On the Maximum Stable Throughput Problem in Random Networks with Directional Antennas *

Christina Peraki
School of Electrical and Computer Engineering
Cornell University
cp87@cornell.edu

Sergio D. Servetto
School of Electrical and Computer Engineering
Cornell University
sds36@cornell.edu

ABSTRACT

We consider the problem of determining rates of growth for the maximum stable throughput achievable in dense wireless networks. We formulate this problem as one of finding maximum flows on random unit-disk graphs. Equipped with the max-flow/min-cut theorem as our basic analysis tool, we obtain rates of growth under three models of communication: (a) omnidirectional transmissions; (b) "simple" directional transmissions, in which sending nodes generate a single beam aimed at a particular receiver; and (c) "complex" directional transmissions, in which sending nodes generate multiple beams aimed at multiple receivers. Our main finding is that an increase of $\Theta(\log^2(n))$ in maximum stable throughput is all that can be achieved by allowing arbitrarily complex signal processing (in the form of generation of directed beams) at the transmitters and receivers. We conclude therefore that neither directional antennas, nor the ability to communicate simultaneously with multiple nodes, can be expected in practice to effectively circumvent the constriction on capacity in dense networks that results from the geometric layout of nodes in space.

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Stochastic Processes, Probabilistic Algorithms; G.2.2 [**Discrete Mathematics**]: Graph Theory—*Network Problems*

General Terms

Algorithms, Design, Performance, Theory.

Keywords

Maximum Stable Throughput, Random Networks, Random Graphs, Network Flow, Directional Antennas, Wireless Networks, Multicommodity Flow.

*This work was supported in part by the National Science Foundation, under awards CCR-0238271 (CAREER), and CCR-0227676.

MobiHoc'03, June 1–3, 2003, Annapolis, Maryland, USA.

1. INTRODUCTION

1.1 Problem Setup

Consider the following network communication problem. n nodes s_i are placed on the closed set $[0,1]\times[0,1]$, at uniformly distributed random locations (x_i, y_i). Each s_i observes and encodes a source of information, and this encoding is to be relayed over the network to a randomly chosen receiver r_i (equal to some other s_k). Each s_i can only send messages to and receive messages from nodes within some distance d_n, for a suitable choice of the common transmission range d_n to be determined later. Links have a fixed finite capacity L, independent of network size. This scenario is illustrated in Fig. 1.

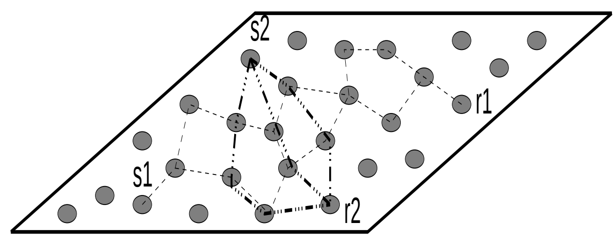

Figure 1: Problem setup. n randomly located transmitters send send data to n randomly chosen receivers, all nodes act as sources/destinations/relays, and nodes can only exchange messages with nearby nodes (within a given transmission range d_n). Optimal routing is assumed. The goal is to find out how much data this network can carry reliably.

Our goal in this paper is to determine the rate of growth for the *maximum stable throughput* (MST) of the network [1, 21]—the total number of packets that all sources can inject into the network, while keeping the size of the largest queue bounded—for a variety of transmitter/receiver models based on directional antennas.

1.2 On Directional Antennas and MST Issues

Why the interest in directional antennas? Because there is a question about wireless networks equipped with such antennas which we believe is very important, and for which we could not find a satisfactory answer in the literature:

- A key result in the analysis of performance of wireless networks states that when n non-mobile nodes are randomly placed in a disk of unit area, traffic patterns are optimally assigned, and the range of each transmission is optimally chosen, the total throughput that the network can carry is $\Theta\left(\sqrt{n/\log n}\right)$ [12].[1] The per-node throughput then is only

[1] In this paper we will use the notation of Graham/Knuth/Patashnik

$\Theta\left(1/\sqrt{n\log(n)}\right)$, i.e., vanishes as the number of nodes in the network increases. The work of [12] sparked significant interest in this problem (see, e.g., [9, 17, 18, 20, 25].

- In a different segment of the research community, the use of *directional* antennas has also received a fair amount of attention in recent times. The rationale is that with omnidirectional antennas, existing MAC protocols require all nodes in the vicinity of a transmission to remain silent. With directional antennas however, it should be possible to achieve higher overall throughput, by means of a higher degree of spatial reuse of the shared medium, and a smaller number of hops visited by a packet on its way to destination (see, e.g., the recent work of [3]). Furthermore, in the context of energy-efficient broadcast/multicast, it has been argued that the ability of a transmitter to reach multiple receivers is an important source of gains to take advantage of in the development of suitable protocols, such as BIP [24].

If we take a step back, careful reading of all these previous results raises an important question: how much exactly is there to gain from the use of directional antennas? Could directional antennas (in which the width of the beams tends to zero as n gets large) be used to effectively overcome the vanishing maximum throughput ($\Theta(1/\sqrt{n\log(n)})$) of [12]? Although we have not been able to find answers to this question in the literature (and that motivated us to start working on this problem in the first place), we have found a couple of related results based on which we can say a-priori that the answer is probably *no*:

- In [12], the authors claim that their result holds irrespective of whether transmissions are omnidirectional or directed, provided that in the case of directed antennas there is some lower bound (independent of network size) on how narrow the beams can be made.
- In [14, 19], for some *regular* networks, it is shown that enabling nodes with Multi-Packet Reception (MPR) capabilities [7] can only increase the total throughput of the network by a constant factor (≈ 1.6), independent of network size.

Given this state of affairs, it seems to us that deciding exactly how much there is to be gained by using directional antennas, and giving some measure of how complex the transmitters/receivers need to be made to achieve those gains, is indeed a topic worth being studied. In this paper, we address both issues.

1.3 Flows on Random Graphs

The main idea behind our approach to this problem is simple: the transport capacity problem posed in [12] is essentially a throughput stability problem—the goal is to determine how much data can be injected by each node into the network while keeping the system stable—, and this throughput stability problem admits a very simple formulation in term of flow networks. Network flow techniques have been proposed to study capacity problems in communication networks before (see, e.g., [2], [5, Ch. 14.10]), and the work carried out in this paper builds on ideas we started developing in [17].

MST is Maximum Multicommodity Flow

The problem of finding the maximum stable throughput of our network is an instance of a *multicommodity flow* problem [4, Ch. 29]:

- There are n commodities: the packets for the transmission from transmitter s_i to receiver r_i.
- The sum of the packets transmitted by all sources cannot exceed the capacity of a link.
- Subject to these constraints, we want to find the largest number of packets that can be injected simultaneously by all sources.

If we represent our network by a graph $G=(V,E)$,[2] the capacity of an edge by $c(u,v)$, and let our optimization variables be $f_i(u,v)$ (the flow along edge (u,v) for the i-th commodity), then the maximum multicommodity flow problem above can be formulated as a linear program, as shown in Table 1.

max	d
subject to:	
$d=\sum_{(s_i,v)\in E} f_i(s_i,v)$,	$1\le i\le n$
$\sum_{i=1}^n f_i(u,v)\le c(u,v)$,	$(u,v)\in E$
$f_i(u,v)=-f_i(v,u)$,	$(u,v)\in E, 1\le i\le n$
$\sum_{v\in V} f_i(u,v)=0$,	$u\in V-\{s_i,r_i\}, 1\le i\le n$

Table 1: Linear programming formulation of the multicommodity flow problem. Note that the flow conservation property is essentially a throughput stability property, in that it ensures that there is no build-up of packets at intermediate nodes (i.e., queues are stable).

A reader familiar with network flows may wonder at this point whether we really need a multicommodity formulation in our problem, or we could get away with considering simpler single commodity flows. We would like to point out that indeed, our problem is essentially a multicommodity problem. The key to see why this is the case is that, in the single commodity problem with multiple sources and sinks, units of flow travel from *any* source to *any* sink, but in our problem, the flow generated by one source has to reach one and only one specific sink. Another observation that the same reader could make is that, in general, maximum multicommodity flow is an NP-hard problem [4, Ch. 29]. But in our setup, the fairness constraint among sources that requires that they all inject the same amount of flow (the first constraint in Table 1) does render the problem computationally tractable.

A Restriction in the Optimization Domain

Not much is known about the structure of optimal solutions to the maximum multicommodity flow problem—the only technique we are aware of for deciding whether a particular amount of flow of each commodity can be supported by the network consists of formulating this problem as a linear program, and then answering the non-emptyness question for its polytope of optimization, using a polynomial time algorithm (e.g., the Ellipsoid method [10]). Therefore, we will not be able to use that formulation to do much about the problem of interest to us. However, there is one special case of our problem which is interesting in its own right, and is described in Fig. 2.

Note that doing this amounts to introducing a restriction in the domain of optimization of the linear program from Table 1: instead of considering all possible network realizations, we only consider those which satisfy the constraints of Fig. 2. The new linear program is shown in Table 2.

Before continuing, we would like to argue that although networks as in Fig. 2 are certainly a restriction of the general case, this restriction is a very interesting one to consider:

for the rate of growth of functions [8, Ch. 9].

[2] The definition of this graph is not trivial, and will be addressed in later sections for different tx/rx architectures. For now, we rely on an intuitive notion of how a network could be represented by a graph. This issue will be dealt with extensively later on.

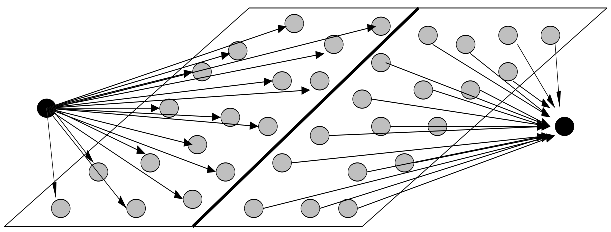

Figure 2: A special case, in which instead of having n nodes that simultaneously play the role of a source/relay/sink, we have $\Theta(n)$ (with probability that tends to 1 as $n \to \infty$) nodes on the left half of the network play the role of source/relay only (no sinks), and $\Theta(n)$ (with probability that tends to 1 as $n \to \infty$) nodes on the right half of the network that play the role of relay/sink only (no sources).

max	$\sum_{u \in V} f(s,u)$
subject to:	
$f(u,v) \le c(u,v),$	$(u,v) \in E$
$f(u,v) = -f(v,u),$	$(u,v) \in E$
$\sum_{v \in V} f(u,v) = 0,$	$u \in V - \{s,t\}$

Table 2: Linear programming formulation of the single commodity flow problem. In this case s is the single (super)source, and t is the single (super)sink.

1. It still captures the key aspect of wireless networks that lead to their constriction of capacity: the need for nodes to share the channel with other neighboring nodes. Packets going from the left to the right must contend for spatially constrained network resources to cross the center cut.

2. We will see later how this formulation can be thoroughly analyzed based on standard (*not* multicommodity) flow methods. This is important because there is a significant amount of analytical machinery available to us to study regular flow problems, that we can bring to bear in this context.

3. Based on this analysis, we find that at least in one network architecture which had been analyzed already (omnidirectional transmissions, [12]), the scaling laws we obtain using our proof techniques coincide with previously known scaling laws. This means that, at least in this case, the optimization problems defining both the multicommodity and the regular flow problems admit the same solution, and hence there is no loss of optimality due to considering this restriction.

Because of these reasons, in the rest of this paper we will be focusing only on this special case. The most important point still left open as of the writing of this paper is actually proving or disproving the equivalence between the two linear programs above.

1.4 Outline of the Proof Techniques

Counting Edges Across a Minimum Cut

Our main task in this paper consists of determining the rate of growth (as a function of network size n) for the solution of the linear program in Table 2. To do so, we make use of a standard result in flow networks: the max-flow/min-cut theorem of Ford and Fulkerson [6]. We solve this problem essentially by counting how many edges, depending on different tx/rx architectures, can be constructed so that they all simultaneously straddle a minimum cut.

A natural question that arises at this point is: what are good cuts that will give us useful estimates for the values of the linear program? And here is where we see once more why the restriction we introduced to turn this problem from a multicommodity flow problem into one involving regular flows only makes sense: there is a "self-evident" cut one should consider under this restriction, the division between the region of transmitters and the region of receivers. Formally, this cut is defined as

$$S = (x_i, y_i) \in V \cap [0, \tfrac{1}{2}) \times [0,1],$$
$$T = (x_i, y_i) \in V \cap [\tfrac{1}{2}, 1] \times [0,1].$$

The choice of cut for analysis is shown in Fig. 3.

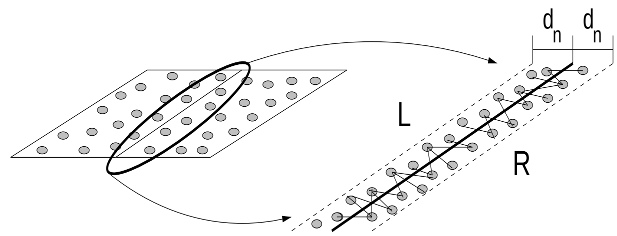

Figure 3: To illustrate the choice of a cut to derive bounds. L and R are sections of the network on each side of the cut boundary, of width d_n, the transmission range.

In the process of determining these counts, we will proceed in two steps. First, we will compute the *expected* number of edges that straddle this cut, with this expectation taken as an ensemble average over all possible network realizations. And then we will prove uniform sharp concentration results: given an arbitrary network realization, with probability 1 as $n \to \infty$, we will show that in this network the actual number of edges that straddle the cut has the exact same rate of growth (in the Θ sense of [8]) as the ensemble mean does.

A Few Tools Used Frequently in this Paper

There are two results we will be using repeatedly throughout our calculations of average number of edges that cross the cut, and so we state them here once:

- What is the smallest possible value of d_n that, under the given statistics for placement of nodes, guarantees that the resulting graph will be connected? This question was answered in [11]. With probability 1 as $n \to \infty$, the graph is connected if and only if d_n is such that

$$\pi d_n^2 = \frac{\log n + \xi_n}{n}, \qquad (1)$$

  for some $\xi_n \to \infty$.

- What is the average number of nodes in a subset $A \subseteq [0,1] \times [0,1]$? A straightforward calculation shows that

$$E(\textit{Number of nodes in } A) = nP(A) = n \int_A f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = n|A|, \qquad (2)$$

  where $|A|$ denotes the area of A.

Also, to prove sharp concentration results, there is one form of the well known Chernoff bounds we will be using repeatedly, and so we state it once here, in general:

- Suppose $X_1 \dots X_n$ are iid and uniformly distributed n points on the $[0,1] \times [0,1]$ plane. Consider we have a number of subsets $A_j \subset [0,1] \times [0,1]$, for $j = 1 ... f(n)$ (the number of subsets may depend on the number of points n), and denote the area of any such subset by $|A_j|$. Now we define some random variables:

$$N_{ij} = \begin{cases} 1, & X_i \in A_j \\ 0, & \text{otherwise.} \end{cases}$$

  Since the X_i's are independent, the N_{ij}'s are also independent.

- Now let N_j be another random variable defining the number of points in A_j, i.e., $N_j = \sum_{i=1}^{n} N_{ij}$. We see in this case that the N_j's, $j = 1 \dots f(n)$ are random variables where each is the sum of n iid binary random variables (but not necessarily independent among the N_j's themselves).

- The expected number of points in A_j is

$$E(N_j) = E\left(\sum_{i=1}^{n} N_{ij}\right) = \sum_{i=1}^{n} E(N_{ij}).$$

  But, since $P(X_i \in A_j) = |A_j|$, we have that $E(N_{ij}) = 1|A_j| + 0(1-|A_j|) = |A_j|$, and hence $E(N_j) = n|A_j|$.

For the family of variables N_j, we have the following standard results, known as the *Chernoff* bounds (see, e.g., [15, Ch. 4]):

1. For any $\delta > 0$:

$$P\left[N_j > (1+\delta)n|A_j|\right] < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{n|A_j|}.$$

2. For any $0 < \delta < 1$:

$$P\left[N_j < (1-\delta)n|A_j|\right] < e^{-\frac{1}{2}n|A_j|\delta^2}$$

Combining the two inequalities we can write for any $0 < \delta < 1$:

$$P\left[\,|N_j - n|A_j|| > \delta n|A_j|\,\right] < e^{-\theta n|A_j|}, \tag{3}$$

where $-\theta = \delta - (1+\delta)\ln(1+\delta)$ in the case of the first bound, and $-\theta = -\frac{1}{2}\delta^2$ in the case of the second bound. Therefore, $\theta > 0$ always, and hence as $n \to \infty$, there exist constants such that deviations from the mean by more than these constants occur with probability 0.

1.5 Main Contributions

Summary of Results

In this paper we apply the proof techniques outlined in 1.4 in the analysis of three different transmitter/receiver architectures: omnidirectional antennas, directional antennas capable of generating a single beam, and directional antennas capable of generating multiple beams simultaneously.

We consider first the case of omnidirectional antennas. Since we have not yet attempted to prove the equivalence of the two linear programs in Tables 1 and 2, it appears necessary to at least provide some kind of evidence that the proof techniques we develop in this work do indeed yield meaningful results. And we do this by showing that, in the case of omnidirectional transmissions, the scaling laws obtained based on our proof method are identical to those of [12]. We are currently investigating the issue of equivalence of these two linear programs.

Then we apply the same proof techniques to the determination of scaling laws for a new architecture, in which transmitter nodes can generate a single and arbitrarily narrow directed beam, and in which receivers can successfully decode multiple transmissions as long as the transmitters are not co-linear. And in this case we find that:

- If only enough power to maintain the network connected is radiated at each node, the maximum stable throughput of this network is $\Theta(\sqrt{n \log n})$.

- If now enough power is radiated to achieve a maximum stable throughput linear in network size (certainly feasible with arbitrarily narrow directed beams), then the number of *resolvable* beams that each node must generate is $\Theta(n)$.

Finally, we consider a node architecture in which each node is able to generate *multiple and arbitrarily narrow* directed beams, simultaneously to all nodes within its transmission range, and receivers operate as above. In this case we find that:

- If only enough power to maintain the network connected is radiated at each node, the maximum stable throughput of this network is $\Theta(\sqrt{n} \log^{\frac{3}{2}} n)$.

- If now enough power is radiated to achieve a maximum stable throughput linear in network size (certainly feasible with arbitrarily narrow directed beams), then the number of *resolvable* beams that each node must generate is $\Theta(n^{\frac{1}{3}})$.

Relevance of the Results

Essentially, our results show that both directional antennas, as well as the ability to communicate simultaneously with multiple nodes, can only achieve modest improvements in terms of achievable MST. While some performance gains are certainly feasible at reasonable complexities (in the order of a low-degree polynomial in $\log n$), the number of *resolvable* beams that need to be generated to actually increase the achievable MST by more than a polylog factor is polynomial in network size, and exponential in the minimum number of beams that would be required to satisfy the basic requirement of keeping the network connected. How many beams need to be resolved is a reasonable measure of complexity, since the higher this number, the narrower these beams need to be made, and hence the higher the complexity of a practical implementation.

On a more conceptual note, we believe that the proof techniques developed in this paper form an interesting contribution in their own right as well. Our results are obtained using only elementary network flow concepts [4], and the calculations involved require only basic probability theory, basic calculus and basic combinatorics. By interpreting the formulation of Gupta and Kumar for random networks presented in [12] as an elementary problem of flows in random graphs, we were able to obtain what we believe is a number of interesting insights into the nature of this problem which were not obvious to us from their proof technique, as well as a set of meaningful generalizations—directional antennas is one presented here, but we have other generalizations currently under investigation as well.

Organization of the Paper

The rest of this paper is organized as follows. In Section 2 we obtain scaling laws for the case of omnidirectional antennas, in Section 3 we do the same for transmitters generating a single directed beam, and in Section 4 we do the same for transmitters capable of generating multiple beams simultaneously. We present concluding remarks in Section 5. We also include an appendix with relevant background material, to keep the paper self-contained.

2. OMNIDIRECTIONAL ANTENNAS

Before we start considering more general node architectures, we show in this section how, for the case of nodes equipped with omnidirectional antennas, using our proof techniques we obtain scaling laws identical to those reported in [12]. This shows that, at least in this case, the value of the two linear programs is identical (to within a constant factor), and therefore that the restriction in the optimization domain does not result in any performance degradation.

2.1 Transmitter/Receiver Model

In [12], transmissions were omnidirectional, and described based on a pure collision model: for a transmission to be successfully decoded, no other transmission has to be in progress within the range of the receiver under consideration. This setup is illustrated in Fig. 4.

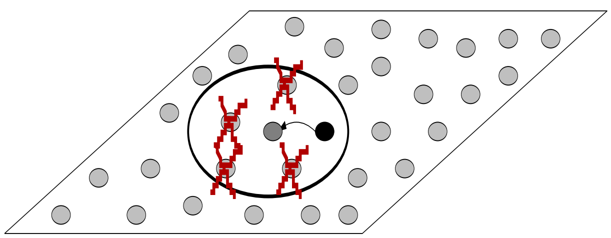

Figure 4: A transmission model based on omnidirectional antennas and pure collisions.

2.2 Average Number of Edges Across the Cut

Our first task is to determine the *average* number of edges that can be simultaneously supported across the cut, average taken over all possible network realizations.

An Upper Bound

For a fixed receiver location (x, y) in R, there can only be one active transmitter within distance d_n of the receiver, for that transmission to be successfully received. Since to obtain an upper bound we only need worry about edges that cross the cut, we first consider all possible locations of one such transmitter in L, by drawing a circle of radius d_n and center (x, y). This region is illustrated in Fig. 5.

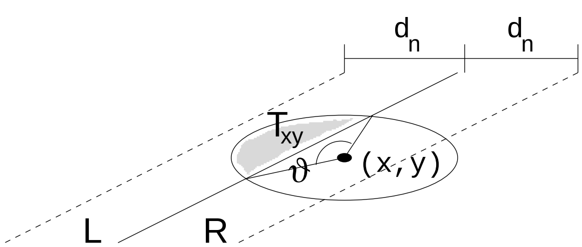

Figure 5: For a receiver at location (x, y), at most one transmitter in the shaded region T_{xy} can send a message (if this message is to be successfully decoded on the other side of the cut).

Denoting by $|T_{xy}|$ the area of the shaded region T_{xy} in Fig. 5, we use eqn. (2) to estimate the number of transmitters located in T_{xy} as $n|T_{xy}|$. However, since only one transmitter located within T_{xy} can transmit successfully to a receiver at (x, y), the number of nodes that are able to transmit at the same time from L to R is upper bounded by

$$\frac{E(\textit{Number of nodes in } L)}{E(\textit{Number of nodes in } T_{xy})} = \frac{nL}{nT_{xy}}$$

This is an *upper bound*, because we are assuming that it is possible to find a set of locations (x, y) in R such that no area in L is wasted—showing that this bound is indeed tight requires proof.

Now, the area of L is d_n. To compute the area of T_{xy}, we have to determine the area of an arc of a circle with angle ϑ, as shown in Fig. 5, and in a computation entirely analogous to that of the calculation of $|Q_p|$ in Section 4. In this case, we have that $\sin(\frac{1}{2}(\pi - \vartheta)) = \frac{x-\frac{1}{2}}{d_n} = \cos(\frac{1}{2}\vartheta)$, and since $\frac{1}{2} \le x < \frac{1}{2} + d_n$ it is clear that we must have $0 < \vartheta \le \pi$ and also $\sin\vartheta \ge 0$. Then, we get $|T_{xy}| = \frac{1}{2}\vartheta d_n^2 - \frac{1}{2}d_n\cos(\frac{1}{2}\vartheta)2d_n\sin(\frac{1}{2}\vartheta) = \frac{1}{2}\vartheta d_n^2 - \frac{1}{2}d_n^2\sin\vartheta$, and therefore, $|T_{xy}| = \frac{1}{2}d_n^2(\vartheta - \sin\vartheta)$. Hence, for each possible value of ϑ, an upper bound on the number of nodes that are able to transmit at the same time from L to R is

$$\frac{nL}{nT_{xy}} = \frac{nd_n}{n\frac{1}{2}d_n^2(\vartheta - \sin\vartheta)} = \frac{2}{d_n(\vartheta - \sin\vartheta)}.$$

Since this upper bound depends on the choice of receiver location (through the angle ϑ), we will make this bound as small as possible by an appropriate choice of ϑ. As noted above, $0 < \vartheta \le \pi$, and $\sin\vartheta \ge 0$. Hence, the number of transmitters in L is smallest when $\vartheta = \pi$ and $\sin\vartheta = 0$, i.e., when the receivers are located close to the cut boundary (as it should be, since it is in this case when receivers "consume" the maximum amount of transmitter area). In this case, we get

$$\min_{0<\vartheta\le\pi}\left[\frac{2}{d_n(\vartheta - \sin\vartheta)}\right] = \frac{2}{\pi d_n}$$

as an upper bound on the number of edges across the cut. Furthermore, in this case we see immediately that to maximize capacity we must keep d_n as small as possible—and we know from eqn. (1) that the smallest possible d_n that will still maintain the network connected is $\Theta(\sqrt{\log n/n})$. Therefore, replacing for the optimal d_n, we finally get an upper bound of $\Theta\left(\sqrt{n/\log n}\right)$.

The Upper Bound is Asymptotically Tight

To verify that the upper bound is tight, we give an explicit flow construction. Consider the placement of disks shown in Fig. 6.

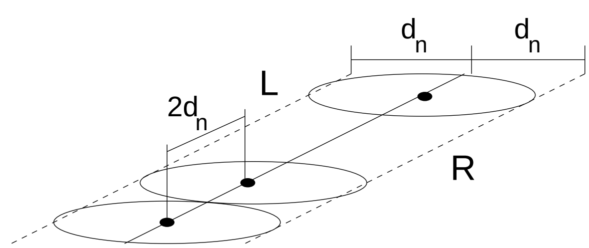

Figure 6: An explicit flow construction.

Since the height of the square is 1, and we are placing nodes at distance $2d_n$ from each other, this guarantees that *if there are nodes in each of the circles to create valid tx/rx pairs*, then the number of successful simultaneous transmissions across the cut is $\frac{1}{2d_n} = \Theta\left(\sqrt{n/\log n}\right)$. Whether all such pairs of nodes can be created simultaneously or not is the issue addressed next.

2.3 Uniform Convergence Issues

Next we prove that when n points are dropped uniformly over the square $[0, 1] \times [0, 1]$, we have that simultaneously (i.e., uniformly) over all $\frac{1}{2d_n}$ circles from Fig. 6, each one of the circles contains $\Theta(\log(n))$ points in almost all network realizations. From this, we conclude that the distribution of the number of edges across the cut is sharply concentrated around its mean, and hence that in a randomly chosen network, with probability approaching 1 as $n \to \infty$, the actual number of straddling edges is indeed $\Theta\left(\sqrt{n/\log(n)}\right)$.

Statement of the Result

Consider we have $\frac{1}{2d_n}$ circles centered along the $x = \frac{1}{2}$ cut as shown in Fig. 6, with centers $y_j = (2j-1)d_n$, $j = 1 \ldots \frac{1}{2d_n}$ and radius d_n—and let the A_j's leading up to eqn. (3) be these circles, and the N_j's be the counts of points contained in these circles. Then, we have the following uniform convergence result:

PROPOSITION 1. *Define* $B_j := [|N_j - \pi \log n| < \delta\pi \log n]$. *Then, as* $n \to \infty$, *and for any* $\delta \in (x, 1)$ $(x \approx 0.6)$, *we have that*

$$\lim_{n\to\infty} P\left[\bigcap_{j=1}^{\sqrt{\frac{n}{\log n}}} B_j\right] = 1.$$

Essentially what this proposition says is that with very high probability and uniformly over j, all A_j's contain $\Theta(\log n)$ nodes.

Proof

Note that $|A_j| = \pi d_n^2 = \pi \frac{\log n}{n}$. Then, invoking eqn. (3), we have that for any $0 < \delta < 1$ we can find a $\theta > 0$ such that

$$P[|N_j - \pi \log n| > \delta\pi \log n] < e^{-\theta\pi \log n} = n^{-\theta\pi}. \qquad (4)$$

Thus, we can conclude that the probability that the values of the random variable N_j deviate by a constant factor from the mean tends to zero as $n \to \infty$. This is a key step in showing that all the events $B_j := [|N_j - \pi \log(n)| < \delta\pi \log(n)]$ occur *simultaneously*, i.e., that we have uniform convergence of the N_j's to their expected values. Now, from the union bound, we have that

$$P\left[\bigcap_{j=1}^{\frac{1}{2d_n}} B_j\right] = 1 - P\left[\bigcup_{j=1}^{\frac{1}{2d_n}} B_j^c\right] \geq 1 - \sum_{j=1}^{\frac{1}{2d_n}} P[B_j^c].$$

But, from eqn. (4), $P[B_j^c] < n^{-\theta\pi}$, and therefore,

$$\sum_{j=1}^{\frac{1}{2d_n}} P[B_j^c] < \sum_{j=1}^{\frac{1}{2d_n}} n^{-\theta\pi} = \frac{n^{-\theta\pi}}{2d_n} = \frac{n^{\frac{1}{2}-\pi\theta}}{2\sqrt{\log n}}$$

Putting everything together, and letting $n \to \infty$, we have

$$P\left[\bigcap_{j=1}^{\frac{1}{2d_n}} B_j\right] \geq 1 - \frac{n^{\frac{1}{2}-\pi\theta}}{2\sqrt{\log n}} \longrightarrow 1,$$

if and only if $\pi\theta > \frac{1}{2}$. And this is true for $\delta \approx 0.6$ and above (this follows from the definition of θ and a simple numerical evaluation).

2.4 Remarks

We would like to conclude this section on omnidirectional antennas with a couple of remarks on the results presented so far.

One deals with the *simplicity* of our arguments: in this section we first computed the average number of edges that cross a cut, then we showed that in any network realization the actual number of nodes is very close to the ensemble mean. In previous work [12], similar results had been obtained based essentially on generalizations of the Glivenko-Cantelli lemma (that adds uniformity to convergence in the law of large numbers), due to Vapnik and Chervonenkis [16, Ch. 2], [22]. Our proof only makes use of simple and much better known results, such as Chernoff bounds.

Another remark is about the fact that indeed, the restriction in the optimization domain introduced to turn the multicommodity flow problem into a regular flow problem is not trivial. We have shown that, on this restricted class of networks (senders on the left, receivers on the right), with omnidirectional antennas the MST we obtain coincides with that of [12]. This suggests that, at least in this case, our restricted model captures well essential aspects of this problem. Of course, to be able to claim an alternative derivation of the results in [12], we need to establish the equivalence of the two linear programs. This is an issue currently under investigation.

3. A SINGLE DIRECTED BEAM

3.1 Transmitter/Receiver Model

In this section we consider the first model based on directional antennas: transmitters can generate a beam of arbitrarily narrow width aimed at any particular receiver, and receivers can accept any number of incoming messages, provided the transmitters are not in the same straight line. This results in a significant increase in the complexity of the signal processing algorithms required at each node, and in this section our goal is to determine if and how much it is possible to increase the achievable MST, compared to the omnidirectional case. This model is illustrated in Fig. 7.

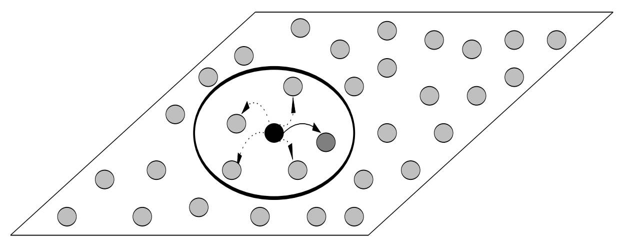

Figure 7: A single beam model for communication between nodes.

3.2 Average Number of Edges Across the Cut

Since at most one edge per transmitter can be active at any point in time, the average number of edges going across the cut can be no larger than nd_n, the average number of transmitters on its left side. Since L and R have the same area, the average number of nodes on each side of the cut is the same (and equal to nd_n), and hence the maximum of nd_n transmissions can actually be received, by "pairing up" every node from one side of the cut with every node on the other side. The pairing of nodes on each side of the cut is illustrated in Fig. 8.

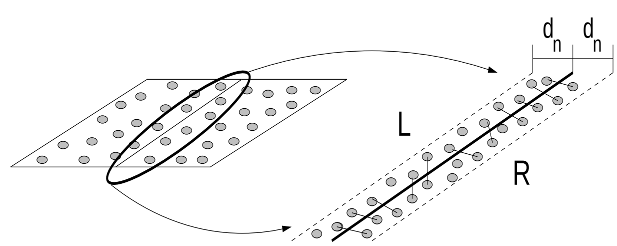

Figure 8: Pairing up one transmitter in L with one receiver in R: at most $n|L| = n|R| = nd_n$ such pairs can be formed.

Finally we note that, under the assumption of arbitrarily narrow and perfectly aligned beams, the only way in which we could have multiple receivers blocked out by a single transmission is by having them all lying in a nearly straight line (i.e., a set of vanishing measure) under the beam of a single transmitter. But then, to have an actual edge count lower than $\Theta(nd_n)$, we would require an increasingly large number of nodes falling in a decreasingly small area: under our statistical model for node placement, this event occurs with vanishing probability, and therefore the average edge count is $\Theta(nd_n)$.

3.3 Sharp Concentration Results

Number of Transmitters in L and Receivers in R

Again, consider n points $X_1 ... X_n$ uniformly distributed over the $[0,1] \times [0,1]$ plane, and consider the area L on the left side of the cut, as shown in Fig. 8. We define variables

$$N_i = \begin{cases} 1, & X_i \in L \\ 0, & \text{otherwise.} \end{cases}$$

and $N = \sum_{i=1}^{n} N_i$. The probability p of $X_i \in L$ is $p = |L| = 1 \cdot d_n$. Hence, $E(N_i) = 1 \cdot p + 0 \cdot (1-p) = p = d_n$, and $E(N) = \sum_{i=1}^{n} E(N_i) = nd_n$. From eqn. (3), we know that

$$P\left(|N - nd_n| > \delta nd_n\right) < e^{-\theta nd_n}.$$

Since $\theta > 0$, we have that as $n \to \infty$, deviations of N from its mean by a constant fraction (independent of n) occur with low probability, provided d_n does not decay too fast. Therefore, we conclude that in almost all realizations of the network, the number of transmitters in L and the number of receivers in R is $\Theta(nd_n)$.

Number of Edges Across the Cut

Knowing that we have $\Theta(nd_n)$ transmitters and receivers within range of each other on each side of the cut is not enough to claim that the number of edges that cross the cut is $\Theta(nd_n)$. This is because, in our model for directional antennas, a receiver can successfully decode two simultaneous incoming transmissions provided the angle formed by the receiver and the two transmitters is strictly positive: if all three are on the same straight line, collisions still occur, and those edges are destroyed. Therefore, we still need to show that the actual number of edges is $\Theta(nd_n)$. And to do this, we need to say something about the location of points that end up in L, and not just count how many. To proceed, we cut the area of L into nd_n rectangles of height $\frac{1}{nd_n}$ and width d_n, as illustrated in Fig. 9. Our goal then becomes to show that in "most" of these rectangles (meaning, in all but a constant fraction of them) we will have nodes capable of forming straddling edges.

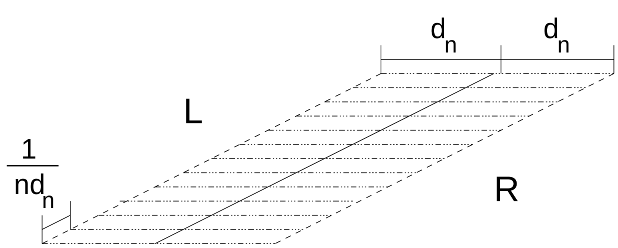

Figure 9: Cutting L and R into rectangles of size $d_n \times \frac{1}{nd_n}$.

Counting how many of the nd_n rectangles in Fig. 9 contain at least one of the $\Theta(nd_n)$ nodes that are dropped in L is an instance of a classical *occupancy* problem, in which k balls are thrown uniformly onto m bins, in the case where $k = m = nd_n$ [15, Ch. 4]. Since $\frac{1}{m}$ is the probability that a ball falls in any particular bin, the probability p of an empty bin after throwing all m balls is $p = (1 - \frac{1}{m})^m$ which, for m large, becomes approximately $\frac{1}{e}$. Therefore, the average number of empty bins is $mp \approx \frac{1}{e}\sqrt{n \log(n)}$. And by eqn. (3) again, we have that

$$P\left(Y - nd_n/e > \delta nd_n/e\right) < e^{-\theta nd_n/e},$$

where Y is the number of empty bins. So, the probability that the number of empty bins is a constant factor away from its mean is small (again, provided d_n does not decay too fast), and hence, for n large, almost all network realizations will have $\Theta(nd_n)$ non-empty rectangles. But since transmitter/receiver pairs in different rectangles are not collinear, the number of edges across the cut is $\Theta(nd_n)$, qed.

3.4 Remarks

MST in a Minimally Connected Network

In this section, we found that the MST achievable by the type of tx/rx pairs considered here depends on the connectivity radius d_n. If we replace d_n with $\frac{1}{\pi}\sqrt{\frac{\log n}{n}}$ (the minimum radius of eqn. (1), from [11]), we get

$$nd_n \approx n\frac{1}{\pi}\sqrt{\frac{\log n}{n}} = \Theta\left(\sqrt{n \log n}\right).$$

Comparing this expression with its equivalent from Section 2, we see that all we gain over the case of omnidirectional antennas is an increase in MST by a factor of $\Theta(\log n)$.

Minimum Connectivity Radius Resuting in MST = $\Theta(n)$

In this tx/rx architecture we are considering the use of arbitrarily narrow and perfectly aligned directed beams. Therefore, it does make sense to consider the use of a possibly larger transmission range than the minimum required to keep the network connected, since in this case a large range does not force other tx/rx pairs to remain silent while a given transmission is in progress. And since by increasing the transmission range now we can increase throughput, our next goal is to determine the minimum range that would be required to achieve MST = $\Theta(n)$.

Solving for d_n in $\Theta(n) = \Theta(nd_n)$, we see that trivially, $d_n = \Theta(1)$. That is, to achieve MST linear in the number of nodes using a single beam in each transmission, the radius of each transmission has to be a constant independent of n.

Minimum Number of Simultaneous Beams

From a practical point of view, does it matter that to achieve linear MST we need to keep the transmission radius constant? In this section we argue that yes it does, very much. To see why this is so, next we count the minimum number γ of narrow beams that a transmitter would have to generate simultaneously, if MST linear in the size of the network is to be achieved: this number gives a measure of the *complexity* of the beamforming transmitter, since $2\pi/\gamma$ is an upper bound on the maximum angle of dispersion of the beam.

Since a node can generate a beam to any receiver within its transmission range (see Fig. 7), again using eqns. (2) and (3), we have that for n large, the number of points within a circle of radius d_n is $\Theta(n \cdot \pi d_n^2)$. In the case of d_n only satisfying the requirement of keeping the network connected,

$$\gamma = n \cdot \pi d_n^2 = n\left(\frac{\pi \log n}{n}\right) = \Theta(\log n).$$

This fact was known already—see [26] for a more complete analysis (constants hidden by the Θ-notation included), including also a number of interesting references on the history of this problem.

But if now we consider a larger d_n satisfying the requirement of achieving linear MST, then

$$\gamma = n \cdot \Theta(1)^2 = \Theta(n).$$

Therefore, we see γ has an *exponential* increase relative to the number required to maintain minimum connectivity—it is on this fact that we base our claim about directional antennas not being able to provide an effective means of overcoming the issue with per-node vanishing throughputs.

4. MULTIPLE DIRECTED BEAMS

4.1 Transmitter/Receiver Model

In this section we consider another model based on directional antennas: transmitters can generate an arbitrary number of beams, of arbitrarily narrow width, aimed at any particular receiver; and receivers can accept any number of incoming messages, provided the transmitters are not in the same straight line. This is perhaps the most complex scheme that could be envisioned based on directed beams. Our goal is to determine if and how much it is possible to increase the achievable MST, compared to the previous two cases. This model is illustrated in Fig. 10.

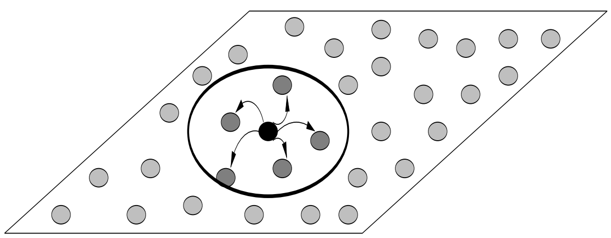

Figure 10: A non-degraded broadcast channel model for communication between nodes: each node is able to send simultaneously a different packet to each one of the nodes within his transmission range. Furthermore, multiple broadcasts (from different transmitters) do not collide, unless the transmitters are perfectly aligned.

4.2 Average Number of Edges Across the Cut

Fix a particular transmitter on the left side of the cut. The number of edges that cross the cut for that one transmitter is exactly the number of nodes in the right side of the cut that can receive this transmission. Therefore, for an arbitrary point $p = (x, y)$ in $L = [\frac{1}{2} - d_n, \frac{1}{2}) \times [0, 1]$, we draw a circle of radius d_n and center (x, y). The points $q = (u, v)$ in $R = [\frac{1}{2}, \frac{1}{2} + d_n] \times [0, 1]$ that are inside the circle are equal to the number of edges we want to count. These points p and q for which an edge exists satisfy the following conditions: (1) $\frac{1}{2} - d_n \leq x \leq \frac{1}{2}$; (2) either (a) $0 \leq y \leq 1$, or (b) $d_n \leq y \leq 1 - d_n$; (3) $\frac{1}{2} < u$; and (4) $(u - x)^2 + (v - y)^2 \leq d_n^2$. The situation is illustrated in Fig. 11.

Number of Receivers per Transmitter

For each $p = (x, y)$, we get the average number of points $q = (u, v)$ within the shaded arc Q_p in Fig. 11 using eqn. (2): E(*Number of points in* Q_p) $= n|Q_p|$.

To compute the area of Q_p (denoted $|Q_p|$), we let ϑ denote the angle of the arc illustrated in Fig. 11. Then, it follows from elementary trigonometric identities that $\sin \frac{\pi - \vartheta}{2} = \frac{\frac{1}{2} - x}{d_n}$, and so $\cos \frac{\vartheta}{2} = \frac{\frac{1}{2} - x}{d_n}$. So, $|Q_p| = \frac{1}{2}\vartheta d_n^2 - \frac{1}{2} d_n \cos\frac{\vartheta}{2} 2 d_n \sin\frac{\vartheta}{2} = \frac{1}{2}\vartheta d_n^2 - \frac{1}{2} d_n^2 \sin\vartheta = \frac{1}{2} d_n^2(\vartheta - \sin\vartheta)$. And plugging this expression into $n|Q_p|$, we get $n|Q_p| = n\frac{1}{2} d_n^2(\vartheta - \sin\vartheta)$. But

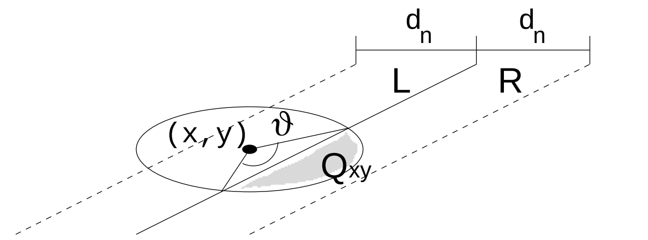

Figure 11: To illustrate constraints on edges. A transmitter at location $p = (x, y)$ can reach receivers within a circle of radius d_n—those on the other side of the cut must lie within the shaded area Q_p. ϑ is the angle of the shaded arc.

now, from the trigonometric identities above, we have that $\vartheta = 2 \arccos \frac{\frac{1}{2} - x}{d_n}$ and hence, $\sin\vartheta = 2 \sin\frac{\vartheta}{2} \cos\frac{\vartheta}{2}$, which implies $\sin^2\vartheta = 4 \sin^2\frac{\vartheta}{2} \cos^2\frac{\vartheta}{2}$, which again implies $\sin^2\vartheta = 4(1 - \cos^2\frac{\vartheta}{2})\cos^2\frac{\vartheta}{2}$. Now, since $0 \leq \vartheta \leq \pi$, $\sin\vartheta = 2\cos\frac{\vartheta}{2}\sqrt{1 - \cos^2\frac{\vartheta}{2}} \geq 0$, and so, finally, we get an expression for $n|Q_p|$ in terms of n, d_n, and the coordinates of the transmitter $p = (x, y)$:

$$\begin{aligned} n|Q_p| &= \frac{1}{2} n d_n^2 (\vartheta - \sin\vartheta) \\ &= \frac{1}{2} n d_n^2 \left(2 \arccos \frac{\frac{1}{2} - x}{d_n} - 2\cos\frac{\vartheta}{2}\sqrt{1 - \cos^2\frac{\vartheta}{2}} \right) \\ &= n d_n^2 \left(\arccos \frac{\frac{1}{2} - x}{d_n} - \frac{\frac{1}{2} - x}{d_n} \sqrt{1 - \frac{(\frac{1}{2} - x)^2}{d_n^2}} \right). \end{aligned}$$

Total Number of Edges

The result above is the average number of edges that cross the cut, starting at a fixed point $p = (x, y)$ in L. To calculate the total number of edges S that cross the cut on average, we need to add up $n|Q_p|$ over all transmitters p, (i.e., compute $S = \sum_{p \in L} n|Q_p|$). And our plan to do this is to approximate this sum by an integral.

The value of $|Q_p|$ is clearly dependent on the location of p: for p's in L near the boundary of the cut ($x <\approx \frac{1}{2}$), $\vartheta \approx \pi$ and hence the shaded area is large; for p's still in L but far from the boundary of the cut ($x >\approx \frac{1}{2} - d_n$), $\vartheta \approx 0$ and hence the shaded area is small. Furthermore, except near the top and bottom boundaries, the area of Q_p is independent of y. Therefore, to obtain a simple expression for the sought sum, our first step consists of dividing L into $\frac{d_n}{\Delta}$ thin strips of height 1 and width Δ (for $\Delta \ll d_n$), and expanding $\sum_{p \in L} n|Q_p|$ in two different ways:

$$S_a = \sum_{k=1}^{d_n/\Delta} n|Q_{xy}| \cdot \underbrace{|\{p = (x, y) \in L : 0 \leq y \leq 1\}|}_{s_a};$$

$$S_b = \sum_{k=1}^{d_n/\Delta} n|Q_{xy}| \cdot \underbrace{|\{p = (x, y) \in L : d_n \leq y \leq 1 - d_n\}|}_{s_b};$$

(in both cases, we take $\frac{1}{2} - d_n + (k-1)\Delta \leq x \leq \frac{1}{2} - d_n + k\Delta$). S_a is an upper bound on $\sum_{p \in L} n|Q_p|$, since we may count edges that end up outside the network; S_b is a lower bound, since we may not count some valid edges close to the network boundary; but as long as $d_n \to 0$ as $n \to \infty$, both bounds become tight and equal to $\sum_{p \in L} n|Q_p|$.

The next step is to observe that once again we can approximate the size estimates s_a and s_b using eqn. (2): $s_a = n\Delta$ and $s_b =$

$n(1-2d_n)\Delta$. Hence we get:

$$\begin{aligned} S_a &= \sum_{k=1}^{d_n/\Delta} n|Q_{xy}| \cdot n\Delta = n^2\Delta \sum_{k=1}^{d_n/\Delta} |Q_{xy}| \\ &\approx n^2 \int_{x=\frac{1}{2}-d_n}^{\frac{1}{2}} \int_{y=0}^{1} |Q_{xy}| \mathrm{d}x\mathrm{d}y; \\ S_b &= \sum_{k=1}^{d_n/\Delta} n|Q_{xy}| \cdot n(1-2d_n)\Delta = n^2(1-2d_n)\Delta \sum_{k=1}^{d_n/\Delta} |Q_{xy}| \\ &\approx n^2 \int_{x=\frac{1}{2}-d_n}^{\frac{1}{2}} \int_{y=d_n}^{1-d_n} |Q_{xy}| \mathrm{d}x\mathrm{d}y, \end{aligned}$$

since $\Delta \sum_{k=1}^{d_n/\Delta} |Q_{xy}|$ is a Riemann sum that, as we let $\Delta \to 0$, converges to the integral over an appropriate region of $|Q_p|$.

And now we are almost done. Since $S_b \leq \sum_{p \in L} n|Q_p| \leq S_a$, and we have that for n large, $S_a \approx S_b \approx n^2 \int_L |Q_p|$, we finally get:

$$\begin{aligned} \sum_{p\in L} n|Q_p| &\approx n^2 \int_L |Q_p| \mathrm{d}p \\ &= n^2 \int_{\frac{1}{2}-d_n}^{\frac{1}{2}} \int_0^1 d_n^2 [\arccos \frac{\frac{1}{2}-x}{d_n} \\ &\quad - \frac{\frac{1}{2}-x}{d_n} \sqrt{1 - \frac{(\frac{1}{2}-x)^2}{d_n^2}}] dydx \\ &= n^2 d_n^2 \int_{\frac{1}{2}-d_n}^{\frac{1}{2}} \int_0^1 \arccos \frac{\frac{1}{2}-x}{d_n} dydx \\ &\quad - n^2 d_n^2 \int_{\frac{1}{2}-d_n}^{\frac{1}{2}} \int_0^1 \frac{\frac{1}{2}-x}{d_n} \sqrt{1 - \frac{(\frac{1}{2}-x)^2}{d_n^2}} dydx \\ &= n^2 d_n^2 \int_{\frac{1}{2}-d_n}^{\frac{1}{2}} \arccos \frac{\frac{1}{2}-x}{d_n} dx \\ &\quad - n^2 d_n^2 \int_{\frac{1}{2}-d_n}^{\frac{1}{2}} \frac{\frac{1}{2}-x}{d_n} \sqrt{1 - \frac{(\frac{1}{2}-x)^2}{d_n^2}} dx \\ &\stackrel{(a)}{=} -n^2 d_n^3 \int_1^0 \arccos u du + n^2 \int_1^0 u\sqrt{1-u^2} du \\ &= n^2 d_n^3 \int_0^1 \arccos u du - n^2 \int_0^1 u\sqrt{1-u^2} du \\ &= n^2 d_n^3 - \tfrac{1}{3} n^2 d_n^3 \\ &= \tfrac{2}{3} n^2 d_n^3, \end{aligned}$$

where (a) follows from the change of variable $\frac{\frac{1}{2}-x}{d_n} = u$.

4.3 Sharp Concentration Results

Our next goal is to show that the actual number of edges straddling the cut in any realization of the network is sharply concentrated around its mean. That is, in almost all networks, the number of edges across the cut is $\Theta(n^2 d_n^3)$,

Number of Receivers per Transmitter

Define a binary random variable N_{ij}, which takes the value 1 if the i-th node is within the transmission range of a node at coordinates (x_j, y_j) on the other side of the cut, as illustrated in Fig. 11:

$$N_{ij} = \begin{cases} 1, & X_i \in Q_{(x_j,y_j)} \\ 0, & \text{otherwise.} \end{cases}$$

Let p denote the probability that X_i is in $Q_{(x_j,y_j)}$ (i.e., that $N_{ij} = 1$). Then, $p = |Q_{(x_j,y_j)}| = \frac{1}{2}d_n^2(\vartheta - \sin(\vartheta))$, with $0 \leq \vartheta \leq \pi$ is as in Fig. 11. Therefore, defining κ_ϑ as $\frac{1}{2}(\vartheta - \sin(\vartheta))$, we have $p = |Q_{(x_j,y_j)}| = \kappa_\vartheta d_n^2 = \kappa_\vartheta \frac{\log n}{n}$.

Define $N_j = \sum_{i=1}^n N_{ij}$ as the number of points in $Q_{(x_j,y_j)}$. In this case, we have $E(N_j) = \sum_{i=1}^n N_{ij} = \sum_{i=1}^n p \cdot 1 + (1-p) \cdot 0 = np = \kappa_\vartheta \log(n)$. Now, again by eqn. (3), we have that

$$P(|N_j - \kappa_\vartheta \log(n)| > \delta \kappa_\vartheta \log(n)) < e^{-\theta\kappa_\vartheta \log(n)} = n^{-\theta\kappa_\vartheta},$$

for θ defined as in previous applications. As $n \to \infty$ this probability tends to zero, and therefore, in almost all network realizations, a transmitter on the left side of the cut will be able to reach $\Theta(\kappa_\vartheta \log(n))$ receivers on the right side.

Total Number of Edges

In a manner analogous to the situation discussed in Section 3, knowing that there are $\Theta(nd_n)$ transmitters on the left side of the cut, and that each transmitter can reach $\Theta(nd_n^2)$ receivers on the other side, is not enough to conclude that the total number of edges going across the cut must be $\Theta(n^2 d_n^3)$. This is because of our requirement that multiple transmitters not be perfectly aligned with a receiver for this receiver to decode all these messages simultaneously. Therefore, we still need to show that the actual number of edges is $\Theta(n^2 d_n^3)$. And to do this, we need to say something about the location of points in R that can be reached from L, and not just count how many. To proceed then, we cut the area of Q_p into $\kappa_\vartheta \log(n)$ slices, each slice of area $\frac{|Q_p|}{\kappa_\vartheta \log(n)} = \frac{1}{n}$, as illustrated in Fig. 12.

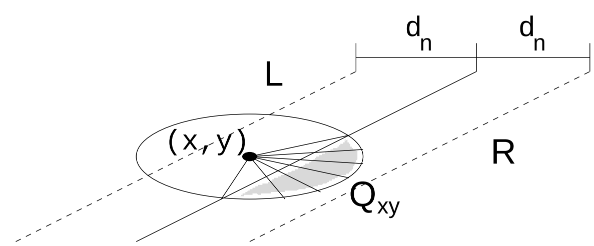

Figure 12: Cutting the shaded arc Q_{xy} into regions of area $\frac{1}{n}$, to formulate this as an occupancy problem analogous to that of Fig. 9.

As in the occupancy problem considered in Section 3, our goal is to show that in "most" of these arc slices (most meaning, in all but a constant fraction of them) we will have nodes capable of forming straddling edges. This is again a problem of throwing k balls uniformly into m bins, where $k = m = \kappa_\vartheta \log(n)$. And again, we have that with probability that tends to 1 as $n \to \infty$, the number of empty bins is $\kappa_\vartheta \log(n)/e$, and hence the number of occupied bins is $\Theta(\kappa_\vartheta \log(n))$.

Consider now a fixed transmitter located at some coordinates (x, y). Any other transmitter located at coordinates $(x', y') \neq (x, y)$ defines a unique straight line that goes through (x, y) and (x', y'). If there is a receiver on the other side of the cut along this line, within reach of both transmitters, then those two edges will be lost—and those will be the *only* lost edges, from among the $\kappa_\vartheta \log(n)$ that each transmitter has. This situation is illustrated in Fig. 13.

And then we are done. We have established that in almost all network realizations, there are $\Theta(nd_n)$ transmitters within each side of the cut, that each transmitter can reach $\Theta(\kappa_\vartheta d_n^2)$ receivers on the other side of the cut, and that integrating out κ_ϑ we obtain exactly $\Theta(n^2 d_n^3)$ edges going across the cut. Therefore, the actual

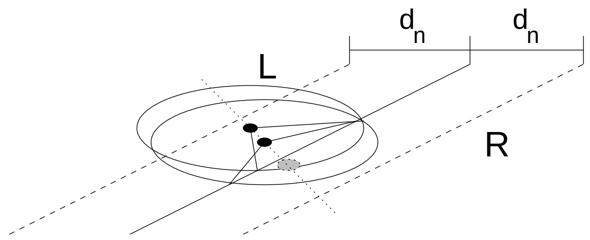

Figure 13: **To illustrate how we could end up losing edges: if the two black transmitters attempt simultaneously to communicate with the gray receiver, a collision will occur, and none of the edges will be created.**

number of edges across the cut is sharply concentrated around its mean, qed.

4.4 Remarks

MST in a Minimally Connected Network

Substituting for $d_n \approx \frac{1}{\sqrt{\pi}}\sqrt{\frac{\log n}{n}}$ from eqn. (1), in $\frac{2}{3}n^2 d_n^3$, we get

$$\frac{2}{3}n^2\left(\frac{\log n}{\pi n}\right)^{\frac{3}{2}} = \frac{2}{3}\sqrt{n}\log^{\frac{3}{2}} n = \Theta\left(\sqrt{n}\log^{\frac{3}{2}}(n)\right)$$

Comparing this expression to the ones obtained in Sections 2 and 3, we see that the MST gain due to the use of multiple simultaneous, arbitrarily narrow beam is, at most, $\Theta\left(\log^2(n)\right)$.

Minimum Connectivity Radius Resuting in MST = $\Theta(n)$

The minimum d_n resulting in linear MST is obtained by solving for d_n in $\Theta(n^2 d_n^3) = \Theta(n)$. Now, for n large enough, there exist constants $c_1 < c_2 \in \mathbb{R}$ ($c_1 > 0$ and $c_2 < \infty$), such that $c_1 n < \frac{2}{3}n^2 d_n^3 < c_2 n$, or equivalently, $c_1 \frac{3}{2} n^{-\frac{1}{3}} < d_n < c_2 \frac{3}{2} n^{-\frac{1}{3}}$. Therefore,

$$d_n = \Theta(n^{-\frac{1}{3}}).$$

Minimum Number of Simultaneous Beams

In Section 3, we said that keeping the transmission range constant resulted in an impractically large number of beams that the receiver needed to generate, if linear MST was to be achieved by increasing the complexity of the signal processing algorithms. But if we generate multiple beams, we have just shown that this minimum radius now is no longer a constant, but instead tends to zero as $\Theta(n^{-\frac{1}{3}})$. However, the situation is not much better compared to the single beam case, and to see this again we compute the minimum number of simultaneous beams that a transmitter would have to generate.

If now we consider the larger d_n satisfying the requirement of achieving maximum stable throughput linear in network size, then

$$\gamma = n \cdot \pi d_n^2 = n\Theta(n^{-\frac{2}{3}}) = \Theta(n^{\frac{1}{3}}).$$

Therefore, we see that while γ is smaller than in the case of the single beam, we still have an *exponential* increase relative to the number required to maintain minimum connectivity—so again, we claim that directional antennas are not able to provide an effective means of overcoming the issue with per-node vanishing throughputs.

5. CONCLUSIONS

In this paper we have presented our first results on the analysis of optimal scaling laws for wireless networks. We first formulated this problem as one of finding the maximum value of a suitably defined multicommodity flow on random unit-disk graphs. Then, due to the complexity of analyzing the multicommodity flow problem, we looked at a subset of all the possible networks of interest, in which this problem becomes a regular maximum flow problem. Working on this restricted class of networks, in one special case we were able to show that the optimal scaling laws for this restriction are identical to the optimal scaling laws without the restriction (the case of omnidirectional transmissions studied by [12]), thus proving the fact that there is no loss of optimality by considering this restriction.

The next step in this work is to formally prove that there is no loss of optimality, by working directly with the linear programs that characterize the solutions of our multicommodity flow problems—that is the current focus of our work. Future work will deal with extending the analysis to a richer class of combinatorial structures, to deal with transmitter/receiver architectures other than directional antennas.

Acknowledgements—The authors would like to thank two of the anonymous reviewers, for unusually useful feedback. This paper reads significantly better now than the abstract submitted for review did, and those improvements came from working on suggestions made primarily by reviewer 4, and also by reviewer 1. The authors would also like to thank T. L. Fine, for pointing them to the text of Pollard [16], a reference that proved to be extremely useful in the process of developing some basic understanding on what VC theory is about [22].

6. REFERENCES

[1] S. Adireddy and L. Tong. Exploiting Decentralized Channel State Information for Random Access. Submitted for publication. Available from `http://acsp.ece.cornell.edu/`.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network Information Flow. *IEEE Trans. Inform. Theory*, 46(4):1204–1216, 2000.

[3] R. R. Choudhury, X. Yang, R. Ramanathan, and N. Vaidya. Medium Access Control in Ad Hoc Networks Using Directional Antennas. In *Proc. ACM MobiCom*, Atlanta, GA, 2002.

[4] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms (2nd ed)*. MIT Press, 2001.

[5] T. M. Cover and J. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., 1991.

[6] L. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, 1962.

[7] S. Ghez, S. Verdú, and S. Schwartz. Optimal Decentralized Control in the Random Access Multipacket Channel. *IEEE Trans. Autom. Control*, 34(11):1153–1163, 1989.

[8] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics—A Foundation for Computer Science*. Addison Wesley, 1994.

[9] M. Grossglauser and D. Tse. Mobility Increases the Capacity of AdHoc Wireless Networks. *IEEE Trans. Networking*, 10(4):477–486, 2002.

[10] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer Verlag, 1988.

[11] P. Gupta and P. R. Kumar. Critical Power for Asymptotic Connectivity in Wireless Networks. In W. M. McEneany, G. Yin, and Q. Zhang, editors, *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming*. Birkhauser, 1998.

[12] P. Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Trans. Inform. Theory*, 46(2):388–404, 2000.

[13] M. V. Marathe, H. Breu, H. B. Hunt III, S. S. Ravi, and D. J. Rosenkrantz. Simple Heuristics for Unit Disk Graphs. *Networks*, 25:59–68, 1995.

[14] G. Mergen and L. Tong. On the Capacity of Regular Wireless Networks with Transceiver Multipacket Communication. In *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Lausanne, Switzerland, 2002.

[15] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[16] D. Pollard. *Convergence of Stochastic Processes*. Springer-Verlag, 1984.

[17] A. Scaglione and S. D. Servetto. On the Interdependence of Routing and Data Compression in Multi-Hop Sensor Networks. In *Proc. ACM MobiCom*, Atlanta, GA, 2002.

[18] S. D. Servetto. Quantization with Side Information: Lattice Codes, Asymptotics, and Applications in Wireless Networks. Submitted to the IEEE Trans. Inform. Theory. Available from `http://people.ece.cornell.edu/servetto/`.

[19] L. Tong, Q. Zhao, and G. Mergen. Multipacket Reception in Random Access Wireless Networks: From Signal Processing to Optimal Medium Access Control. *IEEE Communications Magazine*, 39(11):108–112, 2001.

[20] S. Toumpis and A. J. Goldsmith. Capacity Regions for Wireless Adhoc Networks. IEEE Trans. Wireless Comm., to appear. Available from `http://wsl.stanford.edu/`.

[21] B. S. Tsybakov and V. A. Mikhailov. Ergodicity of a Slotted ALOHA System. *Problemy Peredachi Informatsii*, 15(4):301–312, 1979.

[22] V. N. Vapnik and A. Ya. Chervonenkis. On the Uniform Convergence of Relative Frequencies of Events to their Probabilities. *Theory of Probability and its Applications*, XVI(2):264–280, 1971.

[23] D. West. *Introduction to Graph Theory*. Prentice Hall, 1996.

[24] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides. Energy-Efficient Broadcast and Multicast Trees in Wireless Networks. *ACM/Kluwer Mobile Networks and Applications*, 7(6), 2002.

[25] L.-L. Xie and P. R. Kumar. A Network Information Theory for Wireless Communication: Scaling Laws and Optimal Operation. Submitted to the IEEE Trans. Inform. Theory, April 2002. Available from `http://decision.csl.uiuc.edu/~prkumar/`.

[26] F. Xue and P. R. Kumar. The Number of Neighbors Needed for Connectivity of Wireless Networks. Submitted to Wireless Networks, April 2002. Available from `http://decision.csl.uiuc.edu/~prkumar/`.

APPENDIX

A. BACKGROUND MATERIAL

In this appendix we present some basic definitions and results that make the paper self-contained. Most concepts come almost without changes from [4] or [23].

A.1 Random Unit-Disk Graphs

In this paper, wireless networks are modeled as random *unit disk* graphs. A graph is a unit-disk graph if and only if its vertices can be put in one-to-one correspondence with equisized circles in a plane, in such a way that two vertices are joined by an edge if and only if their corresponding circles intersect [23, Ch. 8].[3] In our model, the set of vertices is given by n nodes randomly placed on a square of unit area:

$$V = \{(x_i, y_i) \in [0,1] \times [0,1], i = 1...n\}.$$

The locations (x_i, y_i) are iid uniformly distributed pairs over $[0,1] \times [0,1]$, with joint distribution:

$$f_{XY}(x,y) = \begin{cases} 1 & \text{if } (x,y) \in [0,1] \times [0,1] \\ 0 & \text{otherwise} \end{cases}$$

The set of edges depends on a parameter d_n, that we will have the freedom to choose for different setups:

$$E = \{(u,v) \in V \times V : |u - v| \leq d_n\},$$

and all edges are assumed to have the same constant capacity $c(u,v) = L > 0$. In order for a node to successfully receive a transmission, the distance between them must be $\leq d_n$. In general, throughout this paper we will always assume n large. The resulting network model is illustrated in Fig. 14.

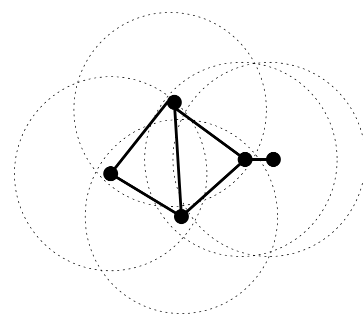

Figure 14: Randomly placed nodes, connected by edges whenever circles centered at each node intersect.

Note that some authors choose to define edges in these graphs when two circles have a non-empty intersection [23], whereas others choose to define connectivity when one node is contained in the circle of another node [12]. We argue that, for the purpose of this paper, both definitions are completely equivalent, and we can use either one of them. That is for the simple reason that the graphs obtained under both models are the same, with twice/half the radius of their circles. But for the asymptotics of interest in this work, changes of the transmission range by a constant factor (2 in this case) are of no consequence.

A.2 Flow Networks

Defintions

A *flow network* $G = (V, E)$ is a graph in which edges $(u,v) \in E$ have a capacity $c(u,v) \geq 0$ in each direction (and if $(u,v) \notin E$, then we assume $c(u,v) = 0$). We distinguish two vertices

[3] In his book [23], West defines *interval* graphs as intersection graphs of collections of intervals on the line—the extension to the plane is straightforward. Marathe et al. [13] have used the definition that we give up here in their work as well.

in the network, the *source* s and the *sink* t. A *flow* in G is a function $f : V \times V \to \mathbb{R}$, which satisfies three constraints: (a) $\forall u, v \in V,\ \ f(u,v) \leq c(u,v)$ (capacity constraint); (b) $\forall u, v \in V,\ \ f(u,v) = -f(v,u)$ (skew symmetry); and (c) $\forall u \in V - \{s,t\},\ \sum_{v \in V} f(u,v) = 0$ (flow conservation). If f is a flow on G, the *value* of f is $|f| = \sum_{u \in V} f(s,u)$. In the *maximum flow problem*, given a flow network G, the goal is to find a flow of maximum value.

A *cut* (S,T) in G is defined as a partition of V into two nonempty sets S and T, i.e., $S, T \neq \emptyset$ satisfy $S \cap T = \emptyset$ and $S \cup T = V$. An edge of G is said to *cross* the cut (S,T) if its two endpoints are on different sides of the cut, i.e., $e = (u,v) \in E$ crosses the cut if $u \in S$ and $v \in T$ (or equivalently, $u \in T$ and $v \in S$). The *capacity* $c(S,T)$ of the cut is the sum of the capacities of all the edges that cross it. Since in our model we assumed $c(u,v) = L$ for all $(u,v) \in E$, for us the capacity of a cut will simplify to $L|(S,T)|$, the number of edges going across the cut times L. If (S,T) is a cut of G, then the net flow across (S,T) is $f(S,T) = |f|$.

For problems like ours, in which there are multiple sources and multiple sinks involved, there is a standard trick to reduce this more general version to the case with a single source and a single sink. We define a new graph $G' = (V', E')$: V' contains the same vertices of V plus two extra nodes, that we refer to as the *supersource* s and the *supersink* t; E' contains the same edges as E plus, for each source s_i and for each sink t_j, edges of the form (s, s_i) and (t_j, t), with $c(s, s_i) = c(t_j, t) = \infty$. Then, it is easy to show that the value of a maximum flow in G' is the same as the value of a maximum flow in G.

The Max-Flow/Min-Cut Theorem

Our interest in flow networks stems from the fact that the notion of a maximum flow is essentially the same as that of maximum stable throughput: in computing the value of a maximum flow, we seek to determine what is the largest amount of flow that can be carried by the network, without violating either link capacity or packet conservation constraints. Therefore, we need to give conditions under which we can make statements about the values of such flows.

In general, we have that the value of *any* flow f from the supersource s to the supersink t in G is bounded from above by the capacity of *any* cut of G for which $s \in S$ and $t \in T$. Indeed, it follows in a straightforward manner from the definitions that

$$|f| = f(S,T) = \sum_{u \in S} \sum_{v \in T} f(u,v) \leq \sum_{u \in S} \sum_{v \in T} c(u,v) = c(S,T).$$

And this gives us a general technique to obtain both upper and lower bounds on the value of a maximum flow:

- To get lower bounds, construct *any* feasible flow in G.
- To get upper bounds, find the capacity of *any* valid cut in G.

How good are the bounds thus obtained? It follows from a classical theorem in the analysis of flow networks that said bounds are tight: according to the *max-flow/min-cut theorem* [4], f is a flow of maximum value *iff* $|f| = c(S,T)$ (for some cut (S,T)).

A.3 Multicommodity Flows

So far we have considered flows "of a single thing": that is, there is one commodity (packets in our problem), of which we want to move the maximum quantity from s to t. A more general version of this problem deals with flows of *multiple* different commodities. In this case, we still have a flow network G with capacity function c, and we also have k different commodities. The i-th commodity is described by (s_i, r_i, d_i): s_i is the source where this commodity is produced, r_i is the sink where it is consumed, and d_i is the *demand* at r_i. We define k flows f_i as in the standard flow case, and then define the *aggregate* flow f as $f(u,v) = \sum_{i=1}^{k} f_i(u,v)$—the aggregate flow is constrained to not exceed the capacity of any edge. In this formulation, there is nothing to maximize: the question of interest is whether the given network is able to support the flow of d_i units of the i-th commodity from s_i to r_i, for $i = 1...k$.

Multicommodity flow problems appear to be simple generalizations of flow problems described above, but they are not—the sophistication in the level of mathematics required to analyze such problems is a notch up compared to single commodity flow problems. For example: whereas there is a simple characterization of maximum flows in terms of cuts (the max-flow/min-cut theorem mentioned above), there is no such thing for maximum multicommodity flows. In fact, this problem is NP-hard, and the only known polynomial time algorithm for answering the question of whether a particular set of commodities can be supported by a given flow network consists of formulating this decision problem as a linear program, and using a polynomial time solver for LPs [4, Ch. 29].