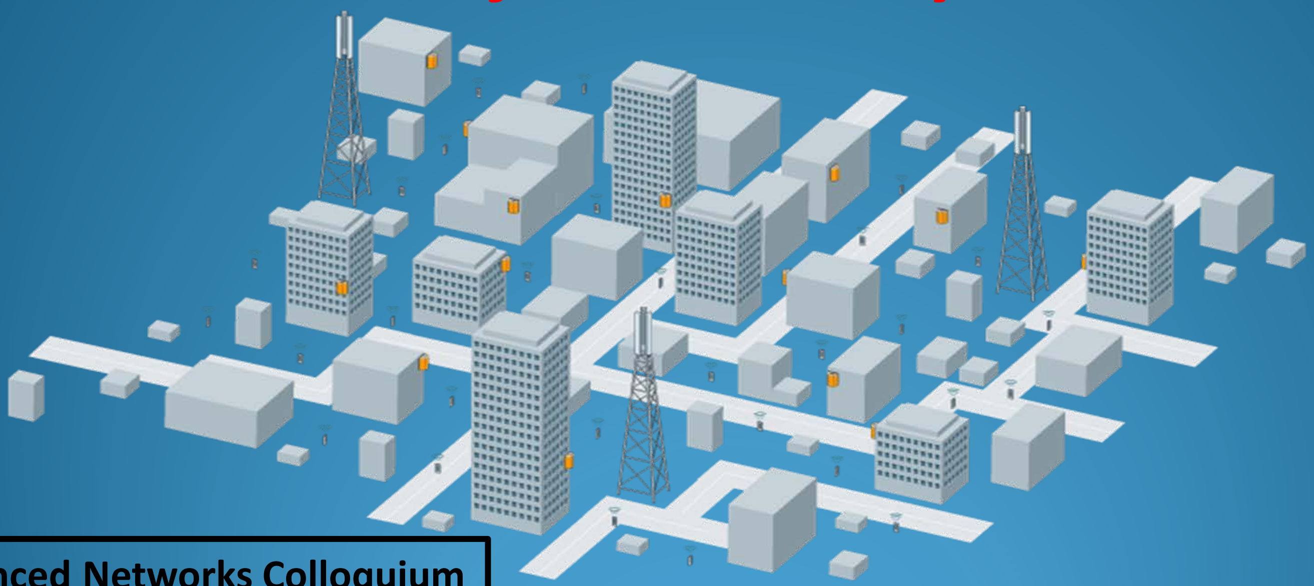
The Road Ahead for Wireless Technology: Dreams and Challenges

Andrea Goldsmith

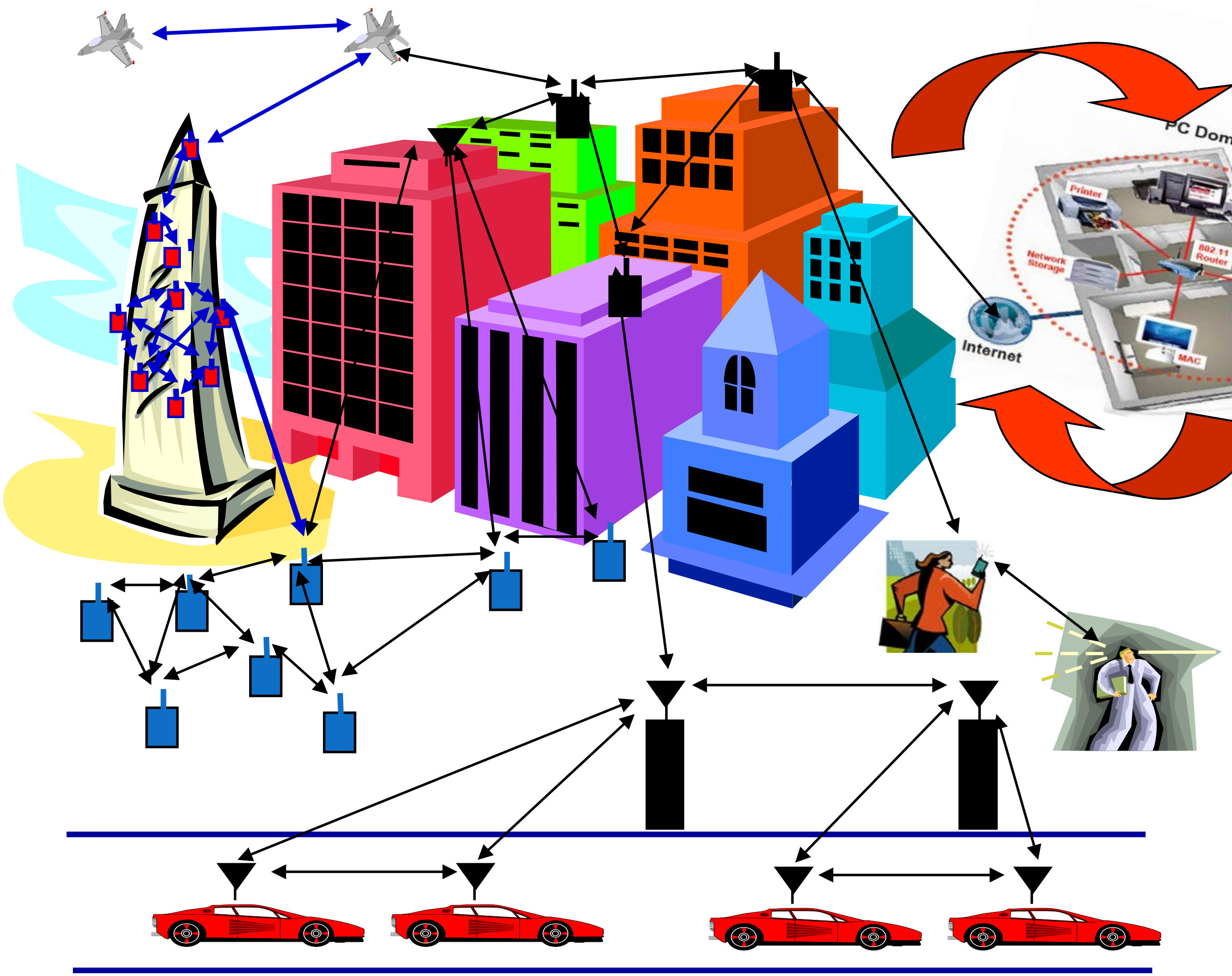
Stanford University



The Advanced Networks Colloquium
Institute for Systems Research
University of Maryland
March 11, 2016

Future Wireless Networks

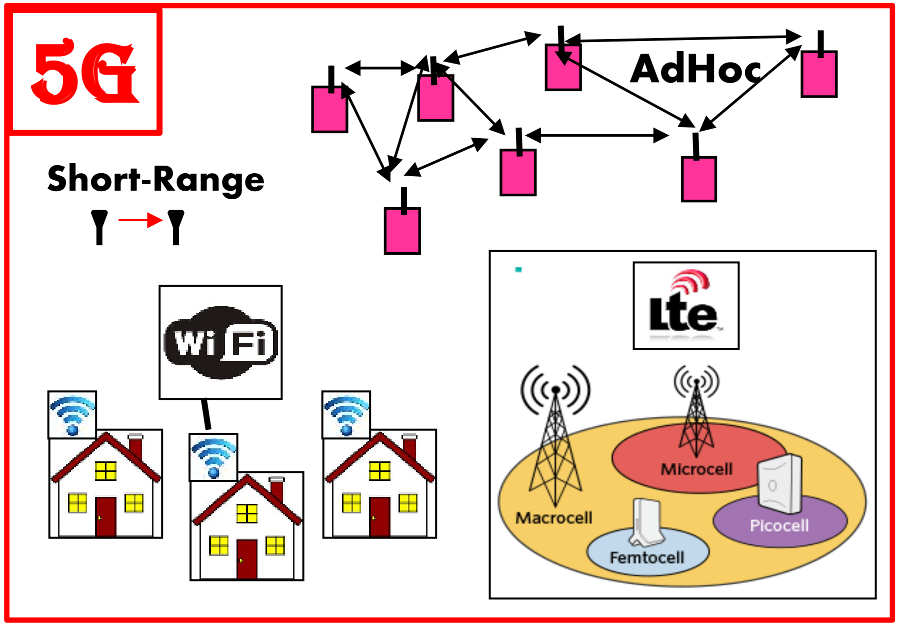
Ubiquitous Communication Among People and Devices



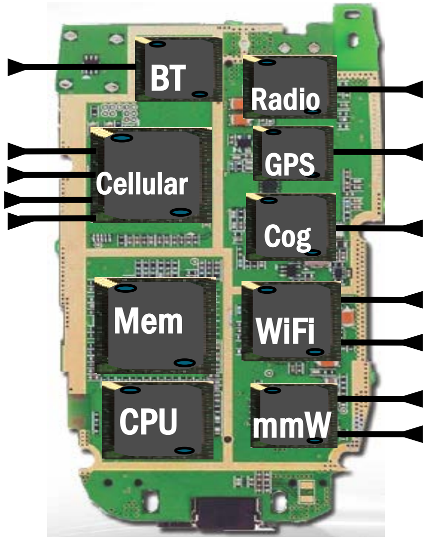
Next-generation Cellular
Wireless Internet Access
Sensor Networks
Smart Homes/Spaces
Automated Highways
Smart Grid
Body-Area Networks
Internet of Things
All this and more …

Challenges

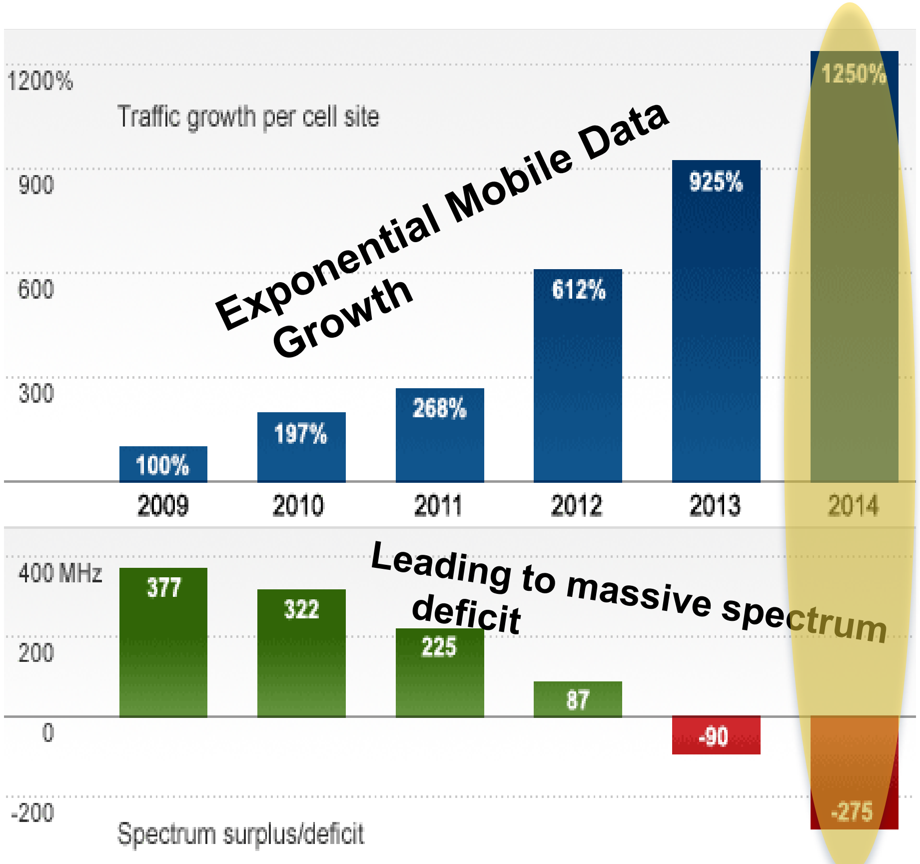
- Network Challenges
  - High performance
  - Extreme energy efficiency
  - Scarce/bifurcated spectrum
  - Heterogeneous networks
  - Reliability and coverage
  - Seamless internetwork handoff



- Device/SoC Challenges
  - Performance
  - Complexity
  - Size, Power, Cost
  - High frequencies/mmWave
  - Multiple Antennas
  - Multiradio Integration
  - Coexistance



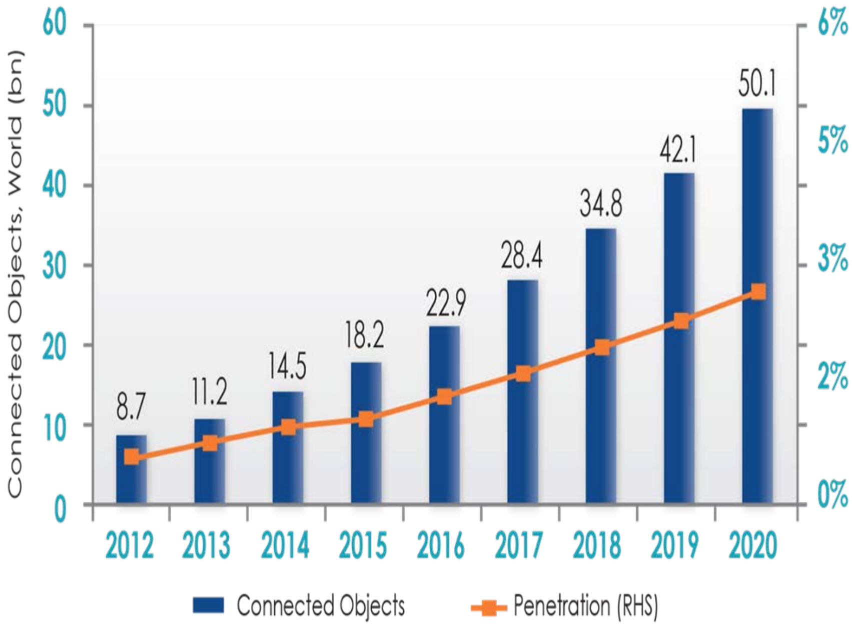
“Sorry America, your airwaves are full*”



Source: FCC

On the Horizon:

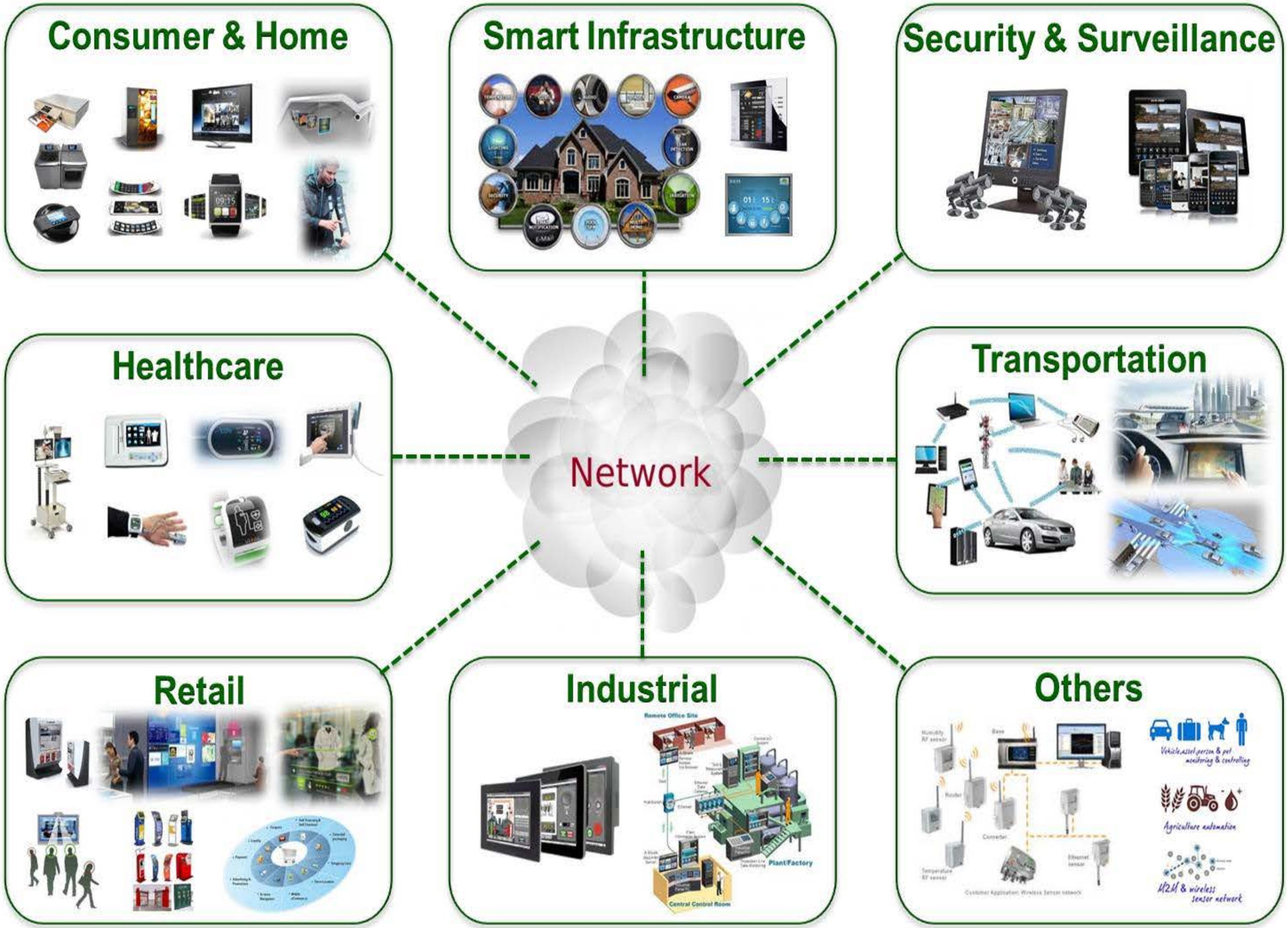
“The Internet of Things”



50 billion devices by 2020

***CNN MoneyTech – Feb. 2012**

IoT is not (completely) hype



Different requirements than smartphones: low rates/energy consumption

Are we at the Shannon limit of the Physical Layer?

We are at the Shannon Limit

- "The wireless industry has reached the theoretical limit of how fast networks can go" *K. Fitcher, Connected Planet*
- "We're 99% of the way" to the "barrier known as Shannon's limit," *D. Warren, GSM Association Sr. Dir. of Tech.*

Shannon was wrong, there is no limit

- "There is no theoretical maximum to the amount of data that can be carried by a radio channel" *M. Gass, 802.11 Wireless Networks: The Definitive Guide*
- "Effectively unlimited" capacity possible via *personal* cells (pcells). *S. Perlman, Artemis.*

What would Shannon say?

We don't know the Shannon capacity of most wireless channels

- Time-varying channels.
- Channels with interference or relays.
- Cellular systems
- Ad-hoc and sensor networks
- Channels with delay/energy/$$$ constraints.

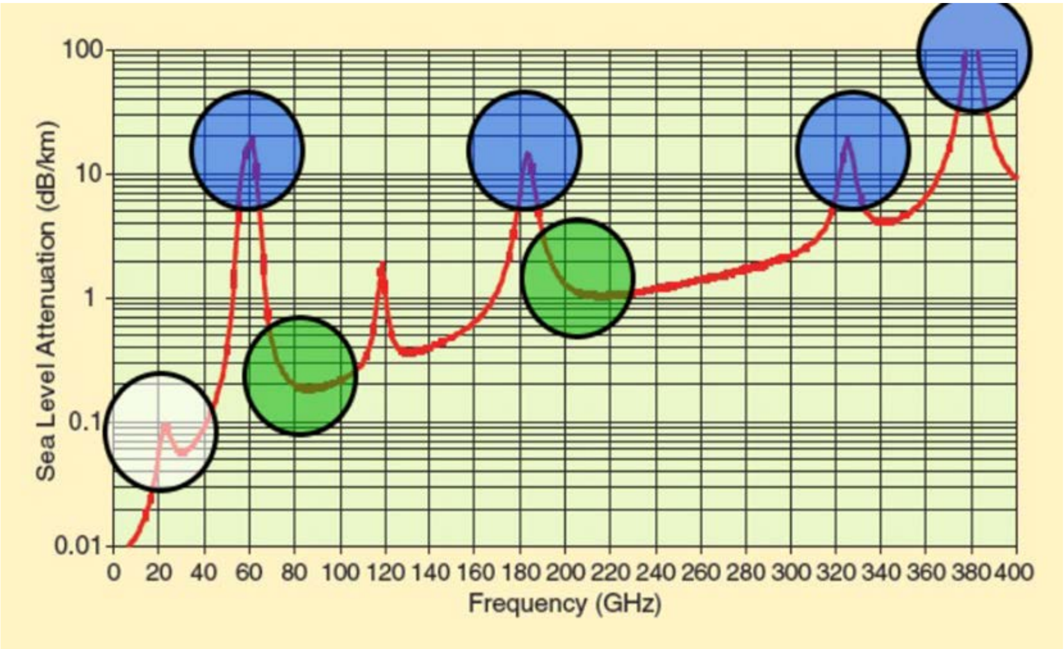
Shannon theory provides design insights and system performance upper bounds

Enablers for increasing wireless data rates

- More spectrum (mmWave)
- (Massive) MIMO
- Innovations in cellular system design
- Software-defined wireless networking
- Cognitive radios

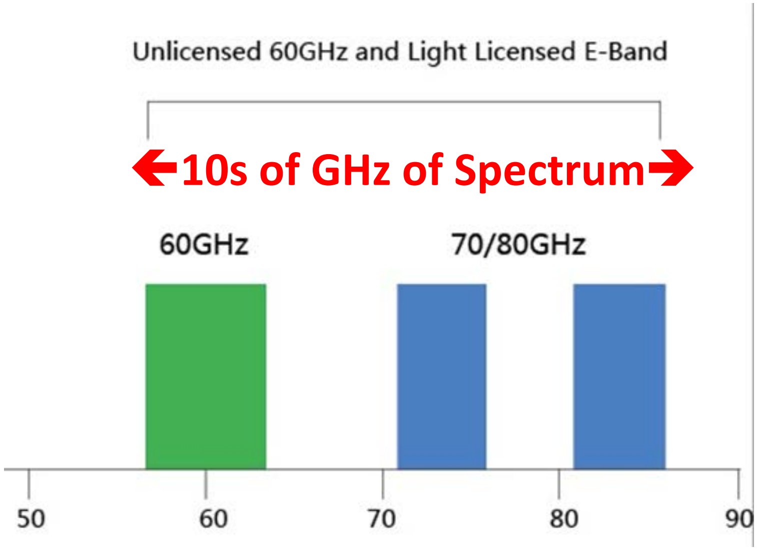
mmW as the next spectral frontier

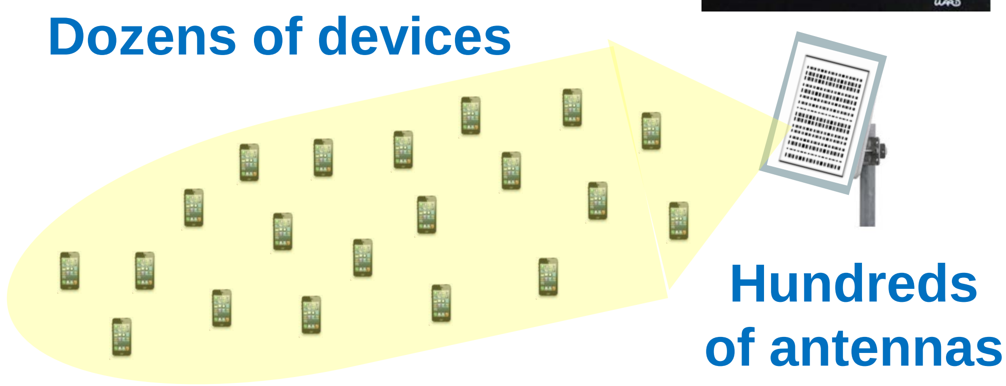
- Large bandwidth allocations, far beyond the 20MHz of 4G
- Rain and atmosphere absorption not a big issue in small cells



  - Not that high at some frequencies; can be overcome with MIMO
- Need cost-effective mmWave CMOS; products now available
- Challenges: Range, cost, channel estimation, large arrays

mmWave Massive MIMO

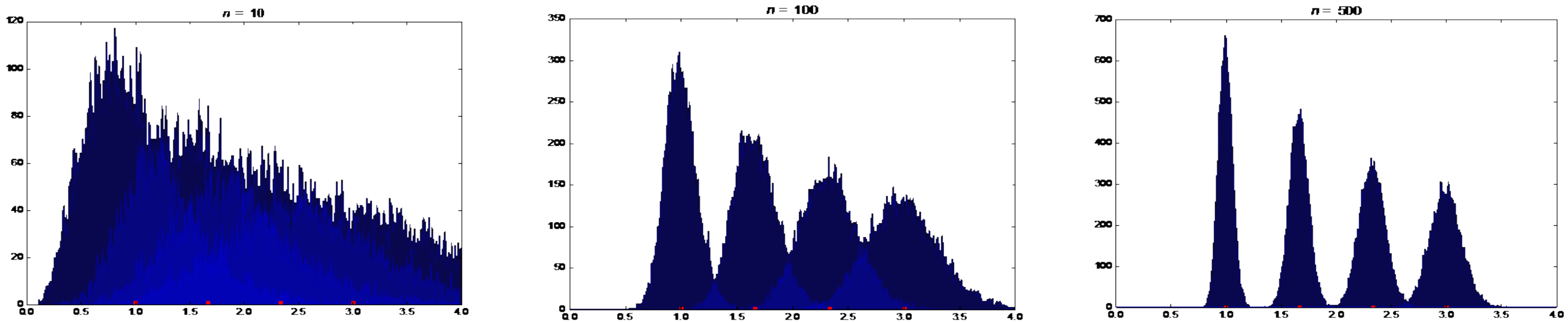




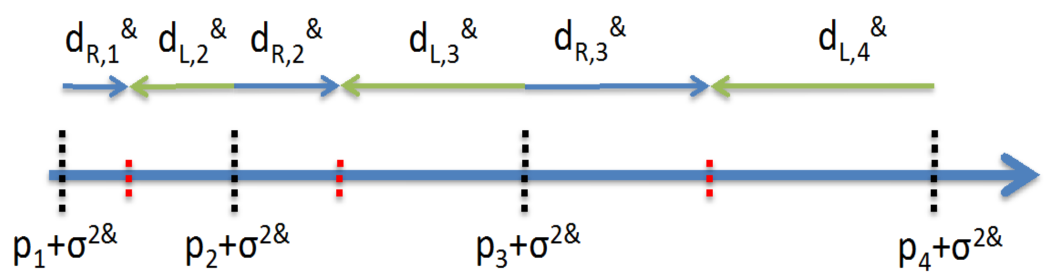
- mmWaves have large attenuation and path loss
- For asymptotically large arrays with channel state information, no attenuation, fading, interference or noise
- mmWave antennas are small: perfect for massive MIMO
- Bottlenecks: channel estimation and system complexity
- Non-coherent design holds significant promise

Non-coherent massive MIMO

- Propose simple energy-based modulation
- No capacity loss for large arrays: $\lim_{n\to\infty} C_{nocsi} = \lim_{n\to\infty} C_{csi}$
  - Holds for single/multiple users (1 TX antenna, n RX antennas)

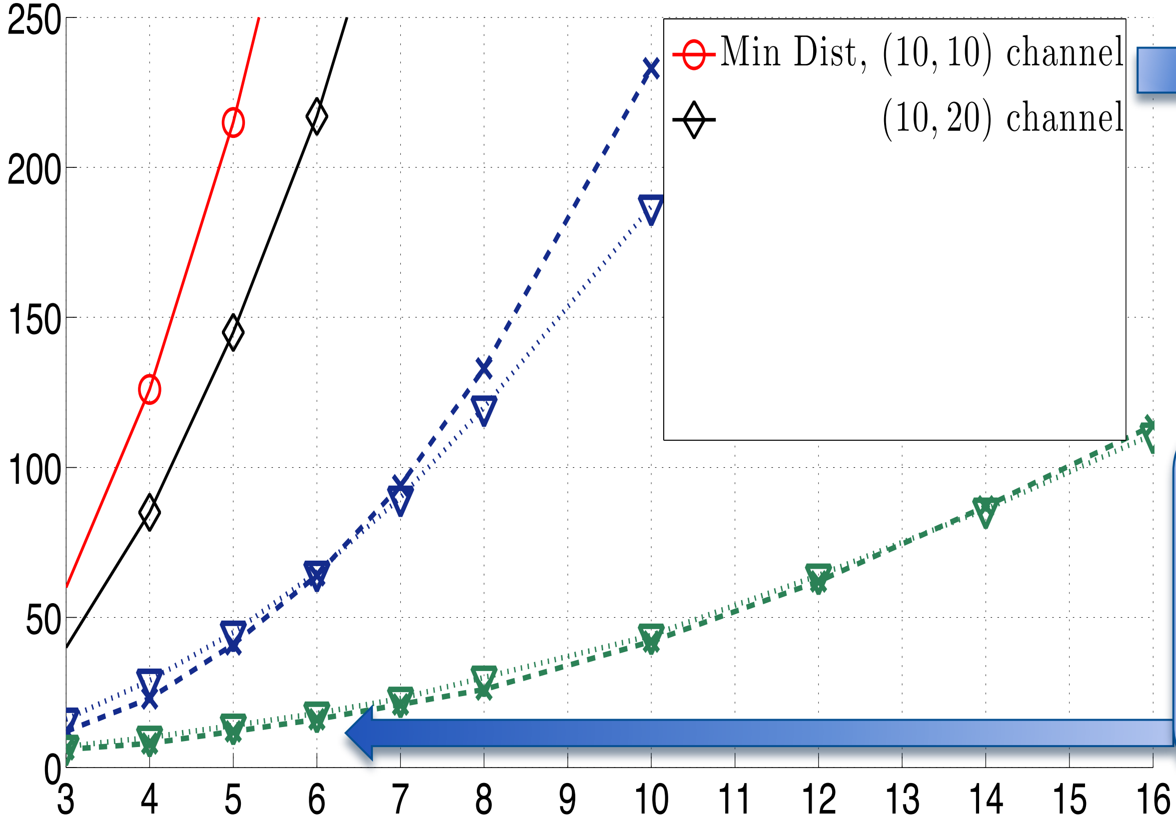


- Constellation optimization: unequal spacing



Joint with M. Chowdhury

Minimum number of Receive Antennas: SER= 10^{-4}

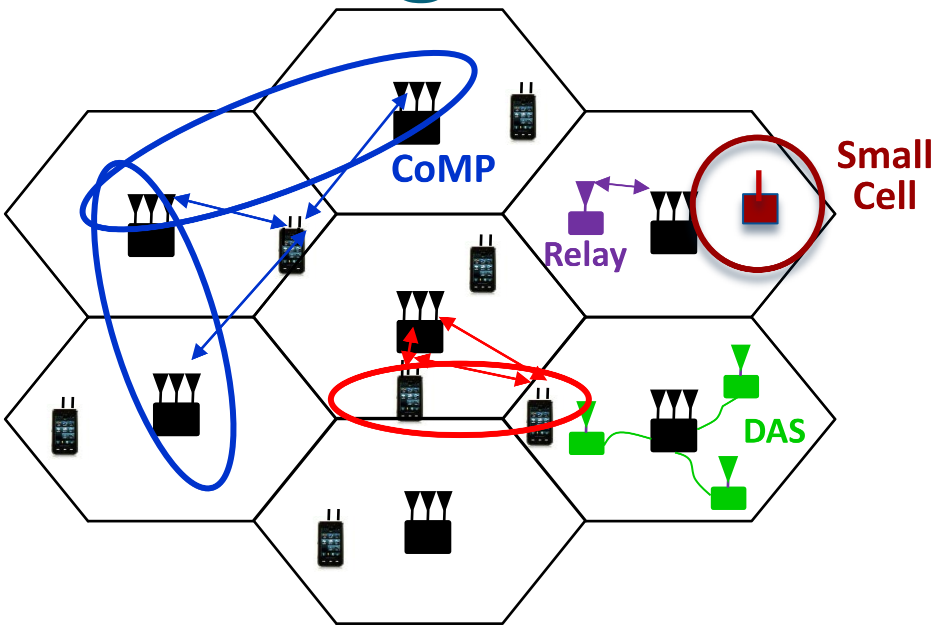


Minimum Distance Design criterion: Significantly worse performance than the new designs.

For low constellation sizes and low uncertainty interval, robust design demonstrates better performance.

Noncoherent communication demonstrates promising performance with reasonably-sized arrays

Rethinking Cellular System Design



How should cellular systems be designed?

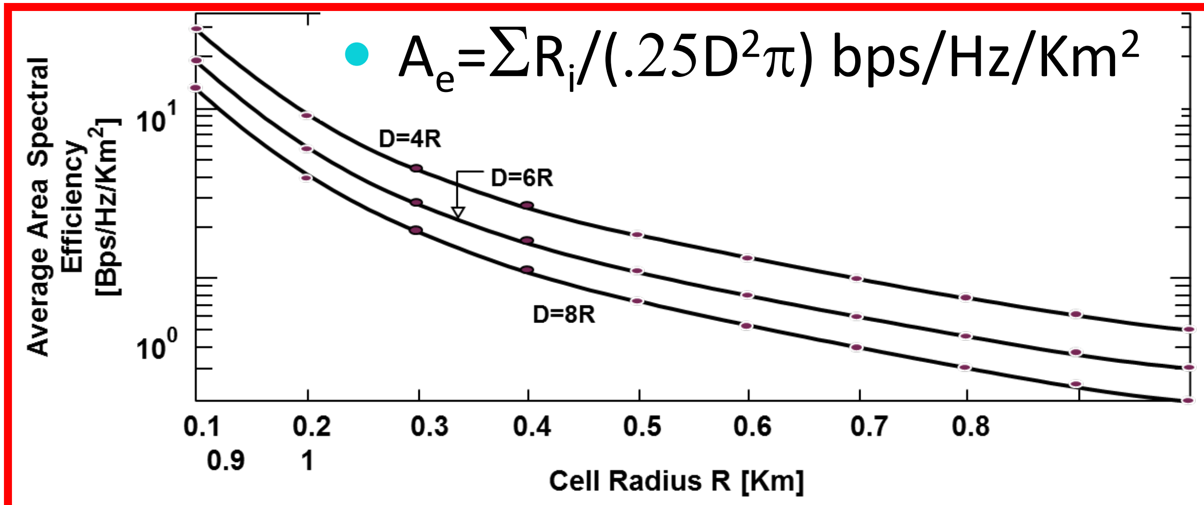
Will gains be big or incremental; in capacity, coverage or energy?

- Traditional cellular design assumes system is “interference-limited”
- No longer the case with recent technology advances:
  - MIMO, multiuser detection, cooperating BSs (CoMP) and relays
- Raises interesting questions such as “what is a cell?”
- Energy efficiency via distributed antennas, small cells, MIMO, and relays
- Dynamic self-organization (SoN) needed for deployment and optimization

Are small cells the solution to increase cellular system capacity?

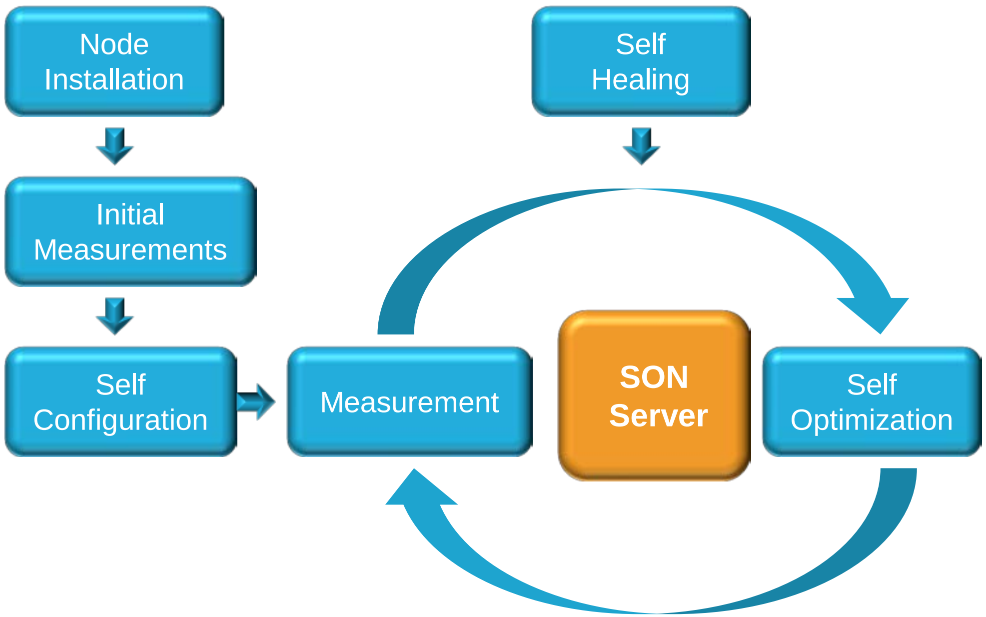
Yes, with reuse one and adaptive techniques (Alouini/Goldsmith 1999)

- $A_e = \Sigma R_i / (.25 D^2 \pi)$ bps/Hz/Km^2

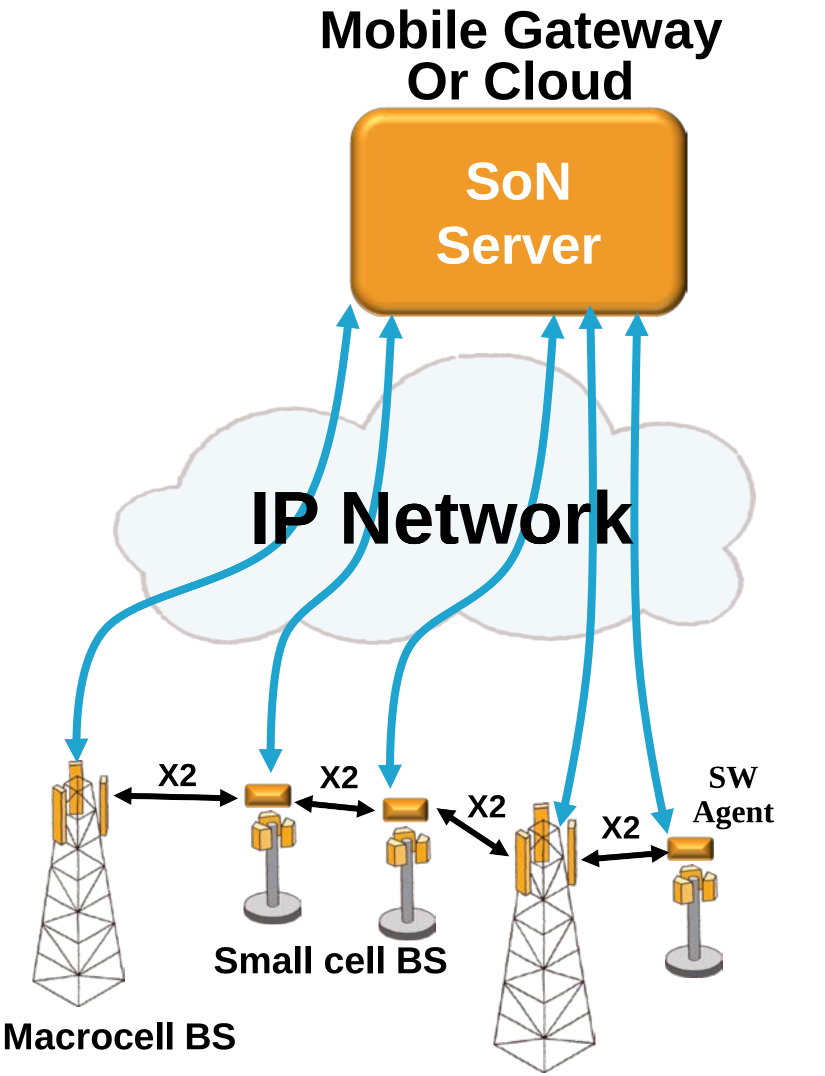


- Future cellular networks will be hierarchical (large and small cells)
  - Large cells for coverage, small cells for capacity/power efficiency
- Small cells require self-optimization (SoN) in the cloud

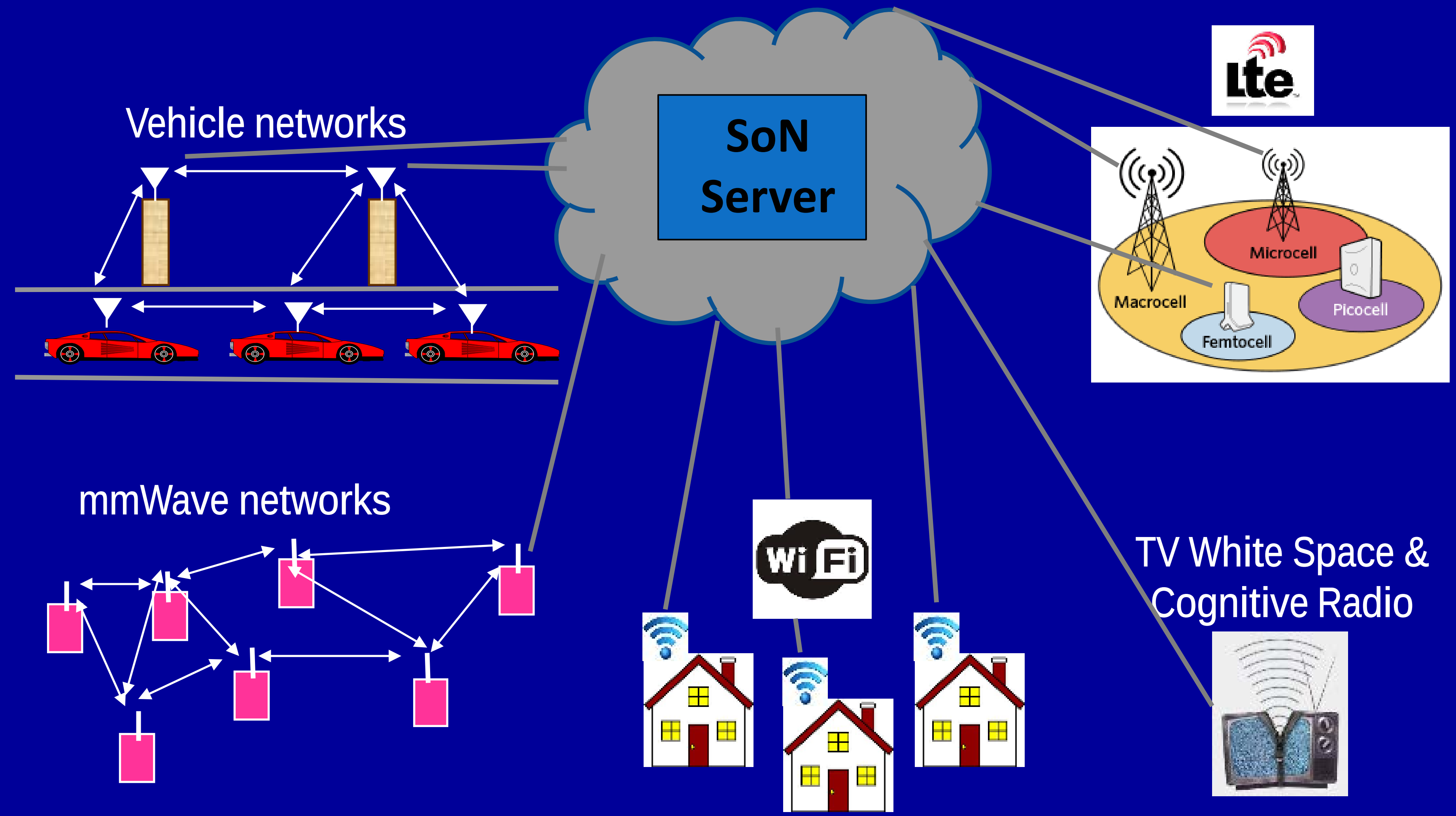
SON Premise and Architecture



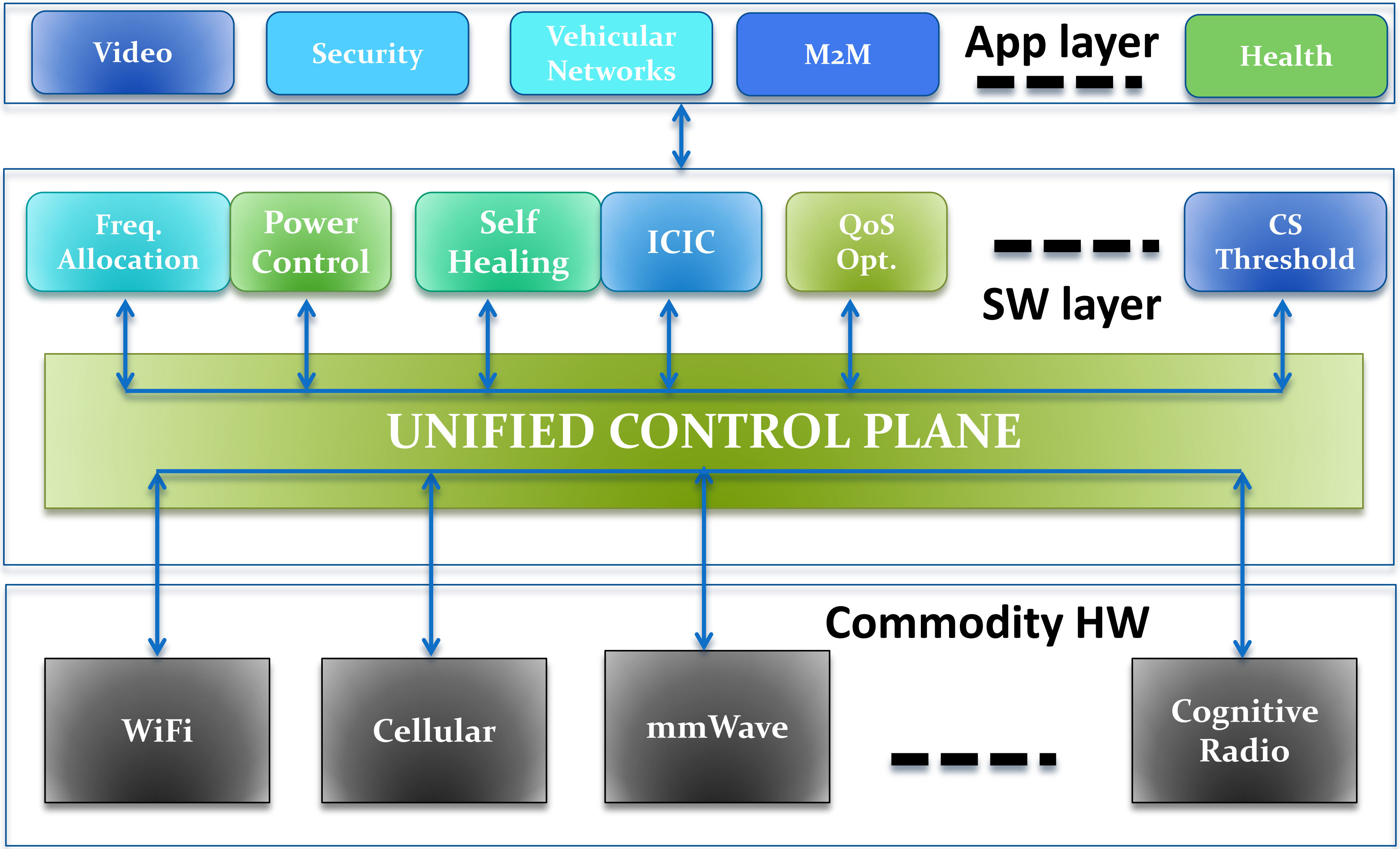
- Small Cell Challenges
  - SoN algorithmic complexity
  - Distributed versus centralized control
  - Backhaul
  - Site Acquisition
  - Resistance from macrocell vendors



Why not use SoN for all wireless networks



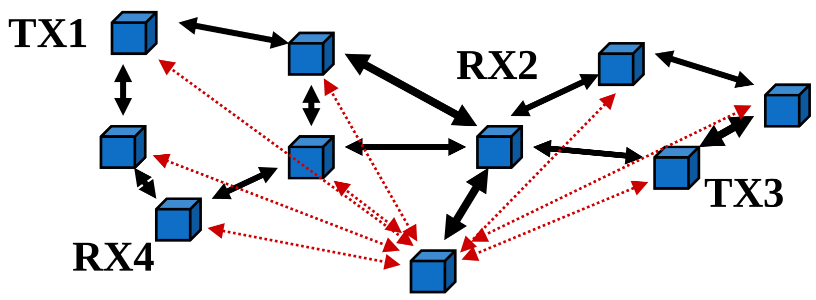
Software-Defined Network Architecture

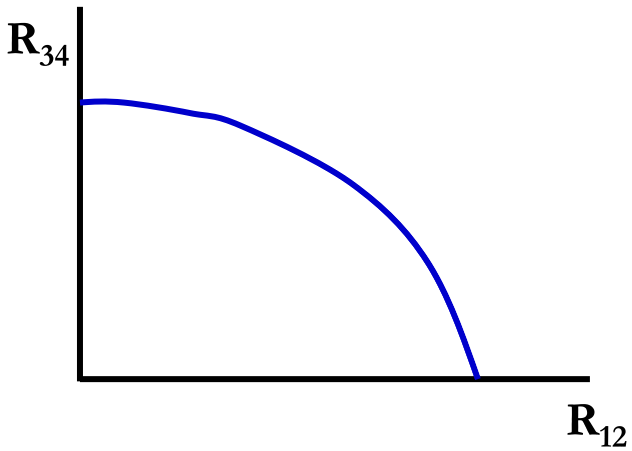


SDWN Challenges

- Algorithmic complexity
  - Frequency allocation alone is NP hard
  - Also have MIMO, power control, CST, hierarchical networks: *NP-really-hard*
  - Advanced optimization tools needed, including a combination of centralized (cloud), distributed, and locally centralized (fog) control
- Hardware Interfaces (especially for WiFi)
- Seamless handoff between heterogenous networks

Ad-hoc Networks and their Capacity

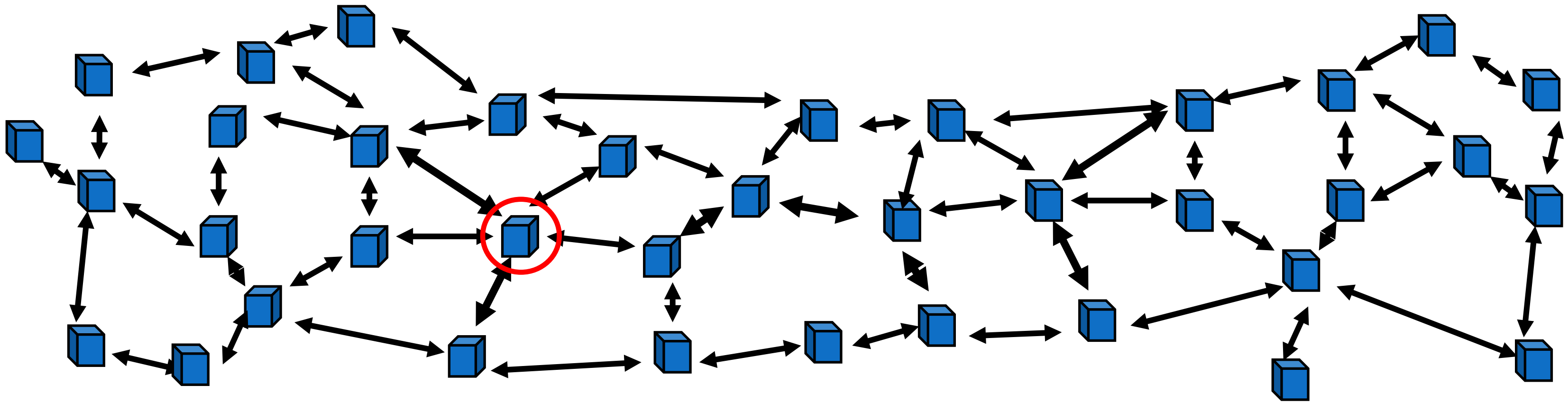




- Ad-hoc networks are fully connected
- Capacity: n(n-1)-dimensional region defining max. data rate between all node pairs with vanishing probability of error
- Ad-hoc network topology combines broadcast, multiple access, interference (IFC) and relay channels
- Lower bounds use coding strategies for these canonical systems
- Good upper bounds have been hard to obtain

Joint with S. Rini

Defining a coding scheme

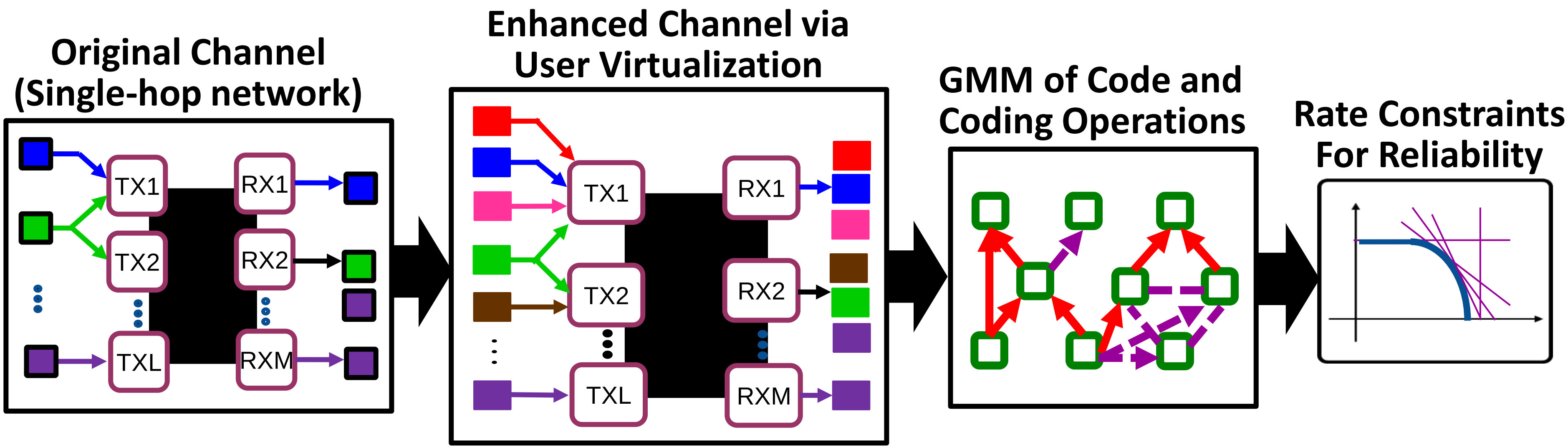


- For each node in the network, scheme indicates
  - To superposition encode, or not
  - To rate split, or not
  - To bin, or not
  - To fully or partially interference decode, or not
  - To time share, or not
  - To relay, or not
- Encoding/decoding at a nodes depend on encoding/decoding at neighbors
- Common messages entail code layering
- Must do this for every node in the network

How to best combine the different techniques?

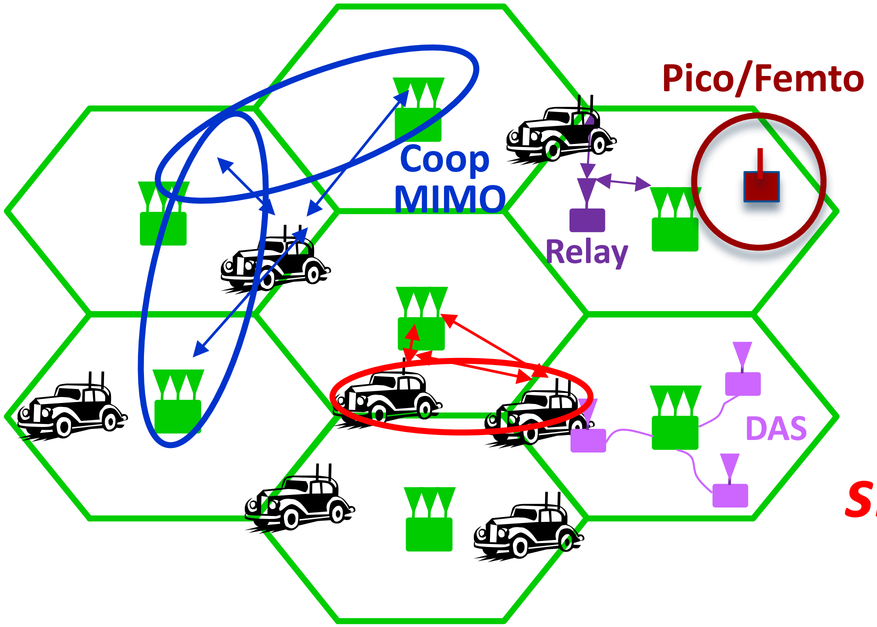
Coding and decoding possibilities grow exponentially with nodes

Unified approach to random coding



- Create virtual users via rate splitting
  - Interference from split message can be decoded and removed
- Use a Graphical Markov Model (GMM) to capture conditional dependencies of codewords due to its set of superposition, splitting and binning operations
- Use packing (max rates in superposition) and covering (rate penalty due to binning) lemmas to define rate bounds for reliable decoding

"Green" Wireless Networks



How should wireless systems be redesigned for minimum energy?

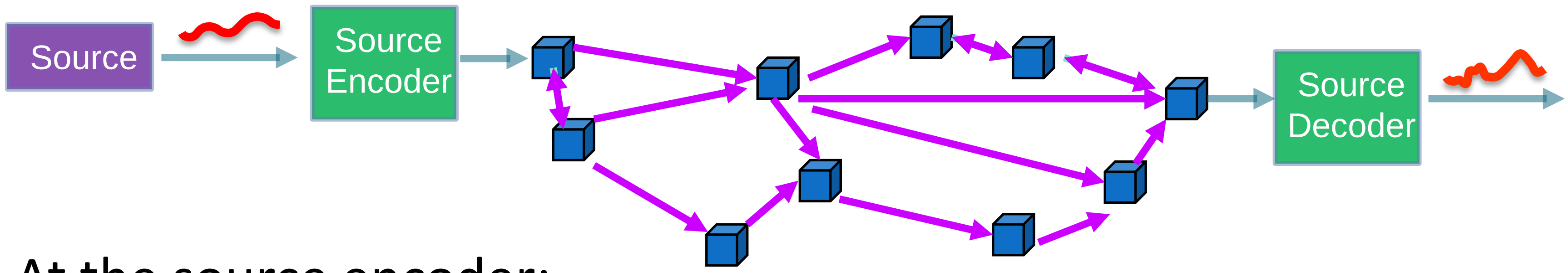
Research indicates that significant savings is possible

- Drastic energy reduction needed (especially for IoT)
  - New Infrastructures: Cell Size, BS/AP placement, Distributed Antennas (DAS), Massive MIMO, Relays
  - New Protocols: Coop MIMO, RRM, Sleeping, Relaying
  - Low-Power (Green) Radios: Radio Architectures, Modulation, Coding, Massive MIMO

Energy-Constrained Radios

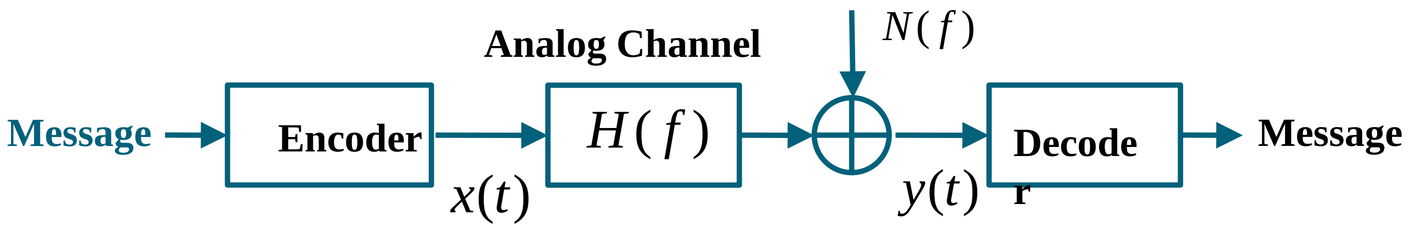
- Transmit energy minimized by sending bits very slowly
  - Leads to increased circuit energy consumption
- Short-range networks must consider both transmit and processing/circuit energy.
  - Sophisticated encoding/decoding not always energy-efficient.
  - MIMO techniques not necessarily energy-efficient
  - Long transmission times not necessarily optimal
  - Multihop routing not necessarily optimal
- Recent work to minimize energy consumption in radios
  - Sub-Nyquist sampling
  - Codes to minimize total energy consumption

Benefits of Sub-Nyquist Sampling



- At the source encoder:
  - Fewer bits to transmit
- At each receiver
  - Fewer bits to process or relay
- We have determined
  - Capacity/optimal transmission for sub-Nyquist-sampled channels
  - Rate-distortion theory for sub-Nyquist sampled sources

Sub-Nyquist Sampled Channels



C. Shannon

Wideband systems may preclude Nyquist-rate sampling!

H. Nyquist

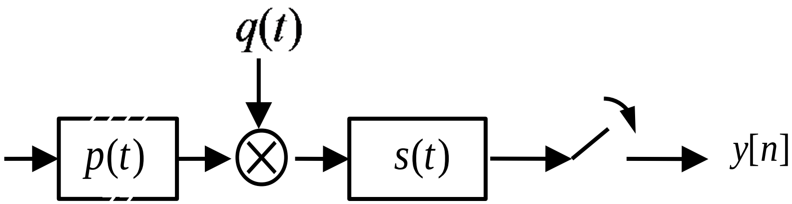
Sub-Nyquist sampling well explored in signal processing

- ***Landau-rate sampling, compressed sensing, etc.***
- ***Performance metric: MSE***

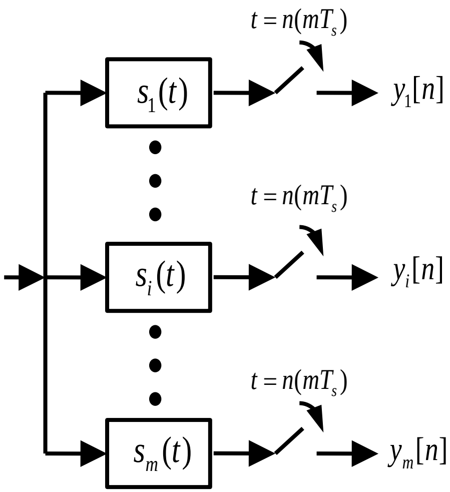
We ask: what is the capacity-achieving sub-Nyquist sampler and communication design

Joint with Y. Chen, Y. Eldar

Optimal Sub-Nyquist Sampling



or



- Also optimal non-uniform sampling technique
- For channel unknown, random sampling optimal

- **Theorem: Capacity of the sampled channel using a bank of *m* filters with aggregate rate f_s**

$$C(f_s) = \frac{1}{2} \max_{Q(f)\in\mathcal{Q}} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \frac{1}{m} \log\left(\det\left(\mathbf{I}_m + f_s^{\dagger} f_h^{\frac{1}{2}} \mathbf{Q} f_h^{\frac{1}{2}} f_s^{\dagger}\right)\right) df$$

Similar to MIMO

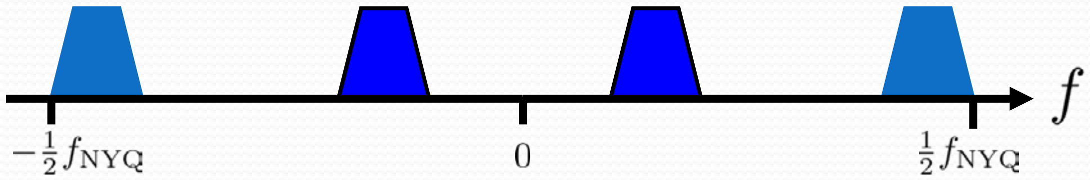
Water-filling over singular values

MIMO – Decoupling

Pre-whitening

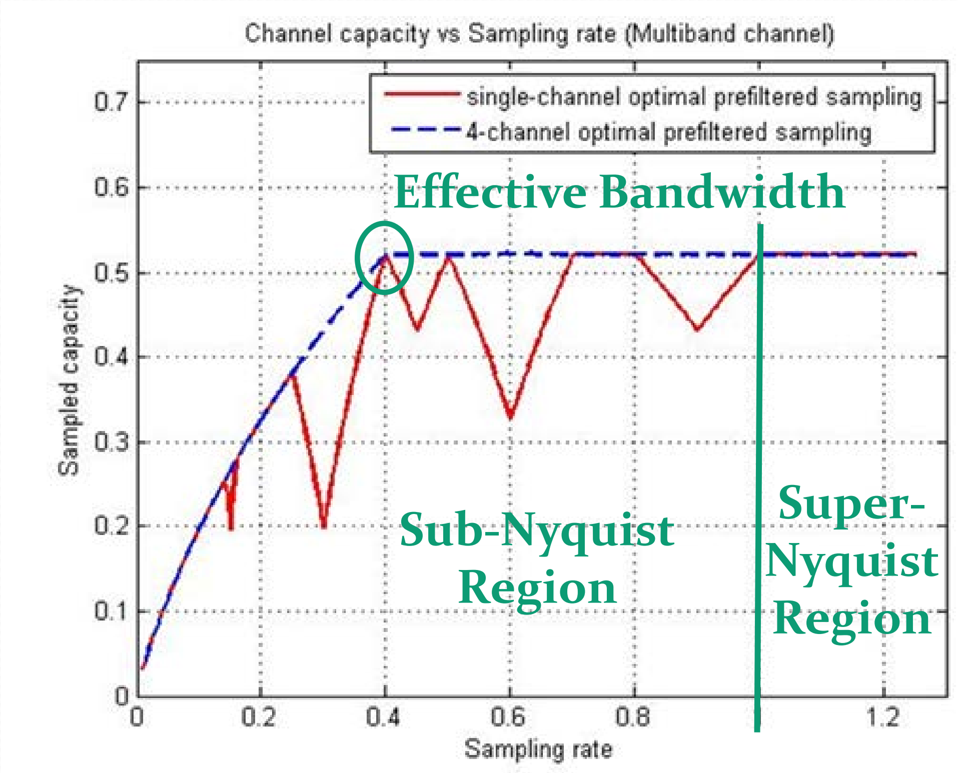
Example: Sparse Channels

- "Sparse" channel model

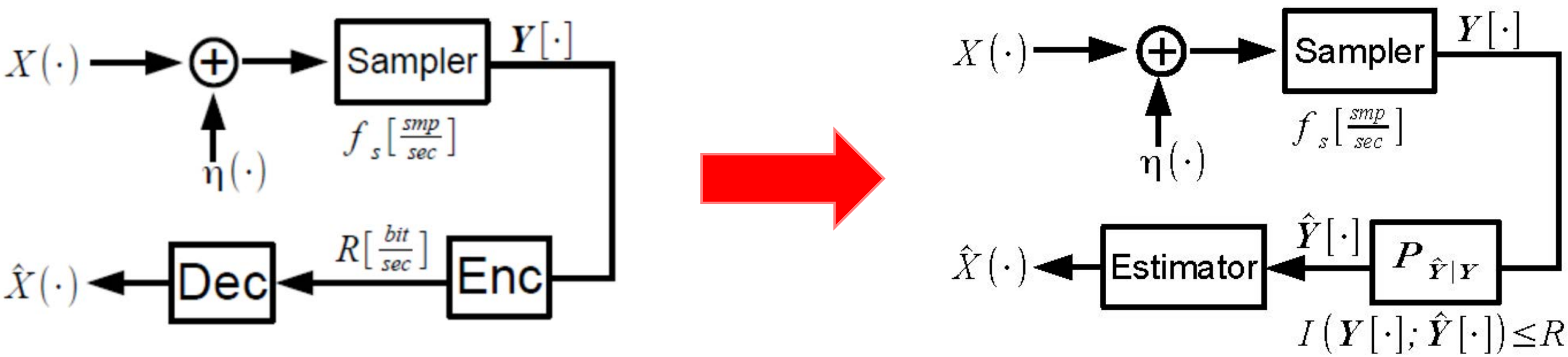


- Capacity not monotonic in f_s for 1 branch

- Capacity monotonic in f_s for enough branches



Unified Rate Distortion/Sampling Theory



- Problem Statement: Find distortion as a function of R and f_s
  - Also find the optimal sampler and source encoder/decoder
  - Noisy Gaussian analog source, sampled and compressed

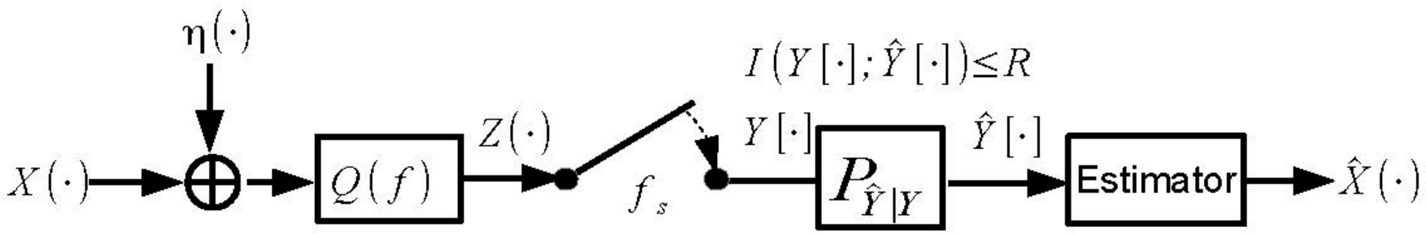
$$R = f_s I(\hat{Y}[\cdot];Y[\cdot]) = \sup_N \frac{f_s}{2N} \sum_{n=-N}^{N} I(\hat{Y}[\cdot];Y_{-N}^{N},\hat{Y}_{-N}^{n})$$

  - Metric: minimum MSE:

$$E\left(d(X(\cdot),\hat{X}(\cdot)\right) = \limsup_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} E\left(X(t)-\hat{X}(t)\right)^2 dt$$

Joint with A. Kipnis, T. Weissman, Y. Eldar

Main Results



Theorem:

$$R(\theta, f_s) = .5\int_{-.5 f_s}^{.5 f_s} \log^{+}\left[J(f)/\theta\right] df$$

Separation

$$D(\theta(R), f_s) = mmse\left(X(\cdot) \mid Y[\cdot]\right) + \int_{-.5 f_s}^{.5 f_s} \min\{J(f), \theta\} df$$

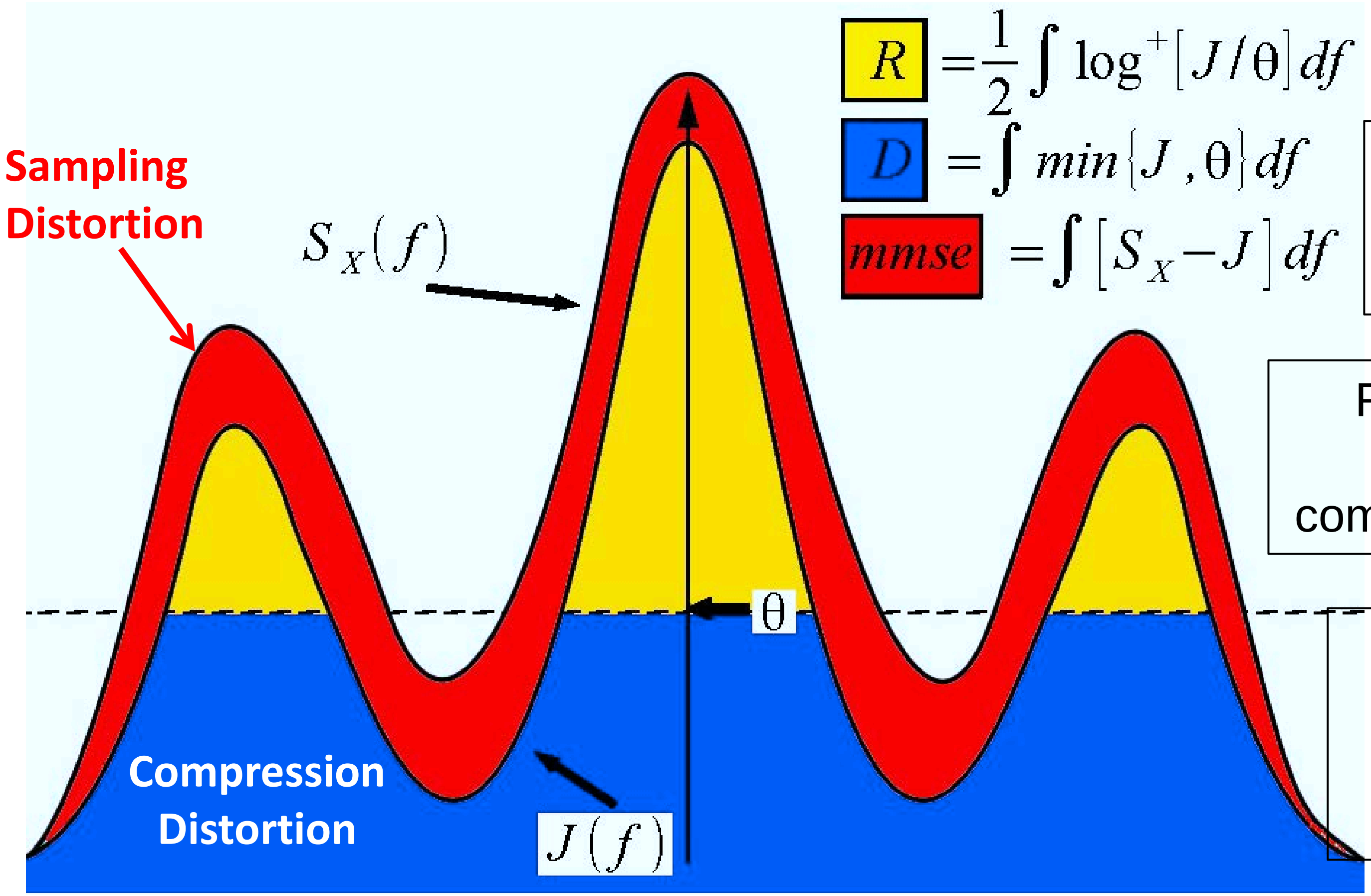
Sampling Component

Compression Component

$$J(f) = \frac{\sum_{k \in Z} S_X^2(f - f_s k)}{\sum_{k \in Z} S_X(f - f_s k)}$$

Optimal sampling rate f_{RD} is below Nyquist

Properties of the Solution



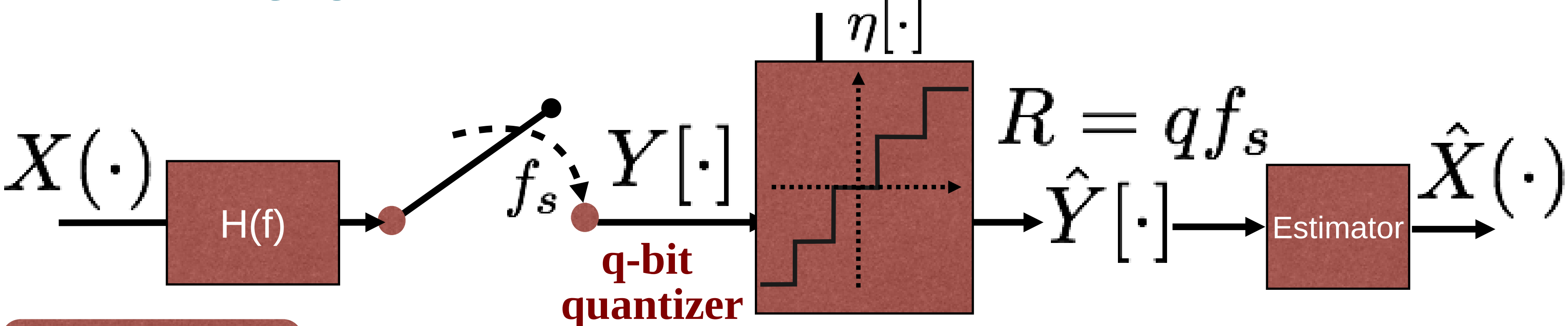
Preserve signal components above "noise floor" θ.

Rate R needed to describe these components: dictates θ

Distortion is signal components below noise floor + sampling distortion

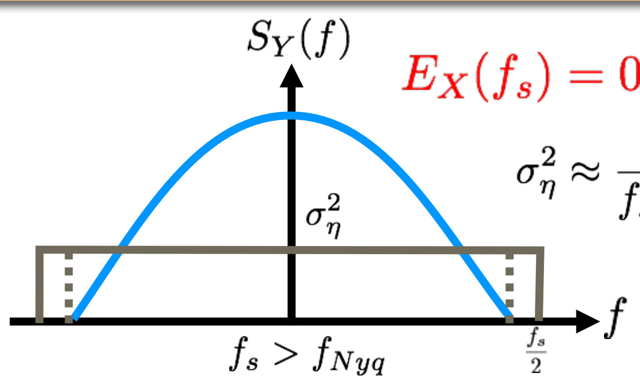
Distortion $= \mathrm{mmse}_{X|Y}(f_s)$+waterfilling over $J(f)$

SubNyquist ADCs with finite bit rate R

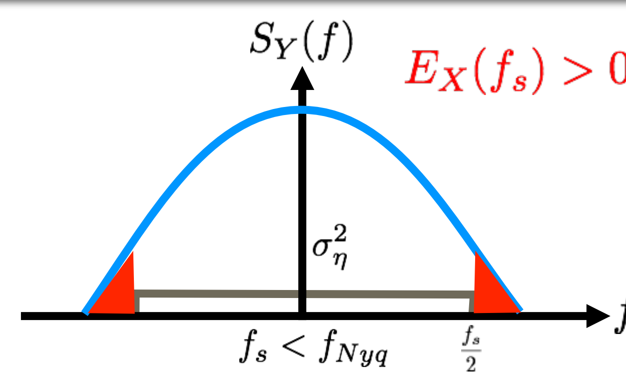


Theorem

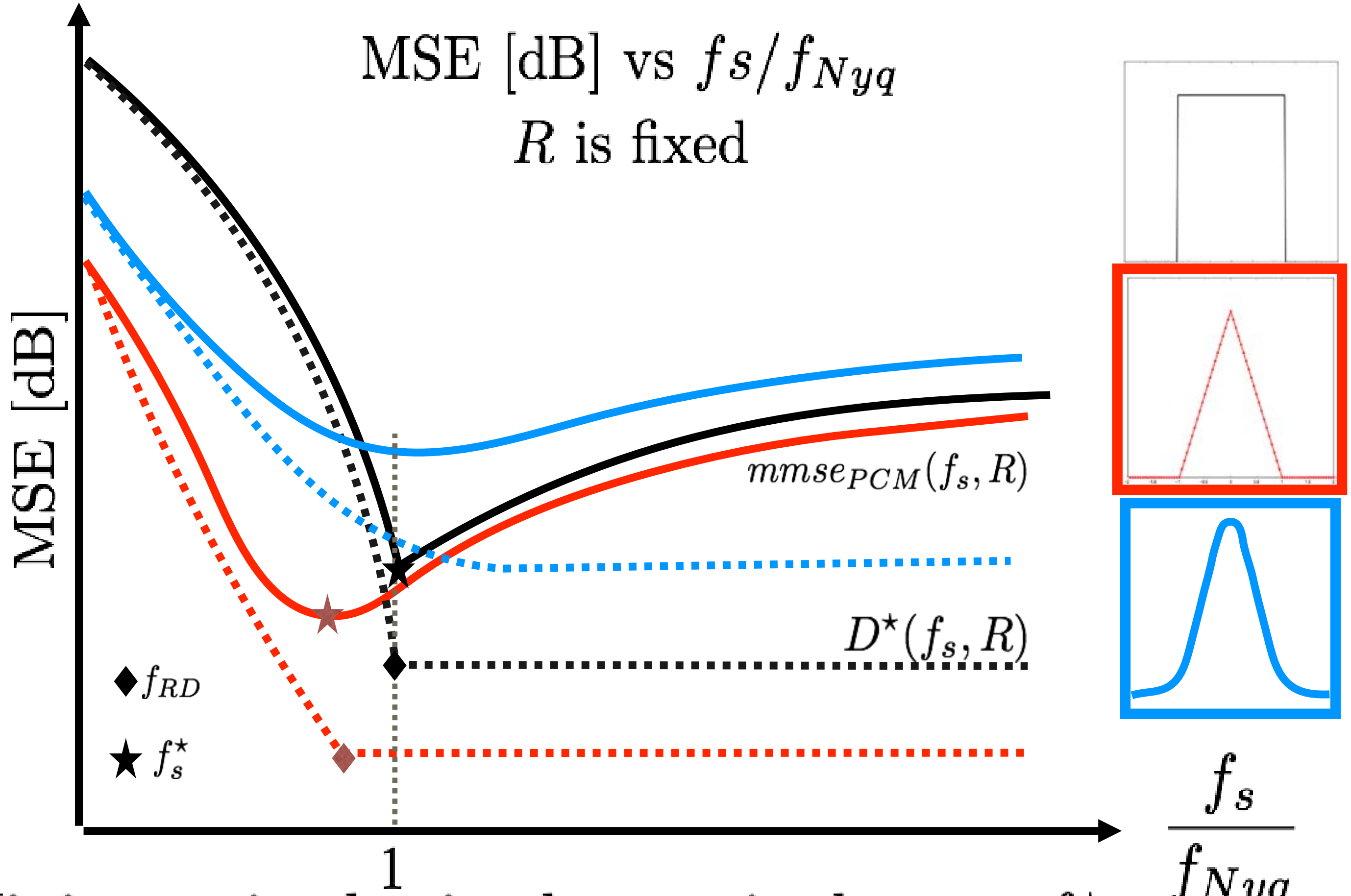
- **Should you sample fast w/low precision (Sigma-Delta A/D)**
- **Or sample below Nyquist with more precision**



$$\sigma_\eta^2 \approx \frac{c_0}{f_s 2^{2q}} = \frac{c_0}{f_0 2^{2R/f_s}}$$



MSE Depends on Input Signal



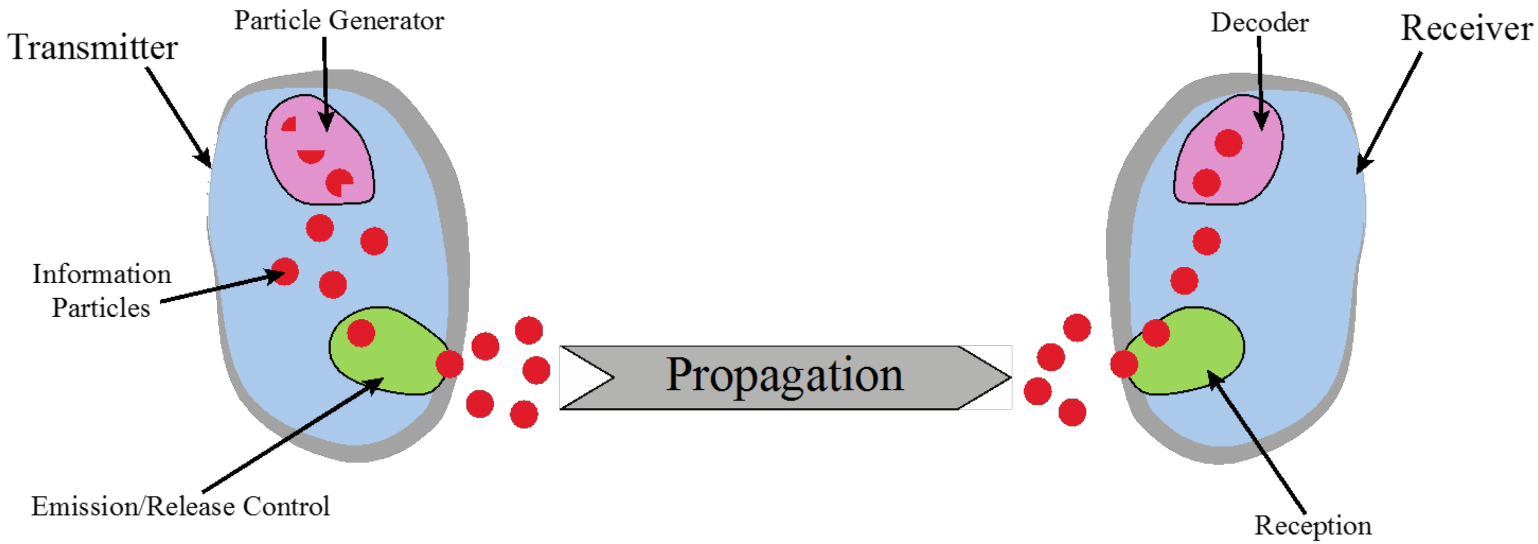
- Minimum is obtained at a single rate $f_s^\star$
- $f_s^\star \leq f_{Nyq}$ depending on $S_X(f)$

Where should energy come from?

- Batteries and traditional charging mechanisms
  - Well-understood devices and systems
- Wireless-power transfer
  - Poorly understood, especially at large distances and with high efficiency
- Communication with Energy Harvesting Devices
  - Intermittent and random energy arrivals
  - Communication becomes energy-dependent
  - Can combine information and energy transmission
  - New principals for communication system design needed.

Applications of Communications and IT to biology, medicine, and neuroscience

Chemical Communications



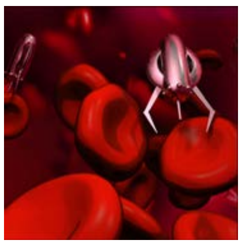
- Can be developed for both macro (>cm) and micro (<mm) scale communications
- Greenfield area of research:
  - Need new modulation schemes, channel impairment mitigation, multiple acces, etc.

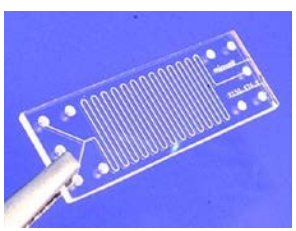
Joint with N. Farsad

Applications

MC Applications

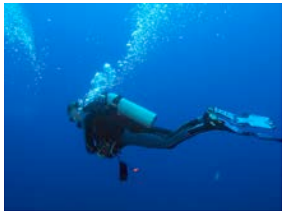
Microscale





- Medicine
- On-Chip MC
- Nano Manufacturing …

Macroscale

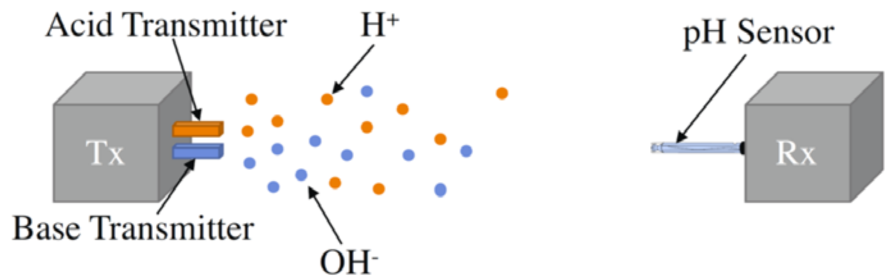


- City infrastructure monitoring
- Search and Rescue
- Underwater …

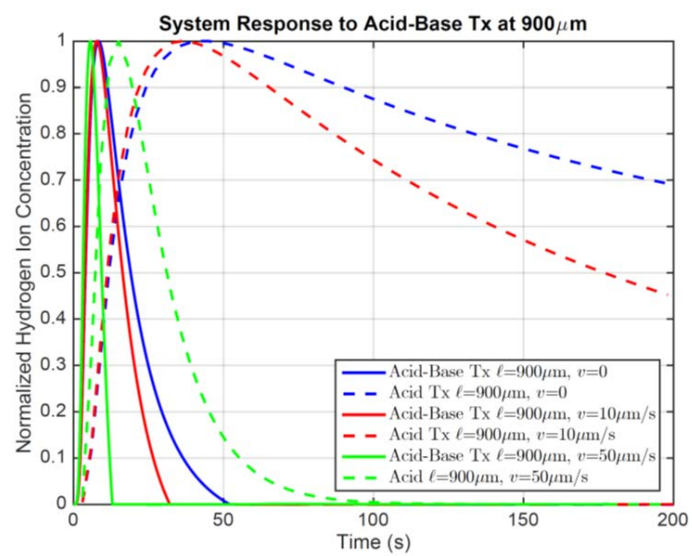
Data rate: .5 bps

“fan-enhanced” channel

Current Work



- Concentration system has limited control on the concentration at the receiver.
- Can use acid/base transmission to decrease concentration (ISI)
- Similar ideas can be applied for multilevel modulation and multiuser techniques

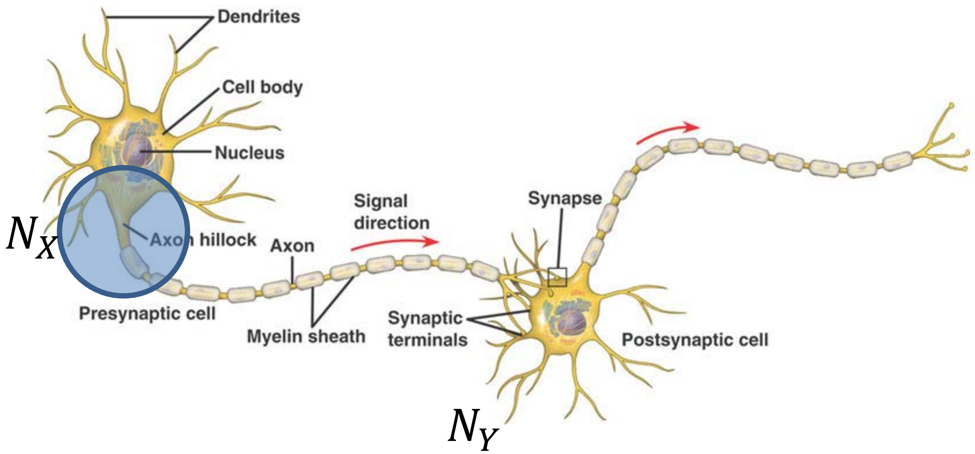


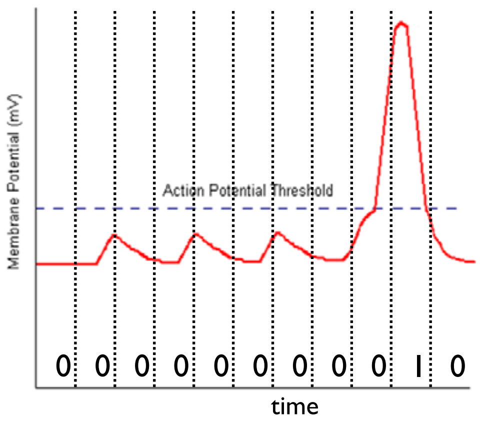
The brain as a network

Joint with N. Soltani, T. Coleman, R. Ma, J. Kim, and J. Parvizi

Neuronal Signaling

- Communication done through action potentials (spikes)





- Observe spike trains

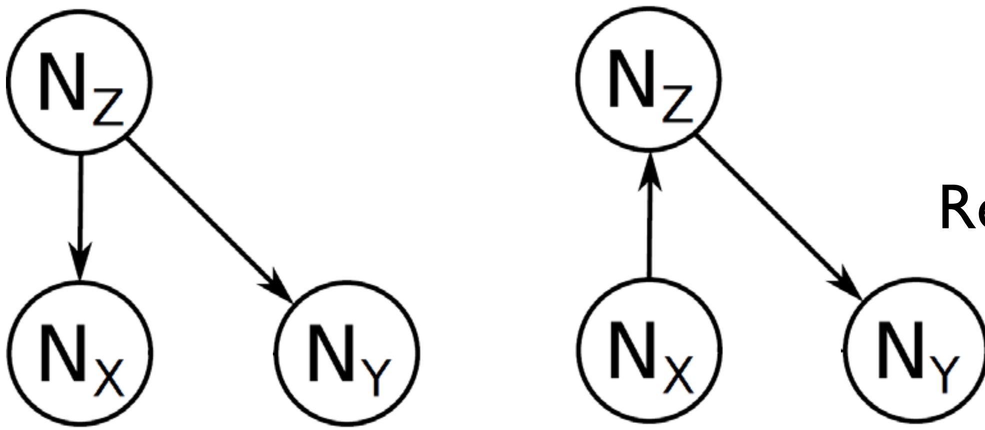
$$X^n = [0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \ \dots]$$
$$Y^n = [0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \ \dots]$$

- <u>Goal</u>: Determine physical connections between neurons
  - Aids in fundamental understanding of how the brain works
  - Can be used to study learning and degeneration

Directed Information

- $I(X^n \rightarrow Y^n) > 0$ necessary for synapse to exist
  - Not sufficient ➔ leads to false positives

Broadcast



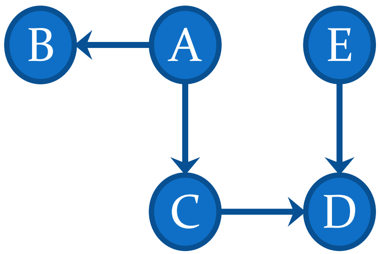
Relay

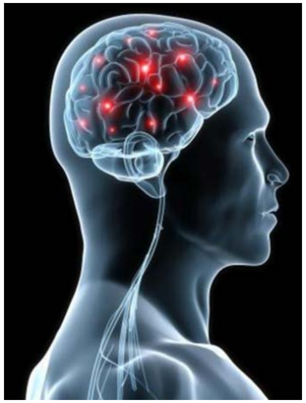
  - Can remove false positives by observing all neurons
    - Like Maximum-likelihood detection
    - But we can't observe all neurons
    - Delay of "relay" can mitigate false positives

Kim et al. (2011),
Quinn et al. (2011)

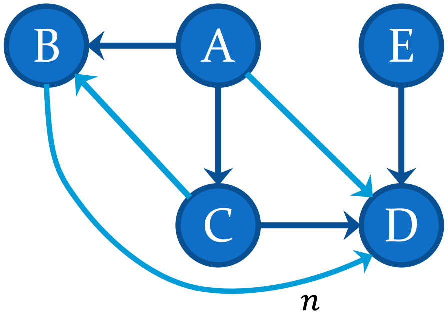
Pathways through the brain

Neuron layout



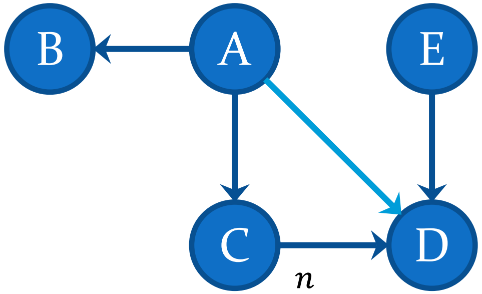


DI inference



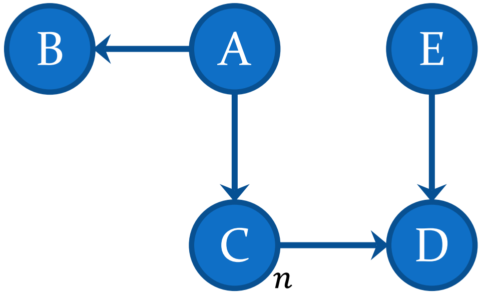
$$I(X^n \to Y^n) = H(Y^n) - \sum_{i=1}^{n} H\left(Y_i | Y^{i-1}, X^i\right)$$

DI inference with delay lower bound



$$I(X^n \to Y^n) = H(Y^n) - \sum_{i=1}^{n} H\left(Y_i | Y^{i-1}, X^{i-D}\right)$$

Constrained DI inference



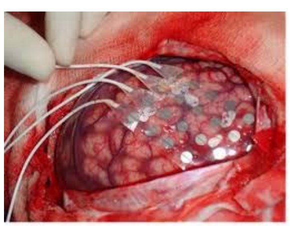
$$I(X^n \to Y^n) = H(Y^n) - \sum_{i=1}^{n} H\left(Y_i | Y^{i-1}, X_{i-D-N}^{i-D}\right)$$

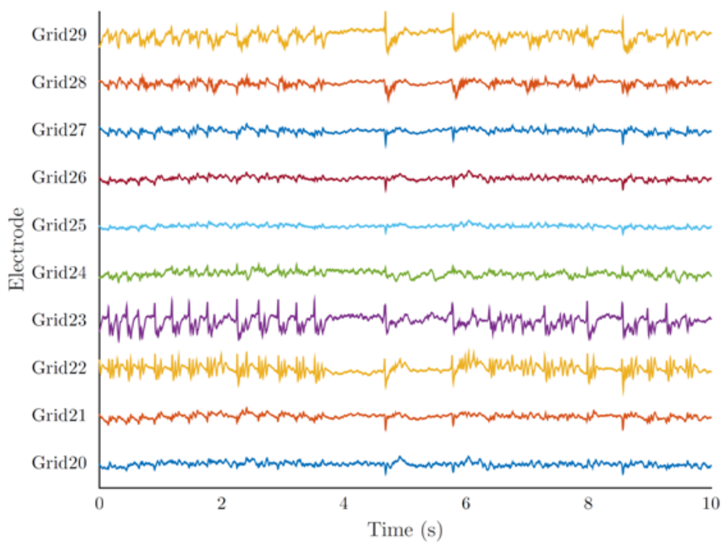
We've developed a DMI model for the leaky integrate-and-fire neuron

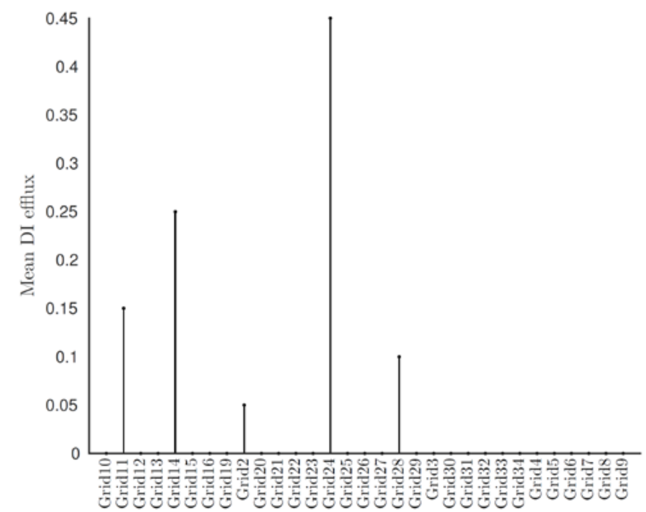
Epileptic Seizure Focal Points

- Seizure caused by an oscillating signal moving across neurons
  - When enough neurons oscillate, a seizure occurs
  - Treatment "cuts out" signal origin: errors have serious implications
- Directed mutual information spanning tree algorithm applied to ECoG measurements estimate the focal point of the seizure
- Application of our algorithm to existing data sets on 3 patients matched well with their medical records

ECoG Data







Electrocortical Silencer

Waveform Generator

F(ε(t))

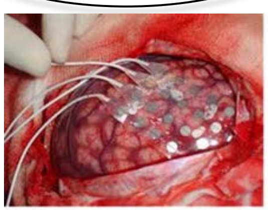
Adaptively change output, ε'(t), based on input, ε(t)

Personalized Electric Current Stimulation ε'(t)

Recorded Voltage ε(t)

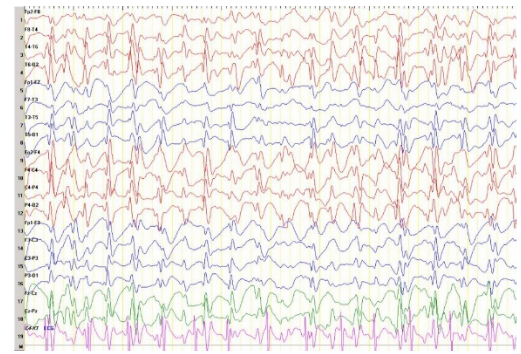
G(ε(t),ε'(t))

Stimulating and Recording Electrode

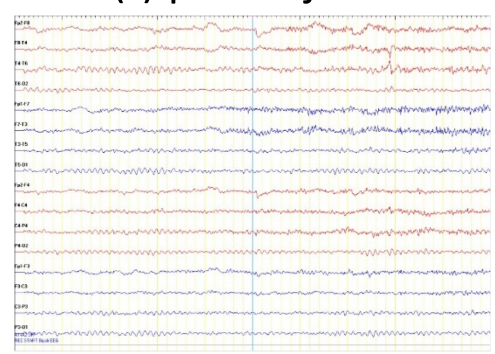


Goal: Silence the epileptic firing in the cortex

ε(t) pre-injection



ε(t) post-injection



Current Status

- Trials on human subjects to start in December
- Also can be used for Parkinson's and depression

Summary

- The next wave in wireless technology is upon us
  - This technology will enable new applications that will change people's lives worldwide
- Future wireless networks must support high rates for some users and extreme energy efficiency for others
  - Small cells, mmWave massive MIMO, Software-Defined Wireless Networks, and energy-efficient design key enablers.
- Communication tools and modeling techniques may provide breakthroughs in other areas of science