CTW 2006
May 24, 2006

Energy and Inference in Wireless Sensor Networks

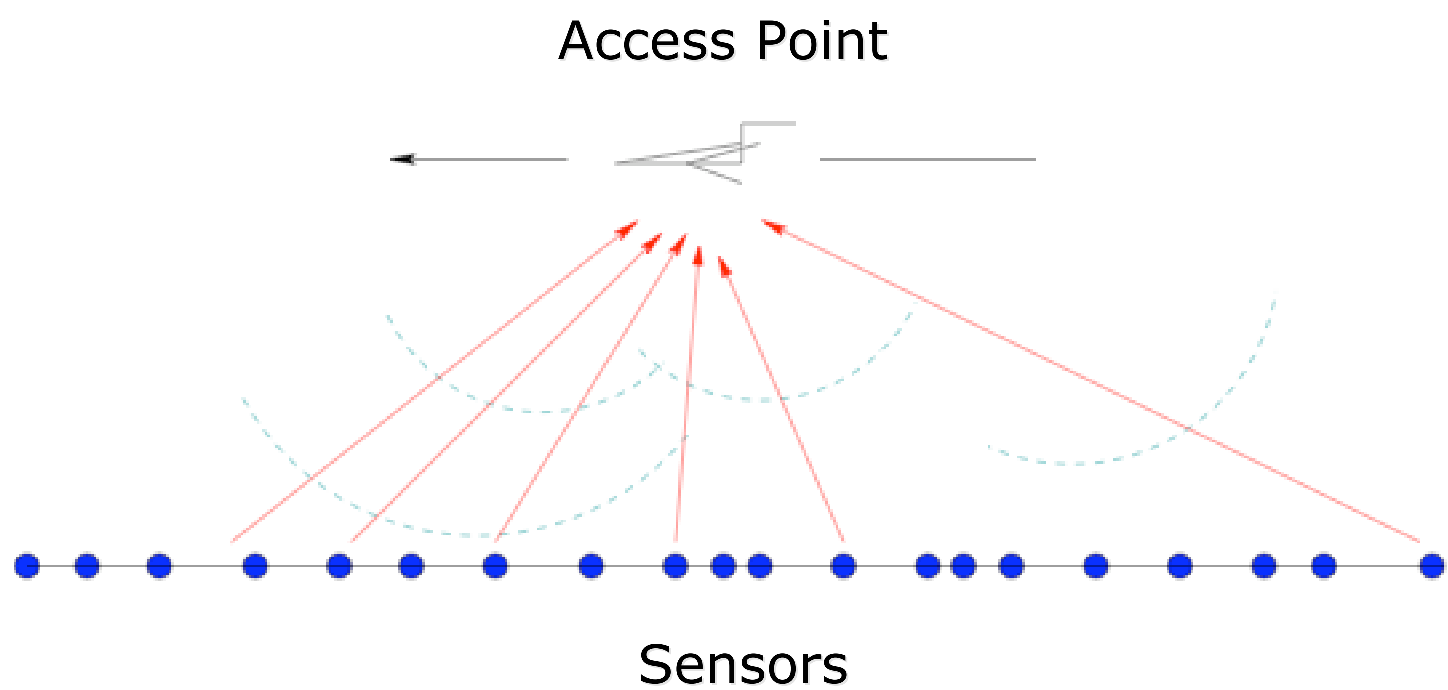
Vince Poor
(poor@princeton.edu)

Motivation

- Salient features of WSNs:
  - The primary application is inference
  - Information at different terminals is often correlated
  - Energy is often severely limited
- Research in WSNs:
  - Networking Issues: capacity, delay, routing, etc.
  - Applications Issues: primarily distributed inference
- Primary goal:
  - Optimize performance within constraints of wireless systems (i.e., "bandwidth & batteries")

Sensor Field

Basic Set Up



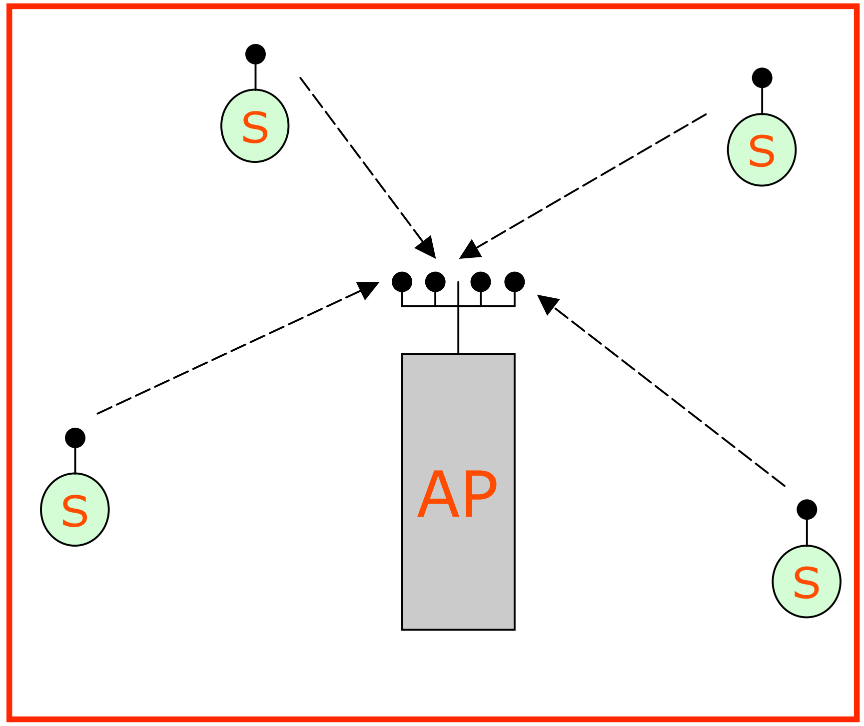
Topics of Today's Talk:

- Energy efficiency in shared-access networks
- Collaborative beamforming
- Energy issues in distributed inference (briefly)

ENERGY EFFICIENCY IN SHARED-ACCESS NETWORKS

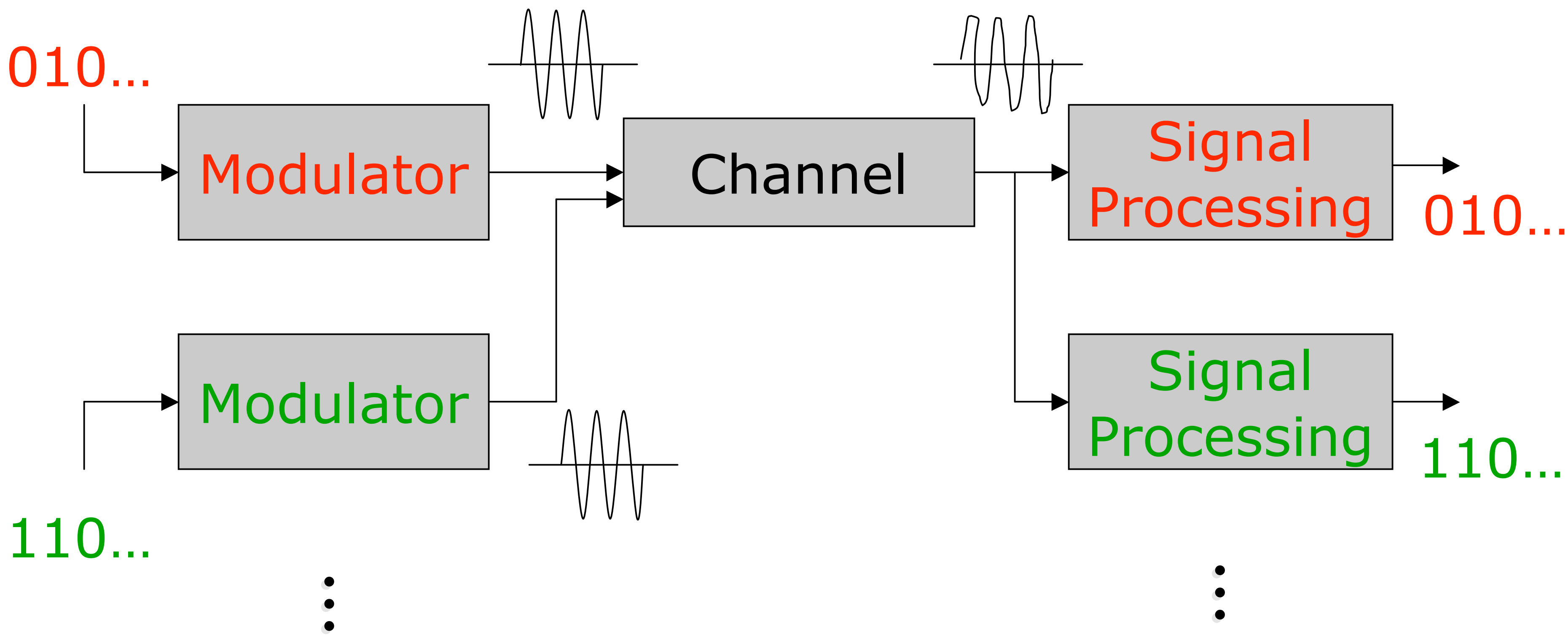
Competition in Shared-Access Networks

- Sensors transmit to an access point via a shared channel.
- Sensors are like players in a game, competing for resources to transmit their data to the AP.
- The action of each sensor affects the others.



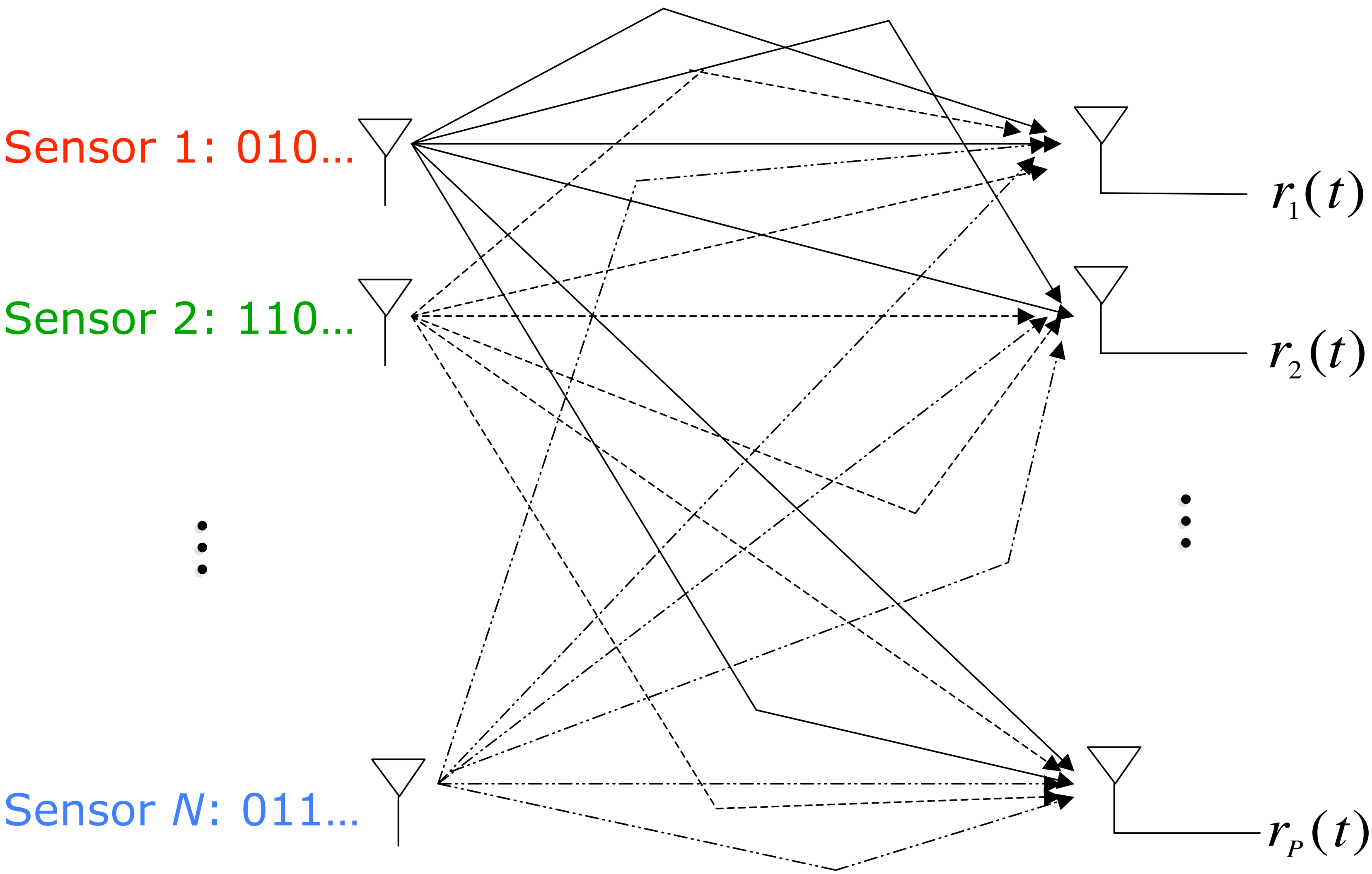
- Can model this as a non-cooperative game, with payoff measured in bits-per-joule.
- First, we digress …

Shared-Access Channel



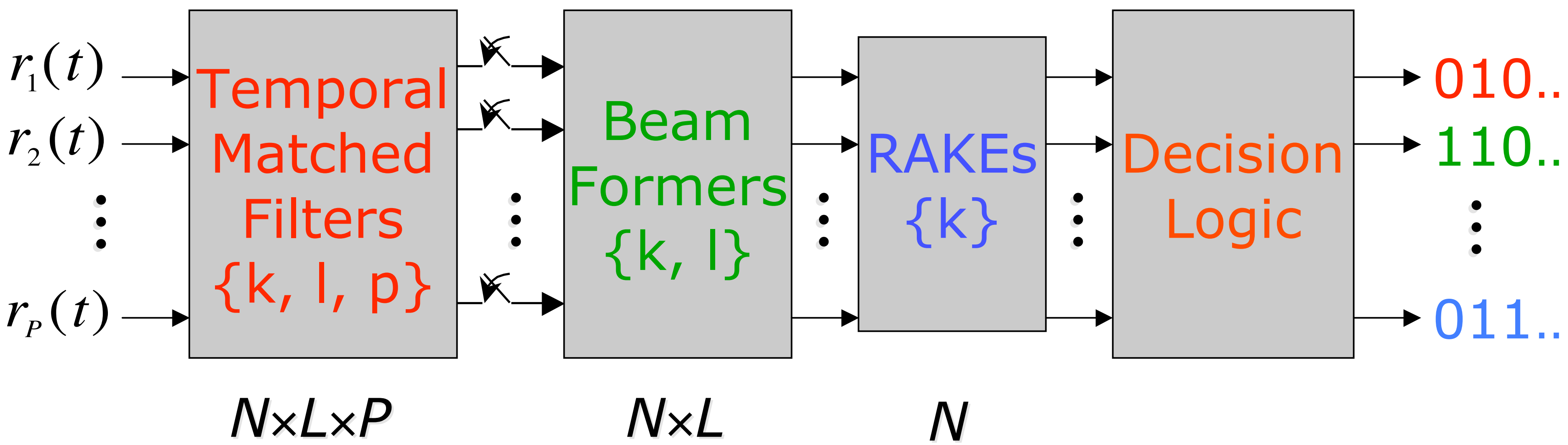
Multiuser Detection: receiver processing for shared-access systems

Multipath, Multi-antenna Case



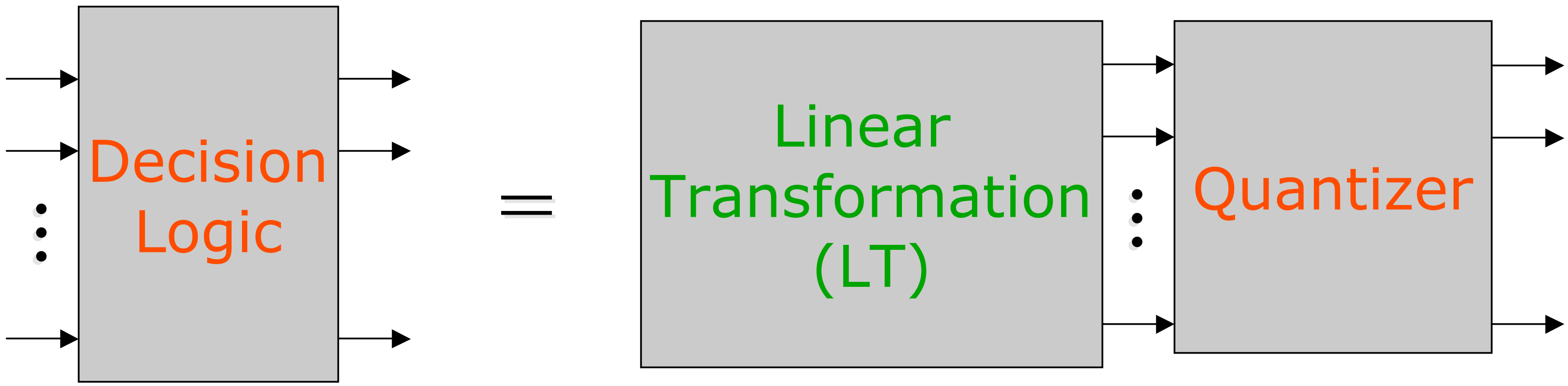
Space-Time MUD Structure

N Sensors; *P* Receive Antennas; *L* Paths/User/Antenna



- XISO (P=1) requires no beam-formers
- Flat fading (L=1)requires no RAKEs
- Decision logic: Optimal (ML, MAP), linear, iterative, adaptive.

Linear MUD



Key Examples:

- Matched Filter/RAKE Receiver: LT = identity
- Decorrelator: LT = channel inverter (i.e., zero-forcing)
- MMSE Detector: LT = MMSE estimate of the transmitted symbols

Game Theoretic Framework

[Meshkati, Poor, Schwartz, Mandayam, *IEEE Trans. COM*, Nov. 2005.]

Game: $G = [\{1,\ldots,N\}, \{A_k\}, \{u_k\}]$

N: total number of sensors

A_k: set of strategies for sensor k

u_k: utility function for sensor k

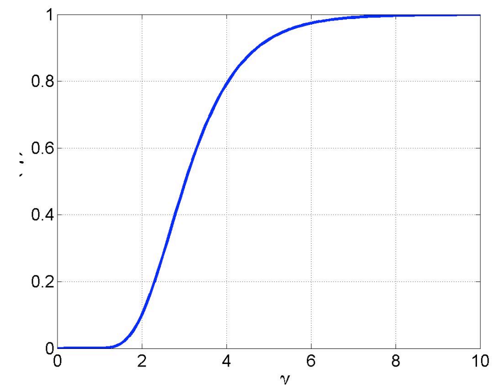
$$u_k = utility = \frac{throughput}{transmit\ power} = \frac{T_k}{p_k} \quad \left[\frac{bits}{Joule}\right]$$

$T_k = R_k f(\gamma_k)$, where $f(\gamma_k)$ is the frame success rate, and γ_k is the received SIR of sensor k.

An Uplink Game

- For a fixed linear MUD at the uplink receiver, each sensor selects its transmit power to maximize its own utility.
- Th'm: f sigmoidal $\Rightarrow$ Nash equilibrium (i.e., no user can unilaterally improve its utility) is reached when each sensor chooses a transmit power that achieves γ^*:

$$\boxed{f(\gamma^*) = \gamma^* f'(\gamma^*)}$$



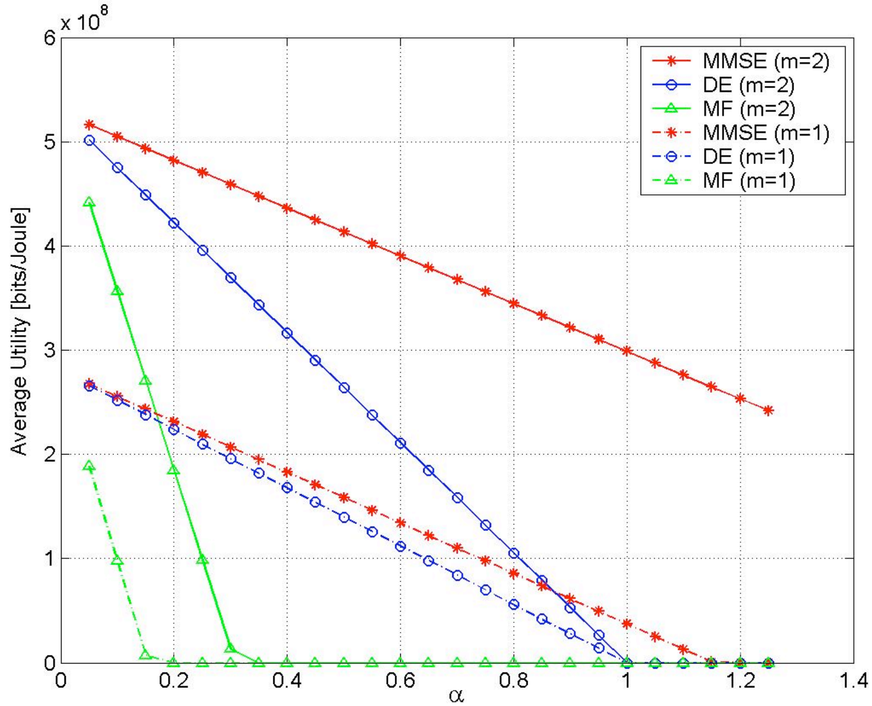
- I.e., Nash equilibrium (NE) requires SIR balancing.

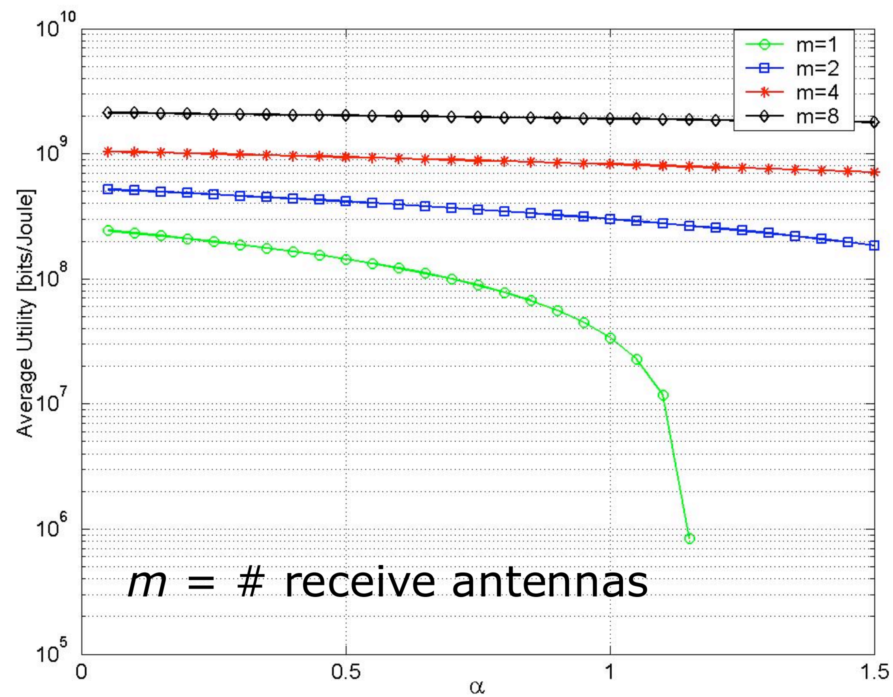
Remarks

- The NE is unique, and can be reached iteratively as the unique fixed point of a nonlinear map.

- Effects of Detector Choice:
  - We can use the NE to examine the effects of uplink receiver choice on energy efficiency.
  - Of interest are the classical matched filter, the (zero-forcing) decorrelator, and the MMSE detector.

Nash Equilibrium
Utility vs. Load (Large-System Limit)





- Random CDMA: *N* sensors; spreading gain *G*
- Load: $\alpha = N/G$ (i.e., the number of users per dimension)
- Large-system limit: $N, G \rightarrow \infty$, with α fixed.

Energy Efficiency in Shared Access Networks

Effects of Delay Constraints

[Meshkati, Poor, Schwartz, ISIT05.]

- For some messages (e.g., alarms), delay is important.
- Delay model (ARQ):
  - X represents the number of transmissions needed for a given packet to be received without error, so that:

$$\boxed{P(X=m) = f(\gamma)\,[1 - f(\gamma)]^{m-1}\ ,\ m = 0, 1, \ldots}$$

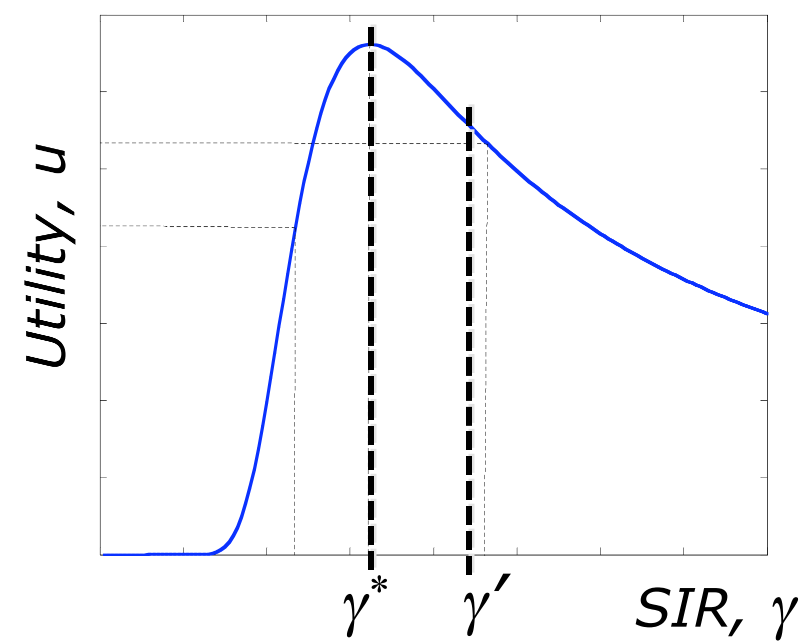
  - *We can represent a delay requirement as a pair (D,β)*:

$$\boxed{P(X \leq D) \geq \beta \Leftrightarrow \gamma \geq \gamma'}$$

  - *Thus, we have a constrained game, with $\gamma_k \geq \gamma_k'$.*

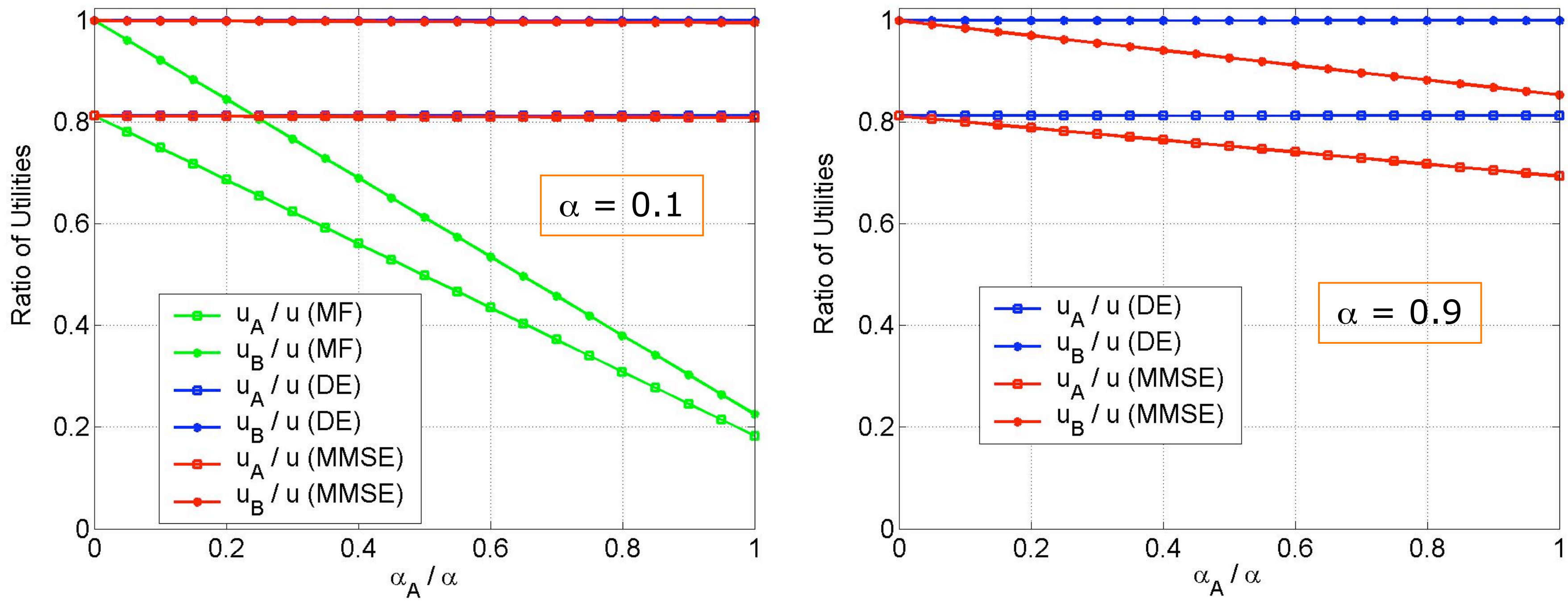
NE for Multiple Delay Classes

- Traffic is typically heterogeneous with multiple delay classes.
- A given delay class c will have its own SIR constraint: γ_c'
- At NE all sensors in class c will SIR-balance to $\max\{\gamma^*, \gamma_c'\}$.



- Tight delay constraints on one class can affect the energy efficiencies of all sensors due to increased interference levels.

2-Class Example: Utility Loss



- RCDMA in the large-system limit: *N, G* $\rightarrow\infty$, with $\alpha = N/G$ fixed.
- Class A: $(D_A, \beta_A) = (1, 0.99)$
- Class B: $(D_B, \beta_B) = (3, 0.90)$

Enhancements

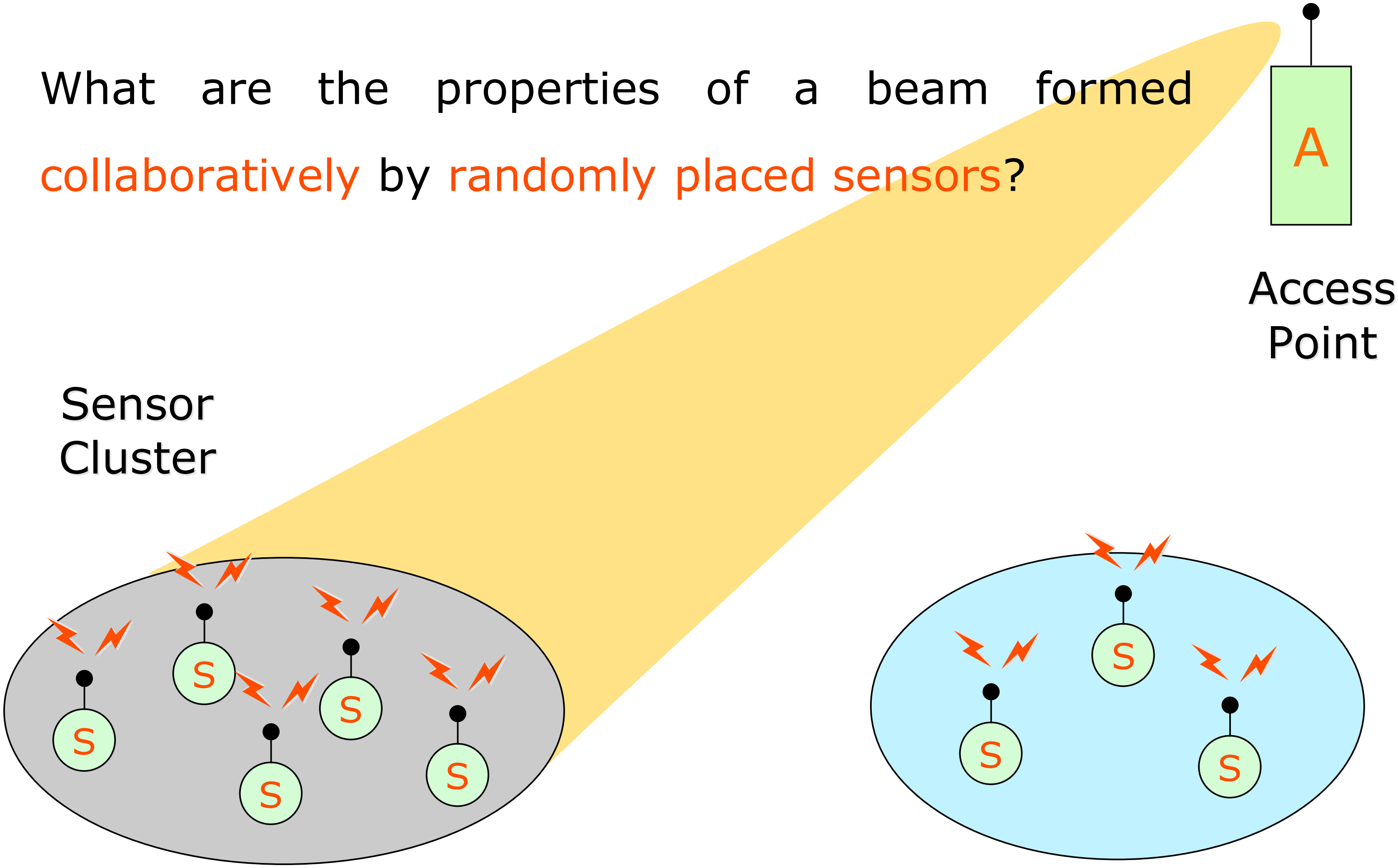
- Nonlinear MUD (ML, MAP, PIC, etc.): SIR-balancing also leads to a Nash equilibrium for certain nonlinear MUDs for RCDMA in the large system limit. [w/ D. Guo; Allerton'05]
- Multicarrier CDMA: Actions also include choice of a carrier; at NE (when it exists) each sensor transmits on its single, best, carrier + SIR balancing. [w/ M. Chiang; *JSAC'*06]
- Delay w/ Finite Backlog: Add queuing. [w/ R. Balan; CITIA Wkshp 06]
- Adaptive Modulation/Coding: Actions also include choice of a modulation. [w/ A. Goldsmith, et al., GLOBECOM 06, submitted]

COLLABORATIVE BEAMFORMING

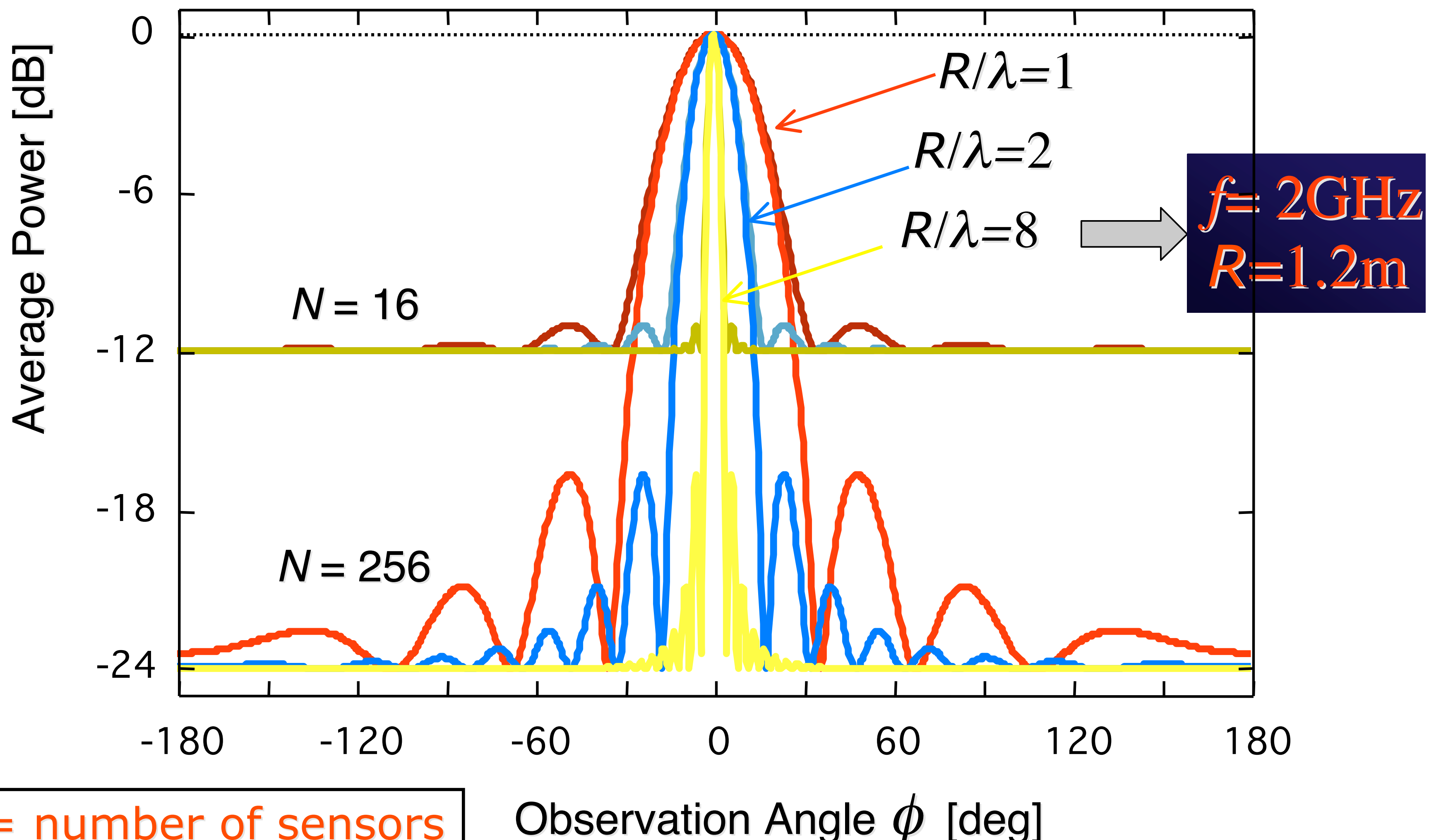
[Ochiai, Mitran, Poor, Tarokh, *IEEE Trans. SP*, Nov. 2005.]

Collaborative Beamforming

What are the properties of a beam formed collaboratively by randomly placed sensors?



Average Beampattern Example



N = number of sensors
R = radius of cluster
λ = wavelength

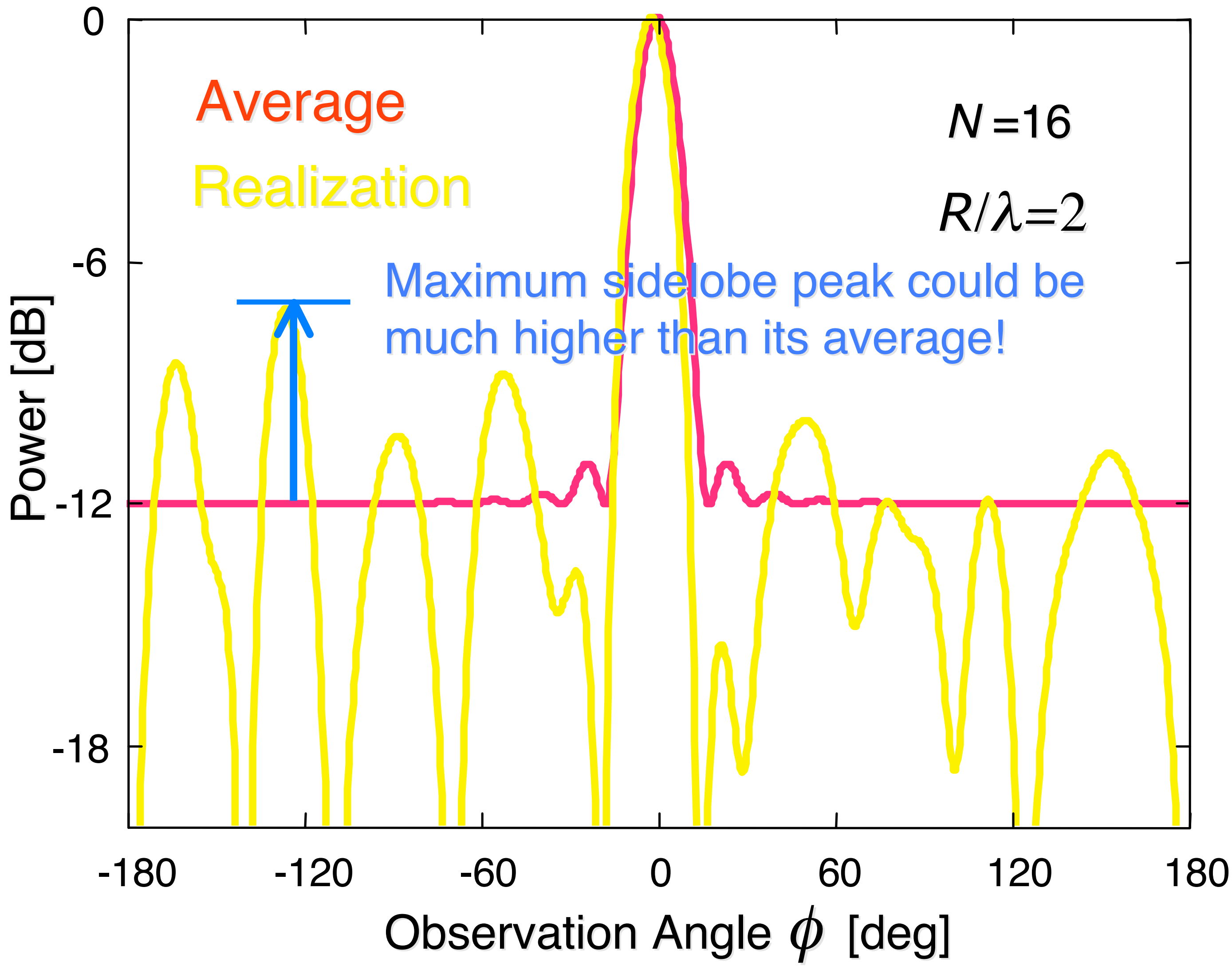
Average Beampattern Properties

- As cluster radius R becomes larger relative to wavelength λ, the main beam becomes sharper.
- Sidelobe level of average beampattern with N sensors is approximately $1/N$.
- Peak ave. sidelobe value does not depend on R/λ, but the peak location does.
- There are no grating lobes.

⇨ Average beam has nice properties. Life is good.

⇨ But, average doesn't represent the *realizations* of the sensor array!

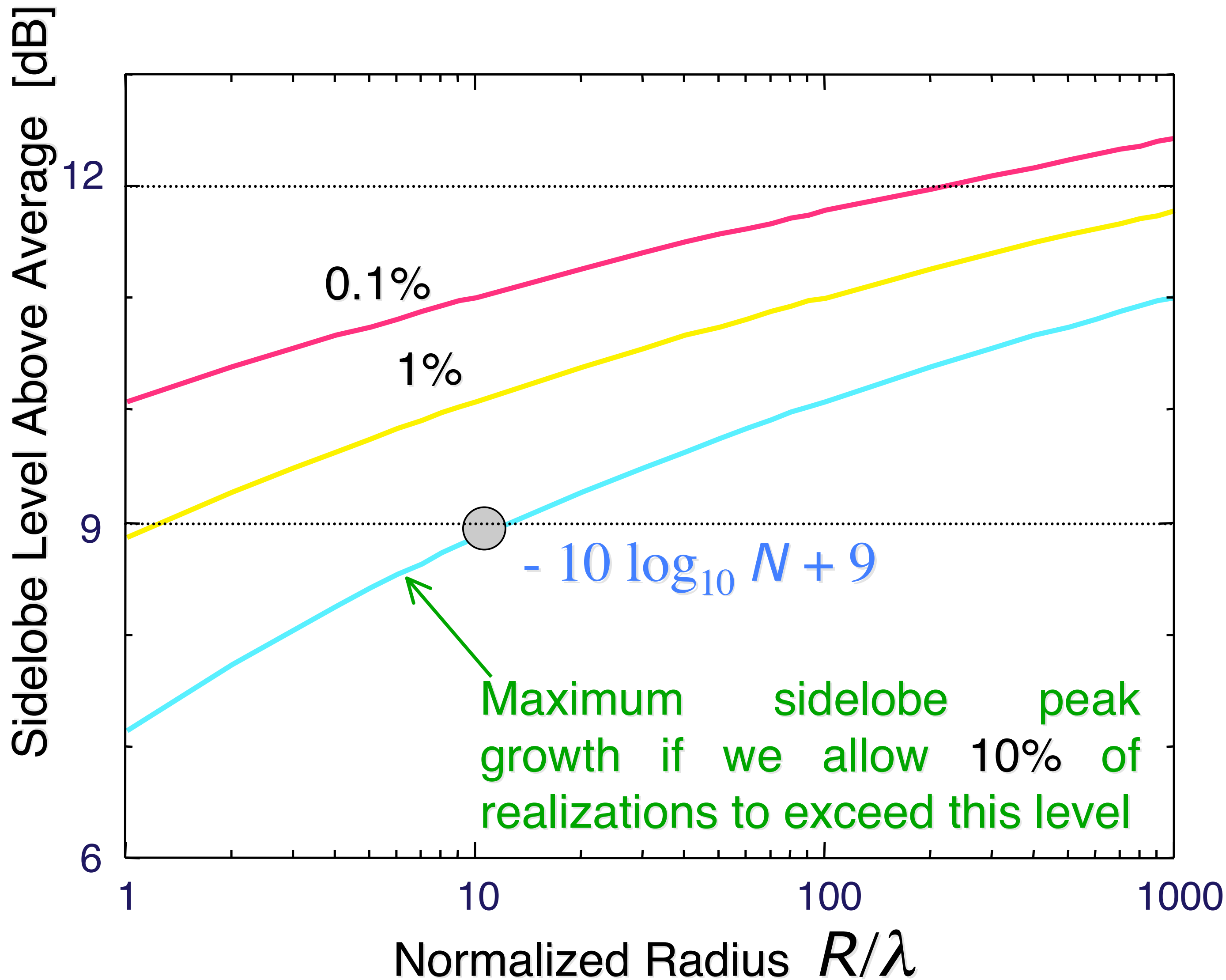
Ave. Beampattern vs. Realization



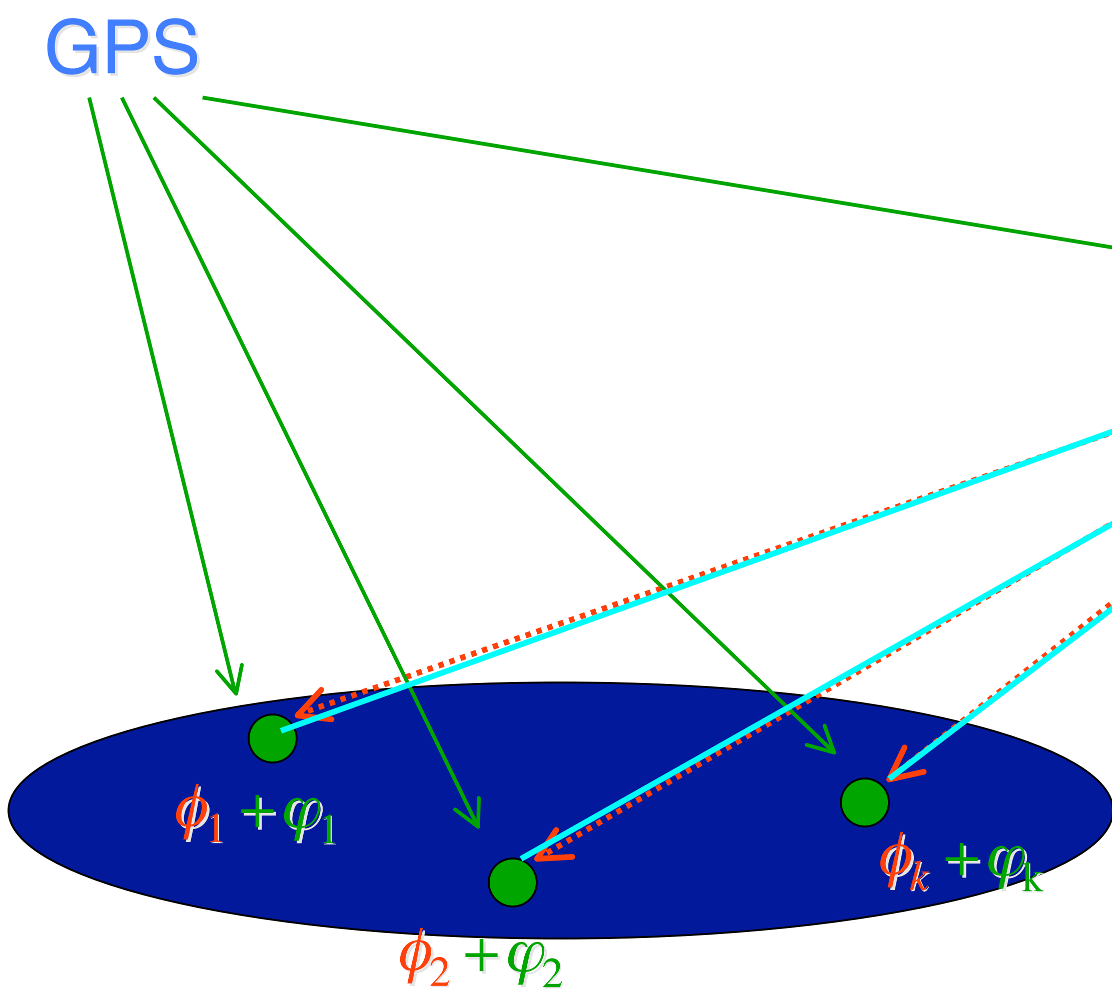
Distribution of Max Sidelobe Peak

- Maximum peak of sidelobe corresponds to worst-case interference.
- We use level-crossing theory to analyze this issue.
- For large *N* the beam is Rice-Nakagami in the sidelobes.
- Modeling sidelobes as a complex stationary Gaussian process, approximate upper bound on sidelobe distribution can be found.
- Simulations show good agreement.

Required Sidelobe Level Margin



Closed-Loop Phase Acquisition (Self-Phasing Arrays)

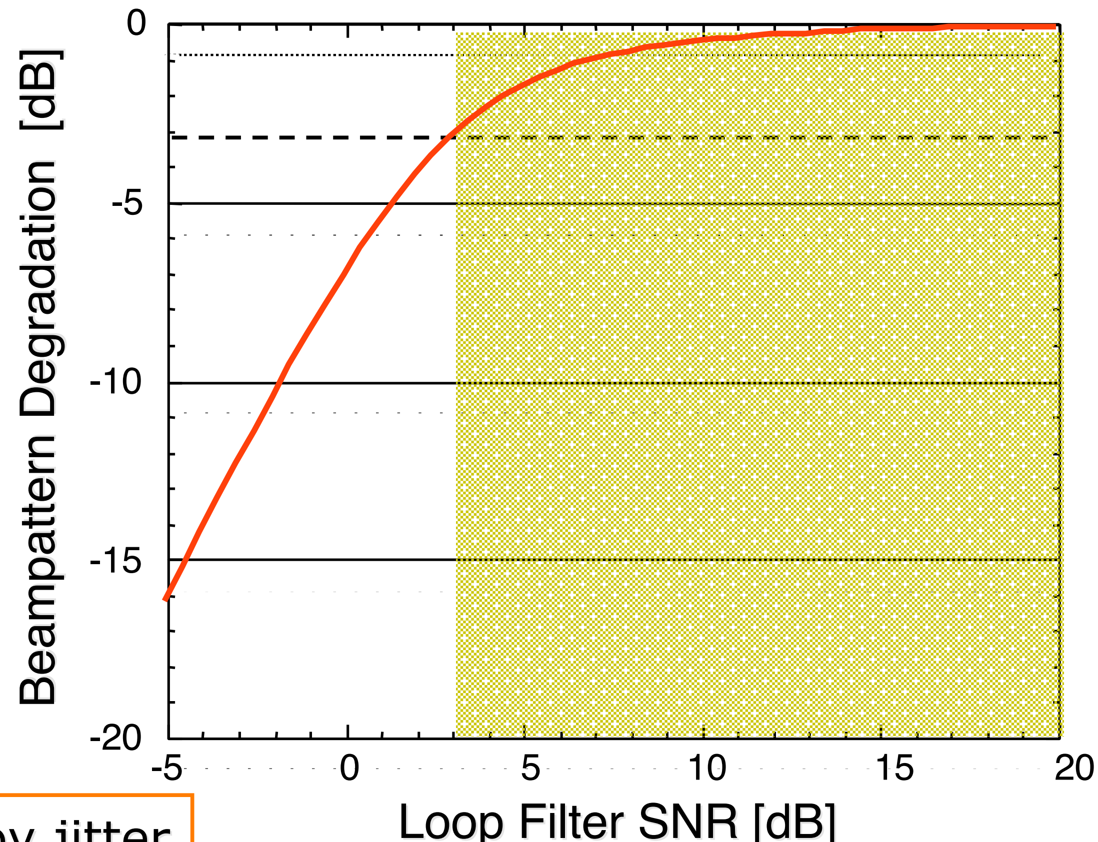


Each node autonomously estimates relative phase offset from pilot signal from target and absolute clock.

Beampattern with Phase Jitter

Average beampattern can be expressed as

$$P_{\mathrm{av}}(\phi) = \frac{1}{N} + \left(1 - \frac{1}{N}\right) \left| 2\frac{J_1(\alpha(\phi))}{\alpha(\phi)} \right|^2 \left| \frac{I_1(\rho)}{I_0(\rho)} \right|^2$$



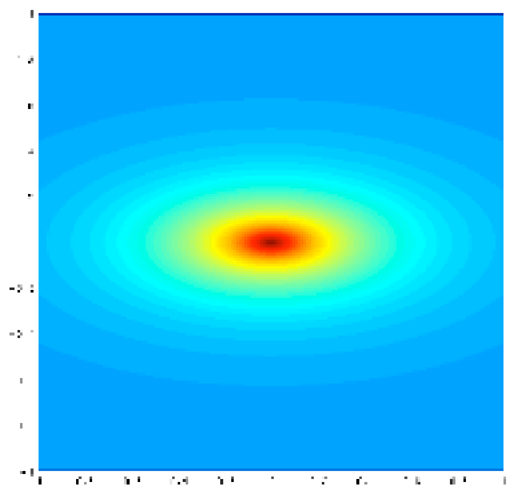
i.i.d.Tikhinov jitter

Collaborative Beamforming

ENERGY ISSUES IN DISTRIBUTED INFERENCE (BRIEFLY)

Energy-Efficient Sensor Scheduling

[Sung, Tong, Poor, *IEEE Trans. IT*, Apr. 2006]

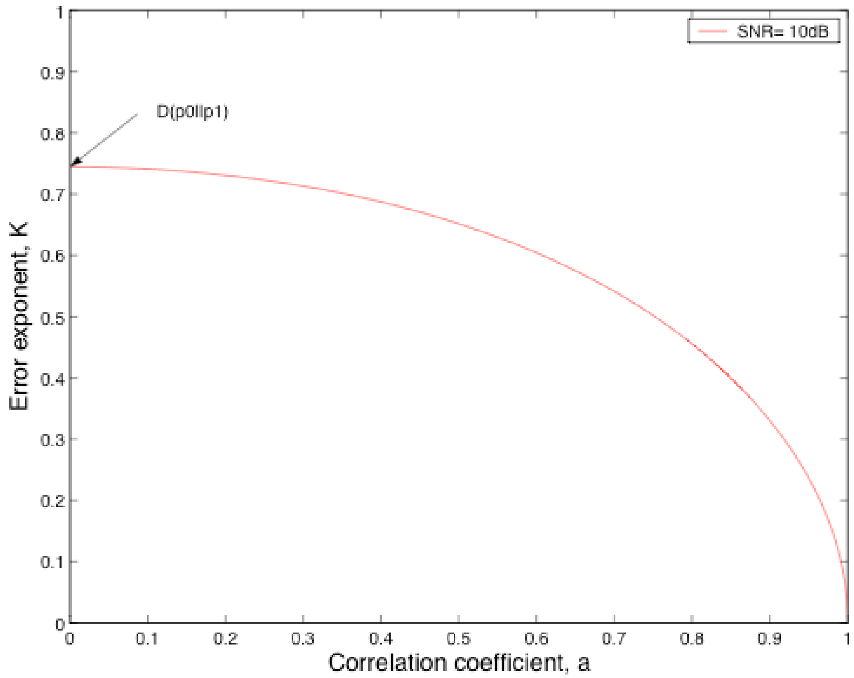


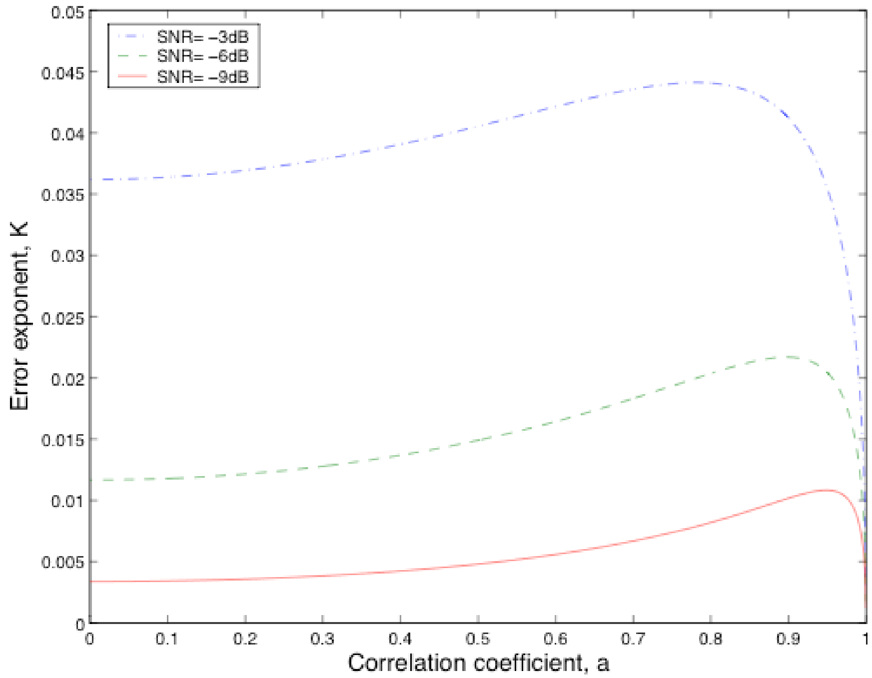
H_1: signal field + *noise* *H_0: noise only*

Neyman-Pearson performance of an N-sensor net measured via the error exponent, K, of the miss prob.:

$$K \sim -\log P_M(N)/N$$

Sensor Scheduling Via K



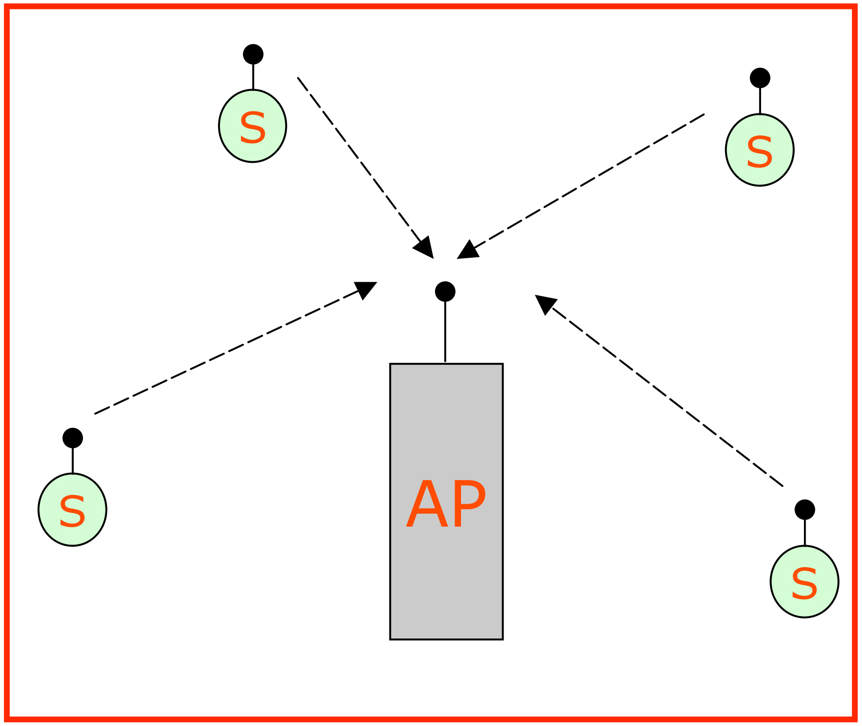


- *K* can be obtained in closed form using state-space model.
- Behavior w.r.t. correlation strength depends on SNR.
- This can be used to schedule sensors to transmit collaboratively for optimal energy efficiency.

Distributed Learning

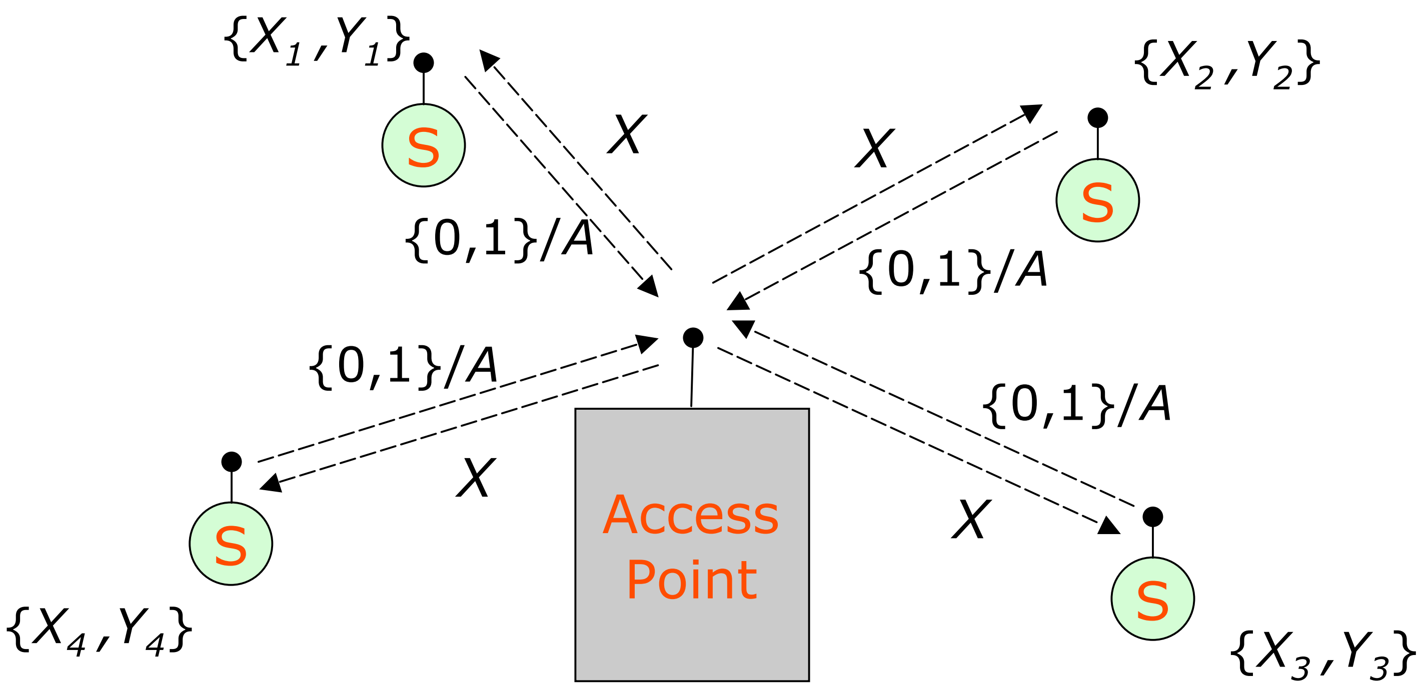
[Predd, Kulkarni, Poor, *IEEE Trans. IT,* Jan. 2006]

- Exemplars are distributed among the sensors in some way.
- Communications capacity between the sensors and fusion center is limited.



- Question: Can we construct algorithms so that optimal inferential functions can be learned consistently ($N \to \infty$) consuming little transmit power?

Distributed-Data Network



We can construct algorithms for the sensors and AP such that:

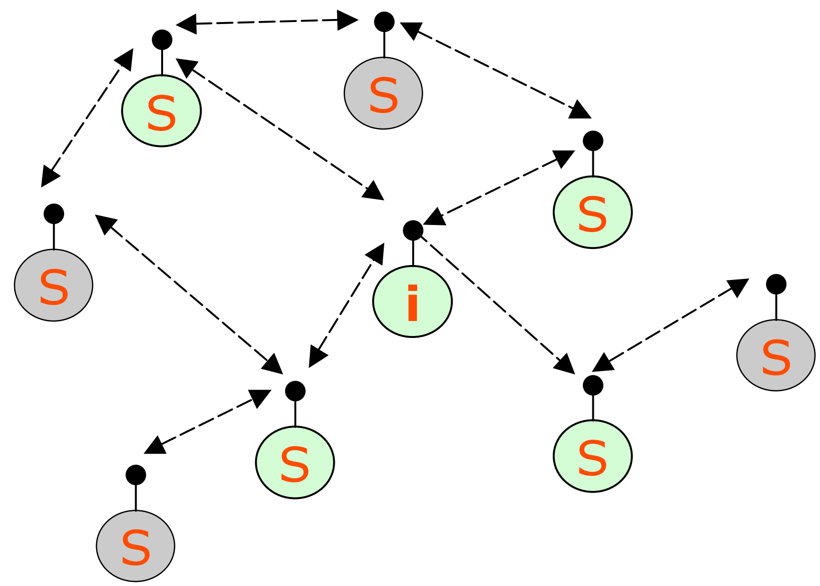
- Classification: transmitting 1 bit/sensor/decision is enough.
- Regression: transmitting $\log_2(3)$ bits/sensor/decision is enough

Collaborative Regression

- *N* sensors at locations $\{x_i\}$ take measurements:

$$y_i = f(x_i) + n_i$$

- Using message-passing-type algorithms, sensors can collaborate with their neighbors to estimate *f*.



E.g., An Algorithm

[Predd, Kulkarni, Poor, ITW06, Uruguay]

- To initialize, the sensors:
  - agree on a kernel $K(.,.)$
  - localize (i.e., estimate x_i)
  - share positions with neighbors
  - measure field locally (i.e. observe y_i)
  - set $z_i = y_i$
- To estimate the field:

  for $t=1,\ldots,T$

  for $s = 1,\ldots, N$

  Query: Sensor s queries z_i from neighbors

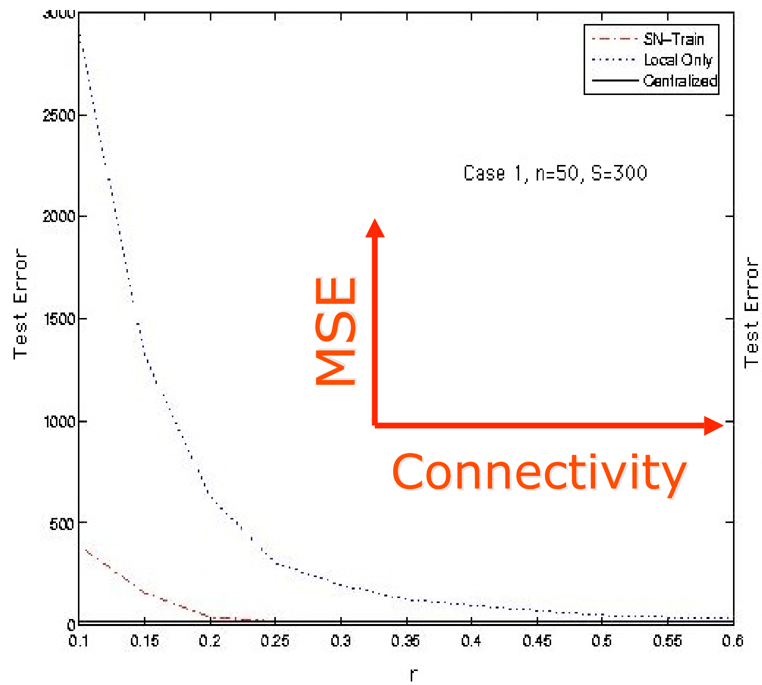
  Compute: $f_{s,t} = \arg\min_{f \in \mathcal{F}} \sum_{j \in N_s} (f(\mathbf{x}_j) - z_{j,t-1})^2 + \lambda_s \| f - f_{s,t-1} \|_{\mathcal{F}}^2$

  Update: Updates neighbors z_i **=** $f_{s,t}(x_i)$

Note: Converges to a relaxation of the centralized RKHS estimator.

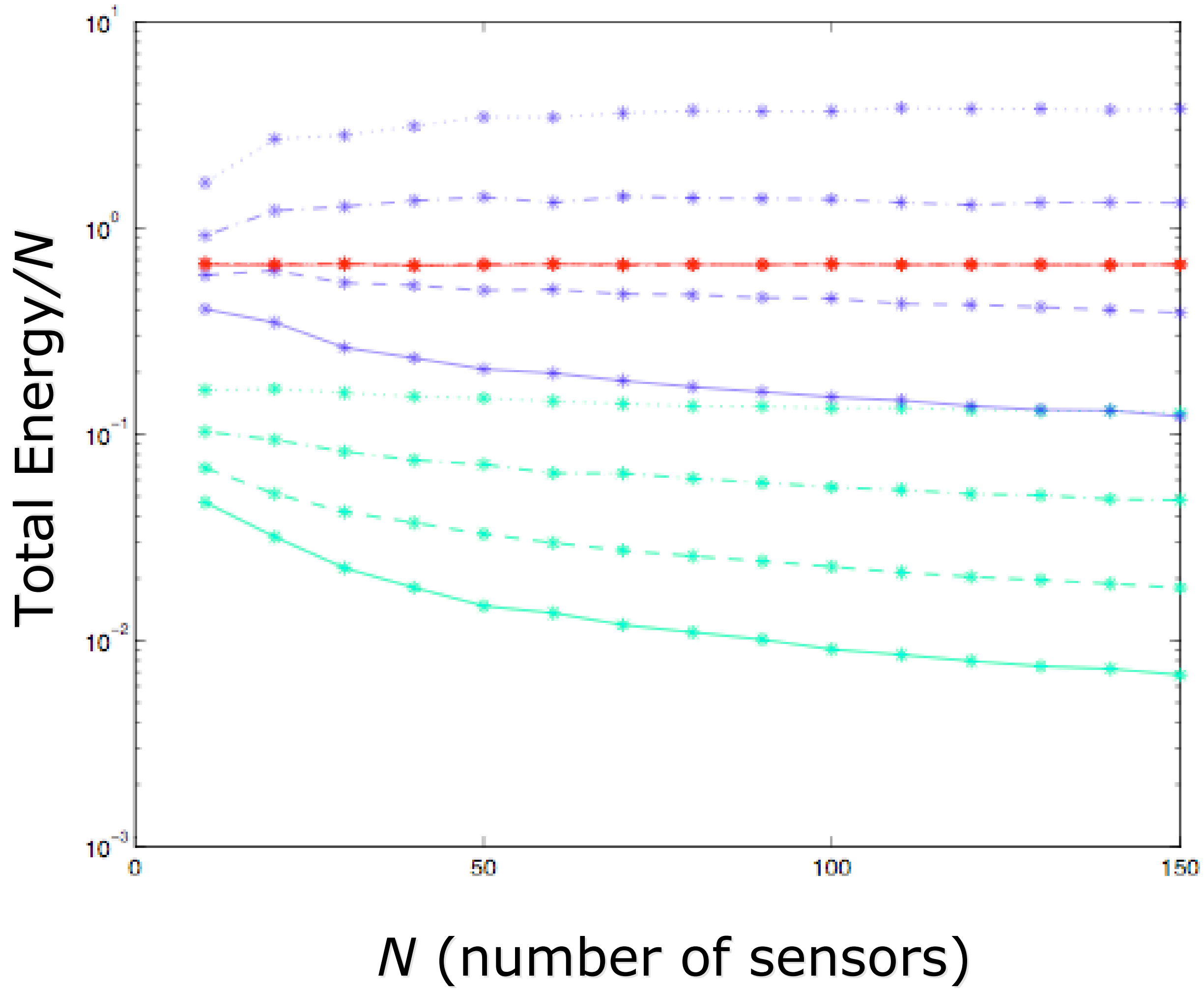
Energy Efficiency

- Overall error decreases with size of the neighborhoods.
- But, energy consumed by message-passing increases with neighborhood size.



- Question: What are the trade-offs?

Energy-per-Sensor vs. N



$r_N = N^\alpha$

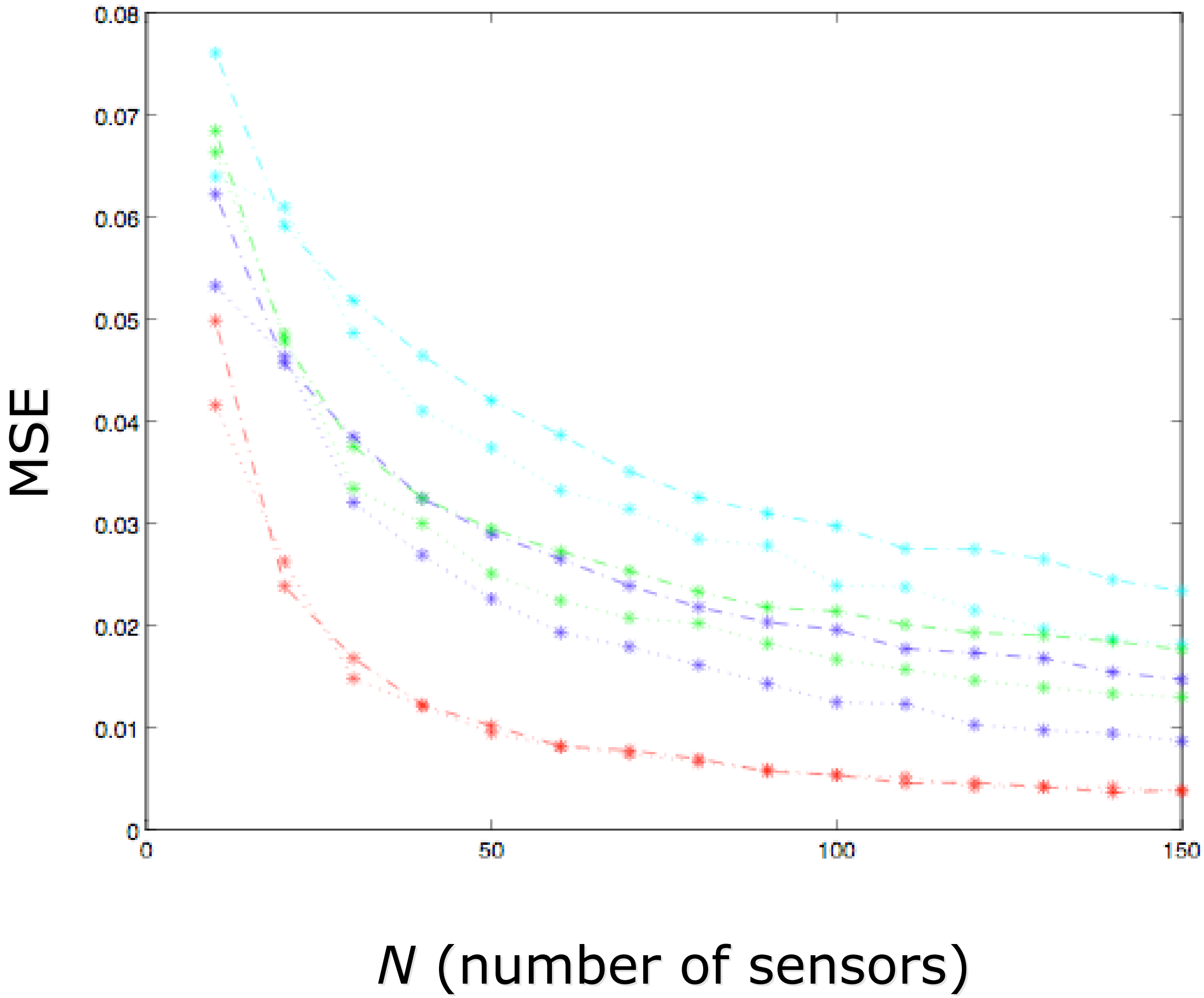
$\alpha = \{.30, .35, .40, .45\}$

PKP

Centralized

Local Averaging

Mean-Square Error vs.N



$r_N = N^{\alpha}$

α ={.30, .35, .40, .45}

PKP

Centralized

Local Averaging

Summary

- We've examined issues (primarily signal processing) affecting the energy efficiency of wireless networks:
  - Energy efficiency in shared access systems
  - Collaborative beamforming
  - Energy issues in distributed inference, briefly

Thank You!