Energy-Efficient Communication in Wireless Networks

Small or massive MIMO?

Dr. Emil Björnson

Seminar at 5GrEEn Summer School, KTH Kista, Stockholm
26 August 2014
Biography

• 1983: Born in Malmö, Sweden

• 2007: Master in Engineering Mathematics, Lund University, Sweden

• 2011: PhD in Telecommunications, KTH, Stockholm, Sweden

• 2012-2014: Joint post-doc at Supélec, Paris, and KTH, based on International postdoc grant “Optimization of Green Small-Cell Networks”

• 2014: Assistant Professor in Communication Systems, Linköping University, Sweden
Outline

• Introduction & Background

• Part 1: Problem Formulation
  - Detailed system model (energy-efficiency, rates and power consumption)

• Part 2: Optimization of Energy-Efficiency
  - Optimal system parameters: Reveal fundamental interplay
  - Numerical results: Single-cell and multi-cell

• Part 3: Massive MIMO
  - Main properties and deployment ideas

• Part 4: Multi-Objective Network Optimization
  - Optimizing energy-efficiency and other metrics \textit{in parallel}
Introduction & Background
Introduction

• Wireless Connectivity
  - A natural part of our lives

- Video streaming
- Web browsing
- Voice call
- Social networks
- Gaming

• Rapid Network Traffic Growth
  - 61% annual growth
  - Exponential increase!
  - Extrapolation:
    - 20x until 2020
    - 200x until 2025
    - 2000x until 2030

Exabytes per Month

2.2 GB/person/month

210 MB/month/person

Source: Cisco VNI Mobile, 2014
Exponential Traffic Growth

• Is this Growth Sustainable?
  - User demand will increase – users expect more for same price
  - Traffic supply – increases only if business models allow it!

• Exponential Growth is Nothing New!
  - $10^6$ increase in last 45 years!

**Martin Cooper’s law**
The number of simultaneous voice/data connections has doubled every 2.5 years since the beginning of wireless

- Coopers law: 32%/year
- New predictions: 61%/year

Source: Personal Communications in 2025, Martin Cooper
Wireless Networks

- Cellular Network Architecture
  - Coverage Area divided into cells
  - One fixed base station per cell
  - Serves all users in the cell

- Different Standards
  - 2G (GSM), 3G (UMTS), 4G (LTE/LTE-A)

More and more focus on data traffic

- Traditional Ways to Handle More Traffic
  - Higher cell density (variable cell sizes)
  - More spectrum (carrier aggregation)
  - Higher spectral efficiency (spatial processing)
High Data Rates

• Traditional Design Metric
  - High peak and/or average rates [bit/s/active user]

• Basic Signal Propagation
  - Signal energy decays with distance
  - Peak rates in cell center
  - Far from peak rates at cell edge

• Traffic Independent of Location
  - Easily satisfied in cell center
  - Highest demand at cell edge!

Need for Additional Metrics!
To optimize and design our networks properly!
Expectations for 5G Networks

• 5G – The Next Network Generation
  - Expected to be introduced by year 2020
  - Design objectives are currently being defined

<table>
<thead>
<tr>
<th>5G Performance Metrics</th>
<th>Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Rate (Mbit/s/active user)</td>
<td>10-100x</td>
</tr>
<tr>
<td>Average Area Rate (Mbit/s/km²)</td>
<td>1000x</td>
</tr>
<tr>
<td>Active devices (per km²)</td>
<td>10-100x</td>
</tr>
<tr>
<td>Energy-Efficiency (Mbit/Joule)</td>
<td>1000x</td>
</tr>
</tbody>
</table>

Source: METIS project (www.metis2020.com)

Parts 1-3
What if we optimize a network only for energy-efficiency?
What will it look like?

Part 4
Is it possible to optimize a network with respect to multiple metrics?
What does “optimality” mean then?
Part 1

Problem Formulation
How to Measure Energy-Efficiency?

• Energy-Efficiency (EE) in bit/Joule

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

• Conventional Academic Approaches:
  - Maximize rates with fixed power
  - Minimize transmit power for fixed rates
  - See for example:

  *Optimal Resource Allocation in Coordinated Multi-Cell Systems*
  
  Book from 2013 by Emil Björnson and Eduard Jorswieck

  *Free to download from my homepage*

**New Problem:** Balance rates and power consumption

*Important to account for overhead signaling and circuit power!*
Basic Information Theory

**Achievable Rate per Active User** [Lower Bound on Shannon Capacity]

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

Signal-to-interference-and-noise ratio (SINR)

- More than One Active User per Cell?
  - Yes, but causes inter-user interference
  - Traditional approach: Orthogonal in time/frequency (TDMA, OFDMA)
  - New multi-antenna approach: Space-division multiple access (SDMA)

**Known as**

Multi-user MIMO (Multiple input multiple output)
Beamforming in Line-of-Sight and Non-Line-of-Sight

- **Line-of-Sight**
  - Adapt signal phases at antennas
  - Steer beam towards receiving user
  - Imperfect beams: inter-user interference

- **Non-Line-of-Sight**
  - Multipath propagation
  - Add components coherently
  - Coherent
  - Non-Coherent
Single-Cell: Optimizing for Energy-Efficiency

- **Clean Slate Design**
  - Single Cell: One base station (BS) with $M$ antennas
  - Geometry: Random distribution for user locations and pathlosses
  - Multiple users: Pick $K$ users randomly and serve with some rate $R$

**Problem Formulation**

Select $(M, K, R)$ to maximize EE!

**Next Step**

Find expression: EE as a function of $M, K, R$. 
System Model: Protocol

- **Time-Division Duplex (TDD) Protocol**
  - Uplink and downlink separated in time
  - Uplink fraction $\zeta^{(ul)}$ and downlink fraction $\zeta^{(dl)}$

- **Coherence Block**
  - $B$ Hz bandwidth = $B$ “channel uses” per second (symbol time $1/B$)
  - Channel stays fixed for $U$ channel uses (symbols) = Coherence block
  - Determines how often we send pilot signals to estimate channels

Assumption: Perfect channel estimation (relaxed later)
System Model: Channels

- **Flat-Fading Channels**
  - Channel between BS and User $k$: $h_k \in \mathbb{C}^M$
  - Rayleigh fading: $h_k \sim CN(0, \lambda_k \mathbb{I})$
  - Channel variances $\lambda_k$: Random variables, pdf $f_\lambda(x)$

- **Uplink Transmission**
  - User $k$ transmits signal $s_k$ with power $\mathbb{E}\{|s_k|^2\} = p_k^{(ul)}$ [Joule/channel use]
  - Received signal at BS:
    \[
    y = h_k s_k + \sum_{i=1, i \neq k}^{K} h_i s_i + n
    \]
  - Recover $s_k$ by receive beamforming $g_k$ as $g_k^H y$:
    \[
    \text{SINR}^{(ul)}_k = \frac{\mathbb{E}\{|s_k|^2 | g_k^H h_k|^2\}}{\sum_{i \neq k} \mathbb{E}\{|s_i|^2 | g_k^H h_i|^2\} + \mathbb{E}\{|g_k^H n|^2\}} = \frac{p_k^{(ul)} |g_k^H h_k|^2}{\sum_{i \neq k} p_i^{(ul)} |g_k^H h_i|^2 + \sigma^2 \|g_k\|^2}
    \]
System Model: Channels (2)

- **Flat-Fading Channels**
  - Channel between BS and User $k$: $h_k \in \mathbb{C}^M$
  - Rayleigh fading: $h_k \sim \mathcal{CN}(0, \lambda_k \mathbf{I})$
  - Channel variances $\lambda_k$: Random variables, pdf $f_\lambda(x)$

- **Downlink Transmission**
  - BS transmits $d_k$ to User $k$ with power $\mathbb{E}\{|d_k|^2\} = p_k^{(dl)}$ [Joule/channel use]
  - Spatial directivity by beamforming vector $\mathbf{v}_k$
  - Received signal at User $k$:

$$y_k = h_k^H \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|} d_k + \sum_{i=1, i \neq k}^{K} h_k^H \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|} d_i + n_k$$

  - Recover $d_k$ at User $k$:

$$\text{SINR}^{(dl)}_k = \frac{p_k^{(dl)} |h_k^H \mathbf{v}_k|^2 / \|\mathbf{v}_k\|^2}{\sum_{i \neq k} p_i^{(dl)} |h_k^H \mathbf{v}_i|^2 / \|\mathbf{v}_i\|^2 + \sigma^2}$$

  - Signals from other users (interference)
  - Noise $\sim \mathcal{CN}(0, \sigma^2)$
System Model: How Much Transmit Power?

- Design Parameter: Gross rate $R$
  
  - Make sure that $R = \begin{cases} B \log_2(1 + \text{SINR}_{k}^{(ul)}) & \text{for all } k \text{ in uplink} \\ B \log_2(1 + \text{SINR}_{k}^{(dl)}) & \text{for all } k \text{ in downlink} \end{cases}$

  - Select beamforming $g_k$ and $v_k$, adapt transmit power $p_k^{(ul)}$ and $p_k^{(dl)}$

- Gives $K$ Equations:

  - \[
  p_k^{(ul)} |g_k^H h_k|^2 = (2^{R/B} - 1)(\sum_{i \neq k} p_i^{(ul)} |g_i^H h_i|^2 + \sigma^2 \|g_k\|^2) \quad \text{for } k = 1, ..., K
  \]

  - \[
  p_k^{(dl)} \frac{|h_k^H v_k|^2}{\|v_k\|^2} = (2^{R/B} - 1)(\sum_{i \neq k} p_i^{(dl)} \frac{|h_i^H v_i|^2}{\|v_i\|^2} + \sigma^2) \quad \text{for } k = 1, ..., K
  \]

- Linear equations in transmit powers $\rightarrow$ Solve by Gaussian elimination!

Total Transmit Power [Joule/s] for $g_k = v_k$

Uplink energy/symbol: $\sigma^2 D^{-H} 1$
Downlink energy/symbol: $\sigma^2 D^{-1} 1$

Same total power: $P_{\text{trans}} = B \mathbb{E}\{\sigma^2 1^H D^{-1} 1\} = B \mathbb{E}\{\sigma^2 1^H D^{-H} 1\}$

where $[D]_{k,l} = \begin{cases} \frac{|h_k^H v_k|^2}{(2^{R/B-1})\|v_k\|^2} & \text{for } k = l \\ -\frac{|h_k^H v_l|^2}{\|v_l\|^2} & \text{for } k \neq l \end{cases}$
System Model: How Much Transmit Power? (2)

- What did we Derive?
  - Optimal power allocation for fixed beamforming vectors

Different Beamforming
- Notation: $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_K]$, $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_K]$, $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$, $\mathbf{P}(ul) = \text{diag}(p_1^{(ul)}, \ldots, p_K^{(ul)})$

- Maximum ratio trans./reception (MRT/MRC): $\mathbf{G} = \mathbf{V} = \mathbf{H}$
- Zero-forcing (ZF) beamforming: $\mathbf{G} = \mathbf{V} = \mathbf{H} (\mathbf{H}^H \mathbf{H})^{-1}$
- Optimal beamforming: $\mathbf{G} = \mathbf{V} = (\sigma^2 \mathbf{I} + \mathbf{H} \mathbf{P}^{(ul)} \mathbf{H}^H)^{-1} \mathbf{H}$

Balance signal and interference (iteratively!)
System Model: How Much Transmit Power? (3)

- Simplified Expressions for ZF ($M \geq K + 1$)
  - Main property: $H^H V = H^H H (H^H H)^{-1} = I$
  - Hence: $[D]_{k,l} = \begin{cases} 
  \frac{|h_k^H v_k|^2}{(2^R/B - 1)\|v_k\|^2} & \text{for } k = l \\
  - \frac{|h_k^H v_l|^2}{\|v_l\|^2} & \text{for } k \neq l
  \end{cases} = \begin{cases} 
  \frac{1}{(2^R/B - 1)\|v_k\|^2} & \text{for } k = l \\
  0 & \text{for } k \neq l
  \end{cases}$

- Total transmit power:
  \[
  P_{\text{trans}} = \mathbb{E}\{B\sigma^2 1^H D^{-1} 1\} = B\sigma^2 (2^R/B - 1) \sum_k \mathbb{E}\{\|v_k\|^2\} = B\sigma^2 (2^R/B - 1) \frac{K}{M - K} \mathbb{E}\{\frac{1}{\lambda}\}
  \]
  
  \[
  = \text{tr}\left((H^H H)^{-1}\right)
  \]

Summary: Transmit Power with ZF

Parameterize gross rate as $R = B \log_2(1 + \alpha(M - K))$ for some $\alpha$

Total transmit power: $P_{\text{trans}} = \alpha B\sigma^2 S_\lambda K$ [Joule/s]
Detailed Power Consumption Model

• What Consumes Power?
  - Not only radiated transmission power
  - Circuits, signal processing, backhaul, etc.
  - Must be specified as functions of $M, K, R$

• Power Amplifiers
  - Amplifier efficiencies: $\eta^{(ul)}, \eta^{(dl)} \in (0,1]$  
  - Average inefficiency: $\bar{\zeta}^{(ul)} / \eta^{(ul)} + \bar{\zeta}^{(dl)} / \eta^{(dl)} = \frac{1}{\eta}$

  Summary: $\frac{P_{\text{trans}}}{\eta}$

• Active Transceiver Chains
  - $P_{\text{FIX}}$ = Fixed power (control signals, oscillator at BS, standby, etc.)
  - $P_{\text{BS}}$ = Circuit power / BS antenna (converters, mixers, filters)
  - $P_{\text{UE}}$ = Circuit power / user (oscillator, converters, mixer, filters)

  Summary: $P_{\text{FIX}} + M \cdot P_{\text{BS}} + K \cdot P_{\text{UE}}$
Detailed Power Consumption Model (2)

- **Signal Processing**
  - Channel estimation and beamforming
  - Efficiency: $L_{BS}, L_{UE}$ arithmetic operations / Joule

- **Channel Estimation:**
  $$\frac{B}{U} \left( \frac{2\tau^{(ul)} MK^2}{L_{BS}} + \frac{4\tau^{(dl)} K^2}{L_{UE}} \right)$$
  - Once in uplink/downlink per coherence block
  - Pilot signal lengths: $\tau^{(ul)} K, \tau^{(dl)} K$ for some $\tau^{(ul)}, \tau^{(dl)} \geq 1$

- **Linear Processing** (for $\mathbf{G} = \mathbf{V}$):
  $$\frac{B}{U} \frac{C_{\text{beamforming}}}{L_{BS}} + B \left( 1 - \frac{\tau^{(ul)} + \tau^{(ul)}}{U} \right) \frac{2MK}{L_{BS}}$$
  - Compute beamforming vector once per coherence block
  - Use beamforming for all $B(1 - (\tau^{(ul)} + \tau^{(ul)})K/U)$ symbols

- **Types of beamforming:**
  - $C_{\text{beamforming}} = \begin{cases} 
  3MK & \text{for MRT/MRC} \\
  3MK^2 + MK + \frac{1}{3}K^3 & \text{for ZF} \\
  Q(3MK^2 + MK + \frac{1}{3}K^3) & \text{for Optimal}
  \end{cases}$
Detailed Power Consumption Model (3)

- **Coding and Decoding:** $R_{\text{sum}} (P_{\text{COD}} + P_{\text{DEC}})$
  - $P_{\text{COD}} = $ Energy for coding data / bit
  - $P_{\text{DEC}} = $ Energy for decoding data / bit

- **Sum rate:**
  
  $$R_{\text{sum}} = K \left( \zeta^{(\text{ul})} - \frac{\tau^{(\text{ul})}K}{U} \right) R + K \left( \zeta^{(\text{dl})} - \frac{\tau^{(\text{dl})}K}{U} \right) R$$

  $$= K \left( 1 - \frac{(\tau^{(\text{ul})} + \tau^{(\text{dl})})K}{U} \right) R$$

- **Backhaul Signaling:** $P_{\text{BH}} + R_{\text{sum}} P_{\text{BT}}$
  - $P_{\text{BH}} = $ Load-independent backhaul power
  - $P_{\text{BT}} = $ Energy for sending data over backhaul / bit
Detailed Power Consumption Model: Summary

- Many Things Consume Power
  - Parameter values (e.g., $P_{BS}, P_{UE}$) change over time
  - Structure is important for analysis

**Generic Power Model**

$$
\frac{P_{trans}}{\eta} + C_{0,0} + C_{0,1}M + C_{1,0}K + C_{1,1}MK + C_{2,0}K^2 + C_{3,0}K^3 + C_{2,1}MK^2 + AK \left( 1 - \frac{(\tau^{(ul)} + \tau^{(dl)})K}{U} \right) R
$$

- Transmit with amplifiers
- Circuit power per transceiver chain
- Cost of signal processing
- Coding/decoding/backhaul

for some parameters $C_{l,m}$ and $A$

- Observations
  - Polynomial in $M$ and $K \rightarrow$ Increases faster than linear with $K$
  - Depends on cell geometry only through $P_{trans}$
Finally: Problem Formulation

- Maximize Energy-Efficiency:

\[
\max_{M, K, R} \frac{K \left(1 - \frac{(τ^{ul}) + (τ^{dl})K}{U}\right)R}{P_{\text{trans}} \frac{\eta}{\eta} + \sum_{i=0}^{3} C_{i,0}K^i + \sum_{i=0}^{2} C_{i,1}MK^i + AK \left(1 - \frac{(τ^{ul}) + (τ^{dl})K}{U}\right)R}
\]

Average Sum Rate [bit/s/cell]

Power Consumption [Joule/s/cell]

Closed Form Expressions with ZF

Recall: \( R = B \log_2(1 + \alpha(M - K)) \) for some \( \alpha \) and \( P_{\text{trans}} = \alpha B \sigma^2 S \lambda K \)

Define: \( τ = τ^{ul} + τ^{dl} \)

\[
\max_{M, K, \alpha} \frac{K \left(1 - \frac{\tau K}{U}\right)B \log_2(1 + \alpha(M - K))}{\frac{\alpha B \sigma^2 S \lambda K}{\eta} + \sum_{i=0}^{3} C_{i,0}K^i + \sum_{i=0}^{2} C_{i,1}MK^i + AK \left(1 - \frac{\tau K}{U}\right)B \log_2(1 + \alpha(M - K))}
\]

Simple ZF expression: Used for analysis, other beamforming by simulation
Why Such a Detailed/Complicated Model?

- Simplified Model \( \rightarrow \) Unreliable Optimization Results
  - Two examples based on ZF
  - Beware: Both has appeared in the literature!

- Example 1: Fixed circuit power and no coding/decoding/backhaul
  \[
  \begin{align*}
  \text{maximize } & \quad K \left(1 - \frac{\tau K}{U}\right) B \log_2(1 + \alpha(M - K)) \\
  & \quad \frac{\alpha B \sigma^2 S \lambda K}{\eta} + C_{0,0}
  \end{align*}
  \]
  - If \( M \to \infty \), then \( \log_2(1 + \alpha(M - K)) \to \infty \) and thus EE \( \to \infty \)!

- Example 2: Ignore pilot overhead and signal processing
  \[
  \begin{align*}
  \text{maximize } & \quad KB \log_2(1 + \alpha(M - K)) \\
  & \quad \frac{\alpha B \sigma^2 S \lambda K}{\eta} + C_{0,0} + C_{1,0}K + C_{0,1}M
  \end{align*}
  \]
  - If \( M, K \to \infty \) with \( \frac{M}{K} \) = constant \( > 1 \), then \( \log_2(1 + \alpha K \left(\frac{M}{K} - 1\right)) \to \infty \) and EE \( \to \infty \)!
Part 1

Questions?
Part 2

Optimization of Energy-Efficiency
Preliminaries

- Our Goal
  - Optimize number of antennas $M$
  - Optimize number of active users $K$
  - Optimize the (normalized) transmit power $\alpha$

- Outline
  - Optimize each variable separately
  - Devise an alternating optimization algorithm

---

**Definition** (Lambert $W$ function)

- Lambert $W$ function, $W(x)$, solves equation $W(x)e^{W(x)} = x$
- The function is increasing and satisfies $W(0) = 0$
- $e^{W(x)}$ behaves as a linear function (i.e., $e^{W(x)} \approx x$):

\[
\frac{x e}{\log_e(x)} \leq e^{W(x) + 1} \leq \frac{x}{\log_e(x)}(1 + e) \quad \text{for} \quad x \geq e.
\]
Solving Optimization Problems

• How to Solve an Optimization Problem?
  - Simple if the function is “nice”:

  **Quasi-Concave Function**

  For any two points on the graph of the function, the line between the points is below the graph.

  Property: Goes up and then down
  Examples: \(-x^2, \log(x)\)

• Maximization of a Quasi-Concave Function \(\varphi(x)\):
  1. Compute the first derivative \(\frac{d}{dx} \varphi(x)\)
  2. Find switching point by setting \(\frac{d}{dx} \varphi(x) = 0\)
  3. Only one solution \(\Rightarrow\) It is the unique maximum!
Optimal Number of BS Antennas

- Find $M$ that maximizes EE with ZF:

$$\max_{M \geq K + 1} \frac{K \left(1 - \frac{\tau K}{U}\right) B \log_2 (1 + \alpha (M - K))}{\frac{\alpha B \sigma^2 S \lambda K}{\eta} + \sum_{i=0}^{3} C_{i,0} K^i + \sum_{i=0}^{2} C_{i,1} MK^i + AK \left(1 - \frac{\tau K}{U}\right) B \log_2 (1 + \alpha (M - K))} + \sum_{i=0}^{3} C_{i,0} K^i + \sum_{i=0}^{2} C_{i,1} MK^i + AK \left(1 - \frac{\tau K}{U}\right) B \log_2 (1 + \alpha (M - K))}$$

**Theorem 1 (Optimal $M$)**

EE is quasi-concave w.r.t. $M$ and maximized by

$$M^* = e \left( \frac{\alpha (B \sigma^2 S \lambda K / \eta + \sum_{i=0}^{3} C_{i,0} K^i)}{e \sum_{i=0}^{2} C_{i,1} K^i} + \frac{\alpha K - 1}{e} \right) + 1$$

- Observations
  - Increases with circuit coefficients independent of $M$ (e.g., $P_{\text{FIX}}, P_{\text{UE}}$)
  - Decreases with circuit coefficients multiplied with $M$ (e.g., $P_{\text{BS}}, 1/L_{\text{BS}}$)
  - Independent of cost of coding/decoding/backhaul
  - Increases with power $\alpha$ approx. as $\frac{\alpha}{\log \alpha}$ (almost linear)
Optimal Transmit Power

- Find $\alpha$ that maximizes EE with ZF:

$$\text{maximize} \quad \frac{K \left(1 - \frac{\tau K}{U}\right) B \log_2(1 + \alpha(M - K))}{\frac{\alpha B \sigma^2 S \lambda K}{\eta} + \sum_{i=0}^{3} C_{i,0} K^i + \sum_{i=0}^{2} C_{i,1} M K^i + AK \left(1 - \frac{\tau K}{U}\right) B \log_2(1 + \alpha(M - K))}$$

**Theorem 2 (Optimal $\alpha$)**

EE is quasi-concave w.r.t. $\alpha$ and maximized by

$$\alpha^* = e^{\frac{\eta (M-K) (\sum_{i=0}^{3} C_{i,0} K^i + \sum_{i=0}^{2} C_{i,1} M K^i) \frac{1}{e}}{B \sigma^2 S \lambda \left(\sum_{i=0}^{3} C_{i,0} K^i + \sum_{i=0}^{2} C_{i,1} M K^i\right) \frac{1}{e}} + 1} - 1$$

- Observations
  - Increases with all circuit coefficients (e.g., $P_{\text{FIX}}, P_{\text{BS}}, P_{\text{UE}}, 1/L_{\text{BS}}$)
  - Independent of cost of coding/decoding/backhaul
  - Increases with $M$ approx. as $\frac{M}{\log M}$ (almost linear)

More circuit power $\rightarrow$ More transmit power
Optimal Number of Users

- Find $K$ that maximizes EE with ZF:

$$
\text{maximize} \quad K \left(1 - \frac{\tau K}{U}\right) B \log_2(1 + \bar{\alpha}(\bar{\beta} - 1))
\frac{\bar{\alpha} B \sigma^2 S_\lambda}{\eta} + \sum_{i=0}^3 C_{i,0} K^i + \sum_{i=0}^2 C_{i,1} \bar{\beta} K^{i+1} + AK \left(1 - \frac{\tau K}{U}\right) B \log_2(1 + \bar{\alpha}(\bar{\beta} - 1))
$$

where $\bar{\alpha} = \alpha K$ and $\bar{\beta} = \frac{M}{K}$ are fixed

**Theorem 3 (Optimal $K$)**

EE is quasi-concave w.r.t. $K$

Maximized by the root of a quartic polynomial:
Closed form for $K^*$ but very “large” expressions

- Observations
  - Increases with fixed circuit power (e.g., $P_{\text{FIX}}$)
  - Decreases with circuit coefficients multiplied with $M$ or $K$ ($P_{\text{BS}}, P_{\text{UE}}, 1/L_{BS}$)
Impact of Cell Size

• Are Smaller Cells More Energy Efficient?
  - Recall: $s_\lambda = \mathbb{E}\left\{\frac{1}{\lambda}\right\}$
  - Smaller cells $\rightarrow \lambda$ is larger $\rightarrow s_\lambda$ is smaller

• For any given parameters $M, \alpha, K$
  - Smaller $s_\lambda$ $\rightarrow$ smaller transmit power $\alpha B \sigma^2 s_\lambda K$
  - Higher EE!

• Expressions for $M^*, \alpha^*, K^*$
  - $M^*$ and $K^*$ increases with $s_\lambda$
  - $\alpha^*$ decreases with $s_\lambda$
  - Smaller cells:
    Less hardware and fewer users per cell
    Use shorter distances to reduce power

Dependence on Other Parameters

Many other observations can be made
Example: Impact of bandwidth $B$, coherence block length $U$, etc.
• Joint EE Optimization
  - EE is a function of $M$, $\alpha$, and $K$
  - Theorems 1-3 optimize one parameter, when the other two are fixed
  - Can we optimize all of them?

**Algorithm:** Alternating Optimization

1. Assume that an initial set $(M, \alpha, K)$ is given
2. Update number of users $K$ (and implicitly $M$ and $\alpha$) using Theorem 3
3. Update number of antennas $M$ using Theorem 1
4. Update transmit power ($\alpha$) using Theorem 2
5. Repeat 2.-5. until convergence

**Theorem 4**
The algorithm convergences to a local optimum to the joint EE optimization problem

**Disclaimer**
$M$ and $K$ should be integers
Theorems 1 and 3 give real numbers → Take one of the 2 closest integers
Single-Cell Simulation Scenario

• Main Characteristics
  - Circular cell with radius 250 m
  - Uniform user distribution
  - Uncorrelated Rayleigh fading
  - Typical 3GPP pathloss model

• Many Parameters in the System Model
  - We found numbers from ≈ 2012 in the literature:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell radius (single-cell): $d_{\text{max}}$</td>
<td>250 m</td>
<td>Fraction of downlink transmission: $\zeta^{(d)}$</td>
<td>0.6</td>
</tr>
<tr>
<td>Minimum distance: $d_{\text{min}}$</td>
<td>35 m</td>
<td>Fraction of uplink transmission: $\zeta^{(u)}$</td>
<td>0.4</td>
</tr>
<tr>
<td>Large-scale fading model: $l(x)$</td>
<td>$10^{-3.53/|x|^{3.76}}$</td>
<td>PA efficiency at the BSs: $\eta^{(d)}$</td>
<td>0.39</td>
</tr>
<tr>
<td>Transmission bandwidth: $B$</td>
<td>20 MHz</td>
<td>PA efficiency at the UEs: $\eta^{(u)}$</td>
<td>0.3</td>
</tr>
<tr>
<td>Channel coherence bandwidth: $B_C$</td>
<td>1800 kHz</td>
<td>Fixed power consumption (control signals, backhaul, etc.): $P_{\text{FIX}}$</td>
<td>18 W</td>
</tr>
<tr>
<td>Channel coherence time: $T_C$</td>
<td>10 ms</td>
<td>Power consumed by local oscillator at BSs: $P_{\text{SYN}}$</td>
<td>2 W</td>
</tr>
<tr>
<td>Coherence block (channel uses): $U$</td>
<td>1800</td>
<td>Power required to run the circuit components at a BS: $P_{\text{BS}}$</td>
<td>1 W</td>
</tr>
<tr>
<td>Total noise power: $B\sigma^2$</td>
<td>-96 dBm</td>
<td>Power required to run the circuit components at a UE: $P_{\text{UE}}$</td>
<td>0.1 W</td>
</tr>
<tr>
<td>Relative pilot lengths: $\tau^{(u)}$, $\tau^{(d)}$</td>
<td>1</td>
<td>Power required for coding of data signals: $P_{\text{COD}}$</td>
<td>0.1 W/(Gbit/s)</td>
</tr>
<tr>
<td>Computational efficiency at BSs: $L_{\text{BS}}$</td>
<td>12.8 Gflops/W</td>
<td>Power required for decoding of data signals: $P_{\text{DEC}}$</td>
<td>0.8 W/(Gbit/s)</td>
</tr>
<tr>
<td>Computational efficiency at UEs: $L_{\text{UE}}$</td>
<td>5 Gflops/W</td>
<td>Power required for backhaul traffic: $P_{\text{BT}}$</td>
<td>0.25 W/(Gbit/s)</td>
</tr>
</tbody>
</table>
Optimal Single-Cell System Design: ZF Beamforming

Optimum

\[ M = 165 \]
\[ K = 104 \]
\[ \alpha = 0.87 \]

User rates:
\[ \approx 64\text{-QAM} \]

Massive MIMO!

Name for multi-user MIMO with very many antennas

Global Optimum:
\[ M = 165, \ K = 104 \]
\[ \text{EE} = 30.7 \text{ Mbit/J} \]

Energy Efficiency [Mbit/Joule]

Number of Antennas \((M)\)

Number of Users \((K)\)
Optimal Single-Cell System Design: “Optimal” Beamforming

**Optimum**

\[
\begin{align*}
M &= 145 \\
K &= 95 \\
\alpha &= 0.91
\end{align*}
\]

User rates: \(\approx 64\)-QAM

**Not optimal!**

Gives optimal beamforming but computations are too costly

![Graph showing energy efficiency](image)

\[Q = 3\]

Global Optimum:
\[M = 145, K = 95, \text{EE} = 30.3 \text{ Mbit/J}\]
Optimal Single-Cell System Design: MRT/MRC Beamforming

**Optimum**

\[ M = 81 \]
\[ K = 77 \]
\[ \alpha = 0.24 \]

User rates: \[ \approx 2\text{-PSK} \]

**Observation**

Lower EE than with ZF

Also Massive MIMO setup

Low rates

Global Optimum:

\[ M = 81, \quad K = 77 \]

EE = 9.86 Mbit/J
Multi-Cell Scenarios and Imperfect Channel Knowledge

- Limitations in Previous Analysis
  - Perfect channel knowledge
  - No interference from other cells

- Consider a Symmetric Multi-Cell Scenario:

<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 3</td>
<td>Cluster 4</td>
<td>Cluster 3</td>
<td>Cluster 4</td>
<td>Cluster 3</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>Cluster 2</td>
<td>Cell under study (Cluster 1)</td>
<td>Cluster 2</td>
<td>Cluster 1</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>Cluster 4</td>
<td>Cluster 3</td>
<td>Cluster 4</td>
<td>Cluster 3</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>Cluster 2</td>
<td>Cluster 1</td>
<td>Cluster 2</td>
<td>Cluster 1</td>
</tr>
</tbody>
</table>

Assumptions
All cells look the same → Jointly optimized
All cells transmit in parallel
Fractional pilot reuse: Divide cells into clusters
Uplink pilot length $\tau_{ul}K$ for $\tau_{ul} \in \{1,2,4\}$
Multi-Cell Scenarios and Imperfect Channel Knowledge (2)

- **Inter-Cell Interference**
  - $\lambda_{jl} = \text{Channel attenuation between a random user in cell } l \text{ and BS } j$
  - $J = \sum_{l \neq j} \mathbb{E}\left\{\frac{\lambda_{jl}}{\lambda_{jj}}\right\}$ is relative severity of inter-cell interference

**Lemma** (Achievable Rate)

Consider same transmit power as before: $P_{\text{trans}} = \alpha B \sigma^2 S \lambda K$

Achievable rate under ZF and pilot-based channel estimation:

$$R = B \log_2 \left( 1 + \frac{\alpha(M - K) I_{PC}}{\alpha(M - K) I_{PC} + \left(1 + I_{PC} + \frac{1}{\alpha K \tau^{(ul)}}\right) (1 + \alpha K J) - \alpha K (1 + I_{PC}^2)} \right)$$

where $I_{PC} = \sum_{l \neq j \text{ only in cluster}} \mathbb{E}\left\{\frac{\lambda_{jl}}{\lambda_{jj}}\right\}$ and $I_{PC}^2 = \sum_{l \neq j \text{ only in cluster}} \mathbb{E}\left\{\left(\frac{\lambda_{jl}}{\lambda_{jj}}\right)^2\right\}$

Pilot contamination (PC) (Strong interference) \quad Intra/inter-cell interference (Weaker)
Multi-Cell Scenarios and Imperfect Channel Knowledge (3)

- Multi-Cell Rate Expression not Amenable for Analysis
  - No closed-form optimization in multi-cell case
  - Numerical analysis still possible

- Similarities and Differences
  - Power consumption is exactly the same
  - Rates are smaller: Upper limited by pilot contamination:
    \[ R = B \log_2 \left( 1 + \frac{\alpha (M-K)}{\alpha \alpha \tau \text{ul} + \frac{1}{\alpha K \text{ul}} (1+K+I_{PC}) + \frac{1}{\alpha K (1+I_{PC})} - \alpha K (1+I_{PC}^2)} \right) \leq B \log_2 \left( 1 + \frac{1}{I_{PC}} \right) \]
  - Overly high rates not possible (but we didn’t get that...)
  - Clustering (fractional pilot reuse) might be good to reduce interference
Optimal Multi-Cell System Design: ZF Beamforming

**Optimum**

\[
M = 123 \\
K = 40 \\
\alpha = 0.28 \\
\tau^{(ul)} = 4
\]

User rates: \(
\approx 4\text{-QAM}
\)

**Massive MIMO!**

Many BS antennas

Note that \(M/K \approx 3\)

Global Optimum:

\[
M = 123, K = 40 \\
\text{EE} = 7.58 \text{ Mbit/J}
\]
Different Pilot Reuse Factors

**Higher Pilot Reuse**

Higher EE *and* rates!

Controlling inter-cell interference is very important!

**Area Throughput**

We only optimized EE

Achieved 6 Gbit/s/km² over 20 MHz bandwidth

METIS project mentions 100 Gbit/s/km² as 5G goal → Need higher bandwidth!
Energy Efficient to Use More Transmit Power?

- Recall from Theorem 2: Transmit power increases $M$
  - Figure shows EE-maximizing power for different $M$

Intuition: More Circuit Power $\rightarrow$ Use More Transmit Power
- Different from $\frac{1}{\sqrt{M}}$ scaling laws in recent massive MIMO literature
- Power per antennas decreases, but only logarithmically
### Summary

- **Optimization Results**
  - EE is a quasi-concave function of \((M, K, \alpha)\)
  - Closed-form optimal \(M, K, \) or \(\alpha\) for single-cell
  - Alternating optimization algorithm

<table>
<thead>
<tr>
<th>Increases with</th>
<th>Decreases with</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antennas (M)</td>
<td>Power (\alpha), coverage area (S_\lambda), and (M)-independent circuit power</td>
</tr>
<tr>
<td>Users (K)</td>
<td>Fixed circuit power (C_{0,0}) and coverage area (S_\lambda)</td>
</tr>
<tr>
<td>Transmit power (\alpha B \sigma^2 S_\lambda K)</td>
<td>Circuit power, coverage area (S_\lambda), antennas (M), and users (K)</td>
</tr>
</tbody>
</table>

Reveals how variables are connected

**Large Cell**
More antennas, users, RF power

**Massive MIMO**
Appears Naturally
Fractional pilot reuse important!

**More Circuit Power**
Use more transmit power

**Limits of \(M, K\)**
Circuit power that scales with \(M, K\)

**Simulations**
Depends on parameters
Download Matlab code to try other values!
Part 2

Questions?
Part 3

Massive MIMO
What is Massive MIMO?

- New Network Architecture
  - Use large arrays at BSs; e.g., $M = 123$ antennas, $K = 40$ users
  - Key: Excessive number of antennas, $M \gg K$
  - Very narrow beamforming
  - Little interference leakage

2013 IEEE Marconi Prize Paper Award


Analytic assumption: $M \to \infty$
What is the Key Difference?

• Number of Antennas?
  - 3G/UMTS: 3 sectors x 20 element-arrays = 60 antennas
  - 4G/LTE-A: 4-MIMO x 60 = 240 antennas

• We Already have Many Antennas!

**Massive MIMO Characteristics**

- Active antennas: Many antenna ports
- Coherent flexible beamforming
- Multi-user MIMO with many users

Image source: gigaom.com

Typical vertical array:
10 antennas x 2 polarizations
Only 1-2 antenna ports

Image source: gigaom.com

3 sectors, 4 vertical arrays per sector
• When to Deploy Massive MIMO?
  - Achieve high energy-efficiency!
  - Improve wide-area coverage
  - Special super-dense scenarios

• Co-located Deployment
  - 1D, 2D, or 3D arrays
  - One or multiple sectors

• Distributed Deployment
  - Remote radio heads
  - Cloud RAN
Original Motivation: Asymptotic Channel Orthogonality

- **Example: Uplink Transmission**
  - Two users channels: $\mathbf{h}_1, \mathbf{h}_2 \sim CN(\mathbf{0}, \mathbf{I}_M)$
  - Signals: $s_1, s_2 \sim CN(0, P)$
  - Noise: $\mathbf{n} \sim CN(\mathbf{0}, \mathbf{I}_M)$
  - Received: $\mathbf{y} = \mathbf{h}_1s_1 + \mathbf{h}_2s_2 + \mathbf{n}$

- **Linear Processing for User 1:** $\tilde{y}_1 = \mathbf{g}_1^H \mathbf{y} = \mathbf{g}_1^H \mathbf{h}_1s_1 + \mathbf{g}_1^H \mathbf{h}_2s_2 + \mathbf{g}_1^H \mathbf{n}$
  - Matched filter: $\mathbf{g}_1 = \frac{1}{M} \mathbf{h}_1$
  - Signal remains: $\mathbf{g}_1^H \mathbf{h}_1 = \frac{1}{M}||\mathbf{h}_1||^2 \xrightarrow{M \to \infty} E[|h_{11}|^2] = 1$
  - Interference vanishes: $\mathbf{g}_1^H \mathbf{h}_2 = \frac{1}{M} \mathbf{h}_1^H \mathbf{h}_2 \xrightarrow{M \to \infty} E[h_{11}^H h_{21}] = 0$
  - Noise vanishes: $\mathbf{g}_1^H \mathbf{n} = \frac{1}{M} \mathbf{h}_1^H \mathbf{n} \xrightarrow{M \to \infty} E[h_{11}^H n_1] = 0$

Asymptotically noise/interference-free communication: $\tilde{y}_1 \xrightarrow{M \to \infty} s_1$
Does This Hold for Practical Channels?

• Initial Measurements: Show similar results


Achievable Rates
Only 10-20% lower than with i.i.d. channels

Main Research Challenges

• Acquisition of Channel State Information
  - Finite coherence block $U \in [100, 10000]$
  - Only $\leq U$ unique pilots $\rightarrow$ Reuse across cells
  - BS cannot tell difference between users
  - Pilot contamination: Correlated estimates
  - This interference doesn’t vanish as $M \rightarrow \infty$

• Not a New Phenomenon
  - Pilot contamination always an issue
  - More pronounced when $M$ and $K$ are large

• Current Solutions:
  - Simple: Fractional pilot reuse
  - Advanced: Exploit spatial correlation
Main Research Challenges (2)

- **Frequency Division Duplex (FDD)**
  - Many systems and spectrum bands are dedicated to FDD
  - Cannot rely on channel reciprocity → Is estimation overhead too large?

- **Computational Complexity**
  - ZF performs better than MRC/MRT but has higher complexity
  - Can complexity be reduced with retained performance?

- **Circuit Design and Hardware Implementation**
  - Cost and power increase in massive MIMO, but as $N$, $\sqrt{N}$, or slower?
  - Can waveforms be design to allow more efficient hardware?
MAMMOET Project

- 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

[Logos of participating institutions]
Part 3

Questions?
Part 4

Multi-Objective Network Optimization
Optimize more than Energy-Efficiency

• Recall: Many Metrics in 5G Discussions
  - Average rate (Mbit/s/active user)
  - Average area rate (Mbit/s/km²)
  - Energy-efficiency (Mbit/Joule)
  - Active devices (per km²)
  - Delay constraints (ms)

• So Far: Only cared about EE
  - Ignored all other metrics

Optimize Multiple Metrics

We want efficient operation w.r.t. all objectives

Is this possible?
For all at the same time?
Basic Assumptions: Multi-Objective Optimization

• Consider $N$ Performance Metrics
  - Objectives to be maximized
  - Notation: $g_1(x), g_2(x), \ldots, g_N(x)$
  - Example: individual user rates, area rates, energy-efficiency

• Optimization Resources
  - Resource bundle: $\mathcal{X}$
  - Example: power, resource blocks, network architecture, antennas, users
  - Feasible allocation: $x \in \mathcal{X}$
Single or Multiple Performance Metrics

• Conventional Optimization
  - Pick one prime metric: $g_1(x)$
  - Turn $g_1(x), g_2(x), ..., g_N(x)$ into constraints

  - Optimization problem:
    \[
    \begin{align*}
    & \text{maximize} \quad g_1(x) \\
    & \text{subject to} \quad x \in \mathcal{X}, \quad g_2(x) \geq C_2, ..., g_N(x) \geq C_N.
    \end{align*}
    \]

    - Solution: A scalar number

    - Cons: Is there a prime metric? How to select constraints?

• Multi-Objective Optimization
  - Consider all $N$ metrics
  - No order or preconceptions!

  - Optimization problem:
    \[
    \begin{align*}
    & \text{maximize} \quad [g_1(x), g_2(x), ..., g_N(x)] \\
    & \text{subject to} \quad x \in \mathcal{X}.
    \end{align*}
    \]

    Solution: A set
    Pareto Boundary
    Improve a metric \rightarrow Degrading another metric
Why Multi-Objective Optimization?

• Study Tradeoffs Between Metrics
  - When are metrics aligned or conflicting?
  - Common in engineering and economics – new in communication theory

**A Posteriori Approach**
Generate region (computationally demanding!)
Look at region and select operating point

![Diagram showing highly conflicting and relatively aligned regions](image)
A Priori Approach

- No Objectively Optimal Solution
  - Utopia point outside of region $\rightarrow$ Only subjectively “good” solutions exist

- System Designer Selects Utility Function $f : \mathbb{R}^N \rightarrow \mathbb{R}$
  - Describes subjective preference (larger is better)

- Examples: Sum performance: $f(g) = \sum_k g_k$
  - Proportional fairness: $f(g) = \prod_k g_k$
  - Harmonic mean: $f(g) = K_r (\sum_k g_k^{-1})^{-1}$
  - Max-min fairness: $f(g) = \min_k g_k$

We obtain a simplified problem:

\[
\begin{align*}
\text{maximize} & \quad f(g_1(x), g_2(x), \ldots, g_N(x)) \\
\text{subject to} & \quad x \in \mathcal{X}
\end{align*}
\]

- Solution: A scalar number (Gives one Pareto optimal point)
- Takes all metrics into account!
Example: Optimization of 5G Networks

- **Design Cellular Network**
  - Symmetric system
  - 16 base stations (BSs)
  - Select:
    - $M = \# \text{ BS antennas}$
    - $K = \# \text{ users}$
    - $P = \text{ power/antenna}$

- Resource bundle:

$$\mathcal{X} = \begin{cases} 
[K \ M \ P]^T : & 
1 \leq K \leq \frac{M}{2}, \\
2 \leq M \leq M_{\max}, \\
0 \leq P \leq M P_{\max} 
\end{cases}$$

250 meters

$M$ transmit antennas

$K$ uniformly distributed users

$500$

$20 \text{ W}$
Example: Optimization of 5G Networks (2)

- **Downlink Multi-Cell Transmission**
  - Each BS serves only its own $K$ users
  - Coherence block length: $U$
  - BS knows channels within the cell (cost: $K/U$)
  - ZF beamforming: no intra-cell interference
  - Interference leaks between cells

- **Average User Rate**

$$R_{\text{average}} = B \left(1 - \frac{K}{U}\right) \log_2 \left(1 + \frac{P}{K} \frac{(M - K)}{S \lambda \sigma^2 + J}\right)$$

- **Power/user**
- **Array gain**
- **Bandwidth (10 MHz)**
- **CSI estimation overhead ($U = 1000$)**
- **Noise / pathloss ($1.72 \cdot 10^{-4}$)**
- **Relative inter-cell interference (0.54)**
Example: Optimization of 5G Networks (3)

• What Consumes Power?
  - Transmit power (+ losses in amplifiers)
  - Circuits attached to each antenna
  - Baseband signal processing
  - Fixed load-independent power

• Total Power Consumption

\[ P_{\text{total}} = \frac{P_{\text{trans}}}{\eta} + C_{0,0} + C_{1,0}K + C_{0,1}M + \frac{B C_{\text{beamforming}}}{U L_{\text{BS}}} \]

- Amplifier efficiency (0.31)
- Fixed power (10 W)
- Circuit power per user (0.3 W)
- Circuit power per antenna (1 W)
- Computing ZF beamforming (\(2.3 \cdot 10^{-6} \cdot MK^2\))
Example: Results

3 Objectives

1. Average user rate
   \[ g_1(x) = R_{\text{average}} \] [bit/s/user]

2. Total area rate
   \[ g_2(x) = \frac{K}{A} R_{\text{average}} \] [bit/s/km²]

3. Energy-efficiency
   \[ g_3(x) = \frac{K R_{\text{average}}}{P_{\text{total}}} \] [bit/J]

Observations

Area and user rates are conflicting objectives

Only energy efficient at high area rates

Different number of users
Example: Results (2)

- Energy-Efficiency vs. User Rates
  - Utility functions normalized by utopia point

**Observations**

Aligned for small user rates

Conflicting for high user rates
Part 4

Questions?
Summary

• What if a Cellular Network is Designed for High Energy-Efficiency?
  - Energy-efficiency [bit/Joule] = $\frac{\text{Average Sum Rate [bit/s/cell]}}{\text{Power Consumption [Joule/s/cell]}}$
  - Necessary: Accurate rate expressions and power consumption
  - Design parameters: Number of users, antennas, and transmit power

• Analytical and Numerical Results
  - Reveals interplay between system parameters
  - Shows that massive MIMO is the energy-efficient solution

• Main Properties of massive MIMO
  - Arrays with many active antennas and relatively many users

• Multi-Objective Optimization
  - Framework to jointly optimize energy-efficiency and other 5G metrics
References


QUESTIONS?

Papers, Presentations, and Simulation Code
All Available on my Homepage:
http://www.commsys.isy.liu.se/en/staff/emibj29