Beyond 4G Cellular Networks: Is Density All We Need?

Jeffrey G. Andrews

Wireless Networking and Communications Group (WNCG)
Dept. of Electrical and Computer Engineering
The University of Texas at Austin

IEEE International Workshop on Emerging Technologies for LTE-Advanced and Beyond @ IEEE Globecom 2013
Atlanta, GA, USA
Dec. 13, 2013

Plausible 5G Requirements
(these are my own opinions)

- Peak Rate: 10 Gbps
  - Peak rate is a marketing number, not an engineering number
  - I expect someone to claim 100 Gbps. This can be ignored.
- **5% Rate: 100 Mbps**
  - This is a real engineering number and very challenging
  - This is what a “typical” 5% user needs to actually achieve
- Latency: 1 millisecond roundtrip
- Cost per bit: 10-100x below 4G
- Power consumption: similar to LTE (thus, requires Joules/bit to drop 10-100x)
- Implicit but crucial: Backhaul that supports all the above

All of these require 10-100x improvement vs. 4G (e.g. LTE Release 10)

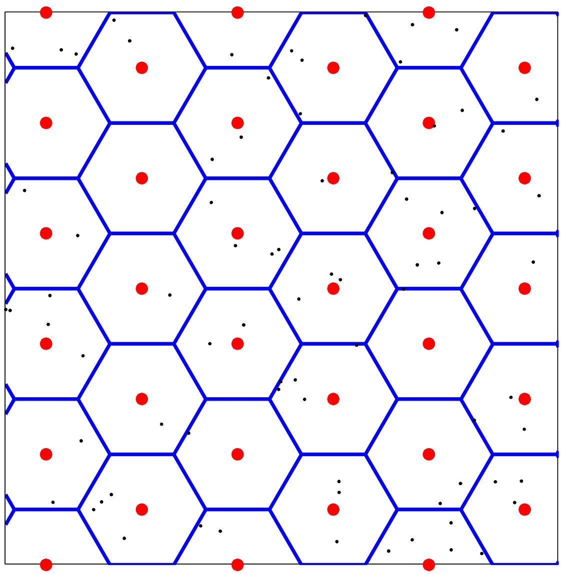
Implications

- To get 100 Mbps for 5% users, we've got to do something pre-log to Shannon's equation.
  - We're nearly achieving log(1+SINR) spectral efficiency in current systems
  - Increasing SINR gives rapidly diminishing gains, unless SINR is very low to begin with (more shortly)
- Unless we've really missed something, that leaves us with three choices:
  1. Extreme Densification via Cell Splitting (load reduction)
  2. Increasing bandwidth, by a lot
  3. Massively parallel communication, namely Massive MIMO (e.g. SDMA) or some other near-magical dimension increasing technique

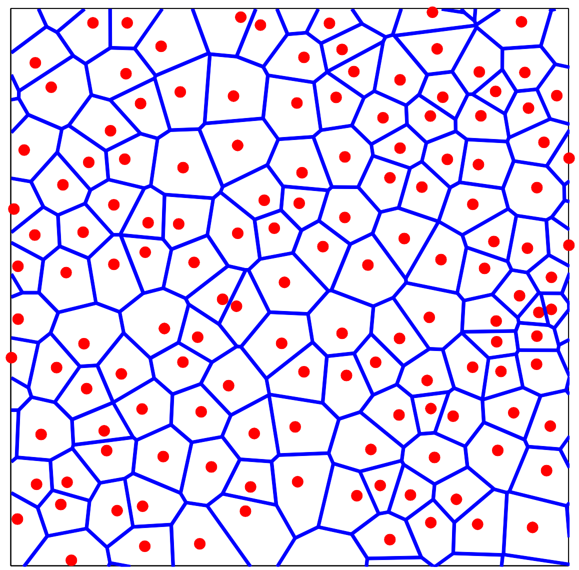
About Those Three Choices…

- More spectrum seems to mean mmWave (30-300 GHz):
  - Blocking and Near-field pathloss are major issues
  - Need large antenna arrays to overcome this, thus cannot use array for Massive MIMO (parallelization).
  - Smaller cells should be helpful here
- Extreme Densification
  - Deployment poses a cost and logistics challenge, but theoretical limit to cell splitting gain seems very (arbitrarily?) large
  - Many challenges (more on this shortly)
  - Interference-limited environment is difficult for Massive MIMO
  - Far fewer users/cell, bad for Massive MIMO (SDMA version)
- Takeaways:
  - Densification and mmWave appear to be very complementary
  - Massive MIMO competes with the other two

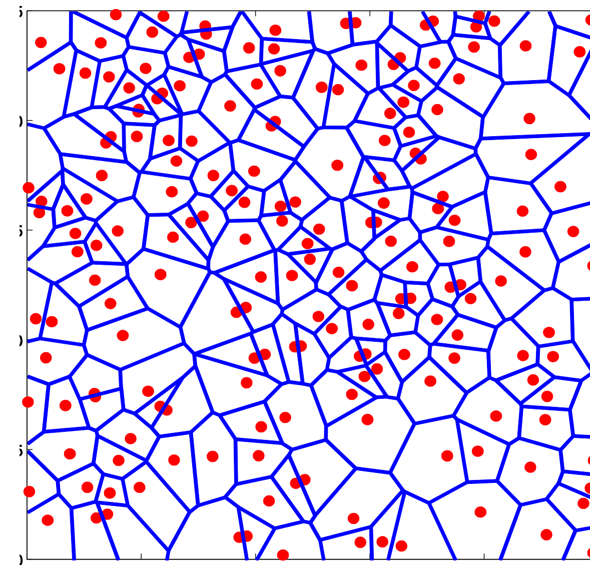
Regardless of geometry: Cellular networks are characterized by **uneven SINR**



Traditional grid model



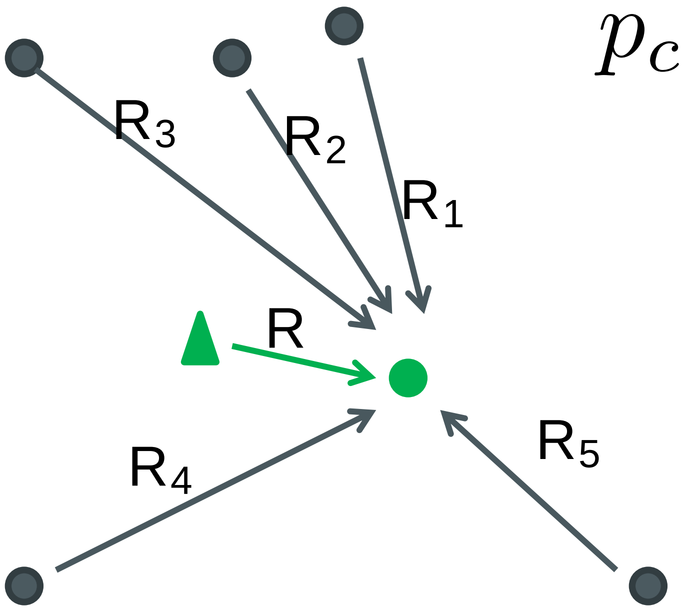
LTE Network in Dallas, TX



Random BSs (PPP)

- High SINR users near BSs have very high SINR (mostly due to high received signal power)
- Cell-edge users have very low SINR (Due both to low received signal and high interference)

Downlink Cellular SINR Analysis



$$p_c(T,\lambda,\alpha) = \mathbb{P}[\mathrm{SINR} > T]$$

$$= \frac{hr^{-\alpha}}{\sigma^2 + I_r}$$

Random channel effects (mean 1/μ accounts for transmit power)

Standard power law path loss

WLOG, aggregate interference can be quantified for MS at the origin as

$$I_r = \sum_{i \in \Phi / b_o} g_i R_i^{-\alpha}$$

Going from an infinite number of random variables to none, in 2 slides

$$p_c(T,\lambda,\alpha) = \mathbb{E}_r\left[\mathbb{P}[\text{SINR} > T \mid r]\right]$$

$f_R(r)$

$$= \int_{r>0} \mathbb{P}[h > Tr^{\alpha}(\sigma^2 + I_r) \mid r]\, 2\pi\lambda r e^{-\pi\lambda r^2} \mathrm{d}r.$$

Using the exponential (Rayleigh) fading distribution to full advantage (crucial only for tractability):

$$\begin{aligned}\mathbb{P}[h > Tr^{\alpha}(\sigma^2 + I_r) \mid r] &= \mathbb{E}_{I_r}\left[\mathbb{P}[h > Tr^{\alpha}(\sigma^2 + I_r) \mid r, I_r]\right] \\ &= \mathbb{E}_{I_r}\left[\exp(-\mu T r^{\alpha}(\sigma^2 + I_r)) \mid r\right] \\ &= e^{-\mu T r^{\alpha}\sigma^2} \mathcal{L}_{I_r}(\mu T r^{\alpha}),\end{aligned}$$

Laplace Transform

Two key steps remove all interference random variables

$$\mathcal{L}_{I_r}(s) = \mathbb{E}_{I_r}[e^{-sI_r}] = \mathbb{E}_{\Phi,g_i}[\exp(-s \sum_{i\in\Phi\setminus\{b_o\}} g_i R_i^{-\alpha})]$$

$$= \mathbb{E}_\Phi\left[\prod_{i\in\Phi\setminus\{b_o\}} \mathbb{E}_g[\exp(-sgR_i^{-\alpha})]\right]$$

$$= \mathbb{E}_\Phi\left[\prod_{i\in\Phi\setminus\{b_o\}} \frac{\mu}{\mu + sR_i^{-\alpha}}\right]$$

MGF of an exponential RV

Using the PGFL of the PPP, a key tool in stochastic geometry gives:

$$\mathcal{L}_{I_r}(s) = \exp\left(-2\pi\lambda \int_r^\infty \left(1 - \frac{\mu}{\mu + sv^{-\alpha}}\right) v\mathrm{d}v\right)$$

$$\mathbb{E}\left[\prod_{x\in\Phi} f(x)\right] = \exp\left(-\lambda \int_{\mathbb{R}^2} (1 - f(x))\mathrm{d}x\right)$$ (PGFL)

Main Result: SINR in Cellular Network

Theorem [Andrews, Bacelli and Ganti 2011]: When the fading power between any two nodes is exponentially distributed with mean μ^{-1}, the coverage probability is

$$p_c(T, \lambda, \alpha) = \pi\lambda \int_0^\infty e^{-\pi\lambda v \rho(T,\alpha) - \mu T \sigma^2 v^{\alpha/2}} \mathrm{d}v$$

where

$$\rho(T, \alpha) = T^{2/\alpha} \int_{T^{-2/\alpha}}^\infty \frac{1}{1 + u^{\alpha/2}} \mathrm{d}u$$

T = SINR threshold; λ = BS density; α = PL exponent; σ^2 = noise variance

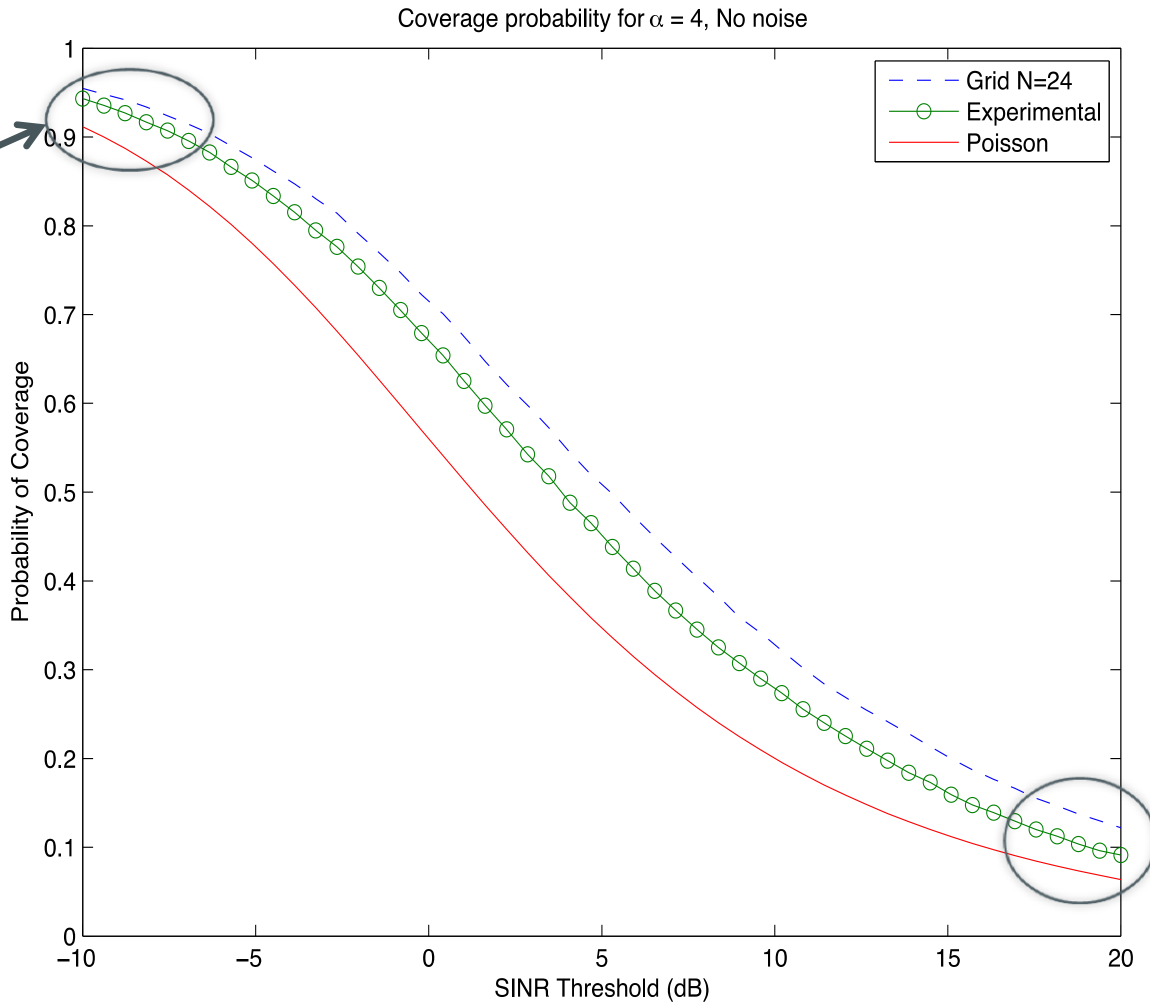
Simplest form gives good rule of thumb

- Theorem 1 (SINR CCDF), with path loss exponent α= 4, and noise << interference:

$$p_c(T,\lambda) = \frac{1}{1+\sqrt{T}\arctan\sqrt{T}}$$

  - Extremely simple expression, just obtained from the Theorem!
  - Includes fading, interference, pathloss, etc.
  - Actually matches real measurements quite well
  - No dependence on BS density, i.e. SINR is "scale invariant"
  - If noise non-negligible, then p_c will be slightly lower and p_c will also improve as BS density increases
- Allows us to observe immediately things like:
  - T = 1 gives $p_c = (1+\pi/4)^{-1}$=0.56
  - T = .1 gives p_c = 0.9
  - **T = .05 (-13 dB) gives p_c = 0.95, the "5% user"**

“Gini coefficient” for cellular SINR is large

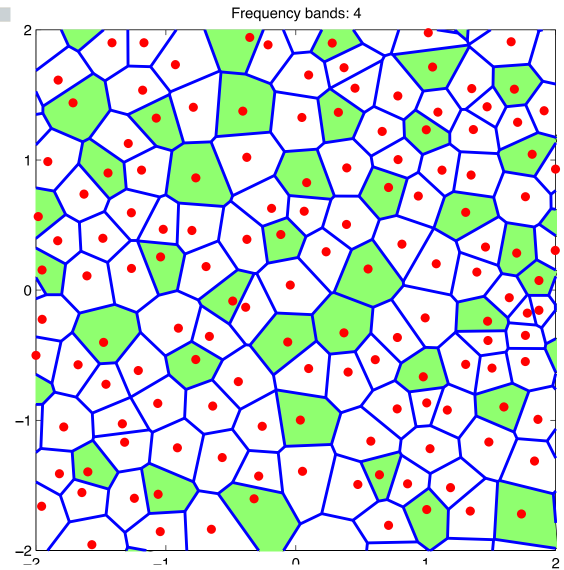


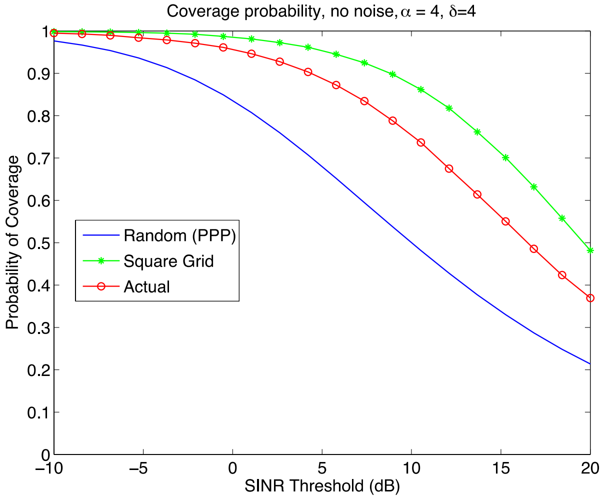
The Poor (-10 dB)

The Rich (20 dB)

Increasing Downlink SINR

- Our field is littered with failed attempts to increase downlink SINR
- Only thing that consistently "works" is frequency reuse/time slot duty cycling
  - This raises everyone's SINRs, especially cell edge
  - This lowers most user's time-averaged rates due to the prelog reuse factor
  - Doesn't improve fairness very much, mainly helps cell edge users get at least something through





Downlink SINR with Small Cell Densification

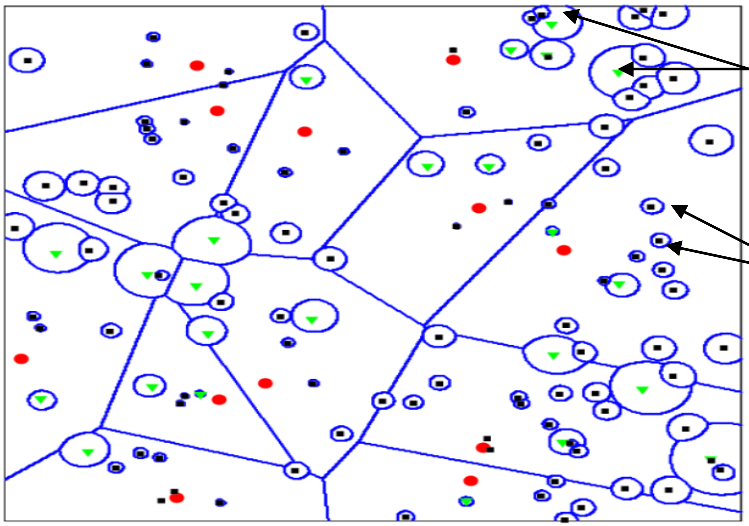
- There seems to still be considerable misunderstanding about what happens when you densify a cellular network
- Key Facts:
  1. When noise-limited towards cell edge, densification always increases SINR
  2. When interference-limited at cell-edge, densification has very little effect on SINR, which is provably true for:
     a. Perfectly regular (grid) networks [easy to show]
     b. Perfectly random (PPP) networks [just showed]
     c. BSs with massive disparities in power (i.e. "HetNets"), as long

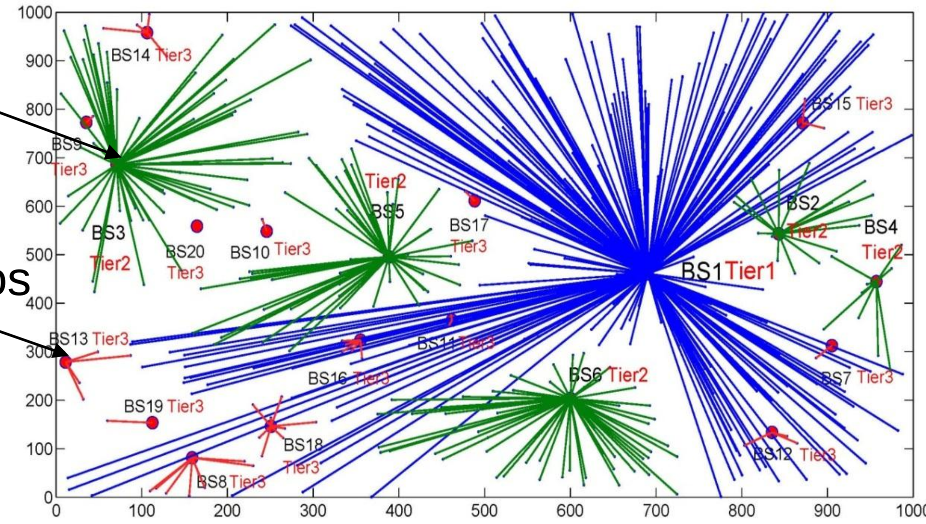
Adding small cells does **NOT** reduce SINR by causing too much interference. That is a myth (true for WiFi/CSMA).

Cell splitting gains

- Cell-splitting gains are the key benefit of densification
- Splitting a cell into two cells doubles the amount of resources, at no SINR cost
- However, the key challenge is being able to make use of those resources

Cell Splitting Gain Depends on Load at Each BS





<u>DL Max-SINR coverage regions</u>

- Small cells have very small DL footprint
- The macrocell DL is the system bottleneck

<u>With UEs now shown:</u>

- Load imbalance is clear
- User-perceived rate is about:

$$R \approx B \log(1+\text{SINR})/N$$

  - N = # of users on the BS

User-perceived rate is more sensitive to load than SINR

Summary of Key Issues for Understanding Densification

1. SINR inequality is unavoidable, will always have a large fraction of low SINR users
2. Cell splitting never hurts, but gain is hard to quantify due to massive disparities in nominal coverage areas
   - Typical loads may vary by 10-100x from macro to pico
   - Many BSs will be very lightly loaded, unhelpful
3. Rate distribution ≠ SINR distribution
4. "Spreading" load across BSs is critical, but:
   - Optimum is complex: 30 BSs and 200 UEs = $O(30^{200})$
   - User distribution is unclear (we'll use uniform, but if they cluster around BSs this helps)

Network-wide load optimization with fractional association

- Utility function: Max-sum rate gives a degenerative solution
- We'll use a max-sum-log rate objective

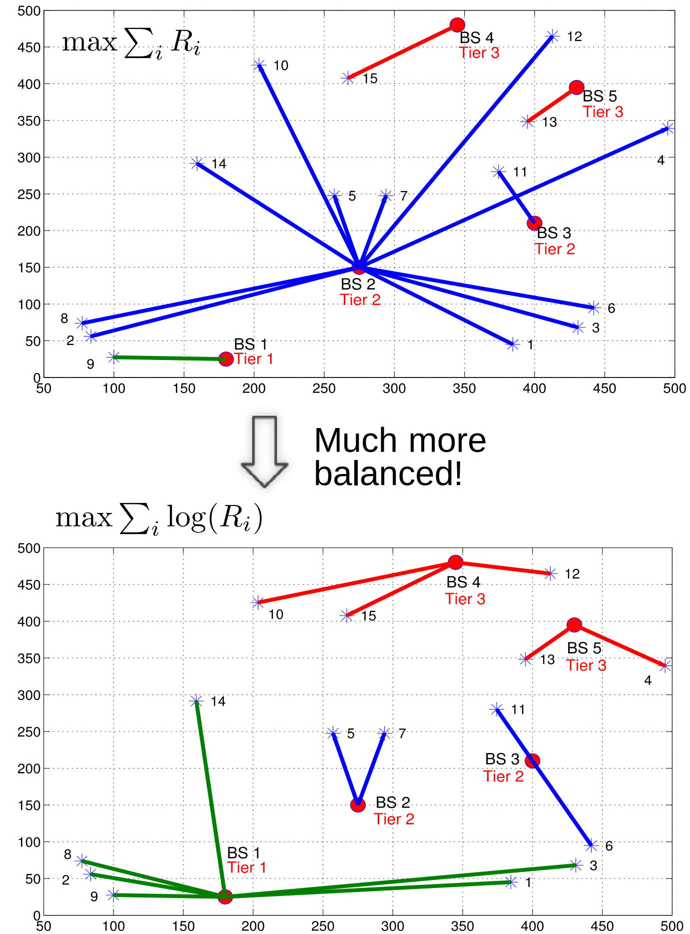
$$\max_{x} \quad \sum_{i}\sum_{j} x_{ij}\log\Big(\sum_{j\in\mathcal{B}} y_{ij}c_{ij}\Big)$$

$$\text{s.t.} \quad \sum_{j} x_{ij} \le 1, \quad \forall i,$$

$$y_{ij} = 1/\sum_{i} x_{ij},$$

$$x_{ij} \in (0,1) \quad \forall i,j.$$

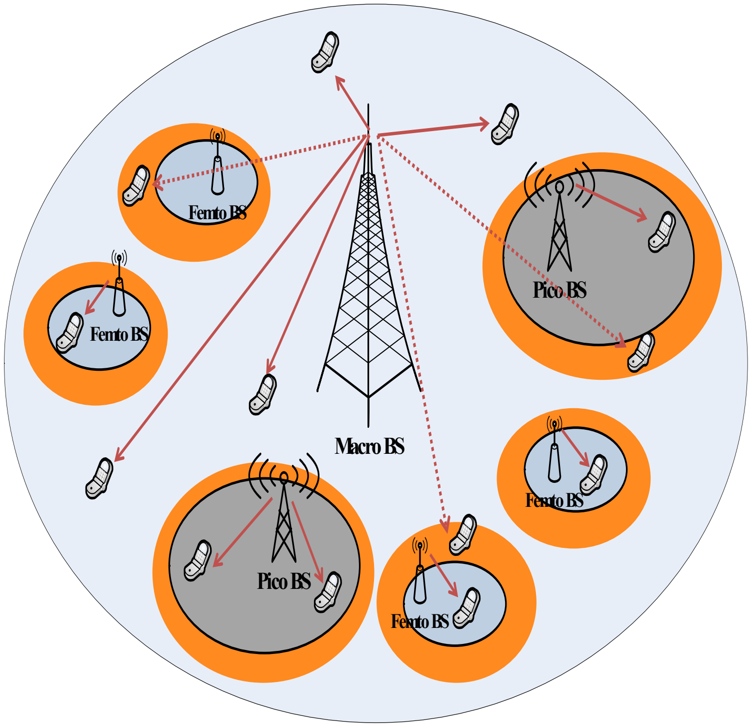
- The fractional association with the BSs allows the problem to become convex
- Can always round it back to an integer (binary) association (with little loss actually)
- Decentralized iterative algorithm based on dual decomposition converges well



Q. Ye et al, 2013.

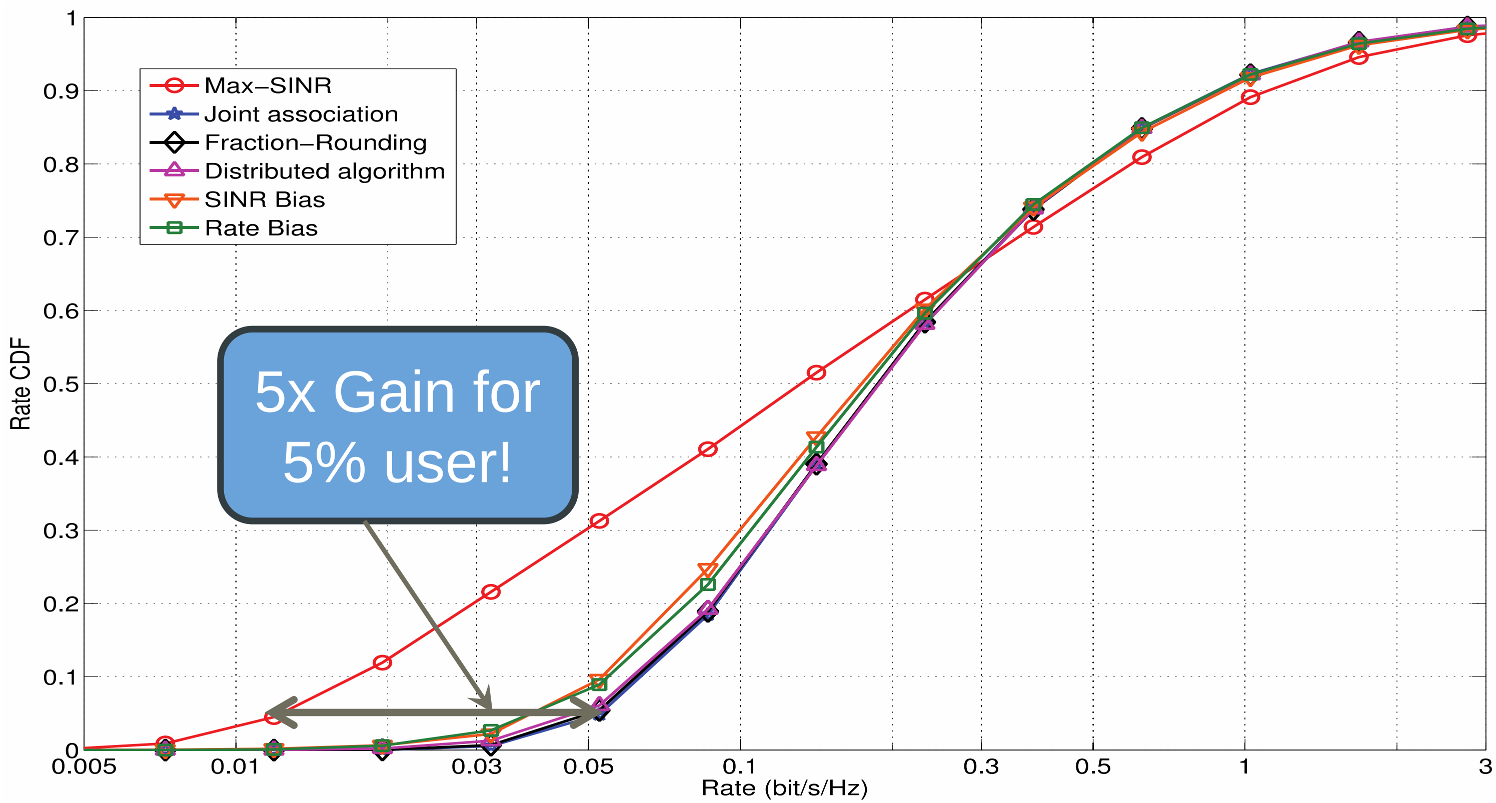
Cell Range Expansion (CRE) – same band

- CRE is a fourth approach
  - Allows for a simple fully uncoordinated decision, no iteration
  - Need only the received power (or rate) from a given BSs
- SINR Bias [used in 3GPP]
  - Assign identical bias value for all BSs in the same tier
  - Simply multiply received SINR by the bias amount, then compare to select
  - "Optimum" values obtained by brute force simulation



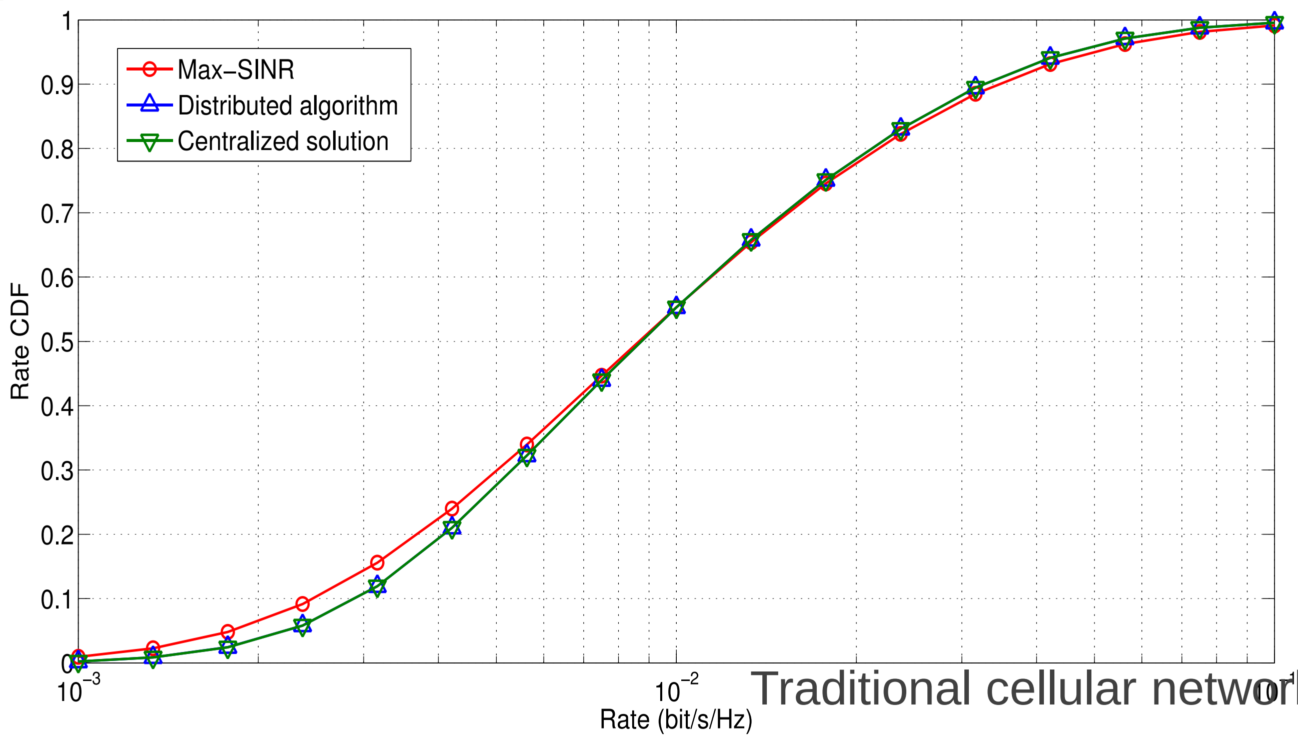
The orange part is the extended coverage area by range expansion.

Very large, and surprisingly similar, gains from all these load-balancing approaches



- 2-10x gain in throughput for bottom half of users (cell-edge)
- Static biasing gets very close to fully centralized optimization!
  - Optimal SINR bias here was [0 6 11] dB for [macro pico femto]

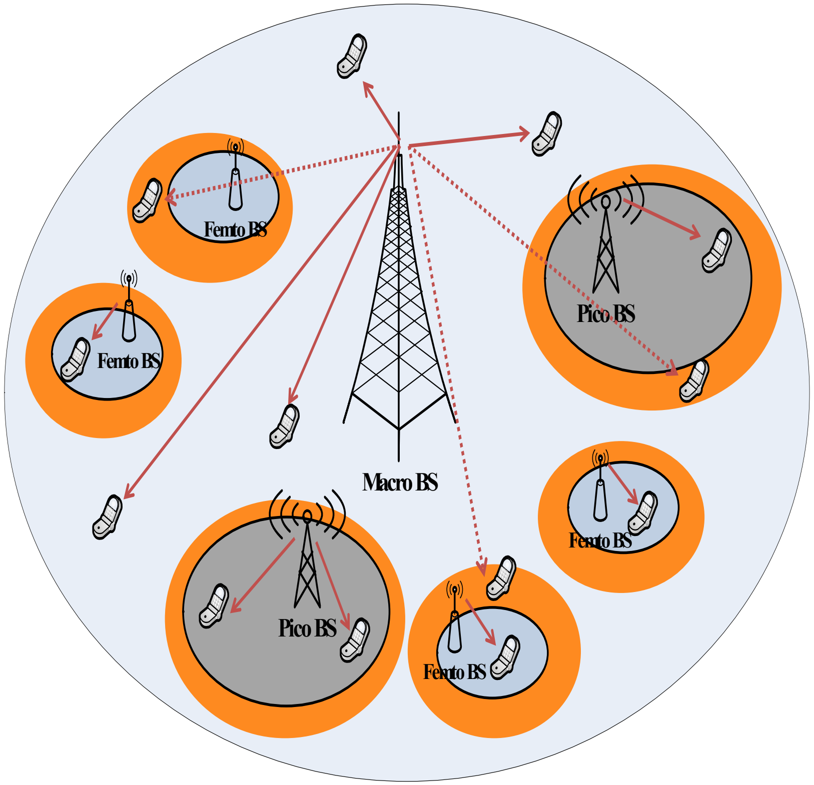
Gain is unique to HetNets



- The observed gain does not materialize in macrocell-only networks, since loads are inherently much more balanced
- <u>Caveat:</u> These results are averaged over the network. Larger gains may still occur locally, particularly if biasing is dynamic.

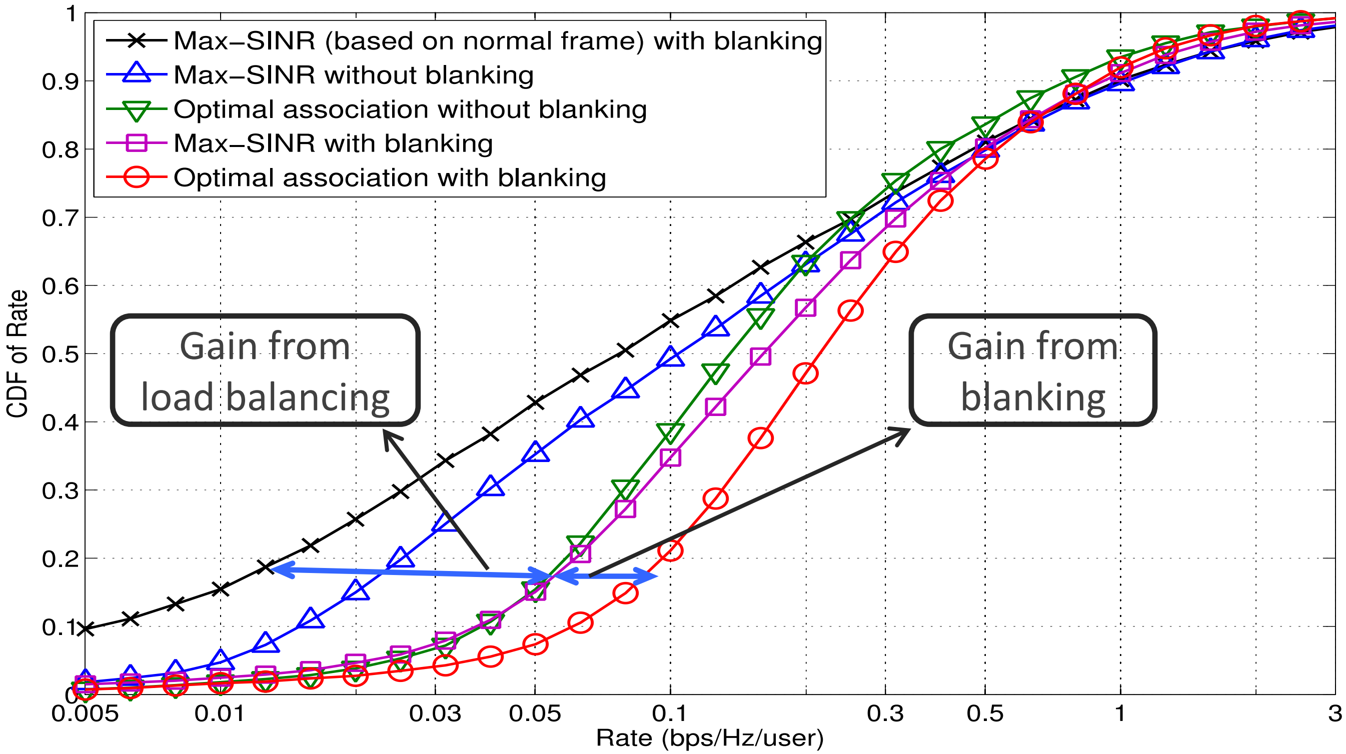
Macrocell "Blanking"

- Consider muting the macrocell for some fraction of the time η
- Called "Almost Blank Subframes" (ABS) in LTE
  - No control signals or data are transmitted (only reference signals)
  - Avoid strong inter-cell interference during range expansion
  - Possibly allow more aggressive offloading
- Is this a good idea?
- If so what should η be?
- How does it affect bias?



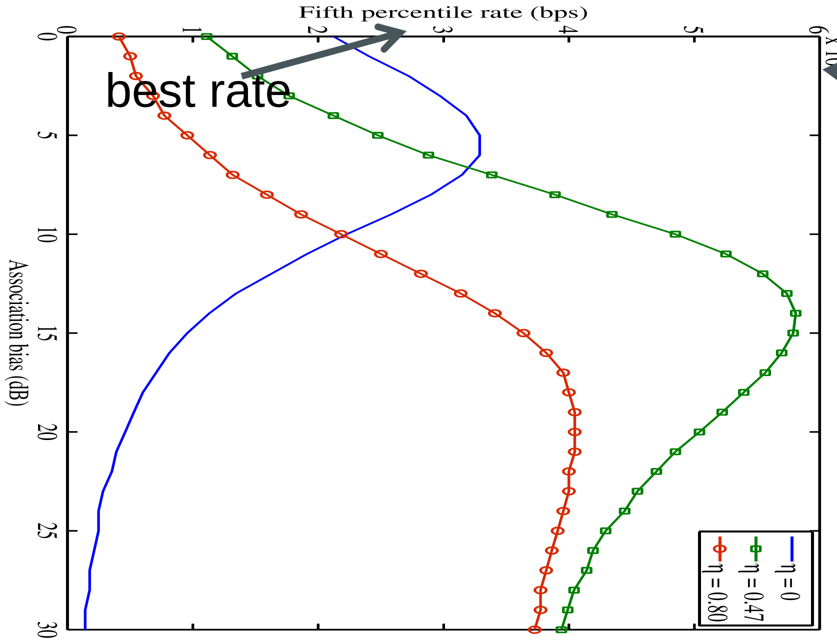
Orange part is the extended coverage area by range expansion.

How much gain is there from blanking?

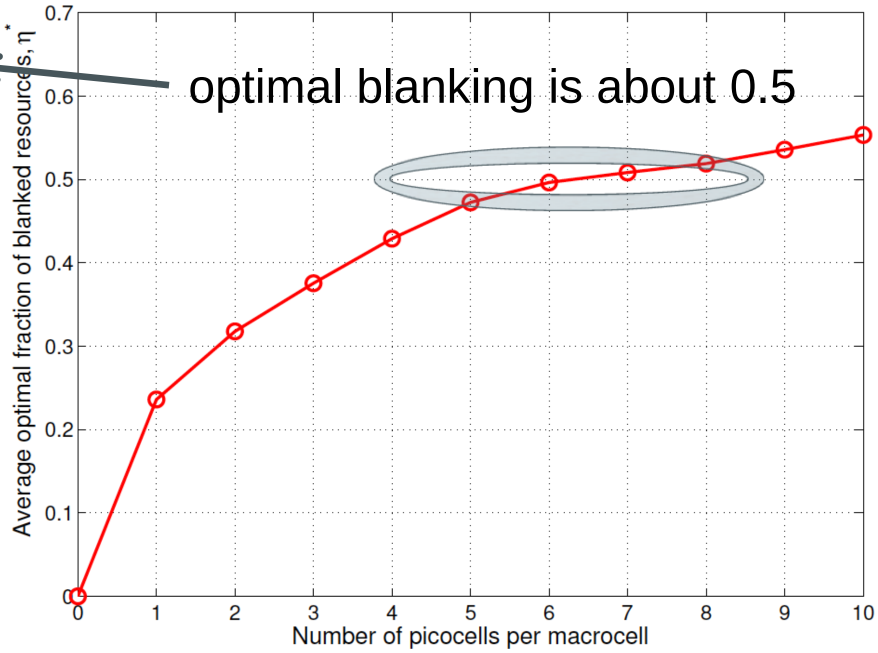


- The performance is improved with ABSs and optimal user association (association is very important)
- This plot is an average, K=3 with densities of [1 5 15], $\eta* = .31$

How many frames should be blanked?



K = 2, on average 5 picos per macro
(stochastic geometry approach)



K = 2, "optimal" association
(optimization approach)

- **Macrocell should be off half the time or more!**
- Blanking demands much more aggressive biasing, due to interference reduction

Interim Conclusions

- Densification clearly provides a great deal of rate gain, and is a big part of any “5G” solution
- Relative locations of BSs and users are largely unknown, but affects things quite a lot
  - In 3GPP models, they drop the users around the picos, and then conclude that biasing doesn’t help much…
  - Blanking and biasing can be very helpful as simple schemes to promote load balancing, reduce inequality
- It remains to be seen whether densification and offloading alone can track 100%/year traffic increases for long, <u>but it’s plausible</u>

Looking Ahead

- The 5% UE rate problem is the tough one for cellular engineers, and will exist regardless of technology
  - Blanking and load balancing (via biasing) seems to be the most promising remedy for now
  - Very little overhead or fragility
- Densification and mmWave are friends
  - mmWave will need a dense network of BSs to overcome blocking & pathloss, and provide multi-point connectivity
  - They also have some of the same enemies (especially the backhaul bottleneck)
- Many related talks in today's workshop: I look forward to learning some new things in this direction!