# Wireless Networks: How do Things Scale?

Babak Hassibi

Department of Electrical Engineering

California Institute of Technology, Pasadena, CA 91125

*ONR Communications Technology Review*

*May 6, 2003*



## Outline

- **Introduction**
- **Cellular Networks**
  - capacity of multi-antenna broadcast channels with partial CSI
  - fairness issues
- **Sensory and Ad-Hoc Wireles Networks**
  - power efficiency
- **Conclusion**

## Introduction and Motivation

- Seamless, robust, high data-rate connectivity is emerging as a major requirement of the communications systems deployed by the US Navy, as well as in the commercial sector
- Evolution is from a point-to-point paradigm to a network-centric one
- Wireless networks, in particular, provide many interesting opportunities and also many challenging research problems
- A variety of technological advances are now available: RF circuitry, multiple antennas, etc.
- In this talk, we will focus on how various performance measures scale in the network size (number of antennas, number of users, etc.)
- This is of crucial importance since it allows us to determine the performance gains as a function of the resources invested in the network



## Other Work

The work I will describe today is a part of my ongoing research in the general area of wireless communications and, in particular

- multi-antenna information theory
- space-time codes
- algorithms
- PMEPR reduction of OFDM signals

This work is supported by

- ONR
- NSF
- Lee center for advanced networking at Caltech
- JPL

Research group: 8 graduate students, 1 post-doc, 1 visiting associate

## Wireless Networks

Wirelesss networks differ from wireline networks in key ways

- 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.



## Brief History

- First wireless networks consisted of a single base station with large coverage in an urban center, along with mobiles (due to interference, very few could be accomodated)
- cellular networks
  - theory developed in Bell Labs in 60's and 70's
  - first system deployed in early 80's (AMPS)
  - key: exploit path-loss to allow for frequency reuse
  - to increase capacity, cells are shrinking (micro-cells, etc.)
  - FDMA, TDMA, CDMA
- WLAN's (802.11's)
- sensory networks, ad-hoc networks

## Basic Problem to be Addressed

*For cellular, sensory and ad-hoc wireless networks we are interested in how various performance measures scale in the number of users/sensors.*

More specifically,

- For cellular networks how does the "network capacity" scale as a function of
  - the number of transmit antennas at the base station?
  - the number of users?
  - the number of receive antennas at each user?
- For sensory and ad-hoc networks how does the
  - network capacity
  - power efficiency

  scale as a function of the number of users?

These are of great practical interest for the design of such networks.



## Network Information Theory

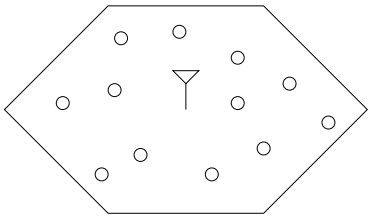
Such questions are related to **Network Information Theory**:

- 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
  - reminiscent of mechanics where three-body problems are open

This is one reason why we will focus on asymptotic results for *large* networks (analogous to statistical mechanics where we have many bodies)

The other reason is that large networks are of practical interest

## Cellular Networks

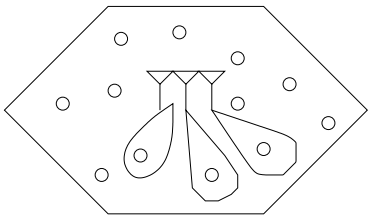


- capacity per cell is fixed (depends on bandwidth and transmit power)
- TDMA, FDMA, CDMA all essentially give same capacity
- to increase network capacity cell sizes are shrinking

Under ideal conditions, there exist sophisticated techniques where capacity scales by $\log \log n$ ($n$ is number of users)



## Increasing Cell Capacity with Multiple Antennas



- idea is to exploit *spatial diversity* to transmit simultaneous beams to more than one user
- requires knowledge of spatial location of different users at base station (also line-of-sight environment)
- some companies attempted this in the 90's (e.g., ArrayComm) but it never really caught on

## Scaling Laws 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:

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

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

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

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

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

  where $T$ is the coherence interval of the channel

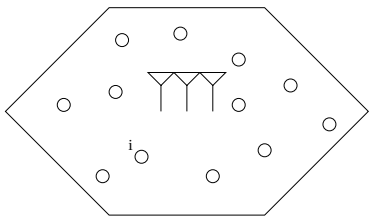


Thus, in point-to-point multiple antennas systems, capacity scales linearly in the minimum of the number of transmit/receive antennas, *irrespective of whether the channel is known at the receiver and/or transmitter*

But what happens in a network setting?...

## Multiple Antennas in Broadcast Channels

Suppose now we have $M$ transmit antennas at the base station and $n$ users each with $N = 1$ receive antennas




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.



## Scaling Laws in Broadcast Channels

- Channel coefficients known to transmitter (Shamai and Caire '02):

$$C = \max_{\sum_i p_i = P} E \log\det\left(I_M + \sum_{i=1}^{N} h_i^* p_i h_i\right)$$

  - achieved by *dirty paper coding* (Costa '80).
  - asymptotically, for fixed $M$ and $n \to \infty$,

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

  - **advantage:** linear growth in $M$
  - **disadvantages:** requires full channel knowledge, sensitive to channel errors, very computationally intensive

- Channel coefficients unknown to transmitter:

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

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

## Conclusion

Thus, the network situation is very different from the point-to-point one.

Lack of knowledge of the channel coefficients brings us down from

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

to

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

- full channel-state-information (CSI) unreasonable
- multiple antennas do not buy us much if no CSI available

What to do?



## Send Random Beams

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

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

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 $i = 1, \ldots, n$ 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 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

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

Due to the symmetry of all the random variables involved:

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

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 $x_i$, $i = 1, \ldots, n$ be iid random variables with distribution $p(x)$. Then if,*

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

*then*

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

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

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

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

Thus, multi-user diversity buys us a lot!



## Some Remarks

- Using partial CSI, our scheme achieves $M \log\log n$, precisely the same as the full CSI case
- Compared to the full CSI case where each user had to feedback $M$ channel coefficients, here each user need only feed back its *best* SINR and the corresponding index.
  - In fact, not all users need to do this. If only users whose best SINR exceeds a threshold, $\eta$, say, feed it back then

    $$C \geq \left(1 - F^{Mn}(\eta)\right) M \log\log n + O(\log\log\log n),$$

    where $F(\cdot)$ is the CDF of the SINR. This reduces the feedback by a factor of $\frac{1}{1-F(\eta)}$.

## How Large can $M$ Be?

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

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 $n \to \infty$, we have

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

- If $\frac{M}{\log n} \to \infty$, then

$$\frac{C}{M} \to 0.$$

Therefore a judicious choice seems to be $M = O(\log n)$



## What if We Add Multiple Receive Antennas?

Suppose we have $N > 1$ receive antennas. Then the best we can do is to treat each antenna independently

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

- Thus, multiple receive antennas do not substantially increase the network capacity
- However, if some fraction of the users have $N$ receive antennas they will receive $N$-times the rate (at the expense of other users in the network)

## Fairness

- 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
- Thus, there appears to be tension between multi-user diversity and fairness
- This is certainly true for single-transmit-antenna systems



## Single Antenna Multi-User Fairness

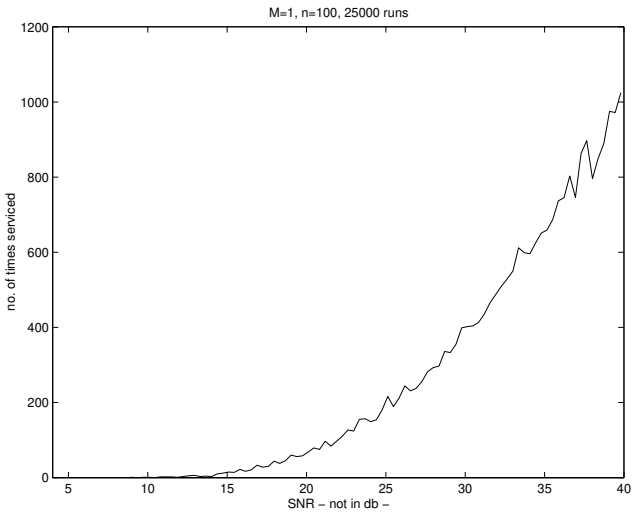



Figure 1: $M = 1$, $n = 100$, SNR 4-40 (6-16db), 25000 runs

## Fairness - Continued

In a heterogenous network, the SINRs now become

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

Note that, if the $\mathrm{SNR}_i$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:**

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

This further illustrates the benefits of having $M = O(\log n)$.



## Multi-Antenna Multi-User Fairness

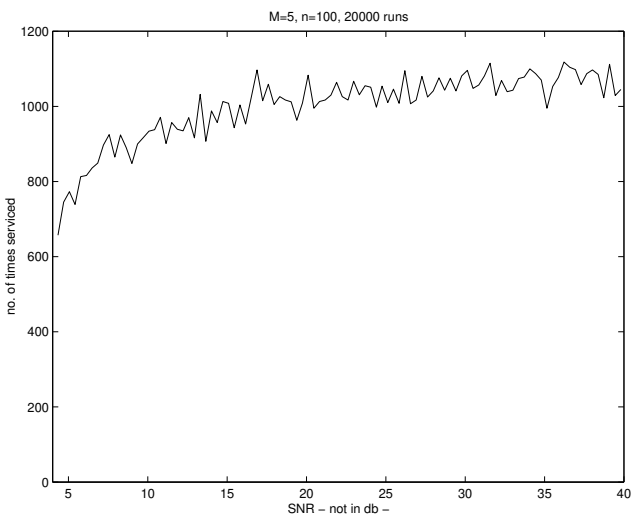



Figure 2: $M = 5$, $n = 100$, SNR 4-40 (6-16db), 20,000 runs

## Sensory Networks and Ad-Hoc Networks

**Sensory Networks:**

- applications: surveillance, military, environmental monitoring, 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/receivers
- applicable in battlefield scenarios



## Scaling Laws for the Capacity

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

$$\eta = \frac{\text{capacity}}{\text{transmit power}}$$

For an AWGN channel this is

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



Clearly,

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

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

  which means we are power in-efficient at high SNR
- as $\sigma_s^2 \to 0$:

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

  which means we are power efficient at low SNR

This was all from a single-user perspective.

What happens in a network setting?...

## Multi-Antenna Power Efficiency

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

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

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

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

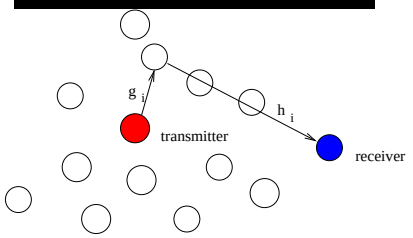
- if the channel matrix *unknown* to the transmitter (Foschini-Telatar):

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

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:

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

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

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

The received signal is

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



The point is that we have done *distributed beam-forming* and so the signal adds up coherently and the noise non-coherently:

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

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

$$\text{SNR} = O(n^{1-2\epsilon}) \quad \text{and} \quad \eta = O(n^{1-\epsilon})$$

Since we must operate at low SNR to be power-efficient, this implies $\epsilon > 1/2$ 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*

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

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

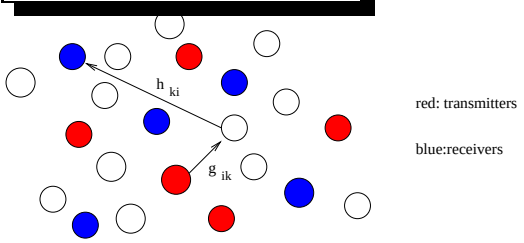


## Power Efficiency of Sensory Networks

- 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!

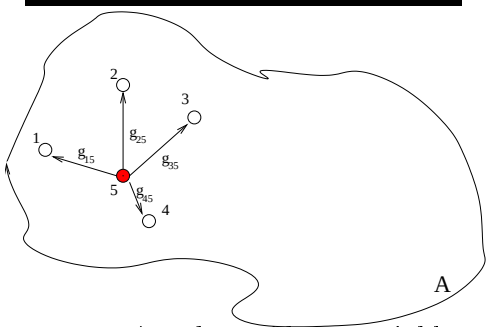
## Ad-Hoc Networks




- $n$ nodes, $r \leq \sqrt{n}$ users transmitting, $r$ users receiving
- gains from transmitters to each node: $g_{ik}, \quad i = 1, \ldots, r \quad k = 1, \ldots, n$
- gains from each node to receivers: $h_{ki}, \quad k = 1, \ldots, n \quad i = 1, \ldots, r$
- **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:

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

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

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

## 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

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

2. The transmit interval: Each of the nodes now transmits the signal

$$t_k = \frac{\sigma_s}{\sqrt{r(\sigma_n^2 + rp)}} \sum_{i=1}^{r} h_{ki}^* g_{ik}^* q_k, \quad k = 1, \ldots, n$$

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



## 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*

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

*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*

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

*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



## Final Remarks and Further Work

We studied various scaling laws for wireless networks.

For cellular networks:

- network capacity depends on the amount of CSI
  - full CSI: $C = M \log\log n + O(1)$
  - no CSI: $C = \log M + O(1)$
- 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 $O(\log n)$
  - exploits multi-user diversity, but also has built-in fairness
- What are forms of partial CSI allow for $C = M \log\log n + O(1)$?

For sensory and ad-hoc networks

- the capacity increase diminishes as the size of the network increases
- if one focuses on power efficiency it apparently does pay off to consider sensory and wireless ad-hoc networks
- the net benefit is larger than the sum of the individual benefits
- to get encouraging asymptotics, rather than attempting to see how the rate scales, the key is to fix the rate and study how the power can be lowered
- interference is, in fact, useful
- it would be interesting to see determine scaling laws for other performance measures such as delay, error probability, bandwidth/power trade-off, etc.