

**Distributed Space-Time Processing**

**in Wireless Networks**

Babak Hassibi

*Department of Electrical Engineering*

*California Institute of Technology*

IEEE Workshop on Signal Processing Advances in Wireless Communications

Lisbon, Portugal, July 13, 2004

## Introduction

The past few years has seen a tremendous outburst of activity in the area of *space-time processing* for MIMO wireless communications systems. Multi-antenna systems can significantly

- increase the capacity, reliability and power efficiency of wireless communication systems

In this talk we shall discuss some of the implications of these results to wireless networks:

- the adjective *space-time* still applies since nodes in the network have different spatial locations and transmit at different times
- the key difference is the *distributed* nature of the problem
  - different nodes do not have access to the same information and/or measurements and their cooperation has a price and is over noisy channels themselves

## Wireless Networks

- wireless is a *shared* medium
  - network is a fully-connected graph
  - bandwidth is precious
- distinguishing features
  - interference
  - path-loss
  - fading

Traditional methods attempt to combat these “deficiencies”.

A better approach is to *exploit* them.

## Studying Wireless Networks is Difficult!

Determining various performance measures (capacity, power-efficiency, etc.) for wireless networks is related **Network Information Theory...**

- single-user information theory well-understood (Shannon)
- multi-user information theory, however, is not
  - computing the capacity of even a three-node network is open
  - situation reminiscent of mechanics where three-body problems have not been solved

Therefore most results focus either on the asymptotic regime of *large* networks (analogous to statistical mechanics where we have many bodies) or on the performance of very specific schemes

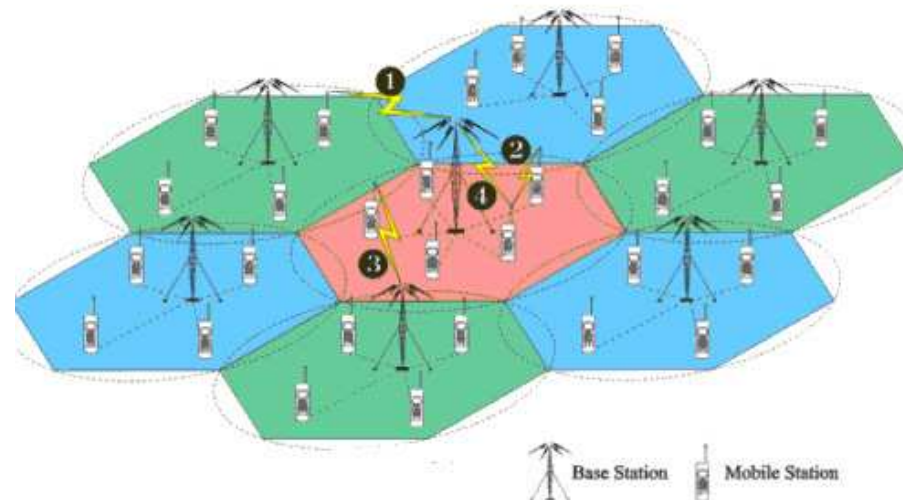
## Outline

We shall study three specific problems.

- **Cellular networks** (joint work with M. Sharif)
  - capacity of multi-antenna broadcast channels with partial CSI
- **Power efficiency in relay networks** (joint work with A. Dana)
  - distributed beam-forming and distributed interference cancellation
- **Diversity in relay networks** (joint work with Y. Jing)
  - distributed space-time coding

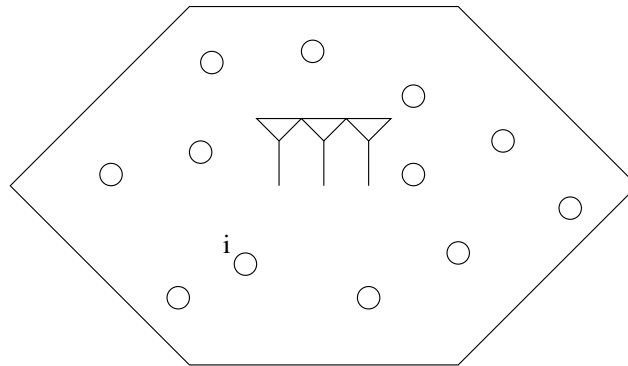
## Introduction to Broadcast Channels

- motivated by down-link scheduling in cellular systems
- to increase network capacity cell sizes are shrinking
- multiple transmit/receive antennas have proven useful in single link communications — what is their role in multi-user networks?



## Multiple Antennas in Broadcast Channels

Assume a system with  $n$  users, where the base station has  $M$  antennas and each user has  $N$  receive antennas.



For example, when  $N = 1$  the channel to the  $i$ -th user is given by an  $M$ -dimensional vector:

$$h_i = \begin{bmatrix} h_{i1} & h_{i2} & \dots & h_{iM} \end{bmatrix}$$

The  $h_i$  are independent random vectors, whose distribution depends on the fading environment.

For Rayleigh fading the entries of  $h_i$  are iid complex Gaussian.

**Question:** *What is the capacity of such a multi-antenna broadcast channel?*

Answer depends on what the transmitter knows about the channels.

To warm up let's first look at the multi-antenna case...

## Multiple Antennas in Single User Links

Consider a single user link with  $M$  transmit and  $N$  receive antennas. Assume that the environment is rich-scattering.

- Transmitter and receiver know the channel:

$$C = \min(M, N) \log \text{SNR} + O(1)$$

- Only receiver knows the channel (Foschini '97, Telatar '97):

$$C = \min(M, N) \log \text{SNR} + O(1)$$

- Neither receiver nor transmitter knows the channel (Zheng and Tse '01, Hassibi and Marzetta '01):

$$C = \min(M, N) \left( 1 - \frac{\min(M, N)}{T} \right) \log \text{SNR} + O(1)$$

where  $T$  is the coherence interval of the channel

## Broadcast Channels with full CSI in Tx/Rx

Capacity with full CSI in the transmitter (Caire and Shamai '03, Viswanath and Tse '03, Goldsmith et al. '03):

$$C^{DP} = E \left\{ \max_{\{P_1, \dots, P_n, \sum P_i = P\}} \log \det \left( I + \sum_{i=1}^n h_i^* P_i h_i \right) \right\}.$$

- Achieved by what is known as *dirty paper coding* (Costa '80).
- We have analyzed this for large number of users, fixed  $M$  and  $N$ , and shown that

$$\lim_{n \rightarrow \infty} \frac{C^{DP}}{M \log \log nN} = 1$$

Dirty paper coding presents two problems:

- It requires full channel knowledge at the transmitter and can be sensitive to channel errors
- It is very computationally intensive (VQ at both Tx/Rx)

## Capacity of MIMO BC with no CSI in Tx

What if we assume no CSI in Tx?

- It is straightforward to show that the Gaussian MIMO BC with no CSI in the transmitter is *degraded* no matter whether the receivers have CSI or not (Amraoui et al '03, Sharif and Hassibi '03).
- For degraded channels the capacity is known and in this case can be shown to be

$$C = \log M + O(1)$$

independent of  $n$ ! (Assuming a fixed transmit power per antenna.)

Thus, when  $N = 1$ , lack of knowledge of the channel coefficients brings us down from  $M \log \log n$  to  $\log M$ .

What to do?

- is there any critical side information that can fill this gap?

## Send Random Beams

Choose  $M$  random orthonormal vectors  $\phi_m$ ,  $m = 1, \dots, M$  (according to an isotropic distribution). At time  $t$ , the  $m$ -th vector is multiplied by the signal  $s_m(t)$  so that the transmitted signal is

$$S(t) = \sum_{m=1}^M \phi_m s_m(t), \quad t = 1, \dots, T$$

where  $T$  is less than the coherence interval of the channel.

After  $T$  channel uses we independently choose another isotropic set of orthonormal vectors  $\{\phi_m\}$ , and so on.

In other words, we are transmitting  $M$  random beams.

Each receiver  $i = 1, \dots, n$  therefore can compute the following  $M$  SINRs

$$\text{SINR}_{im} = \frac{|h_i \phi_m|^2}{1/\text{SNR} + \sum_{n \neq m} |h_i \phi_n|^2}$$

Of course, on average the SINR is (roughly)  $\frac{1}{1/\text{SNR} + M - 1}$ , and so if we randomly assign beams to users we get

$$C \approx M \log \left( 1 + \frac{1}{1/\text{SNR} + M - 1} \right) < \frac{M}{1/\text{SNR} + M - 1},$$

which is pretty lousy.

So what is the point?

## Exploit Multi-User Diversity

Suppose now each user (or, in fact, only those who get favorable SINRs) **feeds back to the transmitter its best SINR.**

Rather than randomly assign the beams, the transmitter assigns signal  $s_m$  to the user with the best SINR for that signal. Therefore

$$C = E \sum_{m=1}^M \log \left( 1 + \max_{i=1, \dots, n} \frac{|h_i \phi_m|^2}{1/\text{SNR} + \sum_{n \neq m} |h_i \phi_n|^2} \right)$$

Due to the symmetry of all the random variables involved:

$$C = ME \log \left( 1 + \max_{i=1, \dots, n} \frac{|h_i \phi_1|^2}{1/\text{SNR} + \sum_{n \neq 1} |h_i \phi_n|^2} \right)$$

Note that the random variables

$$\frac{|h_i \phi_1|^2}{1/\text{SNR} + \sum_{n \neq 1} |h_i \phi_n|^2}$$

are ratios of independent Gaussians, and themselves are independent.

## Maximum of $n$ IID Random Variables

**Theorem 1 (Frechet, 1927, v. Mises 1947, Uzgoren 1954)** *Let  $x_i, i = 1, \dots, n$  be iid random variables with distribution  $p(x)$ . Then if,*

$$\lim_{x \rightarrow \infty} -\frac{p(x)}{p'(x)} = c > 0,$$

*then*

$$\text{Prob} \left( \frac{|x_{\max} - \frac{1}{c} \log n|}{\log \log n} > \epsilon \right) < \frac{\delta}{(\log n)^\alpha}.$$

In our case, the condition of the theorem is met with  $c = 1/\text{SNR}$ .

Therefore, for  $M$  fixed and  $n \rightarrow \infty$

$$C = M \log \log n + O(\log \log \log n)$$

**Thus, multi-user diversity buys us a lot!**

## Some Remarks

- If we have  $N$  receive antennas, it is easy to show that the same scheme achieves  $M \log \log nN$ , the same asymptotic capacity as the full CSI case
- Compared to the full CSI case where each user had to feedback  $2M$  complex numbers, here each user need only feed back its *best* SINR and the corresponding index.
  - In fact, by choosing a suitable threshold, the feedback can be reduced to a single bit

## How Large can $M$ Be?

Our result required that  $M$  be fixed and  $n \rightarrow \infty$ .

In this case,  $\text{SINR}_{\max}$  was obtained when  $\sum_{j \neq m} |h_j \phi_m|^2 \approx 0$  and  $|h_i \phi_m|^2 \approx \log n$ .

However, in practice, we will have a large (but finite  $n$ ) and so it is useful to know how large  $M$  can be to retain a linear increase in capacity.

**Result:** Let  $M = \alpha \log n$ . Then as  $n \rightarrow \infty$ , we have  $\text{SINR}_{\max} \rightarrow 1/\alpha$  and

$$C = M \log \left( 1 + \frac{1}{\alpha} \right) + O(1)$$

Moreover, if  $\frac{M}{\log n} \rightarrow \infty$ , then  $\frac{C}{M} \rightarrow 0$ .

# Sum Rate Throughput of Random Beamforming

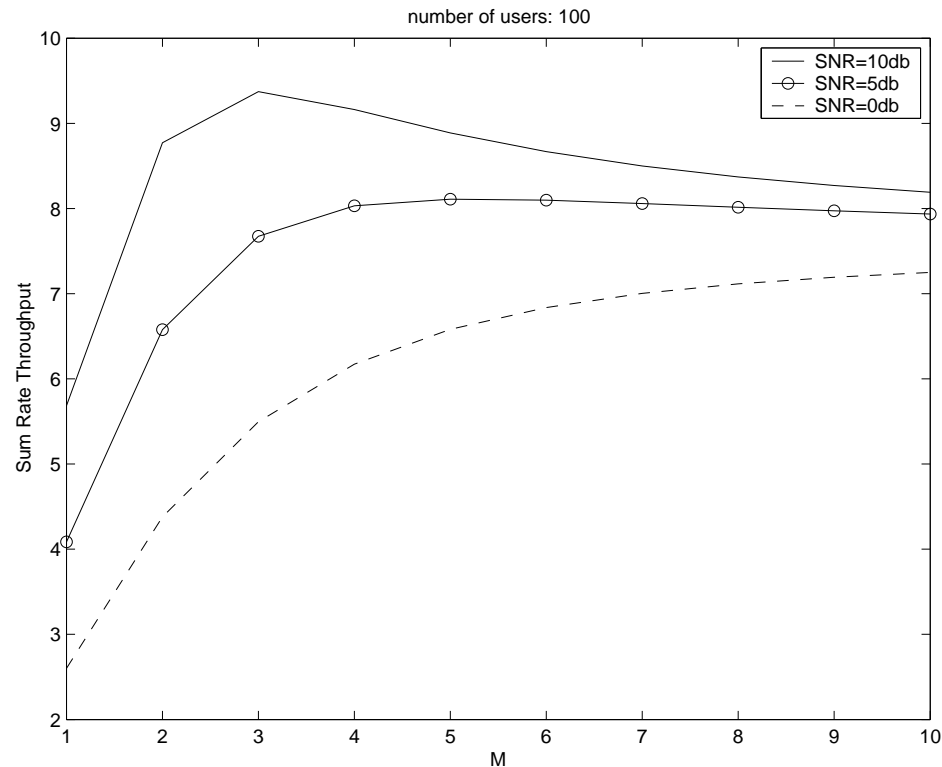


Figure 1: Sum rate throughput for a system with  $n = 100$  users and for different number of transmit antennas

## Fairness of the System

- So far we have assumed that the network is homogenous
- Most networks are heterogenous, in the sense that the SNRs for the different users are different
- Thus, if we transmit to the most favorable users, the system may be dominated by the users with the highest SNR
- This is certainly true for single-transmit-antenna systems

# Single Antenna Multi-User Fairness

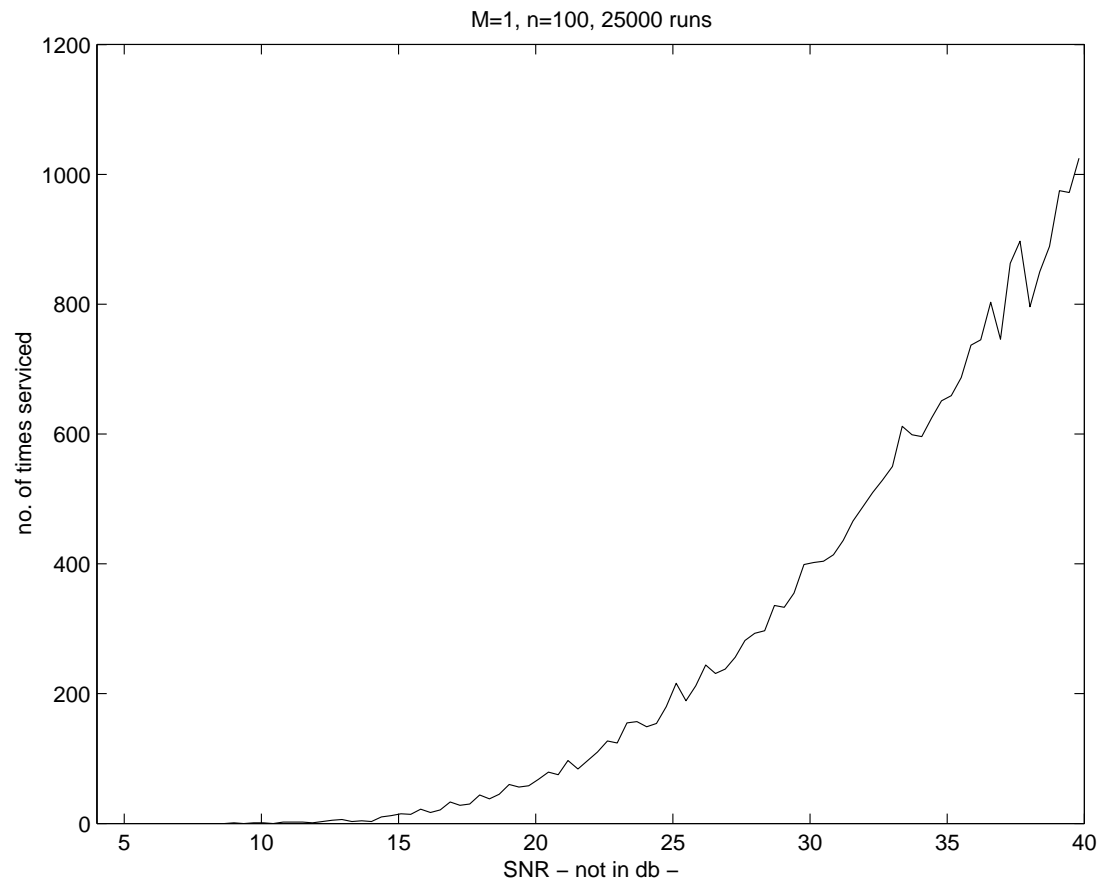


Figure 2:  $M = 1$ ,  $n = 100$ , SNR 4-40 (6-16db), 25000 runs

## Fairness - Continued

In a heterogenous network, the SINRs now become

$$\text{SINR}_{im} = \frac{|h_i \phi_m|^2}{1/\text{SNR}_i + \sum_{n \neq m} |h_i \phi_n|^2}$$

Note that, if the  $\text{SNR}_i$ s are high enough, or if the number of transmit antennas is large enough, then the system is interference-dominated which implies *built-in fairness*.

**Result:**

$$P(\text{choosing user with } \text{SNR}_{min}) \geq \frac{1}{n} e^{-\left(\frac{1}{\text{SNR}_{min}} - \frac{1}{\text{SNR}_{max}}\right) \left(e^{\frac{\log n}{M}} - 1\right)},$$

This further illustrates the benefits of having  $M = \alpha \log n$ .

# Multi-Antenna Multi-User Fairness

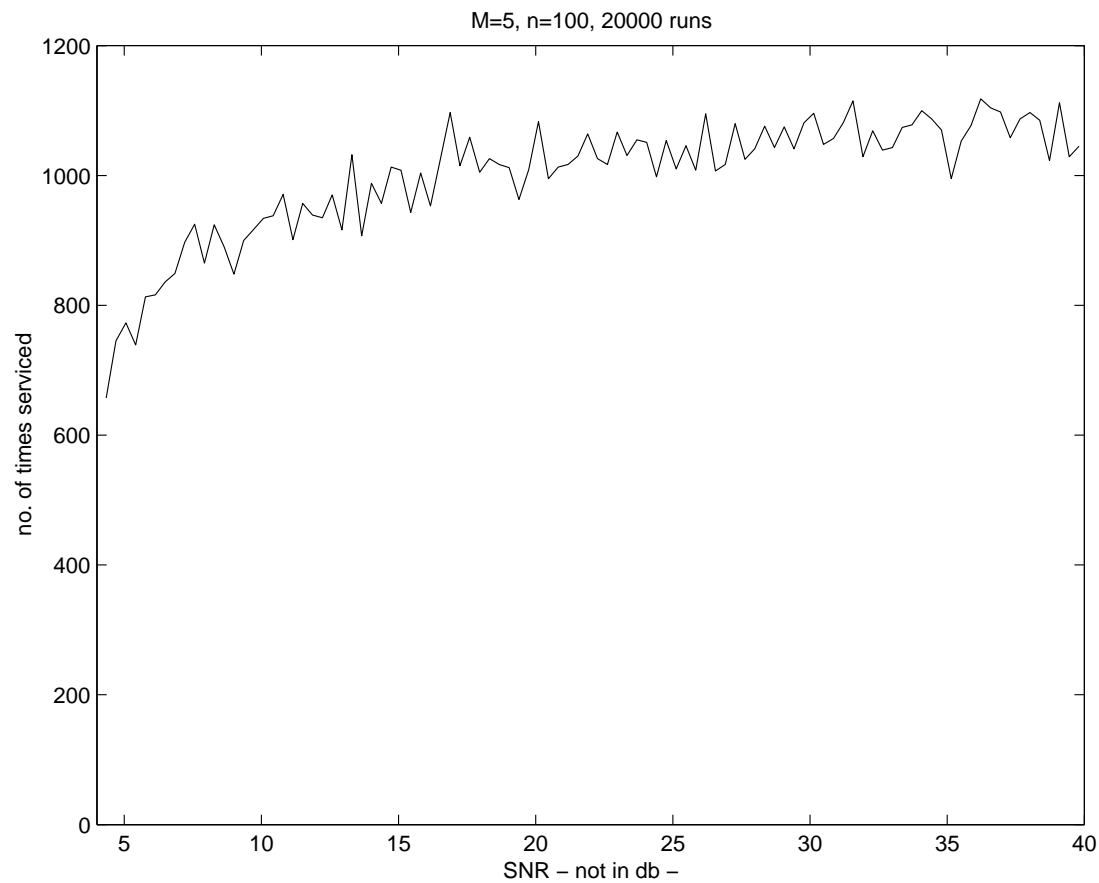


Figure 3:  $M = 5$ ,  $n = 100$ , SNR 4-40 (6-16db), 20,000 runs

Let us now turn to the second part of the talk...

# Sensory Networks and Ad-Hoc Networks

## Sensory Networks:

- applications: environmental, surveillance, military, etc.
- sensor nodes are simple, cheap and consume very little power
- all information intended for a single receiver
- sensors may only need to occasionally transmit information

## Ad-Hoc Networks:

- no infrastructure
- users “somehow” cooperate in a distributed fashion to communicate
- can have many users acting simultaneously as transmitters or receivers

## Sensory and Ad-Hoc Networks: Scaling Laws

How does the capacity of such networks scale in  $n$ , the number of nodes?

- for sensory networks capacity scales as  $O(\log n)$  (Gastpar and Vetterli, 2002)
- for ad-hoc networks capacity scales as  $O(\sqrt{n})$  (Gupta and Kumar, 2000)

Both are overall discouraging results:

- in sensory networks, the capacity per participating node is  $O(\frac{\log n}{n})$
- in ad-hoc networks, the capacity per participating node is  $O(\frac{1}{\sqrt{n}})$

In either case, these are diminishing returns for the nodes' energy investment and participation in the communications

## Power Efficiency

- in wireless networks, especially sensory networks, power consumption is a major bottleneck

Following Verdu (2002) , we shall look at *power efficiency* defined as

$$\eta = \frac{\text{capacity}}{\text{transmit power}} = \frac{\log(1 + \sigma_s^2/\sigma_n^2)}{\sigma_s^2}$$

- as  $\sigma_s^2 \rightarrow \infty$ :

$$\eta \rightarrow \frac{\log \sigma_s^2}{\sigma_s^2} \rightarrow 0,$$

- as  $\sigma_s^2 \rightarrow 0$ :

$$\eta = \frac{\log(1 + \sigma_s^2/\sigma_n^2)}{\sigma_s^2} \rightarrow \frac{\log e}{\sigma_n^2} = O(1),$$

which means we are power efficient at low SNR. *But what happens in a network?*

## Multi-Antenna Power Efficiency

To gain some insight, consider an  $n$ -transmit single-receive multi-antenna channel, with channel matrix:

$$H = \begin{bmatrix} h_1 & h_2 & \dots & h_n \end{bmatrix}, \quad E|h_i|^2 = 1$$

- if the channel matrix is *known* to the transmitter (beamforming):

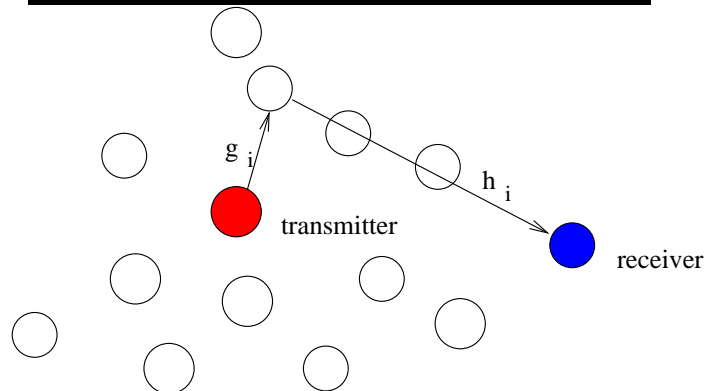
$$\eta = \frac{E \log \left( 1 + \frac{\sigma_s^2}{\sigma_n^2} \left( \sum_{i=1}^n |h_i|^2 \right)^2 \right)}{n\sigma_s^2} \rightarrow \frac{n \log e}{\sigma_n^2} = O(n)$$

- if the channel matrix *unknown* to the transmitter (Foschini-Telatar):

$$\eta = \frac{E \log \left( 1 + \frac{\sigma_s^2}{\sigma_n^2} \sum_{i=1}^n |h_i|^2 \right)}{n\sigma_s^2} \rightarrow \frac{\log e}{\sigma_n^2} = O(1)$$

From a power efficiency point of view, it only pays off to have multiple transmit antennas if the channel is known at the transmitter.

# Sensory Networks



- gains from transmitter to each node:  $g_i, \quad i = 1, \dots, n$
- gains from each node to receiver:  $h_i, \quad i = 1, \dots, n$
- **Assumptions:**
  - $\{g_i, h_i\}$  are subject to independent fading. Averaged over the fading and the point placement of the nodes, they are unit-variance with finite fourth-order moment
  - system is synchronous
  - each node  $i$  knows its local connections,  $g_i$  and  $h_i$ , but not the rest of the network

## A “Listen and Transmit” Protocol

Communication is divided into two intervals:

1. The *listen* interval: the transmitter transmits the signal  $s$ , with power  $p$ , and all other nodes are silent, but listen:

$$q_i = g_i s + v_i, \quad i = 1, \dots, n$$

2. The *transmit* interval: each node transmits with power  $\sigma_s^2$  the signal

$$t_i = \frac{\sigma_s}{\sqrt{\sigma_n^2 + p}} g_i^* h_i^* q_i \quad i = 1, \dots, n$$

The received signal is

$$\begin{aligned} r = \sum_{i=1}^n h_i t_i + w &= \sum_{i=1}^n |h_i|^2 \frac{\sigma_s}{\sqrt{\sigma_n^2 + p}} g_i^* q_i + w \\ &= \sum_{i=1}^n \frac{\sigma_s}{\sqrt{\sigma_n^2 + p}} (|h_i g_i|^2 s + |h_i|^2 g_i^* v_i) + w \end{aligned}$$

Thus the signal adds up coherently and the noise non-coherently:

$$\text{SNR} = \frac{\frac{n^2 \sigma_s^2}{\sigma_n^2 + p}}{\sigma_n^2 \left(1 + \frac{n \sigma_s^2}{\sigma_n^2 + p}\right)} = \frac{n^2 \sigma_s^2 p}{\sigma_n^2 (\sigma_n^2 + p + n \sigma_s^2)}$$

Assume  $p = O(n^{-\epsilon})$  and  $n \sigma_s^2 = O(n^{-\epsilon})$  for some  $\epsilon > 0$ . Then

$$\text{SNR} = \frac{n(n \sigma_s^2)p}{\sigma_n^4} = O(n^{1-2\epsilon}) \quad \text{and} \quad \eta = O\left(\frac{n^{1-2\epsilon}}{n^{-\epsilon}}\right) = O(n^{1-\epsilon})$$

Therefore  $\epsilon > 1/2$  and so the best we can do is:

$$\eta = O(\sqrt{n})$$

## Precise Statement

The above argument was very rough. Here is the precise statement.

**Theorem 2** *Consider a fixed area  $A$  and randomly select  $n$  nodes and a transmit/receive pair in this area. Then*

$$\text{Prob}(\eta > k_1 \sqrt{n}) > 1 - \frac{k_2}{n},$$

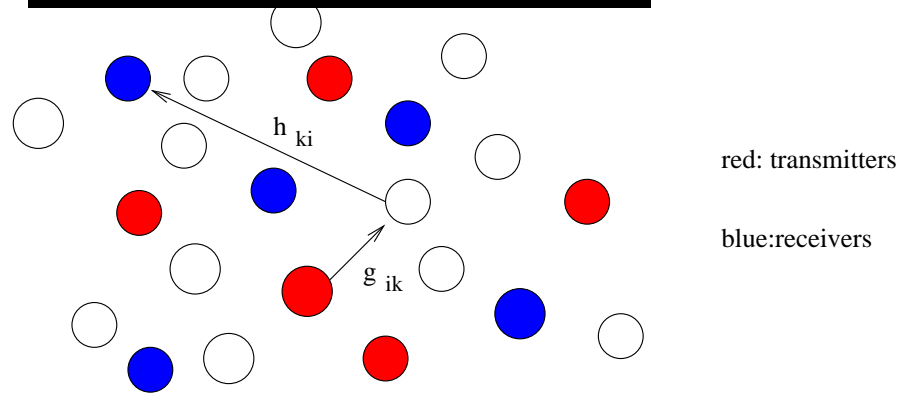
*where  $k_1$  and  $k_2$  are positive constants independent of  $n$ , but that depend on  $A$  and the fading statistics.*

## Power Efficiency of Sensory Networks

- the power efficiency is thus  $O(\sqrt{n})$
- compared to a single link, for a fixed transmission rate, the total power consumption in the network can be reduced by a factor of  $\frac{1}{\sqrt{n}}$
- this is even better than the  $n \times 1$  multi-antenna channel with no channel state information ( $O(1)$ ), though not as good as the  $n \times 1$  case with channel state information ( $O(n)$ )
- the protocol is double-hop and exploits, rather than avoids, interference (it essentially does distributed beam-forming)
- there is built-in fairness: the transmitter transmits with power  $p = O(\frac{1}{\sqrt{n}})$  and all other nodes with power  $\sigma_s^2 = O(\frac{1}{n\sqrt{n}})$

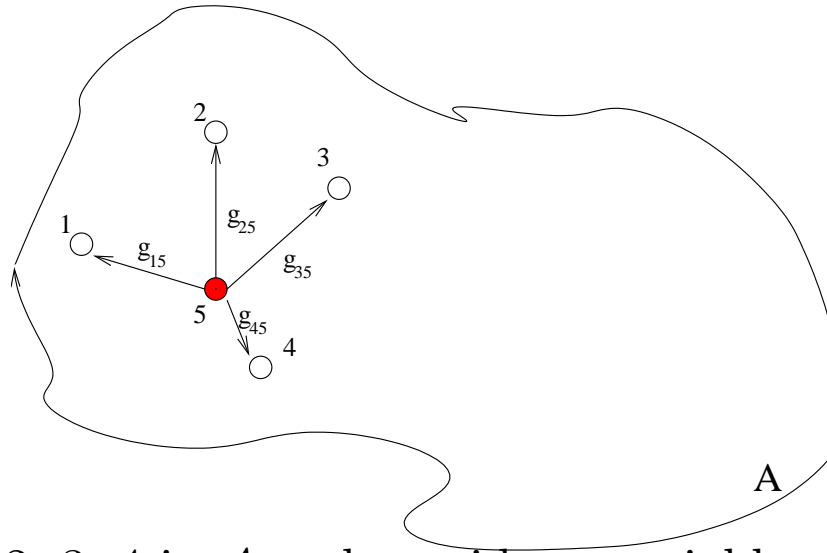
Thus, from a power efficiency point of view, it pays off to network!

# Ad-Hoc Networks



- $n$  nodes,  $r \leq \sqrt{n}$  users transmitting,  $r$  users receiving
- gains from transmitters to each node:  $g_{ik}$ ,  $i = 1, \dots, r$   $k = 1, \dots, n$
- gains from each node to receivers:  $h_{ki}$ ,  $k = 1, \dots, n$   $i = 1, \dots, r$
- **Assumptions:**
  - $\{g_{ik}, h_{ki}\}$  subject to independent fading. Averaged over fading and point placement, unit-variance, finite fourth-order moment
  - system synchronous
  - each node  $k$  knows its local connections,  $g_{ik}$  and  $h_{kj}$ , but not the rest of the network

## Technical Assumption



Fix four points 1, 2, 3, 4 in  $A$  and consider a variable point 5. Then averaged over the fading and the point placement of 5 we assume:

$$E g_{15} g_{25} g_{35} g_{45} = 0.$$

This is true if the  $g_{ij}$  are zero-mean when averaged over the fading:

$$E g_{15} g_{25} g_{35} g_{45} = E_5 \left[ \underbrace{E_{|5} g_{15}}_0 \underbrace{E_{|5} g_{25}}_0 \underbrace{E_{|5} g_{35}}_0 \underbrace{E_{|5} g_{45}}_0 \right] = 0.$$

## Back to “Listen/Transmit” Protocol

We now have interference. Let’s look at our previous protocol.

1. The listen interval: Each of the  $r$  transmit users transmits the signal  $s_i$  with power  $p$ . All other nodes are silent and measure the signals

$$q_k = \sum_{i=1}^r s_i g_{ik} + v_k, \quad i = 1, \dots, r \quad k = 1, \dots, n$$

2. The transmit interval: Each of the nodes now has to transmit a signal. Let us assume that all it can transmit is a scaled version of what it has previously received:

$$t_k = d_k q_k, \quad k = 1, \dots, n$$

such that the transmit power is  $\sigma_s^2$ . But how to choose  $d_k$ ?

Defining the transmit vector  $s = \begin{bmatrix} s_1 & s_2 & \dots & s_r \end{bmatrix}$ , the receive vector  $y = \begin{bmatrix} y_1 & y_2 & \dots & y_r \end{bmatrix}$  and the matrices

$$G = \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{r1} & g_{r2} & \dots & g_{rn} \end{bmatrix}, \quad H = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1r} \\ h_{21} & h_{22} & \dots & h_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \dots & h_{nr} \end{bmatrix}$$

we may write

$$y = s \cdot G \cdot \text{diag}(d_1, \dots, d_n) \cdot H$$

Note that if  $n \geq r^2$ , the  $d_i$  can be chosen such that

$$G \cdot \text{diag}(d_1, \dots, d_n) \cdot H = I_r, \tag{1}$$

which means that the channel is diagonalized and the interference suppressed!

But this can only be done if each node knows all the network gains so that it can solve (1). This is not allowed.

What to do?

- Let each node estimate each of the  $r$  transmitted signals:

$$\hat{s}_i = g_{ik}^* q_k, \quad i = 1, \dots, r \quad k = 1, \dots, n$$

Of course, this will be a lousy estimate since the SINR (signal-to-interference-ratio) is  $\frac{p}{\sigma_n^2 + (r-1)p}$ .

- Let each node attempt to “coherently add” these estimates:

$$t_k = \frac{1}{\sqrt{r}} \cdot \frac{\sigma_s}{\sqrt{\sigma_n^2 + rp}} \sum_{i=1}^r h_{ki}^* \hat{s}_i = \frac{\sigma_s}{\sqrt{r(\sigma_n^2 + rp)}} \sum_{i=1}^r h_{ki}^* g_{ik}^* q_k$$

Note therefore that

$$d_k = \frac{\sigma_s}{\sqrt{r(\sigma_n^2 + rp)}} \sum_{i=1}^r h_{ki}^* g_{ik}^*$$

depends only on *local* knowledge of the network gains.

The received signal at the  $j$ -th receiver is:

$$y_j = \sum_{k=1}^n t_k h_{kj} + w_j.$$

Close inspection of the received signal reveals that it consists of  $n + 1$  noise terms (one from each of the  $n$  nodes) and  $nr^2$  signal terms, of which

- there are  $n$  terms of the desired signal  $s_j$  that add up *coherently* (one from each of the  $n$  nodes)
- there are  $n(r - 1)$  terms of the desired signal  $s_j$  that add up *non-coherently* ( $r - 1$  from each of the  $n$  nodes)
- there are  $nr(r - 1)$  interference terms ( $r$  from each interferer and each node)

Therefore

$$\text{SINR} = \frac{\frac{(n^2 + n(r-1))\sigma_s^2 p}{r(\sigma_n^2 + rp)}}{\frac{nr(r-1)\sigma_s^2 p}{r(\sigma_n^2 + rp)} + \frac{n\sigma_n^2 \sigma_s^2}{r(\sigma_n^2 + rp)} + \sigma_n^2} \approx \frac{rpn^2 \sigma_r^2}{r^2(\sigma_n^2 + rp)(\sigma_n^2 + n\sigma_r^2)}$$

## Precise Statement: Ad-Hoc Network

Assuming the technical condition:

**Theorem 3** Consider a fixed area  $A$  and randomly select  $n$  nodes and  $r \leq O(n^{\frac{1-\epsilon}{2}})$  transmit/receive pairs in this area. Then

$$\text{Prob}(\eta > k_1 \sqrt{n}) > 1 - \frac{k_2}{n^\epsilon},$$

where  $k_1$  and  $k_2$  are positive constants independent of  $n$ , but that depend on  $A$  and the fading statistics.

Without the technical condition:

**Theorem 4** Consider a fixed area  $A$  and randomly select  $n$  nodes and  $r \leq O(n^{\frac{1-\epsilon}{3}})$  transmit/receive pairs in this area. Then

$$\text{Prob}(\eta > k_1 \sqrt{n}) > 1 - \frac{k_2}{n^\epsilon},$$

where  $k_1$  and  $k_2$  are positive constants independent of  $n$ , but that depend on  $A$  and the fading statistics.

## Power Efficiency of Ad-Hoc Networks

- *provided*  $r \leq O(\sqrt{n})$ , the power efficiency is  $O(\sqrt{n})$
- if we fix the rate for each of the  $r$  transmit/receive pairs, then the total power consumption of the network reduces by a factor of  $\frac{1}{\sqrt{n}}$ .
- this is as good as a sensory network
- rather than diminishing rewards, we are reaping benefits from the increase in the size of the network
- there is built-in fairness: the  $r$  transmitters transmit with power  $p = O(\frac{1}{\sqrt{n}})$  and all other nodes with power  $\sigma_s^2 = O(\frac{r}{n} \cdot \frac{1}{\sqrt{n}})$
- again, the protocol is double-hop and exploits, rather than avoids, interference

And now to the last part of the talk...

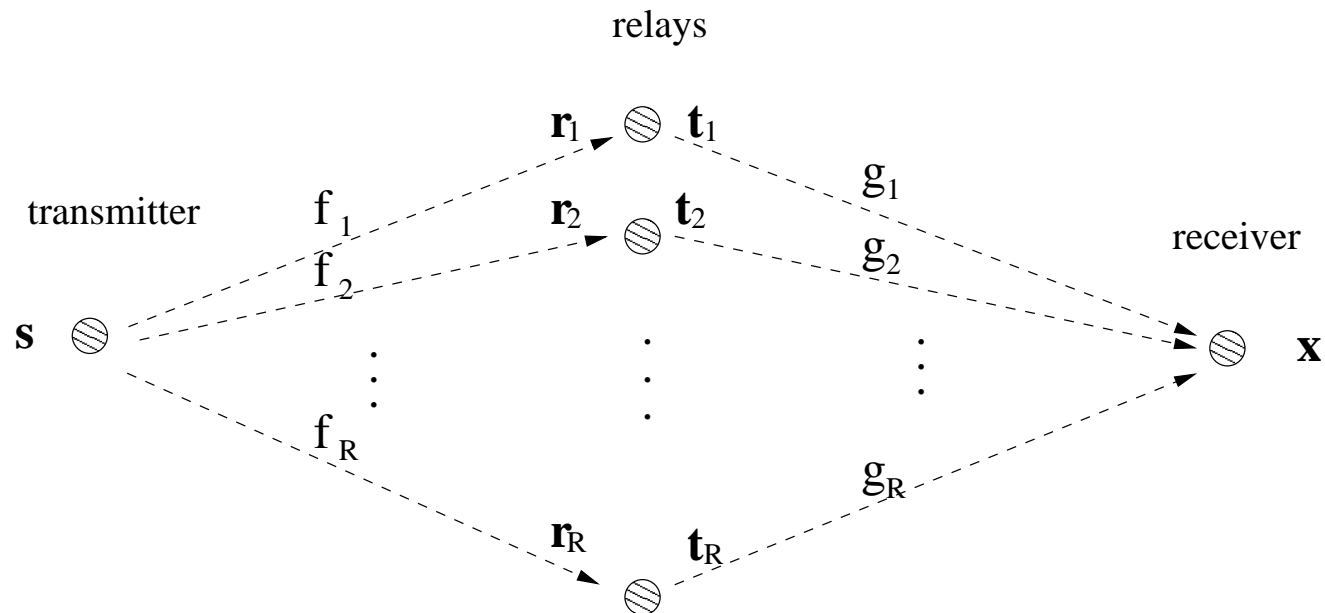
## An Observation on the Role of CSI

- A criticism of the previous results could be that they require that the relay nodes know the local channel coefficients
- This essentially requires that the system be synchronous at the carrier level—something that may be difficult to achieve in a distributed network

It is quite straightforward to show that, without local CSI at the relay nodes, neither power efficiency nor capacity increases with  $n$ .

Thus, is there anything to be had in a nonsynchronous relay network?

# Wireless Relay Network



- single antenna at the transmitter, receiver and all  $R$  relay nodes
- $f_i$ : channel from the transmitter to  $i$ th relay node
- $g_i$ : channel from  $i$ th relay node to receiver
- channel coefficients unknown to relays, but known to receiver

## Modified Listen and Transmit Protocol

**Goal:** send information  $s_1, \dots, s_T$  from transmitter to receiver

**Method:** use a two-step strategy

- *Listen phase:* From time 1 to  $T$ , the transmitter sends  $s_1, \dots, s_T$  to the relay nodes (which are silent).
- *Transmit phase:* From time  $T + 1$  to  $2T$ , the relay nodes transmit signals to the receiver, based on what they have received in the listen phase.

## Notation and System Model

- $\mathbf{s}$ : vector of signals sent from the transmitter at times 1 to  $T$
- $\mathbf{r}_i$ : vector of signals received at relay  $i$  at times 1 to  $T$
- $\mathbf{v}_i$ : vector of noise at relay  $i$  at times 1 to  $T$
- $\mathbf{t}_i$ : vector of signals sent from relay  $i$  at times  $T + 1$  to  $2T$
- $\mathbf{x}$ : vector of signals received at the receiver at times  $T + 1$  to  $2T$
- $\mathbf{w}$ : vector of noise at the receiver at times  $T + 1$  to  $2T$

### System Equations:

$$\mathbf{r}_i = \sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i \quad \text{and} \quad \mathbf{x} = \sum_{i=1}^R g_i \mathbf{t}_i + \mathbf{w}.$$

With  $\text{E tr } \mathbf{s}^* \mathbf{s} = 1$ ,  $P_1$  is the average transmit power at the transmitter.

## Using Space-Time Codes in a Wireless Network

But how to choose the transmit signals  $\mathbf{t}_i$ ?

We will use a linear dispersion code

$$\mathbf{t}_i = \sqrt{\frac{P_2}{P_1 + 1}} A_i \mathbf{r}_i,$$

where  $A_i$  is  $T \times T$  unitary.

- $P_2$  is the average transmit power at each relay node.
- A more general LD code, in which  $\bar{r}_i$  also appears, can also be considered.

## Distributed Space-Time Coding

The received signal vector can be computed to be

$$\mathbf{x} = \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} \underbrace{\begin{bmatrix} A_1 \mathbf{s} & A_2 \mathbf{s} & \cdots & A_R \mathbf{s} \end{bmatrix}}_S \underbrace{\begin{bmatrix} f_1 g_1 \\ f_2 g_2 \\ \vdots \\ f_R g_R \end{bmatrix}}_H + \underbrace{\sqrt{\frac{P_2}{P_1 + 1}} \sum_{i=1}^R g_i A_i \mathbf{v}_i + \mathbf{w}}_W$$

$$\mathbf{x} = \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} SH + W.$$

- $S$  ( $T \times R$ ): works like a space-time code—we call it the *distributed space-time code*
- $H$  ( $T \times 1$ ): the equivalent channel matrix
- $W$  ( $T \times 1$ ): the equivalent noise

## PEP Analysis

**Theorem 5** *The PEP, averaged over the channel coefficients, of mistaking  $\mathbf{s}_i$  by  $\mathbf{s}_j$  has the following Chernoff bound,*

$$PEP \leq \mathbb{E}_{g_i} \det^{-1} \left[ I_R + \frac{P_1 P_2 T}{4 \left( 1 + P_1 + P_2 \sum_{i=1}^R |g_i|^2 \right)} (S_i - S_j)^* (S_i - S_j) G \right]$$

where  $G = \text{diag} \{ |g_1|^2, \dots, |g_R|^2 \}$ ,  $S_i = [A_1 \mathbf{s}_i, \dots, A_R \mathbf{s}_i]$ , and  $S_j = [A_1 \mathbf{s}_j, \dots, A_R \mathbf{s}_j]$ .

For a multiple antenna system (with CSI at the receiver):

$$PEP \leq \det^{-1} \left[ I_R + \frac{PT}{4R} (S_i - S_j)^* (S_i - S_j) \right].$$

- We need to do the expectation over the  $g_i$  (this is the tricky part).
- The “full diversity” condition is the same:  $S_i - S_j$  is full-rank, or equivalently,  $\det(S_i - S_j)^* (S_i - S_j) \neq 0$ .

## PEP Analysis

While the expectation over the  $g_i$  can be performed it is quite tedious. The result simplifies when  $P$  and  $R$  are large

**Theorem 6** *Assume  $T \geq R$ ,  $P \gg 1$  and  $R \gg 1$ . Then the optimal power allocation is*

$$P_1 = \frac{P}{2} \quad \text{and} \quad P_2 = \frac{P}{2R},$$

*i.e., the transmitter uses half the total power and the relays share the other half fairly. Moreover,*

$$Pe \leq \sum_{k=0}^R \left( \frac{8R}{T} \right)^k \sum_{1 \leq i_1 < \dots < i_k \leq R} \det^{-1} [M]_{i_1, \dots, i_k} \frac{\log^k P}{PR},$$

*where  $[M]_{i_1, \dots, i_k}$  is the  $k \times k$  matrix composed by the  $i_1, \dots, i_k$ -th rows and columns of  $M$ .*

## Comparison With Multiple Antenna Systems

- **For relay networks:** The leading order term in the PEP takes the form

$$PEP \leq \left(\frac{8R}{T}\right)^R \det^{-1} M \left(\frac{\log P}{P}\right)^R,$$

which implies a diversity of order  $R(1 - \frac{\log \log P}{\log P})$ .

- **For multi-antenna systems** (with CSI at the receiver):

$$PEP \leq \left(\frac{4R}{T}\right)^R \det^{-1} M \left(\frac{1}{P}\right)^R.$$

The performance of relay network is thus  $(3 + 10 \log_{10} \log P)$ dB worse.

## Discussion of Coding Gain

$$PEP \leq \sum_{k=0}^R \left(\frac{8}{T}\right)^k \sum_{1 \leq i_1 < \dots < i_k \leq R} \det^{-1}[M]_{i_1, \dots, i_k} \frac{\log^k P}{P^R}$$

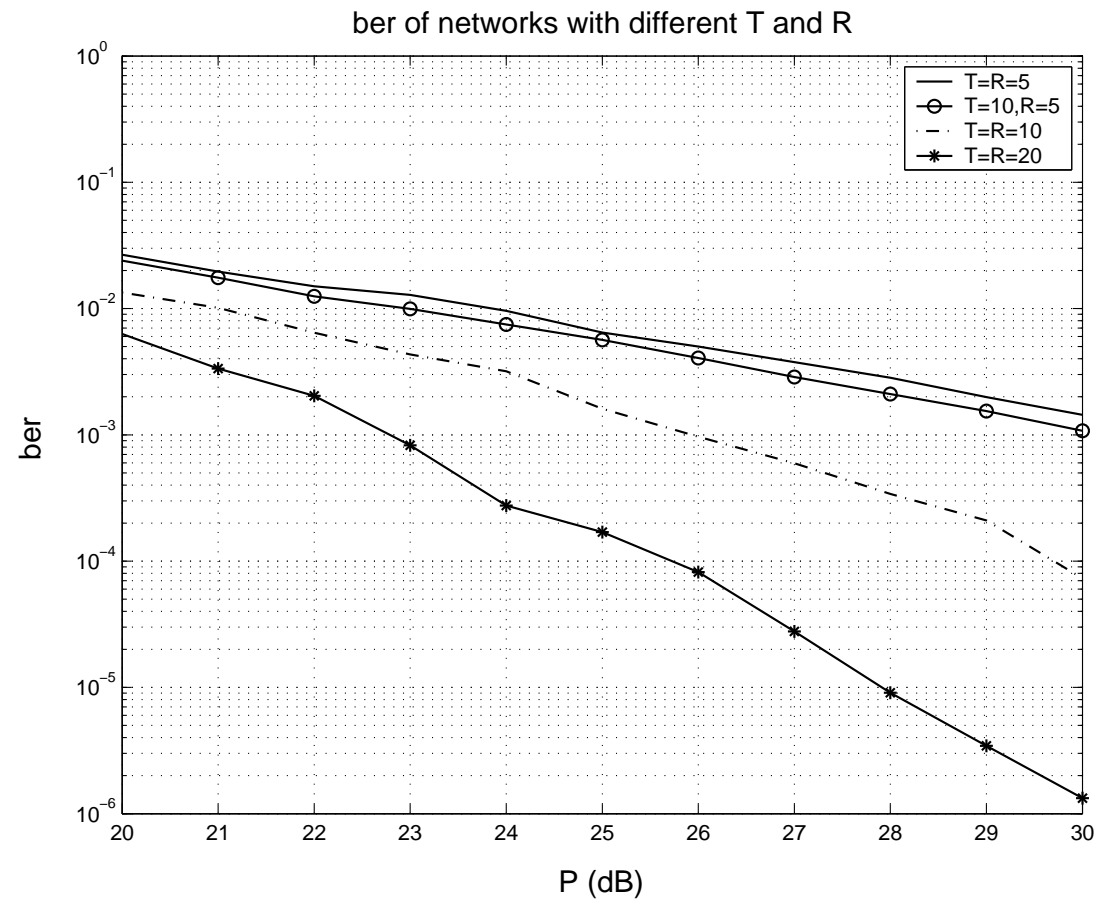
- When  $P$  is very large ( $P \gg \log P$ ), the coding gain is  $\det M$ , which is the same as that of a multiple antenna system.
- When  $P$  is not very large, the second term ( $k = R - 1$ ) and even the  $k = R - 2, \dots$  terms are not neglectable.
- $[M]_{i_1, \dots, i_k} = ([S_i]_{i_1, \dots, i_k} - [S_j]_{i_1, \dots, i_k})^* ([S_i]_{i_1, \dots, i_k} - [S_j]_{i_1, \dots, i_k})$  where  $[S_i]_{i_1, \dots, i_k} = (A_{i_1} \mathbf{s}_i, \dots, A_{i_k} \mathbf{s}_i)$  is the space-time code when only the  $i_1, \dots, i_k$ th relay nodes are working.
- The distributed space-time code should have the property that it is “scale-free” in the sense that it is still a good distributed space-time code when some of the relays are not working.

## Simulations

**Goal:** Compare the performances of relay networks using a distributed space-time code with multi-antenna systems using the same space-time code

- $\mathbf{t}_i = \sqrt{\frac{P_2}{P_1+1}} A_i \mathbf{r}_i$
- $A_i$  are generated randomly based on the isotropic distribution
- $s_1, \dots, s_T$  are designed as independent  $N^2$ -QAM signals
- The rate of the code is  $\frac{1}{T} \log N^{2T} = 2 \log N$
- Error events: bit error rate (ber) and block error rate (bler)
- For multi-antenna system, there are  $R$  transmit antennas and 1 receive antenna

# Diversity Gains of Networks with Different $R$



$$T = R = 10, \text{ Rate} = 2$$

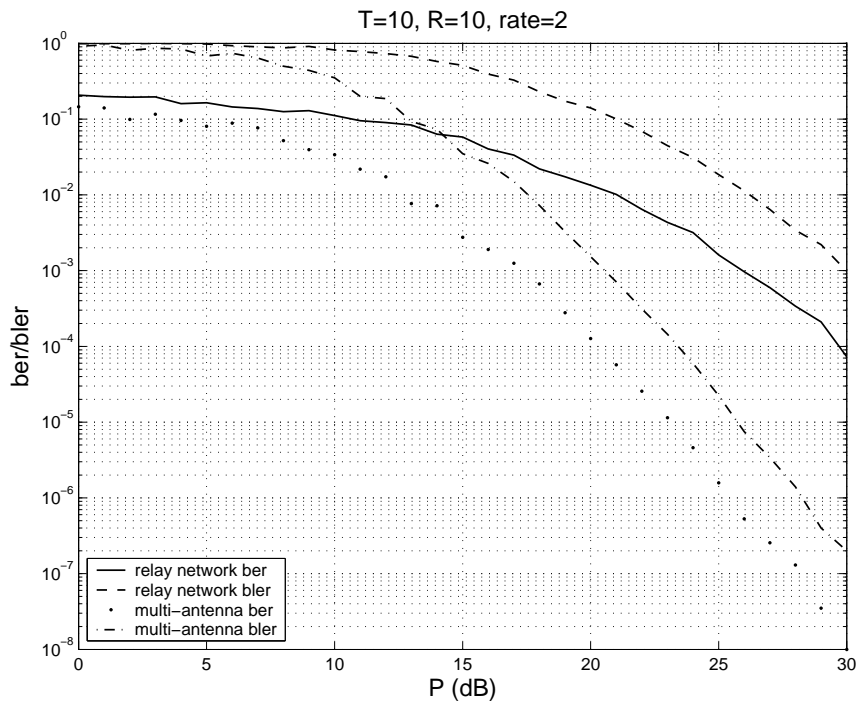


Figure 4: The same power

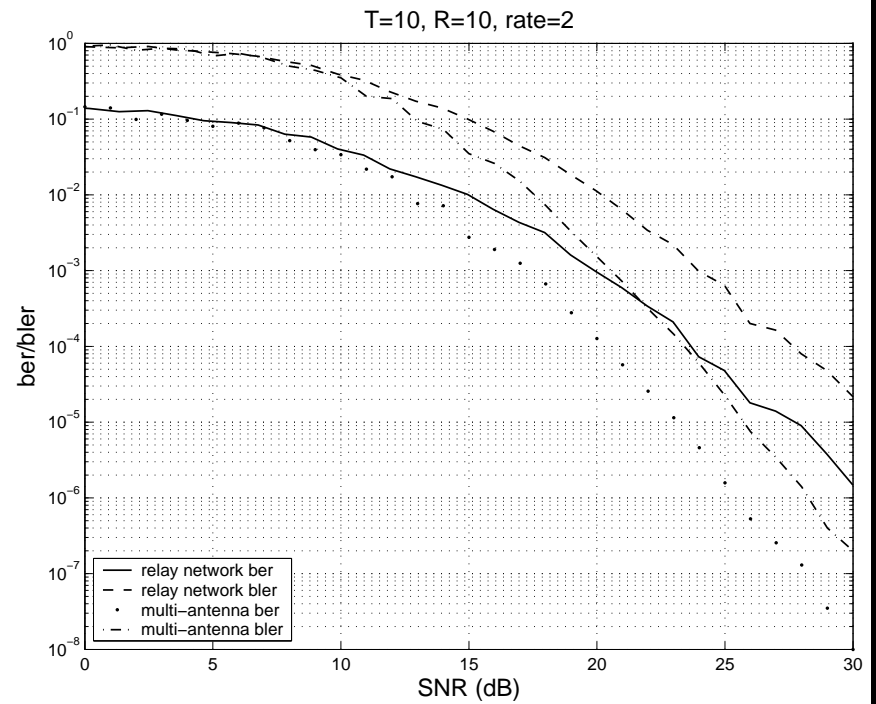


Figure 5: the same receive SNR

## Conculsion and Final Remarks

Studied three problems in wireless networks

- **Broadcast in cellular networks**

- proposed a simple scheme requiring minimal feedback that allows linear increase in capacity in number of transmit antennas
- optimal number of transmit antennas appears to be  $\log n$
- exploits multi-user diversity, but also has built-in fairness
- main question is what side information should be provided to transmitter

- **Power efficiency in sensory and ad hoc networks**

- if relay nodes have local knowledge of the channel coefficients power efficiency scales as  $O(\sqrt{n})$
- key is to use *distributed beamforming* and *distributed interference cancellation*, which exploits the interference

– open question: can one do better than  $O(\sqrt{n})$ ?

- **Diversity of relay networks**

- can achieve a diversity order of  $R(1 - \frac{\log \log P}{\log P})$  via a distributed space-time code, even when the relay nodes have no CSI
- open questions: design of optimal distributed codes, can the diversity be improved? what if receiver has no CSI? (distributed Cayley code?)