KTH ROYAL INSTITUTE
OF TECHNOLOGY

Green 5G Mobile Networks

Cicek Cavdar

15 January 2016
UC-Davis
Friday Semiar organized by Networks Lab

KTH-Royal Institute of Technology

- **RSLab**
- **Wireless@KTH Research Center**
  - Leader: Prof. Jens Zander
  - Staff of 21 includes, 6 PhD students and 15 senior researchers

Projects

<u>EU EIT Projects</u>

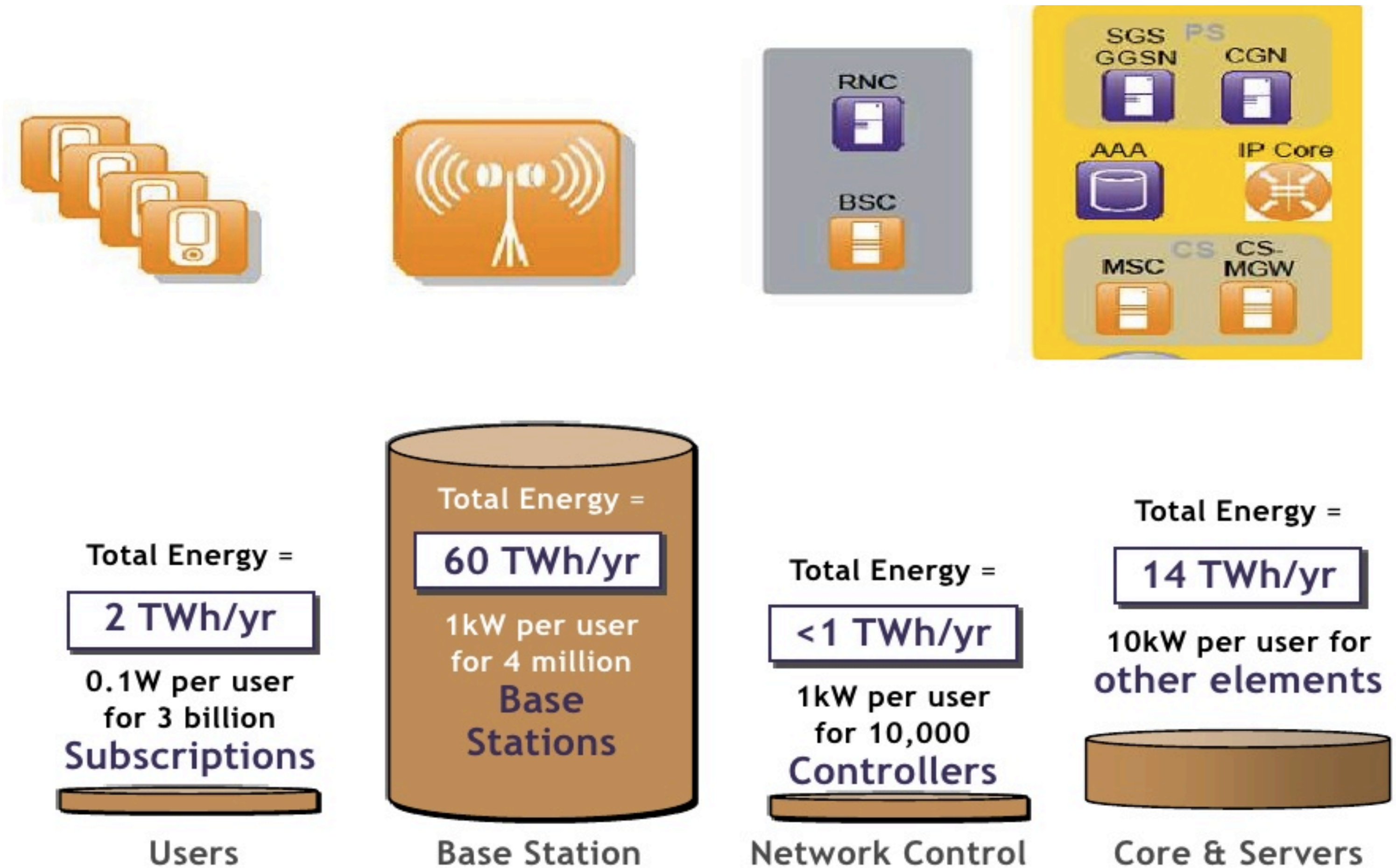
- **5GrEEn** *Green 5G Mobile Networks*
  - **2012-2014,** ≈870k€, 340k€ KTH
- **EXAM** *Energy Efficient Xhaul and M2M* **2015,** ≈100k€ KTH
- **ACTIVE** *Advanced ConnecTIvity platform for IoT VErtical segments 2016-2018 (SDN and Edge-Cloud)* ≈300k€ KTH
- **Seamless DA2GC** *with 5G Radio Technologies in Europe*
  - **2016,** ≈240k€ KTH

<u>EU CELTIC Plus</u>

SooGREEN *Service Oriented Optimization of Green Mobile Networks (Cloud-RAN, CTD, SDN, NFV)*

- **2016-2018,** 750k€ KTH

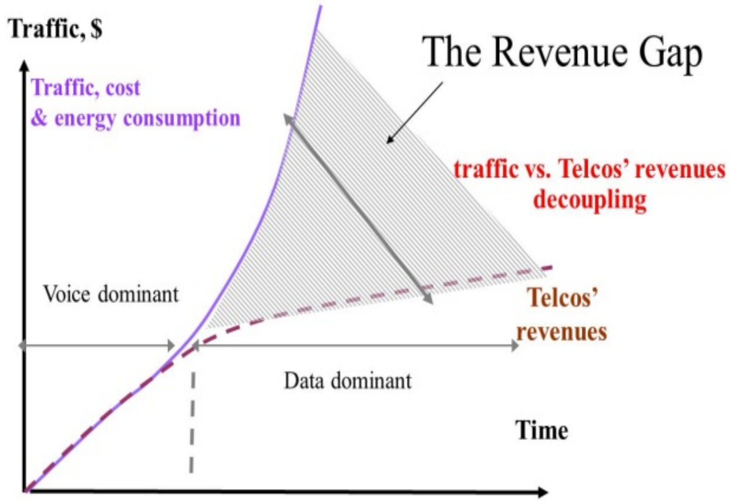
Mobile Communications Power Consumption



Based on: ETSI RRS_05024, NSN, GreenTouch (First presented May 2010, Greentouch presentation at 2012 @ Green Telecom and IT Workshop)

Motivation

- What is the consequence?

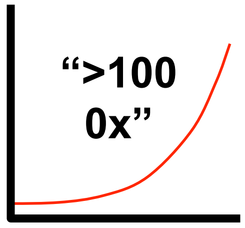


Low energy consumption is key!!

- **Energy Consumption: X2 every 5 years**
  - Densification + New roll out
- **Unit energy cost: X3 in 7 years!**

For Operators ← **For Governments**

5G Challenges and Energy Consumption



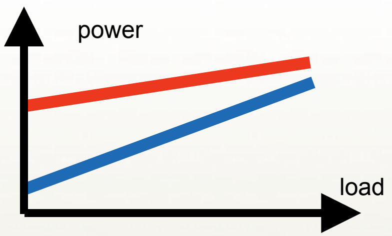
- 5G Challenges:
  - Thousand-fold traffic increase
  - Hundreds of billions of devices
  - Diverse requirements (latency, reliability, spectrum) etc.
  - Affordable, sustainable, and feasible

- 5GrEEn target: Factor of 10 reduction of energy consumption versus today and fulfilling all other requirements!
  - EARTH: Factor of 4 reduction vs 2012 baseline
  - GreenTouch: Factor of 10 reduction vs 2010 baseline

Main areas for improving energy Performance

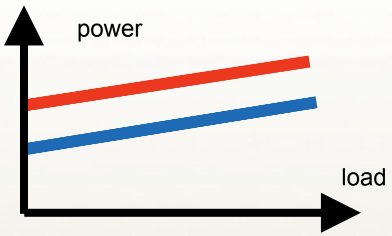
Standardization

"Design energy efficient systems from the start"



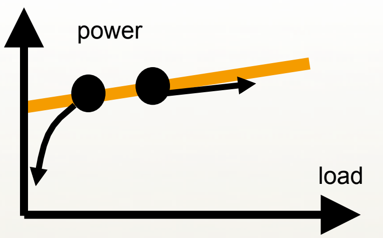
Product Improvements

"State of the art energy lean hardware and software"



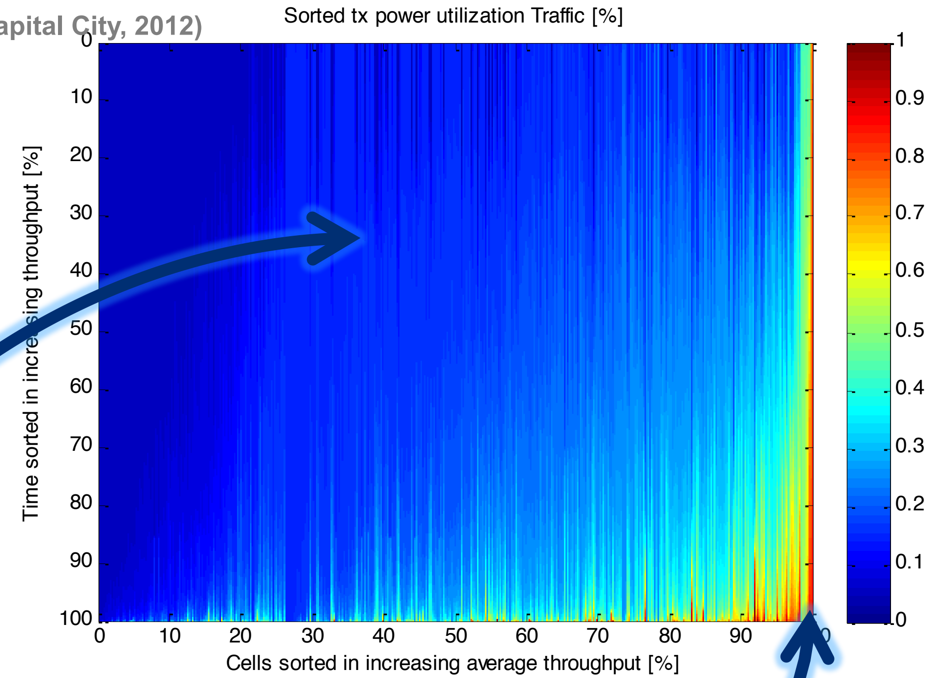
Network management

"Reduce overall energy consumption in case of excess capacity"



Current Network:
Traffic Measurements

(Central parts of Major European Capital City, 2012)



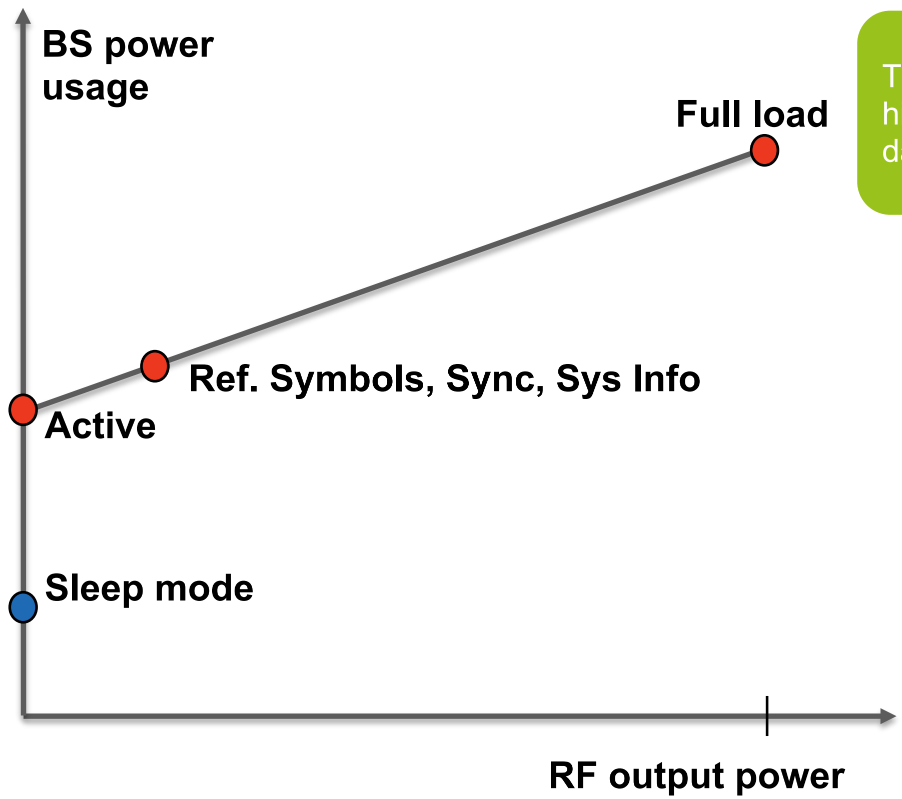
- **High traffic is an exception**
  - Median HS traffic is ≈ 50 kbit/s
  - Median 3G CS voice traffic is ≈ 0.2 Erlang

"Most parking spaces are unused most of the time"

"But not the one I want to use when I want to use it"

Prepared by Pål Frenger, Ericsson AB.

Base Station
Energy performance



Traditional focus on high capacity, peak data rates and delay

Energy performance requires addressing low traffic cases

Base Station
Energy performance

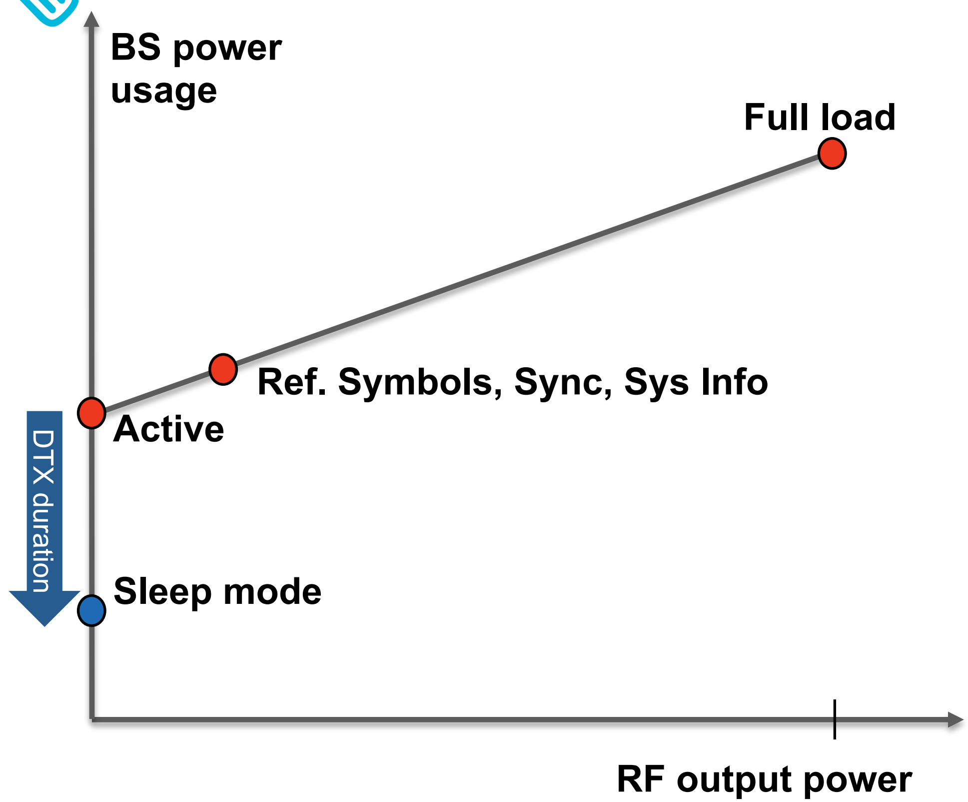
OPEX Cost: Electricity bill

CAPEX Cost: Size of solar panels and battery backup

CO_2 footprint

Type of power solution that is feasible (e.g. solar or diesel)

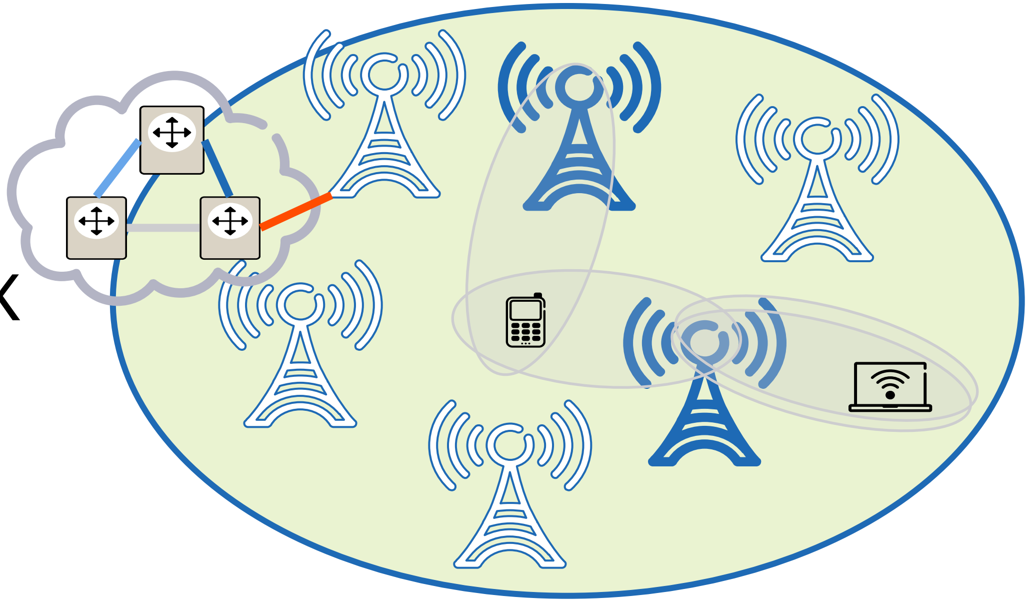
Feasibility of providing low cost wide area coverage



Focus areas & potential solutions: System architecture

- From ***always on***… to ***always available!***

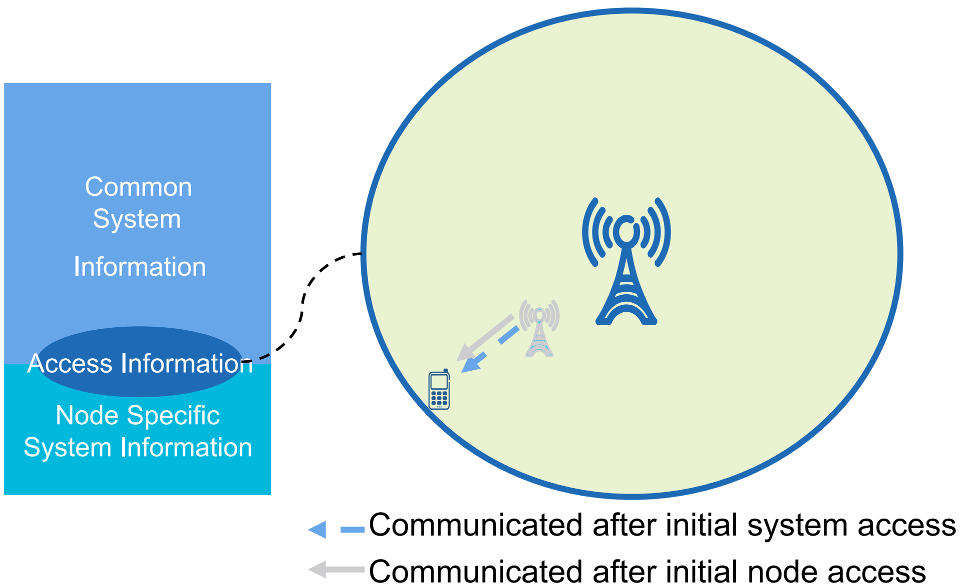
- Logical decoupling of system plane and user plane
  - Cells are dynamically configured to support active users/devices
  - Enables BS DTX/DRX and high gain beamforming



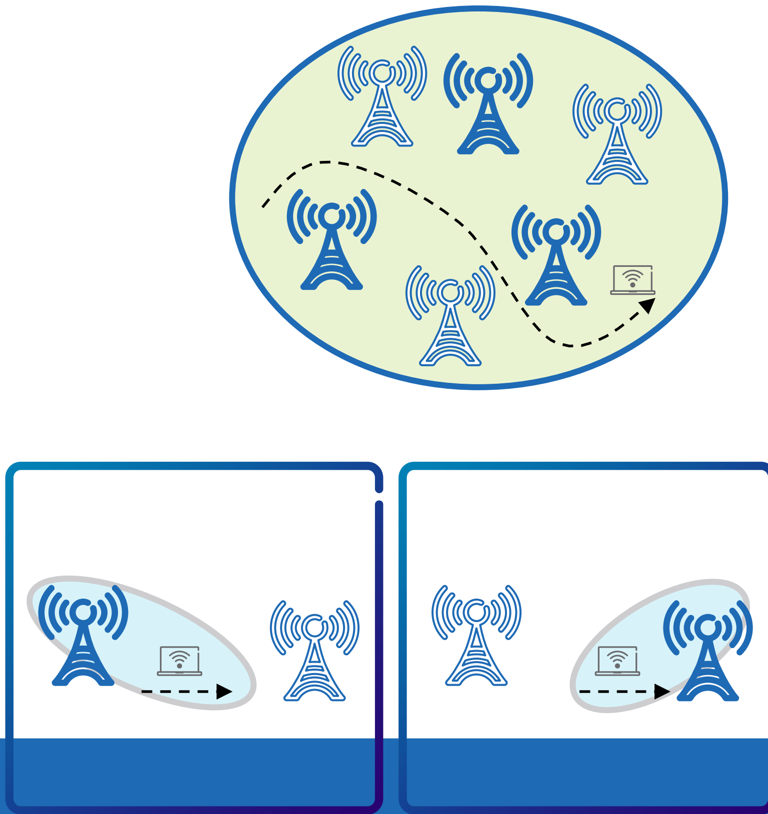
EE When not transmitting data

Ultra-thin 5G-NX Control Plane

Minimize Broadcast



Separate active and idle mode mobility

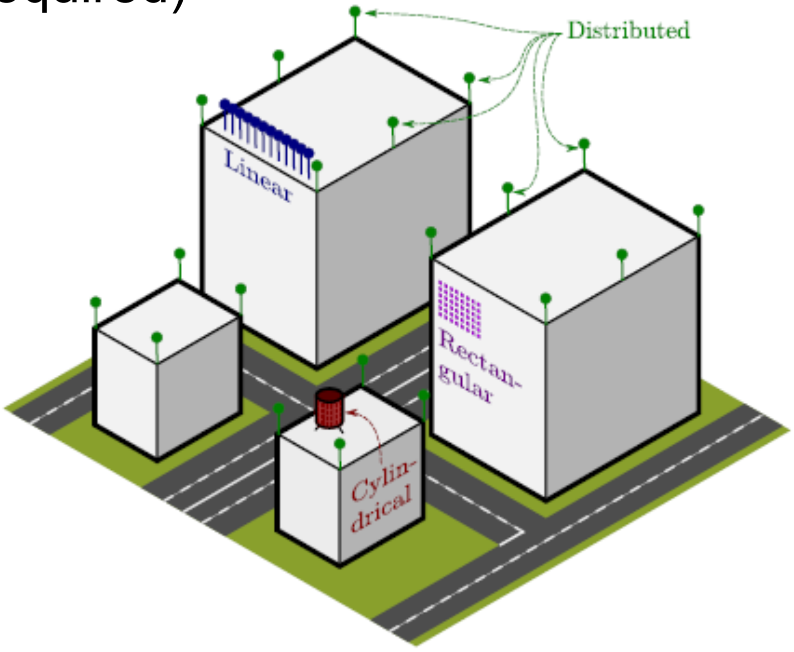


Magnus Olsson, Pål Frenger, Ericsson AB.

EE When transmitting data: Operation

Very Large MIMO

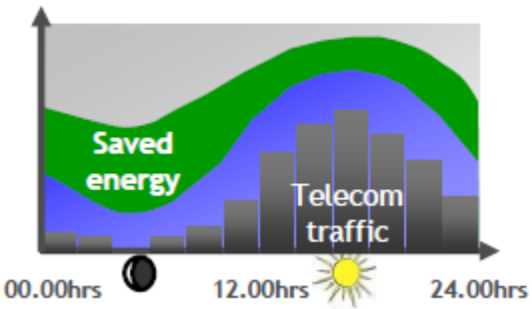
- Why:
  - Focus emitted energy to where the terminals are located
  - Improve data rates (more sleep mode)
  - Reduce interference (less tx-power required)

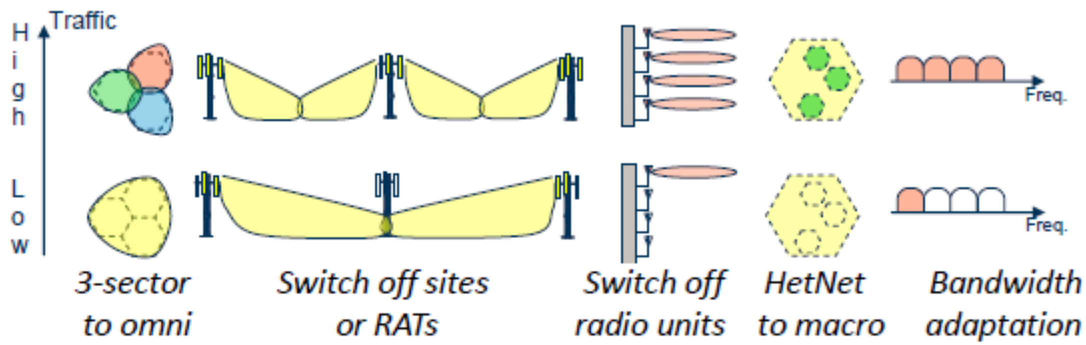


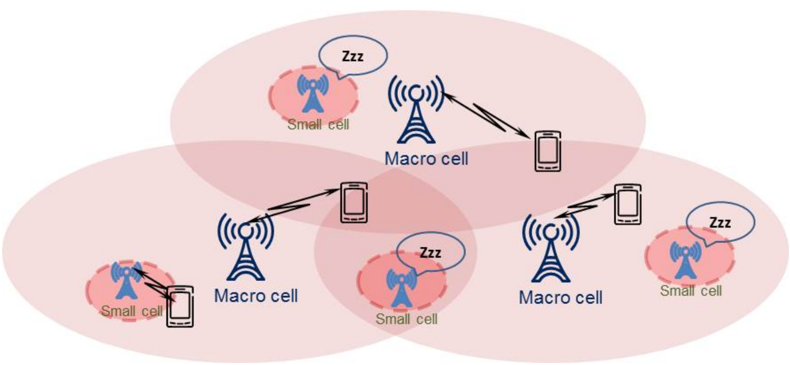
Source: E. G. Larsson et. al. “Massive MIMO for Next Generation Wireless Systems”

When Transmitting Data - Operation -

- **How much** energy we can save at low load scenarios via traffic adaptive macro(~hour)- and micro-level(~ms) sleep techniques?







Outline

- BS Densification Cell Dtx and Small Cell Offload
- Joint BW and Power optimization with QoS Guarantee
- Energy Efficient Load Adaptation in Massive MIMO Systems (optimization of number of antennas per BS to maximize EE over the day)
- Network Sharing Energy Efficiency Benefits

KTH ROYAL INSTITUTE
OF TECHNOLOGY

1/4 Energy Savings with BS Densification, Cell DTX and Small Cell Offloading

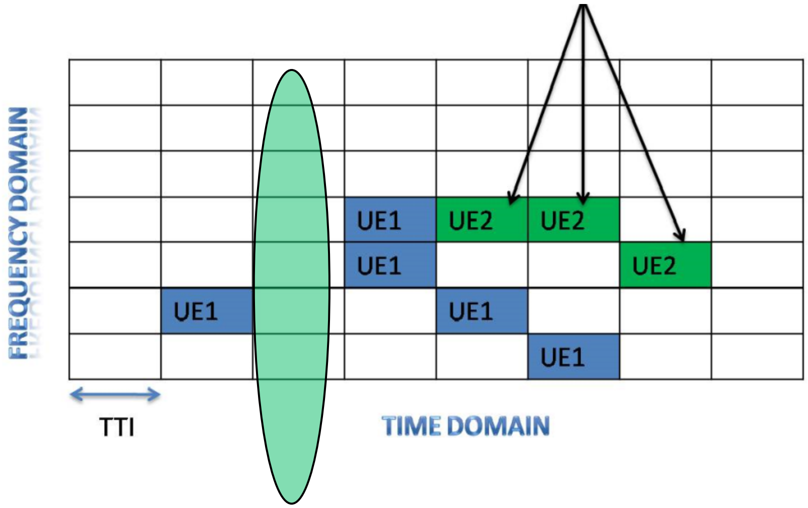
Traffic-Adaptive Network Operation

Micro Sleep – Cell DTX

Solution:

Hardware and software upgrade to enable cell DTX at BS side.

Cell DTX: Switch off the PA during the TTI's that are fully devoted to data transmission when there is no traffic.



Switch off the PA (microsecond level)

Energy saving

Cell DTX in Small Cell Deployments 1/6

Q:

- How much Power we can save by Cell DTX
- How much traffic can we offload from macro layer?
- Can we save power by small cell offloading ?

Important: Calculation of "Cell Activity Factor" by considering interference

- Given cell traffic → interference → data rates → transmitter activity → interference → …

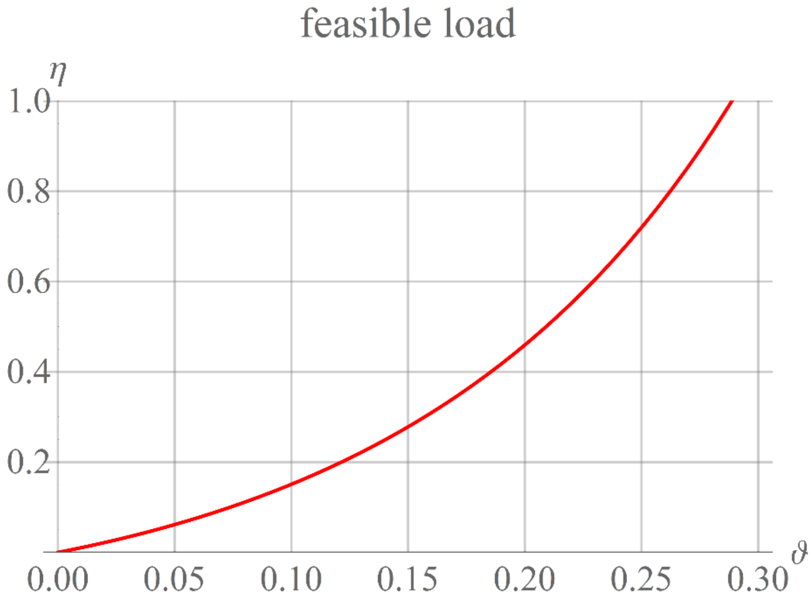
$$\eta = f(\eta)$$

Cell Activity

- Define ”offered load” as a function of N-active-users, file size over an ovservation period / Bandwidth, Max-SE

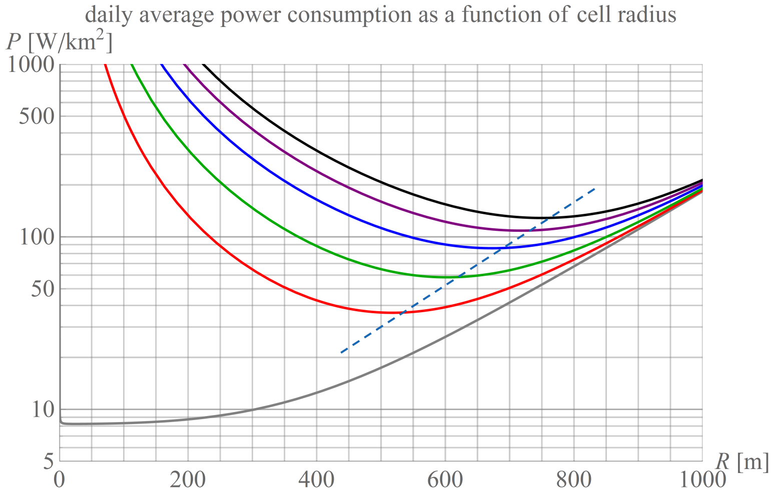
Solve the fixed point

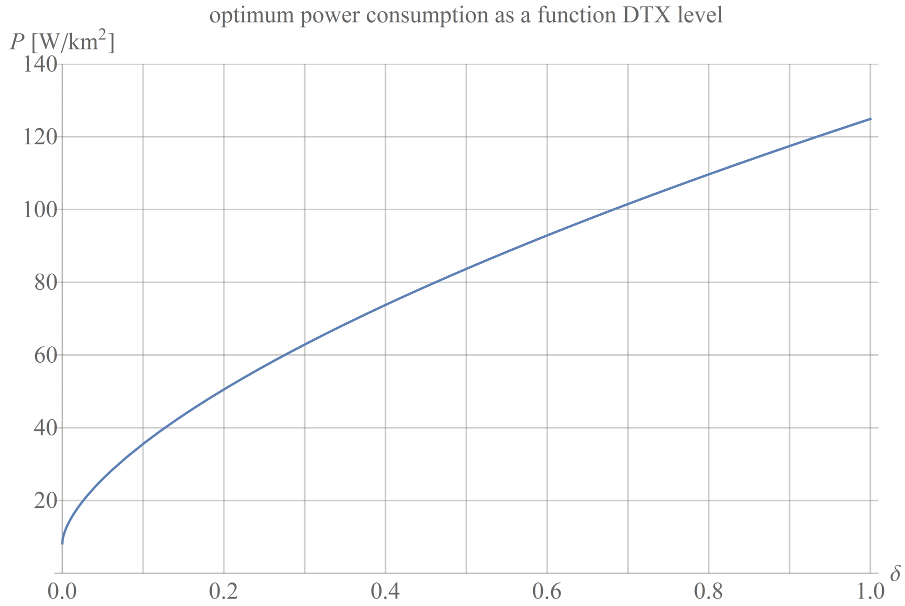
To calculate the Cell Activity



Cell DTX Area Power Savings

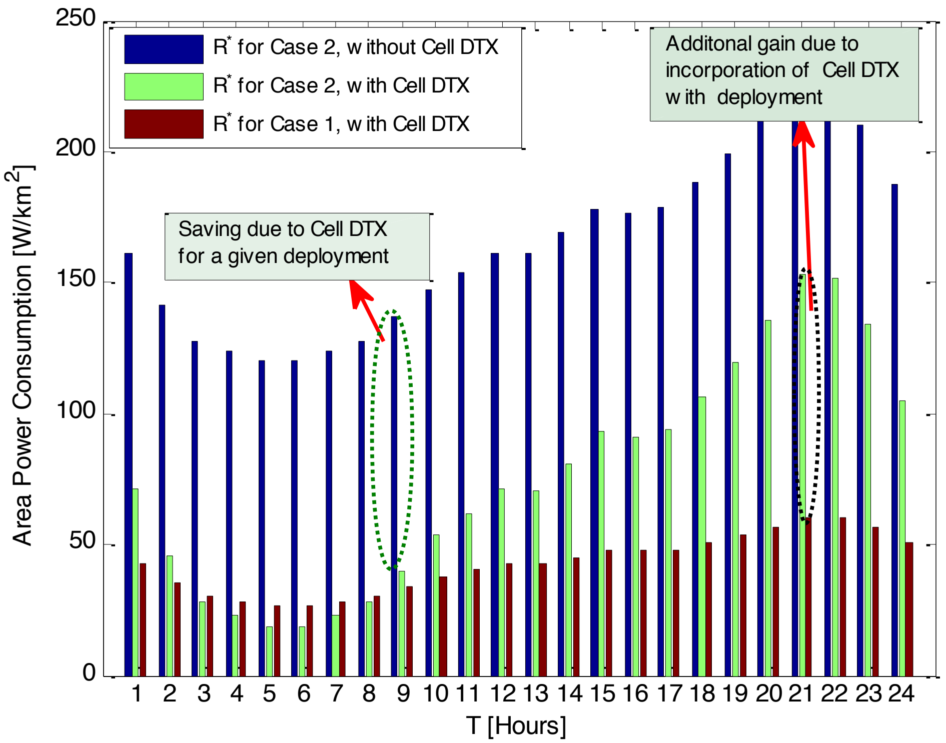
- When we average power consumption expression over 19 cells and changing offered traffic in 24 hours, we get:





- Plotting the Pmin at optimum radius against delta gives the second figure.
- If $\delta = 0.5$, we save 1/3 power compared to $\delta = 1$ (no DTX).

Cell DTX Area Power Savings

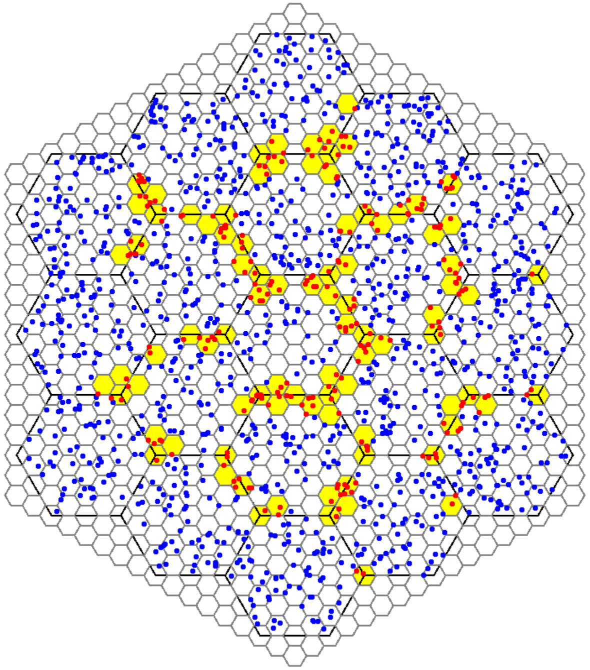


- Cell DTX brings striking energy saving (from blue to green bar) for a given network deployment.
- However, additional <u>42 percent</u> saving is achievable by designing the network under the assumption that cells can be put into DTX mode during idle periods.

Cell DTX in Small Cell Deployments

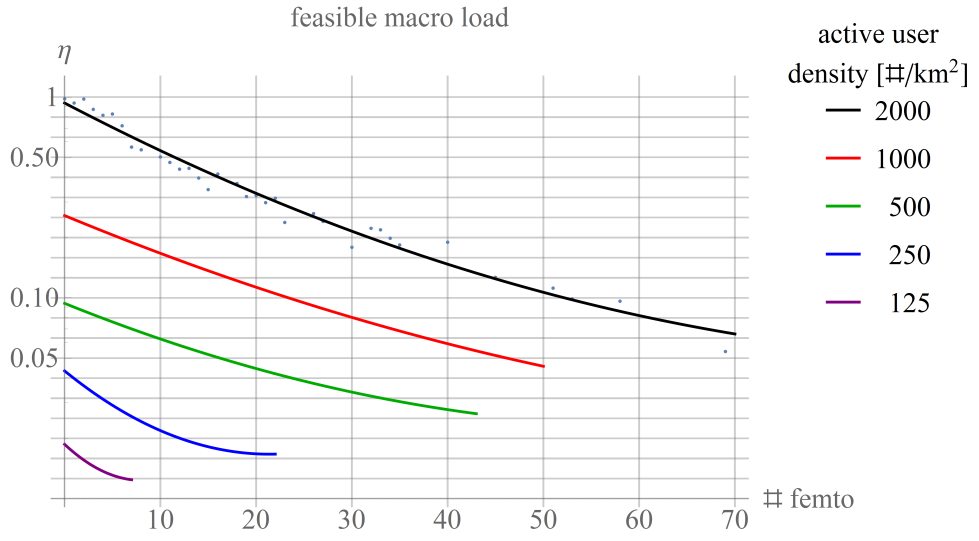
- System model:
  - Macro cell ISD: 500m
  - Femto cell ISD: 50m
  - Offloading femtos deployed where users receive worst rate.
  - Macro TX power: 20W
  - Femto TX power: 0.05W

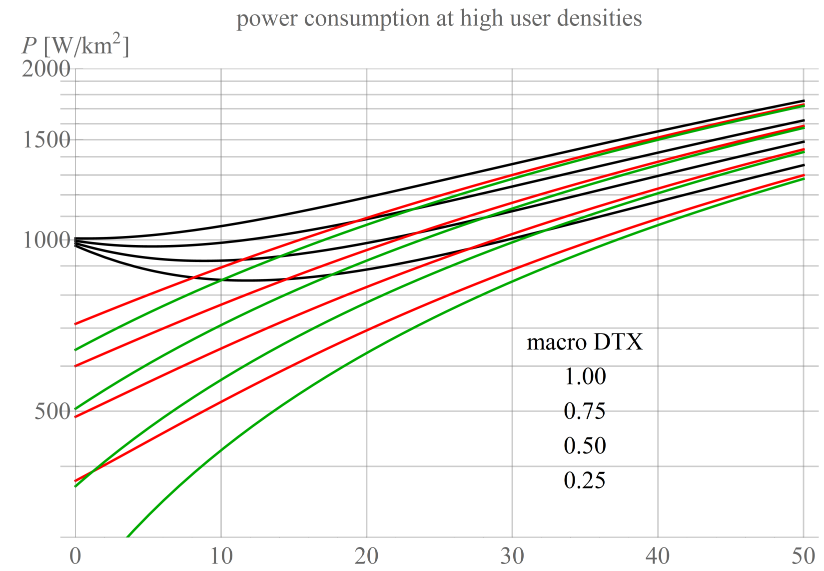
  - No cell-DTX at femto cells.



Cell DTX in Small Cell Deployments

- Adding femtos reduce the time-load of the macro BSs.

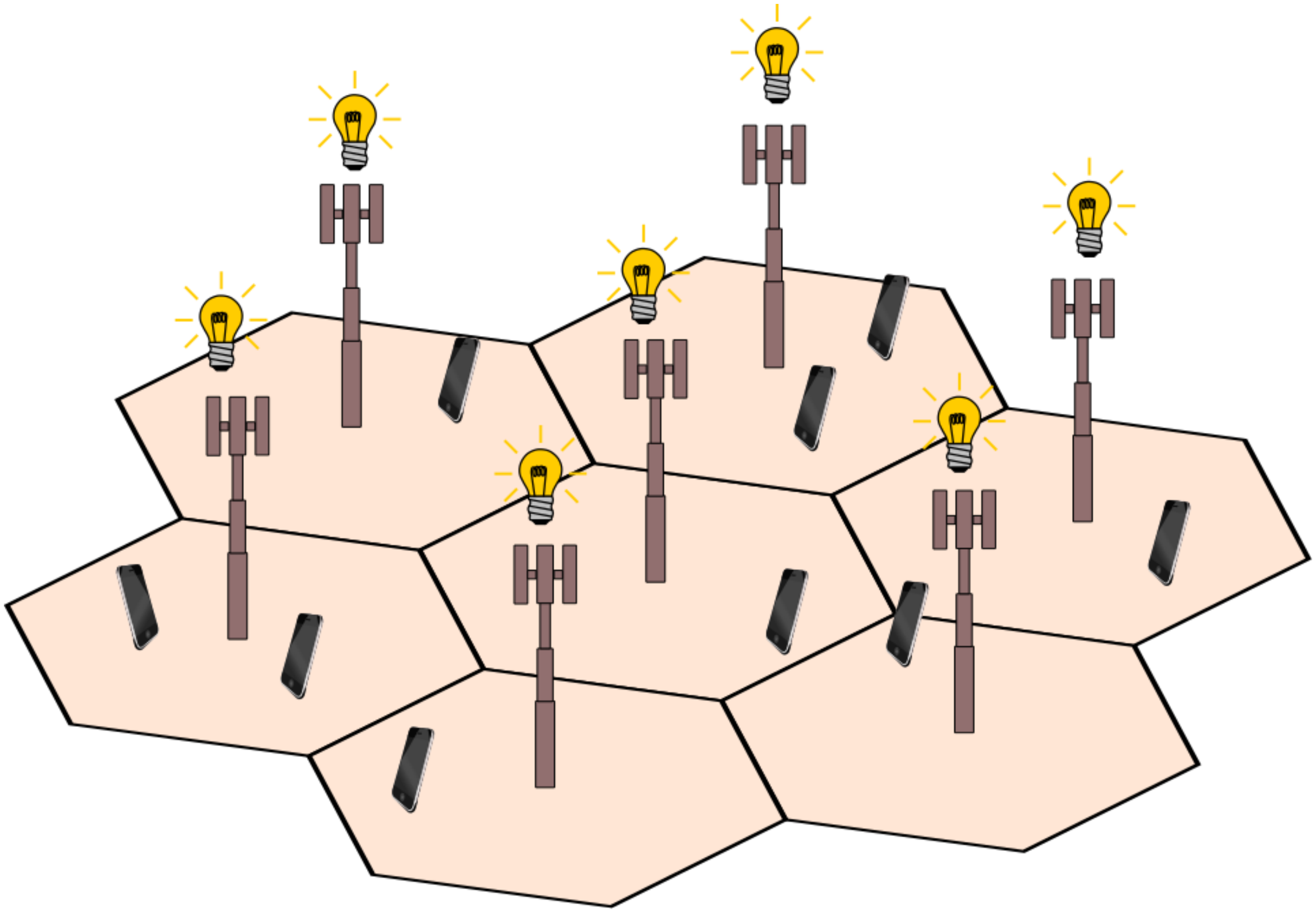




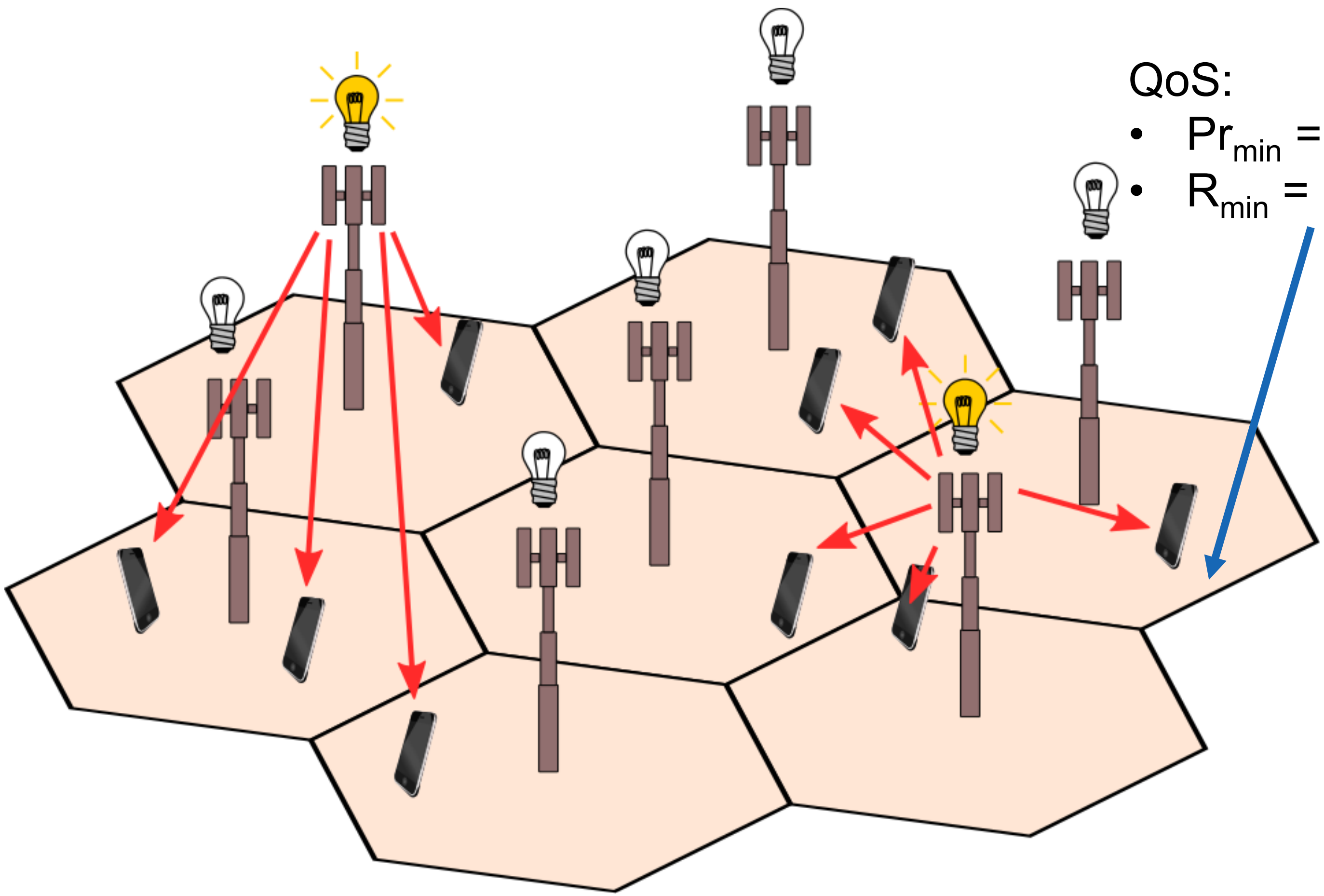
- If we consider area power consumption in second figure:
- Offloading saves power when macro is very loaded.

2/4 Joint BW and Power Allocation with QoS Guarantee

Problem description



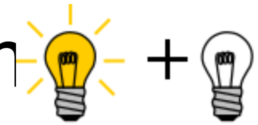
Problem description



QoS:

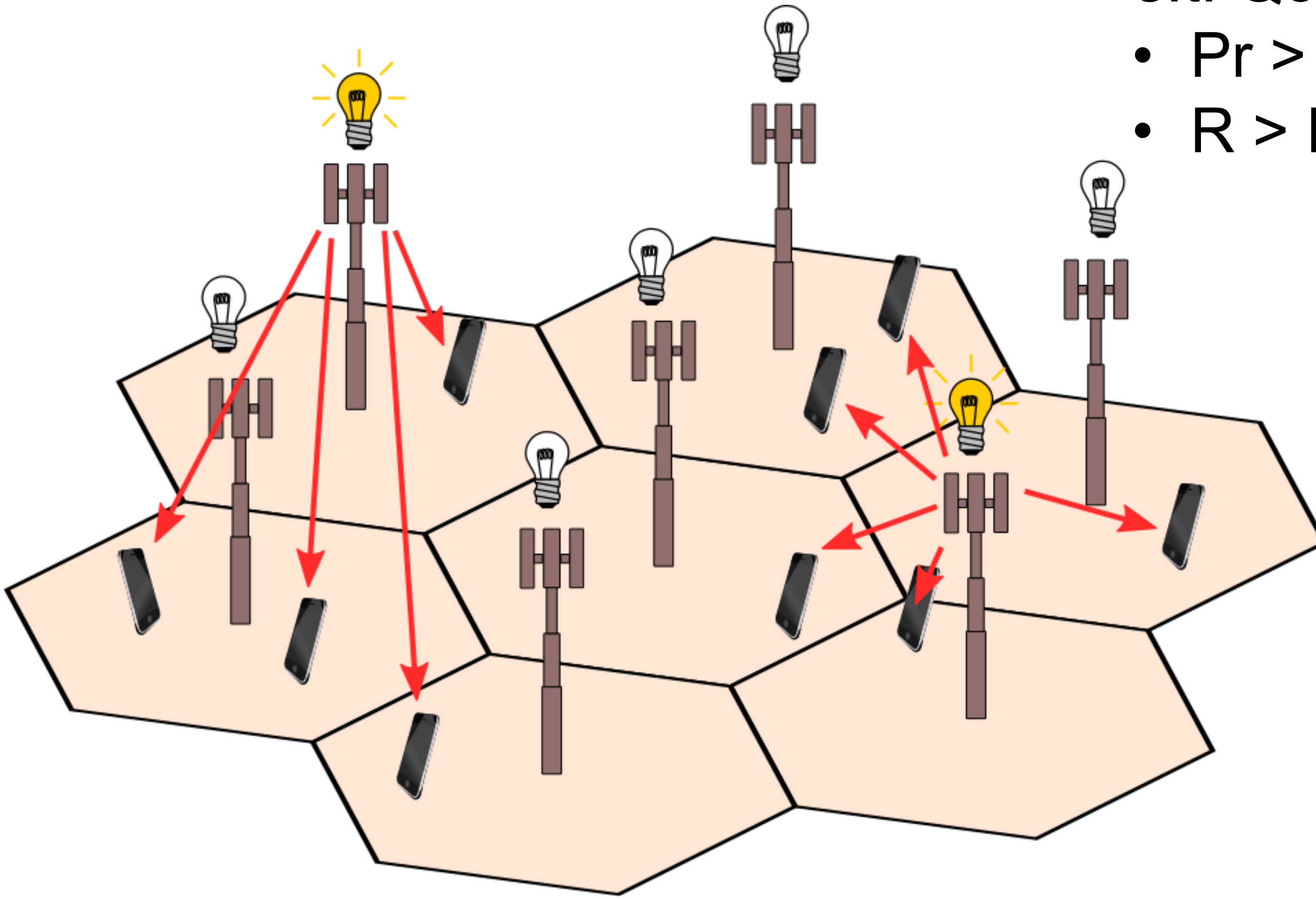
- Pr_{min} = –90dBm
- R_{min} = 500kbps

Problem description

m  +

s.t. QoS satisfied:

- $Pr > Pr_{min}$
- $R > R_{min}$



Solution approach

- Estimate performance of a large system involving an optimization problem.
  ➔ Perform stochastic simulations.
  ➔ Solve optimization problem using IBM CPLEX.

- Trouble: $R = W \log_2(1+SINR)$ ➔ Nonlinear.

- So, perform resource allocation in stages:
  1. Determine BS-UE assignment to minimize power consumption while guaranteeing Pr_{min}
     - MIQP ➔ solvable using CPLEX.
  2. Allocate BW to UEs.
     - Several possible approaches (equal, proportional fair)
  3. Perform power control to improve SINR, thus Rate.
     - Also reduces power consumption.
  4. If all QoS requirements not satisfied, repeat from step 1 using higher Pr_{min} requirement.

BS-UE assignment

$$\arg\min_{\pi,x,\zeta} \sum_{i=1}^{N} [(a \sum_{j=1}^{M} \pi_{ij} x_{ij} + P_0)\zeta_i + (1-\zeta_i)P_{sleep}]$$

Minimize total power consumption.

$$\text{s.t.} \quad \sum_{j=1}^{M} x_{ij} \le N_p \quad \forall i \in \mathcal{B}$$

Each BS serves at most N_{PRB} UEs.

$$\sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij} = M$$

All UEs are served.

$$\sum_{i=1}^{N} x_{ij} = 1 \quad \forall i \in \mathcal{U}$$

Each UE is served by only 1 BS.

$$c_{ij} \le \frac{\pi_{ij} \cdot \sigma_{ij}}{P_{MINj}} \quad \forall i \in \mathcal{B} \quad \forall j \in \mathcal{U}$$

A covered UE has Pr > P_{min}

$$c_{ij} - x_{ij} \ge 0 \quad \forall i \in \mathcal{B} \quad \forall j \in \mathcal{U}$$

A served UE is covered.

$$\zeta_i \ge x_{ij} \quad \forall j \in \mathcal{U} \quad \forall i \in \mathcal{B}$$

If a BS is serving any UE, it is on.

$$\sum_{j=1}^{M} \pi_{ij} \le P_{MAX} \quad \forall i \in \mathcal{B}$$

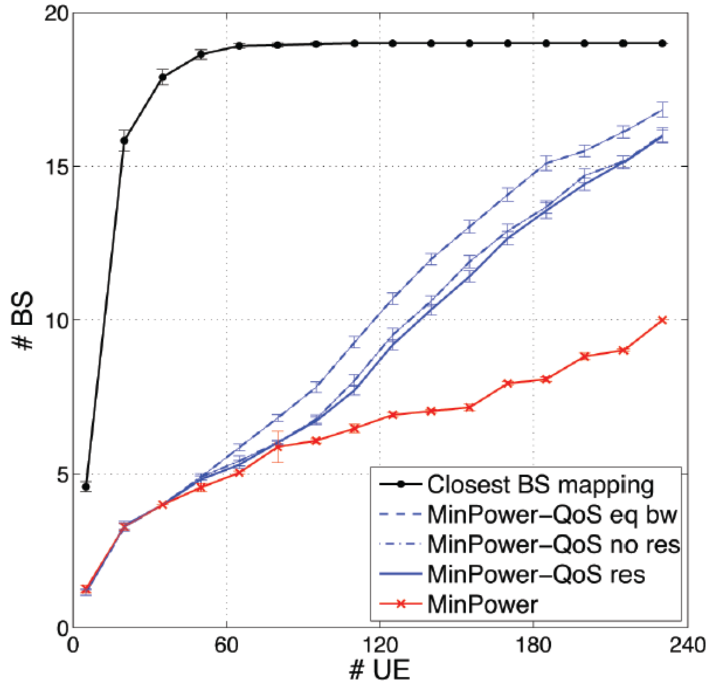
Sum of power allocated by a BS does not exceed BS's power budget.

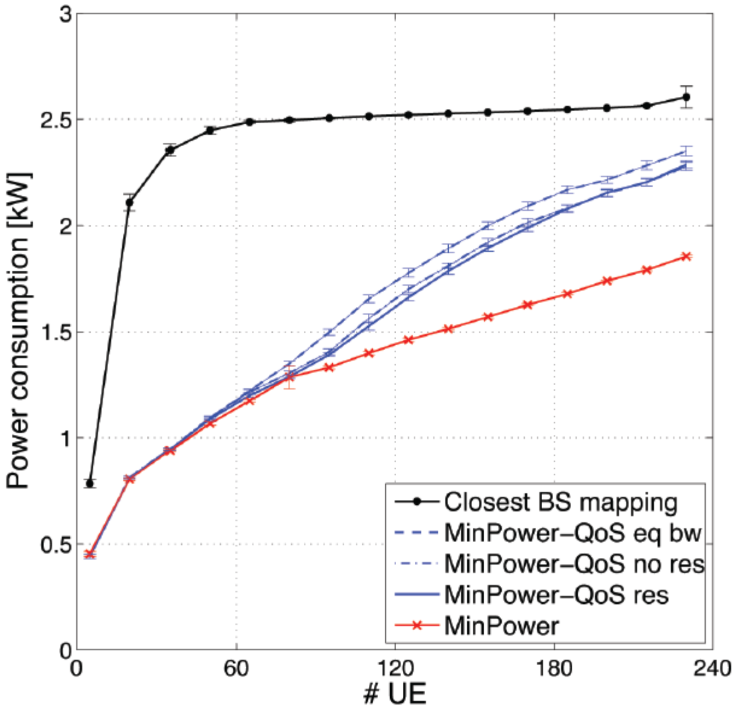
"Energy efficient adaptive cellular network configuration with QoS guarantee"

- Pierpaolo Piunti, Cicek Cavdar, Simone Morosi, Kaleab Ejigayehu Teka, Enrico Del Re, Jens Zander. ICC 2015 in London.

Results

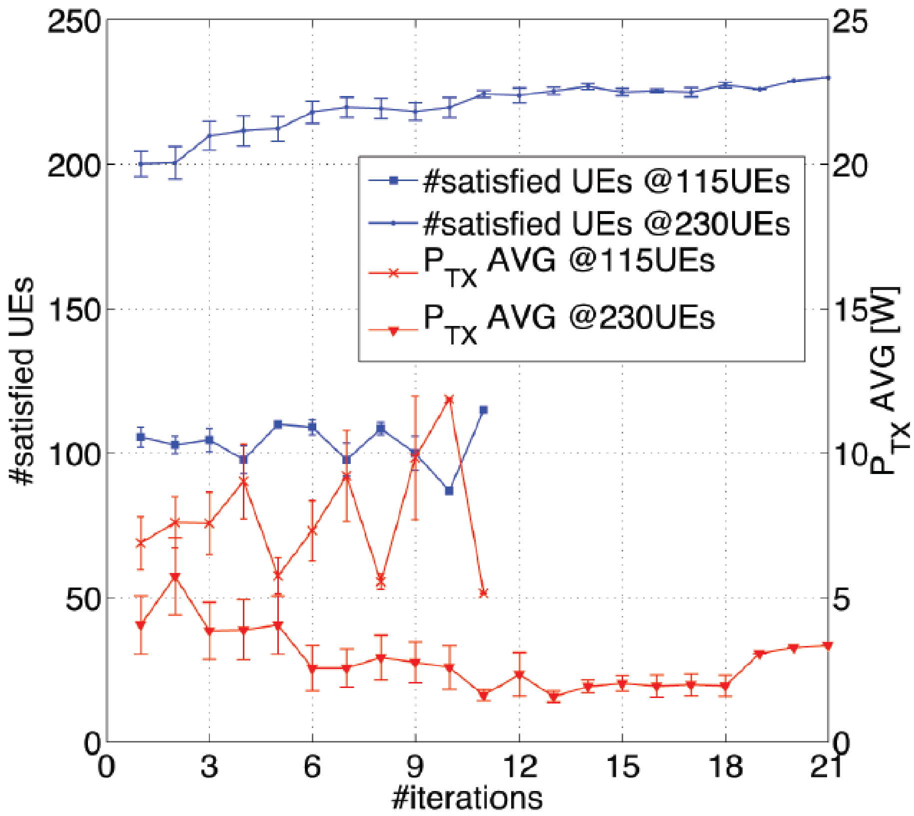
- Total power consumption vs load



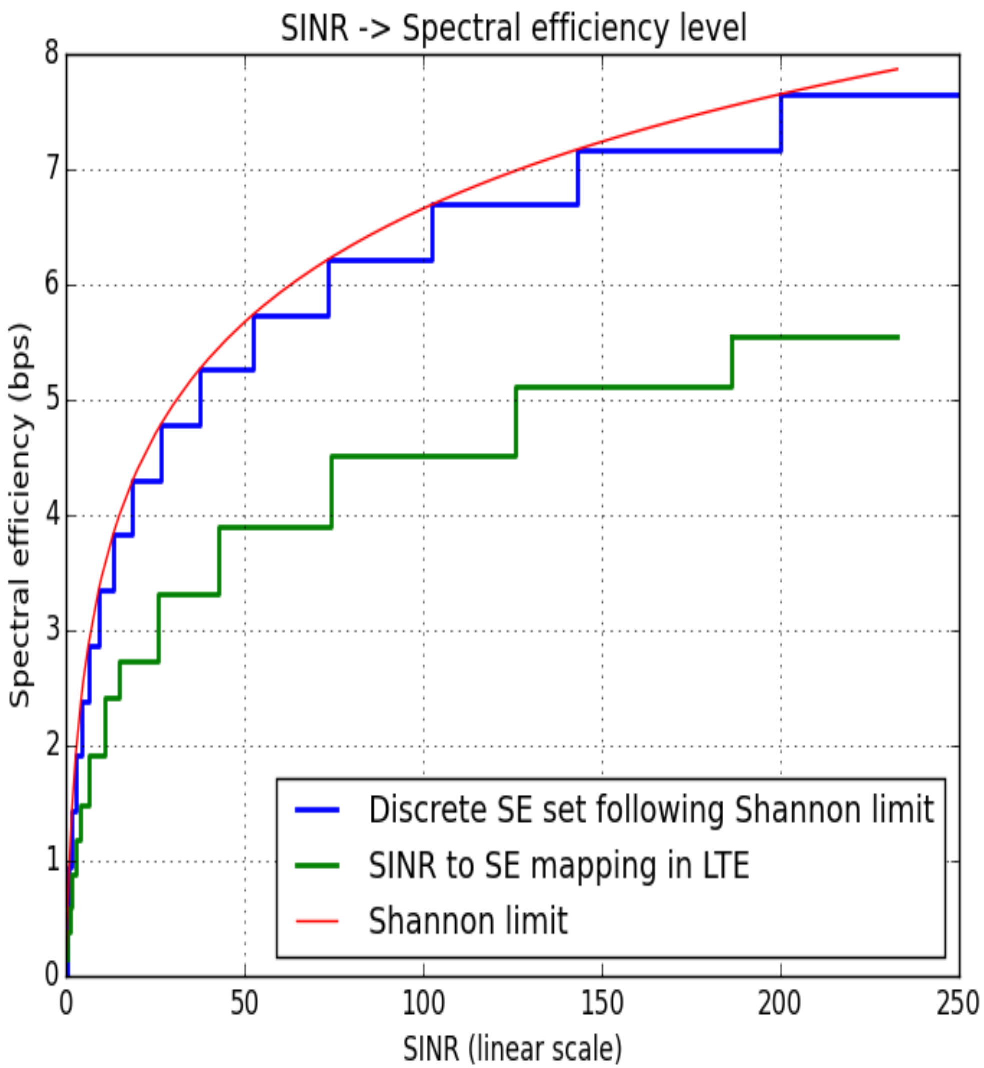


Results

- On rate of convergence of proposed algorithm



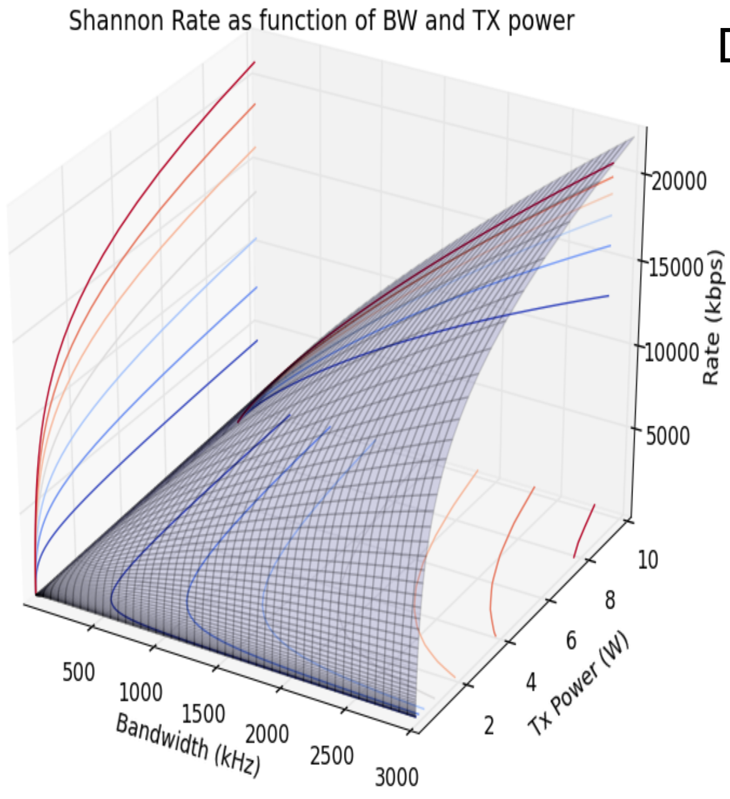
Discrete spectral efficiency set



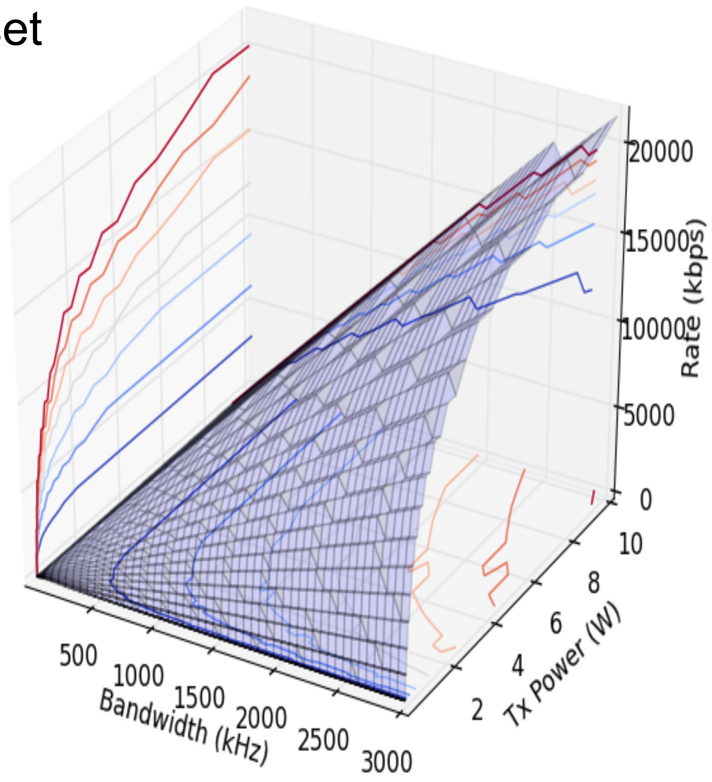
We use a discrete rate set to better represent reality, e.g. an LTE system.

Source:
"Essentials of LTE and LTE-A" A. Ghosh and R. Ratasuk.
P. 98, table 4.7
"The CQI table and reference SINR requirements"

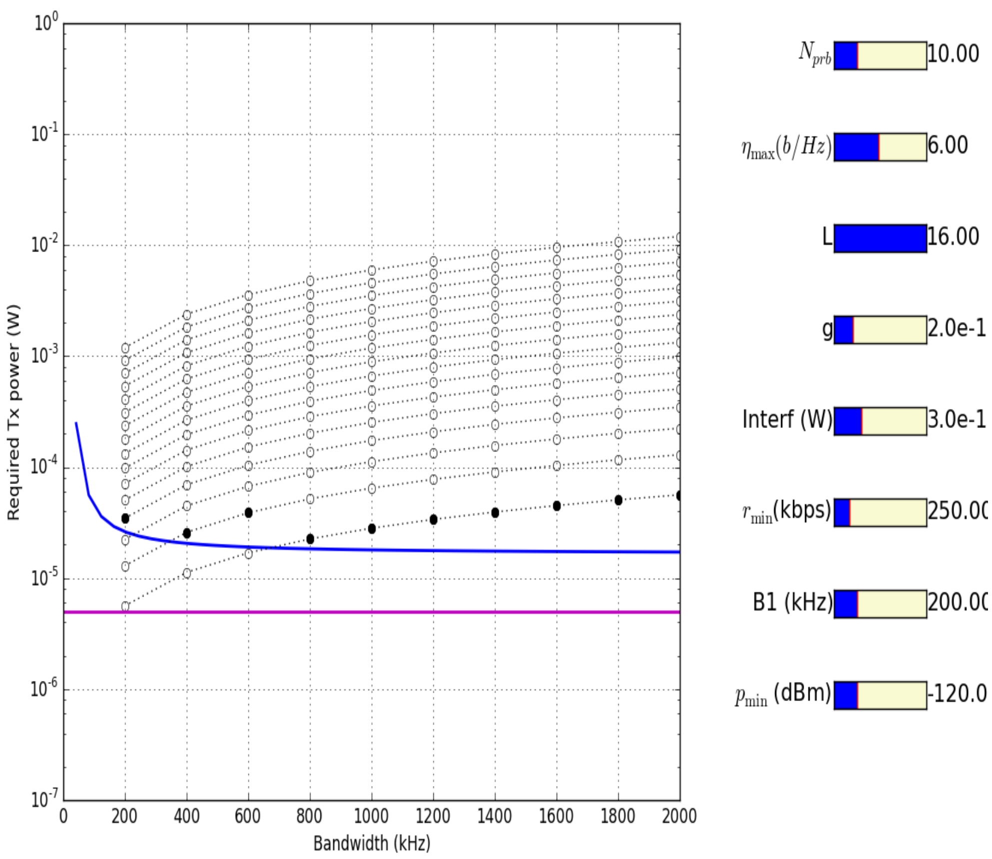
Bandwidth-TX power tradeoff for given rate requirement



Discrete rate set



More BW ≠ Less Tx Power

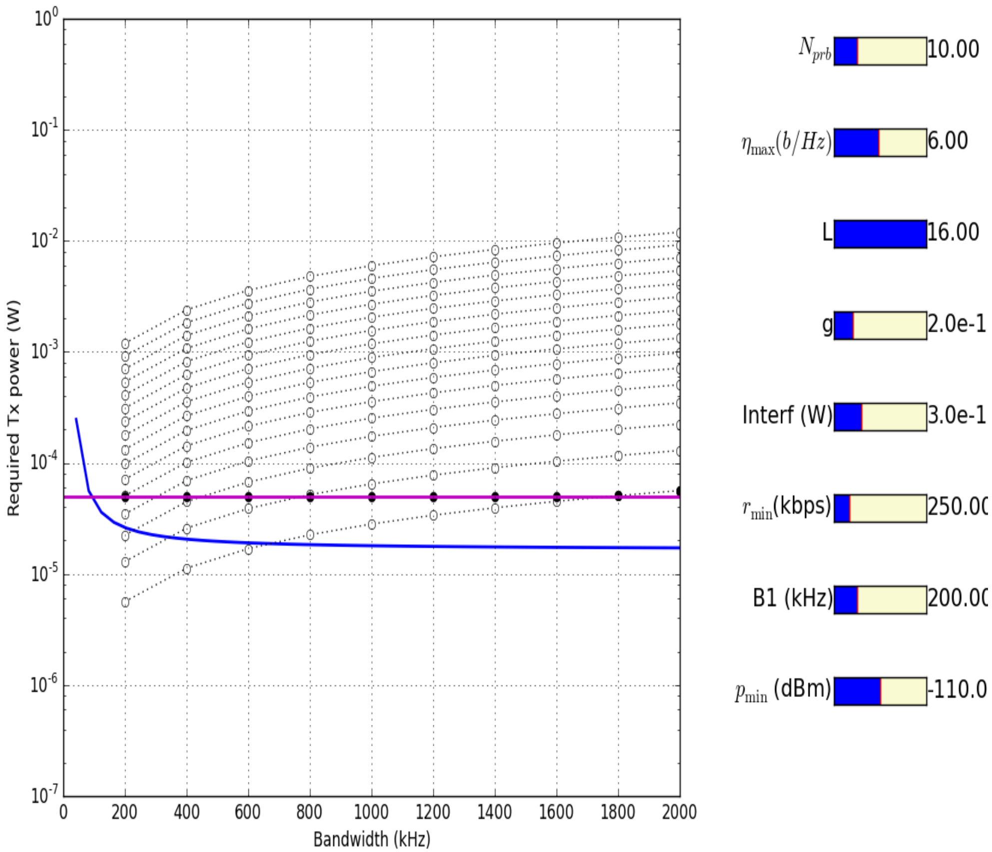


For the propagation and interference experienced by this particular UE, we can reduce Tx power by allocating 2 PRBs (400kHz) instead of 1 PRB (200kHz).

But BW-power relationship is not monotonic.

We need more Tx power to satisfy same rate requirement using 3 PRBs.

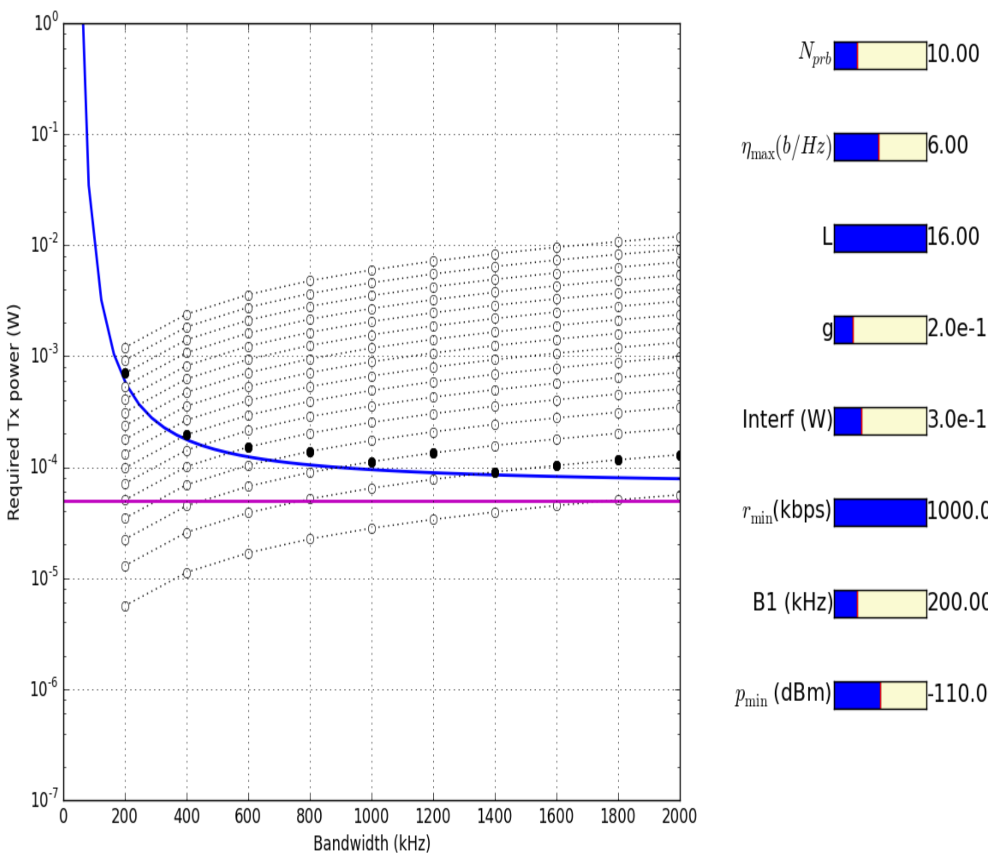
Impact of UE sensitivity



For the same UE, if p_{min} is increased from -120dBm to -110dBm we don't gain anything by allocating more BW.

Note: Even though excess BW may not always reduce Tx power, it can still be used to serve more UEs and switch off BSs.

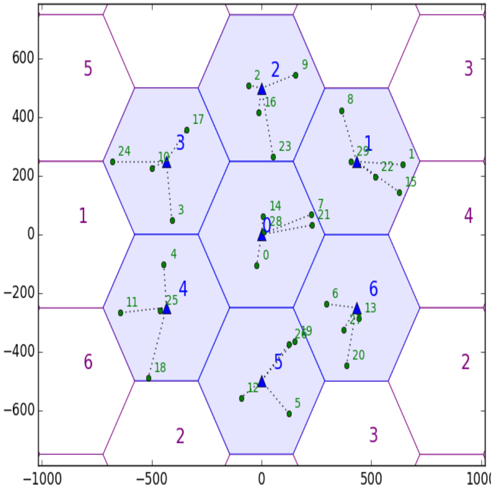
Impact of rate requirement



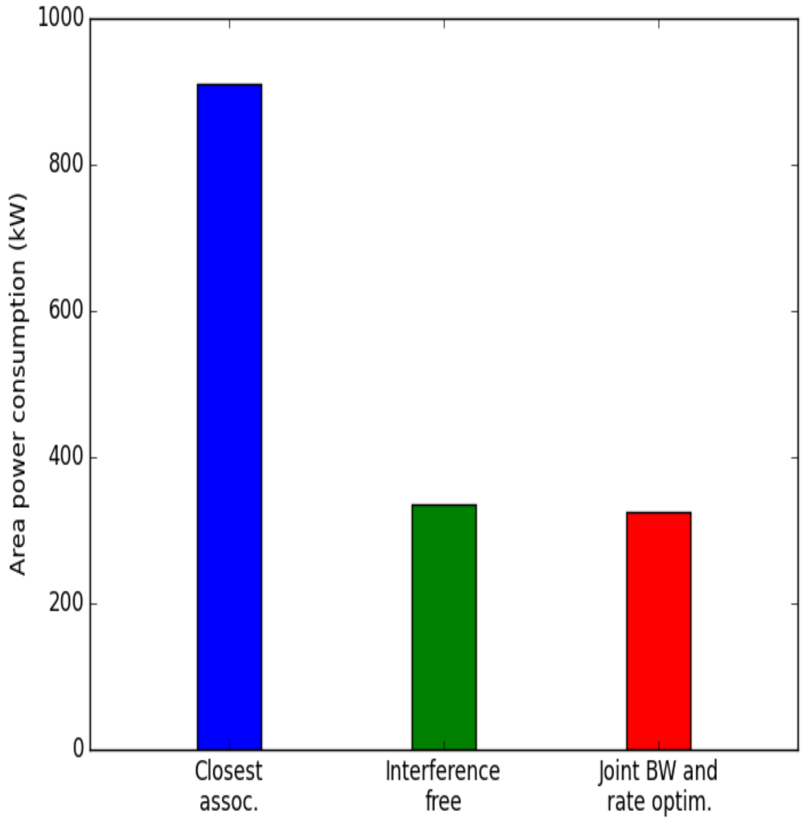
If we increase rate requirement,
e.g. from 250kbps to 1000kbps,
BW-Tx power tradeoff becomes more apparent.

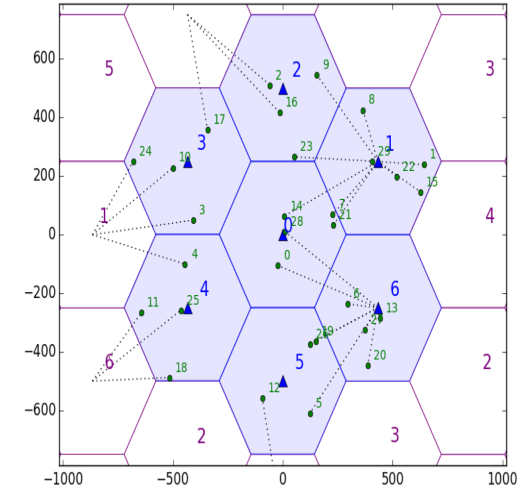
Whether there is a clear tradeoff depends on the relationship between several parameters listed here.

Performance of BW and power allocation algo. with minimum rate constraint



Example System: 7BS, 30UEs, 15 PRBs, Pon= 130W, Poff=13 W, a= 4.7





Baseline: Strongest signal association, all BS are on.

In scenarios we studied, joint BW-power allocation algorithm typically uses the excess bw to switch off BSs rather than reduce TX power. In the figure, only 2 BS are on.

KTH ROYAL INSTITUTE
OF TECHNOLOGY

3/4 Double Auction Based Energy Market for Network Sharing

Aftab Hossain, Cicek Cavdar, Riku Jantti, ICC 2015, London

Motivation

- Network Capacity demand is growing exponentially.
- Around 90% of total energy is wasted by the BSs to ensure coverage.
- Small cell or offloading to wi-fi boosts capacity only.
- Daily network load maximums are 2-10 times higher than the daily minimums.
- Load demands among the operators at a given time serving the same area varies significantly.
- Multi-operator capacity sharing has the potential to reduce energy consumption significantly.

Network Sharing Energy market for MNOs

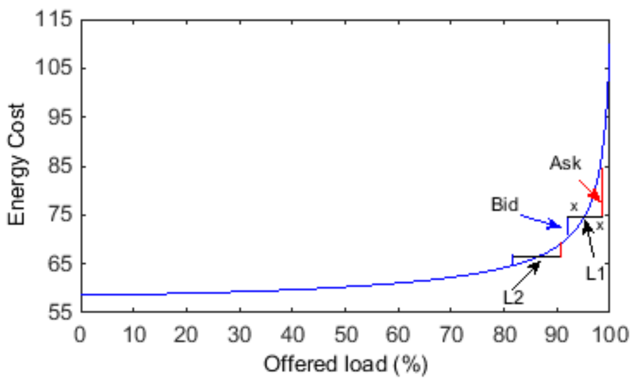


Fig: Bid and ask generation

Operators buy and sell capacity in order to maximize their profit and the clearinghouse minimize overall energy consumption.

Each operator submit both the offer to buy and cell capacity, i.e., ask and bid.

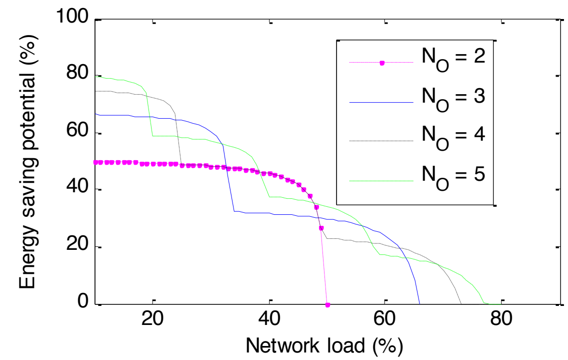


Fig: Energy saving from total offloading by DA

TABLE III
ENERGY SAVING AT HIGH LOAD 2

Operators load					Saving(%)
99	99	99	99	70	17
99	99	99	90	80	13
99	99	90	90	80	10
99	99	90	80	70	12
95	80	80	70	70	2
90	80	80	70	70	1
90	80	70	60	60	12
80	80	80	70	70	9

Fig: Energy saving from partial offloading by DA

- Bid to offload each unit of load is the amount of energy that can be saved by offloading that unit
- Ask for 1 unit of load is the energy cost for accepting 1 unit of extra load.
- Clearinghouse makes the allocation based on the criterion to minimize total energy consumption and also determines the trading price by PMD protocol

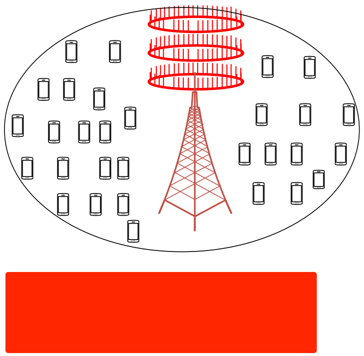
Low load: 50-80 % Energy Saving by total offloading
High Load: 17 – 2 % Energy Saving by partial offloading

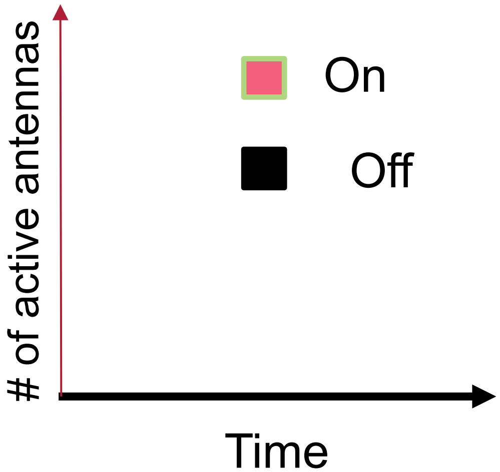
4/4 Energy efficient load-adaptive massive MIMO

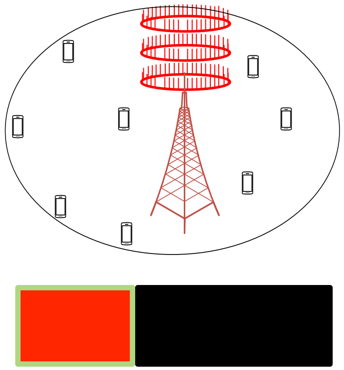
M. M. Aftab Hossain, Cicek Cavdar, Emil Bjornson, Riku Jantti

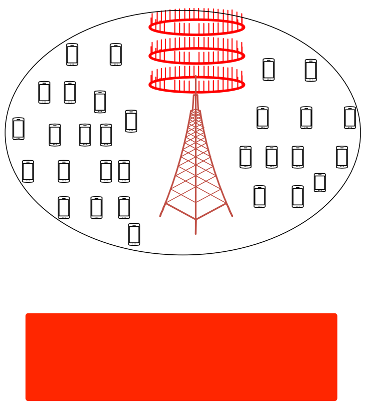
Globecom 2015

Dynamic adaptation of antennas

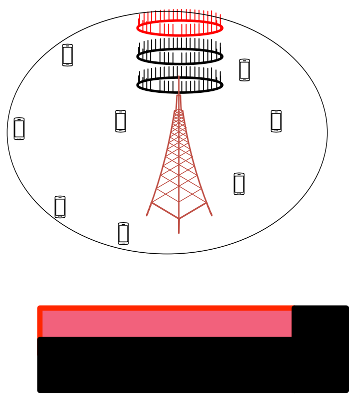








Main contribution is how to adapt the # of antennas to the load dynamically



Massive MIMO and EE

- ”Each BS uses hundreds of antennas to simultaneously serve tens of user equipments (UEs) on the same time-frequency resource.”
- Increasing the number of antenna elements increases capacity. How does energy consumption scales with the number of antennas?
- i.e. is it possible to adjust the number of antenna in order to improve energy efficiency at different network load?

Energy efficiency optimization

- EE
  - the number of bits transferred per Joule of energy
  - the ratio of average sum rate (in bit/second) and the average total power consumption (in Joule/s)
  - $\text{Energy Efficiency (EE)} = \dfrac{\text{Average sum rate}}{\text{Power Consumption}}$

$$= \frac{K_c R_c\,(K_c, M_c, \{M_d\}_{d \neq c})}{P_c^{tot}(K_c, M_c)}$$

- The EE maximization problem for cell c for a particular load

$$\underset{M_c}{\text{maximize:}}\ \frac{K_c R_c\,(K_c, M_c,\ \{M_d\}_{d \neq c})}{P_c^{tot}(K_c,\ M_c)}$$

$$\text{Subject to } M_c \geq K_c + 1$$

Problem formulation

- In order to capture the daily load variation, we model each BS as an $M/G/m/m$ state-dependent queue

- Denoting the steady state probability of the BS c serving n number of users, , i.e., $\Pr[K_c = n]$ during time interval h, by $\pi_c(h,n)$

- The main problem formulation for BS c can be rewritten as

$$\underset{\mathbb{M}_c}{\text{maximize}} \sum_{h=1}^{H} \sum_{n=1}^{m} \pi_c(h,n) \frac{nR_c(n, M_c, \{M_d\}_{d\neq c})}{P_c^{tot}(n, Mc)}$$

$$\text{Subject to } M_c^{(h)}(n) \geq n+1,$$

Where $R_c(K_c, M_c, \{M_d\}_{d\neq c})$ is the average rate per user when there are n users in the cell and $\mathbb{M}_c = [\mathcal{M}_c^{(1)}\ \mathcal{M}_c^{(1)} \ldots \mathcal{M}_c^{(H)}]$ where $\mathcal{M}_c^{(h)}$ is the vector that gives the optimum number of antennas in cell c during the time interval h.

System model

Assumptions:
- BS obtains perfect CSI for its users and
- applies zero forcing precoding i.e. intracell interference is calncelled out.
- power allocation is adapted so that each user gets same rate
- total average transmit power of the BS is fixed.

Array gain

User Rate (average):

$$R_c = B\left(1 - \frac{K_{max}}{T_c}\right)\log_2\left(1 + \frac{\frac{pM_c}{K_c}\left(\mathbf{M_c} - \mathbf{K_c}\right)}{\Lambda_{cc}\,\sigma^2 + \sum_{d \neq c} \Lambda_{cd}\,pM_d)}\right)$$

B = bandwdith,
T_c = coherence time and
K_{max}= maximum number of users in any cell
Λ_{cc} = the channel variance from the serving BS,
$\sum_{d\neq c}\Lambda_{cd}pM_d$= the avearge inter-cell interference power normalized by Λ_{cc},
K_c= number of simultaneously served users
M_c=Number of active antennas

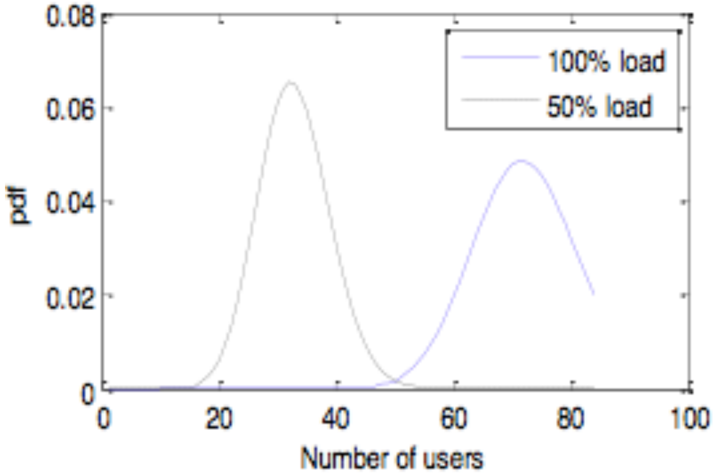
Traffic Model

- We model the massive MIMO system as an $M/G/m/m$ state-dependent queue where maximum m numbers of users are served at a time.
- The steady state probabiliy distribution

$$\pi_c(n) = \left[\frac{\left[\frac{\lambda s}{R_C(1)}\right]^n}{n! f(n)(f(n-1).....f(1)}\right]\pi_c(0), n = 1,2, \ldots m$$

where $\pi_c^{-1}(0) = 1 + \sum_{i=1}^{m}\left\{\frac{\left[\frac{\lambda \sigma}{R_C(1)}\right]^i}{i! f(i) f(i-1).....f(1)}\right\}$

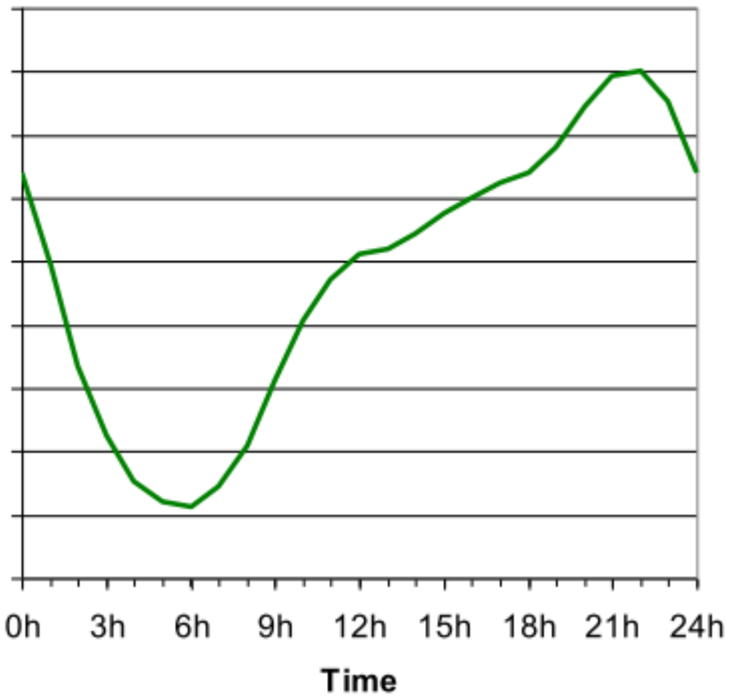
$R_c(1)$= the rate when there is only one user in the system and $f(n) = \frac{R_C(n)}{R_C(1)}$, $R_c(n)$ is the average rate if there are n number of users in the system,



Traffic model-2

- we choose $m = K_{max}$, the number of users being served simultaneously gives global optimum EE and load carried by these number of users is mapped to the highest traffic demand of the DLP.

- For other network loads, we find the corresponding average number of users, e.g., for x% load $\lambda_x = \frac{x}{100} * \lambda_{max}$
- At 100% load we allow at most 2% blocking i.e. $\pi(K_{max}) = 0.02$.



Power consumption model

$$P_{\text{total}} = M_c P_{PA}(p) + P_{BB}(M_c, K_c) + P_{Oth}$$

The baseband processing power consumption is a nonlinear function of K_c but a linear function of M_c and can be summarized as

$$P_{BB}(M_c, K_c) = C_o^{BB} + M_c C_1^{BB}$$

PA consumption: $C_1^{PA} = P_{PA}(p) = \frac{1}{\eta}\sqrt{pP_{max,PA}}$

$$\boldsymbol{P_{total} \approx C_0 + C_1 M_c}$$

where $C_0 = C_0^{BB} + P_{Oth}$, $C_1 = C_1^{PA} + C_1^{BB}$, $P_{max,PA}$ = maximum transmit power of the PA, η= maximum efficiency at $p_{max,PA}$, $a \approx 0.0082$, a PA dependent parameter

EE maximization game

- The objective function when the BS servers n users can be broadly written as

$$E_c = \frac{n\,\beta\,\log(1 - nM_c\gamma_{c,1} + \gamma_{c,1}M_c^2)}{C_0 + C_1 M_c}$$

where $\gamma_{c,1} = \frac{\frac{1}{n}p}{G_{cc}\sigma^2 + \sum_{d\neq c} pG_{cd}M_d}$, is the SINR when using a single antenna.

- We define the EE maximization game, $\boldsymbol{\mathcal{G}}(\boldsymbol{\mathcal{K}}, \boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{E}})$ where
  - the players are the BSs,
  - $S = S1 \times S2 \times \ldots \times S_C$ is the strategy space, i.e., space of number of active antennas,
  - $\boldsymbol{\mathcal{E}} = E_c(\boldsymbol{\mathcal{M}}_c, \mathbf{M}_{-c})$ the utility of the players

  S_c is a function of the number of antennas used by the interfering BSs, $\mathbf{M}_{-c}$

  $S_c(\boldsymbol{\mathcal{M}}_c(\mathrm{n}): \mathrm{n} + 1 \leq \boldsymbol{\mathcal{M}}_\mathrm{c} \leq M_{max}, \forall n \in \boldsymbol{\mathcal{U}}_c$ where $\boldsymbol{\mathcal{U}}_c = \{1, 2, \ldots, m\}$

The best response iteration and convergence

- In game theory, the best response is the strategy (or strategies) which produces the most favorable outcome for a player, taking other players' strategies as given.

- The use of best response strategy gives rise to dynamic system of the form

$$\boldsymbol{\mathcal{M}_c} = \operatorname{argmin}_{\boldsymbol{\mathcal{M}_c} \in S_{c(M_{-c})}} E_c(\boldsymbol{\mathcal{M}_c}, \mathbf{M}_{-c})$$

- The convergence of best response iteration to a unique Nash equilibrium has been proved by showing that this EE maximization problem can be modeled as a S-modular game.

Algorithm

Algorithm 1 Best response iteration

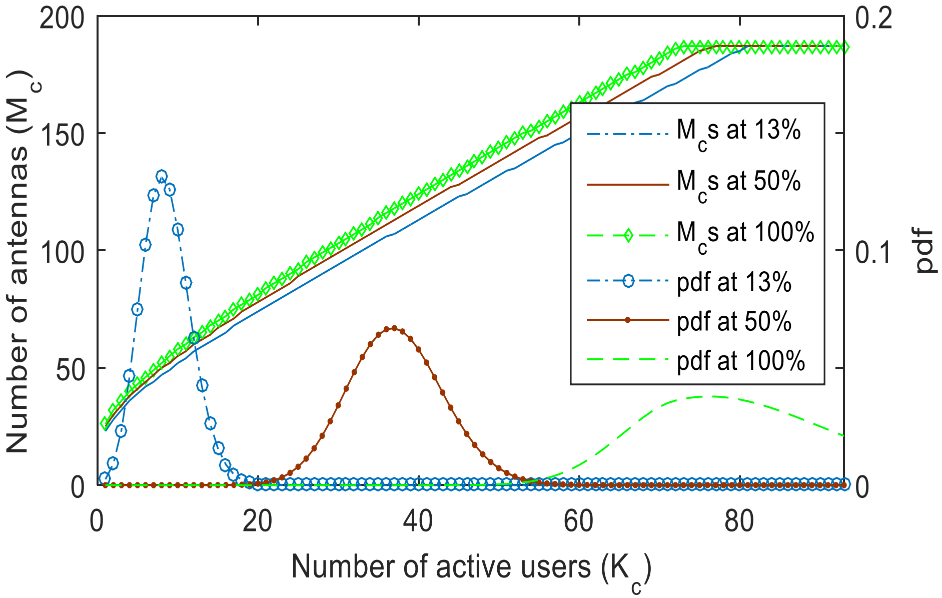
$\mathcal{M}_c \leftarrow M_{\max} \cdot \mathbf{1}, \forall c \in \mathcal{C}.$
maxtol $\leftarrow 1$
while maxtol $\neq 0$ **do**
  for all $c \in \mathcal{C}$ **do**
    $i \leftarrow c$
    $\mathbf{M}_{-c} \leftarrow \sum_n \mathcal{M}_d(n)\pi_n, \forall n \in \mathcal{U}_d, \forall d : d \neq c$
    Define strategy space $\mathcal{S}_c$ based on (16)
    $\mathcal{M}'_c \leftarrow \arg\max_{\mathcal{M}_c \in \mathcal{U}_c(\mathbf{M}_{-c})} E_c(\mathbf{M}_{-c})$
    $\text{tol}_c \leftarrow |\mathcal{M}'_c - \mathcal{M}_c|$
    $\mathcal{M}_c \leftarrow \mathcal{M}'_c$
  end for
  maxtol $\leftarrow \max_c(\text{tol}_c)$
end while

Simulation Parameters

Parameter	Value
Cell radius: d_{max}	500 m
Minimum distance, d_{min}	35 m
Transmission Bandwidth, B	20 MHz
PA maximum efficiency,	80%
BS Fixed power : P_{0th}	18 W
Channel coherence intervals: T_C	12600 symbols
Local oscillator Power: P_{SYN}	2 W
Power required to run the circuit comp. at a BS: P_{BS}	1 W
Total noise power: $B.\sigma^2$	-96 dBm
Power required for coding of data signals:	0:1 W/(Gbit/s)
Power required for decoding of data signals: P_{DEC}	0:8 W/(Gbit/s)
Computational Computation efficiency at L_{BS}	12:8 Gflops/W

Interplay between M_c and K_c

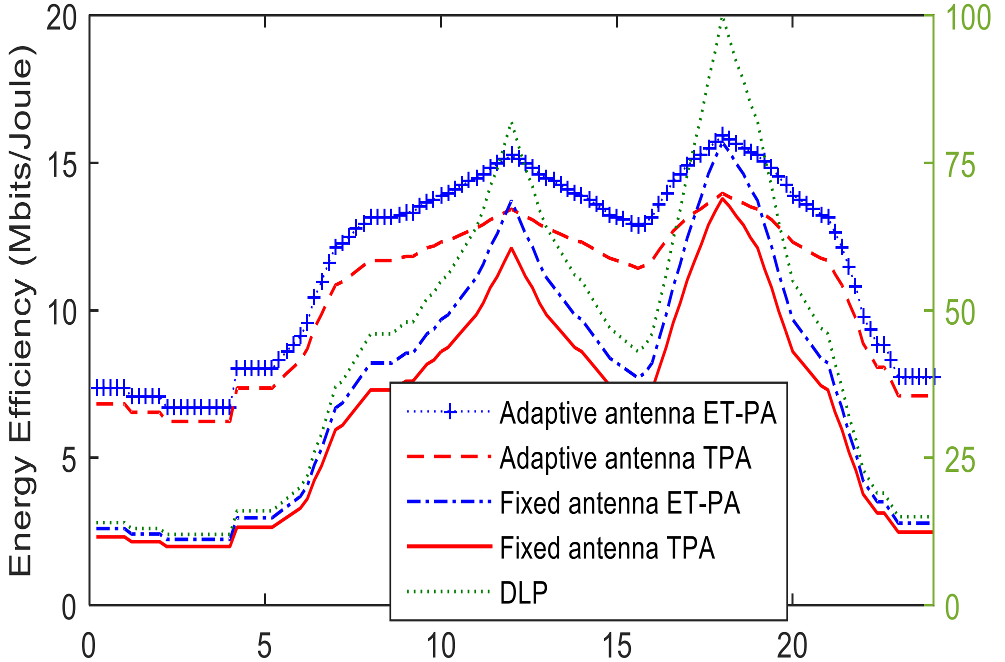
- The relation between M_c and K_c is quite linear for different loads.
- Ratio between M_c and K_c is quite high when BS serves only few users which is around 2 at higher user states.
- The average number of antennas used at different loads vary mainly due to the probability distribution of the users.



- When serving few users, an additional antenna does not consume much energy compared to fixed consumption but contributes singnificantly to increase EE due to higher array gain $(M_c - K_c)$

Energy efficiency improvement

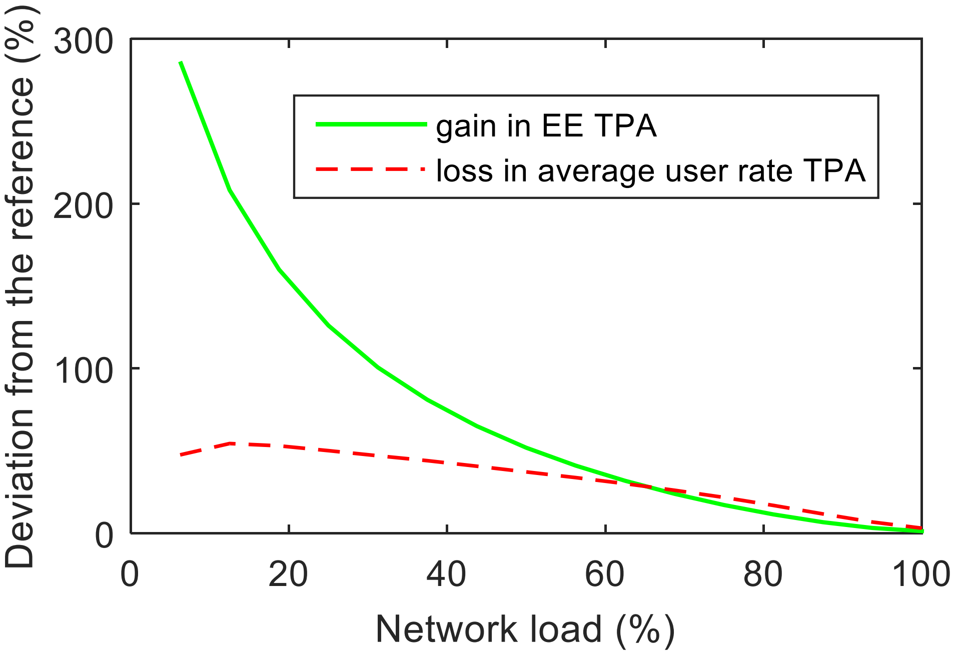
- EE increases with the increase in load for both the reference case and our scheme for both TPA and ET-PA.
- Our scheme attains significantly higher EE compared to the reference case at low load.
- EE gain keeps decreasing with the increase in load.



- At the peak load, the gain is insignificant as the probability of having small number of users which allows EE improvement by reducing antennas is very low.

EE and user rate tradeoff

- At very low load the EE has been increased with around 300% at the cost of around 50% reduction of the average user data rate.
- With the increase of load in the system, both the gain in EE and loss of user rate get reduced.



- Over the 24 hour operation, EE has been found to be improved around 24% at the cost of around 13% reduction of user rate.

Thank you!
cavdar@kth.se