Massive MIMO:
Fundamentals, Opportunities and Challenges

Erik G. Larsson

June 9, 2013

Div. of Communication Systems
Dept. of Electrical Engineering (ISY)
Linköping University
Linköping, Sweden

www.commsys.isy.liu.se
With thanks to my team and collaborators:

- Hien Q. Ngo (LiU, Sweden)
- Antonios Pitarokoilis (LiU)
- Hei Victor Cheng (LiU)
- Daniel Persson (LiU)
- Fredrik Rusek (Lund, Sweden)
- Ove Edfors (Lund)
- Buon Kiong Lau (Lund)
- Fredrik Tufvesson (Lund)
- Thomas L. Marzetta (Bell Labs/Alcatel-Lucent, USA)
- Saif Mohammed (IIT/Delhi, India)
- Christoph Studer (Rice Univ., USA)
Massive MIMO

- Massive multiuser MIMO (MISO):
  - $M \gg K \gg 1$ (think $100 \times 10$ or $500 \times 50$)
  - coherent, but simple, processing
- Potential to dramatically improve rate & reliability
- Potential to drastically scale down TX power
- Not only theory, at least one known testbed ($64 \times \sim 10$)
Large MIMO Deployment Scenarios

- Reduce bulky items (coax)
- Each antenna unit simple (low accuracy)
- Resilience against individual failures (hotswapping)
- Potential economy of scale in manufacturing
- Grid-free powering
Massive MIMO Operation

- Not enough resources for pilots & CSI feedback, so operate in TDD.

- On the uplink,
  - acquire CSI from uplink pilots and/or blindly from data
  - detect symbols
  - \( M \gg K \Rightarrow \) linear processing (MRC, ZF, MMSE) nearly optimal

- On the downlink,
  - use CSI obtained on the uplink
  - make necessary adjustments based on reciprocity calibration
  - apply multiuser MIMO precoding
  - simple precoders desirable (and very good!): MRT, ZF, MMSE, ...

- MRC/MRT operation
  - intracell interference will appear as noise
  - 1 bps/Hz/terminal; \( K \) bps/Hz/terminal total
  - distributed implementation

- ZF/MMSE operation
  - can cancel out intra cell interference
  - computationally more demanding
MRT versus ZF precoding

a) MRT precoding

b) ZF precoding
Massive MIMO Opportunities

- Multiplexing gain $K$, Array power gain $M$ (ideally)

- Rely on the law of large numbers
  - average out fast fading and thermal noise

- Channel hardens $\Rightarrow \|h\|^2/M \approx \text{constant}$
  - no FD scheduling, full BW to all terminals, simple MAC
  - almost no air-interface latency

- In the 1 bps/Hz/t regime, many impairments drown in thermal noise

- $M - K$ unused degrees of freedom (e.g. 500 - 50 = 450)
Massive MIMO Challenges and Questions

- Multiplexing gain can only materialize if channel responses are (nearly) orthogonal.

- Array gain relies on coherency $\Rightarrow$ getting CSI is the main thing.

- HW power consumption (RF) must scale fast enough with $P$.
- HW power consumption (BB) must scale slow enough with $M$.

- Scaling down $P$ $\Rightarrow$ thermal noise eventually limits performance.

- $P$ “large enough,” interference limits performance.
  - intracell interference
  - intercell interference
Limits Imposed by Propagation
“Favorable propagation” in Point-to-Point MIMO

- \( M \times K \) MIMO link \( H \), \( M \geq K \)

![](image)

- **Favorable propagation** (f.p.) if

\[
H^H H \propto I \iff \lambda_1^2 = \ldots = \lambda_K^2
\]

- For \( \|H\|^2 = \text{constant} \), \( \log \left| I + \frac{1}{N_0} H^H H \right| \) max if f.p.

- \( \left( \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \right)^2 \) = extra power needed to use all eigenmodes

- \( H_{mk} \) zero mean & i.i.d., and \( M \gg K \Rightarrow \text{f.p.} \)
Favorable propagation in MU-SIMO

- Favorable propagation if
  \[ \frac{1}{M} g_i^H g_j \approx 0 \text{ for } i \neq j \]

- \( \| g_i \|^2 \) depends on path loss and shadow fading
I.i.d. channels give favorable propagation
Do we have “favorable propagation” in practice?

- Our partners at Lund Univ., Sweden have conducted unique measurements [RPL2013, GERT2011, GTER2012].
- Indoor 128-ant. (4x16 dual-pol.) array. 3 users indoor, 3 outdoor.
- 2.6 GHz CF, 50 MHz BW, 100 snapshots (10m).
- Normalized to retain only small-scale fading.
Lund measurements, example of results

![Graph showing probability of singular values](image)

- **~26 dB**
- **~7 dB**

**Legend:**
- Blue line: meas 6x128
- Red dashed line: meas 6x6

**Axes:**
- Y-axis: Prob(σ ≤ abscissa)
- X-axis: Squared singular values [dB]

**Annotations:**
- Tail
- No tail
Assumptions on f.p. have substantial support in measurements

- l.i.d. model for small-scaled fading appears reasonable

- More details and models in F. Tufvesson’s tutorial next
Limits Imposed by Noise and Intracell Interference with Linear Processing
Single-Cell (Noise-Limited) Uplink

- \( M \) antennas
- \( K \) terminals
- Power (SNR) per terminal: \( P \)
- i.i.d. Rayleigh fading

- Coherence interval: \( T \) symbols (e.g. \( 2 \times 100kHz \times 1ms = 200 \))
- \( K \) mutually orthogonal pilots of length \( \tau \) (\( K \leq \tau \leq T \))

- MMSE channel estimation
Spectral-energy efficiency tradeoff

- Sum-spectral efficiency bounds [NLM2013]:

\[
R = \begin{cases} 
(1 - \frac{\tau}{T}) K \log_2 \left( 1 + \frac{\tau(M-1)P^2}{\tau(K-1)P^2 + (\tau+K)P+1} \right) & \text{, for MRC} \\
(1 - \frac{\tau}{T}) K \log_2 \left( 1 + \frac{\tau(M-K)P^2}{(\tau+K)P+1} \right) & \text{, for ZF}
\end{cases}
\]

- Energy efficiency: \( \eta \triangleq \frac{R}{P} \)

- SISO system: \( K = 1, M = 1 \)

\[
\max_{P,\tau} \eta, \quad \text{s.t.} \quad R = \text{const.}
\]
UL, $T = 200$, SISO reference, $M = K = 1$

![Graph showing Relative Energy-Efficiency and Spectral-Efficiency for various dB levels.](image-url)
Spectral-energy efficiency tradeoff, cont.

- Sum-spectral efficiency bounds [NLM2013]:

\[
R = \begin{cases} 
(1 - \frac{T}{T}) K \log_2 \left( 1 + \frac{\tau(M-1)P^2}{\tau(K-1)P^2 + (\tau+K)P+1} \right), & \text{for MRC} \\
(1 - \frac{T}{T}) K \log_2 \left( 1 + \frac{\tau(M-K)P^2}{(\tau+K)P+1} \right), & \text{for ZF}
\end{cases}
\]

- Energy efficiency: \( \eta = \frac{R}{P} \)

- SISO system: \( K = 1, M = 1 \)

\[
\max_{P,\tau} \eta, \quad \text{s.t.} \quad R = \text{const.}
\]

- Single-user SIMO system: \( K = 1, M = 100 \)

\[
\max_{P,\tau} \eta, \quad \text{s.t.} \quad R = \text{const.}
\]
UL, $T = 200$, SISO and SU-SIMO

-10 dB

0 dB

10 dB

20 dB

$K=1$, $M=100$ (SU-SIMO)

$K=1$, $M=1$ (SISO)

Relative Energy-Efficiency (%/bits/J)/(bits/J)

Spectral-Efficiency (bits/s/Hz)
Spectral-energy efficiency tradeoff, cont.

- Sum-spectral efficiency bounds [NLM2013]:

\[ R = \begin{cases} 
(1 - \frac{\tau}{T}) K \log_2 \left( 1 + \frac{\tau (M-1) P^2}{\tau (K-1) P^2 + (\tau + K) P + 1} \right), & \text{for MRC} \\
(1 - \frac{\tau}{T}) K \log_2 \left( 1 + \frac{\tau (M-K) P^2}{(\tau + K) P + 1} \right), & \text{for ZF} 
\end{cases} \]

- Energy efficiency: \( \eta = \frac{R}{P} \)

- SISO system: \( K = 1, M = 1 \)

\[ \max_{P,\tau} \eta, \quad \text{s.t. } R = \text{const.} \]

- Single-user SIMO system: \( K = 1, M = 100 \)

\[ \max_{P,\tau} \eta, \quad \text{s.t. } R = \text{const.} \]

- Multi-user SIMO system: \( M = 100, K \geq 1 \) adapted

\[ \max_{P,\tau,K} \eta, \quad \text{s.t. } R = \text{const.} \]
UL, $T = 200$, SISO, SU-SIMO & MU-SIMO

The figure illustrates the relative energy-efficiency and spectral-efficiency for different scenarios:

- **SISO (K=1, M=1)**
- **SU-SIMO (K=1, M=100)**
- **MU-SIMO (M=100)**

Different processing methods are shown:

- **ZF processing**
- **MRC processing**

The plots demonstrate the performance at various signal-to-noise ratios (SNR) ranging from -20 dB to 20 dB.
Analysis Methodology - MRC & Perfect CSI

\[ y = \sqrt{P} G x + n \]

\[ z_{mrc} \triangleq G^H y = G^H \left( \sqrt{P} G x + n \right) = \sqrt{P} G^H G x + G^H n \]

\[ \Rightarrow z_{k,mrc} = \sqrt{P} \left\| g_k \right\|^2 x_k + \sqrt{P} \sum_{i \neq k} g_k^H g_i x_i + g_k^H n \]

- Achievable ergodic rate

\[ R_{k,mrc} = \mathbb{E} \left\{ \log_2 \left( 1 + \frac{P \left\| g_k \right\|^2}{P \sum_{i \neq k} \left| \tilde{g}_i \right|^2 + 1} \right) \right\} \geq \log_2 \left( 1 + \left( \mathbb{E} \left\{ \frac{P \sum_{i \neq k} \left| \tilde{g}_i \right|^2 + 1}{P \left\| g_k \right\|^2} \right) \right)^{-1} \right) \]

\[ = \log_2 \left( 1 + \frac{P (M - 1)}{P (K - 1) + 1} \right) \]

- Observe that: \( \tilde{g}_i \triangleq g_k^H g_i \left\| g_k \right\| \sim \mathcal{CN}(0,1) \), indep. of \( g_k \)
- Use \( \mathbb{E} [\log(1 + 1/x)] \geq \log(1 + 1/\mathbb{E}[x]) \)
- Here: \( \mathbb{E} \left\{ \frac{1}{\left\| g_k \right\|^2} \right\} = \frac{1}{M-1} \), since \( \mathbb{E} \left\{ \text{tr} (W^{-1}) \right\} = \frac{m}{n-m} \), if \( W \sim \mathcal{W}_m (n, I_n) \).
CSI from pilots: $\hat{G} = G + \mathcal{E}$, where $\varepsilon_i \sim \mathcal{CN}(0, \frac{1}{P_{p+1}} I_M)$, $\hat{g}_i \sim \mathcal{CN}(0, \frac{P_p}{P_{p+1}} I_M)$.

**MRC**

\[
\begin{align*}
\mathbf{z}_{\text{mrc}} &= \hat{G}^H \mathbf{y} = \hat{G}^H (\sqrt{P} G \mathbf{x} + \mathbf{n}) = \hat{G}^H (\sqrt{P} \hat{G} \mathbf{x} - \sqrt{P} \mathcal{E} \mathbf{x} + \mathbf{n}) \\
&\Rightarrow z_{k,\text{mrc}} = \sqrt{P} \|\hat{g}_k\|^2 x_k + \sqrt{P} \sum_{i \neq k} \hat{g}_k^H \hat{g}_i x_i - \sqrt{P} \sum_{i=1}^K \hat{g}_k^H \varepsilon_i x_i + \hat{g}_k^H \mathbf{n}
\end{align*}
\]

**Achievable ergodic rate per terminal**

\[
R_{k,\text{mrc}} = \left( 1 - \frac{T}{\tau} \right) \mathbb{E} \left\{ \log_2 \left( 1 + \frac{P \|\hat{g}_k\|^2}{P \sum_{i \neq k} |\hat{g}_i|^2 + P \sum_{i=1}^K \frac{1}{\tau P + 1}} + 1 \right) \right\}
\]

\[
\geq \left( 1 - \frac{T}{\tau} \right) \log_2 \left( 1 + \mathbb{E} \left\{ \frac{P \sum_{i \neq k} |\hat{g}_i|^2 + P \sum_{i=1}^K \frac{1}{\tau P + 1}}{P \|\hat{g}_k\|^2} + 1 \right\} \right)^{-1}
\]

\[
= \left( 1 - \frac{T}{\tau} \right) \log_2 \left( 1 + \frac{\tau P^2 (M-1)}{P (\tau P + 1) (K-1) + (\tau + 1) P + 1} \right)
\]

$\hat{g}_k$ and $\hat{g}_i \triangleq \frac{\hat{g}_k^H \hat{g}_i}{\|\hat{g}_k\|}$, $i \neq k$ are independent.
Summary

- $10 \times$ spectral efficiency and $1000 \times$ radiated energy efficiency possible with $M = 100$ antennas and MRC

- Results for DL are similar (no D.o.F penalty for pilot transmission)
Limits Imposed by Intercell Interference
(Probably) must rely on UL pilots to get CSI

On UL, may rely partly on blind/joint channel estimation and data decoding techniques to avoid pilot based estimation

But on DL, need explicit CSI to precode
Pilot Contamination Problem, cont.

- Consider UL training; one interferer

\[ Y_p = G \Phi + N + f \varphi^T \]

- LS estimation, with orthogonal pilots
  \( \Phi \Phi^H = P_p I \), say \( P_p = 1 \)

\[ \hat{G} = Y_p \Phi^H = G + N \Phi^H + f \varphi^T \Phi^H \]

rank−1

- Say \( \hat{g}_1 \) used for MRC data detection: \( y = Gx + fz + n \)

\[
\Rightarrow \frac{1}{M} \hat{g}_1^H y = \frac{1}{M} \left( \hat{g}_1^H Gx + \hat{g}_1^H fz + \hat{g}_1^H n \right) \\
= \frac{1}{M} \hat{g}_1^H Gx + \frac{1}{M} \left( \Phi \varphi^* f^H \right)_{1} fz + \cdots \\
\approx x_1 \\
\approx (\Phi \varphi^*)_{1} z
\]
Pilot Contamination Problem, cont.

- In principle, could project away $f\varphi^T$: 

$$Y_{p\Pi\varphi^T} = G\Phi\Pi\varphi^T + N\Pi\varphi^T$$

and estimate $G$ from $Y_{p\Pi\varphi^T}$. But $\text{rank}(\Pi\varphi^T) = \tau - 1$, so we sacrifice signal space.

- The effect on the downlink is similar.

- Can show: p.c. fundamentally limits performance of massive MIMO if $P$ is fixed and $M \to \infty$ [Mar2010]

- Can show: other cells transmitting data is as bad as other cells sending pilots [NLM2013]

- But currently not clear, how bad the p.c. effect is if other types of CSI estimation is used.
Summary

- Interference from other cells can significantly impair CSI estimation and the ensuing UL detection & DL precoding: “pilot contamination”
- Possible (partial) countermeasures (on UL)
  - “Blind” (partial pilot-free) CSI estimation (NL2011)
  - identify subspace associated with different cells exploiting power control & channel hardening [MVC2013]
  - clever allocation of pilots in different cells [YGFL2013,FAM2013]
  - pilot cont. precoding [AM2012]
Excess Degrees of Freedom in Massive MIMO: A Unique Opportunity for Hardware-Friendly Signal Shaping
Excess Degrees of Freedom

- There are $M - K$ “unused” degrees of freedom.

- In the downlink, these excess DoF may be used to shape the transmitted signals in a hardware-friendly way:
  - Per-antenna constant envelope or low-PAR multiuser precoding
  - Robust to PA non-linearity/less PA backoff
  - Exploit nullspace of $H^T$:

\[
\text{nullspace dimension} = M - K!
\]

\[
y = H^T x + w = H^T (x + z) + w
\]
Beamforming versus Low-PAR Signal Shaping

Erik G. Larsson
Massive MIMO: Fundamentals, Opportunities and Challenges
Constant-env. MU-MIMO, Flat channel [ML2013]

- **K = 10, Proposed CE Precoder (CE)**

Min. reqd. $P_T/\sigma^2$ to achieve a per-user rate of 2 bpcu

Number of Base Station Antennas (M)

-10
-8
-6
-4
-2
0
2
4
6
8

1.7 dB

Erik G. Larsson
Massive MIMO: Fundamentals, Opportunities and Challenges
Constant-env. MU-MIMO, Freq. sel. channel [ML2013a]

![Graph showing the relationship between \( \sigma^2 \) and \( T \)]

- \( L = 1, 2, 4, 8 \), ZF upper bound (TAPC)
- \( L = 1, 2, 4, 8 \), Coop. lower bound (TAPC)
- \( L = 1 \), CE
- \( L = 2 \), CE (\( T \gg \tau \))
- \( L = 4 \), CE (\( T \gg \tau \))
- \( L = 8 \), CE (\( T \gg \tau \))

Erik G. Larsson
Massive MIMO: Fundamentals, Opportunities and Challenges
- $E_k$: energy per *information symbol* $u_k$ (per user)
- $P$: total transmit power
- Received signals:

$$y_k = \sqrt{\frac{P}{M}} \sum_{i=1}^{M} h_{k,i} e^{j\theta_i} + n_k = \sqrt{P} \sqrt{E_k} u_k + \sqrt{P} \left( \frac{\sum_{i=1}^{M} h_{k,i} e^{j\theta_i}}{\sqrt{M}} - \sqrt{E_k} u_k \right) + n_k \triangleq \delta_k$$

- With Gaussian symbols per user:

$$I(y_k; u_k) = h(u_k) - h(u_k \mid y_k) = h(u_k) - h\left( u_k - \frac{y_k}{\sqrt{P} \sqrt{E_k}} \mid y_k \right) \geq h(u_k) - h\left( u_k - \frac{y_k}{\sqrt{P} \sqrt{E_k}} \right)$$

$$\geq h(u_k) - h\left( \frac{\delta_k}{\sqrt{E_k}} + \frac{n_k}{\sqrt{P} \sqrt{E_k}} \right) = \log_2(\pi e) - h\left( \frac{\delta_k}{\sqrt{E_k}} + \frac{n_k}{\sqrt{P} \sqrt{E_k}} \right)$$

$$\geq \log_2(\pi e) - \log_2 \left( \pi e \text{ var} \left[ \frac{\delta_k}{\sqrt{E_k}} + \frac{n_k}{\sqrt{P} \sqrt{E_k}} \right] \right)$$

$$\geq \log_2(\pi e) - \log_2 \left( \pi e \mathbb{E} \left[ \left| \frac{\delta_k}{\sqrt{E_k}} + \frac{n_k}{\sqrt{P} \sqrt{E_k}} \right|^2 \right] \right)$$

$$= \log_2(\pi e) - \log_2 \left( \pi e \left[ \frac{\mathbb{E}[|\delta_k|^2]}{E_k} + \frac{1}{PE_k} \right] \right)$$

- Downlink signal design:

$$\min_{\theta_i \in [-\pi, \pi)} \sum_{k=1}^{K} |\delta_k|^2 \iff \min_{\theta_i \in [-\pi, \pi)} \sum_{k=1}^{K} \left| \sum_{i=1}^{M} h_{k,i} e^{j\theta_i} \sqrt{M} - \sqrt{E_k} u_k \right|^2$$
Precoder design problem:

\[
\begin{align*}
\text{minimize} & \quad \max \left\{ \| \hat{a}_1 \|_{\infty}, \ldots, \| \hat{a}_N \|_{\infty} \right\} \\
\text{subject to} & \quad y_w = H_w s_w + \text{noise}, \quad w \in \mathcal{T} \\
& \quad y_w = \text{only noise}, \quad w \in \mathcal{T}^c.
\end{align*}
\]

- \( H_w \): time-frequency channel
- \( \mathcal{T} \): used subcarriers; \( \mathcal{T}^c \): null subcarriers
- Convex optimization problem, fast algorithm; relaxation
PAR-aware MU-MIMO OFDM DL (99% percentiles)
PAR-aware MU-MIMO OFDM DL (99% percentiles)
Summary

- Low PAR may be desirable to enable high overall energy efficiency & low cost
- Large number of antennas offers entirely new and unique opportunities for low PAR signal design
Massive MIMO—a Goldmine of Research Problems

- Base station receiver processing
  - Partially distributed processing (MRC/ZF/MMSE hybrids)
  - Advanced channel estimation (exploit excess DoF)
  - Synchronization at low SNR
- Base station transmitter processing
  - Partially distributed processing
  - Low-complexity low-PAPR precoders
  - OFDM or single-carrier?
- Effects of hardware imperfections: Will the law of large numbers deliver?
  - phase noise and clock distribution,
  - I/Q imbalance, A/D resolution, PA linearity
- Pilot contamination (both UL & DL) may be mitigated by
  - “Blind” channel estimation algorithms [NL2012]
  - Coordination and planning of pilot reuse [YGFL2013]
- TDD operation using UL channels for DL precoding requires reciprocity calibration
  - Cost? (BW & power), algorithms?, dedicated hardware?
- Other system issues
  - Time and frequency synchronization at low SNR
  - Idle-terminal paging (DL) under non-CSI@TX
Eq.-Free SC vs. OFDM @ 1 bps/Hz/t [PML2012]
Literature


Thank You

Visit the massive MIMO website

www.commsys.isy.liu.se/~egl/vlm/vlm.html