF precoding:

$$\underline{\mathrm{SE}}_{lk} = \log_2\left(1 + \frac{q_{lk} \overbrace{G\gamma_{lk}^l}^{\textit{Array gain (coherent)}}}{\underbrace{\sum_{j=1}^{L}\sum_{i=1}^{K} q_{ji} z_{lk}^j}_{\textit{Non-coherent interf.}} + \underbrace{\sum_{j\in\mathcal{P}_l(f)\setminus\{l\}} q_{jk}\, G\gamma_{lk}^j}_{\textit{Coherent interf.}} + \underbrace{\sigma^2}_{\textit{Noise power}}}\right)$$

where

$$G = \begin{cases} M & \text{with MR} \\ M-K & \text{with ZF} \end{cases} \qquad z_{lk}^j = \begin{cases} \beta_{lk}^j & \text{with MR} \\ \beta_{lk}^j - \gamma_{lk}^j & \text{with ZF} \end{cases}$$

Net Spectral Efficiency

- Coherence Block:

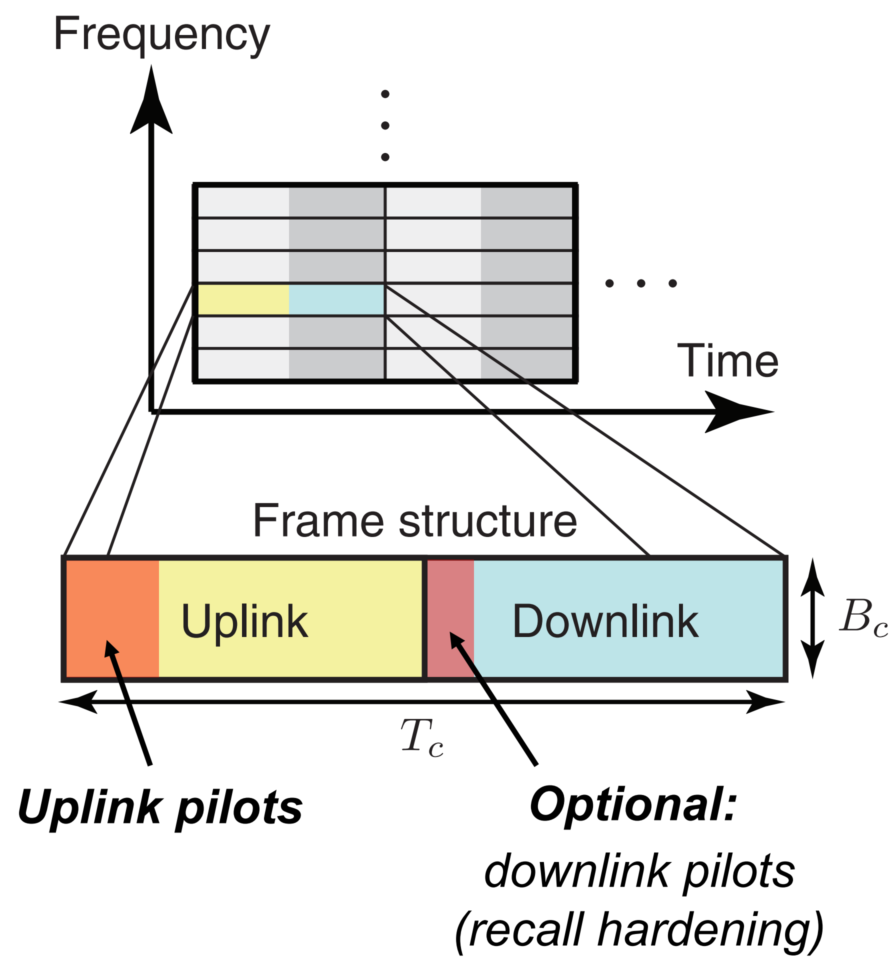
  $$\tau_c = T_c B_c \text{ symbols}$$

- Pilots per Coherence Block:

  $$\tau_p = fK$$

- Payload Transmission
  - $\tau_c - \tau_p$ symbols per block
  - Uplink fraction: ζ^{UL}
  - Downlink fraction: ζ^{DL}

- Net Spectral Efficiency
  - Uplink: $\zeta^{\mathrm{UL}}\left(1-\frac{\tau_p}{\tau_c}\right)\overline{\mathrm{SE}}_{lk}$
  - Downlink: $\zeta^{\mathrm{DL}}\left(1-\frac{\tau_p}{\tau_c}\right)\underline{\mathrm{SE}}_{lk}$



Asymptotic Limits: Many Antennas

- Consider $M \to \infty$ ($G \to \infty$ with MR/ZF)
  - Uplink

$$\overline{\mathrm{SE}}_{lk} = \log_2\left(1 + \frac{\boxed{p_{lk} G \gamma_{lk}^{l}}}{\sum_{j=1}^{L}\sum_{i=1}^{K} p_{ji} z_{ji}^{l} + \sum_{j\in\mathcal{P}_l(f)\setminus\{l\}} \boxed{p_{jk}\, G \gamma_{jk}^{l}} + \sigma^2}\right)$$

$$\to \log_2\left(1 + \frac{p_{lk}\gamma_{lk}^{l}}{\sum_{j\in\mathcal{P}_l(f)\setminus\{l\}} p_{jk}\,\gamma_{jk}^{l}}\right)$$

  - Downlink

$$\underline{\mathrm{SE}}_{lk} = \log_2\left(1 + \frac{\boxed{q_{lk} G \gamma_{lk}^{l}}}{\sum_{j=1}^{L}\sum_{i=1}^{K} q_{ji} z_{lk}^{j} + \sum_{j\in\mathcal{P}_l(f)\setminus\{l\}} \boxed{q_{jk}\, G \gamma_{lk}^{j}} + \sigma^2}\right)$$

$$\to \log_2\left(1 + \frac{q_{lk}\gamma_{lk}^{l}}{\sum_{j\in\mathcal{P}_l(f)\setminus\{l\}} q_{jk}\,\gamma_{lk}^{j}}\right)$$

*Asymptotic SE depends on signal and coherent interference (**pilot contamination**)*

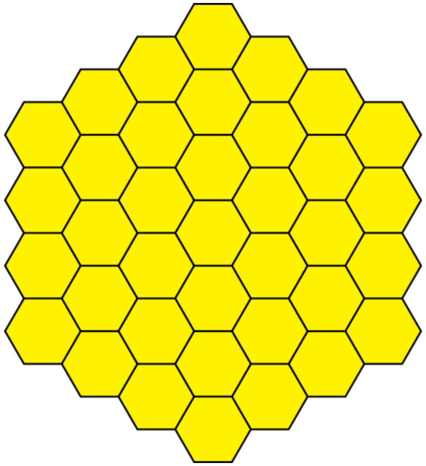
Numerical Results

- Problem Formulation:

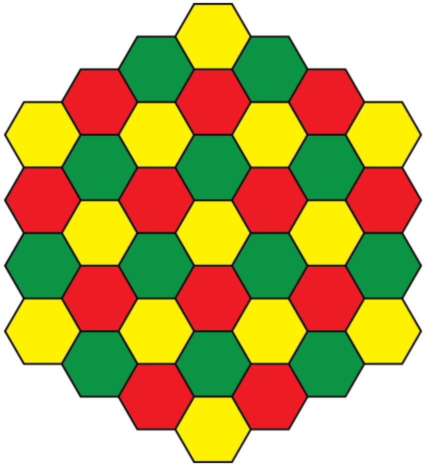
$$\underset{K,f}{\text{maximize}} \quad \text{average sum spectral efficiency} \quad [\text{bit/s/Hz/cell}]$$

for a given M *and* τ_c.

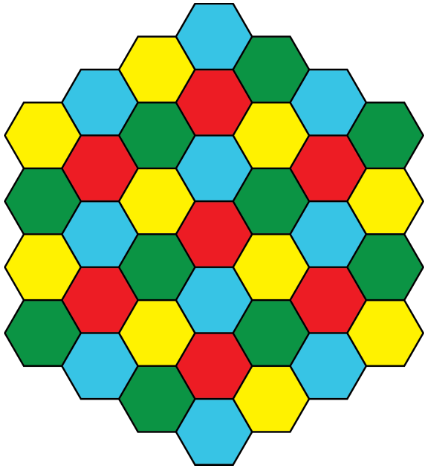
- Large hexagonal cellular network
- Power control to give same SNR to every user
- Average over user locations for different K and f $\quad (\tau_p = fK)$
- Use closed–form expressions



Reuse $f = 1$



Reuse $f = 3$



Reuse $f = 4$

Assumptions

Pathloss exponent: 3.7

Coherence: $\tau_c = 400$

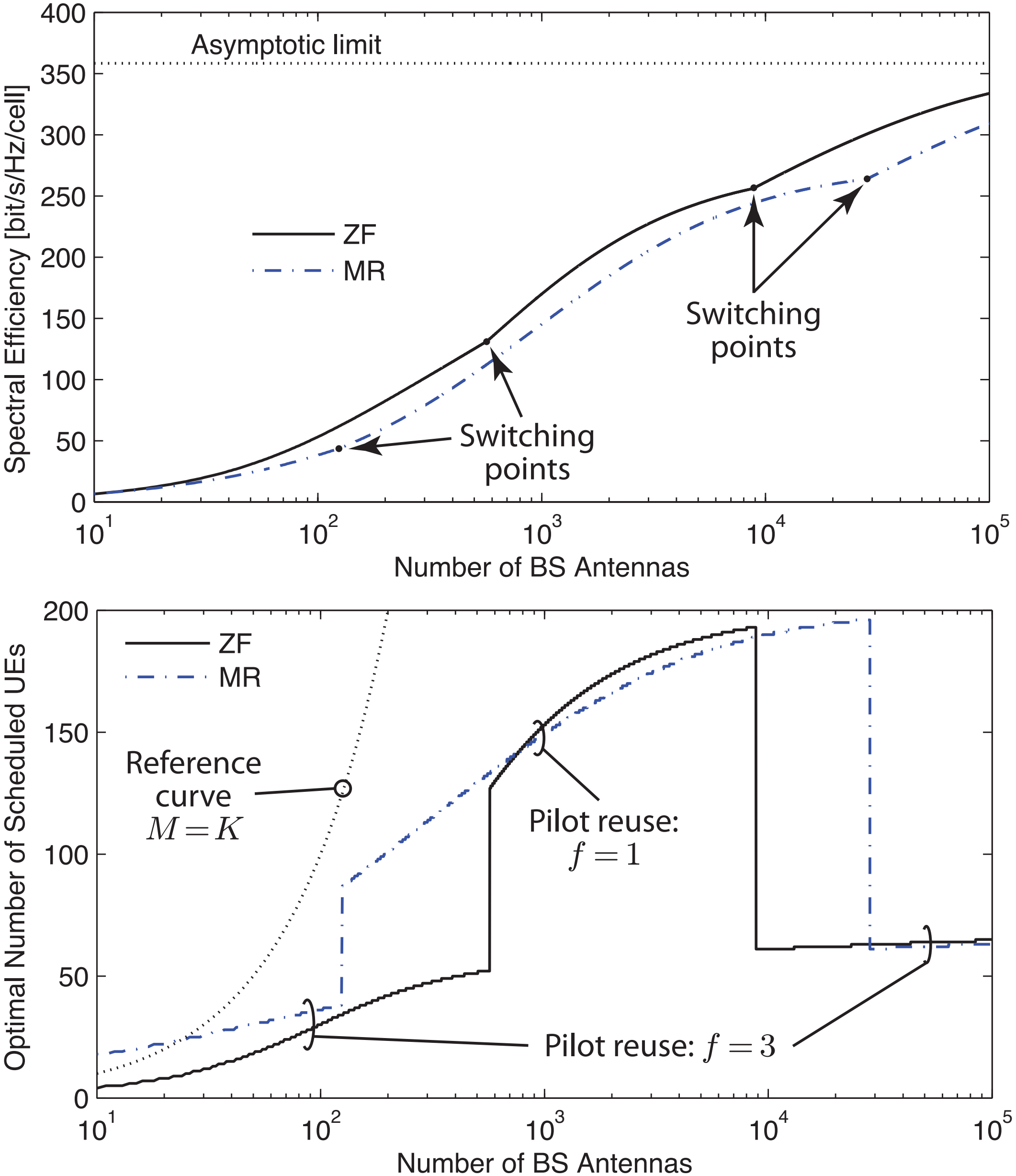
Rayleigh fading

SNR 5 dB

Asymptotic Behaviors (Uplink)

Observations

- *Asymptotic limit not reached*
- *Reuse factor $f = 3$ is desired*
- *K is different for each scheme*
- *Relatively small difference between optimized schemes*
- *Huge spectral efficiency, despite pilot contamination*
- *Same result in downlink by power control (use duality)*



Flexible Number of Users

- SE w.r.t. Number of Users
  - Optimized reuse factors

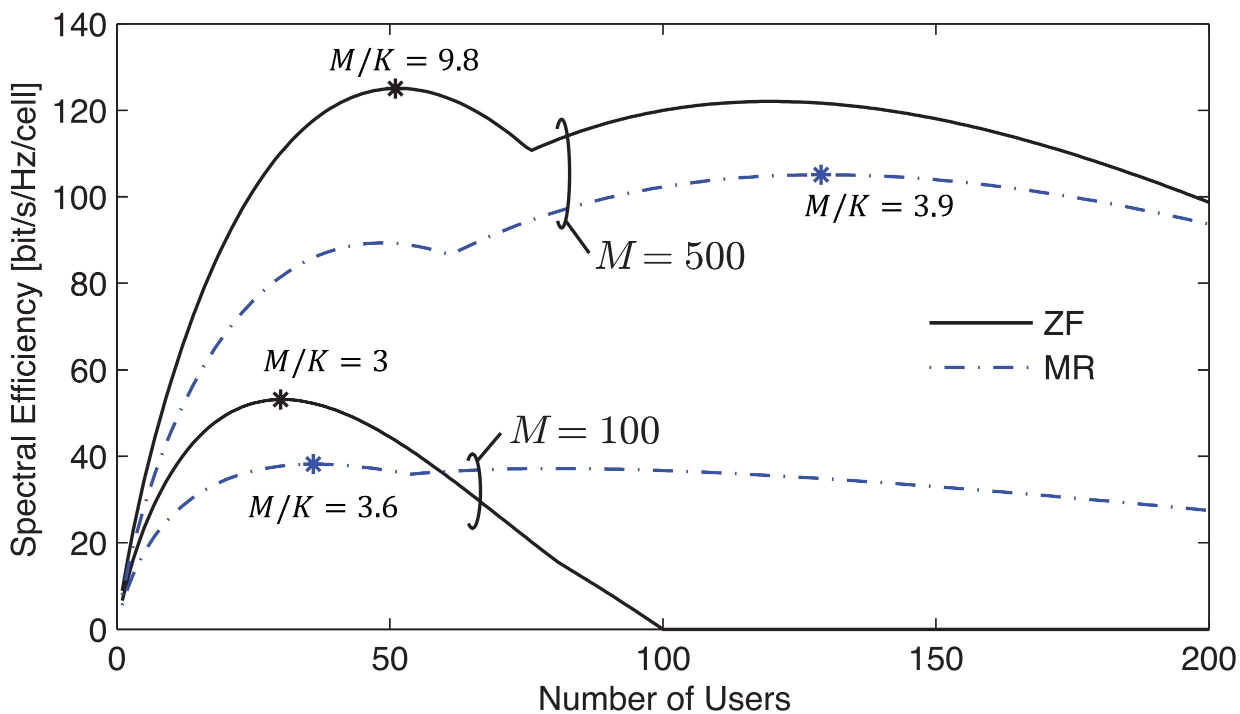
Observations

Stable SE for $K > 10$

Trivial scheduling: Admit everyone

ZF and MR give similar per-cell performance

$M/K < 10$ is just fine!



Spectral Efficiency per User

- User Performance for Optimized System
  - Optimized reuse factors

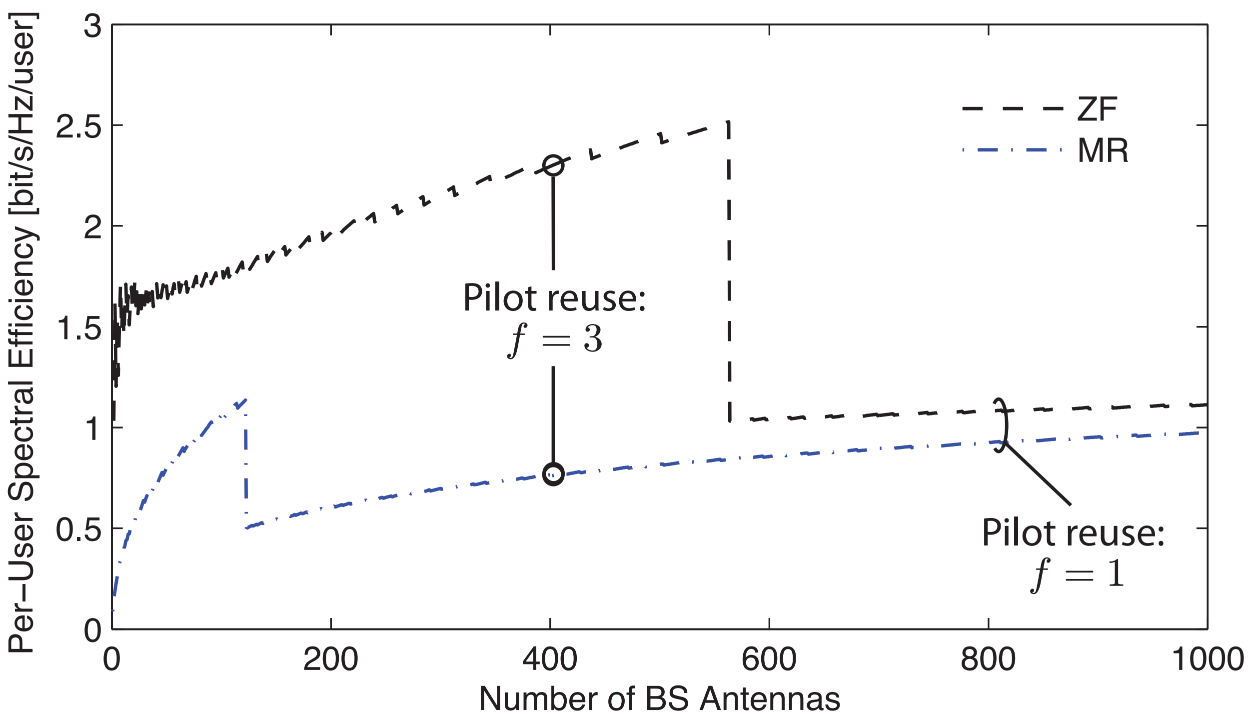
Observations

Modest user performance:

BPSK, Q-PSK, 16-QAM

Schemes for different purposes:

ZF > MR



PART 2: SUMMARY

Summary

- Spectral Efficiency Analysis
  - Channel hardening: Use capacity bounds based on AWGN channels
  - Closed-form for uncorrelated Rayleigh fading with MR or ZF
  - Computed numerically for any channel distribution and processing

- Uplink-Downlink Duality
  - Same performance achievable in uplink and downlink
  - Same precoding/detection vectors, different power allocation

- Numerical Observations
  - Very high spectral efficiency per cell, not per user
  - Non-universal pilot reuse ($f = 3$) preferred
  - MR and ZF prefer different values on K and f
  - “An order of magnitude more antennas than users” is not needed

Part 3

Resource Allocation and Access

LINKÖPING UNIVERSITY

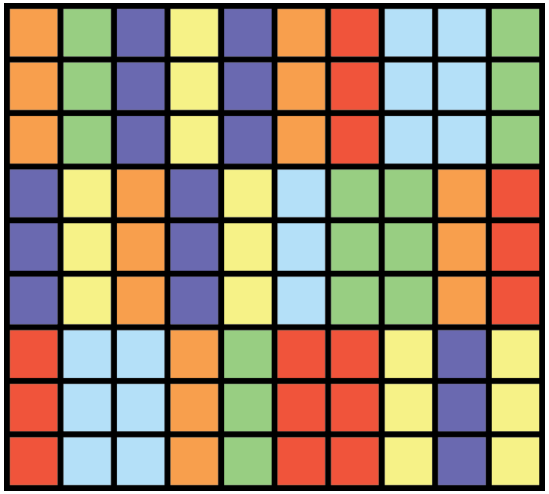
Three Aspects of Cross-Layer Design

- Outline
  - Scheduling and Spatial Correlation
  - Deployment
  - Power control
  - User access

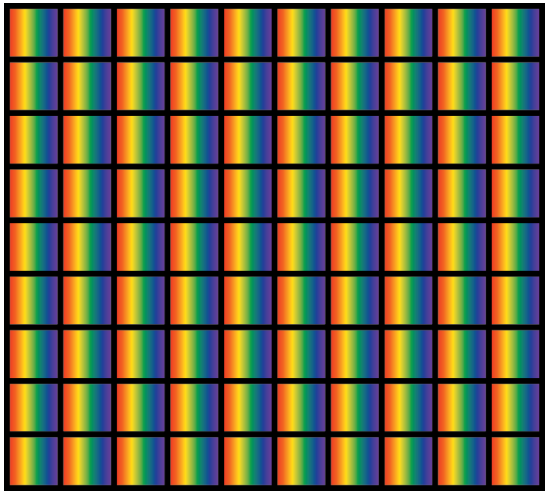
SCHEDULING AND SPATIAL CORRELATION

Simple Scheduling

- Scheduling in 4G
  - Give each time/frequency block to one user
  - Utilize current fading realization
  - *Not needed in Massive MIMO*



- Spatial Resource Allocation
  - Each user get the whole bandwidth, whenever needed!
  - Separate users spatially
  - Same channel quality in all blocks



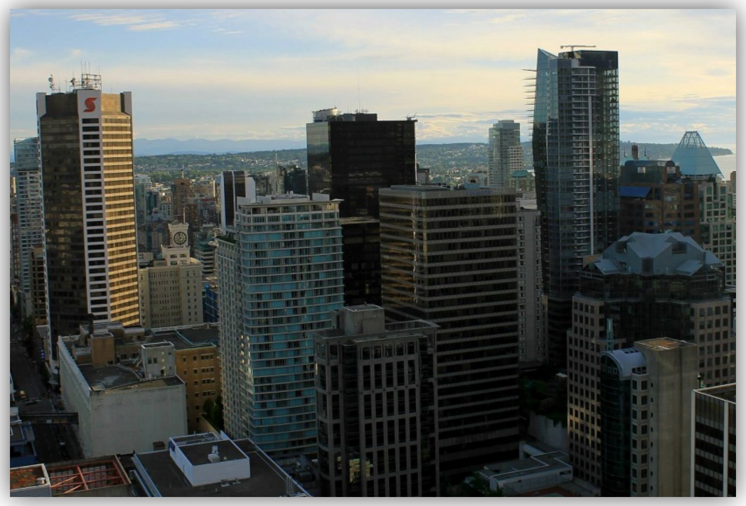
Impact of Spatial Correlation (1)

- Uncorrelated fading is a simplification
  - How do model spatial correlation?

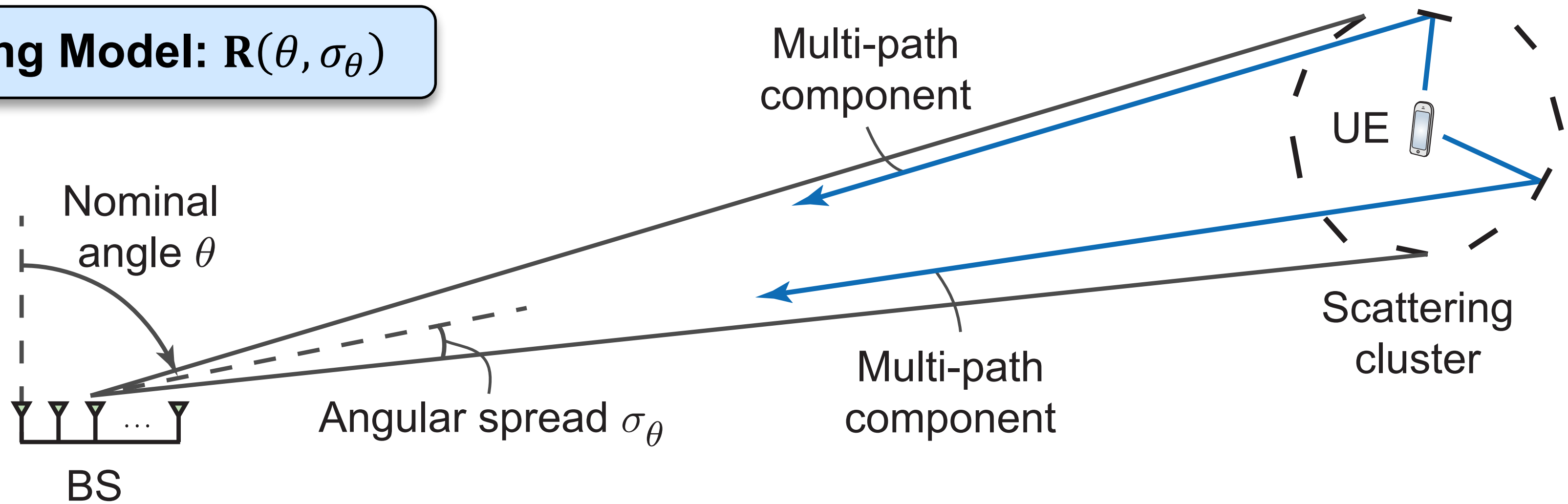
- Correlated Rayleigh fading:

$$\mathbf{h}_{li}^{j} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{li}^{j}\right)$$

  - Covariance matrix $\mathbf{R}_{li}^{j}$
  - Describes probabilities of channel directions



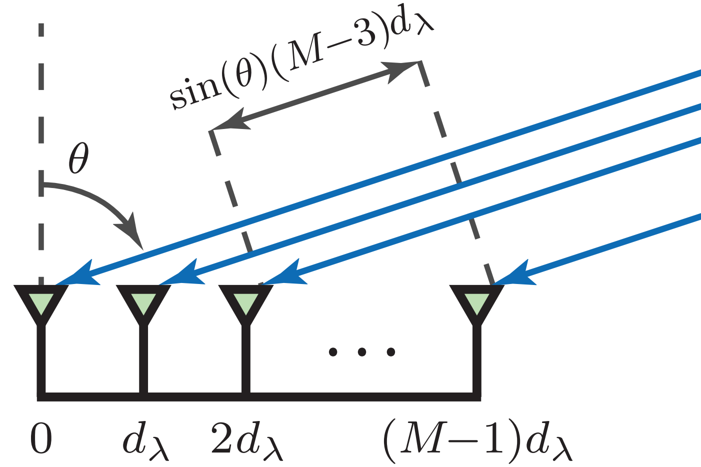
One-Ring Model: $\mathbf{R}(\theta, \sigma_\theta)$



Impact of Spatial Correlation (2)

- One-Ring Model for uniform linear array

$$[\mathbf{R}(\theta,\sigma_\theta)]_{m,l} = \int_{-\pi}^{\pi} e^{2\pi j d_\lambda (l-m)\sin(\psi)} \underbrace{f(\psi,\theta,\sigma_\theta)}_{\textit{Angular multipath distribution}} d\psi$$

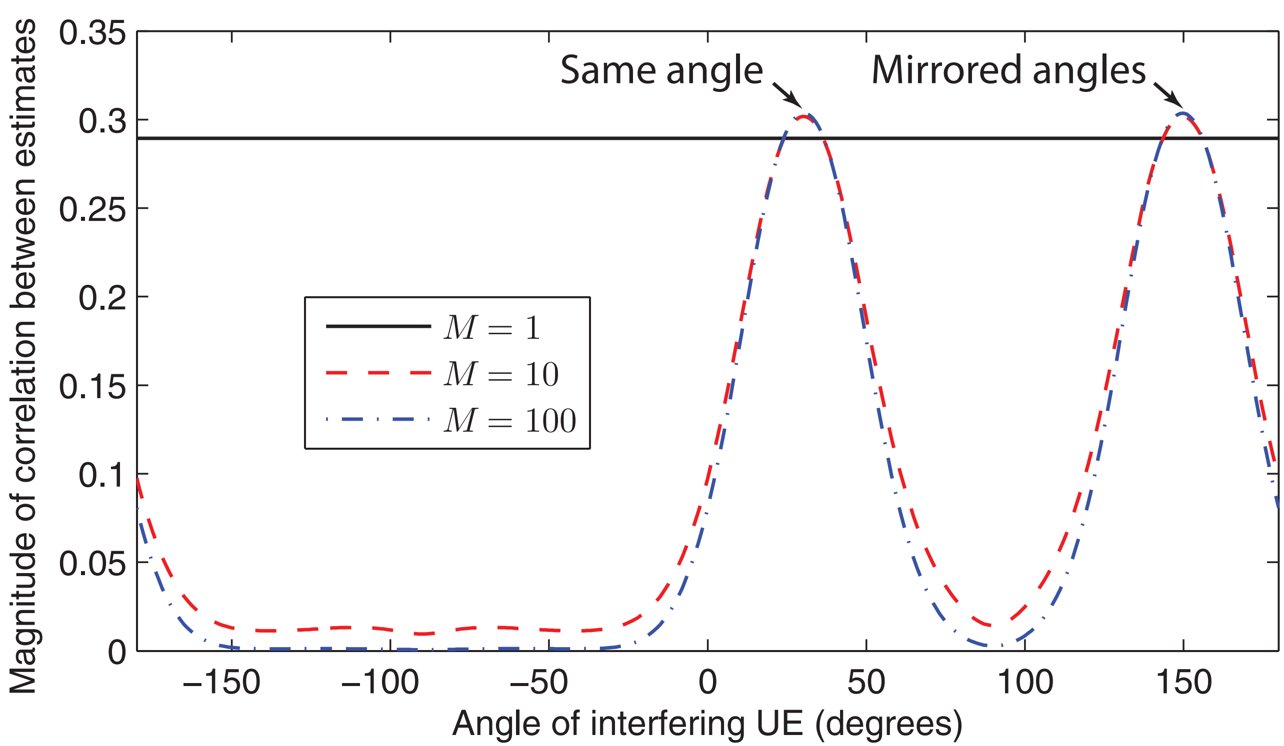


Pilot contamination

Users with same pilot: Correlated channel estimates

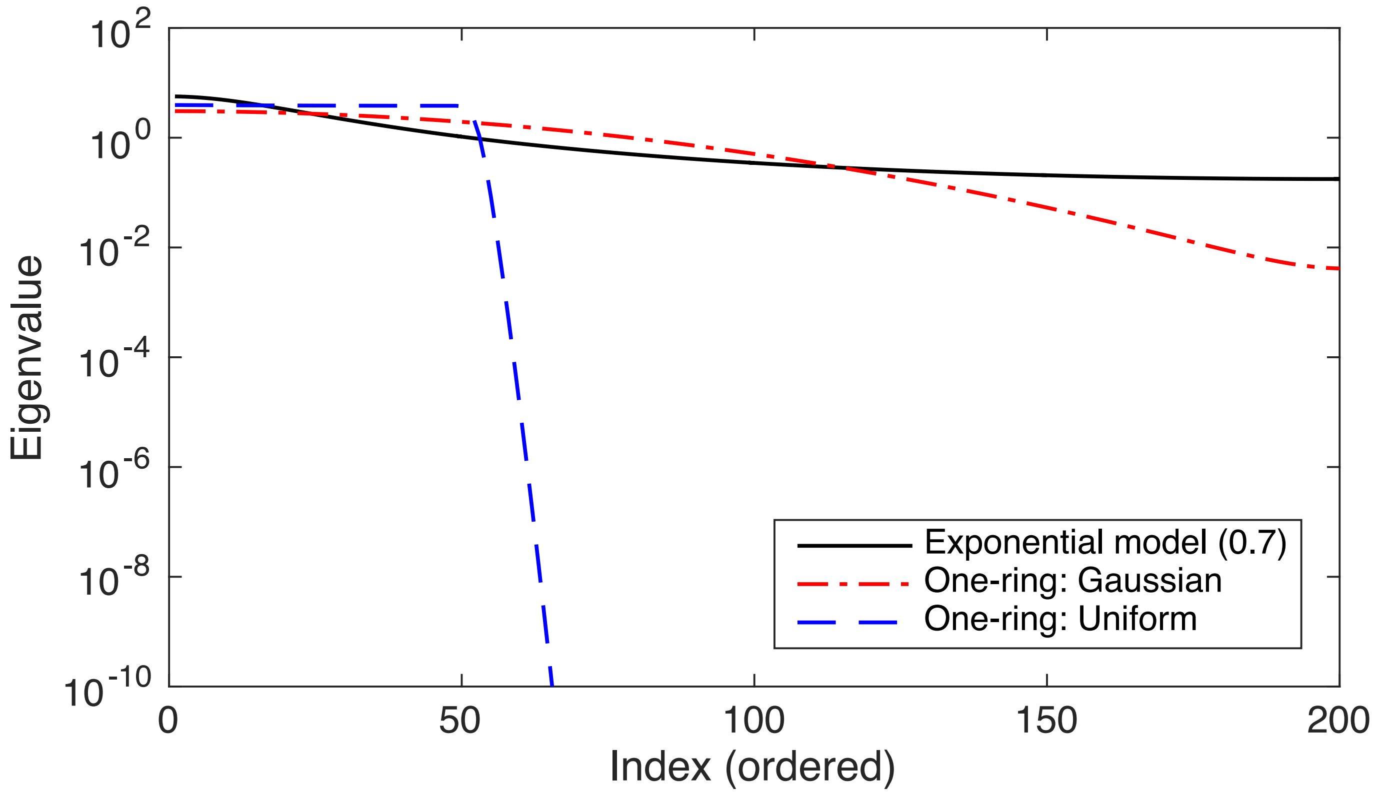
Similar $\mathbf{R}$ *matrices: More contamination*

Different $\mathbf{R}$ *matrices: Less contamination*



Impact of Spatial Correlation (3)

- Are Covariance Matrices Sparse?
  - It depends on the correlation model



Hard to rely on sparsity for simplified channel estimation

Is Spatial Correlation Important? (1)

- Yes, it is very important!
  - Increase interference variation – similar/different covariance matrices

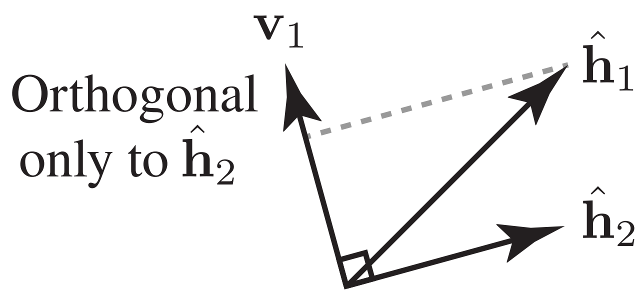
- Pilot contamination is not a fundamental limit any longer
  - Example: UEs share pilot and have channels $\mathbf{h}_1$ and $\mathbf{h}_2$ to a BS
  - Uncorrelated: $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ are parallel
  - Different correlation: $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ are not parallel

Rejection of pilot contamination

Select a vector $\mathbf{v}_1$ such that

$$\mathbf{v}_1^H \hat{\mathbf{h}}_1 \neq 0, \quad \mathbf{v}_1^H \hat{\mathbf{h}}_2 = 0$$

Possible when covariance matrices are linearly independent

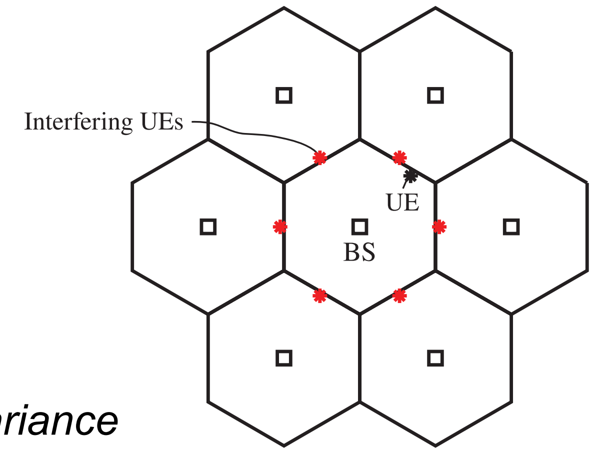


Is Spatial Correlation Important? (2)

- M-MMSE combining:

$$\mathbf{v}_{lk} = \left(\sum_{j=1}^{L} \sum_{i=1}^{K} p_{ji} \hat{\mathbf{h}}_{ji}^{l} \left(\hat{\mathbf{h}}_{ji}^{l} \right)^{H} + \mathbf{Z}_{lk} + \sigma^2 \mathbf{I}_M \right)^{-1} \hat{\mathbf{h}}_{lk}^{l}$$

Based on estimation error covariance



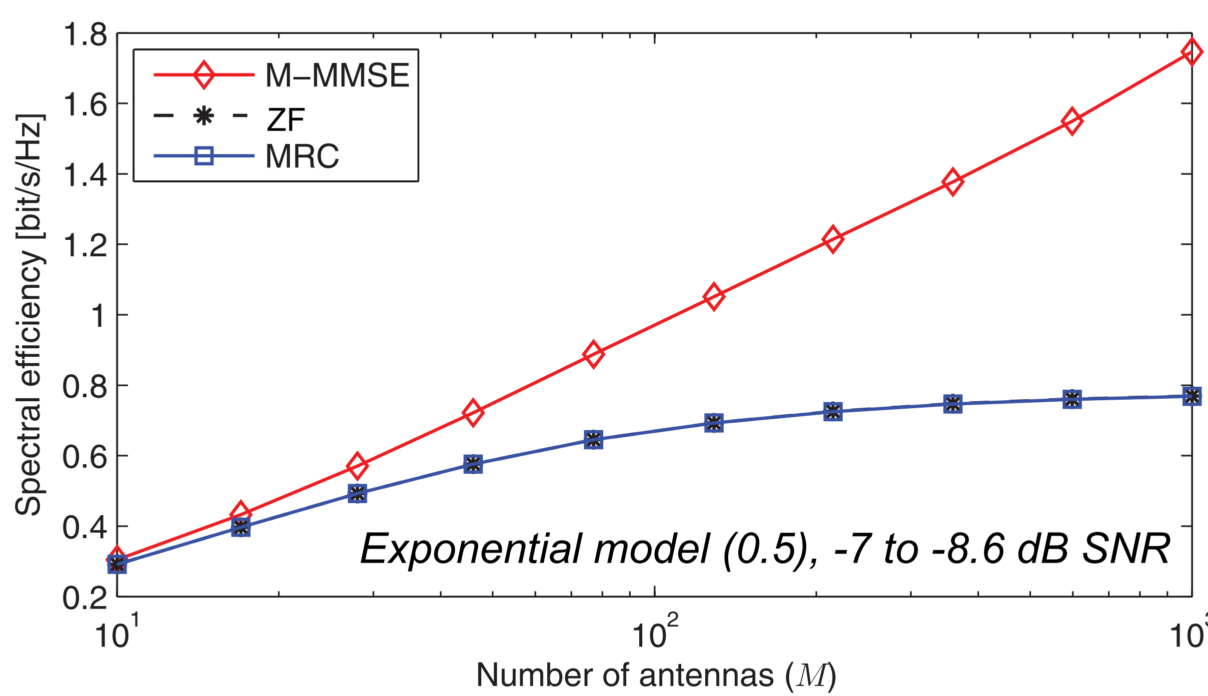
Pilot contamination can be rejected

Except in uncorrelated Rayleigh fading

Or by suboptimal schemes

No sparseness needed!

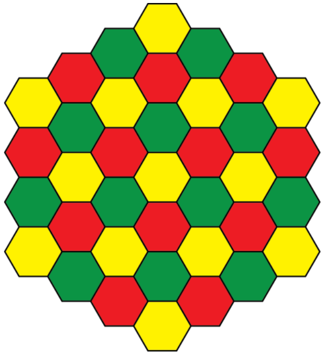
Small differences between covariance matrices is sufficient



DEPLOYMENT

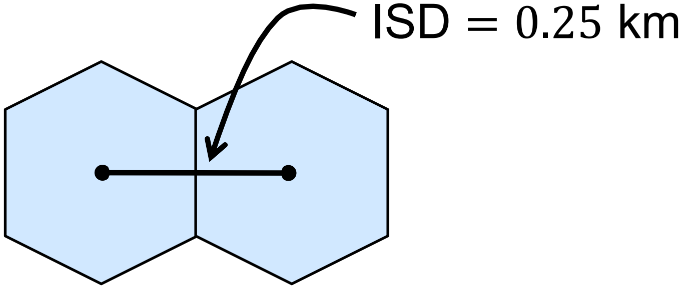
When to Deploy Massive MIMO?

- Any Cellular Scenario with High User Density
  1. Dense urban cities
  2. Wide-area broadband coverage



- Example: Dense Urban Information Society (METIS TC2)
  - User density: Up to 200 000 users / km^2
  - One data packet/minute, one second to transfer

$$\boxed{\text{Average active users } = \frac{\text{All users}}{60}}$$



	10^3 users/km^2	10^4 users/km^2	10^5 users/km^2
Average active users:	0.9	9	90

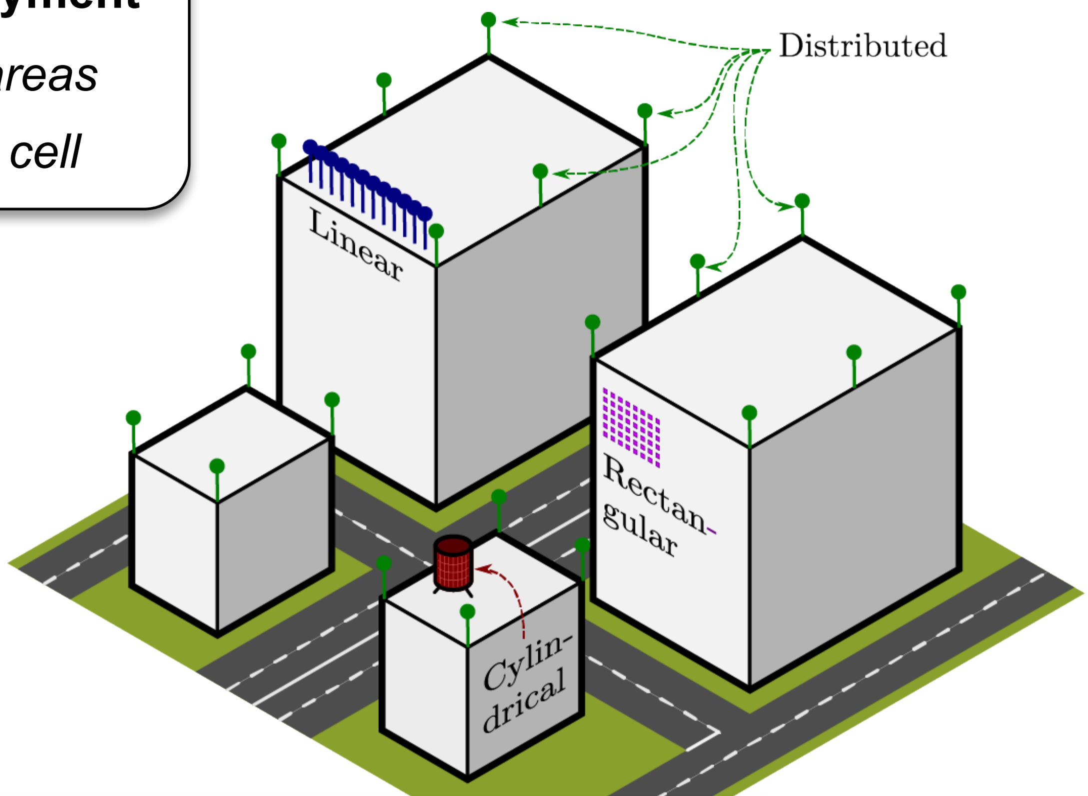
Massive MIMO is needed!

How to Deploy Antennas?

Conventional Cellular Deployment

Non-overlapping coverage areas

One or multiple sectors per cell



- Co-located Deployment
  - 1D, 2D, or 3D arrays
  - No need for sectors
- Distributed Deployment
  - Remote radio heads

Key Benefits of Massive MIMO

Outdoor users: Handle mobility and guarantee coverage

Indoor users: No need to put BSs inside buildings

POWER ALLOCATION

Downlink Power Control (1)

- Downlink Spectral Efficiency [bit/s/Hz] of User k:

$$\underline{\mathrm{SE}}_k = \log_2\left(1 + \frac{q_k G \gamma_k}{\sum_{i=1}^{K} q_i z_k + \sigma^2}\right)$$

with G and z_i given by precoding choice.

Depends on the large-scale fading $\beta_1, \dots, \beta_K$

Not on instantaneous fading realizations

- Opportunity: Implement long-term power control
  - Optimization is hard to perform in every coherence block (msec)
  - Highly feasible when large-scale fading is fixed (sec)

Maximize some utility function $u\left(\underline{\mathrm{SE}}_1, \dots, \underline{\mathrm{SE}}_K\right)$

Downlink Power Control (2)

- **Example**: Max-min SE Optimization

$$u\left(\underline{\mathrm{SE}}_1, \dots, \underline{\mathrm{SE}}_K\right) = \min_k \underline{\mathrm{SE}}_{\mathrm{k}}$$

  - Achieve uniformly good performance for all users
  - Limited transmit power: $q_1 + \cdots + q_K \leq Q$

- Epigraph formulation

$$\underset{R, q_1, \dots, q_K}{\text{maximize}} \quad R \qquad \text{subject to} \quad \underline{\mathrm{SE}}_k \geq R \ \ \forall k, \qquad \sum_{i=1}^{K} q_i \leq Q$$

  - Fix R and solve linear optimization problem:

$$\underset{q_1, \dots, q_K}{\text{minimize}} \ \sum_{i=1}^{K} q_K \ \ \text{subject to} \ \ q_k G \gamma_k \geq \left(\sum_{i=1}^{K} q_i z_k + \sigma^2\right) (2^R - 1) \ \ \forall k$$

  - Make a bisection over R to find the value that gives $\sum_{i=1}^{K} q_i = Q$

***Solvable with polynomial complexity**: Depends on K but not M*

Uplink Power Control

- Similar Optimization Approach Possible
  - Practical constraint: Limited dynamic range in BS receiver
  - Pathloss differences: Up to 50 dB in a cell
  - Power control should reduce this to < 10 dB

- Uplink Spectral Efficiency of User k with MR:

$$\overline{\mathrm{SE}}_k = \log_2\left(1 + \frac{p_k M \beta_k}{\sum_{i=1}^{K} p_i \beta_i + \sigma^2}\right)$$

- Statistical channel inversion: $p_k = p/\beta_k$ for some p
  - Result: All signals are received with power p
  - Equivalent to max-min optimization:

$$\overline{\mathrm{SE}}_k = \log_2\left(1 + \frac{pM}{\sum_{i=1}^{K} p + \sigma^2}\right) = \log_2\left(1 + \frac{pM}{pK + \sigma^2}\right) \quad \text{for } k = 1, \dots, K$$

USER ACCESS

Connection Issue

- *Massiveness* of 5G Networks and Beyond
  - Massive number of user equipments (UEs), intermittent activity
  - Massive total data traffic
  - Massive differences in traffic between UEs (mobile broadband, internet-of-things, etc.)

Can Pre-allocate Pilots with Users

Connected user: Temporarily allocated a dedicated pilot

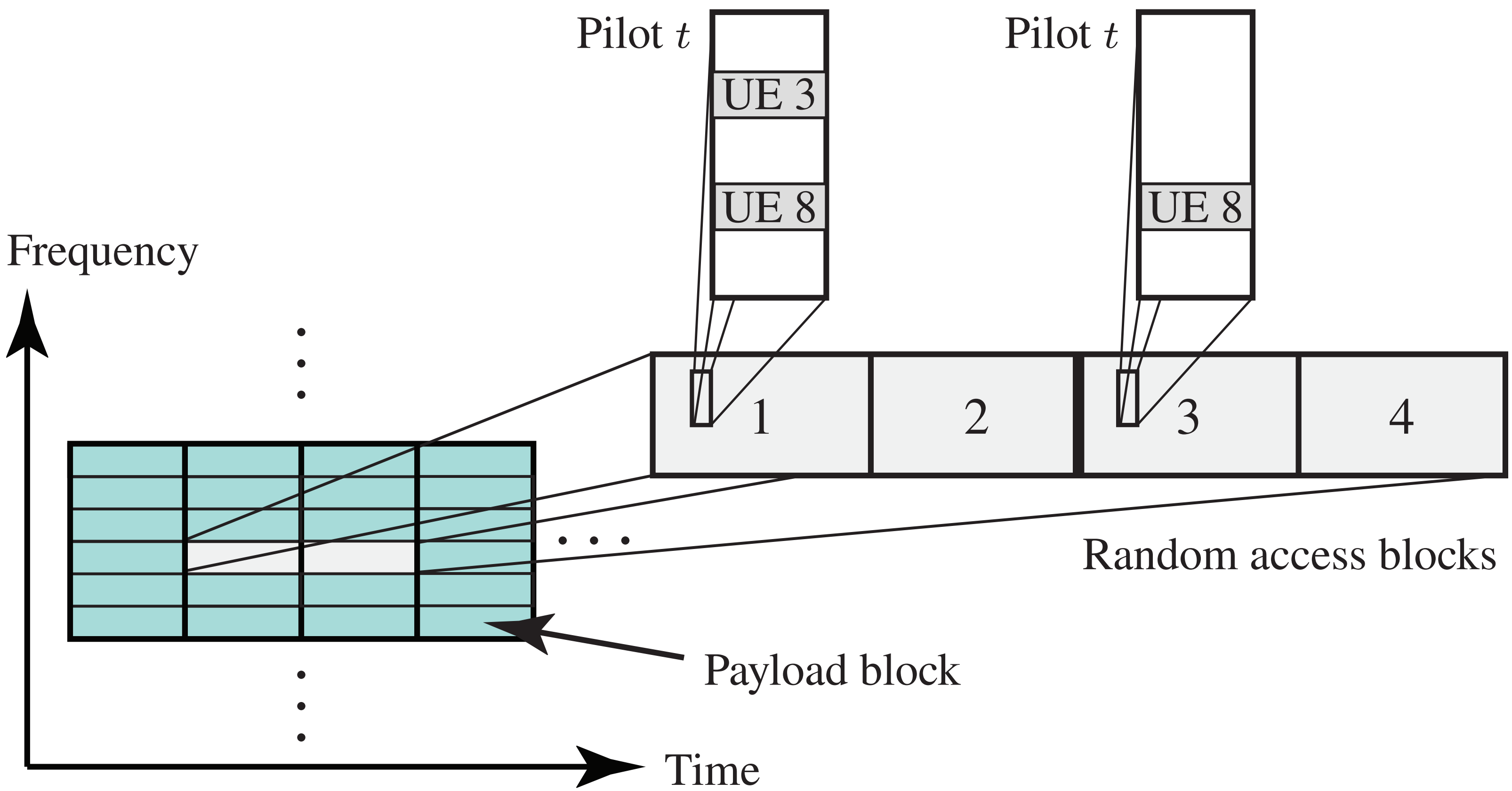
How to connect and disconnect that many UEs from the network?

LTE Random Access Solution Not Enough

More users → More access contention → Requires more overhead

Proposed Frame Structure

- Two types of resource blocks
  1. Payload blocks ← Operated as in classic Massive MIMO
  2. Random access blocks ← Accessing users send a random pilot

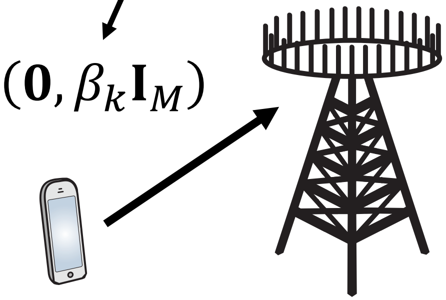


System Model

- Preliminaries
  - K users want to connect
  - Generate $\tau_p \leq \tau_c/2$ pilots for random access
  - P_a: Probability of trying to access
  - Accessing users picks a pilot at random

Channel gain

$$\mathbf{h}_k \sim CN(\mathbf{0}, \beta_k \mathbf{I}_M)$$



UE k

- $\mathcal{S}$: Set of users picking an arbitrary pilot sequence:

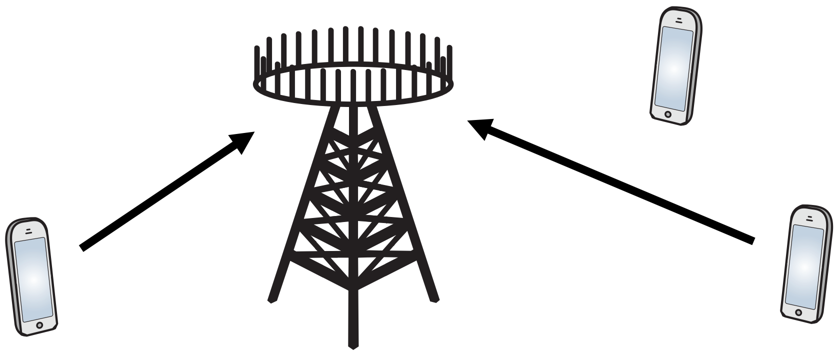
$$|\mathcal{S}| \sim \text{Binomial}\left(K, \frac{P_a}{\tau_p}\right)$$

  - Probability of pilot collision

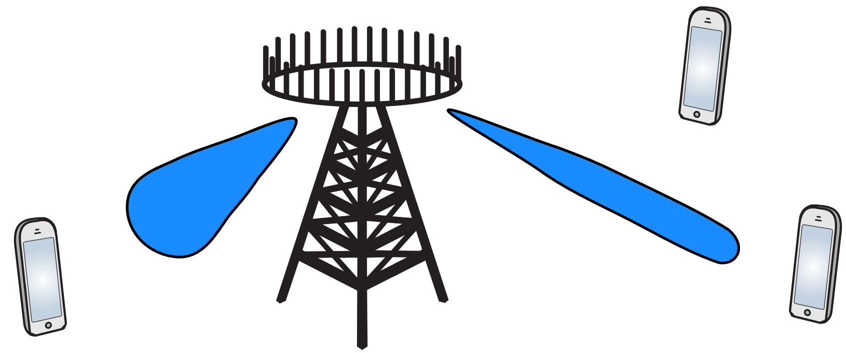
$$\Pr\{|\mathcal{S}| \geq 2\} = 1 - \left(1 - \frac{P_a}{\tau_p}\right)^K - K\frac{P_a}{\tau_p}\left(1 - \frac{P_a}{\tau_p}\right)^{K-1}$$

How to detect and resolve collisions?

Random Access Protocol (1)

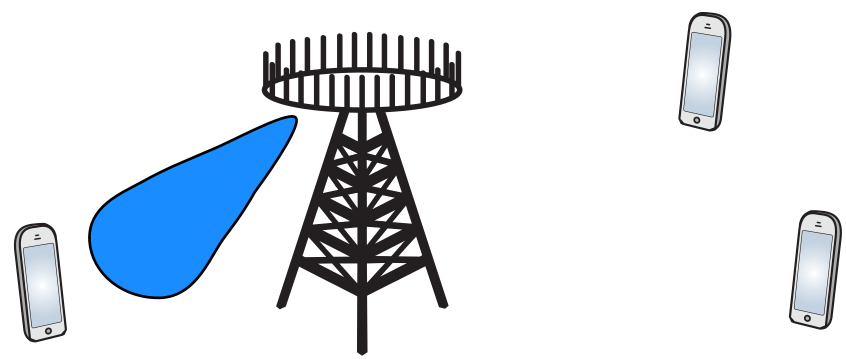


Step 1*: Subset of users send a random pilot*



Step 2*: BS sends precoded response*

Step 3*: Only one user repeats the pilot*



Step 4*: BS responds with connection information*

Random Access Protocol (2)

- User k tries to connect
  - Step 1: Select a pilot at random, transmit in uplink to BS
    Set $\mathcal{S}$ of users selected this pilot, $|\mathcal{S}| > 1$ is a collision
  - Step 2: BS estimates sum of channels to accessing users
    BSs sends a precoded response to the accessing users
  - Step 3: User k observes array gain $G_k = M\beta_k / \sum_{i\in\mathcal{S}} \beta_i$
    Can detect if a collision occurred: $G_k < M$
    A subset $\mathcal{S}_{\text{retrans}} \subseteq \mathcal{S}$ of users retransmit

***Retransmission decision rule** at user k:*

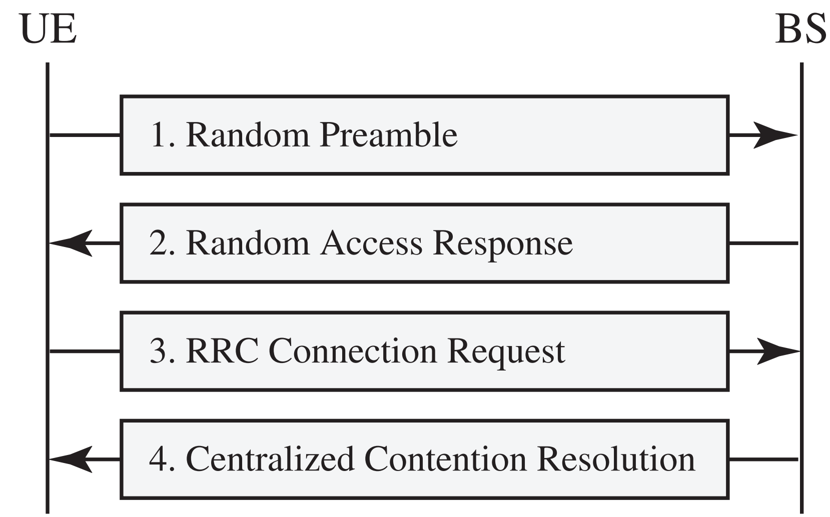
Retransmit:	$G_k > M/2$
Inactive:	$G_k < M/2$

  - Step 4: BS tries to decode signal from retransmitting users
    $|\mathcal{S}_{\text{retrans}}| = 1$: Decoding succeeds
    Admit the user: Give a dedicated pilot for data transmission

Comparison: LTE and New Scheme

LTE

Physical Random Access Channel (PRACH)

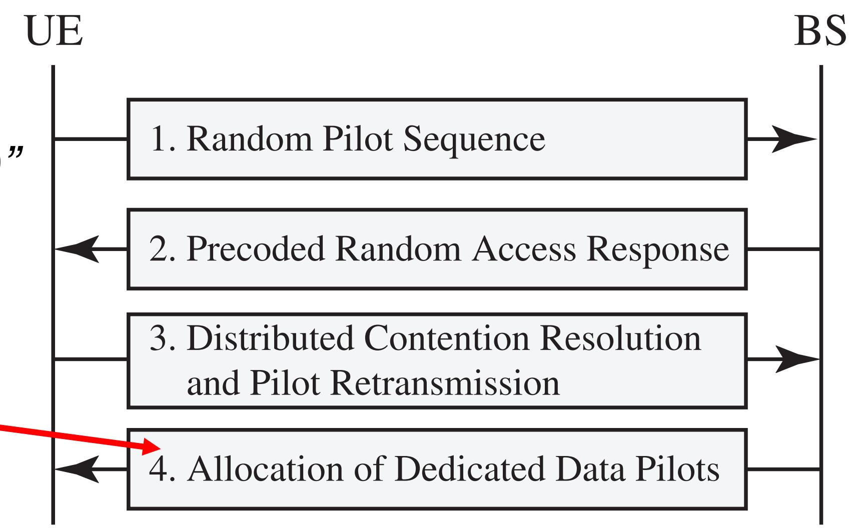


New Scheme

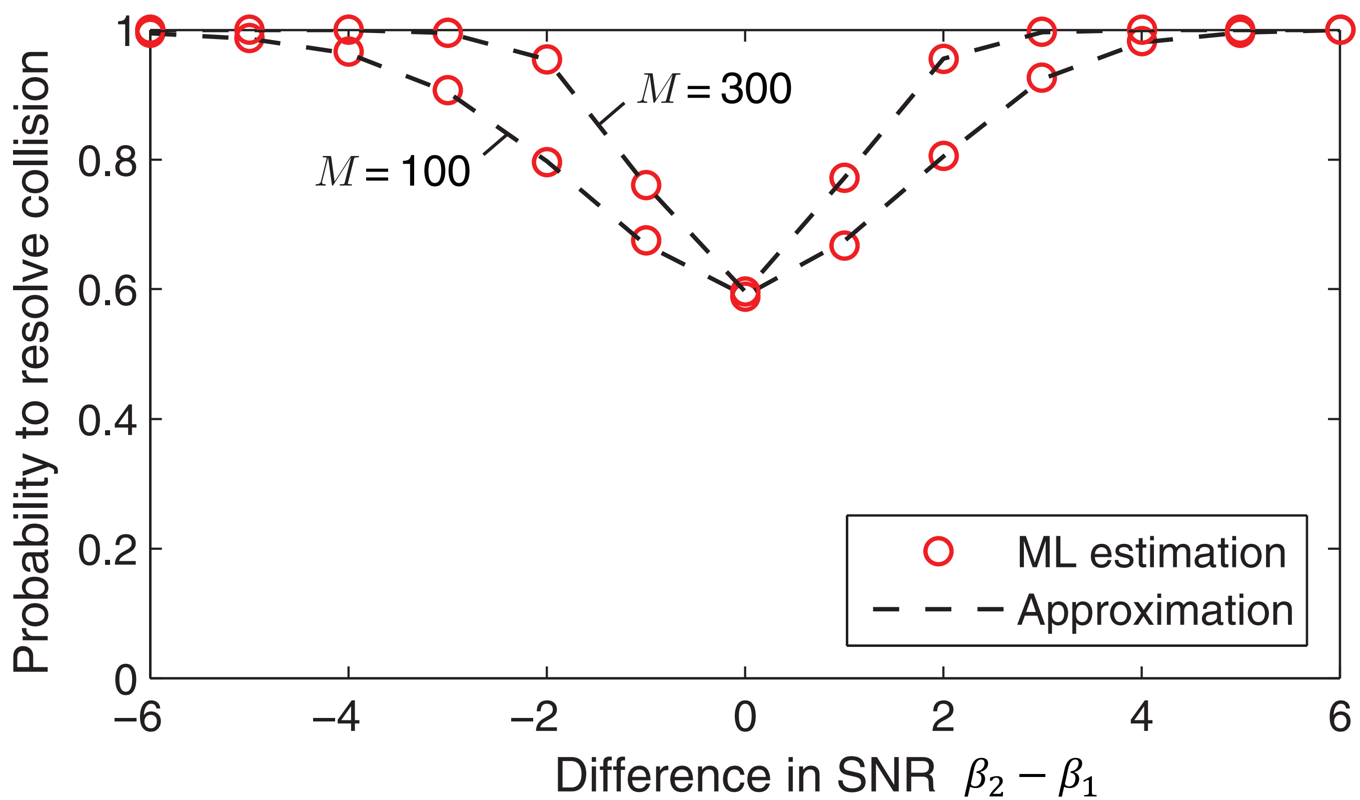
"Strongest-User Collision Resolution (SUCRe)"

Exploits Massive MIMO characteristics

90% of contentions have been resolved at this point!



Basic Test: Resolving a Two-User Collision



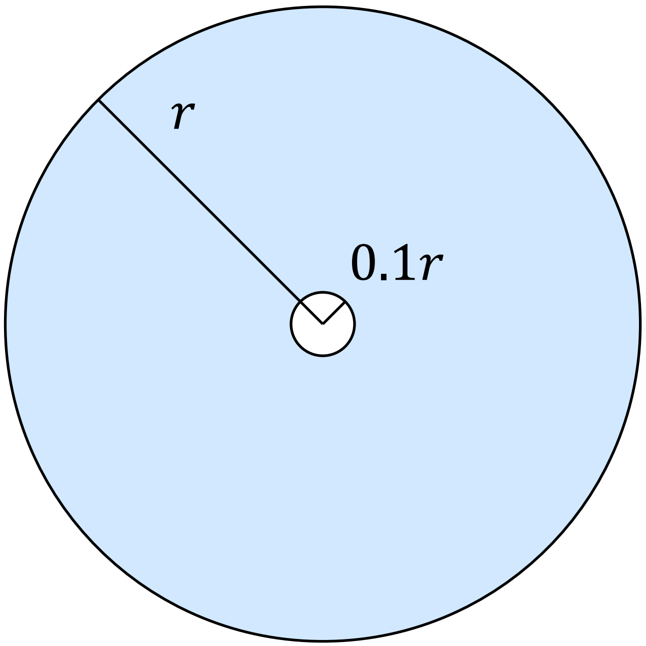
Assumptions

- Pilot SNR of UE k is β_k $(p = q = \sigma^2 = 1)$
- First UE: $\beta_1 = 10$ dB
- Second UE: $\beta_2 = 4$ to 16 dB

Simulation Setup

Scenario

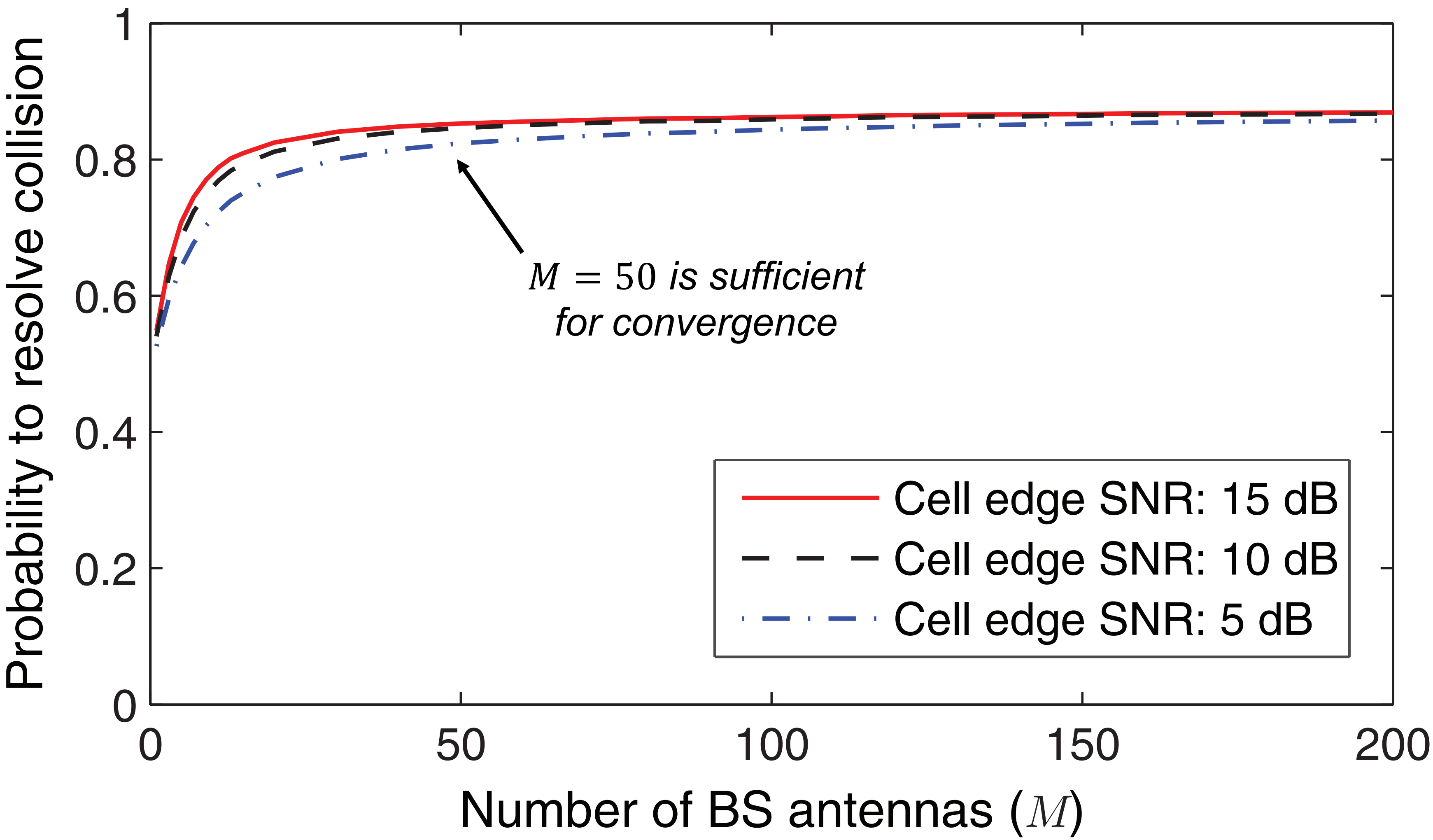
- $\tau_p = 10$ pilot sequences
- $K = 50$ UEs want to connect
- Uniformly distributed, except in cell center
- Pathloss exponent: 3.7
- Shadow fading: 8 dB standard deviation
- Cell edge SNR without shadowing is set



Performance Metric

- Probability to resolve conflicts: $\Pr\{|\mathcal{S}_{\text{retrans}}| = 1\}$

Simulation Result



P_a is optimized for maximal resolution probability

PART 3: SUMMARY

Summary

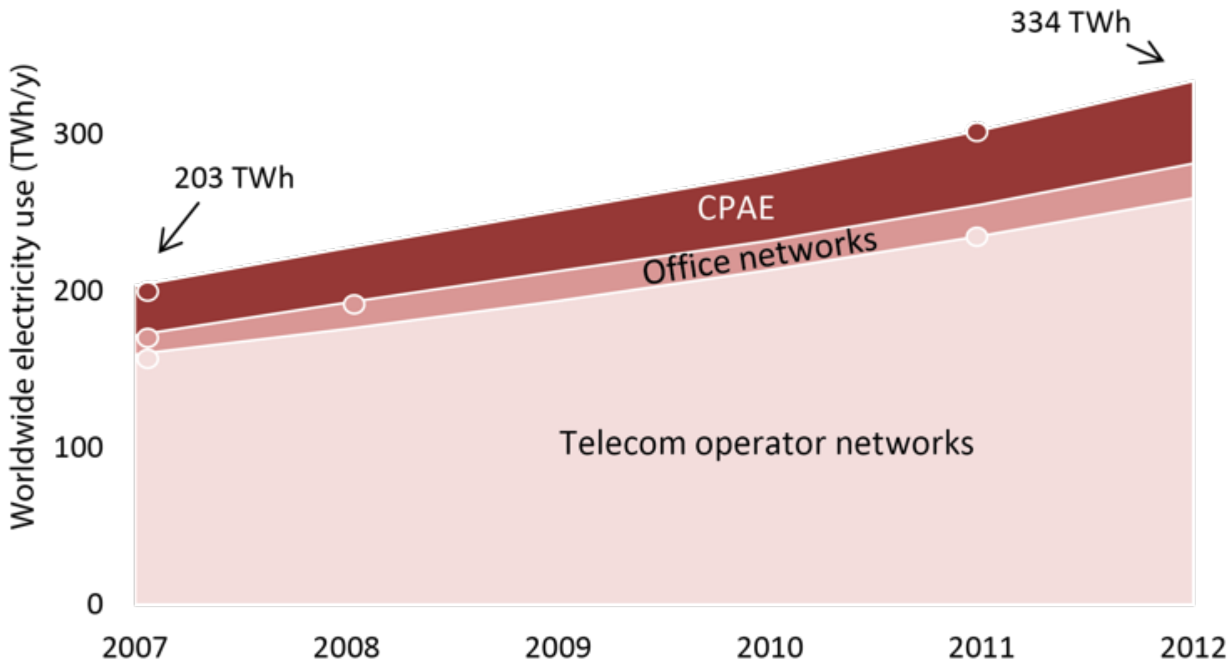
- Resource Allocation
  - Channel hardening: Negligible frequency-selective and small-scale fading
  - Give everyone the whole bandwidth
  - Spatial correlation can be exploited to reduce interference
- Power Control
  - SE expressions only depend on large-scale fading
  - Max-min SE optimization is a simple linear optimization problem
  - Same thing can be shown for other utility functions
  - Uplink power control should take dynamic range into account
- User Access
  - Massive number of UEs – Only allocated dedicated pilots to active UEs
  - Request protected pilots by random access – leads to collisions
  - SUCRe exploits channel hardening to resolve collisions distributively

Part 4

Energy Efficiency

WHAT IS ENERGY EFFICIENCY?

Energy Consumption



Source: Heddeghem et al. "Trends in worldwide ICT electricity consumption from 2007 to 2012"

- Network Electricity Consumption
  - Dominated by network infrastructure – increases continuously
  - 1000x higher data rates: Easy to achieve using 1000x more power
Hard to achieve without using more power
  - Calls for **much higher energy efficiency**!

What is Energy Efficiency?

- Benefit-Cost Analysis of Networks
  - Systematic approach to analyze strengths and weaknesses of networks

Cost:
Power Consumption
[Watt = Joule/s]

→ ***Network*** →

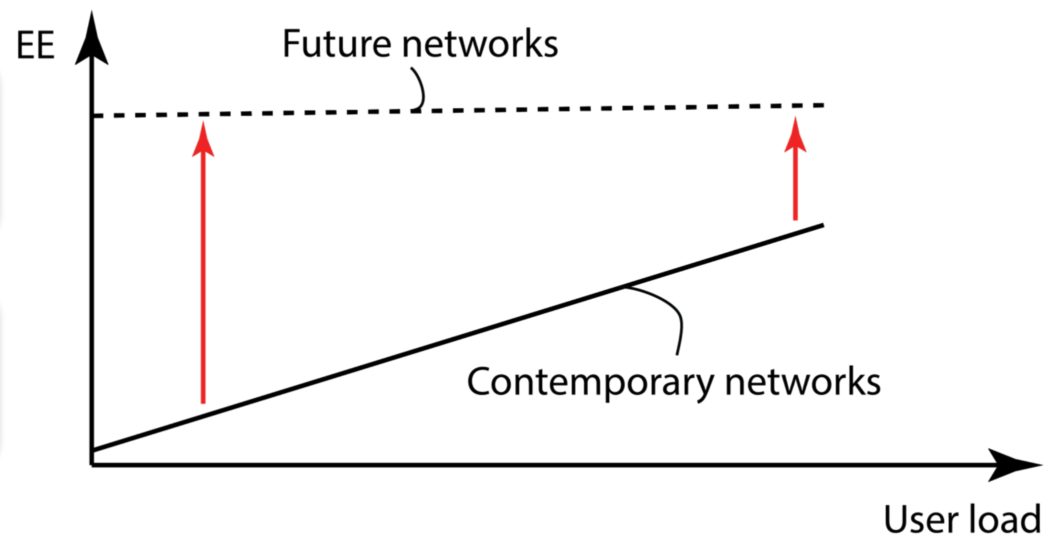
Benefit:
Sum Data Rate
[bit/s]

- Definition: Energy Efficiency (EE):

$$\text{EE [bit/Joule]} = \frac{\text{Average Sum Rate [bit/s/cell]}}{\text{Power Consumption [Joule/s/cell]}}$$

Contemporary networks:
Very inefficient at low load

Future networks:
Must be more efficient at any load



Transmit Power Scaling Law

Power Scaling Law

If the transmit power ρ decreases as $1/M^{\alpha}$ for $\alpha \leq 1/2$:

SE will not go zero as $M \rightarrow \infty$

Example: Set $p_{jk} = p_0/M^{\alpha}$ in $\overline{\mathrm{SE}}_{lk} = \log_2(1 + \mathrm{SINR}_{lk})$:

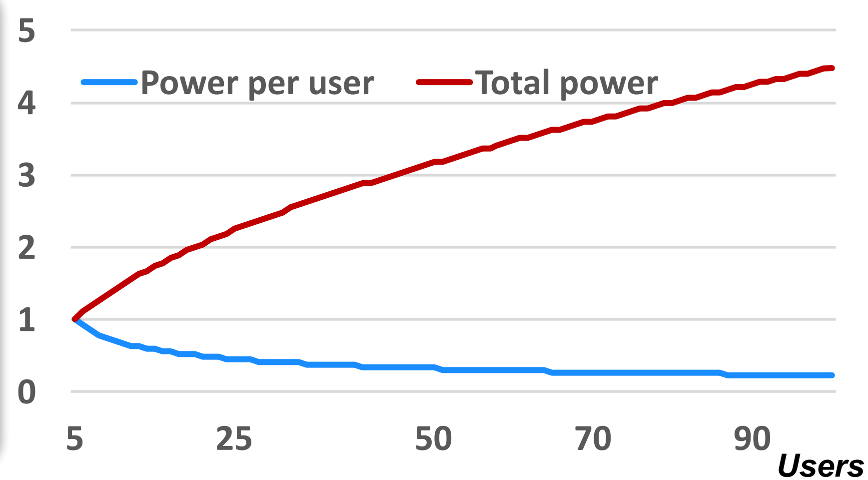
$$\mathrm{SINR}_{lk} = \frac{p_0 M^{1-2\alpha} \tau_p \left(\beta_{lk}^{l}\right)^2}{\left(O(M^{-\alpha}) + \sigma^2\right)^2 + \sum_{j \in \mathcal{P}_l(f)\{l\}} p_0\, \tau_p \left(\beta_{jk}^{l}\right)^2 M^{1-2\alpha}} \rightarrow \frac{p_0 \tau_p \left(\beta_{lk}^{l}\right)^2}{\sum_{j \in \mathcal{P}_l(f)\{l\}} p_0\, \tau_p \left(\beta_{jk}^{l}\right)^2}$$

***Observations** ($\alpha = 1/2$)*

Power per user: Decreases as $\frac{1}{\sqrt{M}}$

Power per BS antenna: Decreases as $\frac{1}{M^{3/2}}$

Sum power: $\frac{K}{\sqrt{M}}$ increases as $\sqrt{M}$ for fixed $\frac{M}{K}$



Radiated Energy Efficiency

- Energy Efficiency in Cell l with Power Scaling:

$$\mathrm{EE}_l = \frac{\text{Sum Rate in Cell } l \text{ [bit/s]}}{\text{Power Consumption [Joule/s]}} = \frac{B\left(1-\frac{\tau_p}{\tau_c}\right)\sum_k \log_2(1+\mathrm{SINR}_{lk})}{\frac{K p_0}{M^{\alpha}}}$$

  - Bandwidth: B Hz

  - Consequence of scaling law as $M \to \infty$:

$$\left.\begin{array}{l}\text{1. Sum rate } \to \text{constant} > 0 \\ \text{2. Transmit power } \to 0\end{array}\right\} \quad \mathrm{EE}_l \to \infty$$

Is Massive MIMO Incredibly Energy Efficient?

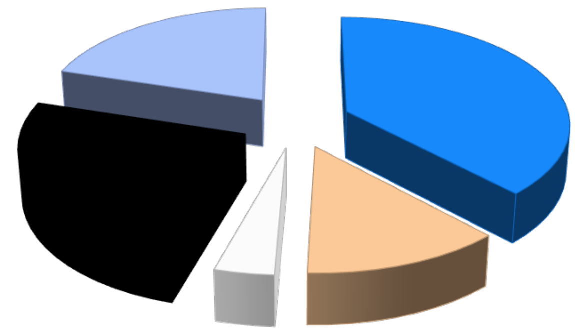
Yes, in terms of bringing down the radiated transmit power

But not all consumed power is radiated!

OPTIMIZING A NETWORK FOR ENERGY EFFICIENCY

Generic Power Consumption Model

- Many Components Consume Power
  - Radiated transmit power
  - Baseband signal processing (e.g., precoding)
  - Active circuits (e.g., converters, mixers, filters)



- Average Power Consumption Model:

$$\mathrm{APC} = \frac{K}{\eta}\mathbb{E}\{p_{lk}\} + C_{0,0} + \underbrace{C_{0,1}M + C_{1,0}K}_{\textit{Circuit power per transceiver chain}} + C_{1,1}MK$$

Power amplifier (η is efficiency)

Fixed power (control signals, backhaul, load-independent processing)

Cost of digital signal processing (e.g., channel estimation and precoding computation)

***Nonlinear increasing function** of M and K*

Many coefficients: $\eta, C_{\mathrm{i,j}}$ *for different* i, j

Optimizing a Cellular Network for High EE

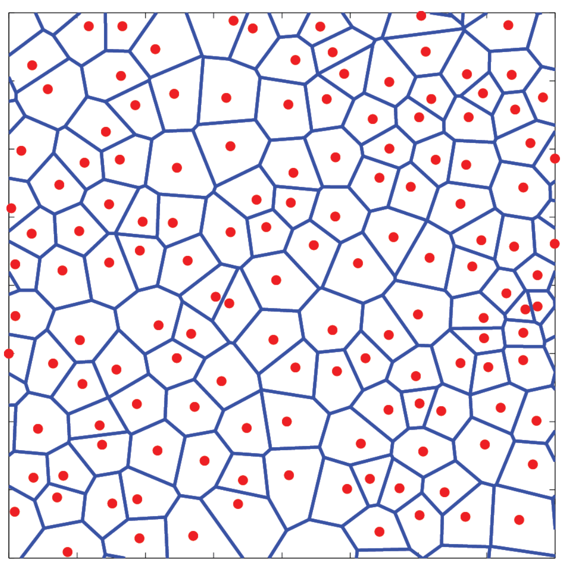
- Clean Slate Network Design
  - Select BS density: λ BSs per km^2
  - Select M and K per cell
  - Asymmetric user load $\rightarrow$ asymmetric deployment

Spatial Point Processes

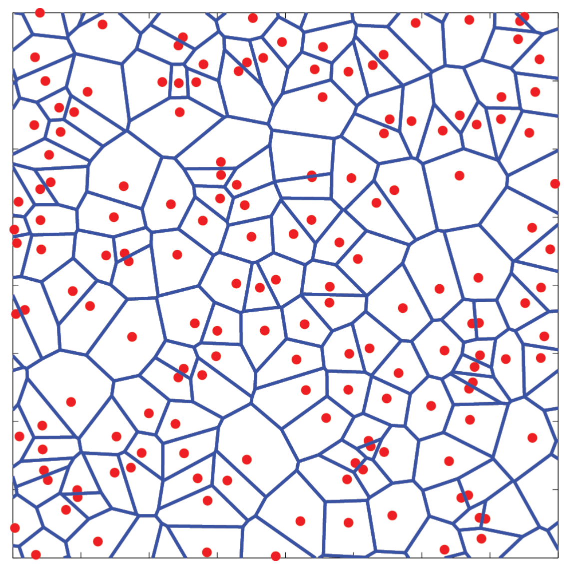
Tractable way to model randomness

Poisson point process (PPP):
Po(λA) *BSs in area of size A km^2*

Random independent deployment:
Lower bound on practical performance



Real BS deployment



Poisson point deployment

Source: Andrews et al. "A Tractable Approach to Coverage and Rate in Cellular Networks"

Average Uplink Spectral Efficiency

Assumptions

BSs distributed as PPP: λ BS/km²

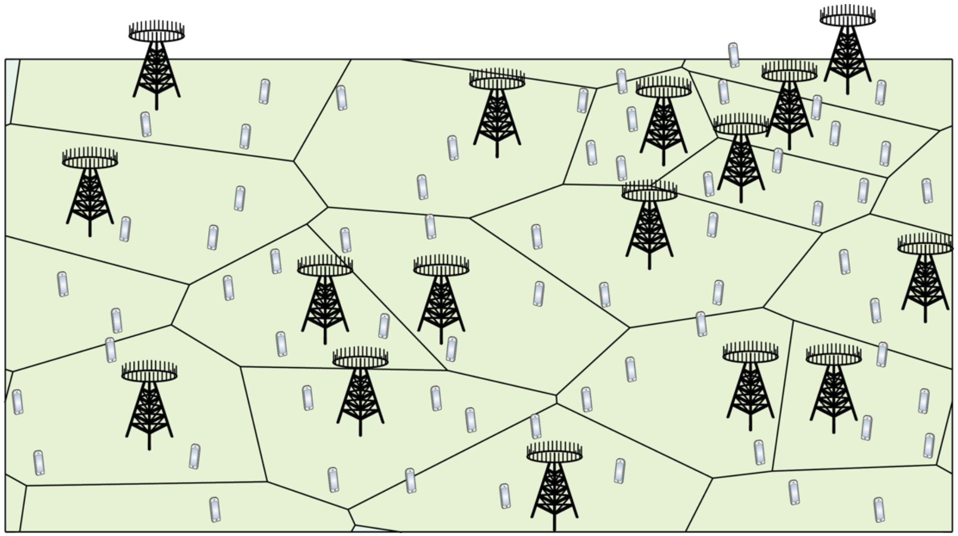
M antennas per BS, K users per cell

Random pilot allocation: $\tau_p = fK$

Statistical channel inversion: p/β_{lk}^{l}

Pathloss over noise:

$$\beta_{lk}^{j} = \omega^{-1}(\text{distance [km]})^{-\alpha}$$



Power per user: $\mathbb{E}\left\{\frac{p}{\beta_{lk}^{l}}\right\} = p\omega \frac{\Gamma(\alpha/2-1)}{(\pi\lambda)^{\alpha/2}}$

Prop. 1: Lower Bound on Average Uplink SE with MR

$$\overline{\text{SE}} = \left(1 - \frac{fK}{\tau_c}\right)\log_2(1+\text{SINR})$$

$$\text{SINR} = \frac{M}{\left(K+\frac{1}{p}\right)\left(1+\frac{2}{f(\alpha-2)}+\frac{1}{p}\right)+\frac{2K}{\alpha-2}\left(1+\frac{1}{p}\right)+\frac{K}{f}\left(\frac{4}{(\alpha-2)^2}+\frac{1}{\alpha-1}\right)+\frac{M}{f(\alpha-1)}}$$

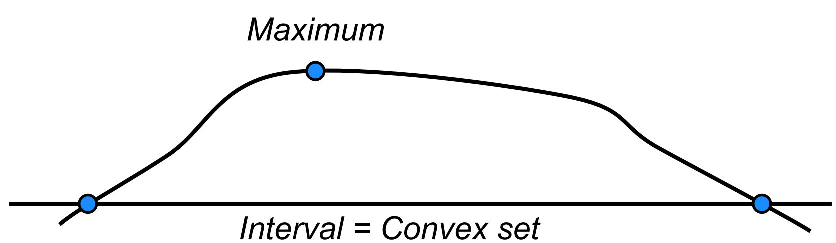
Maximizing Energy Efficiency

$$
\begin{array}{ll}
\underset{M,K,p,\lambda,f}{\text{maximize}} & \dfrac{B \cdot K\left(1-\dfrac{fK}{\tau_c}\right)\log_2(1+\mathrm{SINR})}{\mathrm{APC}} \\
\text{subject to} & \mathrm{SINR} \geq \gamma
\end{array}
$$

- Average SINR constraint γ needed to not get too low SE
- Is the solution small cells (high λ) or Massive MIMO (high M)?

- Main Properties
  1. Can pick f to satisfy SINR constraint
  2. By setting $p = p_0\lambda$, the EE is increasing in λ
  3. Quasi-concave function w.r.t. M and K

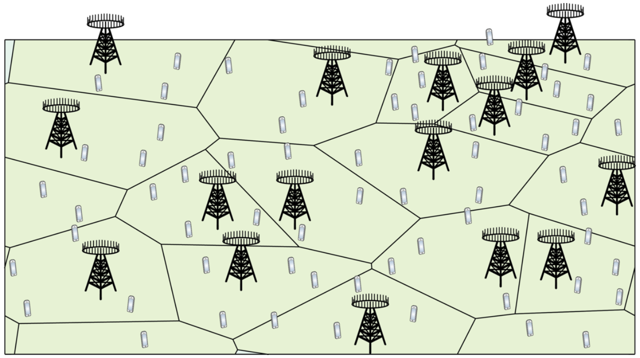
Possible to solve the problem numerically



SIMULATION RESULTS

Simulation Parameters

Parameter	Symbol	Value
Coherence interval	τ_c	400
Pathloss exponent	α	3.76
Pathloss over noise at 1 km	ω	33 dBm
Amplifier efficiency	η	0.39
Bandwidth	B	20 MHz
Static power	$\mathcal{C}_{0,0}$	10 W
Circuit power per active user	$\mathcal{C}_{1,0}$	0.1 W
Circuit power per BS antenna	$\mathcal{C}_{0,1}$	1 W
Signal processing coefficient	$\mathcal{C}_{1,1}$	3.12 mW



I have published simulation code to enable testing of other values!

Impact of BS Density

Simulation

Different BS densities

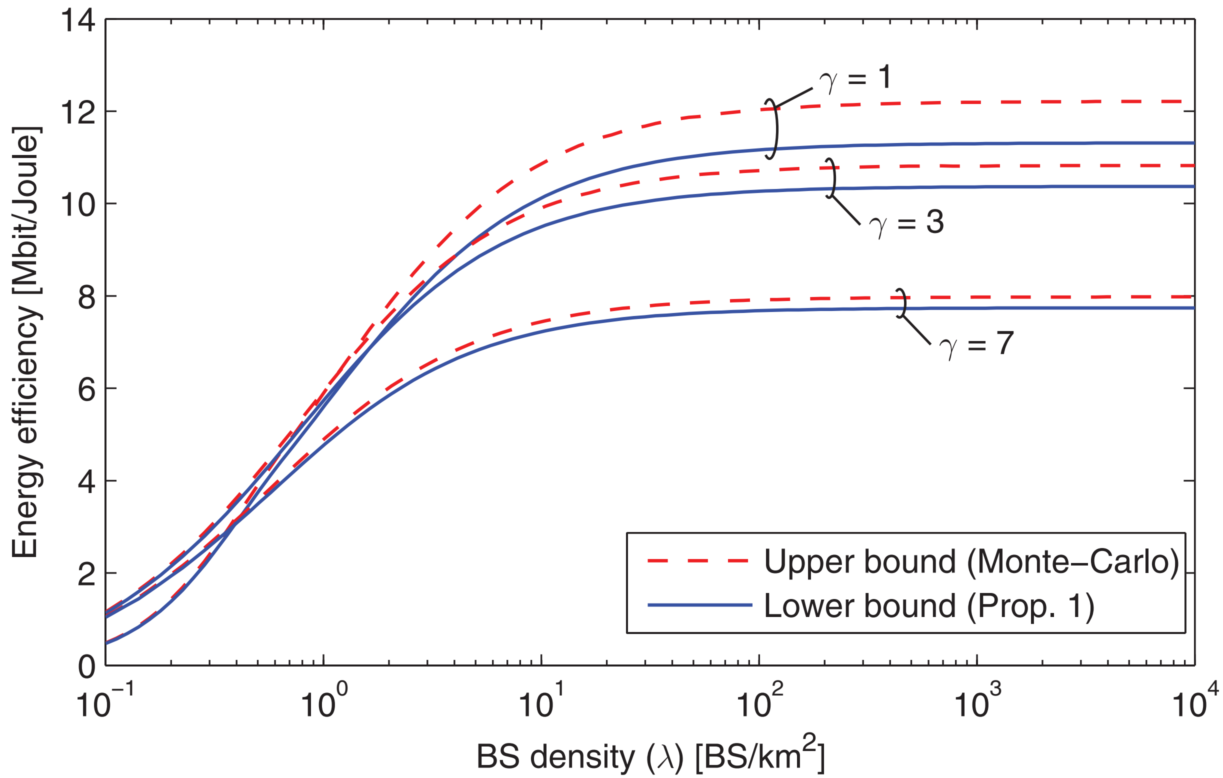
Other variables optimized

Observations

Lower bound is tight

Higher EE with lower γ

EE increases with λ



Saturation Property

EE gain from small cells saturates at $\lambda = 10$

This is satisfied in most urban deployments (300 m between BSs)

We can safely let $\lambda \to \infty$ to simplify analysis

Optimal Number of Antennas and Users

Real-valued Optimization

Optimal $K \in \mathbb{R}$ found in closed-form for fixed M/K

Optimal $M \in \mathbb{R}$ found in closed-form for fixed K

Alternating optimization reaches global maximum

Properties: *Optimal K and M*

$\searrow$: *Decrease as $C_{0,1}$, $C_{1,0}$, and $C_{1,1}$ increase*

$\nearrow$: *Increase as $C_{0,0}$ increases*

Intuition: Activate more hardware if the relative cost is small

Impact of Number of Antennas and Users

Simulation

Optimized f, λ, p

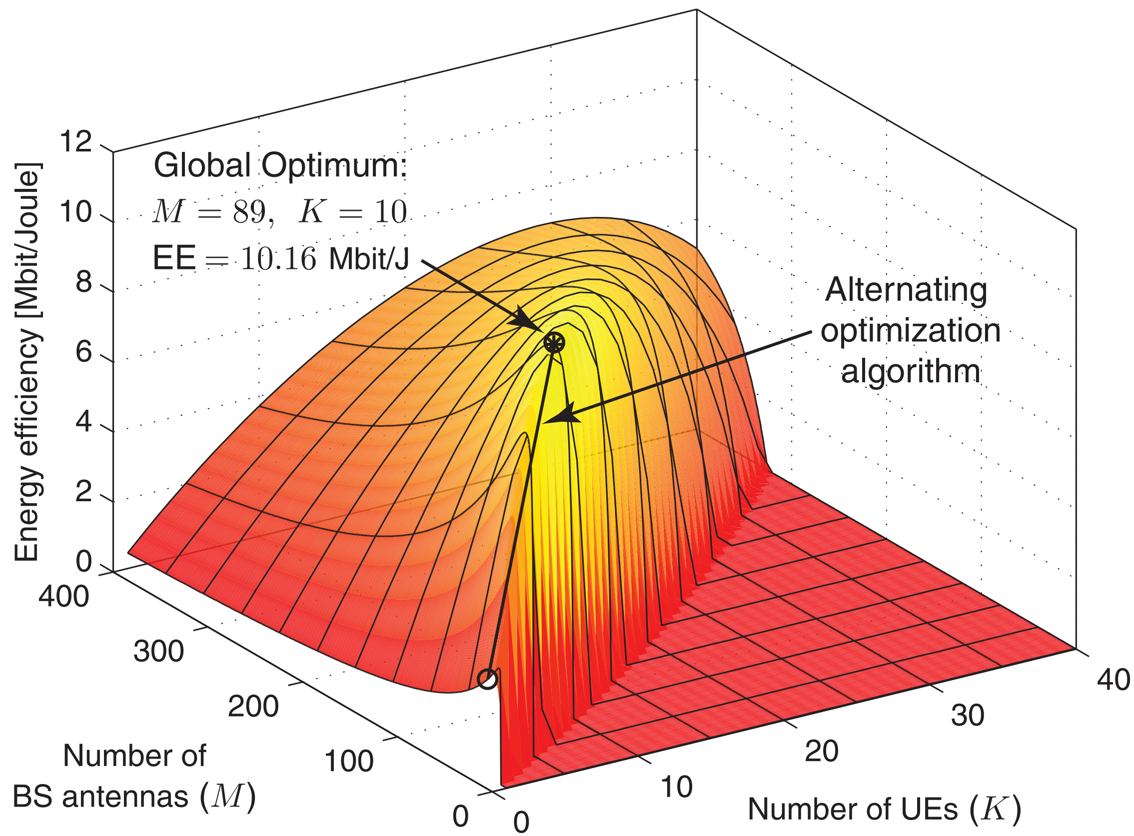
SINR constraint: $\gamma = 3$

Observations

Optimal: $M = 89$, $K = 10$

Massive MIMO with reuse factor $f \approx 7$

Many good solutions



Why is Massive MIMO Energy Efficient?

Interference suppression: Improve SINR, not only SNR as with small cells

Sharing cost: Fixed circuit power costs are shared

Optimization with Given User Density

- User Density
  - So far: K and λ design variables
  - Density: λK users per km^2
  - Heterogeneous user distribution

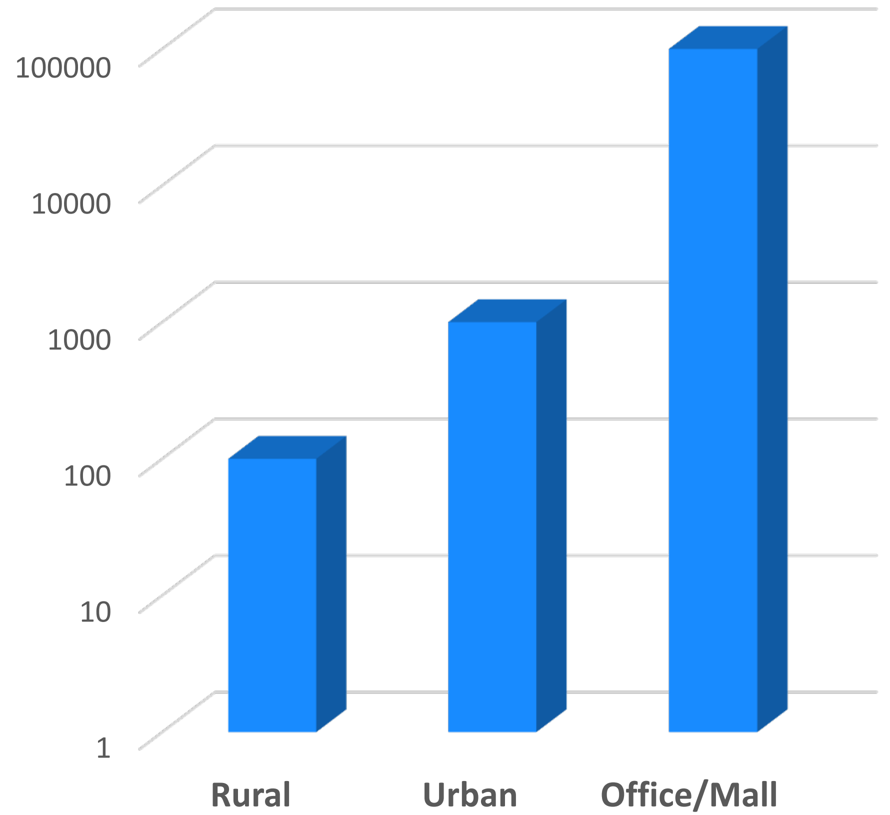
Can we Optimize this Density?

Increase: No, cannot "create" users

Decrease: Yes, by scheduling

Practical User Densities

Rural:	10^2 *per km²*
Urban:	10^3 *per km²*
Office/Mall:	10^5 *per km²*



Source: METIS, "Deliverable D1.1: Scenarios, requirements and KPIs for 5G mobile and wireless system"

Impact of User Density

Simulation

Fixed user density μ users/km²

Rural: $\mu = 10^2$, Malls: $\mu = 10^5$

EE maximization, constraint $K\lambda = \mu$

Low User Density

Many cells with $K \approx 1$

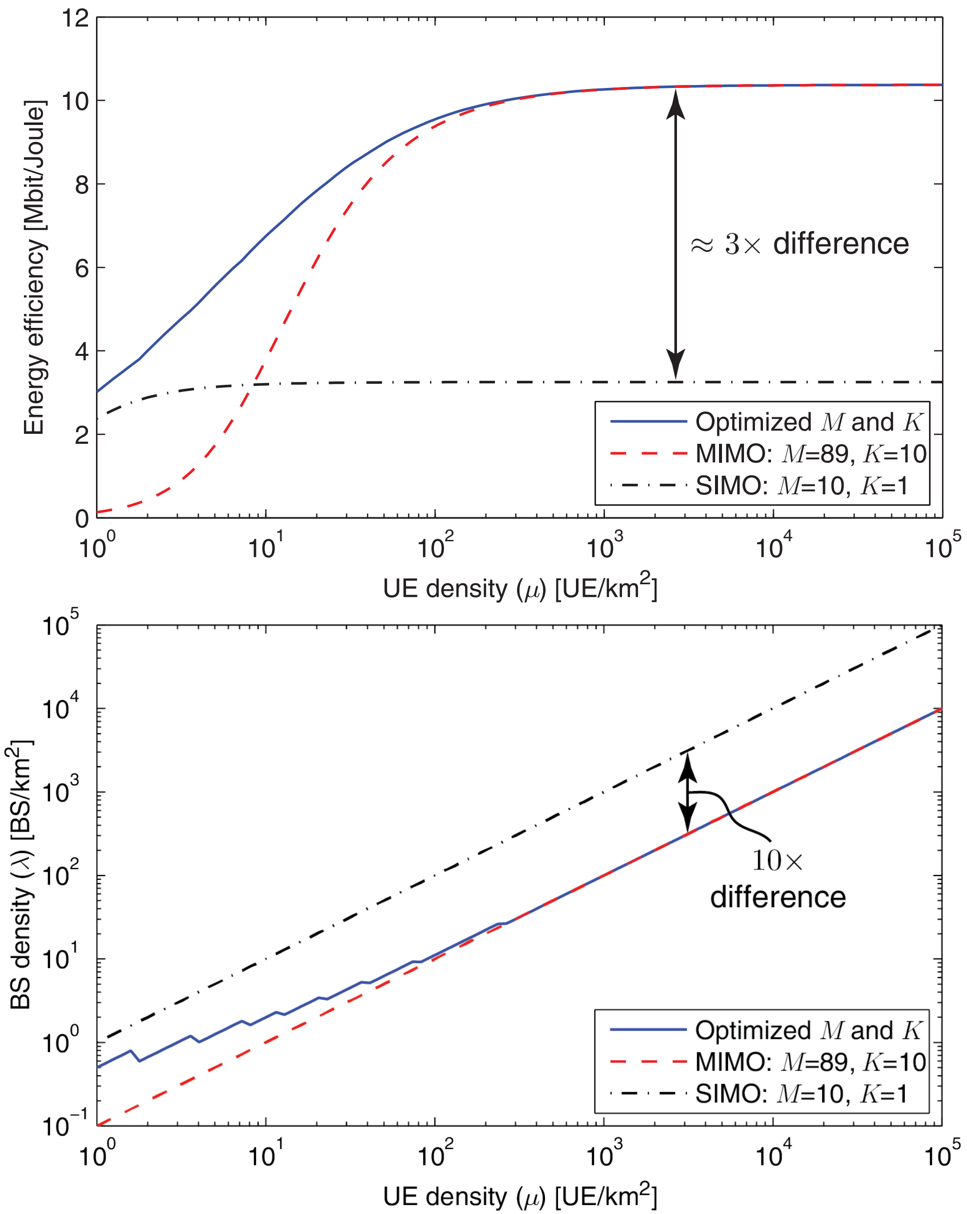
Most important to reduce pathloss

High User Density

Massive MIMO is optimal

Saturation for $\mu \geq 100$:
Covers both rural and shopping malls

Share circuit power and cost over users



PART 4: SUMMARY

Summary

- Transmit Power Scaling Law
  - Reduced as $1/\sqrt{M}$ per user, but sum transmit power might increase
  - Reduced as $1/M^{3/2}$ per BS antenna → Use handset technology?

- Designing Networks for Energy Efficiency
  - Large cells: First step is to reduce cell size
  - Smaller cells: Transmit power only a small part → Use Massive MIMO
  - Intuition: Suppress interference, share circuit power over many users
  - Non-universal pilot reuse is important in random deployments
  - Several Mbit/Joule achieved without coordination

Part 5

Impact of Hardware Impairments

TRANSCEIVER HARDWARE IMPAIRMENTS

Many Antennas and Transceiver Chains

- Many Antenna Elements
  - LTE 8-MIMO: $3 \cdot 8 \cdot 10 = 240$ antennas

    But only 24 transceiver chains!
  - Massive MIMO = M transceiver chains

3 sectors, 4 arrays/sector, 20 antennas/array

- End-to-End Channels
  - Wireless propagation channel
  - Transceiver hardware
  - Simple model:

Can We Afford M High-Grade Transceiver Chains?

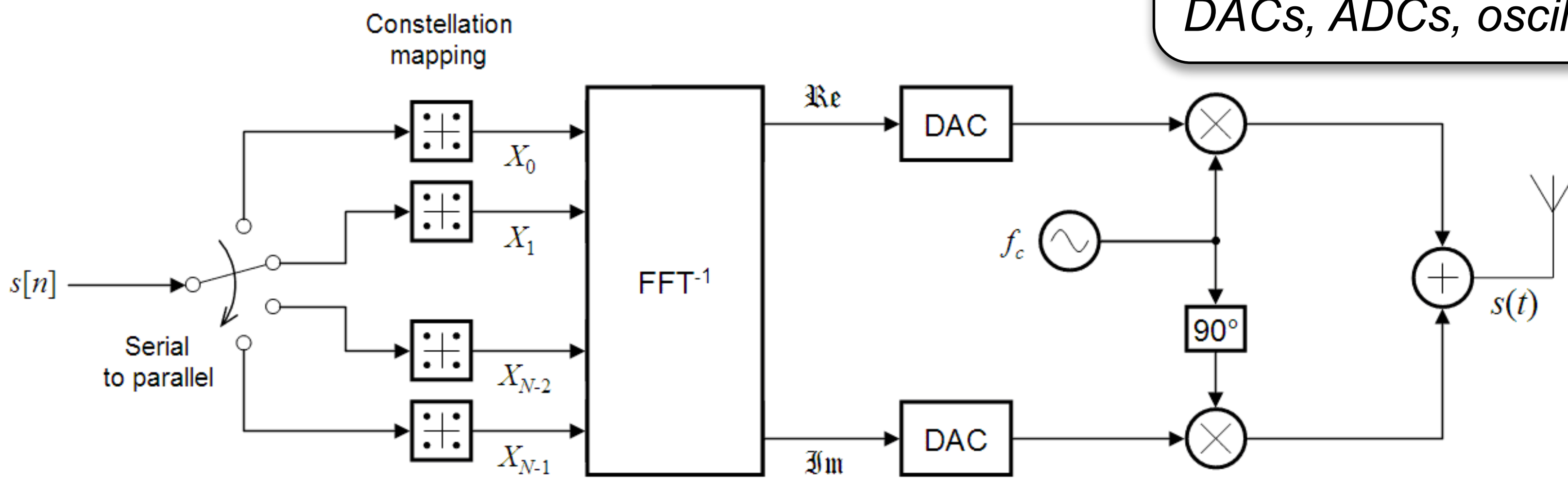
Can Massive MIMO utilize the hardware components more efficiently?

Orthogonal frequency-division multiplexing (OFDM)

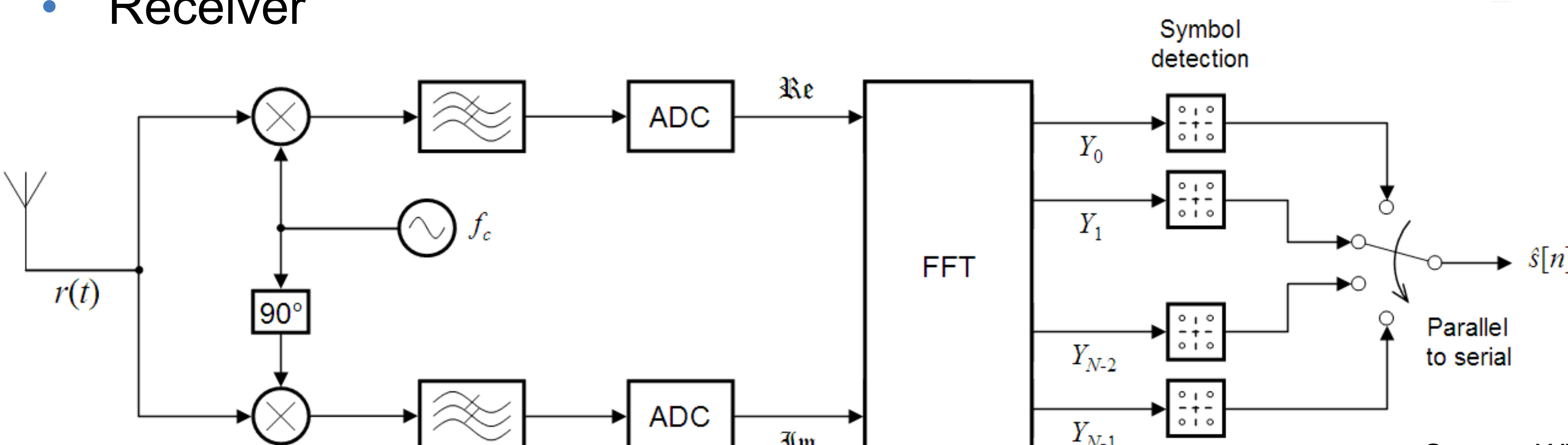
Main Components

Filters, I/Q mixers, DACs, ADCs, oscillators

- Transmitter



- Receiver



Source: Wikipedia

Modeling of Hardware Impairment (1)

- Real Transceivers have Hardware Impairments
  - Phase noise in oscillators
  - I/Q imbalance in mixers
  - Quantization noise in ADCs
  - Non-linearities in amplifiers, etc.

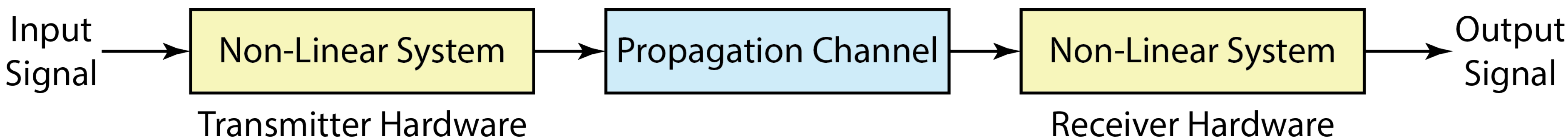
More impairments = Lower price, less power, smaller size

- Impairment Modeling
  - Each impairment can be modeled (for given hardware, waveform etc.)
  - Calibration algorithms can be used to mitigate impairments
  - Challenging to model the residual impairment in detail

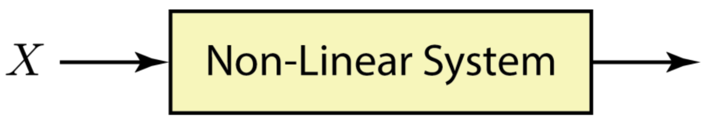
Only impact on communication performance is of importance to us

Modeling of Hardware Impairment (2)

- High-Level Hardware Model:



- Gaussian input signal:



Correlation: $c = \mathbb{E}\{YX^*\}/\mathbb{E}\{X^2\}$

$$Y = cX + \xi$$

Uncorrelated distortion

- Capacity lower bound over a non-linear system:

$$C \geq \log_2\left(1 + \frac{|c|^2\mathbb{E}\{X^2\}}{\mathbb{E}\{Y^2\} - |c|^2\mathbb{E}\{X^2\}}\right) = \log_2\left(1 + \frac{|c|^2\mathbb{E}\{X^2\}}{(1-|c|^2)\mathbb{E}\{X^2\}}\right)$$

Calibration: $\mathbb{E}\{Y^2\} = \mathbb{E}\{X^2\}$

Main hardware characteristics affecting SE are captured by c!

Classical Impact of Hardware Impairments

- Impact on Point-to-Point MIMO
  - Low SNR: Negligible impact on spectral efficiency
  - High SNR: Fundamental upper limit

Error Vector Magnitude

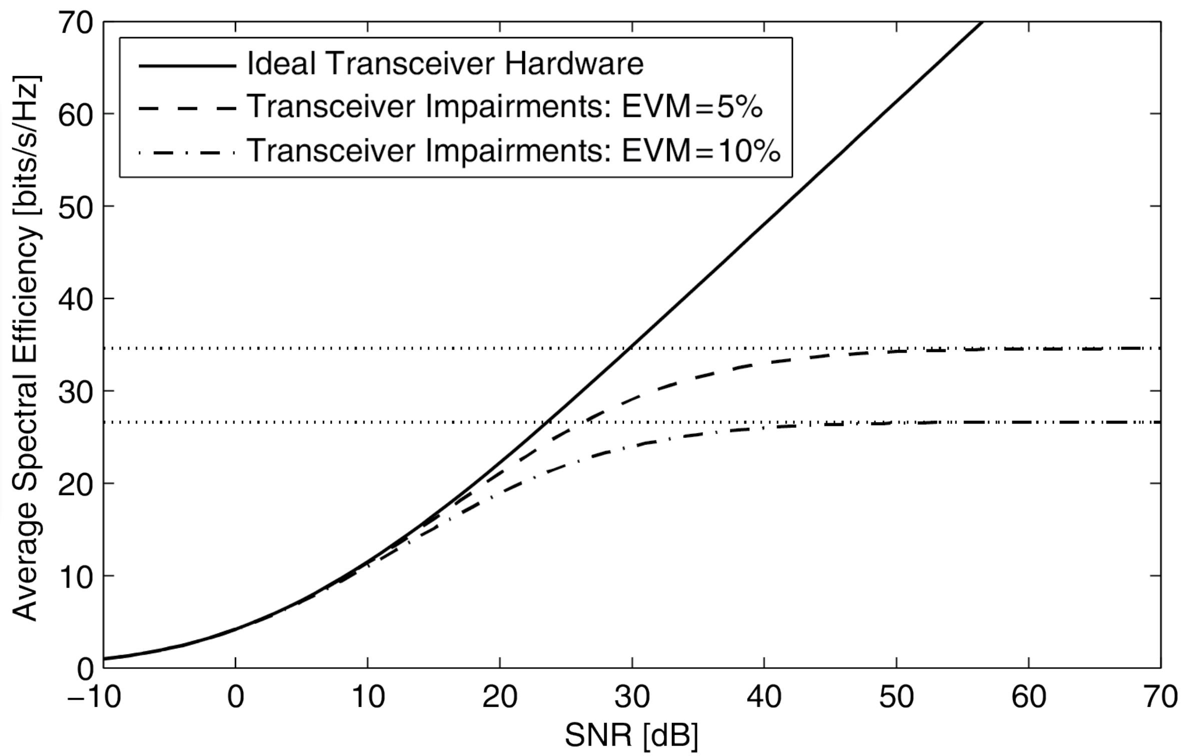
$$\text{EVM} = \frac{\text{Distortion magnitude}}{\text{Signal magnitude}}$$
$$= \frac{1 - |c|^2}{|c|^2}$$

Distortion scales with signal power

LTE EVM limits: 8%-17.5%

What about large M regime?

Large or small impact?



Example: 4x4 point-to-point MIMO, i.i.d. Rayleigh fading

HARDWARE IMPAIRMENTS IN MASSIVE MIMO

Distortion Noise: Definition and Interpretation

- Uplink Signal (conventional):

$$\mathbf{y} = \sum_k \mathbf{h}_k s_k + \mathbf{n}$$

- Uplink Signal (with impairments):

$$\mathbf{y} = c^{\mathrm{rx}} \sum_k \mathbf{h}_k (c_k^{\mathrm{tx}} s_k + \xi_k^{\mathrm{tx}}) + \boldsymbol{\xi}^{\mathrm{rx}} + \mathbf{n}$$

Gain losses

Transmitter distortion

Receiver distortion

Distortion Model

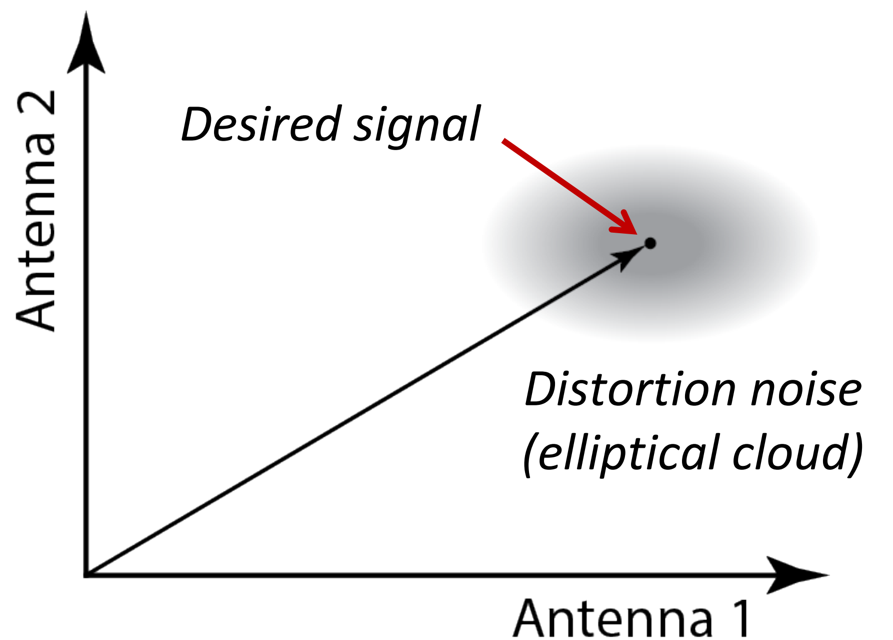
Gaussian distributed

Independent between users and antennas

Error Vector Magnitude

(at transmitter)

$$\mathrm{EVM}^{\mathrm{tx}} = \frac{\sqrt{\mathbb{E}\{|\xi_k^{\mathrm{tx}}|^2\}}}{\sqrt{\mathbb{E}\{|c_k^{\mathrm{tx}} s_k|^2\}}}$$



What is the Impact of Hardware Distortion?

Uplink Single-User Scenario

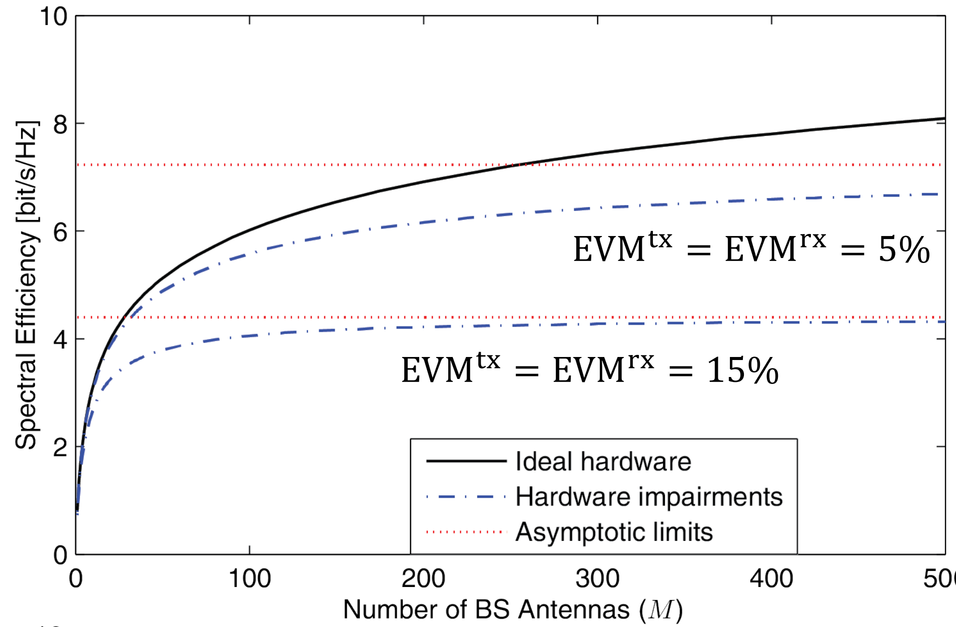
Rayleigh fading, SNR = 5 dB

Observations

Ideal: $\mathrm{SE} = \mathcal{O}(\log M)$

Non-ideal: Asymptotic limits

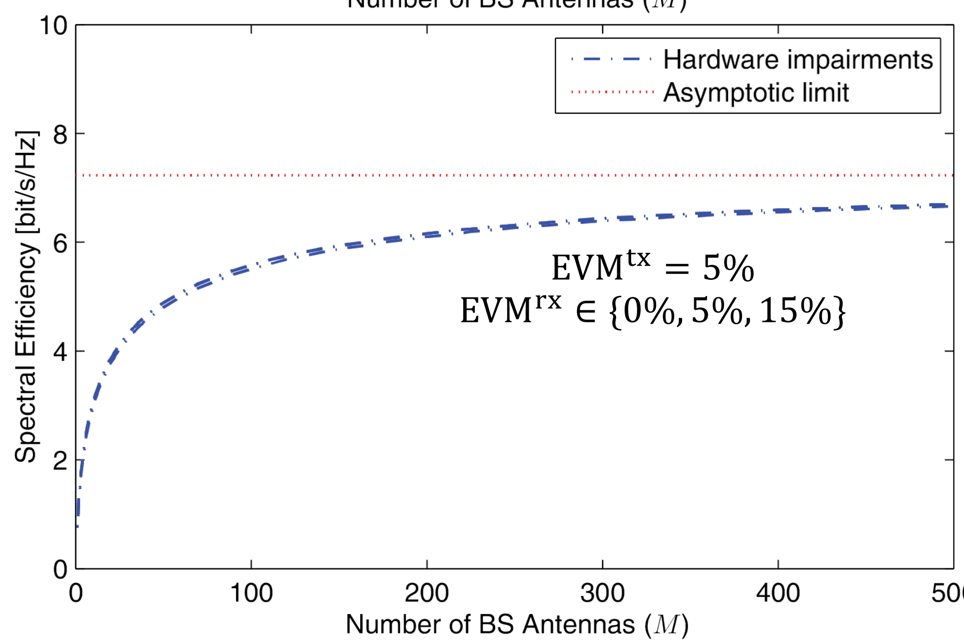
Higher EVM → Lower limit



Observations

Impairments caused by user device determine the limit

Distortion noise caused by BS averages out as $M \to \infty$ *(cf. inter-user interference)*



Multi-Cell Scenario with Distortion Noise

Uplink Multi-Cell Scenario

Rayleigh fading, SNR = 5 dB

$K = 8$ users per cell

MR detection

Hardware Scaling Law

If BS distortion variance increases as M^{κ} for $\kappa \leq 1/2$:

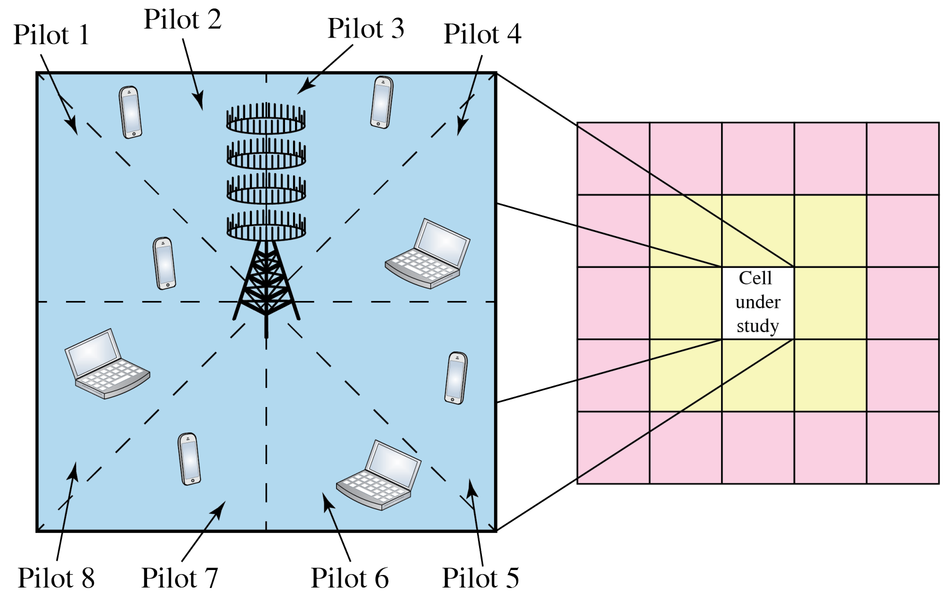
SE will not go zero as $M \rightarrow \infty$

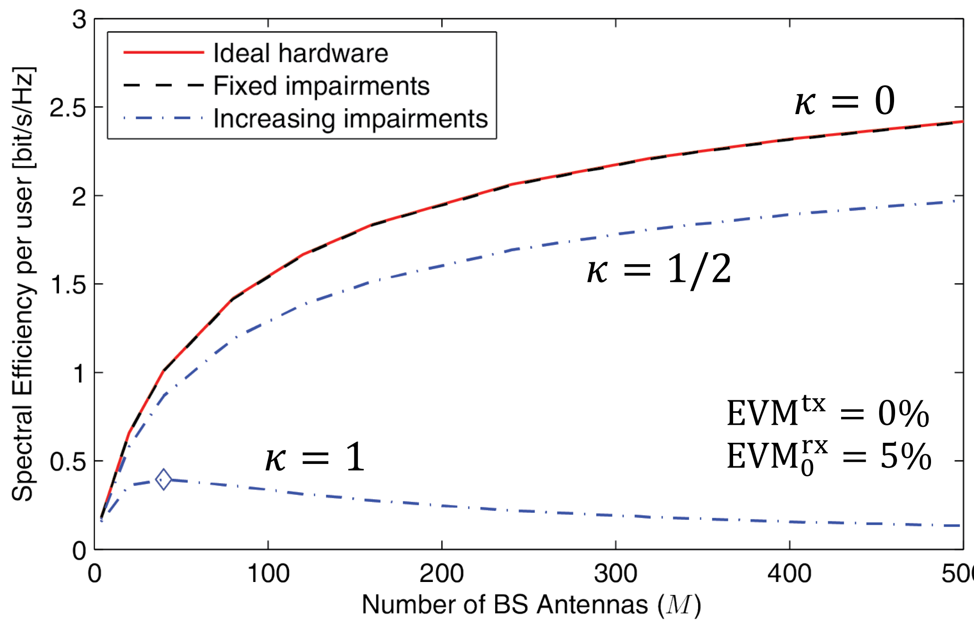
Can be proved rigorously!

Observations

Small loss if law is followed

Otherwise large loss!





Utilizing the Hardware Scaling Law

- Massive MIMO can use Lower-Grade Hardware
  - Reduced cost, power consumption, and size

- Example: Analog-to-Digital Converter (ADC)
  - One b-bit ADC per Transceiver Chain

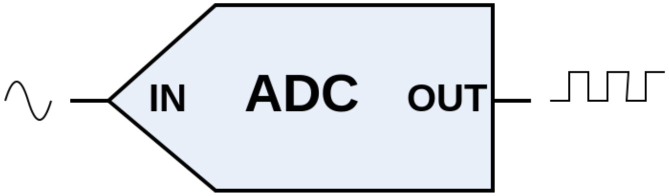


Image source: Wikipedia

  - Adds quantization noise roughly proportional to 2^{-2b}:

$$\sqrt{M} = c_0 \cdot 2^{-2b} \Longrightarrow b = \frac{1}{2}\log_2(c_0) - \frac{1}{4}\log_2(M)$$

Ex: $M = 256$ requires 2 fewer bits than $M = 1$ (even 1-bit ADCs possible)

  - Circuit power roughly proportional to 2^{2b}:

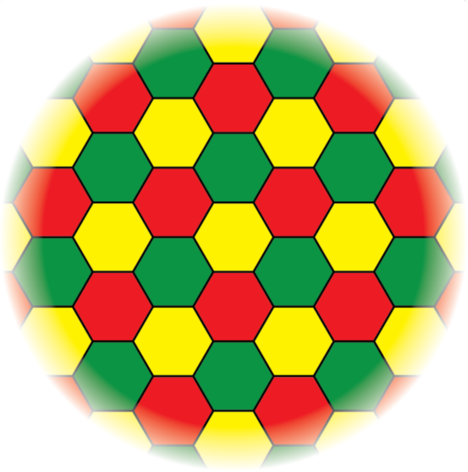
Ex: Power of M ADCs can scale as $\sqrt{M}$ rather than M

Interference Visibility Range

- Only Remaining Interference as $M \to \infty$:
  - Pilot contamination (reuse of pilot resources)
  - Hardware impairments (at user devices)

- Hardware Distortion is Self-Interference
  - Limits the visibility of inter-user interference

Strong self-interference



Weak self-interference

No reason to suppress inter-user interference below self-interference!

PART 5: SUMMARY

Summary

- Any Transceiver is Subject to Hardware Impairments
  - Massive MIMO is resilient to such imperfections
  - Distortion variance at BS may increase as $\sqrt{M}$
  - High-grade BS hardware is not required!
  - User hardware quality is the fundamental limitation

- Further Remarks
  - Analysis with more detailed hardware models show same behavior
  - Phase noise is not worse than in small MIMO systems
  - Reduced transmit power and relaxed impairment constraints
    $\rightarrow$ New compact transceiver designs?

Open Problems

More important things than "pilot contamination"!

- Make Massive MIMO work in FDD mode
  - Long-standing challenge. Is it practically feasible to exploit sparsity?
- Channel measurements, channel modeling, traffic modeling
  - Required for system level simulations
- Implementation-aware algorithmic design
  - Implement ZF with MR-like complexity. Utilize low-resolution hardware.
- Cross-layer design
  - Scalable protocols for random access, control signaling, scheduling
- New deployment characteristics
  - Multi-antenna users, distributed arrays, cell-free (network MIMO)

Bringing an Extraordinary Technology to Reality

- FP7 MAMMOET project (Massive MIMO for Efficient Transmission)
  - Bridge gap between “theoretical and conceptual” Massive MIMO
  - Develop: Flexible, effective and efficient solutions

WP4 Validation and proof-of-concept

WP2 Efficient FE solutions (IC solutions, Comp/Calibration)

WP3 Baseband Solutions (Algorithms, Architectures & Design)

WP1 System approach, scenarios and requirements

MAMMOET • Massive MIMO •

ERICSSON · TECHNIKON · LINKÖPING UNIVERSITY · imec · Infineon · KU LEUVEN · LUND UNIVERSITY · Telefónica

Thanks to my Collaborators

- Erik G. Larsson (LiU, Sweden)
- Hei Victor Cheng, Antonios Pitarokoilis, Marcus Karlsson (LiU)
- Xueru Li (previous visitor at LiU)

- Mérouane Debbah, Marios Kountouris (Huawei, France)
- Luca Sanguinetti (CentraleSupélec and University of Pisa, Italy)

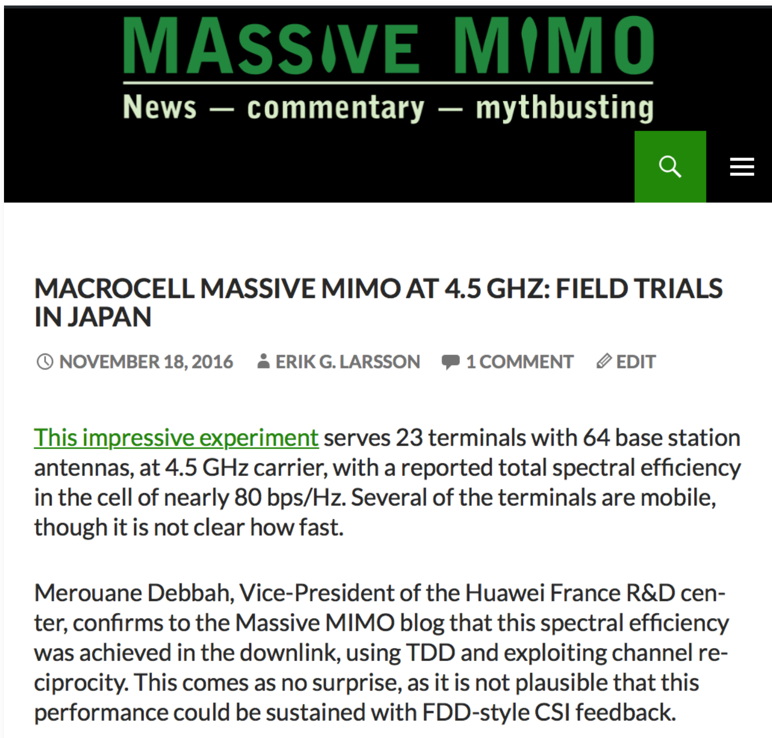
- Thomas L. Marzetta (Bell Labs, USA)
- Jakob Hoydis (former Bell Labs)

- Michail Matthaiou (Queen's University Belfast, UK)

- Björn Ottersten (KTH, Sweden)
- Per Zetterberg (KTH)
- Mats Bengtsson (KTH)

Blog and Books

- Many Massive MIMO Resources
  - Info Point (https://massivemimo.eu): Library with research papers
  - LinkedIn Group: “Massive for 5G”



MASSIVE MIMO

News — commentary — mythbusting

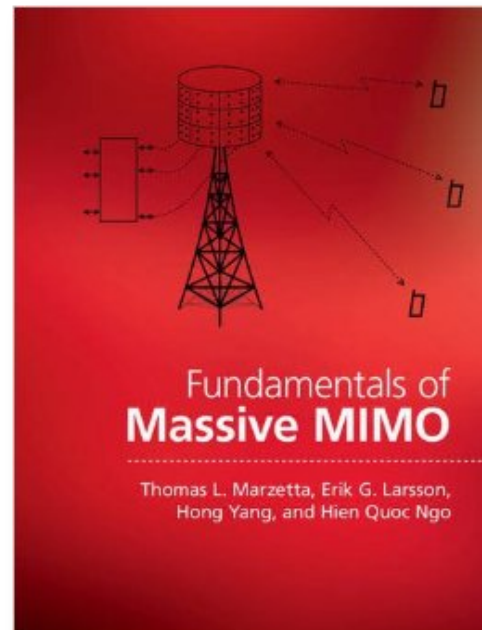
MACROCELL MASSIVE MIMO AT 4.5 GHZ: FIELD TRIALS IN JAPAN

NOVEMBER 18, 2016 ERIK G. LARSSON 1 COMMENT EDIT

This impressive experiment serves 23 terminals with 64 base station antennas, at 4.5 GHz carrier, with a reported total spectral efficiency in the cell of nearly 80 bps/Hz. Several of the terminals are mobile, though it is not clear how fast.

Merouane Debbah, Vice-President of the Huawei France R&D center, confirms to the Massive MIMO blog that this spectral efficiency was achieved in the downlink, using TDD and exploiting channel reciprocity. This comes as no surprise, as it is not plausible that this performance could be sustained with FDD-style CSI feedback.

www.massive-mimo.net



Book by Marzetta, Larsson, et al.

***My own book**: Available 2017*

Thank you!

Emil Björnson

LINKÖPING UNIVERSITY

Slides and papers available online:

http://www.commsys.isy.liu.se/en/staff/emibj29

Key References

Key References (1/4)

Seminal and Overview Papers

1. T. L. Marzetta, "*Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas*," IEEE Trans. Wireless Communications, 2010. **IEEE W.R.G. Baker Prize Paper Award**
2. J. Hoydis, S. ten Brink, M. Debbah, "*Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?*," IEEE Journal on Selected Areas in Communications, 2013. **IEEE Leonard G. Abraham Prize**
3. H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "*Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems*," IEEE Trans. Commun., 2013. **IEEE Stephen O. Rrice Prize**
4. F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "*Scaling up MIMO: Opportunities and Challenges with Very Large Arrays*," IEEE Signal Proces. Mag., 2013.
5. J. Hoydis, K. Hosseini, S. ten Brink, and M. Debbah, "*Making Smart Use of Excess Antennas: Massive MIMO, Small Cells, and TDD*," Bell Labs Technical Journal, 2013.
6. E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "*Massive MIMO for Next Generation Wireless Systems*," IEEE Commun. Mag., 2014.
7. T. L. Marzetta, "*Massive MIMO: An Introduction*," Bell Labs Technical Journal, 2015
8. E. Björnson, E. G. Larsson, T. L. Marzetta, "*Massive MIMO: 10 Myths and One Critical Question*," IEEE Commun. Mag., 2016.
9. T. L. Marzetta, E. G. Larsson, H. Yang, H. Q. Ngo, "*Fundamentals of Massive MIMO*," Cambridge University Press, 2016.

Key References (2/4)

Spectral Efficiency

1. J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "*Pilot Contamination and Precoding in Multi-cell TDD Systems*," IEEE Trans. Wireless Commun., 2011.
2. H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "*Achieving 'Massive MIMO' Spectral Efficiency with a Not-so-Large Number of Antennas*," IEEE Trans. Wireless Communications, 2012.
3. A. Adhikary, N. Junyoung, J.-Y. Ahn, G. Caire, "*Joint Spatial Division and Multiplexing—The Large-Scale Array Regime*," IEEE Trans. Information Theory, 2013.
4. E. Björnson and E. Jorswieck, "*Optimal Resource Allocation in Coordinated Multi-cell Systems*," Foundations and Trends in Communications and Information Theory, 2013.
5. H. Yang and T. Marzetta, "*A Macro Cellular Wireless Network with Uniformly High User Throughputs*," in *Proc. IEEE VTC-Fall*, 2014.
6. E. Björnson, E. G. Larsson, M. Debbah, "*Massive MIMO for Maximal Spectral Efficiency: How Many Users and Pilots Should Be Allocated?*," IEEE Trans. Wireless Communications, 2016.
7. E. Björnson, E. Jorswieck, M. Debbah, B. Ottersten, "*Multi-Objective Signal Processing Optimization: The Way to Balance Conflicting Metrics in 5G Systems*," IEEE Signal Processing Magazine, 2014.
8. E. Björnson, J. Hoydis, L. Sanguinetti, "*Pilot Contamination is Not a Fundamental Asymptotic Limitation in Massive MIMO*," Submitted to ICC 2017.

Key References (3/4)

Energy Efficiency

1. G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "*Energy-efficient wireless communications: tutorial, survey, and open issues*," IEEE Wireless Commun., 2011.
2. E. Björnson, M. Kountouris, M. Debbah, "*Massive MIMO and Small Cells: Improving Energy Efficiency by Optimal Soft-Cell Coordination*," ICT, 2013.
3. H. Yang and T. Marzetta, "*Total energy efficiency of cellular large scale antenna system multiple access mobile networks*," in Proc. IEEE OnlineGreenComm, 2013.
4. W. Liu, S. Han, C. Yang, and C. Sun, "*Massive MIMO or Small Cell Network: Who is More Energy Efficient?*," in Prpc. IEEE WCNCW, 2013.
5. D. Ha, K. Lee, and J. Kang, "*Energy Efficiency Analysis with Circuit Power Consumption in Massive MIMO Systems*," in Proc. IEEE PIMRC, 2013.
6. S. K. Mohammed, "*Impact of Transceiver Power Consumption on the Energy Efficiency of Zero-Forcing Detector in Massive MIMO Systems*," IEEE Trans. Communications, 2014.
7. E. Björnson, L. Sanguinetti, J. Hoydis, M. Debbah, "*Optimal Design of Energy-Efficient Multi-User MIMO Systems: Is Massive MIMO the Answer?*," IEEE Trans. Wireless Commun., 2015.
8. E. Björnson, L. Sanguinetti, M. Kountouris, "*Deploying Dense Networks for Maximal Energy Efficiency: Small Cells Meet Massive MIMO*," IEEE Journal on Selected Areas in Communications, 2016.

Key References (4/4)

Hardware Efficiency

1. M. Wenk, *MIMO-OFDM Testbed: Challenges, Implementations, and Measurement Results*, Series in microelectronics. Hartung-Gorre, 2010.
2. W. Zhang, "*A General Framework for Transmission with Transceiver Distortion and Some Applications*," IEEE Trans. Communications, 2012.
3. E. Björnson, P. Zetterberg, M. Bengtsson, B. Ottersten, "*Capacity Limits and Multiplexing Gains of MIMO Channels with Transceiver Impairments*," IEEE Communications Letters, 2013.
4. S. K. Mohammed, E. G. Larsson, "*Per-Antenna Constant Envelope Precoding for Large Multi-User MIMO Systems*," IEEE Trans. Communications, 2013.
5. E. Björnson, J. Hoydis, M. Kountouris, M. Debbah, "*Massive MIMO Systems with Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits*," IEEE Trans. Information Theory, 2014.
6. A. Pitarokoilis, S. K. Mohammed, E. G. Larsson, "*Uplink Performance of Time-Reversal MRC in Massive MIMO Systems Subject to Phase Noise*," IEEE Trans. Wireless Communications, 2014.
7. U. Gustavsson, C. Sanchéz-Perez, T. Eriksson, F. Athley, G. Durisi, P. Landin, K. Hausmair, C. Fager, L. Svensson, "*On the Impact of Hardware Impairments on Massive MIMO*," Globecom 2014.
8. E. Björnson, M. Matthaiou, M. Debbah, "*Massive MIMO with Non-Ideal Arbitrary Arrays: Hardware Scaling Laws and Circuit-Aware Design*," IEEE Trans. Wireless Commun, 2015.