The Capacity of Wireless Networks

Piyush Gupta, *Student Member, IEEE*, and P. R. Kumar, *Fellow, IEEE*

***Abstract*—When n identical randomly located nodes, each capable of transmitting at W bits per second and using a fixed range, form a wireless network, the throughput $\lambda(n)$ obtainable by each node for a randomly chosen destination is $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ bits per second under a noninterference protocol.**

If the nodes are optimally placed in a disk of unit area, traffic patterns are optimally assigned, and each transmission's range is optimally chosen, the bit–distance product that can be transported by the network per second is $\Theta(W\sqrt{An})$ bit-meters per second. Thus even under optimal circumstances, the throughput is only $\Theta\left(\frac{W}{\sqrt{n}}\right)$ bits per second for each node for a destination nonvanishingly far away.

Similar results also hold under an alternate physical model where a required signal-to-interference ratio is specified for successful receptions.

Fundamentally, it is the need for every node all over the domain to share whatever portion of the channel it is utilizing with nodes in its local neighborhood that is the reason for the constriction in capacity.

Splitting the channel into several subchannels does not change any of the results.

Some implications may be worth considering by designers. Since the throughput furnished to each user diminishes to zero as the number of users is increased, perhaps networks connecting smaller numbers of users, or featuring connections mostly with nearby neighbors, may be more likely to be find acceptance.

***Index Terms*—*Ad hoc* networks, capacity, multihop radio networks, throughput, wireless networks.**

I. INTRODUCTION

WIRELESS networks consist of a number of nodes which communicate with each other over a wireless channel. Some wireless networks have a wired backbone with only the last hop being wireless. Examples are cellular voice and data networks and mobile IP. In others, all links are wireless. One example of such networks is multihop radio networks or *ad hoc* networks. Another possibly futuristic example, see [1], may be collections of "smart homes" where computers, microwave ovens, door locks, water sprinklers, and other "information appliances" are interconnected by a wireless network.

It is to these types of all wireless networks that this paper is addressed. Such networks consist of a group of nodes which communicate with each other over a wireless channel without any centralized control; see Fig. 1. Nodes may cooperate in routing each others' data packets. Lack of any centralized control and possible node mobility give rise to many issues at the network, medium access, and physical layers, which have no counterparts in the wired networks like Internet, or in cellular networks.

At the network layer, the main problem is that of routing, which is exacerbated by the time-varying network topology, power constraints, and the characteristics of the wireless channel; see Ramanathan and Steenstrup [2] for an overview. The choice of medium access scheme is also difficult in *ad hoc* networks due to the time-varying network topology and the lack of centralized control. Use of TDMA or dynamic assignment of frequency bands is complex since there is no centralized control as in cellular networks, FDMA is inefficient in dense networks, CDMA is difficult to implement due to node mobility and the consequent need to keep track of the frequency-hopping patterns and/or spreading codes for nodes in the time-varying neighborhood, and random access appears to be the current favorite. The access problem when many nodes transmit to the same receiver has been much studied in the literature ever since the genesis of the ALOHA network, and bounds on the throughput of successful collision-free transmissions as well as transmission protocols have been devised; see Gallager [3]. Sharing channels in networks does lead to some new problems associated with "hidden" terminals and "exposed" terminals. The protocols MACA and its extension MACAW, see Karn [4] and Bhargavan *et al.* [5] respectively, use a series of handshake signals to resolve these problems to a certain extent. This has been standardized in the IEEE 802.11 protocol, see [6]. At the physical layer, an important issue is that of power control. The transmission power of nodes needs to be regulated so that it is high enough to reach the intended receiver while causing minimal interference at other nodes. Iterative power control algorithms have been devised, see Bambos, Chen, and Pottie [7] and Ulukus and Yates [8].

In this paper we analyze the capacity of wireless networks. We scale space and suppose that n nodes are located in a region of area 1 m^2. Each node can transmit at W bits per second over a common wireless channel. We shall see that it is immaterial

Manuscript received December 3, 1998; revised July 1, 1999. This material is based on work supported in part by the Air Force Office of Scientific Research under Contract AF-DC-5-36128, by EPRI and the U.S. Army Research Office under Subcontract to Cornell University Contracts WO8333-04 and 35352-6086, by the U.S. Army Research Office under Contract DAAH 04-95-1-0090, the Office of Naval Research under Contract N00014-99-1-0696, and the Joint Services Electronics Program under Contract N00014-96-1-0129. Any opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the above agencies.

P. Gupta is with the Department of Electrical and Computer Engineering, and the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801 USA.

P. R. Kumar is with the University of Illinois at Urbana-Champaign, Coordinated Science Laboratory, Urbana, IL 61801 USA.

Communicated by V. Anantharam, Associate Editor for Communication Networks.

Publisher Item Identifier S 0018-9448(00)01358-4.

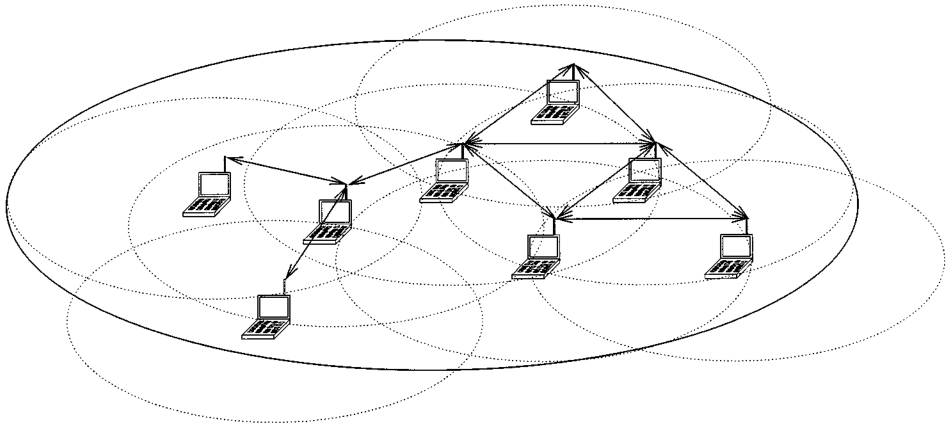

Fig. 1. An *ad hoc* wireless network.

to our results[1] if the channel is broken up into several subchannels of capacity $W_1, W_2, \cdots, W_M$ bits per second, as long as $\sum_{m=1}^{M} W_m = W$. Packets are sent from node to node in a multihop fashion until they reach their final destination. They can be buffered at intermediate nodes while awaiting transmission.

Due to spatial separation, several nodes can make wireless transmissions simultaneously, provided there is no destructive interference of a transmission by others. We will describe in the sequel under what conditions a wireless transmission over a subchannel is received successfully by its intended recipient.

We will consider two types of networks, *Arbitrary Networks*, where the node locations, destinations of sources, and traffic demands, are all arbitrary, and *Random Networks*, where the nodes and their destinations are randomly chosen.

A. Arbitrary Networks: Arbitrarily Located Nodes and Traffic Patterns

In the arbitrary setting we suppose that n nodes are arbitrarily located in a disk of unit area in the plane. Each node has an arbitrarily chosen destination to which it wishes to send traffic at an arbitrary rate; thus the traffic pattern is arbitrary. Each node can choose an arbitrary range or power level for each transmission.

We need to describe when a transmission is received successfully by its intended recipient. We will allow for two possible models for successful reception of a transmission over one hop, called the *Protocol Model* and the *Physical Model*, described below. Let X_i denote the location of a node; we will also use X_i to refer to the node itself.

1) The Protocol Model: Suppose node X_i transmits over the mth subchannel to a node X_j. Then this transmission is successfully received by node X_j if

$$|X_k - X_j| \geq (1+\Delta)|X_i - X_j| \tag{1}$$

for every other node X_k simultaneously transmitting over the same subchannel.

The quantity $\Delta > 0$ models situations where a guard zone is specified by the protocol to prevent a neighboring node from transmitting on the same subchannel at the same time. It also allows for imprecision in the achieved range of transmissions.

Another model which is more related to physical layer considerations is

2) The Physical Model: Let $\{X_k; k \in \mathcal{T}\}$ be the subset of nodes simultaneously transmitting at some time instant over a certain subchannel. Let P_k be the power level chosen by node X_k, for $k \in \mathcal{T}$. Then the transmission from a node X_i, $i \in \mathcal{T}$, is successfully received by a node X_j if

$$\frac{\frac{P_i}{|X_i - X_j|^\alpha}}{N + \sum\limits_{\substack{k \in \mathcal{T} \\ k \neq i}} \frac{P_k}{|X_k - X_j|^\alpha}} \geq \beta. \tag{2}$$

This models a situation where a minimum signal-to-interference ratio (SIR) of β is necessary for successful receptions, the ambient noise power level is N, and signal power decays with distance r as $\frac{1}{r^\alpha}$. We will suppose that $\alpha > 2$, which is the usual model outside a small neighborhood of the transmitter.

3) The Transport Capacity of Arbitrary Networks: Given any set of successful transmissions taking place over time and space, let us say that the network transports one *bit-meter* when one bit has been transported a distance of one meter toward its destination. (We do not give multiple credit for the same bit carried from one source to several different destinations as in the multicast or broadcast cases). This sum of products of bits and the distances over which they are carried is a valuable indicator of a network's *transport capacity*. (It should be noted that when the area of the domain is A square meters rather than the normalized 1 m^2, then all the transport capacity results presented below should be scaled by $\sqrt{A}$). Our main results are the following. Recall Knuth's notation: $f(n) = \Theta(g(n))$ denotes that $f(n) = O(g(n))$ as well as $g(n) = O(f(n))$.

Main Result 1.: The transport capacity of an Arbitrary Network under the Protocol Model is $\Theta(W\sqrt{n})$ bit-meters per second if the nodes are optimally placed, the traffic pattern is optimally chosen, and if the range of each transmission is chosen optimally.

Specifically, an upper bound is $\sqrt{\frac{8}{\pi}}\frac{W}{\Delta}\sqrt{n}$ bit-meters per second for every Arbitrary Network for all spatial and temporal

[1] We are grateful to Kimberly King for asking us to be more explicit about the prospects for routing through multiple technologies.

scheduling strategies, while $\frac{W}{1+2\Delta}\frac{n}{\sqrt{n}+\sqrt{8\pi}}$ bit-meters per second (for n a multiple of four) can be achieved when the nodes and traffic patterns are appropriately chosen, and the ranges and schedules of transmissions are appropriately chosen.

If this transport capacity were to be equitably divided between all the n nodes, then each node would obtain $\Theta\left(\frac{W}{\sqrt{n}}\right)$ bit-meters per second. If, further, each source has its destination about the same distance of 1 m away, then each node would obtain a *throughput capacity* of $\Theta\left(\frac{W}{\sqrt{n}}\right)$ bits per second.

The upper bound on transport capacity does not depend on the transmissions being omnidirectional, as implied by (1), but only on there being some dispersion in the neighborhood of the receiver; see Assumption (A.vi) in Section II.

Main Result 2: For the Physical Model, $cW\sqrt{n}$ bit-meters per second is feasible, while $c'Wn^{\alpha-1/\alpha}$ bit-meters per seconds is not, for appropriate c, c'. Specifically,

$$\frac{1}{\left(16\beta\left(2^{\frac{\alpha}{2}}+\frac{6^{\alpha-2}}{\alpha-2}\right)\right)^{\frac{1}{\alpha}}}\frac{Wn}{\sqrt{n}+\sqrt{8\pi}}$$

bit-meters per second (for n a multiple of 4) is feasible when the network is appropriately designed, while an upper bound is $\frac{1}{\sqrt{\pi}}\left(\frac{2\beta+2}{\beta}\right)^{\frac{1}{\alpha}}Wn^{\frac{\alpha-1}{\alpha}}$ bit-meters per second.

We suspect that an upper bound of order $\Theta(W\sqrt{n})$ bit-meters per second may actually hold. In the special case where the ratio $\frac{P_{\max}}{P_{\min}}$ between the maximum and minimum powers that transmitters can employ is bounded above by β, then an upper bound is in fact

$$\sqrt{\frac{8}{\pi}}\frac{1}{\left(\frac{\beta P_{\min}}{P_{\max}}\right)^{\frac{1}{\alpha}}-1}W\sqrt{n}$$

bit-meters per second.

It is worth noting that both bounds suggest that transport capacity improves when α is larger, i.e., when the signal power decays more rapidly with distance.

B. Random Networks: Randomly Located Nodes and Traffic Patterns

In a random scenario, n nodes are randomly located, i.e., independently and uniformly distributed, either on the surface S^2 of a three-dimensional sphere of area 1 m^2, or in a disk of area 1 m^2 in the plane. Our purpose in studying S^2 is to separate edge effects from other phenomena. Each node has a randomly chosen destination to which it wishes to send $\lambda(n)$ bits per second. The destination for each node is independently chosen as the node nearest to a randomly located point, i.e., uniformly and independently distributed. (Thus destinations are on the order of 1 m away on average.)

In this random setting, we will assume that the nodes are homogeneous, i.e., all transmissions employ the same nominal range or power. As for Arbitrary Networks, we will allow for both a Protocol Model as well as a Physical Model for interference.

1) The Protocol Model: All nodes employ a common *range* r for all their transmissions. When node X_i transmits to a node X_j over the mth subchannel, this transmission is successfully received by X_j if

i) The distance between X_i and X_j is no more than r, i.e.,

$$|X_i - X_j| \le r. \tag{3}$$

ii) For every other node X_k simultaneously transmitting over the same subchannel

$$|X_k - X_j| \ge (1+\Delta)r. \tag{4}$$

2) The Physical Model: All nodes choose a common power level P for all their transmissions. Let $\{X_k; k \in \mathcal{T}\}$ be the subset of nodes simultaneously transmitting at some time instant over a certain subchannel. A transmission from a node X_i, $i \in \mathcal{T}$, is successfully received by a node X_j if

$$\frac{\frac{P}{|X_i-X_j|^\alpha}}{N+\sum\limits_{\substack{k\in\mathcal{T}\\k\neq i}}\frac{P}{|X_k-X_j|^\alpha}} \ge \beta. \tag{5}$$

3) The Throughput Capacity of Random Networks: The notion of throughput is defined in the usual manner as the time average of the number of bits per second that can be transmitted by every node to its destination.

Definition: Feasible Throughput: A throughput of $\lambda(n)$ bits per second for each node is *feasible* if there is a spatial and temporal scheme for scheduling transmissions, such that by operating the network in a multihop fashion and buffering at intermediate nodes when awaiting transmission, every node can send $\lambda(n)$ bits per second on average to its chosen destination node. That is, there is a $T < \infty$ such that in every time interval $[(i-1)T, iT]$ every node can send $T\lambda(n)$ bits to its corresponding destination node.

Whether a particular throughput level is feasible may depend on the locations of the nodes. These locations are random. So is the destination for the traffic entering each node. As in PAC Learning Theory (see Valiant [9]), given the randomness involved in the problem statement, we allow for vanishingly small probabilities when defining the "throughput capacity."

Definition: The Throughput Capacity of Random Wireless Networks: We say that the *throughput capacity* of the class of Random Networks is of order $\Theta(f(n))$ bits per second if there are deterministic constants $c > 0$ and $c' < +\infty$ such that

$$\lim_{n\to\infty} \text{Prob}\,(\lambda(n) = cf(n) \text{ is feasible}) = 1$$
$$\liminf_{n\to\infty} \text{Prob}\,(\lambda(n) = c'f(n) \text{ is feasible}) < 1.$$

Our main results are the following.

Main Result 3.: In the case of both the surface of the sphere and a planar disk, the order of the throughput capacity is

$$\lambda(n) = \Theta\left(\frac{W}{\sqrt{n\log n}}\right)$$

bits per second for the Protocol Model. For the upper bound we actually prove the sharp cutoff phenomenon that for some c'

$$\lim_{n\to\infty} \text{Prob}\left(\lambda(n) = c'\frac{W}{\sqrt{n\log n}} \text{ is feasible}\right) = 0.$$

Specifically, there are deterministic constants c'' and c''' not depending on n, Δ, or W, such that

$$\lambda(n) = \frac{c''W}{(1+\Delta)^2\sqrt{n\log n}}$$

bits per second is feasible, and

$$\lambda(n) = \frac{c'''W}{\Delta^2\sqrt{n\log n}}$$

bits per second is infeasible, both with probability approaching one as $n \to \infty$. Since routing hot spots may form at the center in the case of a disk on the plane, and yet the order of throughput capacity is the same as on the surface of the sphere, it shows that the cause of the throughput constriction is not the formation of hot spots, but is the pervasive need for all nodes to share the channel locally with other nodes.

Main Result 4: For the Physical Model a throughput of $\lambda(n) = \frac{cW}{\sqrt{n\log n}}$ bits per second is feasible, while $\lambda(n) = \frac{c'W}{\sqrt{n}}$ bits per second is not, for appropriate c, c', both with probability approaching one as $n \to \infty$. Specifically, there are deterministic constants c'' and c''' not depending on n, N, α, β, or W, such that

$$\lambda(n) = \frac{c''}{\left(2\left(c'''\beta\left(3+\frac{1}{\alpha-1}+\frac{2}{\alpha-2}\right)\right)^{\frac{1}{\alpha}}-1\right)^2}\frac{W}{\sqrt{n\log n}}$$

bits per second is feasible with probability approaching one as $n \to \infty$. If $\overline{L}$ is the mean distance between two points independently and uniformly distributed in the domain (either surface of sphere or planar disk of unit area), then there is a deterministic sequence $\epsilon(n) \to 0$, not depending on N, α, β or W, such that

$$\sqrt{\frac{8}{\pi}}\frac{W}{\overline{L}(\beta^{\frac{1}{\alpha}}-1)}\frac{1+\epsilon(n)}{\sqrt{n}}$$

bit-meters per second is infeasible with probability approaching one as $n \to \infty$.

C. Some Possible Implications

The results in this paper allow for a perfect scheduling algorithm which knows the locations of all nodes and all traffic demands, and which coordinates wireless transmissions temporally and spatially to avoid collisions which would otherwise result in lost packets. Also, the nodes are not mobile. If such perfect node location information is not available, or if nodes move, or traffic demands are not known, then the capacity can only be even smaller.

There are some implications of these results which designers may want to consider. The decrease in throughput with n may be regarded as unacceptable by users when the number n of nodes is large. Perhaps designers should target their efforts at networks for smaller numbers of users, rather than try to develop large wireless networks.

A feasible scenario is where nodes need to communicate only with nearby nodes. Then the scaled distance between sources and destinations is only $O\left(\frac{1}{\sqrt{n}}\right)$ meters. Thus all nodes can transmit data to nearby neighbors at a bit rate that does not decrease with n. Such a scenario can arise, for example, in collections of "smart homes," each home having sensors and actuators communicating by wireless means.

Another implication concerns the power consumption by each node for transmission. Consider Random Networks. The fraction of time that a modem is busy, whether relaying traffic or sending packets originating at the node, is only $\Theta\left(\frac{1}{\log n}\right)$. Not only that, the scaled range of each transmission is about $O\left(\sqrt{\frac{\log n}{n}}\right)$. The bounds for the Physical Model suggest that a faster rate of decay of signal power with distance, i.e., a larger α, allows greater transport and throughput capacity.

One more implication follows from the constructive proof of capacity. It shows that one can group the nodes into small clusters or "cells," where in each cell one can designate one specific node to carry all the burden of relaying multihop packets, if so desired. Thus a division of labor is possible, were this to be found profitable. Moreover, it would further reduce the transmission power consumed by the vast majority of other nodes. This may offer some suggestive guidelines for designers of routing protocols.

It should be noted that dividing the channel into subchannels does not change any of the results.

Yet another issue concerns the use of relay nodes.[2] Consider a Random Network with n source nodes. Then the throughput that can be furnished to each of them is only $\Theta\left(\frac{W}{\sqrt{n\log n}}\right)$ under the Protocol Model. Suppose m additional homogeneous nodes are deployed as pure relays in random positions, with no independent traffic needs of their own, i.e., they are not sources. Then the throughput that can be furnished to each of the n sources is $\Theta\left(\frac{(n+m)W}{n\sqrt{(n+m)\log(n+m)}}\right)$. There is, however, a severe cost of providing this increase in throughput. The number of additional relay nodes that need to be deployed to gain an appreciable increase in capacity for the source nodes may be very large. When there are $n = 100$ active nodes, to make $\frac{(n+m)}{n\sqrt{(n+m)\log(n+m)}}$ equal to five times its value at $m = 0$, m will have to be equal to at least 4476. The addition of kn nodes to serve as pure relays provides a less than $\sqrt{k+1}$-fold increase in this term.

One way to overcome the barrier of wireless networks is to do what is done in cellular telephony—connect the base stations by a wired network. If, however, nondirected wireless links are used for connecting the base stations, then the capacity limitation of wireless networks remains with us, though in less obvious ways. For example, suppose a high-power base station is chosen in each cell, which communicates with other distant base stations by a wireless channel. Then the set of base stations inherits the same capacity limitation. A set of b wire-

[2] We are grateful to Chip Elliott for raising this issue.

lessly connected base stations can provide a throughput of only $\Theta\left(\frac{1}{\sqrt{b \log b}}\right)$ for each base station.

D. A Discussion of the Tradeoffs Involved

Why does the throughput capacity diminish as the number of nodes increases? For an insight into some of the tradeoffs involved, consider Random Networks. Let the mean distance to be traversed by a packet be $\overline{L}$, and denote by $r(n)$ the common range of all transmissions. Then the mean number of hops taken by packets is no less than $\frac{\overline{L}}{r(n)}$. Thus each node generates at least $\frac{\overline{L}\lambda(n)}{r(n)}$ bits per second of traffic for other nodes. Since the total number of nodes is n, the total traffic is no less than $\frac{\overline{L}n\lambda(n)}{r(n)}$ bits per second. This has to be served by n nodes each capable of W bits per second. Thus one needs $\frac{\overline{L}n\lambda(n)}{r(n)} \le nW$. An upper bound on the throughput is therefore $\lambda(n) \le \frac{Wr(n)}{\overline{L}}$. Since the term on the right side grows linearly in $r(n)$, it might appear that to increase the throughput by reducing the number of hops traversed by each packet, and thus the burden on other nodes serving as relays, one should increase the range $r(n)$ of each node. However, the expression above is not an achievable upper bound as a function of $r(n)$. The reason is that we have neglected the reduction in capacity due to spatial concurrency constraints, since nodes close to a receiver are required to be idle to avoid collisions which cause the loss of packets. In fact, the loss from increasing $r(n)$ is quadratic due to the area of the conflict involved. Therefore, the desire to reduce the multihop burden and the desire to increase spatial concurrency and frequency reuse are in conflict. It turns out that when we consider both issues together, we find that one really needs to reduce the value of $r(n)$ to as small a value as possible. However, there is a limit to how small one can make $r(n)$. When the range $r(n)$ of transmissions is too small, the wireless network loses connectivity. In a precursor result, see [10], the critical range for connectivity of networks formed by randomly located nodes on a disk in the plane has been determined. Consider the graph with random vertices uniformly and independently distributed in a disk of unit area. Join two vertices by an edge whenever they are within a distance $r(n)$ from each other. The critical radius for connectivity is $\sqrt{\frac{\log n}{\pi n}}$ in the sense that the graph with $r(n) = \sqrt{\frac{\log n + \kappa_n}{\pi n}}$ is connected with probability approaching one as $n \to \infty$ if and only if $\kappa_n \to +\infty$.

For Arbitrary Networks under the Protocol model, just three constraints—the length of routes, the consumption of valuable two-dimensional area by transmissions, and the total number of nodes—are enough to force the transport capacity to be no more than $O(W\sqrt{n})$ bit-meters per second.

The rest of this paper is organized as follows. In Section II we exhibit upper bounds on the transport capacity of the form $cW\sqrt{n}$ bit-meters per second and $c'Wn^{\frac{\alpha-1}{\alpha}}$ bit-meters per second, under the Protocol and Physical Models, respectively, for Arbitrary Networks. In Section III we show that a transport capacity of $c''W\sqrt{n}$ bit-meters per second is also feasible for Arbitrary Networks. In Section IV we construct a scheduling and routing scheme which achieves a throughput of $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ bits per second for Random Networks on S^2. In Section V we show that $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ bits per second and $\Theta\left(\frac{W}{\sqrt{n}}\right)$ bits per second are upper bounds on the throughput for Random Networks on S^2, under the Protocol and Physical Models, respectively. In Section VI we show that the above results for Random Networks also hold for a disk in the plane.

II. Arbitrary Networks: An Upper Bound on Transport Capacity

We consider the setting on a planar disk of unit area. Consider the following (nearly) minimal set of assumptions:

(A.i) There are n nodes arbitrarily located in a disk of unit area on the plane. (The results carry over to any domain of unit area in R^2 which is the closure of its interior.)

(A.ii) The network transports λnT bits over T seconds.

(A.iii) The average distance between the source and destination of a bit is $\overline{L}$. Note that, together with (A.ii), this implies that a transport capacity of $\lambda n\overline{L}$ bit-meters per second is achieved.

(A.iv) Each node can transmit over any subset of M subchannels with capacities W_m bits per second, $1 \le m \le M$, where $\sum_{m=1}^{M} W_m = W$.

(A.v) Transmissions are slotted into synchronized slots of length τ seconds. (This assumption can be eliminated, but makes the exposition easier.)

(A.vi) While retaining the restriction (2) for the case of the Physical Model, we can either retain (1) in the Protocol Model or consider an alternate restriction as follows: If a node X_i transmits to another node X_j located at a distance of r units on a certain subchannel in a certain slot, then there can be no other receiver within a radius of Δr around X_j on the same subchannel in the same slot. This alternate restriction addresses situations where the transmissions are not omnidirectional, but nevertheless there is some dispersion in the neighborhood of the receiver.

Theorem 2.1:

i) In the Protocol Model, the transport capacity $\lambda n\overline{L}$ is bounded as follows:

$$\lambda n\overline{L} \le \frac{\sqrt{8}}{\pi}\,\frac{1}{\Delta}\,W\sqrt{n} \text{ bit-meters per second.}$$

ii) In the Physical Model

$$\lambda n\overline{L} \le \left(\frac{2\beta+2}{\beta}\right)^{1/\alpha}\frac{1}{\sqrt{\pi}}\,Wn^{\alpha-1/\alpha} \text{ bit-meters per second.}$$

iii) If the ratio $\frac{P_{\max}}{P_{\min}}$ between the maximum and minimum powers that transmitters can employ is strictly bounded above by β, then

$$\lambda n\overline{L} \le \sqrt{\frac{8}{\pi}}\,\frac{1}{\left(\frac{\beta P_{\min}}{P_{\max}}\right)^{1/\alpha} - 1}\,W\sqrt{n} \text{ bit-meters per second.}$$

iv) When the domain is of A square meters rather than 1 m2, then all the upper bounds above are scaled by $\sqrt{A}$.

Proof: Consider bit b, where $1 \leq b \leq \lambda n T$. Let us suppose that it moves from its origin to its destination in a sequence of $h(b)$ hops, where the hth hop traverses a distance of r_b^h. Then from (A.iii)

$$\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} r_b^h \geq \lambda n T \overline{L}. \tag{6}$$

Note now that in any slot at most $n/2$ nodes can transmit. Hence for any subchannel m and any slot s

$$\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} 1(\text{The } h\text{th hop of bit } b \text{ is over subchannel } m \text{ in slot } s) \leq \frac{W_m \tau n}{2}.$$

Summing over the subchannels and the slots, and noting that there can be no more than $\frac{T}{\tau}$ slots in T seconds, yields

$$H := \sum_{b=1}^{\lambda n T} h(b) \leq \frac{WTn}{2}. \tag{7}$$

Consider now the Protocol Model. Suppose that X_j is receiving a transmission from X_i over the mth subchannel at the same time that X_ℓ is receiving a transmission from X_k over the same subchannel. Then from the triangle inequality and (1)

$$\begin{aligned} |X_j - X_\ell| &\geq |X_j - X_k| - |X_\ell - X_k| \\ &\geq (1+\Delta)|X_i - X_j| - |X_\ell - X_k|. \end{aligned}$$

Similarly,

$$|X_\ell - X_j| \geq (1+\Delta)|X_k - X_\ell| - |X_j - X_i|.$$

Adding the two inequalities, we obtain

$$|X_\ell - X_j| \geq \frac{\Delta}{2} \left(|X_k - X_\ell| + |X_i - X_j|\right).$$

Hence disks of radius $\frac{\Delta}{2}$ times the lengths of hops centered at the receivers over the same subchannel in the same slot are essentially disjoint. (Note that this conclusion directly follows when (1) is replaced by the alternate restriction of Assumption (A.vi)). Allowing for edge effects where a node is near the periphery of the domain, and noting that a range greater than the diameter of the domain is unnecessary, we see that at least a quarter of such a disk is within the domain. Since at most $W_m \tau$ bits can be carried in slot s from a receiver to a transmitter over the mth subchannel, we have

$$\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} 1(\text{The } h\text{th hop of bit } b \text{ is over subchannel } m \text{ in slot } s) \frac{\pi \Delta^2}{16} (r_b^h)^2 \leq W_m \tau. \tag{8}$$

Summing over the subchannels and the slots gives

$$\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} \frac{\pi \Delta^2}{16} (r_b^h)^2 \leq WT.$$

This can be rewritten as

$$\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} \frac{1}{H} (r_b^h)^2 \leq \frac{16WT}{\pi \Delta^2 H}. \tag{9}$$

Note now that the quadratic function is convex. Hence

$$\left(\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} \frac{1}{H} r_b^h\right)^2 \leq \sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} \frac{1}{H} (r_b^h)^2. \tag{10}$$

Combining (9) and (10) yields

$$\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} r_b^h \leq \sqrt{\frac{16WTH}{\pi \Delta^2}}. \tag{11}$$

Now substituting (6) in (11) gives

$$\lambda n T \overline{L} \leq \sqrt{\frac{16WTH}{\pi \Delta^2}}. \tag{12}$$

Substituting (7) in (12) yields the result.

Now turn to the Physical Model. The difference stems from the need to replace (8) by a different expression. Suppose X_i is transmitting to $X_{j(i)}$ over the mth subchannel at power level P_i at some time, and let $\mathcal{T}$ denote the set of all simultaneous transmitters over the mth subchannel at that time. Including the signal power of X_i also in the denominator, the signal-to-interference requirement (2) for $X_{j(i)}$ can be written as

$$\frac{\frac{P_i}{|X_i - X_{j(i)}|^\alpha}}{N + \sum_{k \in \mathcal{T}} \frac{P_k}{|X_k - X_{j(i)}|^\alpha}} \geq \frac{\beta}{\beta + 1}.$$

Hence

$$\begin{aligned} |X_i - X_{j(i)}|^\alpha &\leq \frac{\beta+1}{\beta} \frac{P_i}{N + \sum_{k \in \mathcal{T}} \frac{P_k}{|X_k - X_{j(i)}|^\alpha}} \\ &\leq \frac{\beta+1}{\beta} \frac{P_i}{N + \left(\frac{\pi}{4}\right)^{\alpha/2} \sum_{k \in \mathcal{T}} P_k} \\ &\qquad \left(\text{since } |X_k - X_{j(i)}| \leq \frac{2}{\sqrt{\pi}}\right). \end{aligned}$$

Summing over all transmitter-receiver pairs

$$\begin{aligned} \sum_{i \in \mathcal{T}} |X_i - X_{j(i)}|^\alpha &\leq \frac{\beta+1}{\beta} \frac{\sum_{i \in \mathcal{T}} P_i}{N + \left(\frac{\pi}{4}\right)^{\alpha/2} \sum_{k \in \mathcal{T}} P_k} \\ &\leq 2^\alpha \pi^{-(\alpha/2)} \frac{\beta+1}{\beta}. \end{aligned}$$

Summing over all slots and subchannels gives

$$\sum_{b=1}^{\lambda n T} \sum_{h=1}^{h(b)} r^\alpha(h,b) \leq 2^\alpha \pi^{-\frac{\alpha}{2}} \frac{\beta+1}{\beta} WT.$$

The rest of the proof proceeds along lines similar to the Protocol Model, invoking the convexity of r^α instead of r^2.

For the consideration of the special case where $\frac{P_{\max}}{P_{\min}} < \beta$, we start with (2). From it, it follows that if X_i is transmitting to X_j at the same time that X_k is transmitting to X_ℓ, both over the same subchannel, then

$$\frac{\frac{P_i}{|X_i - X_j|^\alpha}}{\frac{P_k}{|X_k - X_j|^\alpha}} \geq \beta.$$

Thus

$$\begin{aligned} |X_k - X_j| &\geq \left(\frac{\beta P_{\min}}{P_{\max}}\right)^{\frac{1}{\alpha}} |X_i - X_j| \\ &= (1+\Delta)|X_i - X_j| \end{aligned}$$

where $\Delta := \left(\frac{\beta P_{\min}}{P_{\max}}\right)^{\frac{1}{\alpha}} - 1$. Thus the same upper bound as for the Protocol Model carries over with Δ defined as above. $\square$

III. Arbitrary Networks: A Constructive Lower Bound on Transport Capacity

We will now show that the order of the upper bound in the previous section is sharp for the Protocol Model, by exhibiting a scenario where it is achieved. This scenario is also feasible for the Physical Model.

Theorem 3.1: There is a placement of nodes and an assignment of traffic patterns such that the network can achieve $\frac{W}{1+2\Delta} \frac{n}{\sqrt{n}+\sqrt{8\pi}}$ bit-meters per second under the Protocol Model, and

$$\frac{1}{\left(16\beta\left(2^{\frac{\alpha}{2}} + \frac{6^{\alpha-2}}{\alpha-2}\right)\right)^{\frac{1}{\alpha}}} \frac{Wn}{\sqrt{n}+\sqrt{8\pi}}$$

bit-meters per second under the Physical Model, both whenever n is a multiple of 4.

Proof: Consider the Protocol Model. Define

$$r := \frac{1}{1+2\Delta} \frac{1}{\sqrt{\frac{n}{4}} + \sqrt{2\pi}}.$$

Recall that the domain is a disk of unit area, i.e., of radius $\frac{1}{\sqrt{\pi}}$ in the plane. With the center of the disk located at the origin, place transmitters at locations

$$(j(1+2\Delta)r \pm \Delta r, k(1+2\Delta)r)$$

and

$$(j(1+2\Delta)r, k(1+2\Delta)r \pm \Delta r)$$

where $|j+k|$ is even. Also place receivers at

$$(j(1+2\Delta)r \pm \Delta r, k(1+2\Delta)r)$$

and

$$(j(1+2\Delta)r, k(1+2\Delta)r \pm \Delta r)$$

where $|j+k|$ is odd. Each transmitter can transmit to its nearest receiver, which is at a distance r away, without interference from any other transmitter–receiver pair. It can be verified that there are at least $\frac{n}{2}$ transmitter–receiver pairs all located within the domain. (This is done by noting that for a tessellation of the plane by squares of side s, all squares intersecting a disk of radius $R - \sqrt{2}s$ are entirely contained within a larger concentric disk of radius R. The number of such squares is greater than $\frac{\pi(R-\sqrt{2}s)^2}{s^2}$. Now take $s = (1+2\Delta)r$ and $R = \frac{1}{\sqrt{\pi}}$.) Restricting attention to just these pairs, there are a total of $\frac{n}{2}$ simultaneous transmissions, each of range r, and each at W bits per second. This achieves the transport capacity indicated.

For the Physical Model, a calculation of the SIR shows that it is lower-bounded at all receivers by $\frac{(1+2\Delta)^\alpha}{16\left(2^{\frac{\alpha}{2}} + \frac{6^{\alpha-2}}{\alpha-2}\right)}$. Choosing Δ to make this lower bound equal to β yields the result. $\square$

The above lower bounds on feasible transport capacity can be sharpened. The following bounds may be useful in the design of networks with small numbers of nodes.

Lemma 3.1: In the Protocol Model, there is a placement of nodes and an assignment of traffic patterns such that the network can achieve

$$\frac{2W}{\sqrt{\pi}} \text{ bit-meters per second,} \qquad \text{for } n \geq 2$$

$$\frac{4W}{\sqrt{\pi}(1+\Delta)} \text{ bit-meters per second,} \qquad \text{for } n \geq 8$$

$$\frac{W}{1+2\Delta} \frac{n}{\sqrt{n}+\sqrt{8\pi}} \text{ bit-meters per second,} \qquad \text{for } n = 2, 3, 4, \cdots, 19, 20, 21$$

and

$$\frac{W}{1+2\Delta} \frac{4\lfloor \frac{n}{4} \rfloor}{\sqrt{4\lfloor \frac{n}{4} \rfloor} + \sqrt{8\pi}} \text{ bit-meters per second,} \quad \text{for all } n.$$

Proof: With at least two nodes, clearly $\frac{2W}{\sqrt{\pi}}$ bit-meters per second can be achieved by placing two nodes at diametrically opposite locations. This verifies the formula for the bound for $n \leq 8$. With at least eight nodes, four transmitters can be placed at the opposite ends of perpendicular diameters, and each can transmit toward its receiver located at a distance $\frac{1}{\sqrt{\pi}(2+2\Delta)}$ toward the center of the domain. This yields $\frac{4W}{\sqrt{\pi}(1+\Delta)}$ bit-meters per second, verifying the formula up to $n = 21$. $\square$

These bounds can be further improved slightly by tessellating the domain into hexagons, at the expense of more unwieldy expressions.

IV. Random Networks: A Constructive Lower Bound on Throughput Capacity

Now we turn to Random Networks. Even though the setting of the problem is very different, the proof of throughput capacity is somewhat reminiscent of traditional information-theoretic arguments. We provide a constructive scheme to show that one can spatially and temporally schedule transmissions in a random graph so that when each randomly located node has a randomly chosen destination, each source–destination pair can indeed be guaranteed a "virtual channel" of capacity $\frac{cW}{(1+\Delta)^2\sqrt{n \log n}}$ bits per second with probability approaching 1 as $n \to \infty$, for an appropriate constant $c > 0$. We will show how to route traffic efficiently through the random graph so that no node is overloaded. The routing scheme will utilize a Voronoi tessellation of S^2 with some special properties. The size of each Voronoi cell is chosen carefully in relation to the number of nodes. Every cell should also be neither too thin nor too fat. The routing will be over nearly straight-line paths, which assures that it is efficient. To show that the load is balanced uniformly over the entire network, we calculate the Vapnik–Chervonenkis dimension for certain geometrically defined random variables on the plane and the sphere, which are connected with the tessellations and routes used. We will need to ensure that the routes are independently and identically distributed. This will require us to circumvent the possible pitfall that knowledge of one route provides information on the locations of the source, destination, and intermediate relay nodes, thus possibly introducing dependencies with other routes which may depend on the locations of these nodes.

We begin the constructive proof of the lower bound on the throughput capacity for Random Networks. Our treatment will be directed at the Protocol Model. Where appropriate we will comment on the arguments required for the Physical Model.

A. A Spatial Tessellation

We use a Voronoi tessellation of the surface S^2 of the sphere. Recall the definition of a Voronoi tessellation, see Okabe, Boots and Sugihara [11]. Let $\{a_1, a_2, \cdots, a_p\}$ be a set of p points on S^2 (or any other set for that matter). The Voronoi cell $V(a_i)$ is the set of all points which are closer to a_i than to any of the other a_j's, i.e.,

$$V(a_i) := \{x \in S^2 : |x - a_i| = \min_{1 \le j \le p} |x - a_j|\}.$$

Above and throughout, distances are measured on the surface S^2 of the sphere by segments of great circles connecting two points; see Stilwell [12]. The point a_i is called the generator of the Voronoi cell $V(a_i)$. Fig. 2 shows an example of a tessellation of S^2. Unfortunately, the surface of the sphere does not allow any regular tessellation where all cells look the same, except for the platonic solids; see Lyndon [13]. These latter tessellations cannot be made as fine as we need to make them. Moreover, our Voronoi tessellations will also need to be not too eccentrically shaped. We exhibit tessellations with these two special properties in the following lemma, the proof of which is constructive.

Lemma 4.1: For every $\epsilon > 0$, there is a Voronoi tessellation of S^2 with the property that every Voronoi cell contains a disk of radius ϵ and is contained in a disk of radius 2ϵ.

Proof: Denote by $D(x, \epsilon)$ a disk of radius ϵ centered at x. Choose a_1 as any point in S^2. Suppose that $a_1, \cdots, a_p$ have already been chosen such that the distance between any two a_j's is at least 2ϵ. There are two cases to consider.

Suppose there is a point x such that $D(x, \epsilon)$ does not intersect any $D(a_i, \epsilon)$. Then x can be added to the collection: Define $a_{p+1} := x$. Otherwise, we stop.

This procedure has to terminate in a finite number of steps since the addition of each a_i removes the area of a disk of radius $\epsilon > 0$ from S^2. When we stop we will have a set of generators such that they are at least 2ϵ units apart, and such that all other points on S^2 are within a distance of 2ϵ from one of the generators. The Voronoi tessellation arising from this set of generators has the desired properties. □

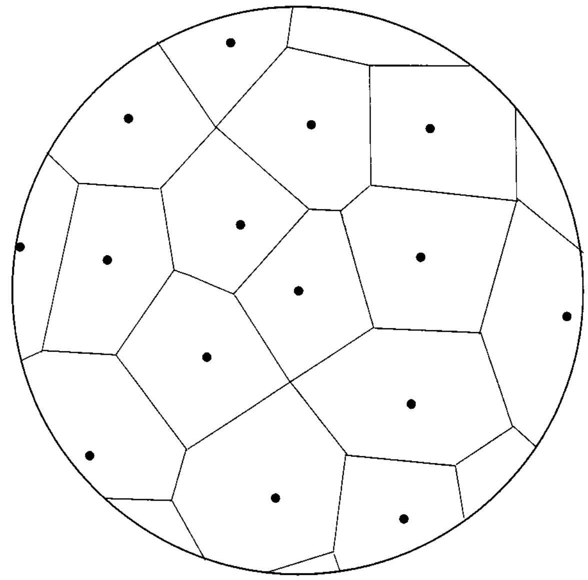

Fig. 2. A tessellation of the surface S^2 of the sphere.

In the sequel we will use a Voronoi tessellation $\mathcal{V}_n$ for which

(V.i) Every Voronoi cell contains a disk of area $100 \log n/n$. Let

$$\rho(n) := \text{radius of a disk of area } \frac{100 \log n}{n} \text{ on } S^2. \tag{13}$$

(Note that the area of a disk of radius ρ on S^2 is less than $\pi\rho^2$).

(V.ii) Every Voronoi cell is contained in a disk of radius $2\rho(n)$.

We will refer to each Voronoi cell $V \in \mathcal{V}_n$ as simply a "cell."

B. Adjacency and Interference

Note that all Voronoi cells are polygons since they are formed as finite intersections of hemispheres on S^2 (or halfspaces in the case of R^2).

Definition: Adjacent Cells: Say that two cells are *adjacent*, if they share a common point. (Recall that every cell is a closed set).

Let us choose the range $r(n)$ of each transmission so that

$$r(n) = 8\rho(n). \tag{14}$$

This range allows direct communication within a cell and between adjacent cells.

Lemma 4.2: Every node in a cell is within a distance $r(n)$ from every node in its own cell or adjacent cell.

Proof: The diameter of cells is bounded by $4\rho(n)$; see (V.ii). The range of a transmission is $8\rho(n)$. Thus the area covered by the transmission of a node includes adjacent cells. □

Definition: Interfering Neighbors: We say that two cells are *interfering neighbors* if there is a point in one cell which is within a distance $(2 + \Delta)r(n)$ of some point in the other cell.

As the name implies, the interpretation is this: If two cells are not interfering neighbors, then in the Protocol Model a transmission from one cell cannot collide with a transmission from the other cell.

C. A Bound on the Number of Interfering Neighbors of a Cell

An important property of the constructed Voronoi tessellation $\mathcal{V}_n$ is that the number of interfering neighbors of a cell is uniformly bounded. This will be exploited in the next section in constructing a spatial transmission schedule which allows for a high degree of spatial concurrency and thus frequency reuse. From now on c_i's will be used to denote deterministic constants not depending on n.

Lemma 4.3: Every cell in $\mathcal{V}_n$ has no more than c_1 interfering neighbors. c_1 depends only on Δ and grows no faster than linearly in $(1+\Delta)^2$.

Proof: Let V be a Voronoi cell. If V' is an interfering neighboring Voronoi cell, there must be two points, one in V and the other in V', which are no more than $(2+\Delta)r(n)$ units apart. From (V.ii), the diameter of a cell is bounded by $4\rho(n)$. Hence V', and similarly every other interfering neighbor in the Protocol Model, must be contained within a common large disk D of radius $6\rho(n)+(2+\Delta)r(n)$.

Such a disk D cannot contain more than $c_2\frac{(6\rho(n)+(2+\Delta)r(n))^2}{\rho^2(n)}$ disks of radius $\rho(n)$. By (V.i), there can therefore be no more than this number of cells within D. This therefore is an upper bound on the number of interfering neighbors of the cell V. The result follows from the magnitudes of $\rho(n)$ and $r(n)$ chosen as in (14). □

D. A Bound on the Length of an All-Cell Inclusive Transmission Schedule

The bounded number of interfering neighbors for each cell allows the construction of a schedule of bounded length which allows one opportunity for each cell in the tessellation $\mathcal{V}_n$ to transmit.

Lemma 4.4:

i) In the Protocol Model there is a schedule for transmitting packets such that in every $(1+c_1)$ slots, each cell in the tessellation $\mathcal{V}_n$ gets one slot in which to transmit, and such that all transmissions are successfully received within a distance $r(n)$ from their transmitters.

ii) There is a deterministic constant c not depending on n, N, α, β, or W such that if Δ is chosen to satisfy

$$(1+\Delta)^2 > \left(2\left(c\beta\left(3+\frac{1}{\alpha-1}+\frac{2}{\alpha-2}\right)\right)^{\frac{1}{\alpha}}-1\right)^2$$

then for a large enough common power level P, the above result i) holds even for the Physical Model.

Proof: First we show the result for the Protocol Model. This follows from a well-known fact about vertex coloring of graphs of bounded degree: A graph of degree no more than c_1 can have its vertices colored by using no more than $(1+c_1)$ colors, with no two neighboring vertices have the same color; see Bondy and Murthy [14]. One can therefore color the cells with no more than $(1+c_1)$ colors such that no two interfering neighbors have the same color. This gives a schedule of length at most $(1+c_1)$, where one can transmit one packet from each cell of the same color in a slot.

For the Physical Model we will show that under the same schedule as above, the required SIR of β is obtained if each transmitter chooses an identical power level P that is high enough, and Δ is large enough.

Note first that any two nodes transmitting simultaneously are separated by a distance of at least $(2+\Delta)r(n)$. Hence disks of radius $(1+\frac{\Delta}{2})r(n)$ around each transmitter are disjoint. The area of each such disk is at least $c_3\pi(1+\frac{\Delta}{2})^2r^2(n)$. (In the case of disks on the plane $c_3=1$, but it is smaller for disks on the surface of the sphere).

Consider a node X_i transmitting to a node X_j at a distance less than $r(n)$. The signal power received at X_j is at least $\frac{P}{r^\alpha(n)}$. Now we look at the interference power due to all the other simultaneous transmissions. Consider the annulus of all points lying within a distance between a and b from X_j. A transmitter within this annulus has the disk centered at itself and of radius $(1+\frac{\Delta}{2})r(n)$ entirely contained within a larger annulus of all points lying between a distance $a-(1+\frac{\Delta}{2})r(n)$ and $b+(1+\frac{\Delta}{2})r(n)$. The area of this larger annulus is no more than

$$c_4\pi\left\{\left[b+\left(1+\frac{\Delta}{2}\right)r(n)\right]^2-\left[a-\left(1+\frac{\Delta}{2}\right)r(n)\right]^2\right\}.$$

Each transmitter above "consumes" an area of at least $c_3\pi(1+\frac{\Delta}{2})^2r^2(n)$, as noted earlier. Hence the annulus of points at a distance between a and b from the receiver X_j cannot contain more than

$$\frac{c_4\pi\left\{\left[b+\left(1+\frac{\Delta}{2}\right)r(n)\right]^2-\left[a-\left(1+\frac{\Delta}{2}\right)r(n)\right]^2\right\}}{c_3\pi\left(1+\frac{\Delta}{2}\right)^2r^2(n)}$$

transmitters. Furthermore, the received power at X_j from each such transmission is at most P/a^α. Noting that there can be no other simultaneous transmitter within a distance $(1+\Delta)r(n)$ of X_j, and taking $a=k(1+\frac{\Delta}{2})r(n)$ and $b=(k+1)(1+\frac{\Delta}{2})r(n)$ for $k=1,2,3,\cdots$, we see that the SIR at X_j is lower-bounded by

$$\frac{\frac{P}{r^\alpha(n)}}{N+\sum_{k=1}^{+\infty}\frac{c_4((k+2)^2-(k-1)^2)}{c_3}\frac{P}{k^\alpha\left(1+\frac{\Delta}{2}\right)^\alpha r^\alpha(n)}}=\frac{\frac{P}{N}}{r^\alpha(n)+\frac{c_4}{c_3\left(1+\frac{\Delta}{2}\right)^\alpha}\frac{P}{N}\sum_{k=1}^{+\infty}\frac{6k+3}{k^\alpha}}.$$

Since $\alpha>2$, the sum in the denominator converges, and is in fact smaller than $(9+\frac{3}{\alpha-1}+\frac{6}{\alpha-2})$. When Δ is as specified and $P\to\infty$, the lower bound on the SIR converges to a value greater than β. □

E. The Source-Destination Pairs

Each node wishes to communicate with the node nearest to a randomly chosen location. Let Y_i be a randomly chosen location such that X_i and Y_i are independently and uniformly distributed (i.i.d.) on S^2, and that the sequence $\{(X_i,Y_i)\}_{i=1}^n$ is i.i.d. The *destination node* $X_{\text{dest}(i)}$ for the traffic generated at node X_i is chosen as the node X_j which is closest to Y_i.

Denote by L_i the straight-line segment connecting X_i and Y_i. Above, and in the rest of the paper, by a "straight-line" segment we actually mean a segment of the great circle on the surface S^2 of the sphere; see [12]. There is one significant property enjoyed by the sequence of straight lines $\{L_i\}_{i=1}^n$.

Lemma 4.5: The random sequence of straight-line segments $\{L_i\}_{i=1}^n$ is i.i.d.

This has the powerful consequence of allowing us to apply the law of large numbers to the i.i.d. straight-line segments. It will be useful since the route followed by each origination–destination pair will approximate the corresponding straight-line segment, as described in the next section.

F. The Routes of Packets

We will choose the routes of packets to approximate these straight-line segments. The straight-line segment L_i will intersect many cells in the tessellation $\mathcal{V}_n$. Let V_i denote the particular cell which contains X_i, and V_i' the cell which contains Y_i.

Packets originating at X_i will be relayed from the cell V_i to the cell V_i' in a sequence of hops. In each hop, the packet is transferred from one cell to another in the order in which they intersect the line. (If two cells are both "next" cells, then either can be chosen arbitrarily). Finally, after reaching the cell V_i' containing Y_i, the packets will be sent on to their final destination, which we shall show later in Section IV-G to be no more than one hop away with high probability.

Note that this is a randomized algorithm for choosing routes. It can be thought of as a load balancing scheme with some rather powerful uniformity properties, as shown in Section IV-I.

G. Each Cell Contains at Least One Node

To make relaying of traffic from one cell to an adjacent cell feasible, we need to first ensure that *every* cell V in $\mathcal{V}_n$ contains at least one node. For this we use uniform convergence in the weak law of large numbers. Note that uniformity is required over all cells in $\mathcal{V}_n$. We recall the following definitions; see Vapnik and Chervonenkis [15] and Vapnik [16]. Let $\mathcal{F}$ be a set of subsets. A finite set of points A is said to be *shattered* by $\mathcal{F}$ if for every subset B of A there is a set $F \in \mathcal{F}$ such that $A \cap F = B$. The *VC-dimension of* $\mathcal{F}$, denoted by VC-$d(\mathcal{F})$, is defined as the supremum of the sizes of all finite sets that can be shattered by $\mathcal{F}$. For sets of finite VC-dimension, one has uniform convergence in the weak law of large numbers.

The Vapnik–Chervonenkis Theorem: If $\mathcal{F}$ is a set of finite VC-dimension VC-$d(\mathcal{F})$, and $\{X_i\}$ is a sequence of i.i.d. random variables with common probability distribution P, then for every ϵ, $\delta > 0$

$$\text{Prob}\left(\sup_{F\in\mathcal{F}}\left|\frac{1}{N}\sum_{i=1}^{N} I(X_i \in F) - P(F)\right| \le \epsilon\right) > 1-\delta$$

whenever

$$N > \max\left\{\frac{\text{VC-}d(\mathcal{F})}{\epsilon}\log\frac{16e}{\epsilon}, \frac{4}{\epsilon}\log\frac{2}{\delta}\right\}.$$

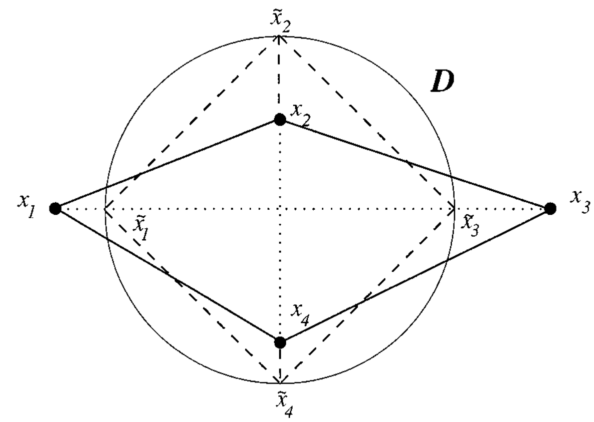

Fig. 3. Proof that the vertices of a quadrilateral cannot be shattered by the set of disks.

First we will consider the case where $\mathcal{F}$ is the set of all disks on the plane. Later we will consider the case where the disks are located on S^2. In the planar case we can make use of results from Euclidean geometry. The following result may perhaps be known already, though we have been unable to find it in the literature.

Lemma 4.6: The Vapnik–Chervonenkis dimension of the set of disks in R^2 is 3.

Proof: It is easy to see that there is a three-point set that can be shattered by the set of disks. An example is the set of vertices of an equilateral triangle.

Suppose there is a set $\{x_1, x_2, x_3, x_4\}$ of four points that is shattered by the set of disks. If any one of the X_i's lies in the convex hull of the other three points, then there is no disk which can contain the others without containing X_i too. Hence we can assume without loss of generality that the convex hull of the four points is a quadrilateral.

Again, we obtain a contradiction as follows. Without loss of generality, suppose that the angles of the quadrilateral at x_1 and x_3 sum to at least 180°, i.e.,

$$\angle x_1 + \angle x_3 \ge 180°.$$

Suppose D is a disk which contains x_2 and x_4, but not x_1 or x_3; see Fig. 3. Extend the diagonal x_2x_4 outwards in both directions till it meets the circumference of D at the points $\tilde{x}_2$ and $\tilde{x}_4$. Simultaneously, let $\tilde{x}_1$ and $\tilde{x}_3$ be the points of intersection of the diagonal x_1x_3 with the circumference of D. Then $\tilde{x}_1\tilde{x}_2\tilde{x}_3\tilde{x}_4$ is a cyclic quadrilateral. However,

$$\begin{aligned}\angle\tilde{x}_1 + \angle\tilde{x}_3 &> \angle x_1 + \angle x_3\\ &\ge 180°.\end{aligned}$$

This is a contradiction since the sum of the opposite angles of a cyclic quadrilateral is exactly 180°. □

Now we address the problem of determining the VC-dimension of disks on the surface of a sphere. It is sufficient for us to restrict attention to disks strictly smaller than hemispheres.

To convert results from the plane to S^2, we use a mapping called the "inversion map" which maps the punctured surface of the sphere onto the plane. Since the radius of the sphere is

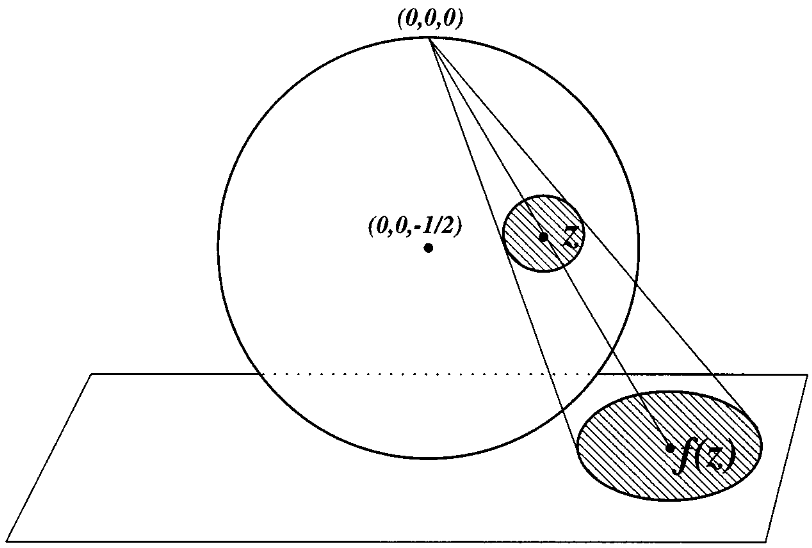

Fig. 4. The inversion mapping maps disks on S^2 into disks on R^2.

immaterial for the remainder of this discussion, we consider a sphere of radius $\frac{1}{2}$, centered at the point $(0,0,-\frac{1}{2})$. Let us refer to it temporarily as S^2. Also let us refer to the plane $z_3 = -1$ as H. Then the mapping

$$f(z) := \frac{z}{||z||^2}$$

where $||\cdot||$ is the Euclidean norm, has several useful properties (see [11]).

(i) It maps the punctured surface S^2 (i.e., S^2 except for the origin) onto the plane H. In fact, each point z on S^2 is mapped to the point obtained by extending the ray from the origin to z until it hits the plane H.

(ii) $f^{-1}(z) = f(z)$.

(iii) It maps disks on S^2 not containing the origin into disks on the plane H. See Fig. 4.

For our purposes, the last property is most important. It is used in the following lemma.

Lemma 4.7: The VC-dimension of the set of disks on S^2 strictly smaller than hemispheres is 3.

Proof: The proof parallels the contradiction argument of Lemma 4.6. Suppose that there is a set of four points $\{x_1, x_2, x_3, x_4\}$ which is shattered by such disks. They all have to be contained in a disk smaller than a hemisphere. Let x_1 and x_3 be opposite vertices of the quadrilateral formed. Since the set is shattered, there are two disks, each of radius less than that necessary to form a hemisphere, one of which contains x_1 and x_3 but excludes x_2 and x_4, while the other contains x_2 and x_4 but excludes x_1 and x_3. Since each disk is strictly less than a hemisphere, there is a point in the complement of their union. Rotate the sphere so that this point is at the top.

Without loss of generality we can scale the sphere so that its radius is $\frac{1}{2}$, and then translate it so that its top is at the origin. Applying the inversion map shows that there is a disk on the plane H which contains $f(x_1)$ and $f(x_3)$ and excludes $f(x_2)$ and $f(x_4)$, and another disk on H which contains $f(x_2)$ and $f(x_4)$ and excludes $f(x_1)$ and $f(x_3)$. However, we have seen the impossibility of this happening on the plane in Lemma 4.6. □

Since each cell V in the tessellation $\mathcal{V}_n$ contains a disk of area $100\ \log n/n$ (from V.i), we can appeal to uniform convergence in the law of large numbers.

Lemma 4.8: There is a sequence $\delta(n) \to 0$ such that

$$\text{Prob(Every cell } V \in \mathcal{V}_n \text{ contains a node)} \geq 1 - \delta(n).$$

Proof: Let $\mathcal{F}$ denote the class of disks of area $100 \log n/n$. Note that the VC-dimension of $\mathcal{F}$ is also 3. Hence

$$\text{Prob}\left(\sup_{D\in\mathcal{F}} \left|\frac{\text{Number of nodes in } D}{n} - \frac{100\log n}{n}\right| \leq \epsilon(n)\right) > 1 - \delta(n)$$

whenever

$$n > \max\left\{\frac{24}{\epsilon(n)}\ \log\frac{16e}{\epsilon(n)}, \frac{4}{\epsilon(n)}\ \log\frac{2}{\delta(n)}\right\}.$$

This is satisfied when

$$\epsilon(n) = \delta(n) = \frac{50\log n}{n}.$$

Since each cell V in $\mathcal{V}_n$ contains a disk of area $100 \log n/n$, we have

$$\text{Prob(Number of nodes in } V \geq 50\log n, \text{ for every } V \in \mathcal{V}_n) > 1 - \delta(n).$$

The result follows. □

Hence every cell in $\mathcal{V}_n$ contains at least one node to relay the traffic (with probability exceeding $(1 - \frac{50\log n}{n})$). Moreover, every such node has enough range to communicate with all nodes in any adjacent cell (see Lemma 4.2). Hence packets can be relayed from one cell intersecting a line L_i to the next cell intersecting the line. Hence the routing scheme given above can indeed work as planned with probability exceeding $1 - \frac{50\log n}{n}$.

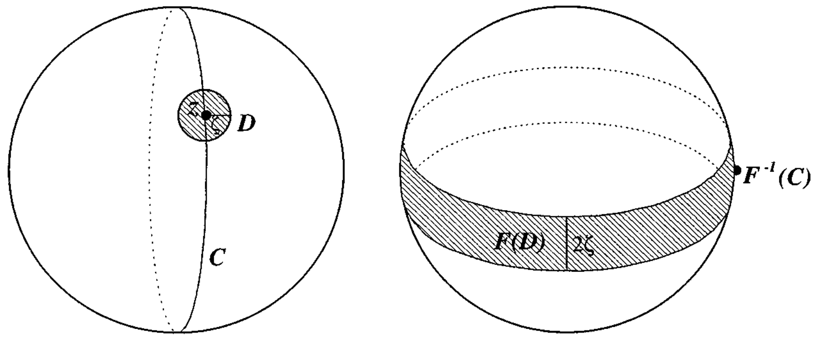

Fig. 5. Transforming great circles intersecting disks into points lying in equatorial bands.

From now on we will use the phrase "with high probability," abbreviated as *whp* to stand for "with probability approaching 1 as $n \to \infty$." The multihop relaying scheme can therefore function as planned *whp*.

H. The Mean Number of Routes Served by Each Cell

Recall that the straight line L_i connects X_i and Y_i, where X_i and Y_i are independently and uniformly distributed on S^2. By our assumption (V.ii) on the tessellation $\mathcal{V}_n$, each cell $V \in \mathcal{V}_n$ is contained in a disk of radius no more than $\sqrt{\frac{400 \log n}{\pi n}}$. (Note that the area of a disk of radius ρ on S^2 is less than $\pi\rho^2$). This allows us to bound the probability that a line L_i intersects a given cell V in $\mathcal{V}_n$.

Lemma 4.9: For every line L_i and cell $V \in \mathcal{V}_n$

$$\text{Prob}(\text{Line } L_i \text{ intersects } V) \le c_5\sqrt{\frac{\log n}{n}}.$$

Proof: As noted above, from property (V.ii) of the tessellation, every cell $V \in \mathcal{V}_n$ is contained in a disk of radius $\sqrt{\frac{400 \log n}{\pi n}}$. If X_i lies at a distance x from the disk, then the angle α subtended at X_i by the disk is no more than $\frac{c_6}{x}\sqrt{\frac{\log n}{n}}$. The area of the sector so formed is no more than $\frac{c_7 \alpha}{2\pi}$. If Y_i does not lie in this sector, then the line L_i joining X_i and Y_i cannot intersect the disk containing the cell V. Hence, for a point X_i at a distance x from the disk of radius $\sqrt{\frac{400 \log n}{\pi n}}$ containing the cell V, the probability that the line connecting X_i and Y_i intersects the disk is no more than $\frac{c_8}{x}\sqrt{\frac{\log n}{n}}$.

Since X_i is uniformly distributed on S^2, the probability density that it is at a distance x from the disk is bounded above by $2c_9\pi(x + \sqrt{\frac{400 \log n}{\pi n}})$. Integrating, we obtain

$$\begin{aligned}\text{Prob}\,(L_i \text{ intersects } V) \le& \int_{\sqrt{400 \log n/n}}^{\sqrt{\pi}/2} \left(\frac{c_8}{x}\sqrt{\frac{\log\ n}{n}}\right) \\ &\cdot 2c_9\pi\left(x + \sqrt{\frac{400 \log\ n}{\pi n}}\right)\,dx. \\ \le& c_5\sqrt{\frac{\log n}{n}}. \qquad \square\end{aligned}$$

Let C_i denote the great circle containing the line L_i, i.e., the extension of the line so that it wraps around the sphere. The same proof technique shows the following.

Lemma 4.10: For every great circle C_i and cell $V \in \mathcal{V}_n$

$$\text{Prob}\,(\text{Great circle } C_i \text{ intersects } V) \le c_{10}\sqrt{\frac{\log n}{n}}.$$

There being a total of n lines $\{L_i\}_{i=1}^n$, one connecting each X_i with Y_i, the mean number of lines or great circles passing through a cell is bounded as follows:

$$\begin{aligned}&E[\text{Number of lines in } \{L_i\}_{i=1}^n \text{ intersecting a cell } V] \\ &\qquad\le c_5\sqrt{n \log n} \\ &E[\text{Number of great circles in} \{C_i\}_{i=1}^n \text{ intersecting a cell } V] \\ &\qquad\le c_{10}\sqrt{n \log n}.\end{aligned}$$

I. The Actual Traffic Served by Each Cell

Above, since routes follow lines, we have bounded the mean number of routes passing through each cell. However, what we need to bound is the *actual* random number of routes served by every cell.

To do this we make use of the critical property that the sequence $\{(X_i, Y_i)\}$ is i.i.d. Hence, so are the straight lines L_i. This allows us to exploit uniform convergence in the law of large numbers.

Recall that each cell $V \in \mathcal{V}_n$ is contained in a disk of radius $2\rho(n)$. We will bound the number of great circles C_i intersecting such disks of radius $2\rho(n)$. This is clearly an upper bound on the number of lines L_i passing through cells.

We transform the problem of counting "intersections" of disks of radius ϵ with great circles into a "shattering" problem as follows. For every point z on S^2 let $F(z)$ denote the (unique) great circle containing all points equidistant from it. This is akin to associating an equator with a pole.

Given a great circle C, the inverse of this map is not well defined since every equator has two poles. However, we arbitrarily choose one of these two poles and designate it as the inverse $F^{-1}(C)$. Consider a disk D of radius ζ centered at a point z on S^2. Let $F(D) := \cup_{x \in D} F(x)$ denote the set of all points which are within a distance ζ from $F(z)$; it is a band of width 2ζ around the great circle $F(z)$. See Fig. 5.

Let $\mathcal{D}$ denote the set of all disks on S^2. It is easy to see the following lemma and corollary.

Lemma 4.11: The great circle C intersects the disk D if and only if the point $F^{-1}(C)$ is contained in the band $F(D)$.

Corollary 4.1: Let $\mathcal{C}(D)$ denote the set of all great circles which intersect $D \in \mathcal{D}$. The VC-dimension of $\{\mathcal{C}(D) : D \in \mathcal{D}\}$ is the same as the VC-dimension of $\{F(D) : D \in \mathcal{D}\}$.

Let $\mathcal{D}'$ denote the set of all disks strictly smaller than hemispheres. To appeal to uniform convergence in the law of large numbers we only have to show that the VC-dimension of $\{F(D) : D \in \mathcal{D}'\}$ is bounded. Note that for $D \in \mathcal{D}'$, each band $F(D)$ is the intersection of two disks, each strictly larger than a hemisphere. It is trivial that the VC-dimension of a class of sets is the same as the VC-dimension of the class formed by the complements of the sets. It is also known (see Vidyasagar [17]) that if $\mathcal{A}$ is a set of sets, and $\mathcal{B}$ consists of sets which are each obtained by intersecting two sets in $\mathcal{A}$, then

$$\text{VC-}d(\mathcal{B}) \leq 10 \text{ VC-}d(\mathcal{A}).$$

Hence we obtain the following lemma.

Lemma 4.12: The VC-dimension of $\{F(D) : D \in \mathcal{D}'\}$ is no more than ten times the VC-dimension of $\mathcal{D}'$.

In Lemma 4.7, we have already shown that the VC-dimension of $\mathcal{D}'$ is 3. Hence uniform convergence in the weak law of large numbers holds, and we obtain the following.

Lemma 4.13: There is a $\delta'(n) \to 0$ such that

$$\text{Prob}\Big(\sup_{V\in\mathcal{V}_n} \text{(Number of lines } L_i \text{ intersecting } V) \leq c_5\sqrt{n \log n}\Big) \geq 1 - \delta'(n).$$

Note that if a cell contains Y_i, it needs to forward the packet to its final destination $X_{\text{dest }(i)}$. This final destination is at most one hop away whp. Else, if a cell does not contain Y_i, then the traffic is relayed to the next cell. Hence the traffic handled by a cell is proportional to the number of lines passing through it. Since each line L_i carries traffic of rate $\lambda(n)$ bits per second, we have obtained the following bound.

Lemma 4.14: There is a $\delta'(n) \to 0$ such that

$$\text{Prob}\Big(\sup_{V\in\mathcal{V}_n} \text{(Traffic needing to be carried by cell } V) \leq c_5\lambda(n)\sqrt{n \log n}\Big) \geq 1 - \delta'(n).$$

J. Lower Bound on Throughput Capacity of Random Networks

From Lemma 4.4 we know that there exists a schedule for transmitting packets such that in every $(1 + c_1)$ slots, each cell in the tessellation $\mathcal{V}_n$ gets one slot to transmit, and such that each transmission is received within a range $r(n)$ of the transmitter. Thus the rate at which each cell gets to transmit is $W/1 + c_1$ bits per second.

On the other hand, the rate at which each cell *needs* to transmit is less than $c_5\lambda(n)\sqrt{n \log n}$ *whp*. With high probability, this rate can be accommodated by all cells if it is less than the rate available, i.e., if

$$c_5\lambda(n)\sqrt{n \log n} \leq \frac{W}{1 + c_1}.$$

Moreover, within a cell, the traffic to be handled by the entire cell can be handled by any one node in the cell, since each node can transmit at rate W bits per second whenever necessary. In fact, one can even designate one node in each cell as a "relay" node. This node can handle all the traffic needing to be relayed. The other nodes can simply serve as sources or sinks.

We have proved the following theorem, noting the linear growth of c_1 in $(1 + \Delta)^2$ in Lemma 4.3, and the choice of Δ in Lemma 4.4 for the Physical Model.

Theorem 4.1:

(i) For Random Networks on S^2 in the Protocol Model, there is a deterministic constant $c > 0$ not depending on n, Δ, or W, such that

$$\lambda(n) = \frac{cW}{(1 + \Delta)^2\sqrt{n \log n}}$$

bits per second is feasible *whp*.

ii) For Random Networks on S^2 in the Physical Model, there are deterministic constants c' and c'' not depending on n, N, α, β, or W, such that

$$\lambda(n) = \frac{c'}{(2(c''\beta(3 + \frac{1}{\alpha-1} + \frac{2}{\alpha-2}))^{\frac{1}{\alpha}} - 1)^2} \frac{W}{\sqrt{n \log n}}$$

bits per second is feasible *whp*.

It should be noted that these throughput levels have been attained without subdividing the wireless channel into subchannels of smaller capacity.

V. Random Networks: An Upper Bound on Throughput Capacity

Now we turn to the proof of the upper bound on the capacity for Random Networks.

First we will show that that when the range is too small not every source will be able to communicate with its desired destination.

A. Asymptotic Probability of an Isolated Node

From [10] we know that a necessary condition for connectivity *whp* for the problem of n nodes strewn on a disk of unit area in the plane is $r(n) = \sqrt{\frac{\log n + \kappa_n}{\pi n}}$, where $\kappa_n \to +\infty$. The S^2 setting here requires a slightly different treatment. The area of a disk of radius r on S^2 is not πr^2. A saving grace in comparison to a disk on the plane is that there is no need to consider the tedious issue of edge effects.

Another subtle issue is that we may not need connectivity of the entire graph. Strictly speaking, we only need that every source be able to communicate with its chosen destination. What we will show below is that disconnectedness manifests itself by the presence of isolated nodes. These nodes will then be unable to communicate with any other node. Hence the absence of isolated nodes is indeed a necessary condition for feasibility of any throughput.

We recall two results from [10].

Lemma 5.1:

(i) For any $p \in [0, 1]$

$$(1-p) \leq e^{-p}.$$

(ii) For any given $\theta \geq 1$, there exists $p_0 \in [0, 1]$, such that

$$e^{-\theta p} \leq (1-p), \qquad \text{for all } 0 \leq p \leq p_0.$$

If $\theta > 1$, then $p_0 > 0$.

Lemma 5.2: If $\pi r^2(n) = \frac{\log n + \kappa}{n}$, then, for any fixed $\theta < 1$ and for all sufficiently large n

$$n(1-\pi r^2(n))^{n-1} \geq \theta e^{-\kappa}.$$

Given the n nodes, denote by $\mathcal{G}(n, r(n))$ the graph which results from connecting nodes separated by a distance less than $r(n)$ by an edge. Let $P^{(k)}(n, r(n)),\ k = 1, 2, \cdots$ denote the probability that a graph $\mathcal{G}(n, r(n))$ has at least one order-k component, i.e., a set of k nodes which form a connected set, but which are not connected with any other node. Also, let $P_d(n, r(n))$ denote the probability that $\mathcal{G}(n, r(n))$ is disconnected.

The main necessary condition for the absence of a single isolated node, and consequently also for connectivity, is the following.

Lemma 5.3: If $\pi r^2(n) = \frac{\log n + \kappa_n}{n}$ where

$$\lim_{n\to\infty} \kappa_n = \kappa < +\infty$$

then

$$\liminf_{n\to\infty} P^{(1)}(n, r(n)) \geq e^{-\kappa}\left(1 - e^{-\kappa}\right)$$

and

$$\liminf_{n\to\infty} P_d(n, r(n)) \geq e^{-\kappa}\left(1 - e^{-\kappa}\right).$$

Proof: Consider first the case where $\pi r^2(n) = \frac{\log n + \kappa}{n}$ for a fixed κ. Consider $P^{(1)}(n, r(n))$, the probability that $\mathcal{G}(n, r(n))$ has *at least* one order-1 component. Then

$$\begin{aligned}
&P^{(1)}(n, r(n)) \\
&\quad \geq \sum_{i=1}^{n} P(\{i \text{ is the only isolated node in } \mathcal{G}(n, r(n))\}) \\
&\quad \geq \sum_{i=1}^{n} \Big(P(\{i \text{ is an isolated node in } \mathcal{G}(n, r(n))\}) \\
&\qquad - \sum_{j \neq i} P(\{i \text{ and } j \text{ are isolated nodes in } \mathcal{G}(n, r(n))\})\Big) \\
&\quad \geq \sum_{i=1}^{n} P(\{i \text{ is isolated in } \mathcal{G}(n, r(n))\}) \\
&\qquad - \sum_{i=1}^{n} \sum_{j \neq i} P(\{i \text{ and } j \text{ are isolated in } \mathcal{G}(n, r(n))\}).
\end{aligned} \tag{15}$$

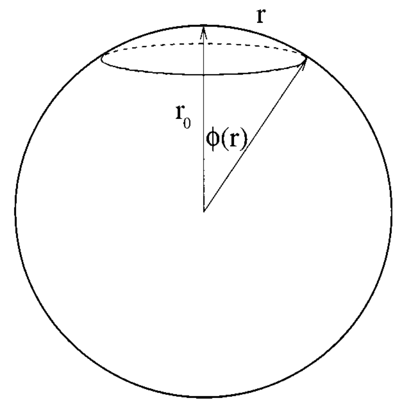

Fig. 6. Computing $A(r)$, the area of a disk of radius r on a sphere of unit surface area.

Next we compute the area $A(r)$ of a disk of radius r on S^2. Note that the radius of the sphere itself is $r_0 = \frac{1}{\sqrt{4\pi}}$. From $\phi(r) := r/r_0$ as indicated in Fig. 6, we get

$$\begin{aligned}
A(r) &= \int_0^{\phi(r)} 2\pi r_0 (\sin\phi) r_0 \, d\phi \\
&= 2\pi r_0^2 (1 - \cos\phi(r)) \\
&= \frac{1}{2}\left(\frac{\phi^2(r)}{2} - \frac{\phi^4(r)}{4!} + \cdots\right) \\
&= \pi r^2 - \frac{\pi^2 r^4}{3} + \cdots.
\end{aligned} \tag{16}$$

Hence

$$\pi r^2 - \frac{\pi^2 r^4}{3} < A(r) < \pi r^2. \tag{17}$$

Now

$$\begin{aligned}
P(\{i \text{ is isolated in } \mathcal{G}(n, r(n))\}) &= (1 - A(r(n)))^{n-1} \\
&> (1 - \pi r^2(n))^{n-1}.
\end{aligned} \tag{18}$$

Also

$$\begin{aligned}
&P(\{i \text{ and } j \text{ isolated in } \mathcal{G}(n, r(n))\}) \\
&\quad < (A(2r(n)) - A(r(n)))(1 - \tfrac{3}{2} A(r(n)))^{n-2} \\
&\qquad + (1 - A(2r(n)))(1 - 2A(r(n)))^{n-2}
\end{aligned} \tag{19}$$

where the first term on the right-hand side above takes into account the case where the distance between i and j is between $r(n)$ and $2r(n)$. Substituting (18) and (19) in (15) and using (17), we get

$$\begin{aligned}
P^{(1)}(n, r(n)) &\geq n(1 - \pi r^2(n))^{n-1} - n(n-1) \\
&\quad \cdot \left(\left(3\pi r^2(n) + \frac{\pi^2 r^4(n)}{3}\right)\right. \\
&\quad \cdot \left(1 - \frac{3}{2}\left(\pi r^2(n) - \pi^2 r^4(n)/3\right)\right)^{n-2} \\
&\quad \left. + \left(1 - 2(\pi r^2(n) - \pi^2 r^4(n)/3)\right)^{n-2}\right).
\end{aligned}$$

Using Lemmas 5.1 and 5.2, for $\pi r^2(n) = \frac{\log n + \kappa}{n}$, and any fixed $\theta > 1$ and $\epsilon, \epsilon' > 0$ we have

$$\begin{aligned} P^{(1)}(n, r(n)) \geq & \, \theta e^{-\kappa} - n(n-1) \\ & \cdot (3(1+\epsilon')\pi r^2(n) e^{-(3/2)(n-2)\pi r^2(n)} \\ & + (1+\epsilon') e^{-2(n-2)\pi r^2(n)}) \\ \geq & \, \theta e^{-\kappa} - (1+\epsilon) e^{-2\kappa}, \qquad \text{for all } n > N(\epsilon, \theta, \kappa). \end{aligned}$$

Now, replace κ by κ_n where $\lim_{n\to\infty} \kappa_n = \overline{\kappa}$. Then, for any $\epsilon > 0$, $\kappa_n \leq \overline{\kappa} + \epsilon$ for all $n \geq N'(\epsilon)$. Also, the probability of an isolated node is monotone decreasing in κ. Hence

$$P^{(1)}(n, r(n)) \geq \theta e^{-(\overline{\kappa}+\epsilon)} - (1+\epsilon) e^{-2(\overline{\kappa}-\epsilon)}$$

for $n \geq \max\{N(\epsilon, \theta, \overline{\kappa}+\epsilon), N'(\epsilon)\}$. Taking limits

$$\liminf_{n\to\infty} P^{(1)}(n, r(n)) \geq \theta e^{-(\overline{\kappa}+\epsilon)} - (1+\epsilon) e^{-2(\overline{\kappa}-\epsilon)}.$$

Since this holds for all $\epsilon > 0$ and $\theta < 1$, and since

$$P^{(1)}(n, r(n)) \leq P_d(n, r(n))$$

the results follow. □

Corollary 5.1: The asymptotic probability that graph $\mathcal{G}(n, r(n))$ has an isolated node and is disconnected is strictly positive if $\pi r^2(n) = \frac{\log n + \kappa_n}{n}$ and $\limsup_n \kappa_n < +\infty$.

B. Upper Bound on Throughput Capacity of Random Networks

The key to the upper bound, as in the case of Arbitrary Networks, is to note that each transmission consumes valuable area.

Lemma 5.4: The number of simultaneous transmissions on any particular subchannel is no more than

$$\frac{4}{c_{11}\pi\Delta^2 r^2(n)}$$

in the Protocol Model.

Proof: Suppose node X_i in Fig. 7 transmits successfully to node X_j on the mth subchannel. Then no other node X_k within a distance $\Delta r(n)$ of X_j can be simultaneously receiving a separate transmission on the same subchannel due to the requirements (3) and (4) and the triangle inequality.

Hence disks of radius $\frac{\Delta r(n)}{2}$ centered at each receiver on the mth subchannel are disjoint. Since the area of each such disk is $\frac{c_{11}\pi\Delta^2 r^2(n)}{4}$, it follows that the network can support no more than $\frac{4}{c_{11}\pi\Delta^2 r^2(n)}$ simultaneous transmissions on the mth subchannel. □

Noting that each transmission over the mth subchannel is of W_m bits per second, by adding all the transmissions taking place at the same time over all the M subchannels, we see that they cannot total more than

$$\frac{4}{c_{11}\pi\Delta^2 r^2(n)} \sum_{m=1}^{M} W_m = \frac{4}{c_{11}\pi\Delta^2 r^2(n)} W$$

bits per second in the Protocol Model.

Now let $\overline{L}$ denote the mean length of a line connecting two independently and uniformly distributed points on S^2. Then the mean length of the path of packets is at least $\overline{L} - o(1)$ since there is always a node within a distance $o(1)$ of a point on the sphere *whp*. (This was shown in Lemma 4.8). Thus the mean number of hops taken by a packet is at least $\frac{\overline{L}-o(1)}{r(n)}$. Since each source generates $\lambda(n)$ bits per second, there are n sources, and each bit needs to be relayed on the average by at least $\frac{\overline{L}-o(1)}{r(n)}$ nodes, it follows that the total number of bits per second served by the entire network needs to be at least $\frac{\overline{L}-o(1))n\lambda(n)}{r(n)}$. To ensure that all the required traffic is carried, we therefore need

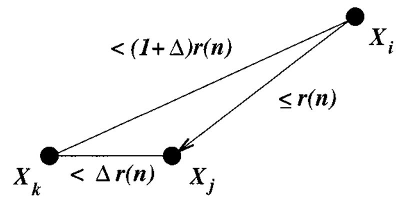

Fig. 7. X_k cannot receive at the same time as X_j on the same subchannel.

$$\frac{(\overline{L} - o(1))n\lambda(n)}{r(n)} \leq \frac{4W}{c_{11}\pi\Delta^2 r^2(n)}.$$

Thus

$$\lambda(n) \leq \frac{c_{12} W}{\Delta^2 n r(n)}.$$

From the previous section we know that $r(n) > \sqrt{\frac{\log n}{\pi n}}$ is necessary to guarantee connectivity *whp*. Hence we obtain the following upper bound.

Theorem 5.1: For Random Networks on S^2 under the Protocol Model, there is a deterministic constant $c' < +\infty$, not depending on n, Δ, or W, such that

$$\lim_{n\to\infty} \text{Prob}\left(\lambda(n) = \frac{c' W}{\Delta^2 \sqrt{n \log n}} \text{ is feasible}\right) = 0.$$

Note that just as in Theorem 4.1 the number of subchannels is irrelevant.

For the Physical Model, the upper bound is as follows.

Theorem 5.2: For Random Networks on S^2 under the Physical Model, there is a deterministic sequence $\epsilon(n) \to 0$, not depending on N, α, β, or W, such that

$$\lim_{n\to\infty} \text{Prob}\left(\lambda(n) = \sqrt{\frac{8}{\pi}} \frac{W}{\overline{L}(\beta^{1/\alpha} - 1)} \frac{1+\epsilon(n)}{\sqrt{n}} \text{ is feasible}\right) = 0$$

where $\overline{L}$ is the mean distance between two points independently and uniformly distributed on the unit area surface of the sphere.

Proof: In Section II we have shown that $\sqrt{\frac{8}{\pi}} \frac{W}{\Delta} \sqrt{n}$ bit-meters per second is an upper bound on the transport capacity for an Arbitrary Network under the Protocol Model. We will now show that any upper bound on the transport capacity for Arbitrary Networks under the Protocol Model is also an upper bound on the transport capacity for Random Networks under the Physical Model. This will prove the assertion since there are n nodes, each having its destination at least $\overline{L} - o(1)$ meters away on average.

Consider any set of successful simultaneous transmissions under the Physical Model for Random Networks. If X_i is successfully transmitting to X_j over a the mth subchannel, at the same time that X_k is also successfully transmitting to X_ℓ over the same subchannel, then from (5)

$$\frac{\frac{P}{|X_i-X_j|^\alpha}}{\frac{P}{|X_k-X_j|^\alpha}} \geq \beta$$

and so

$$|X_k - X_j| \geq (1+\Delta)|X_i - X_j|$$

where $\Delta := (\beta^{\frac{1}{\alpha}} - 1)$. Hence any set of simultaneous transmissions feasible for Random Networks under the Physical Model is also feasible in the Protocol Model for Arbitrary Networks. Thus the upper bound on the transport capacity for the latter also holds for the former. □

VI. Throughput Capacity of Random Networks on Planar Disk

The reader may wonder if the capacity is much different when the network is located on a disk in the two-dimensional plane, rather than on the surface of a sphere. The key issue is whether hot spots created at the center of the domain by several origin–destination pairs routing their traffic through the center will make it a bottleneck. The answer is no. The order of the capacity is unchanged for the Protocol Model, and the earlier orders for the lower and upper bounds for the Physical Model continue to hold.

Clearly, the arguments for the earlier upper bounds still hold, in view of the same necessary condition on the radius for connectivity (see [10]) in Random Networks under the Protocol Model, and the same reduction of Random Networks under the Physical Model to Arbitrary Networks under the Protocol Model.

The critical issue is to show that the earlier lower bounds can still be achieved. We show this by using the same tessellation-based scheme as on S^2. Let $\mathcal{G}$ be the disk of unit area on the plane on which the nodes are randomly located. Note that just as on S^2, the probability that a randomly chosen line on $\mathcal{G}$ intersects a disk of radius $2\rho(n)$ is no more than $c_5\sqrt{\frac{\log n}{n}}$. This applies even to disks of radius $2\rho(n)$ in the center of $\mathcal{G}$. Thus no unduly hot spots are expected to occur at the center of the domain $\mathcal{G}$. The key result to show however is that with high probability no hot spots are created *anywhere*. That is, we need to show the analog of Lemma 4.13 that the number of lines intersecting every cell is less than $c_5\sqrt{n\log n}$ *whp*. Lemma 4.11 and Corollary 4.1 are not applicable any more since we are not on S^2. However, we can circumvent this problem as follows.

We map $\mathcal{G}$ into a large sphere of radius M by using an inversion map $f(\cdot)$. Consider a straight line L on $\mathcal{G}$. Let $f(L)$ denote the curve on S^2 which is the image of the line, and let $g(L)$ denote the corresponding geodesic on S^2 connecting the two end points. When M is large enough, every such $f(L)$ deviates from $g(L)$ by no more than a distance $\rho(n)$. That is, the distortion between the images of straight lines on the disk and the geodesics is very small.

Consider now a cell $V \subset \mathcal{G}$ of the tessellation $\mathcal{V}_n$ of $\mathcal{G}$. It is contained in a disk D of radius $2\rho(n)$. This disk is mapped into another disk $A = f(D) \subset S^2$. Let $A' \subset S^2$ be a disk in S^2 with the same center as A, but with a radius $2\rho(n)$ larger than that of A. It follows that a straight L_i on $\mathcal{G}$ intersects the disk D only if the corresponding geodesic $g(L_i)$ on S^2 intersects the disk A'. (The reason is that the enlargement of the radius of A accounts for the distortion involved in replacing the images of straight line by geodesics). We have already shown in Section IV-I that the uniform law of large numbers holds for the probability of randomly chosen geodesics intersecting disks. Mapping back into D on the plane shows that the uniform upper bound on the number of straight lines passing through the disks of radius $2\rho(n)$ applies with high probability.

Thus the same results for the capacity continue to hold.

Theorem 6.1: For Random Networks on a planar disk of unit area, the results of Theorems 4.1, 5.1, and 5.2 continue to hold, except that in Theorem 5.2, $\overline{L}$ is the mean distance between two points independently and uniformly distributed in the planar disk of unit area.

VII. Concluding Remarks

We have shown that under a Protocol Model of noninterference, the capacity of wireless networks with n randomly located nodes each capable of transmitting at W bits per second and employing a common range, and each with randomly chosen and therefore likely far away destination, is $\Theta\left(\frac{W}{\sqrt{n\log n}}\right)$. This is true whether the nodes are located on the surface of a three-dimensional sphere or on a planar disk. Even when the nodes are optimally placed in a disk of unit area, and the range of each transmission is optimally selected, a wireless network cannot provide a throughput of more than $\Theta\left(\frac{W}{\sqrt{n}}\right)$ bits per second to each node for a distance of the order of 1 m away. In fact, summing over all the bits transported, a wireless network on a disk of unit area in the plane cannot transport a total of more than $\Theta(W\sqrt{n})$ bit-meters per second, irrespective of how the load is distributed. Under a Physical Model of noninterference, the lower bounds are the same as those above for the Protocol Model, while the upper bounds on throughput are $\Theta\left(\frac{W}{\sqrt{n}}\right)$ for Random Networks and $\Theta\left(\frac{W}{n^{\frac{1}{\alpha}}}\right)$ for Arbitrary Networks.

Splitting the channel into several subchannels does not change any of these results.

These results have some implications that designers may want to consider. Perhaps efforts should be targeted at designing networks with small numbers of nodes.

On the positive side, the results show that modulo further medium access or adaptive routing restrictions, communication with nearby neighbors at constant bit rates can be provided in a dense clusters of nodes, since the source–destination distances then shrink in scaled length as $O\left(\frac{1}{\sqrt{n}}\right)$. This shows that scenarios envisaged in collections of smart homes, or networks with mostly close-range transactions and sparse long-range demands, are feasible.

We have not considered in this paper the additional burden in coordinating access to the wireless channel, and the additional burden caused by mobility and link failures and the consequent need to route traffic in a distributed and adaptive way. These can only further throttle capacity. It would be useful to quantify these additional burdens.

Another issue to be studied is delay. This will arise when the traffic is bursty or when nodes are mobile. These two sources of delay are markedly different.

Finally, spatial directivity in the antennas or beamforming will be advantageous in increasing the spatial concurrency of transmissions, since wireless networks can then behave like wired ones. Ephremides [18] has analyzed the medium access problem for a single channel and shown that when only ternary feedback from the channel can be used to schedule transmissions, the throughput of collision-free successful transmissions is the same as in the usual omnidirectional case. When node locations and demands are known and do not have to be figured out purely from ternary feedback, transmissions can be advantageously scheduled so that collisions are avoided, and the throughput can consequently be increased. However, this is a challenging proposition since transmissions from nodes will have to be carefully orchestrated. Such schemes may pose some technological challenges though for low-cost networks. Finally, there is the challange of a more information-theoretic formulation.

REFERENCES

[1] "What you will want next and the really smart house," *Newsweek*, May 30, 1999.

[2] S. Ramanathan and M. Steenstrup, "A survey of routing techniques for mobile communication networks," *Mobile Networks and Appl.*, vol. 1, no. 2, pp. 89–104, 1996.

[3] R. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Inform. Theory (Special Issue on Random Access Communications)*, vol. IT-31, Mar. 1985.

[4] P. Karn, "MACA: A new channel access method for packet radio," in *Proc. 9th Computer Networking Conf.*, Sept. 1990, pp. 134–140.

[5] V. Bharghavan, A. Demers, S. Shenkar, and L. Zhang, "MACAW: A media access protocol for wireless LANs," in *Proc. SIGCOMM'94 Conf. on Communications Architectures, Protocols and Applications*, Aug. 1994, pp. 212–225.

[6] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11–1997, IEEE Computer Society LAN MAN Standards Committee, Ed., 1997.

[7] N. Bambos, S. Chen, and G. Pottie, "Radio link admission algorithms for wirelessd networks with power control and active link quality protection," in *Proc. IEEE INFOCOM*, Boston, MA, 1996.

[8] S. Ulukus and R. Yates, "Stochastic power control for cellular radio systems", Preprint, 1996.

[9] L. Valiant, "A theory of the learnable," *Commun. ACM*, vol. 27, pp. 1134–1142, Nov. 1984.

[10] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming*, W. M. McEneany, G. Yin, and Q. Zhang, Eds. Boston, MA: Birkhauser, 1998, pp. 547–566.

[11] A. Okabe, B. Boots, and K. Sugihara, *Spatial Tessellations Concepts and Applications of Voronoi Diagrams*. New York: Wiley, 1992.

[12] J. Stilwell, *Geometry of Surfaces*. New York: Springer-Verlag, 1992.

[13] R. C. Lynwood, *Groups and Geometry*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[14] J. A. Bondy and U. Murthy, *Graph Theory with Applications*. New York: Elsevier, 1976.

[15] V. N. Vapnik and A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory Probab. its Applic.*, vol. 16, no. 2, pp. 264–280, 1971.

[16] V. N. Vapnik, *Estimation of Dependences Based on Empirical Data*. New York: Springer-Verlag, 1982.

[17] M. Vidyasagar, *A Theory of Learning and Generalization*. London, U.K.: Springer-Verlag, 1997.

[18] A. Ephremides, "Some wireless networking problems with a theoretical conscience," in *Codes, Curves, and Signals: Common Threads in Communications*, A. Vardy, Ed., Aug. 1998.