Capacity Bounds for Three Classes of Wireless Networks: Asymmetric, Cluster, and Hybrid

Stavros Toumpis
Telecommunications Research Center Vienna
Donau-City-Straße 1/3., Vienna, Austria, A-1030
toumpis@ftw.at

ABSTRACT

We present capacity results for three classes of wireless ad hoc networks, using a general framework that allows their unified treatment. The results hold with probability going to 1 as the number of nodes in the network approaches infinity, and under a general model for channel fading.

We first study *asymmetric networks* that consist of n source nodes and around n^d destination nodes, communicating over a wireless channel. Each source node creates data traffic that is directed to a destination node chosen at random. When $\frac{1}{2} < d < 1$, an aggregate throughput that increases with n as $n^{\frac{1}{2}}$ is achievable. If, however, $0 < d < \frac{1}{2}$, bottlenecks are formed and the aggregate throughput can not increase faster than n^d.

We also consider *cluster networks*, that consist of n client nodes and around n^d cluster heads, communicating over a wireless channel. Each of the clients wants to communicate with one of the cluster heads, but the particular choice of cluster head is not important. In this setting, the maximum aggregate throughput is on the order of n^d, and it can be achieved with no transmissions taking place between client nodes.

We conclude with the study of *hybrid* networks. These consist of n wireless nodes and around n^d access points. The access points are equipped with wireless transceivers, but are also connected with each other through an independent network of infinite capacity. Their only task is to support the operation of the wireless nodes. When $\frac{1}{2} < d < 1$, an aggregate throughput on the order of n^d is achievable, through the use of the infrastructure. If, however, $0 < d < \frac{1}{2}$, using the infrastructure offers no significant gain, and the wireless nodes can achieve an aggregate throughput on the order of $n^{\frac{1}{2}}$ by using the wireless medium only.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless Communication*

General Terms

Performance, Design, Theory.

Keywords

Asymmetric traffic, capacity, cluster networks, hybrid networks, infrastructure support, mobile ad hoc networks, wireless access.

1. INTRODUCTION

Ongoing research on wireless ad hoc networks can be classified in two main areas: on the one hand, there is intense activity toward the development of *protocols*, for example on the media access [2], the routing [3], and higher [11] layers. On the other hand, there is also significant interest toward the establishment of the *capacity* of such networks, i. e., theoretical bounds on how much traffic they can support [8, 10, 9, 16, 14, 13, 6, 7, 17]. The work of this paper falls into this second line of investigation.

Research in the capacity of wireless networks has been significantly accelerated by the publication of the seminal paper of Gupta and Kumar [8]. There, the authors investigate the asymptotic behavior of the capacity of a class of two-dimensional random networks as the number of nodes n approaches infinity, under a uniform traffic assumption. The authors present a scheme that achieves **with high probability (*w. h. p.*)**, i.e., with probability approaching 1 as n approaches infinity, a rate of communication equal to $(n \log n)^{-\frac{1}{2}}$, up to a multiplicative constant, from each node to its randomly chosen destination. The authors also show that, with high probability, the n nodes cannot send data to their destinations with a per-node rate of communication greater than $(n \log n)^{-\frac{1}{2}}$, up to a (different) multiplicative constant.

In [8], and in other works that follow the same approach, such as [16, 13, 7], it is assumed that each of the n nodes picks one of the other $n-1$ nodes as the destination for its traffic. This implies that there are approximately as many destinations as there are sources. This is a good assumption for many scenarios. For example, if the network is used for supporting unicast two-way communication between a large number of users, it is indeed true that there will be roughly as many destinations as there are sources. On the other hand, there are many conceivable applications for which the assumption is inappropriate. As an example, consider a sensor network consisting of n sensor nodes and m actuator nodes, to whom the sensor nodes would like to send information, upon the sensing of an event. If m is dramatically different from n, then the induced traffic patterns will be fundamentally very different from those of [8].

Continuing along the tangent of [8], the authors of [10] and [9] consider the case of wireless ad hoc networks with infrastructure support. These networks consist of a collection of all-wireless nodes and a second collection of nodes that can also transmit and receive over the wireless channel but in addition are connected with each other through a wired infrastructure of infinite capacity. These nodes do not have any traffic requirements of their own, but are there to support the communication of the wireless nodes, in the same way that base stations support the communication of mobile terminals in cellular network. Such networks are also called hybrid, as they clearly share feature of both pure ad hoc networks and cellular networks. They are of great theoretical and also practical interest, as it is expected that future generation wireless systems will move from a purely cellular to a hybrid topology.

In this work we study three classes of wireless ad hoc networks, following the line of investigation initiated in [8] and continued in [10, 9]. The three classes are studied under a general framework that permits their unified treatment. In all cases we assume a general model for channel fading. In the spirit of [8], all our results hold with high probability, i. e., with probability approaching 1 as the number of nodes in the network approaches infinity.

We start by considering **asymmetric networks**. These consist of n source nodes and m destination nodes, each equipped with a wireless transceiver and communicating over a wireless channel, without the help of any infrastructure. Each source node creates data (with rate λ, common for all source nodes) that must be delivered to one of the destination nodes, chosen randomly. Because m is different from n, on the average the volume of traffic leaving each source node is different from the volume of traffic arriving at a destination node. As a result, when m is around n^d with $0 < d < \frac{1}{2}$, bottlenecks form around the destinations, and the aggregate throughput is only around n^d. On the other hand, when $\frac{1}{2} < d < 1$, the aggregate throughput can increase like $n^{\frac{1}{2}}$, and no bottlenecks are formed. Rather, as in the case of [8], all part of the network are more or less equally congested.

We proceed to study **cluster networks**. These consist of n client nodes and m cluster heads. Each of the n client nodes needs to maintain a bi-directional data traffic (of rate λ, common for all client nodes) with one of the cluster heads, but the choice of cluster head is not important. We show that if m is around n^d the maximum achievable aggregate throughput is around n^d for all $0 < d < 1$. Furthermore, this throughput can be achieved without the clients having to forward each other's traffic. Finally, the performance of the network is limited by the formation of bottlenecks around the cluster heads.

We conclude by studying **hybrid networks**, continuing the work in [10, 9]. These consist of n wireless nodes and m access points. The access points are equipped with wireless transceivers, but are also connected with each other through an infrastructure network of infinite capacity. The n wireless nodes would like to communicate with each other, but are free to use the infrastructure network to support their communication needs. We show that when m is on the order of n^d with $\frac{1}{2} < d < 1$, an aggregate throughput on the order of n^d is achievable. This throughput can be achieved by a scheme that does not require communication between wireless nodes. If, however, $0 < d < \frac{1}{2}$, it is best for the wireless nodes to ignore the presence of the infrastructure and route their packets exclusively using each other. In that case, an aggregate throughput on the order of $n^{\frac{1}{2}}$ is achievable. These results are similar to those reported in [10], however they are different in a few critical ways: Firstly, we require that the aggregate throughput is divided equally among the wireless nodes, so that no node is deprived of bandwidth in order to maximize the aggregate throughput. Also, we assume a more realistic channel model that includes a general model of fading, and we assume a totally random topology.

The rest of the work is organized as follows: In Section 2 we present detailed models for the three types of networks, along with an overview of our results. In Sections 3, 4, and 5 we provide proofs for the results, for asymmetric networks, cluster networks, and hybrid networks respectively. We conclude in Section 6.

2. NETWORK MODELS AND OVERVIEW OF RESULTS

2.1 Signal Propagation and Transceivers

In the following, we study networks of nodes equipped with transceivers used for communication over a wireless channel of bandwidth W. Here we list our assumptions regarding the signal propagation and the transceivers.

We assume that nodes cannot transmit and receive simultaneously (in other words, communication is half-duplex). Each node Z_i can transmit with any power $P_i \leq P_0$, where P_0 is a global maximum. When node Z_i transmits with power P_i, node Z_j receives the transmitted signal with power $G_{ij}P_i$, where $G_{ij} = Kf_{ij}|Z_i - Z_j|^{-\alpha}$. K is a constant, the same for all nodes, $|Z_i - Z_j|$ is the distance between nodes Z_i and Z_j, $\alpha > 2$ is the **decay exponent**, and the factor f_{ij} is the **fading coefficient**, a non-negative random variable that models fading, and does not change with time.

We assume that the expectation $E[f_{ij}] = 1$, and that $f_{ij} = f_{ji}$. Distinct fading coefficients are independent and identically distributed (iid). We also assume that their complementary cumulative distribution function $F^c(x)$ has a thin, exponentially decaying tail. Formally:

$$F^c(x) \equiv P[f_{ij} > x] \leq \exp[-qx] \; \forall x > x_1, \tag{1}$$

for some real and positive parameters q, x_1. In addition, we assume that there is a median value $f_m > 0$ such that $P[f_{ij} \geq f_m] \geq \frac{1}{2}$. Both of these assumptions are satisfied by most distributions used to model fading, for example the Nakagami, Ricean and Rayleigh distributions, and the trivial distribution, for which $P[f_{ij} = 1] = 1$.

Let $\{Z_t : t \in \mathcal{T}\}$ be the set of transmitting nodes at a given time, each node Z_t transmitting with power P_t. Let us assume that node Z_j, $j \notin \mathcal{T}$ is receiving a data packet from Z_i, $i \in \mathcal{T}$. Then the **signal to interference and noise ratio (SINR)** at node Z_j will be

$$\gamma_j = \frac{G_{ij}P_i}{\eta + \sum_{k \in \mathcal{T},\ k \neq i} G_{kj}P_k},$$

where η is the thermal noise power at the receiver, which is assumed the same for all nodes. We assume that the

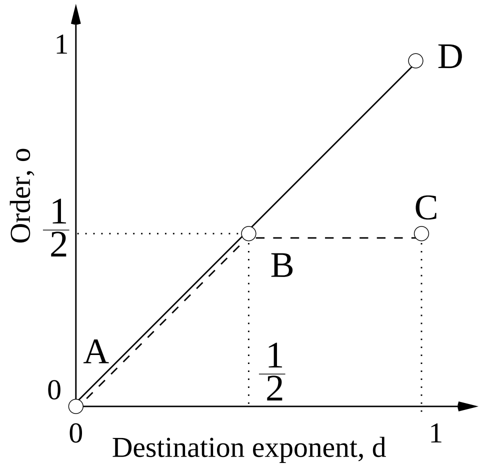

Figure 1: Lower and upper bounds on the order o of asymmetric networks, versus the destination exponent d. The open line intervals AB and BC represent lower bounds on the order and come from (5), and the open line interval AD represents an upper bound and comes from (6).

transmission of the packet will be successful if and only if, for the whole period of transmission, the transmission rate used, R_j, satisfies the inequality

$$R_j \leq f_R(\gamma_j) \equiv W \log_2(1 + \frac{1}{\Gamma}\gamma_j) \qquad (2)$$

where $\log_2(x)$ denotes the base-2 logarithm of x. With $\Gamma = 1$, the receiver achieves Shannon's capacity. With $\Gamma > 1$, (2) approximates the maximum rate that meets a given BER requirement under a specific modulation and coding scheme such as coded MQAM [4].

2.2 Asymmetric Networks

We consider a set of n **source nodes** $X_1, X_2, \ldots, X_n$, and m **destination nodes** $Y_1, Y_2, \ldots, Y_m$, placed randomly, uniformly and independently, in the two-dimensional area $\{(x, y) : |x|, |y| \leq \frac{1}{2}\}$ [1].

We use the symbols $<_a$, $>_a$, $\leq_a$, $\geq_a$ to denote that the corresponding inequality only holds **asymptotically**, i.e., for sufficiently large n. For example, $f(n) <_a g(n)$ means that there is a n_0 such that $f(n) < g(n)$ for all $n > n_0$. Using this notation, we make the assumption that:

$$D_1 n^d \leq_a m \leq_a D_2 n^d, \qquad (3)$$

where $0 < d < 1$ and $0 < D_1 < D_2$. We call d the **destination exponent**.

Each source node is creating data traffic with a fixed data rate λ bps, that must be delivered to one of the destination nodes. Each source selects its destination randomly, uniformly, and independently of the others. Both types of nodes are allowed to transmit and receive, as well as store data packets on their way to their destination.

The fundamental difference of this network from previously considered networks, such as the one in [8], is not that there are two types of nodes (sources and destinations), but the fact that their numbers n and m are different. In fact, because m is polynomially smaller than n, the traffic is asymmetric: on the average more packets are arriving at each destination, than there are leaving each source. Therefore, we will call such networks **asymmetric networks**. In this work we determine the effects of the relation of n and m, as expressed in (3), on the performance of asymmetric networks.

In particular, we define the **capacity** $C(n)$ **of the network** as the supremum of all rates $\lambda(n)$ that are uniformly achievable by all sources in the network, multiplied by their number n. In other words, the capacity is the supremum of all achievable aggregate throughputs, under the requirement that source nodes create traffic with the same rate. Since the locations of the nodes, the destination of each data stream, and the fading coefficients are random, the capacity of the network is a random variable. We also define the **order** o of the network to be the supremum of all exponents s for which $C(n) \geq n^s$ with high probability:

$$o = \sup\{s : \lim_{n\to\infty} P[C(n) \geq n^s] = 1\}. \qquad (4)$$

In other words, the order of the network is the largest exponent with which the capacity is guaranteed to be increasing, ignoring factors that are smaller than polynomial, such as poly-logarithmic factors of the form $k_1(\log n)^{k_2}$. Under this setting, the following theorem can be shown to hold:

THEOREM 1. *In asymmetric networks the capacity $C(n)$ is bounded with high probability as follows:*

$$C(n) \geq \left[\frac{3\alpha - 6}{3\alpha - 5}\right]\left[\frac{Wqf_m 5^{-\frac{\alpha}{2}}}{676\Gamma \log 2}\right] \times \begin{cases} \left[\frac{D_1}{4k_1^{\frac{1}{2}}(D_1 + 2D_2)}\right]\frac{n^{\frac{1}{2}}}{(\log n)^{\frac{3}{2}}} & \text{if } \frac{1}{2} < d < 1, \\ \left[\frac{D_1^2(1-2d)}{9D_2}\right]\frac{n^d}{\log n} & \text{if } 0 < d < \frac{1}{2}, \end{cases} \qquad (5)$$

$$C(n) \leq \left[\frac{4\alpha D_2 W}{\log 2}\right] n^d \log n. \qquad (6)$$

Consequently, we have that:

$$o = d, \text{ when } 0 < d < \frac{1}{2},$$
$$\frac{1}{2} \leq o \leq d, \text{ when } \frac{1}{2} < d < 1.$$

As will be shown in the proof of the theorem, the upper bound on the capacity originates from the need of the nodes to share the area around the destinations. When $d < \frac{1}{2}$, bottlenecks are formed around the destinations, limiting the capacity of the network. If, however, $\frac{1}{2} < d < 1$, no bottlenecks are formed around the destinations, and the capacity increases at least as fast as $n^{\frac{1}{2}}$, but not faster than n^d.

An important practical implication of Theorem 1 is that designers of networks should avoid traffic patterns that are characterized by an extreme convergence of traffic streams. Although this is intuitively clear without any need for math, Theorem 1 provides a quantitative rule: the number of destinations should be at least on the order of $n^{\frac{1}{2}}$, where n is the number of sources.

[1] A preliminary study of these networks appeared in [17].

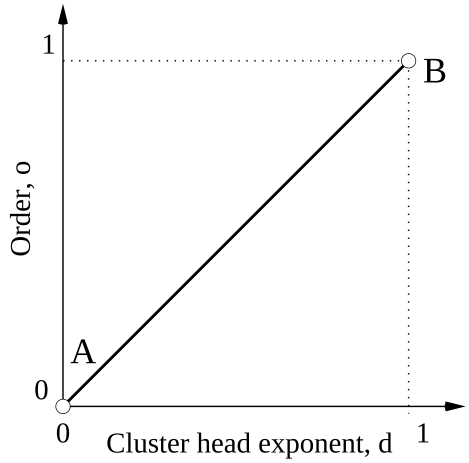

Figure 2: The order of cluster networks versus the cluster head exponent d.

2.3 Cluster Networks

Cluster networks consist of a set of n **client nodes** X_1, X_2,..., X_n, and m **cluster heads** Y_1, Y_2, ..., Y_m, placed randomly, uniformly and independently, in the area $\{(x, y) : |x|, |y| \leq \frac{1}{2}\}$, and communicating over a wireless channel of bandwidth W. Regarding the relative numbers of mobile nodes and access points, we assume that (3) continues to hold:

$$D_1 n^d \leq_a m \leq_a D_2 n^d,$$

where $0 < d < 1$ and $0 < D_1 < D_2$. We now call d the **cluster head exponent**.

Regarding the traffic model, we assume that each client wants to establish a bi-directional communication (with rate λ(n) in each direction) with *any* of the cluster heads. This model approximates well the traffic patterns that exist in wireless networks that operate using hierarchical, clustering protocols, as for example Bluetooth [15]. It also approximates well a scenario in which each of the cluster heads is connected to the outside world through a data connection of infinite capacity, and clients are interested in sending to and receiving data packets from the outside world.

We define the **capacity** $C(n)$ **of the network** as the supremum of all rates $\lambda(n)$ that are uniformly achievable by all data streams in the network, multiplied by their number $2n$ (there are two data streams for each client, one for the uplink connection to its selected cluster head, and another in the opposite direction). As in the previous case, the capacity is a random variable, whose statistics depend on n. The **order** o of cluster networks is defined as:

$$o = \sup\{s : \lim_{n\to\infty} P[C(n) \geq n^s] = 1\}.$$

THEOREM 2. *In cluster networks the capacity is bounded with high probability as follows:*

$$C(n) \geq \left[\frac{W D_1 q f_m}{900(\log 2) k_1 \Gamma}\right] \left[\frac{3\alpha - 6}{3\alpha - 5}\right] 5^{-\frac{\alpha}{2}} \frac{n^d}{(\log n)^2}, \qquad (7)$$

$$C(n) \leq \left[\frac{4\alpha D_2 W}{\log 2}\right] n^d \log n. \qquad (8)$$

Therefore, the order o is equal to the cluster head exponent d: $o = d$.

The theorem shows that, ignoring poly-logarithmic factors of the form $k_1(\log n)^{k_2}$, the capacity increases with n roughly as n^d. The upper bound of (8) comes from the need of the network to share the area around the cluster heads. Therefore, the larger d is, the faster capacity increases with n.

In the context of networks that use clustering, the theorem suggests that, from a capacity perspective, it is important that the size of clusters remains bounded. However, that would imply that the number of clusters increases linearly with n. If network designers are not willing to accept such a large number of clusters, they should be ready to sacrifice part of the capacity. The exact tradeoff is very simple, and is captured by Theorem 2.

In the context of networks where the cluster heads are gateways to the outside world, the theorem suggests that there is no limit to how many access points are needed: the greater the investment of the network provider (i. e., the larger d is), the larger the capacity is going to be. Again, the tradeoff is very simple and is captured by Theorem 2.

2.4 Hybrid Networks

Hybrid networks consist of a set of n **wireless nodes** X_1, X_2,..., X_n, and m **access points** Y_1, Y_2, ..., Y_m, placed randomly, uniformly and independently, in the same two-dimensional area $\{(x, y) : |x|, |y| \leq \frac{1}{2}\}$, and communicating over a wireless channel of bandwidth W. Regarding the relative numbers of wireless nodes and access points, we assume that (3) continues to hold:

$$D_1 n^d \leq_a m \leq_a D_2 n^d,$$

where $0 < d < 1$ and $0 < D_1 < D_2$. We now call d the **access point exponent**. We assume that the access points are connected with each other through a data link of infinite capacity that does not consume any of the available bandwidth W.

Regarding the traffic model, each wireless node is the source of a single stream, and the destination of a single stream. A node cannot be the source and destination of the *same* stream. Apart from this restriction, all other combinations of sources and destinations are equally probable. The access points do not have any communication needs of their own, but are there to support the communication of the wireless nodes.

This network shares important common characteristics with both pure wireless ad hoc networks and also pure cellular networks: On the one hand, it partly consists of a large number of wireless nodes that communicate over a wireless channel and can route each other's traffic, as in wireless ad hoc networks. On the other hand, the wireless nodes are supported by a set of access points that form an independent network with infinite capacity and do not have any traffic needs of their own; their role is similar to that of base stations in cellular networks.

As in the previous cases, we define the **capacity** $C(n)$ **of the network** as the supremum of all rates λ that are uniformly achievable by all data streams in the network, multiplied by their number n. The **order** o of hybrid networks is defined as:

$$o = \sup\{s : \lim_{n\to\infty} P[C(n) \geq n^s] = 1\}.$$

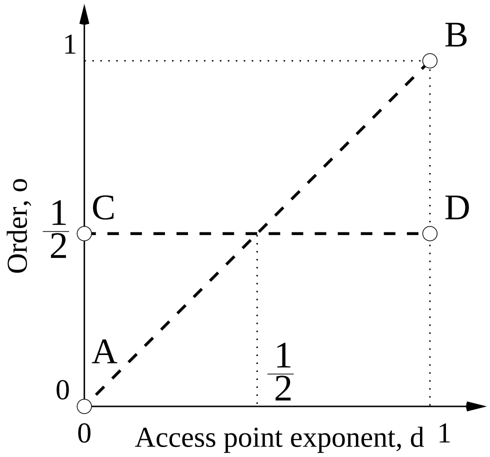

Figure 3: Lower bounds on the order of hybrid networks versus the access point exponent d. The open interval AB comes from (9), and the open interval CD comes from (10).

THEOREM 3. *In hybrid networks the capacity is bounded with high probability as follows:*

$$C(n) \geq \frac{1}{2}\left[\frac{WD_1 q f_m}{900(\log 2)k_1\Gamma}\right]\left[\frac{3\alpha-6}{3\alpha-5}\right]5^{-\frac{\alpha}{2}}\frac{n^d}{(\log n)^2}, \quad (9)$$

$$C(n) \geq \left[\frac{10^{-\frac{a+3}{2}}}{648}\frac{3\alpha-6}{3\alpha-5}\frac{Wqf_m}{\Gamma}\right]n^{\frac{1}{2}}(\log n)^{-\frac{3}{2}}. \quad (10)$$

Consequently, the order is bounded as follows:

$$o \geq \max\{d, \frac{1}{2}\}.$$

The theorem suggests that there is a minimum investment that is required in order for the infrastructure to have an effect on the performance of the network. In particular, if there are n wireless nodes, more than $n^{\frac{1}{2}}$ access points are needed. A very similar result was first reported in [10]. Our setup, however, is different in a number of critical ways: Firstly, we require that all wireless nodes are guaranteed the same throughput, and no nodes are allowed to starve in order for the aggregate throughput to be maximized. In addition, the locations on the access points are random, and we assume a more realistic channel model, that includes a general fading model.

3. ASYMMETRIC NETWORKS

In this section we prove Theorem 1. We first show the lower bound of (5), by constructing a scheme that works w. h. p. and whose aggregate throughput exceeds that lower bound. We then show the upper bound of (6) by calculating the aggregate throughput that would be hypothetically achieved if interference could be removed from the useful signal of each receiver. We start with two technical lemmas:

3.1 Technical Lemmas

A simple question to ask is how many source nodes select a certain destination node as the sink for their packets. Clearly, since there are n sources and m destinations, on the average $\frac{n}{m}$ sources will be sending packets to each destination. In fact, something much stronger holds:

LEMMA 1. *(Number of sources per destination) Let b_i be the number of sources that have selected destination Y_i. Then for any $\epsilon > 0$, w. h. p.,*

$$\forall i \quad \frac{1-\epsilon}{D_2}n^{1-d} \leq b_i \leq \frac{1+\epsilon}{D_1}n^{1-d}.$$

Proof: We make use of Chernoff's bounds [12]: Let X be a binomially distributed random variable, with parameters k (the number of Bernoulli experiments) and p (the probability of success of each Bernoulli experiment). Then, for any $\delta \in (0, 1]$,

$$P[X < (1-\delta)kp] < \exp(-kp\frac{\delta^2}{2}). \quad (11)$$

Also, for any $\delta > 0$,

$$P[X > (1+\delta)kp] < \exp[-kpf(\delta)], \quad (12)$$

where $f(\delta) = (1+\delta)\log(1+\delta) - \delta$. By calculating the derivative, we immediately have that $f(\delta) > 0$ for $\delta > 0$.

Since each source chooses its destination independently of the others, b_i follows the binomial distribution, with number of experiments equal to n and probability of success equal to $\frac{1}{m}$. We have

$$\begin{aligned} & P\left[b_i < \frac{1-\epsilon}{D_2}n^{1-d}\right] \\ \leq_a \quad & P\left[b_i < (1-\epsilon)\frac{n}{m}\right] \quad \text{(using (3))} \\ < \quad & \exp(-\frac{n}{m}\frac{\epsilon^2}{2}) \quad \text{(using (11))} \\ \leq_a \quad & \exp(-\frac{\epsilon^2}{2D_2}n^{1-d}). \quad \text{(using (3))} \qquad (13) \end{aligned}$$

Similarly, but using (12) instead of (11), we arrive at:

$$P\left[b_i > \frac{1+\epsilon}{D_1}n^{1-d}\right] <_a \exp[-\frac{f(\epsilon)}{D_2}n^{1-d}]. \quad (14)$$

We note the basic inequality $P(\cup_{i=1}^n E_i) \leq \sum_{i=1}^n P(E_i)$, typically referred to as the *union bound*. Then:

$$\begin{aligned} & P[\frac{1-\epsilon}{D_2}n^{1-d} \leq b_i \leq \frac{1+\epsilon}{D_1}n^{1-d} \ \forall i] \\ \geq \quad & 1 - \sum_{i=1}^m \left\{ P\left[b_i < \frac{1-\epsilon}{D_2}n^{1-d}\right] \right. \\ & \left. + P\left[b_i > \frac{1+\epsilon}{D_1}n^{1-d}\right]\right\} \quad \text{(union bound)} \\ = \quad & 1 - m\left\{P\left[b_1 < \frac{1-\epsilon}{D_2}n^{1-d}\right] + \right. \\ & \left. P\left[b_1 > \frac{1+\epsilon}{D_1}n^{1-d}\right]\right\} \quad \text{(symmetry)} \\ \geq \quad & 1 - m\left\{\exp(-\frac{\epsilon^2}{2D_2}n^{1-d}) + \right. \\ & \left. \exp[-\frac{f(\epsilon)}{D_2}n^{1-d}]\right\} \quad \text{(using (13), (14))} \\ \to \quad & 1. \quad \text{(for } n \to \infty) \end{aligned}$$

Note that this lemma is closely related to the well-known Coupon Collector's Problem [5]. To see the connection, let n be the number of coupons a collector purchases and m be the number of distinct types of coupons that exist. The lemma gives uniform upper and lower bounds on the number of coupons collected from each type, when m is polynomially smaller than n.

□

Another simple question to ask is how large the fading coefficients f_{ij} (where $1 \leq i < j \leq n+m$) between the $n+m$ source and destination nodes are expected to be. Although arbitrarily large values are possible, because the complementary distribution function of the fading coefficients has an exponentially thin tail, w. h. p. all coefficients will be relatively small, as the next lemma shows:

LEMMA 2. *(Bound on fading coefficients) W. h. p.,*

$$\max_{1\leq i<j\leq n+m}\{f_{ij}\} \leq \frac{3}{q}\log n.$$

Proof: Let $F_{ij}(x) = \{f_{ij} > x\log n\}$. Then:

$$\begin{aligned} &P[\max_{1\leq i<j\leq n+m}\{f_{ij}\} \leq x\log n] \\ &= 1 - P[\cup_{1\leq i<j\leq n+m} F_{ij(x)}] \\ &\geq 1 - \sum_{1\leq i<j\leq n+m} P[F_{ij}(x)] \quad \text{(union bound)} \\ &\geq_a 1 - \frac{(n+m)(n+m-1)}{2} n^{-qx} \quad \text{(symmetry, (1))} \\ &\geq_a 1 - n^{2-qx}. \quad \text{(using (3))} \end{aligned}$$

Setting $x = \frac{3}{q}$, the result follows. □

3.2 Cell Lattice

Let $\lfloor x \rfloor$ be the greatest integer that is less than or equal to x. Let $\lfloor(\frac{n}{k_1\log n})^{\frac{1}{2}}\rfloor = r$, where k_1 is a constant to be specified later. As shown in Fig. 4, we divide the square region $\{(x,y) : |x|,|y| \leq \frac{1}{2}\}$, where all the nodes are placed, in a regular lattice of $g(n) = r^2$ cells. The following will hold:

$$\frac{n}{2k_1\log n} <_a g(n) = r^2 \leq \frac{n}{k_1\log n}. \quad (15)$$

Indeed, $g(n) = \lfloor(\frac{n}{k_1\log n})^{\frac{1}{2}}\rfloor^2 \leq \frac{n}{k_1\log n}$. Also, let $x(n) = (\frac{n}{k_1\log n})^{\frac{1}{2}} - \lfloor(\frac{n}{k_1\log n})^{\frac{1}{2}}\rfloor$, so that $0 \leq x(n) < 1$. We then note that $g(n) = [1 - x(n)(\frac{n}{k_1\log n})^{-\frac{1}{2}}]^2 \frac{n}{k_1\log n} >_a \frac{n}{2k_1\log n}$.

We denote the cells by $c_1, c_2, \ldots, c_{g(n)}$. In addition, each cell is identified by its coordinates (v_1, v_2) in the lattice, where $1 \leq v_1,\ v_2 \leq r$; the cell on the lower left corner has coordinates $(1,1)$. We define the coordinates of a node to be the coordinates of the cell in which the node lies (so two nodes may have the same coordinates). We call two cells **neighbors** if they share a common boundary edge, so that each cell has at most four neighbors. We call two nodes **neighbors**, if they lie in the same or neighboring cells.

Since there are n sources and $g(n)$ cells, each cell will contain $\frac{n}{g(n)}$ sources on the average. By applying the Chernoff bounds, we can derive a much stronger result:

LEMMA 3. *(Number of source nodes in cells) Let s_i be the number of source nodes in cell c_i. Then:*

$$P[\frac{k_1\log n}{2} \leq s_i \leq 4k_1\log n \quad \forall i] >_a 1 - \frac{2n^{1-\frac{k_1}{8}}}{k_1\log n}.$$

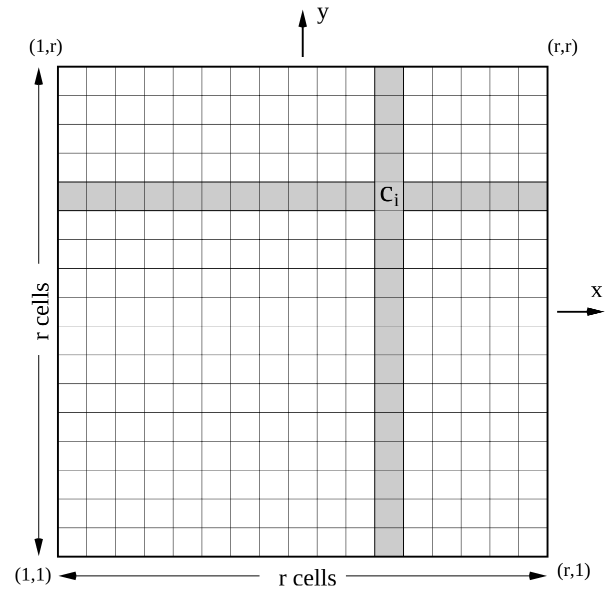

Figure 4: Partition of the square region $\{(x,y) : |x|,|y| \leq \frac{1}{2}\}$ into a regular lattice of r^2 cells. We define s_i as the number of source nodes in cell c_i, M_i as the number of source nodes lying in cells who share the same x-coordinate with c_i (the shaded cell column) and N_i as the number of destination nodes lying in cells who share the same y-coordinate with c_i (the shaded cell row).

Proof: Let us concentrate on a particular cell c_i. Since each source is placed in the cell with probability $\frac{1}{g(n)}$ and independently of the others, s_i is binomially distributed. Then:

$$\begin{aligned} &P[s_i < \frac{k_1\log n}{2}] \\ &\leq P[s_i < \frac{n}{2g(n)}] \quad \text{(using (15))} \\ &< \exp[-\frac{n}{8g(n)}] \quad \text{(using (11) with } \delta = \frac{1}{2}) \\ &\leq \exp[-\frac{k_1}{8}\log n] \quad \text{(using (15))} \\ &= n^{-\frac{k_1}{8}}. \end{aligned} \quad (16)$$

In a similar manner, but using (12) instead of (11), it may be shown that

$$P[s_i > 4k_1\log n] <_a n^{-k_1 f(1)} < n^{-\frac{k_1}{8}}. \quad (17)$$

Combining (16) and (17) with the union bound, and using (15), we arrive at the result. □

The fading coefficient between a source node and a destination node lying in the same cell may be arbitrarily low. However, the next lemma shows that, for a sufficiently large value of k_1, a second source node lying in the same cell that can act as a relay (because its fading coefficients with the other two nodes are large enough) will be available w. h. p.

LEMMA 4. *(Intra-cell communication) Let E_i, where $i = 1, \dots, m$ be the event that there is a source node X_j lying in the same cell with destination node Y_i, such that there is no third source node X_k lying in the same cell with the fading coefficients $f_{X_j X_k}, f_{X_k Y_i} \geq f_m$. Then:*

$$P(\cup_{i=1}^m E_i) \leq (4k_1 D_2 \log n) n^{d - \frac{k_1}{2}(\log 4 - \log 3)} + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n}. \tag{18}$$

Proof: Let the event

$$A = \left\{ \frac{k_1 \log n}{2} \leq s_i \leq 4k_1 \log n \quad \forall i \right\}.$$

We have:

$$\begin{aligned}
&P[\cup_{i=1}^m E_i] \\
&= P[\cup_i^m E_i | A] P(A) + P[\cup_i^m E_i | A^c] P(A^c) \\
&\leq P[\cup_{i=1}^m E_i | A] + P[A^c] \\
&\leq_a D_2 n^d P(E_1 | A) + P[A^c] \\
&\qquad \text{(union bound, symmetry, (3))} \\
&\leq_a D_2 n^d P(E_1 | A) + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n} \quad \text{(using Lemma 3)} \\
&\leq D_2 n^d (4k_1 \log n) (\frac{3}{4})^{\frac{k_1 \log n}{2}} + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n} \\
&= (4k_1 D_2 \log n) n^{d - \frac{k_1}{2}(\log 4 - \log 3)} + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n}.
\end{aligned}$$

The last inequality comes from noting that there are at most $4k_1 \log n$ source nodes for which a relay to Y_i may or may not exist. However, there are also at least $\frac{k_1 \log n}{2}$ source nodes, and the probability that each of them will be able to act as a relay is at least $\frac{1}{4}$. □

By Lemma 3, for a sufficiently large value of k_1, there are around $\log n$ source nodes per cell w. h. p. Therefore, a source node that wants to send a packet to a neighboring cell, will always be able to find a source node in that cell such that the fading coefficient between the two nodes is greater or equal to the median:

LEMMA 5. *(Inter-cell communication) Let F_i with $i = 1, \dots, n$ be the event that source node X_i has a neighboring cell in which all source nodes have fading coefficients with node X_i which are strictly smaller than the median f_m. Then:*

$$P[\cup_i^n F_i] \leq_a 4n^{1-\frac{k_1 \log 2}{2}} + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n}. \tag{19}$$

Proof: We will use conditioning on the event A, as in Lemma 4. Then:

$$\begin{aligned}
&P[\cup_i^n F_i] \\
&\leq P[\cup_i^n F_i | A] + P(A^c) \\
&\leq_a P[\cup_i^n F_i | A] + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n} \quad \text{(using Lemma 3)} \\
&\leq 4n (\frac{1}{2})^{\frac{k_1}{2} \log n} + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n} \\
&= 4n^{1-\frac{k_1 \log 2}{2}} + \frac{2n^{1-\frac{k_1}{8}}}{k_1 \log n}.
\end{aligned}$$

The last inequality comes by noting that there are n sources, each with at most 4 neighboring cells, and that under the conditioning on A each of these neighbors will have at least $\frac{k_1 \log n}{2}$ source nodes. □

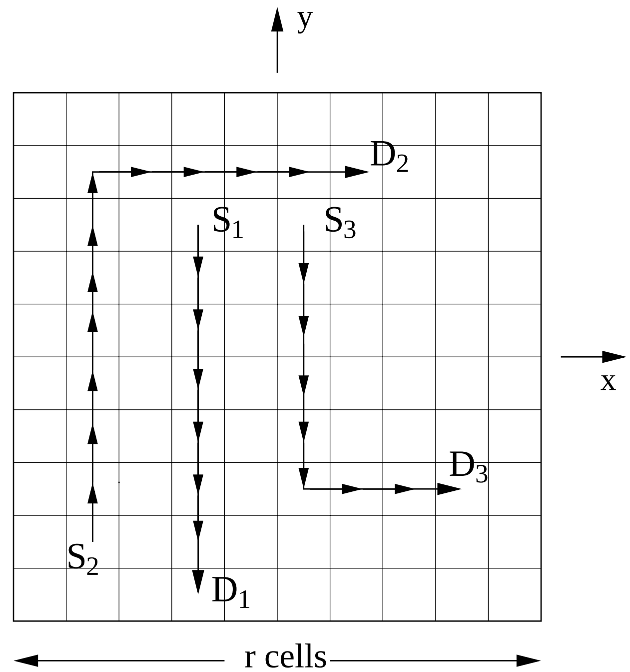

Figure 5: Examples of routes that follow the rules of Section 3.3

3.3 Routing Protocol

We now set $k_1 = 9 \geq \max\{8, \frac{2d}{\log 4 - \log 3}, \frac{2}{\log 2}\}$. For this selection of k_1, the right hand sides of both (18) and (19) go to 0. Therefore, w. h. p. the following routing rules are acceptable:

(i) If a source node X_j has data packets (possibly not created at X_j) that must be delivered to a destination node Y_i lying in the same cell, and $f_{X_j Y_i} < f_m$, X_j will transmit the data packets to another source node X_k lying in the same cell, for which $f_{X_j X_k} \geq f_m$ and $f_{X_k Y_i} \geq f_m$. Node X_k will then transmit the packet to the destination node Y_i. By Lemma 4, such a node exists w. h. p.

(ii) If the destination node Y_j of a source node X_i lies in a different cell from X_i, the packets of X_i are routed through intermediate cells. In particular, only communication between source nodes who lie in neighboring cells and whose mutual fading coefficient is greater than the median is allowed. In addition, the packets are first transmitted along cells whose x-coordinate is the same as the x-coordinate of the source, until they arrive at a cell whose y-coordinate is the same as the y-coordinate of the destination. Then, the packets are transmitted along cells whose y-coordinate is the same as the y-coordinate of the destination, until they arrive at a source node lying in the same cell with the destination. The existence of relays is guaranteed by Lemma 5, and our selection of k_1. Once the packets arrive at the cell of the destination, they are delivered to the destination as specified by rule **(i)**.

Following the spirit of [8], our routing protocol is using transmissions over as small distances as possible. In addition, as in [16], the routing protocol counters fading by taking advantage of the large number of source nodes that exist in each cell. To evaluate its performance, we must calculate the load that the routing protocol creates for each cell. The next two lemmas address this issue:

LEMMA 6. *Let M_i be the number of source nodes whose x-coordinate is the same as the x-coordinate of cell c_i. Let N_i be the number of destination nodes whose y-coordinate is the same as the y-coordinate of cell c_i. The following inequalities will hold w. h. p., for all i:*

$$M_i \leq 4(k_1 n \log n)^{\frac{1}{2}}, \tag{20}$$

$$N_i \leq \begin{cases} 4D_2(k_1 \log n)^{\frac{1}{2}} n^{d-\frac{1}{2}} & \text{if } \frac{1}{2} < d < 1, \\ \frac{4D_2}{(1-2d)D_1} & \text{if } 0 < d < \frac{1}{2}. \end{cases} \tag{21}$$

Proof: For M_i to be greater than $4(k_1 n \log n)^{\frac{1}{2}}$ for some i, it is necessary that the number of nodes in one of the cells is greater than $4k_1 \log n$. By Lemma 3 and our choice of k_1, this does not occur, w. h. p., so (20) follows immediately.

Regarding N_i, we note that because the destination nodes are placed independently and uniformly in the cells, N_i follows the binomial distribution, with number of tries m and probability of success $\frac{1}{r}$ (there are r^2 cells and success is declared if a destination node is placed in a row containing r of them).

We consider first the case $\frac{1}{2} < d < 1$, and we focus on a particular cell c_i. We have:

$$\begin{aligned} &P[N_i > 4D_2(k_1 \log n)^{\frac{1}{2}} n^{d-\frac{1}{2}}] \\ &\leq_a \quad P[N_i > \frac{2m}{r}] \quad \text{(using (3), (15))} \\ &< \quad \exp[-\frac{f(1)m}{r}] \quad \text{(using (12))} \\ &\leq \quad \exp\left\{-f(1)D_1[k_1 \log n]^{\frac{1}{2}} n^{d-\frac{1}{2}}\right\} \quad \text{(using (3), (15))}. \end{aligned}$$

By coupling the last inequality with the union bound, and noting that there are polynomially many cells, we arrive at (21) for the case $\frac{1}{2} < d < 1$.

For the case $0 < d < \frac{1}{2}$, we again use the Chernoff bound of (12), which can be written in the following different but equivalent form:

$$P[X > (1+\delta)kp] < \frac{\exp[\delta kp]}{(1+\delta)^{(1+\delta)kp}}. \tag{22}$$

We set

$$(1+\delta) = \frac{B}{D_2 n^d}\left[\frac{n}{2k_1 \log n}\right]^{\frac{1}{2}}, \tag{23}$$

where B will be specified shortly. We note that $\delta >_a 0$ and that $k_2 \equiv \frac{D_1 B}{D_2\sqrt{2}} \leq_a (1+\delta)\frac{m}{r} <_a B$. We then have:

$$\begin{aligned} &P[N_i > B] \\ &\leq_a \quad P\left[(1+\delta)\frac{m}{r}\right] \\ &\leq_a \quad e^B\left[\frac{D_2\sqrt{2k_1}}{B}\right]^{k_2} (\log n)^{\frac{k_2}{2}} n^{k_2(d-\frac{1}{2})} \quad \text{(by (22), (23))} \\ &\leq_a \quad (\log n)^{k_2} n^{k_2(d-\frac{1}{2})}. \end{aligned}$$

Combining the last inequality with the union bound, we have that

$$P[N_i \leq B \ \forall i] \geq 1 - \frac{n}{k_1 \log n}(\log n)^{k_2} n^{k_2(d-\frac{1}{2})}.$$

Equation (21) for the case $0 < d < \frac{1}{2}$ follows by setting $B = \frac{4D_2}{(1-2d)D_1}$. □

LEMMA 7. *(Number of routes arriving at a cell) Let r_i be the number of routes (each corresponding to a source) arriving, and possibly terminating, at cell c_i. Then w. h. p., for all i, the following bound holds:*

$$r_i \leq r_{\max}(n) \equiv \begin{cases} 4(1+\frac{2D_2}{D_1})(k_1 n \log n)^{\frac{1}{2}} & \text{if } \frac{1}{2} < d < 1, \\ \frac{9D_2}{(1-2d)D_1^2} n^{1-d} & \text{if } 0 < d < \frac{1}{2}. \end{cases}$$

Proof: Let r_{i1} be the number of routes that cross c_i while on their vertical leg (see Fig. 5). The sources of those routes share a common x-coordinate with c_i. Also, let r_{i2} be the number of routes that cross c_i while on their horizontal leg. Clearly, the destination nodes of these routes share a common y-coordinate with c_i. Each route crossing c_i will belong to one or both of the two types of routes, so necessarily $r_i \leq r_{i1} + ri2$. Therefore, it suffices to bound both r_{i1} and $ri2$ uniformly for all cells c_i.

Bounding r_{i1} is straightforward: since each node is the source of a single stream, $r_{i1} \leq M_i$. To bound r_{i2}, we note that, by Lemma 1, at most $\frac{1+\epsilon}{D_1} n^{1-d}$ routes can be terminating at each destination, with high probability. Therefore $r_{i2} \leq \frac{1+\epsilon}{D_1} n^{1-d} N_i$ w. h. p. Combining these inequalities we have that $r_i \leq M_i + \frac{(1+\epsilon)n^{1-d}}{D_1} N_i$ w. h. p., for all cells c_i. The result follows by setting $\epsilon = 2$ and using the bounds of Lemma 6. □

Since there are n routes, each requiring around $(\frac{n}{\log n})^{\frac{1}{2}}$ hops, and the total number of hops must be shared by $\frac{n}{\log n}$ cells, on the average each cell will be required to relay around $(n \log n)^{\frac{1}{2}}$ routes. Therefore, Lemma 7 implies that when $d > \frac{1}{2}$, no cell will have to carry more that its 'fair share' of the traffic, up to a multiplicative constant. If, however, $d < \frac{1}{2}$, then there are so few destinations, that a few unlucky cells (those on the same column of cells with a destination) will be required to serve around n^{1-d} routes, which is much more than their 'fair share' of traffic. In those cells, bottlenecks will form.

3.4 Time Division

We divide the $g(n) = r^2$ cells into nine regular sub-lattices, such that any two cells belonging in the same sub-lattice are separated by at least two cells belonging to different sub-lattices. In Fig. 6 we have shaded the cells belonging to one of the 9 sub-lattices.

We divide time into frames, and each frame into nine slots, each slot corresponding to a sub-lattice. At any time during that slot, only one node from each cell of the corresponding sub-lattice is allowed to *receive* (but many nodes in that cell may receive consecutively in the same slot). Because of the way we constructed the routing protocol, the transmitter of that transmission will have to lie in the same cell, or in one of the four neighboring cells. All transmissions will be with the maximum power P_0.

LEMMA 8. *(Lower bound on the SINR) The SINR γ_j at any source or destination node Z_j that is receiving is lower bounded w. h. p. by*

$$\gamma_j >_a \gamma_{min}(n) \equiv 5^{-\frac{\alpha}{2}} \left[\frac{3\alpha-6}{3\alpha-5}\right] \left[\frac{qf_m}{25}\right] \frac{1}{\log n}. \qquad (24)$$

Proof: We first bound the interference I_j. For this, we first note that by Lemma 2, w. h. p. no fading coefficient is greater than $\frac{3}{q}\log n$. Next, let $x_0 = \frac{1}{r}$ be the length of the sides of the cells, and let c_k be the cell in which the receiving node lies. Working as in [1], we note that the rest of the cells in the same sub-lattice are located along the perimeters of concentric squares, whose center is cell c_k. For example, there are 8 cells along the perimeter of the first square (fewer if the cell is at the edge of the network). Irrespective of the coordinates of c_k, all the cells of its sub-lattice are located along the perimeters of at most $\lfloor\frac{r-1}{3}\rfloor$ squares. There are at most $8i$ interferers corresponding to the i-th square, whose distance from the receiver will be at least $x_0(3i-2)$. Consequently, the interference at the receiver is upper bounded by

$$\begin{aligned} I_j &\le \left[\frac{3}{q}\log n\right] \sum_{i=1}^{\lfloor\frac{r-1}{3}\rfloor} \frac{8iKP_0}{[x_0(3i-2)]^\alpha} \\ &\le \left[\frac{3}{q}\log n\right] \frac{8KP_0}{x_0^\alpha}\ [1+\sum_{i=2}^{r}(3i-2)^{1-\alpha}] \\ &< \left[\frac{3}{q}\log n\right] \frac{8KP_0}{x_0^\alpha}\ [1+\int_0^r (3x+1)^{1-\alpha}\ dx] \\ &\le \left[\frac{3}{q}\log n\right] \frac{8KP_0}{x_0^\alpha} \left[\frac{3\alpha-5}{3\alpha-6}\right].\quad (\alpha>2) \qquad (25) \end{aligned}$$

We also need a lower bound on the power of the useful signal. Clearly, since the maximum possible distance that the useful signal will need to travel, under the routing assumptions, is $\sqrt{5}x_0$, and the fading coefficient between the transmitter and the receiver is greater than f_m, the power of the useful signal S_j is bounded w. h. p. as follows:

$$S_j \ge KP_0 f_m(\sqrt{5}x_0)^{-\alpha}. \qquad (26)$$

Combining (25) with (26), and noting that the thermal noise remains bounded, and therefore becomes negligible as $n \to \infty$, we arrive at (24). □

We now assume that all transmitters transmit with rate $f_R(\gamma_{\min})$. By Lemma 8, w. h. p. all transmissions will be successful.

3.5 Lower Bound

Proof of (5): We first calculate the throughput that is achieved by the scheme we have developed. We start by noting that the nodes of each cell are allowed to receive during only 1 out of 9 slots, and with rate equal to $f_R(\gamma_{\min}(n))$. The number of routes that will be crossing each cell c_i is upper bounded by $r_{\max}(n)$, determined by Lemma 7. Most of these routes will require one reception, however a few of these, in particular those whose destination lies in cell c_i, will require three transmissions. Therefore, each route, and its associated source node, is guaranteed a rate of communication $\lambda(n) = f_R(\gamma_{\min})\,[3 \times 9 \times r_{\max}(n)]^{-1}$. Multiplying by n, and substituting for $\gamma_{\min}(n)$ and $r_{\max}(n)$ from Lemmas 8 and 7 respectively, we see that our scheme achieves an aggregate throughput equal to the lower bound of (5). Since the capacity is the supremum of the aggregate throughputs of *all* possible schemes, it will necessarily be greater than the aggregate throughput of our scheme, and the result follows. □

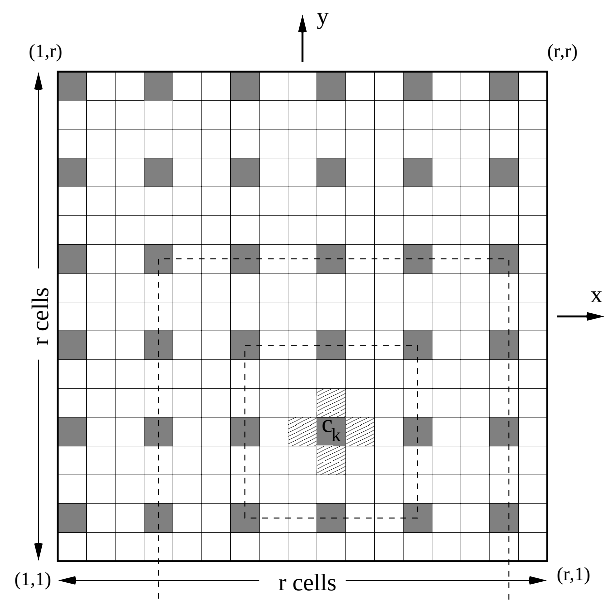

Figure 6: One of the 9 sub-lattices of cells appears shaded. Only nodes in that sub-lattice are allowed to receive in the corresponding slot, and only from nodes in the same or neighboring cells. The neighbors of cell c_k are lightly shaded. The cells belonging to the same sub-lattice as cell c_k may be placed in at most $\lfloor\frac{r-1}{3}\rfloor$ concentric squares of increasing size, centered at c_k.

One might think that our scheme performs very poorly in the case $d < \frac{1}{2}$, because our simplistic routing protocol creates bottlenecks that a more sophisticated protocol would avoid. However, the upper bound of (6) shows that this is not the case: Any other scheme would also not avoid the formation of bottlenecks, and would not perform better by more than a poly-logarithmic factor.

3.6 Upper Bound

Having determined the lower bound of (5), we now establish the upper bound of (6). We start with a straightforward technical lemma:

LEMMA 9. *Let $d_{\min}$ be the minimum of all distances between all mn source-destination pairs. Then $P[d_{\min} \le x] \le mn\pi x^2$.*

Proof: Let $H_{ij}(x)$ be the event $\{|X_i - Y_j| \le x\}$. Then:

$$\begin{aligned} P[d_{\min} \le x] &= P[\cup_{i,j} H_{ij}(x)] \\ &\le \sum_{i=1}^{n}\sum_{j=1}^{m} P[H_{ij}(x)] \quad \text{(union bound)} \\ &= nmP[H_{11}(x)] \quad \text{(using symmetry)} \\ &\le nm\pi x^2. \end{aligned}$$

The last inequality comes from noting that the nodes are placed in a square with surface area equal to 1, and that nodes X_1 and Y_1 will be within distance x of each other if Y_1 is placed on the intersection of the square with a disk of radius x, centered at node X_1. □

Proof of (6): The capacity is less than the aggregate throughput $T(n)$ that would have been achieved if all destination nodes were receiving, for all time, using the whole bandwidth, and without experiencing interference from competing transmissions. Lemmas 2 and 9 allow us to bound $T(n)$ in a straightforward manner, and w. h. p.:

$$\begin{aligned} T(n) &\le mW\log_2\!\left(1 + \frac{1}{\Gamma}\frac{K d_{\min}^{-\alpha}\frac{3}{q}\log n}{\eta}\right) \\ &\le mW\log_2\!\left(1 + \frac{3K}{\eta q \Gamma} n^{3\alpha}\log n\right) \\ &\le_a \left[\frac{4\alpha D_2 W}{\log 2}\right] n^d \log n. \end{aligned}$$

The first inequality comes from assuming that all destinations receive all the time and with no interference, and by using the bound on the value of the fading coefficients of Lemma 2. The second inequality comes by applying Lemma 9 with $x = n^{-3}$. The last one comes from (1) and using simple properties of the logarithm function. Since $C(n) \le T(n)$, the result follows. □

4. CLUSTER NETWORKS

In this section we prove Theorem 2. We start by noting that, because of the similarities between asymmetric and cluster networks, no real work is needed to prove (8). Indeed, the technique used in the proof of (6) can be used almost verbatim, by simply considering upper bounds on the aggregate throughput received at the *cluster heads*, as opposed to the *destination nodes*.

We next present a proof of (7). As was the case with (5), the proof is constructive, i.e., we define a communication scheme and show that the scheme achieves a prescribed aggregate throughput, w. h. p.

As in Section 3, we divide the square region $\{(x,y) : |x|,|y| \le \frac{1}{2}\}$, where all the nodes are placed, in the regular lattice of $g(n) = r^2$ cells of Fig. 4. However, we now set $r = \lfloor (\frac{D_1 n^d}{k_1 \log n})^{\frac{1}{2}} \rfloor$. It is straightforward to show that

$$\frac{D_1 n^d}{2k_1 \log n} <_a g(n) \le \frac{D_1 n^d}{k_1 \log n}. \tag{27}$$

Our first step is to uniformly bound the number of nodes per cell. Let s_i and d_i be the numbers of client nodes and cluster heads respectively in cell c_i. Then:

LEMMA 10. *(Number of nodes in cells)*

$$P[\frac{1}{2}\frac{k_1}{D_1}n^{1-d}\log n \le s_i \le 4\frac{k_1}{D_1}n^{1-d}\log n \quad \forall i] \to_{n\to\infty} 1, \tag{28}$$

$$P[\frac{k_1 \log n}{2} \le d_i \le 4\frac{k_1 D_2}{D_1}\log n \quad \forall i] >_a 1 - \frac{2D_1}{k_1}\frac{n^{d-\frac{k_1}{8}}}{\log n}. \tag{29}$$

Proof: Again, we make use of Chernoff's bounds of (11) and (12). To prove (28), let us concentrate on a particular cell c_i. Then:

$$\begin{aligned} &P[s_i < \frac{1}{2}\frac{k_1}{D_1}n^{1-d}\log n] \\ &\le P[s_i < \frac{n}{2g(n)}] \quad \text{(Using (27))} \\ &< \exp[-\frac{n}{8g(n)}] \quad \text{(Using (11) with } \delta = \tfrac{1}{2}\text{)} \\ &\le \exp[-\frac{k_1}{8D_1}n^{1-d}\log n]. \quad \text{(Using (27))} \end{aligned} \tag{30}$$

In a similar manner, but using (12) instead of (11), it may be shown that

$$P[s_i > 4\frac{k_1}{D_1}n^{1-d}\log n] <_a \exp[-f(1)\frac{k_1}{D_1}n^{1-d}\log n]. \tag{31}$$

We then note that

$$\begin{aligned} &P[\frac{1}{2}\frac{k_1}{D_1}n^{1-d}\log n \le s_i \le 4\frac{k_1}{D_1}n^{1-d}\log n \quad \forall i] \\ &\ge 1 - \sum_{i=1}^{g(n)}\Big\{P[s_i < \frac{1}{2}\frac{k_1}{D_1}n^{1-d}\log n] \\ &\quad + P[s_i > 4\frac{k_1}{D_1}n^{1-d}\log n]\Big\} \quad \text{(by the union bound)} \\ &\ge 1 - \frac{D_1 n^d}{k_1 \log n}\Big\{\exp[-\frac{k_1}{8D_1}n^{1-d}\log n] + \\ &\quad \exp[-f(1)\frac{k_1}{D_1}n^{1-d}\log n]\Big\} \quad \text{(by (27), (30), (31))} \\ &\to 1. \quad \text{(for } n \to \infty\text{)} \end{aligned}$$

Equation (29) follows in a similar manner. □

Lemma 10 guarantees that, for sufficiently large k_1, the number of cluster heads in each cell will be on the order of $\log n$. Therefore, client nodes may be restricted to transmit to only cluster heads lying in the same cell with them, and still be able to find a cluster head with which they have a sufficiently strong channel, despite the presence of fading. The next lemma formalizes this idea:

LEMMA 11. *Let F_i be the event that there is no cluster head Y_j lying in the same cell as client node X_i, such that the fading gain between them is equal or greater than f_m. The probability of the event $\cup_{i=1}^n F_i$ is upper bounded by*

$$P[\cup_{i=1}^n F_i] \le_a n^{1-\frac{k_1 \log 2}{2}} + \frac{2D_1}{k_1}\frac{n^{d-\frac{k_1}{8}}}{\log n}. \tag{32}$$

Proof: Let $A = \{\frac{k_1 \log n}{2} \leq d_i \quad \forall i\}$.

$$\begin{aligned} & P[\cup_{i=1}^n F_i] \\ \leq \quad & P[\cup_{i=1}^n F_i | A] + P(A^c) \\ <_a \quad & P[\cup_{i=1}^n F_i | A] + \frac{2D_1}{k_1} \frac{n^{d-\frac{k_1}{8}}}{\log n} \quad \text{(using (29))} \\ \leq \quad & n \left(\frac{1}{2}\right)^{\frac{k_1 \log n}{2}} + \frac{2D_1}{k_1} \frac{n^{d-\frac{k_1}{8}}}{\log n} \\ \leq \quad & n^{1-\frac{k_1 \log 2}{2}} + \frac{2D_1}{k_1} \frac{n^{d-\frac{k_1}{8}}}{\log n}. \end{aligned}$$

The third inequality comes from the union bound, and by noting that conditioning on A means that there are at least $\frac{k_1}{2 \log n}$ cluster heads in the same cell with client X_i, and the fading coefficient between any of them and X_i will be less than the median with probability at most $\frac{1}{2}$. □

We now set $k_1 = 9 > \max\{8d, \frac{2}{\log 2}\}$. For this choice of k_1, the right hand side of (29) goes to 1 and the right hand side of (32) goes to 0. We are now ready to prove our lower bound:

Proof of (7): It suffices to specify a transmission scheme and show that w. h. p. the aggregate throughput $T(n)$ of that scheme exceeds the lower bound.

By Lemma 11, w. h. p. for every client node there is a cluster head in the same cell, such that the fading coefficient between the two is equal or greater than f_m. We restrict that client node to transmit to and receive from only that cluster head.

In addition, we impose on the nodes the time division scheme of Section 3.4: time is divided in frames, and each frame in 9 slots. At any time during a slot, only a single node (either a cluster head or a client node) from each cell of the corresponding sub-lattice is allowed to transmit, and with maximum power. Since the receiver necessarily lies in the same cell, the lower bound on the SINR of Lemma 8 continues to hold. Therefore, if the transmitter transmits with rate $f_R(\gamma_{\min}(n))$, where $\gamma_{\min}(n)$ is given by (24), w. h. p. all transmissions will be successful.

By (28), w. h. p. there are less than $[4\frac{k_1}{D_1}n^{1-d}\log n]$ client nodes in each slot. We divide each slot in $2\times[4\frac{k_1}{D_1}n^{1-d}\log n]$ time intervals, each of which is devoted to the transmission of a packet either from or to a client node. Some of these time intervals will be wasted, however (28) shows that the aggregate throughput is not reduced by more than a factor of 8 (the ratio of the upper and lower bounds) because of this underutilization.

Each stream of data is guaranteed a rate of communication equal to $\lambda(n) = f_R(\gamma_{\min}(n))[2 \times 9 \times 4\frac{k_1}{D_1}n^{1-d}\log n]^{-1}$. Multiplying by $2n$ for the total number of streams, and substituting for $\gamma_{\min}(n)$ from Lemma 8, we arrive at the needed inequality. □

Ignoring constants, the upper and lower bounds of Theorem 2 differ only by a poly-logarithmic factor $(\log n)^3$, so both bounds are relatively tight. This implies that we can achieve an aggregate throughput very close to the capacity without resorting to multihop routing between the source nodes over the wireless channel. Therefore, our scheme is more closely related to traditional cellular schemes than to pure ad hoc schemes such as the one introduced in [8]. On the other hand, in contrast to traditional cellular schemes, the scheme requires that each source node is allowed to communicate with any of around $\log n$ cluster heads, in order to bypass the effects of fading.

5. HYBRID NETWORKS

In this section we prove the bounds (9) and (10) of Theorem 3. In fact, no new work is needed: Because of similarities of hybrid networks with previously studied networks, we can apply already known results.

For example, because of the similarities between cluster and hybrid networks, the wireless nodes can use for their communication the scheme that was developed in Section 4, for proving the lower bound of (7). In particular, wireless nodes do not transmit to each other, but rather transmit directly to an access point near by. The packet is then transmitted through the infinite capacity network to an access point close to its destination, and is then transmitted one more time through the use of the wireless interface to the destination. All the analysis of Section 4 goes through, if we substitute client nodes with wireless nodes and cluster heads with access points. The only difference is that, because each packet must be transmitted twice, the aggregate throughput is one half of the throughput achieved in cluster networks. Equation (9) follows immediately.

To derive (10), we consider the opposite extreme. In particular, we note that the n wireless nodes are free to ignore the wireline infrastructure of the access points, and establish a communication scheme using only themselves. This case has been studied independently in [16]. In that work it is shown (Eq. (4) of [16]) that, under uniform traffic conditions, it is possible to achieve a per-node throughput equal to:

$$\lambda_2(n) = \left[\frac{10^{-\frac{a+3}{2}}}{648} \frac{3\alpha - 6}{3\alpha - 5} \frac{Wqf_m}{\Gamma}\right] n^{-\frac{1}{2}} (\log n)^{-\frac{3}{2}}.$$

Multiplying by the number of nodes n, we derive the lower bound of (10).

6. CONCLUSIONS

We present capacity results for three classes of wireless ad hoc networks: asymmetric, cluster and hybrid networks. Although the three networks have fundamentally different topologies and therefore different applications, they have important underlying similarities, that permit a largely unified treatment.

We first consider asymmetric networks, that consist of n source nodes and m destination nodes, communicating over a wireless medium, without any help from a wired infrastructure. When m is around n^d with $\frac{1}{2} < d < 1$, an aggregate throughput that increases like $n^{\frac{1}{2}}$ is achievable. If, on the other hand, $d < \frac{1}{2}$, the maximum aggregate throughput increases like n^d, and the performance of the network is constrained by the existence of bottlenecks around the destinations. Although we do not prove this, it is intuitively clear that similar results hold when we have n destinations and n^d sources.

This result has an important implication that network designers should consider. Specifically, a certain amount of asymmetry on the traffic does not have an adverse effect, but beyond a certain point, the capacity of the network is reduced by the formation of bottlenecks. For applications in which the number of destinations m is a design parameter, and it is useful to minimize m (because, for example, destinations are more expensive) the network has a "sweet-spot": m should be around $n^{\frac{1}{2}}$. Using more destinations will not improve the performance significantly, but using fewer will severely reduce it.

We next consider cluster networks, in which n client nodes are interested with communicating with n^d cluster heads, and the choice of cluster head for each client is not important. Many interesting applications fall within this framework. For example, this traffic model approximates well the control traffic that is induced by the use of clustering protocols, such as Bluetooth [15]. Another application is a sensor network that consists of n sensors and n^d databases containing identical information, that the sensors would like to access. A final application is next generation cellular networks, in which n mobile users would like to communicate with the outside world through n^d base stations.

In this context, we make two discoveries: Firstly, the maximum aggregate throughput that is possible is around n^d. In other words, the network has no "sweet-spot": the larger the investment in cluster heads, the better the performance. Ideally, we would like the number of cluster heads to scale linearly with n, i.e., be on the order of n, so that the capacity scales also linearly with n. If, however, network designers are not prepared to consider such a large investment on cluster heads, they should be ready to sacrifice part of the capacity. The second discovery is that, irrespective of the value of d, the capacity can be achieved even if clients do not transmit to each other. In other words, advanced ad hoc routing protocols are of no use here, and designers should focus on the efficient handling of bottlenecks around the cluster heads. Those bottlenecks are *unavoidable*.

We finally consider hybrid networks. These consist of n wireless nodes, and an infinite-capacity network of n^d access points, that are also equipped with wireless transceivers. The access points have no communication need of their own, but are there to facilitate the communication between network nodes. Such networks were first studied in [10, 9], and are of great practical interest, as it is expected that future generation cellular systems will be using this hybrid topology.

Our main find is that more than $n^{\frac{1}{2}}$ access points are needed for the infinite-capacity infrastructure to have any effect on the performance of the network. When $d < \frac{1}{2}$, the wireless nodes should ignore the available infrastructure, and use each other for communication. That way, an aggregate throughput on the order of $n^{\frac{1}{2}}$ is achievable. If, however, $\frac{1}{2} < d < 1$, the wireless nodes should not depend on each other for routing their traffic, but rather should make heavy use of the infrastructure. A similar result was first reported in [10], however our result is different in a number of important ways: for example, we require that all nodes share the resources equally, we adopt a random topology, and we use a more realistic channel model that incorporates a general model for fading.

7. REFERENCES

[1] N. Bansal and Z. Liu. Capacity, delay and mobility in wireless ad-hoc networks. In *Proc. IEEE INFOCOM*, pages 1553–1563, San Francisco, CA, April 2003.

[2] T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad hoc networks. In *Proc. IEEE INFOCOM*, volume 2, pages 976–985, New York, NY, June 2002.

[3] J. Eriksson, M. Faloutsos, and S. Krishnamurthy. Scalable ad hoc routing: The case of dynamic addressing. In *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

[4] M. V. Eyoboğlu and J. G. D. Forney. Trellis precoding: Combined coding, precoding and shaping for intersymbol interference channels. *IEEE Trans. Inform. Theory*, 38(2):301–314, 1992.

[5] W. Feller. *An Introduction to Probability Theory and its Applications*, volume I. New York, NY: John Wiley and Sons, 3rd edition, 1975.

[6] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

[7] M. Grossglauser and D. Tse. Mobility increases the capacity of ad-hoc wireless networks. In *Proc. IEEE INFOCOM*, volume 3, pages 1360–1369, Anchorage, AL, Apr. 2001.

[8] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Trans. Inform. Theory*, 46(2):388–404, Mar. 2000.

[9] U. C. Kozat and L. Tassiulas. Throughput capacity of random ad hoc networks with infrastructure support. In *Proc. ACM MobiCom*, pages 55–65, San Diego, CA, Sep. 2003.

[10] B. Liu, Z. Liu, and D. Towsley. On the capacity of hybrid wireless networks. In *Proc. IEEE INFOCOM*, volume 2, pages 1543–1552, San Francisco, CA, April 2003.

[11] W. Lou, W. Liu, and Y. Fang. SPREAD: Enhancing data confidentiality in mobile ad hoc networks. In *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

[12] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge, UK: Cambridge University Press, 1st edition, 1995.

[13] R. Negi and A. Rajeswaran. Capacity of power constrained ad-hoc networks. In *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

[14] C. Peraki and S. D. Servetto. On the maximum stable throughput problem in random networks with directional antennas. In *Proc. ACM MobiHoc*, pages 76–87, Annapolis, MD, June 2003.

[15] T. Salonidis, Bhagwat, L. Tassiulas, and R. LaMaire. Distributed topology construction of bluetooth personal area networks. In *Proc. IEEE INFOCOM*, volume 3, pages 1577–1586, Anchorage, AL, Apr. 2001.

[16] S. Toumpis and A. J. Goldsmith. Large wireless networks under fading, mobility, and delay constraints. In *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

[17] S. Toumpis, A. J. Goldsmith, and J. Sayir. Capacity results for asymmetric wireless networks. In *Proc. Int. Zurich Sem.*, pages 180–183, Zurich, Switzerland, Feb. 2004.