Foundations and Trends® in
Networking
Vol. 1, No 2 (2006) 145–270

DOI: 10.1561/1300000002

Scaling Laws for Ad Hoc Wireless Networks: An Information Theoretic Approach

Feng Xue[1] and P. R. Kumar[2]

[1] *Communications Technology Laboratory/Intel Research, Intel Corporation, Santa Clara, California 95052, USA, feng.xue1@intel.com*

[2] *Coordinated Science Laboratory, and Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, Urbana, Illinois 61801, USA, prkumar@uiuc.edu*

Abstract

In recent years there has been significant and increasing interest in ad hoc wireless networks. The design, analysis and deployment of such wireless networks necessitate a fundamental understanding of how much information transfer they can support, as well as what the appropriate architectures and protocols are for operating them. This monograph addresses these questions by presenting various models and results that quantify the information transport capability of wireless networks, as well as shed light on architecture design from a high level point of view. The models take into consideration important features such as the spatial distribution of nodes, strategies for sharing the wireless medium, the attenuation of signals with distance, and how information is to be transferred, whether it be by encoding, decoding, choice of power level, spatio-temporal scheduling of transmissions, choice of multi-hop routes, or other modalities of cooperation between nodes. An important aspect of the approach is to characterize how the information hauling capacity scales with the number of nodes in the network.

The monograph begins by studying models of wireless networks based on current technology, which schedules concurrent transmissions to take account of interference, and then routes packets from their sources to destinations in a multi-hop fashion. An index of performance, called transport capacity, which is measured by the bit meters per second the network can convey in aggregate, is studied. For arbitrary networks, including those allowing for optimization of node locations, the scaling law for the transport capacity in terms of the number of nodes in the network is identified. For random networks, where nodes are randomly distributed, and source-destination pairs are randomly chosen, the scaling law for the maximum common throughput capacity that can be supported for all the source-destination pairs is characterized. The constructive procedure for obtaining the sharp lower bound gives insight into an order optimal architecture for wireless networks operating under a multi-hop strategy.

To determine the ultimate limits on how much information wireless networks can carry requires an information theoretic treatment, and this is the subject of the second half of the monograph. Since wireless communication takes place over a shared medium, it allows more advanced operations in addition to multi-hop. To understand the limitations as well as possibilities for such information transfer, wireless networks are studied from a Shannon information-theoretic point of view, allowing any causal operation. Models that characterize how signals attenuate with distance, as well as multi-path fading, are introduced. Fundamental bounds on the transport capacity are established for both high and low attenuation regimes. The results show that the multi-hop transport scheme achieves the same order of scaling, though with a different pre-constant, as the information theoretically best possible, in the high attenuation regime. However, in the low attenuation regime, superlinear scaling may be possible through recourse to more advanced modes of cooperation between nodes. Techniques used in analyzing multi-antenna systems are also studied to characterize the scaling behavior of large wireless networks.

1

Introduction

Over the past few years there has emerged a network information theory motivated by the twin goals of applicability as well as tractability vis-à-vis the rapidly emerging field of wireless networking. A central aspect of this theory is that the spatial aspects of the system, including locations of nodes and signal attenuation with distance, are explicitly modeled. Also, distance is intimately involved even in the performance measure of transport capacity that is analyzed. This theory has been used to develop bounds on the distance hauling capacity of wireless networks, feasibility results, scaling laws for network capacity as the number of nodes increases, and also suggest some insight into architectures. It establishes relationships for information transport in wireless networks between phenomena such as how radio signals attenuate with distance and the information hauling capacity of networks. It also connects the more recent field of networking with its emphasis on architecture and protocols with the more traditional field of communication theory with its emphasis on signals, transmitters and receivers. This text provides an account of the salient results.

The focus of this text is on ad hoc wireless networks, a topic which has aroused much interest in recent years. These are wireless networks without infrastructure; see Figure 1.1. Examples of technologies

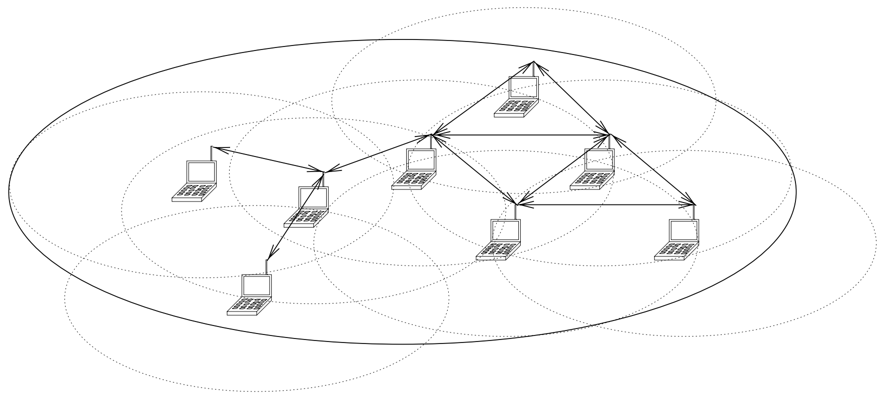

Fig. 1.1 An ad hoc wireless network.

envisioning such wireless networks are Bluetooth [29] through the use of scatter-nets, e.g., [13], IEEE802.11 [30] through the use of the Distributed Coordinated Function, e.g., [3], IEEE802.15.4 ZigBee networks [34, 33] deployed as multi-hop sensor networks, and IEEE802.16 [31, 32] deployed as mesh networks.

Such ad hoc wireless networks have been proposed to be operated in multi-hop mode: packets are relayed from node to node in several short hops until they reach their destinations. The top layer shown in the protocol stack in Figure 1.2, the *Transport Layer*, can address functionalities such as end-to-end reliable delivery of information, as well as regulating the rate at which data packets are pumped into the network so as to match it to the rate at which the network can carry information. The choice of the sequence of nodes along which to relay is the *routing* problem, and is addressed in the *Network Layer*. At each hop, a medium access control protocol is employed so that the reception at the receiver is not interfered with by another nearby transmitter, as well as to ensure that packets are retransmitted repeatedly, at least a few times, until an acknowledgement is received from the receiver. Another important functionality is *Power Control*. This addresses the power level at which a packet on a hop is transmitted. Proposals have been made to address it at the Network Layer [20, 16] or the Medium Access Control Layer [19, 15]. The *Physical Layer* addresses issues related to modulation, etc..

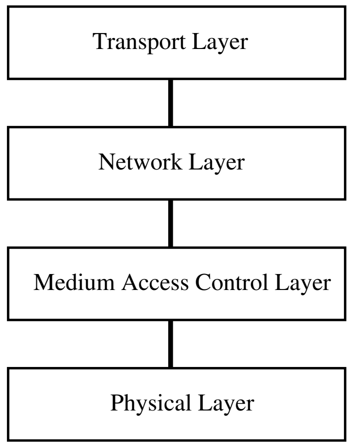

Fig. 1.2 A proposed protocol stack for wireless ad hoc networks.

One of the goals of this text is to present results on what the capacity of wireless networks is under the multi-hop model. Another question is whether the multi-hop mode of information transfer is indeed an appropriate mode. This is motivated by the fact that there are several alternative ways in which the wireless medium can be used. To study such alternatives takes us into the domain of network information theory, an area in which several apparently simple-looking problems have continued to defy characterization even after several decades of research. Another question of interest is how the information transfer capability of wireless networks scales as the number of nodes increases. This leads into the issue of scaling laws for wireless networks. Such results can help in understanding the complicated interactions in wireless networks, and to shed light on operating and designing more efficient networks.

It is to these questions that this text is addressed. At the same time, there are several issues that are excluded. Except peripherally, issues such as energy lifetime, latency, fairness, etc., are not centrally addressed.

The results presented in this text are as follows.

Sections 2–4 consider arbitrary wireless networks, where node locations are allowed to be optimized for network performance.

Section 2 introduces the definition of transport capacity, which takes into account not only the throughputs supported for source–destination pairs, but also the distances between sources and their destinations. Specifically it is the sum of the rates for source–destination pairs weighted by their distance. It introduces two variants of a so-called Protocol Model. This postulates a model for successful packet reception at a receiver, by specifying either a guard zone around a receiver or an interference footprint around a transmitter. This section presents results on the scaling behavior of the transport capacity of arbitrary wireless networks under the Protocol Model. It exhibits a square-root scaling law in the number of nodes in the network for the transport capacity.

Section 3 introduces the concepts of Exclusion Region and Interference Region. Building on them, it presents improved bounds on the transport capacity of arbitrary wireless networks under the Protocol Model.

Section 4 introduces the Physical Model, which models successful reception in terms of the received signal-to-noise-plus-interference ratio at a receiver. This section shows that there is a correspondence between the Protocol Model and the Physical Model, and then presents similar results on the scaling behavior of the transport capacity for arbitrary wireless networks under the Physical Model.

Sections 5–7 study the performance of random wireless networks, where nodes are distributed randomly over a domain and destinations for sources are randomly chosen. Section 5 considers homogenous random wireless networks under the Protocol Model, where every node employs a common transmission range and wishes to transmit at a common rate. Results on the throughput capacity, the maximum common rate achievable for every source, are presented. It is shown that the common throughput that can be furnished to all the n source–destination pairs is $\Theta(\frac{1}{\sqrt{n \log n}})$. Since the factor $\sqrt{\log n}$ grows very slowly, it shows that random networks so operated are close to best case. An auxiliary consequence is that utilization of a common range for all transmissions is nearly optimal.

Section 6 considers homogenous random wireless networks under the Physical Model, where every node employs a common power for transmission. Similar results, as in Section 5, on the throughput capacity for such networks are presented for the Physical Model.

Section 7 considers random wireless networks with node locations generated by a Poisson point process, and operating under the Protocol Model. Nodes are allowed to use different transmission ranges (a common power can be used too). A constructive scheme shows that with such flexibility a better common rate for each node is achievable elimination the factor $\sqrt{\log n}$, compared to the one achieved in Section 5.

Sections 8 to Section 11 delve into an information theoretic framework, attempting to characterize fundamental limits on the performance of wireless networks, under any causal strategy, without making presuppositions about the manner in which information is sought to be communicated.

Section 8 first specifies a model for signal attenuation with distance. The central result of the section is that, in the high attenuation regime, the scaling behavior of the transport capacity of arbitrary wireless networks is similar to that of networks studied in previous sections, after an appropriate scaling of area. This shows that in this attenuation regime the proposed multi-hop information architecture towards which many current design efforts are targeted is an order-optimal architecture. In this sense information theory provides strategic guidance to designers on the architecture for information transport in wireless networks.

Section 9 addresses the scaling behavior of transport capacity in the low attenuation regime. It presents results showing that the scaling behavior can be very different compared to that in the high attenuation regime. This shows that the architecture for information transport in wireless networks under very low attenuation can indeed need to be quite different from the higher attenuation case. Thus there is a connection between the attenuation property of the medium and the architecture that needs to be adopted. Also different strategies for information transport emerge as of interest.

Section 10 studies the transport capacity for wireless networks in the presence of multi-path fading. The results show that in the high attenuation regime, for many fading cases, the scaling behavior is the same as that in the no-fading environment. So in this attenuation regime there is no difference in the transport capacity achievable, at least up to a preconstant.

Section 11 presents results showing how techniques from Multi-input Multi-output (MIMO) systems can be applied to study the performance of large wireless networks.

Model	Node Locations	Constraint on Power/Range	Power Attenuation	Major Results	Chap-ters	Note
Protocol model (Defined in terms of geometry)	Arbitrary in a region of area A	Arbitrary	No assumption	T(n) = $\Theta(W\sqrt{An})$ bit*meter/sec	2,3	W is the maximum single link rate
	Randomly distributed in a finite domain	Common transmission range r	No assumption	$\lambda(n) = \Theta(\frac{W}{\sqrt{n\log n}})$ bit/sec	5	Every node has a randomly chosen destination node
	Poisson process with density n in a unit domain	Different ranges are allowed	No assumption	$\lambda(n) = \Theta(W/\sqrt{n})$	7	Every node is only one destination node
Physical model (Defined in terms of SINR)	Arbitrary in a region of area A	$P_{\max} \le \Theta((nA)^{\alpha/2})$	$1/d^{\alpha}$ with $\alpha > 2$	T(n)= $\Theta(W\sqrt{An})$	4	A Generalized Physical model is also considered
	Randomly distributed in a unit square	Common transmission power P	$1/d^{\alpha}$ with $\alpha > 2$	$\Theta(\frac{W}{\sqrt{n\log n}}) \le \lambda(n) \le \Theta(W/\sqrt{n})$	6	Every node has a randomly chosen destination node
	Poisson process with density n in a unit domain	No assumption	$1/d^{\alpha}$ with $\alpha > 2$	$\lambda(n) = \Theta(W/\sqrt{n})$	7	Every node is only one destination node
Information theoretical model (Defined based on physical signals)	Arbitrary, but with a minimum separation $r_{\min}>0$	Every node is subject to a power constraint P	$Ge^{-r\gamma}/d^{\delta}$ with $\gamma > 0$ or $\delta > 3/2$	T(n)= $\Theta(\sqrt{n})$	8	
	Arbitrary, but with a minimum separation $r_{\min}>0$		G/d^{δ} with δ small	Large T(n) can be supported with fixed total power; super-linear T(n) is achievable under individual power constraint	9	A coherent relaying with interference subtraction (CRIS) scheme is used
	Arbitrary, but with a minimum separation $r_{\min}>0$	Every node is subject to a power constraint P	Multi-path fading with $\gamma > 0$ or $\delta > 3$	T(n)= $\Theta(\sqrt{n})$	10	
	MIMO techniques are shown to upper bound the transport capacity and study large network behavior				11	

Fig. 1.3 A summary of the main models and results.

Section 12 concludes this text.

For the convenience of readers, we summarize the major models and results in Figure 1.3. Note that we assume that there are n nodes in the network, and $T(n)$ denotes the transport capacity (bit-meters/sec), and $\lambda(n)$ the per-node throughput (bits/second).

2

The Transport Capacity of Arbitrary Wireless Networks – Protocol Model

This section, as well as Sections 3–5, considers arbitrary networks operating under a multi-hop mode of information transfer, where the locations of the nodes, the choice of source–destination pairs, rates along each hop, transmission time slots, and routing, can be jointly optimized.

For wireless networks envisaged to operate in such a multi-hop mode of data transfer, the following question is fundamental: how much information can they transport? More specifically, how will the transport capacity scale with the network size, the number of nodes in the network? This is the question we try to answer from this section to Section 5.

In this section, we first introduce the definition of transport capacity and the network model. Then we present a model for successful transmissions: Protocol Model, which specifies the condition under which a packet can be successfully received.

Given the Protocol Model, assuming that its nodes are located in a region of area $A\,\mathrm{m}^2$, we show that the transport capacity cannot grow faster than $\Theta(\sqrt{An})$ for a network of n nodes[1]. On the other hand,

[1] Given two functions f and g, we say that $f = O(g)$ if $\sup_n |f(n)/g(n)| < \infty$. We say that $f = \Omega(g)$ if $g = O(f)$. If both $f = O(g)$ and $f = \Omega(g)$, then we say that $f = \Theta(g)$.

networks with grid distribution and neighbor-only transmissions are shown to achieve $\Theta(\sqrt{An})$ bit-meters per second. This shows transport capacity scales as $\Theta(\sqrt{An})$ for arbitrary networks under protocol model. To understand the importance of this result suppose that all nodes share equally in this transport capacity. Then each obtains only $\Theta(\frac{\sqrt{A}}{\sqrt{n}})$ bit-meters/second, when the area is held fixed. To interpret this result note that if each node wants to communicate to a distant node at a distance of $\Omega(\sqrt{A})$ meters away, then it can only obtain a rate of $\frac{1}{\sqrt{n}}$. On the other hand, if each node only wishes to communicate with its nearest neighbor that is to a distance of $O(\frac{\sqrt{A}}{\sqrt{n}})$ meters away, then it can do so at a non-vanishing rate.

2.1 Network and the Protocol Model

Consider an *arbitrary network*, where n nodes are arbitrarily located in *a disk of area A* on the plane. The scenario when nodes locations are random will be discussed in Sections 5–7. Let X_i denote the location, as well as the identity, of a node. At each node X_i, there is originating traffic of rate λ_{ij} which is destined for a remote node j. Each node can choose an arbitrary range or power level for any transmission.

We model how networks are intended to work under current technology and protocol development efforts. The common scheme is that packets reach destinations by traveling over a multi-hop path. At each hop, the receiving node decodes an incoming packet, and then re-transmits to the next node, unless the receiving node is the ultimate destination of the packet[2]. Packets can be buffered at intermediate nodes while awaiting transmission. Note that the choice of the sequence of nodes along which a packet is sent from its origin to its final destination is the routing problem.

We assume that each node can transmit at W bits per second over a common wireless channel shared by all nodes. It will be shown that it will not change the ensuring capacity results if the channel is broken up

[2] Whether alternatives such as amplifying-and-forwarding, or advanced techniques such as multi-user decoding, or some other choices, would make a significant difference, is the topic of Sections 5–9.

into several sub-channels of capacity $W_1, W_2, \cdots, W_M$ such that $W = W_1 + \cdots + W_M$.

Each node can choose an arbitrary range or power level for each transmission. In order for a transmission to be successful, i.e., for a receiver to correctly decode a packet intended for it, certain requirements need to be satisfied. Here we describe two versions of a Protocol Model, where the conditions for successful transmissions are postulated geometrically (A different model based on more physical considerations will be treated in Sections 4, 6 and 7).

2.1.1 The Protocol Model

Suppose node X_i transmits over the m-th sub-channel to a node X_j. Then this transmission at rate W_m bits/sec is assumed to be successfully received by node X_j if

$$|X_k - X_j| \geq (1 + \Delta)|X_i - X_j|, \tag{2.1}$$

for every other node X_k simultaneously transmitting over the same sub-channel.

The quantity $\Delta > 0$, or more properly a circle of radius $(1 + \Delta)|X_i - X_j|$ quantifies a guard zone required around the receiver to ensure that there is no destructive interference from neighboring nodes transmitting on the same m-th sub-channel at the same time.

An alternative is to specify an interference footprint around a transmitter, which gives rise to a slightly different model.

2.1.2 A variant of the Protocol Model

Suppose node X_i transmits over the m-th sub-channel to a node X_j at rate W_m bits/sec. Then this transmission is postulated to be successfully received by node X_j if

$$|X_k - X_j| \geq (1 + \Delta)|X_k - X_l|, \tag{2.2}$$

for every other node X_k simultaneously transmitting over the same m-th sub-channel, with X_l denoting intended the recipient of node X_k's transmission.

Here the quantity $\Delta > 0$ models situations where an interference footprint is specified by the protocol around the transmitter. The circle of radius $(1 + \Delta)|X_k - X_l|$ models the interference footprint created by a transmission of range $|X_k - X_l|$ originating at node k, within which region no other concurrent reception is possible. We will see that both variants of the Protocol Model lead to the same capacity results.

We note that the choice of the range of a packet is achieved by power control. The choice of the time transmission is made is achieved by the medium access control protocol.

2.2 Definition of transport capacity

Suppose that based on a certain overall scheme, ultimately b bits are successfully transmitted from a node i to an intended node j that is at a distance $d_{ij} = |X_i - X_j|$. Then we say that the network has pumped bd_{ij} bit-meters. Notice only the distance between the *original* source and the *final* destination counts; extra distance travelled due to, say, non-straight line routing is not counted.

Definition 2.1. *The transport capacity of a specific network* is defined as the maximum bit·meters per second the network can achieve in aggregate. *The transport capacity of n nodes* is the maximum of all achievable transport capacities networks with n nodes in a disk of area A – the difference is that in this latter case the locations of the n nodes are also allowed to be optimized, as are the choices of source–destination pairs.

Thus, if a network is able to support a rate of λ_{ij} bits per second from each node i to each node j, then the transport capacity of the network is the supreme of $\sum_{i \neq j} \lambda_{ij} |X_i - X_j|$ over all such supportable rate vector $\{\lambda_{ij} : 1 \leq i, j \leq n\}$.

2.3 Main results

The following result is the main result of this section for arbitrary networks.

Theorem 2.2. The transport capacity of an Arbitrary Network of n nodes under the Protocol Model is $\Theta(W\sqrt{An})$ bit-meters/sec.

This is achievable when the locations of the nodes and the source–destination pairs are chosen optimally, and the network is optimally operated. By the phrase "optimally operated" we mean optimized over the choice of a route, or a multiple set of routes to be used for each source–destination pair, as well as optimal timing of all transmissions, i.e., spatio-temporal optimization of all transmissions.

Note that $g(n) = \Theta(f(n))$ if for some constants $c_1 > 0$ and $c_2 < +\infty$, $c_1 f(n) \leq g(n) \leq c_2 f(n)$. Specifically, an upper bound is $\sqrt{\frac{8}{\pi}}\frac{W}{\Delta}\sqrt{An}$ bit-meters/sec for every Arbitrary Network for all spatial and temporal scheduling strategies, while $\frac{W\sqrt{A}}{1+2\Delta}\frac{n}{\sqrt{n}+\sqrt{8\pi}}$ bit-meters/sec (for n a multiple of four) is an achievable lower bound, when the node locations and the transmissions are chosen appropriately.

One implication of this result is that each node, on average, obtains $\Theta\left(\frac{W\sqrt{A}}{\sqrt{n}}\right)$ bit-meters/sec. Since this quantity diminishes to zero as n goes large, we see that there is a law of diminishing returns in our model where the area of the domain is fixed, while the number of nodes is allowed to grow.

2.4 Main ideas behind the proof

The essential idea to upper-bound the transport capacity in the proof below is to observe that successful transmissions "consume" area as they happen. Moreover the radius of such a consumed area is proportional to the transmission range. Since the sum of such areas is upper-bounded by the limited total area A, it follows from invoking the convexity of quadratic function, that one can transform this upper-bound into an upper-bound for the bit-meters per second – the transport capacity.

To show the achievability of $\Theta(\sqrt{An})$ bit-meters/second, one can first arrange the n nodes in grid-like positions, then choose $n/2$ nodes as senders with each of them transmitting only to one of its nearest neighbors.

2.5 Upper-bounding the transport capacity

In this subsection, we establish the upper-bound for the transport capacity for networks under the Protocol Model (2.1)[3]. To recall, the following are the basic assumptions on the model being considered:

(A.i) There are n nodes arbitrarily located in a disk of area A on the plane.

(A.ii) Each node can transmit W_m bits/sec over sub-channel m, $1 \leq m \leq M$, and $\sum_{m=1}^{M} W_m = W$.

Let us consider sub-channel m, and a time instant $t \in [0,T]$ within a time interval of duration T.

Suppose node X_i is transmitting successfully to node X_j over this sub-channel at time t, and suppose that node X_k is also transmitting successfully to some node X_l at the same time. Let $\mathcal{T}_m(t)$ denote the set of all such transmissions ongoing at time t over sub-channel m.

From the triangle inequality and (2.1), we have

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

Similarly,

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

Adding the two inequalities, we obtain

$$|X_\ell - X_j| \geq \frac{\Delta}{2}\left(|X_k - X_\ell| + |X_i - X_j|\right). \tag{2.3}$$

This can be interpreted as saying that two disks, one of radius $\frac{\Delta}{2}|X_i - X_j|$ centered at X_j, and the other of radius $\frac{\Delta}{2}|X_k - X_l|$ centered at X_l, are essentially[4] disjoint; as shown in Figure 2.1.

Hence, for all transmissions occurring at time t, disks of radius $\frac{\Delta}{2}$ times the transmission range, centered at the receivers over the same

[3] For networks under Protocol Model (2.2), it can be shown similarly.

[4] We use the word "essentially" since the two disks have been defined as closed sets and may intersect at their boundaries.

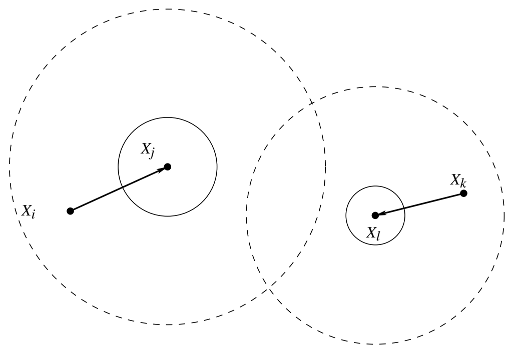

Fig. 2.1 The exclusion disk around each receiver node.

sub-channel m, are essentially disjoint. We call these disks "exclusion disks."

Noting that a transmission range is less than the diameter of the circular domain, it can be shown that at least a quarter of each exclusion disk is within the domain. (This happens when the range of the transmission is exactly equal to the diameter of the domain). Hence we get

$$\sum_{(i,j)\in\mathcal{T}_m(t)} \frac{1}{4}\pi(\frac{\Delta}{2}d_{i,j})^2 = \sum_{(i,j)\in\mathcal{T}_m(t)} d_{i,j}^2 \cdot \frac{\pi\Delta^2}{16} \leq A, \tag{2.4}$$

where we recall that $\mathcal{T}_m(t)$ denotes all the effective transmissions at time t over sub-channel m, and d_{ij} is the distance between node i and j.

Since at most half the nodes can be transmitting at any time while the other half nodes are receiving, there are at most $n/2$ concurrent transmissions at time t, i.e., $|\mathcal{T}_m(t)| \leq n/2$. By the Cauchy-Schwarz inequality we have

$$\begin{aligned}\sum_{(i,j)\in\mathcal{T}_m(t)} d_{i,j} &\leq \sqrt{\sum_{(i,j)\in\mathcal{T}_m(t)} d_{i,j}^2 \cdot \sum_{(i,j)\in\mathcal{T}_m(t)} 1^2} \\ &\leq \sqrt{\sum_{(i,j)\in\mathcal{T}_m(t)} d_{i,j}^2 \frac{n}{2}} \leq \sqrt{\frac{8An}{\pi\Delta^2}}.\end{aligned}$$

So the instantaneous rate in bit-meters/second at time t, over sub-channel m, is upper-bounded by

$$W_m \sum_{(i,j)\in\mathcal{T}_m(t)} d_{i,j} \leq \sqrt{\frac{8A}{\pi}}\frac{W_m}{\Delta}\sqrt{n}.$$

This is true at any time t and for every sub-channel m. Hence summing over all the sub-channels, by Assumption (A.ii), we deduce that the transport capacity is upper-bounded by $\sqrt{\frac{8A}{\pi}}\frac{W}{\Delta}\sqrt{n}$ bit-meters per second. □

2.6 A constructive lower bound

We now show that for n a multiple of 4, $\frac{W}{1+2\Delta}\frac{n\sqrt{A}}{\sqrt{n}+\sqrt{8\pi}}$ bit-meters/sec is indeed achievable under the Protocol Model, for an appropriate arrangement of n nodes in a disk of area A.

Define $r := \frac{1}{1+2\Delta}\frac{\sqrt{A}}{\sqrt{\frac{n}{4}}+\sqrt{2\pi}}$. Center the domain of radius $\frac{\sqrt{A}}{\sqrt{\pi}}$ at the origin. Then 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, as shown in Figure 2.2. 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 chooses its nearest neighboring node as its receiver, which is at a distance r away. A simple calculation shows that all the transmitters can transmit concurrently, there is no interference from any other

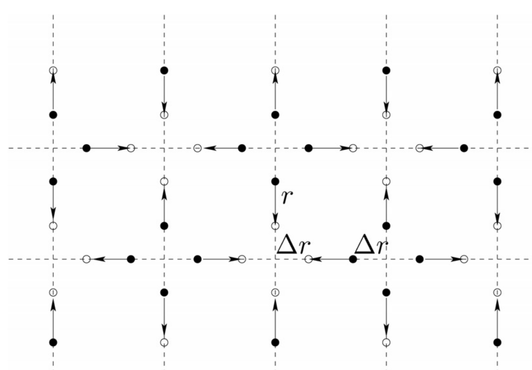

Fig. 2.2 The nodes arrangement for achieving the lower bound.

transmitter-receiver pair. It is easy to verify that one can put $n/2$ transmitter – receiver pairs 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 $\pi(R - \sqrt{2}s)^2/s^2$. Now take $s = (1 + 2\Delta)r$ and $R = \frac{\sqrt{A}}{\sqrt{\pi}}$). For this n-node network, there are $n/2$ concurrent transmissions feasible with the same range r and rate W; thus $\frac{W}{1+2\Delta}\frac{n\sqrt{A}}{\sqrt{n}+\sqrt{8\pi}}$ bit-meters/sec is achieved. Notice also that in achieving this, there is no need to divide the original into m separate sub-channels.

A more careful design will be presented in Section 3 which leads to an even higher achievable transport capacity.

2.7 Notes

This section is mainly based on Sections II–III of [12]; though a simplified proof for the upper bound for the transport capacity is presented here.

3

Sharpening the Bounds on Transport Capacity of Arbitrary Networks – Protocol Model

In this section, more carefully developing notions of *exclusion region* and *interference region* [1] associated with each active sender-receiver pair, we present improved upper and lower bounds to those presented in Section 2. The new bounds bracket the transport capacity to within a factor of $\sqrt{8}$.

For simplicity, throughout the section, we assume that the network of n nodes is located arbitrarily in a unit area disk. It can be seen that under the Protocol Model shrinking the area from A square meters to 1 simply involves rescaling all distances by a factor of $\frac{1}{\sqrt{A}}$ without effecting what transmissions are concurrently feasible. Thus the only difference is that all transport capacity results are to multiplied by the factor $\sqrt{A}$.

For clarity, we also assume that there is only one channel with rate W bits/second. It is easy to show, as in Section 2, that the results also hold for the multi-channel case when the channel is divided into separate sub-channels with sum bandwidth $W_1 + W_2 + \cdots + W_m = W$.

The proofs will be tailored to networks under the second variant of the Protocol Model. Similar proofs can be used for networks under the first variant as well.

3.1 Protocol Model

Let us recall the Protocol Model variants for postulating a successful transmission from node X_i to node $X_{R(i)}$:
Protocol Model:

$$|X_k - X_{R(i)}| \geq (1 + \Delta)|X_i - X_{R(i)}|, \quad \forall k \in \{\text{Active transmitters}\}, k \neq i. \tag{3.1}$$

The above model specifies a guard zone around the *receiver*. In this section, we provide the proofs for the following variant, which specifies an interference footprint around the *transmitter*.
Protocol Model Variant:

$$|X_k - X_{R(i)}| \geq (1 + \Delta)|X_k - X_{R(k)}|, \quad \forall k \in \{\text{Active transmitters}\}, k \neq i. \tag{3.2}$$

Remark 3.1. As will be easily seen from the proofs, the bounds on the transport capacity of wireless networks hold for the other model too.

3.2 Exclusion region and interference region

Recall that the basic idea to upper bound the transport capacity in Section 2 is to show that the disks, centered at each active receiver with radius $\frac{\Delta}{2}$ times the transmission range, are essentially disjoint. This is an example of the notion of *exclusion region*. In this section we further develop this concept of exclusion region based on [1]. We also introduce a concept of generalized *interference region*.

Definition 3.2. *Exclusion region*: For a particular configuration of transmitters and receivers in a network, an exclusion region of an active transmitter-receiver pair is an associated area such that, for the transmission to be successful, it must be kept disjoint from every other exclusion region at that time and over the same sub-channel.

The following result shows that there is a capsule-shaped exclusion region around each transmitter-receiver pair for the Protocol Model, as shown in Figure 3.1.

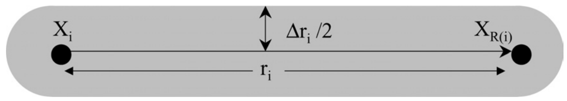

Fig. 3.1 The $\Delta r/2$ neighborhood of a transmitter-receiver pair.

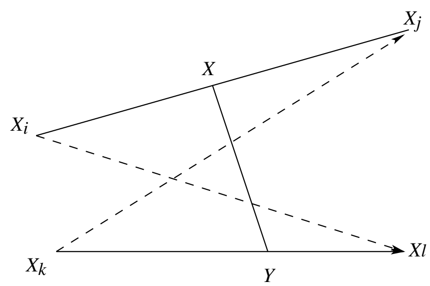

Fig. 3.2 Two transmitter-receiver pairs.

Theorem 3.3. In a wireless network under the protocol model, if (X_i, X_j) is an active transmitter-receiver pair, then the $\Delta d_{ij}/2$ radius neighborhood of the line joining them is an exclusion region for (i, j).

Proof. Let (X_i, X_j) and (X_k, X_l) be any two concurrently active transmitter-receiver pairs. Let X and Y be two points on the line segments X_iX_j and X_kX_l, respectively; see Figure 3.2. By the triangle inequality, we have

$$|X_i - X| + |X - Y| + |Y - X_l| \geq |X_i - X_l|, \tag{3.3}$$

$$|X_k - Y| + |Y - X| + |X - X_j| \geq |X_k - X_j|. \tag{3.4}$$

Adding them, we get

$$|X_i - X| + |X - X_j| + |X_k - Y| + |Y - X_l| + 2|X - Y| \\ \geq |X_i - X_l| + |X_k - X_j|.$$

Since $|X_i - X| + |X - X_j| = |X_i - X_j|$, the above becomes

$$|X_i - X_j| + |X_k - X_l| + 2|X - Y| \geq |X_i - X_l| + |X_k - X_j|.$$

According to the variant of the Protocol Model (3.2), we get

$$|X_i - X_j| + |X_k - X_l| + 2|X - Y| \geq (1 + \Delta)(|X_i - X_j| + |X_k - X_l|),$$

which simplifies to

$$|X - Y| \geq \frac{\Delta}{2}(d_{ij} + d_{kl}).$$

This suffices to show that the $\frac{\Delta}{2}d_{ij}$ neighborhood of X_iX_j, and the $\frac{\Delta}{2}d_{kl}$ neighborhood of X_kX_l, are disjoint. Otherwise there exist point Z, point X on X_iX_j, and point Y on X_kX_l, such that $|Z - X| < \frac{\Delta}{2}d_{ij}$ and $|Z - Y| < \frac{\Delta}{2}d_{kl}$. Then we would get

$$|X - Y| < |Z - X| + |Z - Y| < \frac{\Delta}{2}(d_{ij} + d_{kl}),$$

which is a contradiction. □

Now we introduce the definitions of *interference region* and the *generalized protocol transmission model*.

Definition 3.4. Generalized Protocol Model and the Interference Region: Suppose that associated with each transmitter-receiver pair $(X_k, X_{R(k)})$ there is an area I_k, such that for a transmission from X_i to $X_{R(i)}$ to be successful, it is necessary that

$$X_{R(i)} \notin I_k, \quad \forall k \in \{\text{Active transmitters}\}, \; k \neq i. \tag{3.5}$$

Such an area I_k will be called an *Interference Region*.

Remark 3.5. Interference Regions corresponding to the variant of the Protocol Model (3.2) are $I_k := \{X : |X - X_k| < (1 + \Delta)|X_k - X_{R(k)}|\}$, which is a disk of radius $(1 + \Delta)|X_k - X_{R(k)}|$ centered at X_k. Also, it is easy to see that a simple exclusion region is a disk centered at each X_k with radius $\eta/2$, where $\eta = \inf_{X \notin I_k} |X_{R(k)} - X|$.

Now, we present an Exclusion Region for the Generalized Protocol Model.

Theorem 3.6. The set $E_k := \{X : |X - X_{R(k)}| < |X - Y|, \forall Y \notin I_k\}$, is a valid exclusion region for each k, and it is convex.

Proof. First we show that E_k is a valid exclusion region. If not, then there exist distinct i and j (denoting nodes by their indices for brevity), both concurrently active, such that $E_i \cap E_j \neq \phi$. Assume $X \in E_i \cap E_j \neq \phi$. Then

$$\begin{aligned} |X - X_{R(i)}| &< |X - Y|, \quad \forall Y \notin I_i; \\ |X - X_{R(j)}| &< |X - Y|, \quad \forall Y \notin I_j. \end{aligned} \tag{3.6}$$

By (3.5), we know $X_{R(i)} \notin I_j$ and $X_{R(j)} \notin I_i$. So setting $Y = X_{R(j)}$ in the first inequality of (3.6), and $Y = X_{R(i)}$ in the second, we get

$$\begin{aligned} |X - X_{R(i)}| &< |X - X_{R(j)}|, \\ |X - X_{R(j)}| &< |X - X_{R(i)}|. \end{aligned}$$

This is a contradiction, and hence E_k is a valid exclusion region for every active k.

Now we show that E_k is convex. If not, let $X, Y \in E_k$ and $Z \notin E_k$, with Z lying on the line segment XY. By the definition of E_k we know

$$\begin{aligned} |X - X_{R(k)}| &< |X - Q|, \quad \forall Q \notin I_k \\ |Y - X_{R(k)}| &< |Y - Q|, \quad \forall Q \notin I_k, \end{aligned} \tag{3.7}$$

while for some $Q' \notin I_k$,

$$|Z - X_{R(k)}| \geq |Z - Q'|. \tag{3.8}$$

Substituting $Q = Q'$ into (3.7), we get

$$\begin{aligned} |X - X_{R(k)}| &< |X - Q'|, \\ |Y - X_{R(k)}| &< |Y - Q'|. \end{aligned} \tag{3.9}$$

Now define $\Gamma := \{X : |X - X_{R(k)}| < |X - Q'|\}$, which is a convex set since it is an open half-plane. Inequalities (3.9) imply that $X, Y \in \Gamma$, which implies further that $Z \in \Gamma$, by the convexity of Γ. This is a contradiction to (3.8). □

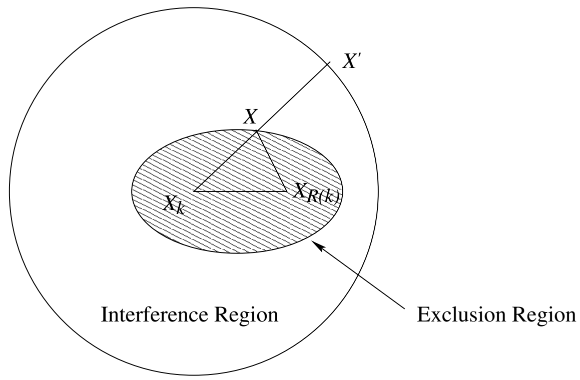

Fig. 3.3 The exclusion region when I_k is a disk.

3.2.1 An exclusion region for the variant of the Protocol Model (3.2)

Now Theorem 3.6 can be applied to the variant (3.2) of the Protocol Model; it shows that E_k is an ellipse; see Figure 3.3.

Theorem 3.7. Suppose I_k is a disk of radius $(1+\Delta)|X_k - X_{R(k)}|$ centered at X_k. Then E_k is the region inside the ellipse with X_k and $X_{R(k)}$ as the foci, and eccentricity $\frac{1}{1+\Delta}$, i.e.,

$$E_k = \{X : |X - X_{R(k)}| + |X - X_k| \leq (1+\Delta)|X_{R(k)} - X_k|\}. \quad (3.10)$$

Proof. As in Figure 3.3, for a point X in the exclusion region, extend X_kX towards X, and assume X' is the point crossing the boundary of I_k. Since I_k is the open interior of a disk, it is easy to verify that

$$\inf_{Y \notin I_k} |X - Y| = |X - X'|.$$

Hence we have

$$\begin{aligned}
&|X - X_{R(k)}| < |X - Y|, \quad \forall Y \notin I_k \\
&\quad \Leftrightarrow |X - X_{R(k)}| \leq |X - X'| \\
&\quad \Leftrightarrow |X - X_{R(k)}| + |X - X_k| \leq |X - X'| + |X - X_k| = |X' - X_k| \\
&\qquad = (1+\Delta)|X_{R(k)} - X_k| \\
&\quad \Leftrightarrow |X - X_{R(k)}| + |X - X_k| \leq (1+\Delta)|X_{R(k)} - X_k|.
\end{aligned}$$

The last expression represents an ellipse with $X_k, X_{R(k)}$ as the foci, and eccentricity $\frac{1}{1+\Delta}$. □

3.3 Improved bounds for transport capacity

In this subsection we apply the results on the exclusion regions to give improved upper and lower bounds for the transport capacity of arbitrary wireless networks under the Protocol Model.

Lemma 3.8. With at least two active transmitter-receiver pairs, at least a quarter of the area of the exclusion region, identified in Theorem 3.7, must lie inside the unit disk.

Proof. Two active transmitter-receiver pairs implies the existence of a pair $(X_k, X_{R(k)})$ such that at least one point on the boundary of its exclusion region – the ellipse with $(X_k, X_{R(k)})$ as focus – must be inside the unit disk. This, by geometric argument, guarantees that the unit disk contains at least a quarter of the area of the ellipse. The detailed proof can be found in Lemma 3.2 of [1]. □

Combining the results developed in this section, we obtain the following improved upper bound.

Theorem 3.9. Under the variant (3.2) of the Protocol Model, the transport capacity of arbitrary wireless networks is upper bounded by

$$\sqrt{\frac{8}{\pi}} \frac{1}{\sqrt{(1+\Delta)\sqrt{\Delta}\sqrt{2+\Delta}}} W \cdot \sqrt{n}. \tag{3.11}$$

Proof. Denote by $\mathcal{T}(t)$ the set of active transmitter-receiver pairs at time t[1]. By Theorem 3.7, each transmitter-receiver pair $(k, R(k))$ has an exclusion region which is an ellipse with major axis $2a = r_k(1+\Delta)$

[1] As noted at the beginning of this section, the result holds even if the channel consists of m sub-channels with $W_1 + \cdots + W_m = W$.

and minor axis $2b = r_k\sqrt{\Delta^2 + 2\Delta}$. The area of the ellipse is πab. Since at least one-fourth of the ellipse is within the unit disk domain, and the exclusion regions are disjoint, we have

$$\begin{aligned}
&\frac{1}{4}\sum_{(k,R(k))\in\mathcal{T}(t)} \frac{1}{4}\pi r_k(1+\Delta)\cdot r_k\sqrt{\Delta^2+2\Delta} \\
&= \sum_{(k,R(k))\in\mathcal{T}(t)} \frac{\pi}{16}(1+\Delta)\sqrt{\Delta^2+2\Delta}\cdot r_k^2 \\
&\leq 1.
\end{aligned}$$

Now, we can use this refined bound to replace the bound in (2.4). The rest of the proof remains the same. The instantaneous bit-meters rate at time t is thus upper-bounded by

$$W\sum_{(i,j)\in\mathcal{T}(t)} d_{i,j} \leq \sqrt{\frac{8}{\pi}}\frac{W}{\sqrt{(1+\Delta)\Delta(2+\Delta)}}\sqrt{n}.$$

This is true for every t and hence it is an upper bound on the transport capacity. □

One can also construct an improved lower bound for the transport capacity based on the results on exclusion regions.

Theorem 3.10. There exists an arrangement of node locations and traffic patterns such that under the variant of the Protocol Model (3.2) with parameter Δ, the network can achieve

$$\sqrt{\frac{1}{\pi}}\frac{W}{\sqrt{(1+\Delta)\sqrt{\Delta}\sqrt{2+\Delta}}}\sqrt{n}\ \textit{bit-meters/sec.} \tag{3.12}$$

Proof. Suppose square S, with edge length $L = \sqrt{2/\pi}$, is a square inscribed in the unit disk. We will only consider putting nodes inside S.

Let $r =: \sqrt{\frac{2L^2}{n(1+\Delta)\sqrt{\Delta}\sqrt{2+\Delta}}}$, $a =: r(1+\Delta)$, and $b =: r\sqrt{\Delta}\sqrt{2+\Delta}$.

Now tessellate S with an axis-parallel grid of rectangles of size $a \times b$. There are $L/a =$ columns and L/b rows, and more than $n/2$ rectangles

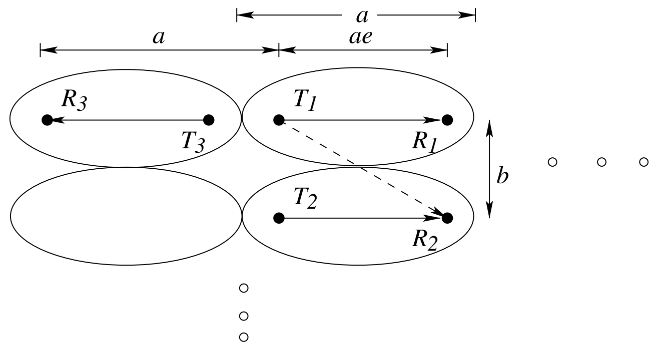

Fig. 3.4 The arrangement of transmitter-receiver pairs.

because

$$L/a \times L/b = \frac{L^2}{r^2(1+\Delta)\sqrt{\Delta}\sqrt{2+\Delta}} = \frac{n}{2}.$$

Now insert into each rectangle an ellipse with major axis a and minor axis b, and a pair of nodes on the foci, with the left one being a transmitter and the right one being its receiver; see Figure 3.4. Notice the eccentricity of the ellipse is $e = 1/(1+\Delta)$ and the distance between each pair is $ae = r$.

Now we show that every pair can successfully transmit and receive under the variant of the Protocol Model (3.2). This is readily seen by noticing that

$$|T_1 - R_2|^2 = (ae)^2 + b^2 = a^2e^2 + a^2(1-e^2) = a^2 = r^2(1+\Delta)^2, \text{ and}$$
$$|T_1 - R_3| = a = r(1+\Delta).$$

Since there are $n/2$ pairs of transmitter-receiver pairs, the bit-meters/sec transport specified in (3.12) is achieved. □

3.4 Notes

This section is based on [1]. The concept of Exclusion Regions can be generalized to subsets of transmitter-receiver pairs. A set of exclusion regions for networks using directional antennas can be found in Section 4 of [1].

4

The Transport Capacity of Arbitrary Wireless Networks – Physical Model

In the previous Sections 2–3, the criterion for a successful transmission is specified by geometric constraints: a guard zone around a receiver or an interference footprint around a transmitter. In this section, we consider a more physical criterion for successful reception specified by the requirement on the signal to interference plus noise ratio (SINR) – called the Physical Model. Since common radio technology for decoding a packet works only if the SINR is sufficiently large, this model is more faithful to physical considerations. In this section, the scaling behavior of the transport capacity for arbitrary networks is characterized. We show that there exists a correspondence between networks under the Protocol Model and networks under the Physical Model. Thus by simply translating the result for the Protocol Model, $\Theta(\sqrt{n})$ bit-meters per second is achieved for networks under the Physical Model. Next, we show that the upper bound is also of order $\Theta(\sqrt{n})$, by carefully examining interference generated by transmitters and bounding it. Thus we see that the results for the simple geometric interference model are robust and continue for more physically realistic models.

We consider networks with nodes lying in a *square*[1] of area A. Networks distributed in different domains can be considered similarly.

4.1 Physical Model and the Generalized Physical Model

Let $\{X_k; k \in \mathcal{T}(t)\}$ be the subset of nodes simultaneously transmitting at some time instant t over a certain sub-channel m. Let P_k be the power level chosen by node X_k, for $k \in \mathcal{T}(t)$. Then the transmission between node X_i, $i \in \mathcal{T}(t)$, and $X_{R(i)}$ is successful if

$$\frac{\frac{P_i}{|X_i - X_j|^\alpha}}{N + \sum_{\substack{k \in \mathcal{T} \\ k \neq i}} \frac{P_k}{|X_k - X_j|^\alpha}} \geq \beta, \tag{4.1}$$

where N is the ambient noise power level. Once this is satisfied, we assume that the data rate over the link is W bits/second. This models the situation where a minimum signal to interference plus noise ratio (SINR) of β is required for successful receptions. The signal power is assumed to decay with distance r as $1/r^\alpha$. The exponent α is call the "path loss exponent" for the power. In this section we assume $\alpha > 2$.

A. Generalized Physical Model

The above model assumes that a transmission can only occur at one of two rates: W bits/sec if the SINR exceeds β, and 0 bits/sec otherwise. This model of data rate can be generalized to be continuous in SINR, based on Shannon's capacity formula for the additive Gaussian noise channel. (This will however require that the modulation or coding scheme for the transmission to be adapted to the existent SINR, a fact that entails further overhead for coordinating between the nodes.) In this case the data rate from transmitter X_k to its receiver $X_{R(k)}$ is assumed to be

$$W_k = H_m \log \left(1 + \frac{\frac{P_k}{|X_k - X_{R(k)}|^\alpha}}{N H_m + \sum_{i \in \mathcal{T}, i \neq k} \frac{P_i}{|X_i - X_{R(k)}|^\alpha}} \right) \tag{4.2}$$

[1] The reason for considering a square, instead of a disk, is only for the sake of clarity in proofs. The results hold true for a disk domain and in fact any domain which is the closure of its interior.

where H_m is the bandwidth of channel m in hertz, such that the total bandwidth is finite: $\sum_m H_m \leq H_0$; and $N/2$ is the noise spectral density in watts/hertz.

4.2 Correspondence between Protocol Model and Physical Model, and a lower bound

The Protocol Model stipulates local, geometric constraints, while in the Physical Model every transmission influences every reception. They thus seem very different from each other, but, interestingly, as shown by the following result, there is a correspondence between them.

Theorem 4.1. Let $\Delta(\beta) := (48\frac{2^{\alpha-2}}{\alpha-2}\beta)^{1/\alpha}$. Suppose that for $\Delta > \Delta(\beta)$ the Protocol Model allows simultaneous transmissions for all transmitter-receiver pairs in a set $\mathcal{T}(t)$. Then there exists a power assignment $\{P_i, 1 \leq i \leq n\}$ allowing the same set of transmissions under the Physical Model with threshold β.

Proof. By Theorem 3.7, we know that an exclusion region for an active transmitter-receiver pair $(X_k, X_{R(k)})$ is an ellipse with $X_k, X_{R(k)}$ being the foci:

$$E_k = \{P : |P - X_{R(k)}| + |P - X_k| = (1 + \Delta)|X_{R(k)} - X_k|\}.$$

Denote by $D_k, D_{R(k)}$, disks of radius $\Delta r_k/2$ around X_k and $X_{R(k)}$ respectively. Then all such disks for active transmitters and receivers are disjoint.

Now we show that a power assignment $P_k = c\Delta^2 r_k^2$ with $c \geq \frac{N}{c_\alpha}(\Delta\sqrt{2A})^{\alpha-2}$ suffices, where $c_\alpha := \frac{24}{\alpha-2}2^{\alpha-2}$, and N is the ambient noise power.

First, we consider the interference I_k at a receiver $X_{R(k)}$.

Let disk D'_j be the disk centered at $X_{R(j)}$ with radius $|X_i - X_{R(j)}|$, and $D_{ij} := \{x : x \in D_i \cap D'_j\}$; see Figure 4.1. Denote by B, C the two points where D_i intersects D'_j. Since the radius of D_i is less than that of D'_j, we know $\angle CX_iX_{R(j)} = \angle BX_iX_{R(j)} \geq \pi/3$. So the area of D_{ij} is

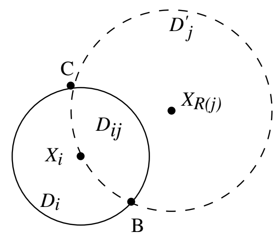

Fig. 4.1 Calculating the interference.

at least one third of D_i. Letting $d\mathcal{A}$ be the area element, we now have

$$\begin{aligned} I_k &:= \sum_{i\in\mathcal{T}, i\neq k} \frac{P_i}{|X_i - X_{R(k)}|^\alpha} \\ &= \sum_{i\in\mathcal{T}, i\neq k} \frac{4c}{\pi} \int_{D_i} \frac{d\mathcal{A}}{|X_i - X_{R(k)}|^\alpha} \\ &\leq \sum_{i\in\mathcal{T}, i\neq k} \frac{12c}{\pi} \int_{D_{ik}} \frac{d\mathcal{A}}{|X_i - X_{R(k)}|^\alpha} \\ &\leq \sum_{i\in\mathcal{T}, i\neq k} \frac{12c}{\pi} \int_{D_{ik}} \frac{d\mathcal{A}}{|x - X_{R(k)}|^\alpha} \\ &= \frac{12c}{\pi} \int_{\cup_{i\in\mathcal{T}, i\neq k} D_{ik}} \frac{d\mathcal{A}}{|x - X_{R(k)}|^\alpha}. \end{aligned}$$

Since $(\cup_{i\in\mathcal{T}, i\neq k} D_{ik}) \cap D_k = \phi$, we can bound the right hand side of the last expression such that

$$\begin{aligned} I_k &\leq \frac{12c}{\pi} \int_{|x - X_{R(k)}| \geq \Delta r_k/2} \frac{d\mathcal{A}}{|x - X_{R(k)}|^\alpha} \\ &= \frac{12c}{\pi} \int_{\Delta r_k/2}^{\infty} \frac{2\pi r dr}{r^\alpha} \\ &= \frac{24c}{\alpha - 2} (\frac{2}{\Delta r_k})^{\alpha - 2} = c c_\alpha (\Delta r_k)^{2-\alpha}. \end{aligned}$$

Noting that $r_k \leq \sqrt{2A}$, the SINR received at $X_{R(k)}$ is

$$\begin{aligned}\frac{P_k/r_k^\alpha}{N+I_k} &\geq \frac{c\Delta^2 r_k^{2-\alpha}}{N+cc_\alpha(\Delta r_k)^{2-\alpha}} \\ &= \frac{\Delta^\alpha}{\frac{N}{c}\Delta^{\alpha-2}r_k^{\alpha-2}+c_\alpha} \\ &\geq \frac{\Delta^\alpha}{2c_\alpha} \geq \beta.\end{aligned}$$ □

This shows that any feasible set of active transmitter receiver pairs in the Protocol Model with sufficiently large Δ, dependent on β, admits a power assignment for the same set of nodes to transmit under the Physical Model. This allows one to get feasibility results for the Physical Model based on those for the Protocol Model.

In particular, consider the network involving a grid of ellipses, as in Section 3, Figure 3.4. Then based on Theorem 3.10, noting that L corresponds to $\sqrt{A}$ now given the nodes' layout, we get the following achievability result for the Physical Model.

Theorem 4.2. There exists an arrangement of node locations and traffic patterns so that a network of n nodes under the Physical Model with SINR threshold β can achieve

$$\sqrt{\frac{A}{2}}\frac{W\sqrt{n}}{\sqrt{(1+\Delta)\sqrt{\Delta}\sqrt{2+\Delta}}} \text{ bit-meters/sec,}$$

where $\Delta := (2c_\alpha\beta)^{1/\alpha}$.

Remark 4.3. Note that the above theorem also establishes a similar lower bound, differing by only a constant ratio, for networks under the Generalized Physical Model. This is because a successful transmission in a network under the Physical Model requires that the SINR be larger than a threshold, which enables a constant rate between these two nodes under the Generalized Physical Model.

4.3 Upper bound

In this subsection, we present an upper bound on the transport capacity for arbitrary networks under the Generalized Physical Model (4.2). We note that this also establishes an upper bound, differing by only a constant ratio, for networks under the Physical Model. This is because if a network under the Physical Model can realize a successful transmission between X_i and $X_{R(i)}$ at rate W bits/sec, the SINR must be no less than β. Then in that case, the Generalized Physical Model also allows a constant rate transmission corresponding to the SINR β in Shannon's formula between nodes X_i and $X_{R(i)}$.

Theorem 4.4. Consider the Generalized Physical Model (4.2) with $N > 0$, $\alpha > 2$, and available total bandwidth $H_0 = \sum_{m=1}^{M} H_m$, where H_m is the bandwidth used in the m-th channel. If the maximum power that a node can employ on the m-th sub-channel is bounded by $P_{max} = H_m N n^{\alpha/2}$, then the transport capacity of an n node network located in a *unit* square and under the Generalized Physical Model is upper bounded by $H_0 \mu_\alpha \sqrt{n}$ bit-meters per second, where $\mu_\alpha := \log_2(e)(\alpha(13\sqrt{2}+6) + 4((\sqrt{2}+1)2)^\alpha)$.

To prove this result, one needs to carefully examine the interference experienced by each receiver. The following lemma is needed in the interference analysis.

Lemma 4.5. Consider an $m \times m$ square Q, tessellated by m^2 unit squares, where m is a power of 2. Suppose there are k ($\leq m^2$) points $\{Y_i, 1 \leq i \leq k\}$ in Q, and each unit square contains at most one point. Let there be a total ordering "$\prec$" imposed on the k points. For $1 \leq i \leq k$, define Y_i's "in-order distance" with respect to Q, as $d_i := min(\{m\sqrt{2}\} \cup \{|Y_j - Y_i| : 1 \leq j \leq k, j \neq i, Y_i \prec Y_j\})$, i.e., the distance to the nearest higher ordered point. Then $\sum_{i=1}^{k} d_i \leq 3\sqrt{2}m^2 - 2\sqrt{2}m$.

Proof. Let $f(m)$ denote the upper bound on $\sum_{i=1}^{k} d_i$ over any choice of the point locations and their ordering. We will obtain a recurrence for $f(m)$.

Note first that $f(1) = \sqrt{2}$. For $m \geq 4$, divide the grid into four equal quadrants of size $m/2 \times m/2$. For any point Y_i, say in quadrant Q', let $d'_i = min(\{\frac{m}{2}\sqrt{2}\} \cup \{|Y_j - Y_i| : 1 \leq j \leq k, j \neq i, Y_j \in Q', Y_i \prec Y_j\})$ be the in-order distance of Y_i with respect to Q'.

Next note that for any two points Y_i and Y_j in the same quadrant, $|Y_i - Y_j| \leq \frac{m}{2}\sqrt{2}$. So either d_i or d_j is no more than $\frac{m}{2}\sqrt{2}$. Hence there is at most one point in each quadrant with in-order distance $m\sqrt{2}$ with respect to Q. Such a point can only be the *maximal* element in the quadrant under the ordering "$\prec$", implying $d'_i = \frac{m}{2}\sqrt{2}$. For point Y_j that is not the maximal element in the quadrant, $d_j \leq d'_j$ since d_j involves taking the minimum over a larger set.

We thus have $\sum_{i=1}^{k} d_i \leq \sum_{i=1}^{k} d'_i + 4(m\sqrt{2} - \frac{m}{2}\sqrt{2})$, which gives

$$f(m) \leq 4f(m/2) + 2m\sqrt{2}.$$

Now we can prove the lemma by induction. First, for $m = 1$, $f(1) = \sqrt{2}$, so the base condition is satisfied. For $m \geq 2$, we get

$$\begin{aligned} f(m) &\leq 4f(m/2) + 2m\sqrt{2} \\ &\leq 4(3\sqrt{2}\frac{m^2}{4} - 2\sqrt{2}\frac{m}{2}) + 2m\sqrt{2} \\ &= 3\sqrt{2}m^2 - 2\sqrt{2}m. \end{aligned}$$

□

Now we prove Theorem 4.4.

Proof of Theorem 4.4: Suppose at a time instant t over the m-th sub-channel, $\{(X_k, X_{R(k)}) : k \in \mathcal{T}(t)\}$ is the set of active transmitter-receiver pairs. Define $r_k := |X_k - X_{R(k)}|$ and $r_{ki} := |X_k - X_{R(i)}|$. Let w be a node transmitting at power level $\max_{i \in \mathcal{T}(t)}\{P_i\}$. Let $\mathcal{T}'(t) = \mathcal{T}(t) \backslash \{w\}$. Then the bit-meters per second achieved at this time instant over sub-channel m, with W_i denoting the data rate obtained by X_i, is given by:

$$\begin{aligned} \sum_{i \in \mathcal{T}(t)} r_i W_i &= \sum_{i \in \mathcal{T}(t)} r_i H_m \log_2 (1 + \frac{P_i / r_i^{\alpha}}{N H_m + \sum_{k \in \mathcal{T}(t), k \neq i} \frac{P_k}{r_{ki}^{\alpha}}}) \\ &\leq r_w H_m \log_2 (1 + \frac{P_w / r_w^{\alpha}}{N H_m}) \\ &\quad + \sum_{i \in \mathcal{T}'(t)} r_i H_m \log_2 (1 + \frac{P_i / r_i^{\alpha}}{\sum_{k \in \mathcal{T}(t), k \neq i} \frac{P_k}{r_{ki}^{\alpha}}}). \end{aligned}$$

Because $\ln(1+x^a) \le ax$, for all $a > 0, x > 0$, and $P_w \le P_{max} \le NH_m n^{\frac{\alpha}{2}}$, we have

$$\begin{aligned}
&\sum_{i\in\mathcal{T}(t)} r_i W_i \\
&\quad \le H_m \log_2(e)\left(\alpha(\frac{P_w}{NH_m})^{1/\alpha} + \sum_{i\in\mathcal{T}'(t)} r_i \ln(1 + \frac{P_i/r_i^\alpha}{\sum_{k\in\mathcal{T}(t),k\ne i}\frac{P_k}{r_{ki}^\alpha}})\right) \\
&\quad \le H_m \log_2(e)\left(\alpha\sqrt{n} + \sum_{i\in\mathcal{T}'(t)} r_i \ln(1 + \frac{P_i/r_i^\alpha}{\sum_{k\in\mathcal{T}(t),k\ne i}\frac{P_k}{r_{ki}^\alpha}})\right). \qquad (4.3)
\end{aligned}$$

Let $\mathbb{S}$ be the the summation in (4.3), i.e.,

$$\mathbb{S} := \sum_{i\in\mathcal{T}'(t)} r_i \ln(1 + \frac{P_i/r_i^\alpha}{\sum_{k\in\mathcal{T}(t),k\ne i}\frac{P_k}{r_{ki}^\alpha}})$$

We will bound it by grouping transmissions first into classes based on their range. Then, for each class, only the interference coming from transmitters in the same class is considered. This certainly is lower than the actual interference, but it is still enough to lead to a good upper bound for the achievable rate along each link.

Consider all the transmitters $\{i \in \mathcal{T}'(t)\}$. Divide them into classes based on their transmission ranges $\{r_i := |X_i - X_{R(i)}|\}$, as follows. For $j \ge 0$, define

$$\begin{cases} C_0 := \{i : i \in \mathcal{T}(t), r_i < L_0\}, \\ C_j := \{i : i \in \mathcal{T}(t), L_{j-1} \le r_i < L_j\}, \end{cases}$$

where $L_j := 2^{j-\lceil \log\sqrt{n}\rceil}$. Note that $\frac{2^{j-1}}{\sqrt{n}} < L_j \le \frac{2^j}{\sqrt{n}}$.

Now consider the transmissions in class C_j.

Tessellate the unit square using small squares of side length L_j. Suppose the C_j-*class* receivers are distributed in G_j small squares: $F_g, 1 \le g \le G_j$, such that each contains at least one such receiver. Assume in each small square F_g, there are n_g C_j-*class* receivers. Denote by (X_{gh}, r_{gh}, P_{gh}) the triple for receiver, range, and transmission power, for $0 \le h \le n_g$, and suppose that they are sequenced according to P_{gh} with P_{g0} being the largest.

First we consider the contribution from $\{X_{g0}, 1 \le g \le G_j\}$ to $\mathbb{S}$.

A. Contribution from $\{X_{g0}, 1 \le g \le G_j\}$

Define a total ordering "$\prec$" in the set by ordering them according to P_{g0}, and breaking ties arbitrarily (so that $P_{g_1 0} < P_{g_2 0} \Rightarrow X_{g_1 0} \prec X_{g_2 0}$). Define $l_g := \min(\{\sqrt{2}\} \cup \{|X_{g_1 0} - X_{g0}| : 1 \le g_1 \le G_j, \;\; g_1 \ne g, \;\; X_{g0} \prec X_{g_1 0}\})$. Thus, from the definition of l_g, for any receiver X_{g0}, there must be a transmitter within a distance l_g of X_{g0}'s sender, which is transmitting with power at least P_{g0}. (Note for $g = G_j$, the node w that is transmitting at the globally maximum power level, is within a distance $\sqrt{2}$.)

Thus, the interference at node X_{g0}, I_{g0}, is at least

$$I_{g0} \ge \frac{P_{g0}}{(d_g + L_j)^\alpha}.$$

So the contribution from $\{X_{g0}, 1 \le g \le G_j\}$ is upper bounded by

$$\begin{aligned}
\sum_{g=1}^{G_j} r_{g0} \ln\left(1 + \frac{\frac{P_{g0}}{{r_{g0}}^\alpha}}{I_{g0}}\right) &\le \sum_{g=1}^{G_j} r_{g0} \ln\left(1 + \frac{\frac{P_{g0}}{{r_{g0}}^\alpha}}{\frac{P_{g0}}{(d_g+L_j)^\alpha}}\right) \\
&\le \sum_{g=1}^{G_j} \alpha(d_g + L_j) \quad (\text{Because } \ln(1 + x^\alpha) \le \alpha x) \\
&\le \alpha\left(L_j\left(3\sqrt{2}\frac{1}{{L_j}^2}\right) + L_j G_j\right) \\
&\le \frac{\alpha}{L_j}(3\sqrt{2} + 1), \qquad (4.4)
\end{aligned}$$

where the third inequality is by Lemma 4.5, and the last is because $G_j \le 1/L_j^2$.

B. Contribution from $\{X_{gh}, 1 \le g \le G_j, h \ge 1\}$

For each small square F_g, define the total C_j-class power as $P_g := \sum_{h=0}^{n_g - 1} P_{gh}$. Then the interference seen by X_{gh}, I_{gh}, is at least $\frac{P_g - P_{gh}}{(\sqrt{2}L_j + L_j)^\alpha}$, by considering only those transmitters whose receivers are within small square F_g, and upper bounding the distance of such transmitters. Since P_{g0} is the largest, we have $P_g \ge 2P_{gh}$, implying that

$P_g - P_{gh} \geq \frac{P_g}{2}$. Thus the interference I_{gh} is bounded as

$$I_{gh} \geq \frac{P_g/2}{(\sqrt{2}L_j + L_j)^\alpha}.$$

So, the bit-meters per second from $\{X_{gh}, 1 \leq g \leq G_j, h \geq 1\}$ is upper bounded as

$$\begin{aligned} &\sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} r_{gh} \ln\left(1 + \frac{P_{gh}/r_{gh}^\alpha}{I_{gh}}\right) \\ &\quad \leq \sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} r_{gh} \ln\left(1 + \frac{2P_{gh}}{P_g} \frac{((\sqrt{2}+1)L_j)^\alpha}{r_{gh}^\alpha}\right). \end{aligned} \tag{4.5}$$

For class C_0, since $2P_{gh} \leq P_g$ and $\ln(1 + x^\alpha) \leq \alpha x$, we have

$$\begin{aligned} &\sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} r_{gh} \ln\left(1 + \frac{2P_{gh}}{P_g} \frac{((\sqrt{2}+1)L_j)^\alpha}{r_{gh}^\alpha}\right) \\ &\quad \leq \sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} r_{gh} \ln\left(1 + \frac{((\sqrt{2}+1)L_j)^\alpha}{r_{gh}^\alpha}\right) \\ &\quad \leq \sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} \alpha(\sqrt{2}+1)L_0 \\ &\quad \leq \alpha(\sqrt{2}+1)L_0 \cdot \frac{1}{L_0^2} \\ &\quad \leq \alpha(\sqrt{2}+1)\sqrt{n}. \end{aligned} \tag{4.6}$$

For class C_j with $j > 0$, since $L_{j-1} \leq r_{gh} < L_j$, we have

$$\begin{aligned} &\sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} r_{gh} \ln(1 + \frac{2P_{gh}}{P_g} \frac{((\sqrt{2}+1)L_j)^\alpha}{r_{gh}^\alpha}) \\ &\quad \leq \sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} L_j \ln(1 + \frac{2P_{gh}}{P_g} \frac{((\sqrt{2}+1)L_j)^\alpha}{L_{j-1}^\alpha}) \\ &\quad = \sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} L_j \ln(1 + \frac{2P_{gh}}{P_g}((\sqrt{2}+1)2)^\alpha) \end{aligned}$$

$$\leq \sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} L_j \frac{2P_{gh}}{P_g}((\sqrt{2}+1)2)^{\alpha}$$

$$\leq L_j((\sqrt{2}+1)2)^{\alpha} \sum_{g=1}^{G_j} 2 \quad (\text{Because } \textstyle\sum_h P_{gh} = P_g)$$

$$\leq L_j((\sqrt{2}+1)2)^{\alpha} 2 \frac{1}{L_j^2} = ((\sqrt{2}+1)2)^{\alpha} \frac{2}{L_j}. \tag{4.7}$$

Finally, combining (4.4, 4.6, 4.7), we have

$$\begin{aligned}
\mathbb{S} &= \sum_{i \in T'} r_i \ln \left(1 + \frac{P_i/r_i^{\alpha}}{\sum_{k \in T, k \neq i} \frac{P_k}{r_{ki}^{\alpha}}}\right) \\
&= \sum_{j \geq 0} \sum_{i \in C_j} r_i \ln (1 + \frac{P_i/r_i^{\alpha}}{\sum_{k \in T, k \neq i} \frac{P_k}{{r_{ki}}^{\alpha}}}) \\
&= \sum_{j \geq 0} \sum_{g=1}^{G_j} r_{g0} \ln (1 + \frac{P_{g0}/{r_{g0}}^{\alpha}}{I_{g0}}) + \sum_{j \geq 0} \sum_{g=1}^{G_j} \sum_{h=1}^{n_g-1} r_{gh} \ln (1 + \frac{P_{gh}/r_{gh}^{\alpha}}{I_{gh}}) \\
&\leq \sum_{j \geq 0} \frac{\alpha}{L_j}(3\sqrt{2}+1) + (\alpha(\sqrt{2}+1)\sqrt{n}) + \sum_{j > 0} (((\sqrt{2}+1)2)^{\alpha} \frac{2}{L_j}) \\
&= \frac{\alpha}{L_0}(3\sqrt{2}+1) \sum_{j \geq 0} 2^{-j} + (\alpha(\sqrt{2}+1)\sqrt{n}) \\
&\qquad + (((\sqrt{2}+1)2)^{\alpha} \frac{2}{L_0}) \sum_{j > 0} 2^{-j} \\
&< \alpha 2\sqrt{n}(3\sqrt{2}+1)(2) + (\alpha(\sqrt{2}+1)\sqrt{n}) + (((\sqrt{2}+1)2)^{\alpha} 4\sqrt{n}) \cdot 1 \\
&= \sqrt{n}(\alpha(13\sqrt{2}+5) + 4((\sqrt{2}+1)2)^{\alpha}.
\end{aligned} \tag{4.8}$$

Given the above bound, by (4.3), we now have

$$\begin{aligned}
\sum_{i \in T} r_i W_i &\leq H_m \log_2(e)(\alpha\sqrt{n} + \sqrt{n}(\alpha(13\sqrt{2}+5) + 4((\sqrt{2}+1)2)^{\alpha})) \\
&= H_m \log_2(e)\sqrt{n}(\alpha(13\sqrt{2}+6) + 4((\sqrt{2}+1)2)^{\alpha}) \\
&= H_m \mu_{\alpha} \sqrt{n}.
\end{aligned}$$

Summing over all the sub-channels proves the theorem. □

Remark 4.6. It is interesting to notice that the bound does not require that there be a minimum separation distance between two nodes. This is in contrast to the results under the information theoretic framework we will present in Sections 8–10, where, usually, a minimum separation is needed. Note that the attenuation function in the model has a singular point at the origin. When nodes can be arbitrary close to each other, for example in so called "dense networks model, it becomes very subtle as to how to model attenuation. Indeed for physical reasons nodes may have to necessarily be separated by a minimum positive distance, which is the model we assume in the information theoretic treatment in Sections 8–11.

Given Theorem 4.4, by considering a square of area A and shrinking it to a unit square using the mapping $P_i' = P_i A^{-\alpha/2}$ for the power levels, we know the following is also true.

Theorem 4.7. Consider n-node networks within a square of area A, such that $\alpha > 2$, $N > 0$, and power level is bounded by $P_{max} = H_m N (nA)^{\frac{\alpha}{2}}$ for any channel with bandwidth H_m. Then the transport capacity scales as $\Theta(\sqrt{An})$.

4.4 Notes

This section is based on [2] and [12]. The transport capacity under the Physical Model was first studied in [12], which gives a constructive $\Theta(\sqrt{n})$ lower bound, and an $O(n^{\frac{\alpha-1}{\alpha}})$ upper bound, where α is the path loss exponent. This upper bound was improved to $\Theta(\sqrt{n})$ in [2] using the techniques presented here, closing the gap between lower and upper bounds.

5

The Throughput Capacity of Random Wireless Networks – Protocol Model

In this section, we study wireless networks where nodes are *randomly* distributed.

Often the locations of transmitter-receiver pairs cannot be chosen, or unknown *a priori*. One is therefore interested in how random node locations will influence the performance of wireless networks. We present results that address this question through the study of the scaling behavior of the *throughput capacity*, which is the guaranteed rate that can be supported uniformly for all source–destination pairs. For simplicity we assume that every node has originating traffic with a single randomly chosen destination. Under appropriate assumptions, the results can be generalized to other contexts where not all nodes are source, or a source may have (differing) traffic for several destination nodes.

We describe the model for random networks and first consider the Protocol Model[1]. In addition to (2.1) or (2.2), in this section we impose a further constraint on the range of transmitting nodes. We will suppose that all nodes employ a common transmission range. Clearly this can only lead to the diminishment of a network's capacity; however we show that the effect is not significant. We define the throughput

[1] The Physical Model will be discussed in Section 6.

capacity and show that it is of order $\Theta(\frac{W}{\sqrt{n \log n}})$. Thus these results suggest that random networks are nearly best case, since even by optimally placing nodes one cannot obtain a common throughput exceeding $O(\frac{1}{\sqrt{n}})$ bits/sec when destinations lie at non-vanishing distances from their sources. Moreover it also suggests that not too much is lost by homogeneous operation constrained to employ a common range for all transmissions. The results also provide a construction of an order optimal architecture, which is useful for designers.

5.1 Random networks, Protocol Model, and throughput capacity

Consider a *random network* where n nodes are uniformly and independently distributed in a *unit square*. Each node has a random destination – a node – it wishes to send packets to. The destination for node $X_i, i \in \{1, \cdots, n\}$, is chosen as follows. A position is first picked uniformly from within the unit square, then the node nearest to it is chosen as node X_i's destination.

We consider the case when the transmission range and the traffic pattern are homogeneous for each node, defined as follows.

Definition 5.1. *The Protocol Model*: All nodes employ a common transmission range r for all their transmissions. Node X_i can successfully transmit to node $X_{R(i)}$ if

i) The distance between the transmitter and the receiver is no more than r, i.e., $|X_i - X_{R(i)}| \leq r$.
ii) For every node X_k, $k \neq i$, transmitting at the same time, $|X_k - X_{R(i)}| \geq (1 + \Delta)r$.
iii) The data rate between such successful transmitter-receiver pair is W bits/second.

Remark 5.2. It is easy to extend all the proofs we present to networks under the variant of the Protocol Model (3.2) to get the same results. Also, the results hold even if the channel is split into several sub-channels, but with the same sum-rate of W bits/second.

Now we define the per-node throughput of random wireless networks.

Definition 5.3. *Feasible Throughput:* A throughput of $\lambda(n)$ bits per second for a random network of n nodes is said to be *feasible* if there is a choice of the common range r, and a scheme to schedule transmissions and choice of routes between sources and destinations, so that *every* node can achieve data transfer to its destination at such rate. The scheme needs to specify, for each link, when it should be active under the transmission model (Definition 5.1), and whose packets it should send. A packet can be delayed at intermediate nodes before reaching its destination.

Since we are considering random networks, the transmission arrangement and hence the feasible throughput will be influenced by the randomness. We introduce the following definition for the throughput capacity of wireless networks.

Definition 5.4. *Order of The Throughput Capacity of Random Wireless Networks:* The throughput capacity of random wireless networks is said to be of order $\Theta(f(n))$ bits per second if there exist deterministic positive constants c and c' such that

$$\begin{aligned} &\lim_{n\to\infty} \text{Prob}(\ \lambda(n) = cf(n) \text{ is feasible }) = 1; \text{ while} \\ &\lim_{n\to\infty} \text{Prob}(\ \lambda(n) = c'f(n) \text{ is feasible }) < 1. \end{aligned} \tag{5.1}$$

5.2 Main results

We have the following main result.

Theorem 5.5. The order of the throughput capacity of random wireless networks under the Protocol Model is

$$\lambda(n) = \Theta\left(\frac{W}{\sqrt{n\log n}}\right) \textit{ bits/sec.}$$

In fact the condition (5.1) can be strengthened to exhibit a sharp cutoff phenomenon: There are positive constants c_1 and c_2, not depending on n, Δ, or W, such that $\lambda(n) = \frac{c_1 W}{(1+\Delta)^2\sqrt{n \log n}}$ bits/sec is feasible, and $\lambda(n) = \frac{c_2 W}{\Delta^2 \sqrt{n \log n}}$ bits/sec is infeasible, both with probability approaching one as $n \to \infty$. Thus the limit of the left hand side in (5.1) is actually zero.

5.3 A constructive lower bound on throughput capacity

In this subsection, we present a scheme that achieves $\lambda(n) = \frac{c_1 W}{(1+\Delta)^2\sqrt{n \log n}}$ bits/sec for every node in the network to its chosen destination, with probability approaching one as $n \to \infty$.

First the unit square is divided into small cells of such a size that each of them holds at least one but no more than $O(\log n)$ nodes. Second, the cells are divided into a finite number of non-interfering groups which can take turns in transmission without causing interference. Finally we show that a simple routing strategy – forwarding a packet from cell to cell “along” the line connecting the originating cell to the destination cell containing each node and its randomly chosen end point – can fulfill the job.

Remark 5.6. In the proof, cells are chosen to be squares as an example mainly to simplify presentation. Based on results in Section IV of [12], any Voronoi tessellation (See Figure 5.1) satisfying the following property can be used to achieve the same result:

- Every Voronoi cell contains a disk of area $\tilde{K}_1 \log n/n$ for some $\tilde{K}_1 > 1$; and
- Every Voronoi cell is contained in a disk of area $\tilde{K}_2 \log n/n$ for some $\tilde{K}_2 > \tilde{K}_1$.

5.3.1 Tessellating the unit square by small squares

We tessellate the unit square by square cells of side $s_n = \sqrt{\frac{K \log n}{n}}$; see Figure 5.2, and consider the number of nodes within each of them.

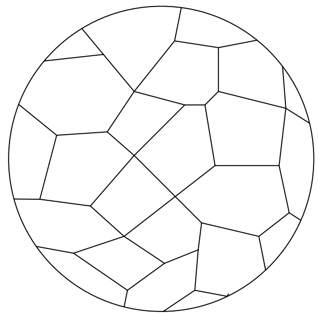

Fig. 5.1 A Voronoi Tessellation.

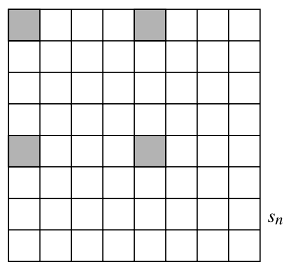

Fig. 5.2 A tessellation of the unit square into small square cells. The shaded cells are concurrently transmitting cells with $M = 4$.

Lemma 5.7. For any $K > 1$, we have that, with probability going to one, each cell holds at least one but no more than $Ke\log n$ nodes.

Proof. Observe that, for a particular cell, a particular node X_i, $1 \leq i \leq n$, falls into it with probability $p_n := s_n^2 = \frac{K\log n}{n}$, a Bernoulli event.

So the probability that a particular cell is empty is $(1 - p_n)^n$. By the union bound, the probability that at least one cell is empty is upper-bounded by the total number of cells times $(1 - p_n)^n$, which is $\frac{1}{s_n^2}(1 - p_n)^n = \frac{n}{K\log n}(1 - \frac{K\log n}{n})^n$. Since $1 - x \leq \exp(-x)$, we know $(1 - \frac{K\log n}{n})^n \leq n^{-K}$, and thus $\frac{1}{s_n^2}(1 - p_n)^n \to 0$ as $n \to \infty$ whenever $K > 1$.

Now consider the upper-bound for the number of nodes, N_n, in a particular cell. The number of nodes has a binomial distribution with parameters (p_n, n). So by the Chernoff bound we have,

$$\Pr(N_n > Ke\log n) \le \frac{E\exp(N_n)}{\exp(Ke\log n)}.$$

Since $E\exp(N_n) = (1+(e-1)p_n)^n \le n^{K(e-1)}$ (because $1+x \le \exp(x)$), we have $\Pr(N_n > Ke\log n) \le n^{-K}$. As long as $K > 1$, by the union bound we know

$$\Pr(\text{Some small cell has more than } Ke\log n \text{ nodes}) \le n\frac{1}{n^K} \to 0,$$

as n goes to infinity. □

Remark 5.8. It is easy to show by the Chernoff Bound that, for $K > 1/(1-e^{-1})$, there exists $a = a(K) > 0$ such that Pr(Every cell contains at least $a\log n$ nodes)$\to \infty$, as $n \to \infty$. So in fact each cell contains $\Theta(\log n)$ nodes.

In the sequel we fix K to be some constant larger than 1.

5.3.2 Concurrently transmitting cells, transmission schedule, and interference

There are $1/s_n \times 1/s_n = \sqrt{n/(K\log n)} \times \sqrt{n/(K\log n)}$ cells in the tessellation. For simplicity, assume $\frac{1}{s_n}$ is an integer. Index them as $S_{i,j}$, with i denoting the column number and j the row number.

For a positive integer M, let $C(k_1,k_2) := \{S_{i,j} : i \mod M = k_1,\ j \mod M = k_2\}$ for $0 \le k_1, k_2 \le M-1$. All the cells in $C(k_1,k_2)$ are called "non-interfering cells", and $C(k_1,k_2)$ is called a "non-interfering group;" see Figure 5.2. (The name anticipates the property established in the sequel that after choice of an appropriate range r, one node in each of the cells in $C(k_1,k_2)$ can transmit, with all the concurrent transmissions being successful.)

Now we define the transmission range and schedule.

Transmission Schedule:

All nodes choose a common transmission range $r_n = 2\sqrt{2}s_n$, so that every node can cover all its neighboring cells. All the M^2 non-interfering groups $C(i, j)$ take turns to transmit in a round robin manner. If group $C(i_0, j_0)$ is scheduled in a time slot, one node from each of the cells of $C(i_0, j_0)$ can transmit (simultaneously) to its neighboring cells.

As will be shown in later subsections, any node in an active cell can be the active node for a particular slot. The chosen node will then send data packets to a node in a neighboring cell. This will be specified in more detail in the sequel when elaborating on the routes followed by packets.

One node in each of the cells in a non-interfering group can transmit successfully in the same active slot. Following from the Protocol Model (Definition 5.1), it is easy to verify that if M is large enough then there will be no interference from nearby cells.

Lemma 5.9. There exists $c_3 > 0$ such that, if the network transmits according to the above transmission schedule with $M = c_3(1 + \Delta)$, then every transmission will be successful.

5.3.3 The routing

According to the model, each node X_i, $1 \leq i \leq n$, generates data packets at rate $\lambda(n)$ with an end destination chosen as the node nearest to a randomly chosen location Y_i. Denote by $X_{dest(i)}$ the node nearest to Y_i, and by L_i the straight-line segment connecting X_i and Y_i. The packets generated by X_i will be forwarded toward $X_{dest(i)}$ in a multi-hop manner, from cell to cell in the order that they are intersected by L_i. In each hop, the packet is transmitted from one cell to the next cell intersecting L_i. Any node in the cell can be chosen as a receiver. Finally, after reaching the cell containing Y_i, the packet will be forwarded to $X_{dest(i)}$ in the next active slot for that cell. This can be done because $X_{dest(i)}$ is within a range of r_n to any node in that cell, by Lemma 5.7.

A. A Bound on the Number of Routes Each Cell Needs to Serve

First, we bound the probability that a line will intersect a particular cell.

Lemma 5.10. There exists a constant $c_4 > 0$ such that for every line L_i and cell $S_{k_0 j_0}$,

$$p := \text{Prob}\{\text{Line } L_i \text{ intersects } S_{k_0,j_0}\} \leq c_4 \sqrt{\frac{\log n}{n}}.$$

Proof. Note $S_{k_0 j_0}$ is contained in a disk of radius $d_n =: \frac{1}{\sqrt{2}} s_n = \sqrt{\frac{K \log n}{2n}}$ centered at s_n's center D; see Figure 5.3. Suppose X_i is at distance x from the disk. Extend the two tangent lines originating from X_i equally such that $|X_i A| = |X_i B|$ and $|X_i C| = \sqrt{2}$, where C is the mid-point of AB.

Then L_i intersects $S_{k_0 j_0}$ only if Y_i is in the shaded area. Its area is less than the minimum of 1 and the area of the triangle, which is $\sqrt{2} \cdot \frac{\sqrt{2}}{\sqrt{(x+d_n)^2 - d_n^2}}$, less than $2d_n/x$.

Since X_i is uniformly distributed, the probability density that it is at a distance x away from the disk is bounded above by $c_6 \pi (x + d_n)$, for some constant $c_6 > 0$. So we get

$$\begin{aligned} \text{Prob}(L_i \text{ intersects } S_{k_0 j_0}) &\leq \int_{d_n}^{\sqrt{2}} \left(\frac{2d_n}{x} \ \vee\ 1\right) \cdot c_6 \pi (x + d_n) dx \\ &\leq c_4 \sqrt{\frac{\log n}{n}}. \end{aligned}$$

□

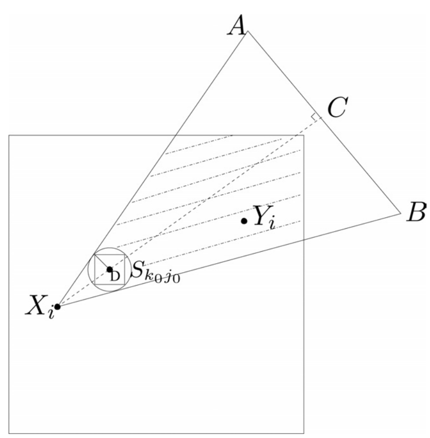

Fig. 5.3 The exclusion region when I_k is a disk.

Based on the above lemma, we can show the following uniform bound on the number of routes served by each cell.

Lemma 5.11. We have, for any constant $c_5 > c_4$,

$$\text{Prob}\left(\sup_{(k,j)} \{ \text{ Number of lines } L_i \text{ intersecting } S_{k,j} \} \leq c_5\sqrt{n \log n}\right) \to 1.$$

Proof. First we bound the number of routes served by one particular cell S_{k_0,j_0}. Define iid random variables I_i, $1 \leq i \leq n$, as follows:

$$I_i = \begin{cases} 1, & \text{if } L_i \text{ intersects } S_{k_0,j_0}; \\ 0, & \text{if not.} \end{cases}$$

Then $Pr(I_i = 1) = p$, $\forall i$, where p is defined as in Lemma 5.10.

Denote by Z_n the total number of routes served by S_{k_0,j_0}. Then $Z_n := I_1 + \cdots + I_n$. Thus by the Chernoff Bound, for all positive m and a, $Pr(Z_n > m) \leq \frac{Ee^{aZ_n}}{e^{am}}$. Because $1 + x \leq e^x$, we have

$$\begin{aligned} Ee^{aZ_n} = (1 + (e^a - 1)p)^n &\leq \exp(n(e^a - 1)p) \\ &\leq \exp(c_4(e^a - 1)\sqrt{n \log n}). \end{aligned}$$

Now choosing $m = c_5\sqrt{n \log n}$, we get

$$Pr(Z_n > c_5\sqrt{n \log n}) \leq \exp(\sqrt{n \log n}(c_4(e^a - 1) - ac_5)).$$

Since $c_5 > c_4$, one can choose a small enough such that

$$Pr(Z_n > c_5\sqrt{n \log n}) < \exp(-\epsilon\sqrt{n \log n}), \tag{5.2}$$

for some constant $\epsilon > 0$.

Thus by the union bound, we have

$$\begin{aligned} &\text{Prob}\left(\text{Some cell intersects more than } c_5\sqrt{n \log n} \text{ lines}\right) \\ &\quad \leq \sum_{k,j} \text{Prob}\left(\text{Cell } S_{jk} \text{ intersects more than } c_5\sqrt{n \log n} \text{ lines}\right) \\ &\quad \leq \frac{1}{s_n^2} \exp(-\epsilon\sqrt{n \log n}) \\ &\quad = \frac{n}{K \log n} \exp(-\epsilon\sqrt{n \log n}). \end{aligned}$$

The right hand side goes to zero as n goes to infinity. □

5.3.4 Lower bound on throughput capacity of random networks

From Lemma 5.9 we know that there exists a transmitting schedule such that in every $M^2 = (c_3(1+\Delta))^2$ slots, each cell gets one slot to transmit at rate W bits/second with transmission range r_n. So the rate at which each cell gets to transmit is $W/(c_3(1+\Delta))^2$ bits/second.

By Lemma 5.11, each cell needs to transmit at rate less than $\lambda(n)c_5\sqrt{n\log n}$, with probability approaching one. This can therefore be accommodated by all cells if

$$\lambda(n)c_5\sqrt{n\log n} \leq \frac{W}{(c_3(1+\Delta))^2}.$$

Thus we have proven the achievability part of Theorem 5.5. □

Remark 5.12. In each cell, the traffic passing through that cell can be handled by any designated node in that cell.

5.4 Upper bound on throughput capacity

We now establish the upper bound on throughput capacity. First, for a given choice of common range r, we study the probability that there is at least one isolated node. From this a lower bound is determined on the common range that nodes need to employ in order for no node to be isolated. Then a similar technique, as for arbitrary networks, is used to show the upper bound.

A. The Probability of an Isolated Node

In [21, 11] the following result is proved on the appearance of an isolated node in a random network.

Theorem 5.13. For all $\epsilon > 0$, if $r_n = (1-\epsilon)\sqrt{\frac{\log n}{n\pi}}$,then Prob(There is an isolated node) $\to 1$, as n goes to infinity; while if $r_n = (1+\epsilon)\sqrt{\frac{\log n}{n\pi}}$, then Prob(There is an isolated node) $\to 0$, as n goes to infinity.

Since each node does need to communicate with some other node, no node can be allowed to be an isolated node. Given this result, it follows that r_n should be asymptotically larger than $\sqrt{\frac{\log n}{n\pi}}$.

B. Number of Simultaneous Transmissions

Under the Protocol Model, as established in the upper bound for arbitrary networks (see (2.3) of Section 2) we know that each transmission "consumes" area. Specifically, disks of radius $\frac{\Delta}{2}r_n$ around every transmitter are disjoint. Notice that the area of each disk is $\frac{\pi\Delta^2 r_n^2}{4}$, and at least $1/(4\pi)$ portion is within the unit square. Thus we get the following lemma.

Lemma 5.14. The maximum number of simultaneous transmissions feasible is no more than $\frac{1}{\frac{1}{4\pi}\frac{\pi\Delta^2 r_n^2}{4}} = \frac{16}{\Delta^2 r_n^2}$.

C. The Upper Bound for Throughput Capacity

Let $\bar{L}$ be the expected distance between two uniformly and independently chosen points within the unit square. Then the expected length from a node to its destination is $\bar{L} - o(1)$ since there is always a node within $\Theta(\sqrt{\frac{\log n}{n}})$ distance from any point (Lemma 5.7). Thus on average each packet needs $\frac{\bar{L}-o(1)}{r_n}$ hops to the destination. Since each node generates packets at rate $\lambda(n)$, this means the bits per second being transmitted by the whole network is at least $n\lambda(n)\frac{\bar{L}-o(1)}{r_n}$. Hence by Lemma 5.14 we have

$$n\lambda(n)\frac{\bar{L}-o(1)}{r_n} \leq W\frac{16}{\Delta^2 r_n^2}.$$

By Theorem 5.13 we know $\sqrt{\log n/n}$ is a lower bound for r_n; so we have

$$\lambda(n) \leq \frac{16W}{n\Delta^2 r_n(\bar{L}-o(1))} \leq \frac{c_7 W}{\Delta^2\sqrt{n\log n}}.$$

Thus we have shown the upper bound for the throughput capacity of random wireless networks. □

5.5 Notes

These results on the throughput capacity of random wireless networks are from [12]. We have used the Chernoff bound rather than Vapnik-Chervonenkis theory.

6

The Throughput Capacity of Random Wireless Networks – Physical Model

In this section, we study the performance of random wireless networks under the Physical Model.

The focus is on the maximum common throughput that can be supported for every node simultaneously.

We first define the Physical Model and then study the scaling behavior of the throughput capacity. We show that a common throughput of order $\Theta(\frac{W}{\sqrt{n\log n}})$ bits per second is feasible; while $\Theta(\frac{W}{\sqrt{n}})$ bits per second is not. This again shows that the results under the simple geometric model are fairly robust to assumptions that are more physically realistic.

6.1 Random networks, Physical Model, and throughput capacity

As in the last section, we consider a random network where n nodes are uniformly and independently distributed in a *unit square*. Each node has a random destination node it wishes to send packets to. The destination for node $X_i, i \in \{1, \cdots, n\}$, is chosen as in Section 5. A position is first picked uniformly from within the unit square, then the node nearest to it is chosen as node X_i's destination.

We consider the case where the power level, rather than the transmission range, and the traffic requirements, are homogeneous for each node, as follows.

Definition 6.1. The Physical Model: All nodes employ a common transmission power P for all their transmissions. Node X_i can successfully transmit to node $X_{R(i)}$ if

i) The signal to interference plus noise ratio (SINR) at the receiver is no less than a threshold β, i.e., assuming $\{X_k; k \in \mathcal{T}(t)\}$ is the subset of nodes simultaneously transmitting at time t, then

$$\frac{\frac{P}{|X_i - X_{R(i)}|^\alpha}}{N + \sum_{\substack{k \in \mathcal{T}(t) \\ k \neq i}} \frac{P}{|X_k - X_{R(i)}|^\alpha}} \geq \beta, \tag{6.1}$$

where N is the ambient noise power level. The signal power is assumed to decay with distance r as $\frac{1}{r^\alpha}$, with $\alpha > 2$.

ii) The data rate between every successful transmitter-receiver pair is W bits/second.

The definition of throughput under the Physical Model is the same as in the last section. A throughput of $\lambda(n)$ bits per second for a random network of n nodes is feasible if there is a scheme to schedule transmissions so that every node can send data to its destination at such a rate. A packet can be delayed at intermediate nodes before reaching its destination.

We are interested in finding the maximum feasible throughput for a network.

Definition 6.2. *The Order of the Throughput Capacity of Random Wireless Networks:* The throughput capacity of random wireless networks is said to be of order $\Theta(f(n))$ bits per second if there exist deterministic positive constants c and c' such that

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

6.2 Main results

The following theorem establishes an upper bound and a separate lower bound for the throughput capacity.

Theorem 6.3. For the Physical Model, there exist positive constants c, c' such that a throughput of $\lambda(n) = \frac{cW}{\sqrt{n \log n}}$ bits/sec is feasible, while $\lambda(n) = \frac{c'W}{\sqrt{n}}$ bits/sec is not, both with probability approaching one as $n \to \infty$.

Specifically, the lower bound can be achieved by using a similar strategy as in the achievability part for the Protocol Model, where the small square cells are divided into M^2 non-interfering groups and at each time slot only one group is allowed to transmit. It can be shown that, for any common power level $P > 0$, one can choose M sufficiently large (a function of P, α, β only) to achieve $\Theta(\frac{W}{\sqrt{n \log n}})$ bits/sec. On the other hand, suppose $\bar{L}$ is the mean distance between two points independently and uniformly distributed in the domain, then there is a deterministic sequence $\epsilon(n) \to 0$, not depending on N, α, β or W, such that $\sqrt{\frac{8}{\pi}} \frac{W}{\bar{L}(\beta^{\frac{1}{\alpha}} - 1)} \frac{1+\epsilon(n)}{\sqrt{n}}$ bit-meters/sec is infeasible with probability approaching one as $n \to \infty$.

6.3 A constructive lower bound

A strategy similar to that of the last section on the Protocol Model is used. First the unit square is tessellated into small square cells, and non-interfering groups are defined. Then, for any $P > 0$, we show that the SINR for successful transmission can be achieved if the number of such groups is sufficiently large (but still constant). Thus $\Theta(\frac{W}{\sqrt{n \log n}})$ bits/sec throughput is achieved.

A. Square Tessellation and non-interfering groups

Tessellate the unit square by small squares of side $s_n = \sqrt{\frac{K \log n}{n}}$ for some $K > 1$; see Figure 5.2.

There are $1/s_n \times 1/s_n = \sqrt{n/(K \log n)} \times \sqrt{n/(K \log n)}$ square cells. Furthermore, by Lemma 5.7, there is at least one node, but no more than $Ke \log n$ nodes, in each cell s, with probability approaching one as $n \to \infty$.

Index the cell s as $S_{i,j}$, with i denoting the column number and j the row number. For a positive integer M, let $C(k_1, k_2) := \{S_{i,j} : i \mod M = k_1,\ j \mod M = k_2\}$, for all $0 \le k_1, k_2 \le M - 1$. All the cells in $C(k_1, k_2)$ are called "non-interfering cells"; see Figure 5.2.

The transmission schedule is the same as in the last section, which we briefly summarize as follows.

B. Transmission Schedule

In every M^2 time slots, every cell gets one slot to transmit. When a cell has an active slot in which it is allowed to transmit, one node in it transmits to a node in one of the neighboring cells according to the routing algorithm specified in last section. The receiver node can be any node in the neighboring cell.

If each transmission can be guaranteed to be successful, then $\Theta(\frac{W}{\sqrt{n \log n}})$ bits/sec is achievable, just as in the analysis in last section.

Now we show that successful transmission, i.e., a SINR ratio larger than a given β, can be achieved if M is large enough (but still a constant as $n \to \infty$).

C. Adjusting M to Make $SINR \ge \beta$

Note that every node employs common transmission power P for transmission.

Consider an active sender X_i and its intended receiver $X_{R(i)}$, lying in cell S_{j_0,k_0}. First, the distance between them is no more than $2\sqrt{2}s_n$, i.e., twice the diagonal of a cell. Second, all the interference comes from the same non-interfering group to which S_{j_0,k_0} belongs. So, if one draws squares of size $2k \cdot Ms_n$, $k = 1, 2, \cdots$, centered at the lower left corner of S_{j_0,k_0}, then there are at most $8k$ interfering cells from the k-th square; see Figure 5.2. Furthermore, the distance from a interfering cell to S_{j_0,k_0} is at least $k \cdot Ms_n - s_n$.

Now we can calculate a lower bound on the achieved SINR at a receiver as follows:

$$\frac{\frac{P_i}{|X_i - X_{R(i)}|^\alpha}}{N + \sum_{\substack{k\in\mathcal{T}(t)\\ k\neq i}} \frac{P_k}{|X_k - X_{R(i)}|^\alpha}} \geq \frac{\frac{P}{(2\sqrt{2}s_n)^\alpha}}{N + \sum_{k=1}^\infty 8k \frac{P}{(kMs_n - s_n)^\alpha}} = \frac{\frac{P}{(2\sqrt{2})^\alpha}}{Ns_n^\alpha + \frac{8P}{M^\alpha}\sum_{k=1}^\infty \frac{k}{(k-\frac{1}{M})^\alpha}}. \tag{6.2}$$

Since $\sum_{k=1}^\infty \frac{k}{(k-\frac{1}{M})^\alpha}$ converges when $\alpha > 2$, and $s_n \to 0$, we see that by letting M sufficiently large, the SINR can be made larger than the specified β. Thus successful transmissions are guaranteed for such a M, and $\Theta(\frac{W}{\sqrt{n\log n}})$ bits/sec is indeed achieved.

Remark 6.4. The sum $\sum_{k=1}^\infty \frac{k}{(k-\frac{1}{M})^\alpha}$ can be bounded as follows:

$$\begin{aligned}
\sum_{k=1}^\infty \frac{k}{(k-\frac{1}{M})^\alpha} &= \sum_{k=1}^\infty \frac{1}{(k-\frac{1}{M})^{\alpha-1}} + \frac{1}{M}\sum_{k=1}^\infty \frac{1}{(k-\frac{1}{M})^\alpha} \\
&\leq \frac{1}{(1-\frac{1}{M})^{\alpha-1}} + \int_{1-\frac{1}{M}}^\infty \frac{1}{x^{\alpha-1}} \\
&\quad + \frac{1}{M}\left(\frac{1}{(1-\frac{1}{M})^\alpha} + \int_{1-\frac{1}{M}}^\infty \frac{1}{x^\alpha}\right) \\
&= \frac{1}{(1-\frac{1}{M})^{\alpha-1}} + \frac{(1-\frac{1}{M})^{-(\alpha-2)}}{\alpha-2} + \frac{1}{M}\frac{1}{(1-\frac{1}{M})^\alpha} \\
&\quad + \frac{1}{M}\frac{(1-\frac{1}{M})^{-(\alpha-1)}}{\alpha-1}.
\end{aligned}$$

6.4 Upper bound

In Section 2 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 arbitrary wireless networks under the Protocol Model. We can show this 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, and the average distance between a source node and its destination is $\bar{L} - o(1)$.

Suppose that, X_i is concurrently successfully transmitting to $X_{R(i)}$, while X_j is to $X_{R(j)}$. From the Physical Model 6.1, we have

$$\frac{\frac{P}{|X_i - X_{R(i)}|^\alpha}}{\frac{P}{|X_j - X_{R(i)}|^\alpha}} \geq \beta,$$

which means that

$$|X_j - X_{R(i)}| \geq (1 + \Delta)|X_i - X_{R(i)}|,$$

with $\Delta := (\beta^{\frac{1}{\alpha}} - 1)$. Hence any set of transmissions feasible for random networks under the Physical Model corresponds to a set of successful transmissions for an arbitrary network, under the Protocol Model. Thus the upper bound on the transport capacity for the latter also holds for the former.

6.5 Notes

These results are from [12].

7

Improved Throughput for Random Wireless Networks Using Differentiated Transmission Ranges

In Section 6 we considered the throughput capacity of random wireless networks under the Physical Model, assuming every node employs the same transmission power. The constructive lower bound shows that $\Theta(\frac{W}{\sqrt{n \log n}})$ bits per second for each node is achievable; while $\Theta(\frac{W}{\sqrt{n}})$ bits per second is an upper bound.

This section addresses this gap for a slightly different model by showing a constructive scheme achieving $\Theta(\frac{W}{\sqrt{n}})$ bits per second for every node. In this model, the network is generated as a Poisson random process with density n in the unit square, and each node is only the destination of one other node. In contrast to the scheme used in the proceeding section, different ranges are used for different transmissions. The import of this is that even the factor $\sqrt{\log n}$ present in the earlier results employing a common range can be dispensed with, however at the expense of a more complicated architecture for packet transport, as shown below.

By using techniques from percolation theory [10], one can show that a "highway system" can be formed in the random network. The claimed rate is then achieved by routing local traffic onto the highway system, and then distributing packets off later to their destinations.

The results can also be generalized to random networks under the Protocol Model when different transmission ranges are allowed.

7.1 Model

Consider a domain consisting of a unit square, where nodes are randomly distributed according to a Poisson point process with density n. Each node X_i uniformly picks a destination node as a destination for the traffic originating at that node. However, an additional restriction is imposed that each node serves as exactly one destination[1]. Thus the source–destination pairs yield a permutation matrix. Communication uses a multi-hop mode, and a node X_i may select a power level P_i for its transmission of a particular packet. Whether a transmission is successful depends on the received SINR, and we recall the Physical Model here.

Definition 7.1. *The Physical Model:* Node X_i can successfully transmit to node $X_{R(i)}$ if

i) The signal to interference plus noise ratio (SINR) at the receiver is no less than a threshold β, i.e., assuming $\{X_k; k \in \mathcal{T}\}$ is the subset of nodes simultaneously transmitting,

$$\frac{\frac{P_i}{|X_i - X_{R(i)}|^\alpha}}{N + \sum_{\substack{k \in \mathcal{T} \\ k \neq i}} \frac{P_k}{|X_k - X_{R(i)}|^\alpha}} \geq \beta, \tag{7.1}$$

where N is the ambient noise power level. The signal power decays with distance r as $\frac{1}{r^\alpha}$ with $\alpha > 2$.

ii) The data rate between every successful transmitter-receiver pair is W bits/second.

---

[1] The destination selection here is slightly different to the one in last section; however the difference is not fundamental. One can also easily show the average source–destination distance here is also a positive constant.

7.2 Overview of the constructive scheme

The main idea is to show first that a "highway system" can be formed. It consists parallel horizontal highways, as well as parallel vertical highways, that go from one side of the unit square to the opposite side. Specifically, to form the horizontal highways, the unit square is divided into horizontal rectangles of width $\frac{\sqrt{2}c}{\sqrt{n}} \log \frac{n}{\sqrt{2}c}$. Then using techniques from percolation theory, one can show at least $\gamma \log \frac{n}{\sqrt{2}c}$ horizontal highways can be formed within each rectangle, with each hop being no more than $2\sqrt{2}c/\sqrt{n}$, with probability approaching one as $n \to \infty$. By symmetry, the vertical highways can be formed similarly.

Once the highway system is formed, one can show that a time sharing scheme allows each highway to route packets at a constant rate. The idea is to use non-interfering groups, which allow nearby nodes to share the transmission slots, as in Section 5.

Now, the routing of a packet from its source to the destination can be divided into three stages. In the first stage, the source node forwards the packet to a nearby highway – it is no more than $c\sqrt{2}\log(\sqrt{n}/\sqrt{2}c)/\sqrt{n}$ away. Then the packet is routed along the highway system, first horizontally and then vertically, towards its end destination. In the last stage, the packet is delivered to the destination from the highway nearby, again with distance no more than $c\sqrt{2}\log(\sqrt{n}/\sqrt{2}c)/\sqrt{n}$. All the above happen with probability going to one, with $\Theta(\frac{W}{\sqrt{n}})$ bits/sec rate for each node, by carefully choosing all the parameters involved.

7.3 Highway system

In this section, first the construction of the highway system will be shown, then analysis will be presented showing that constant bit rate for each highway can be indeed achieved.

A. Construction

Tessellate the unit square first using 45°-angled squares of size $c/\sqrt{n}$, as depicted in the left-hand side of Figure 7.1. For each positive c, we have

$$\text{Prob(A square contains at least one node)} = 1 - e^{-c^2} := p.$$

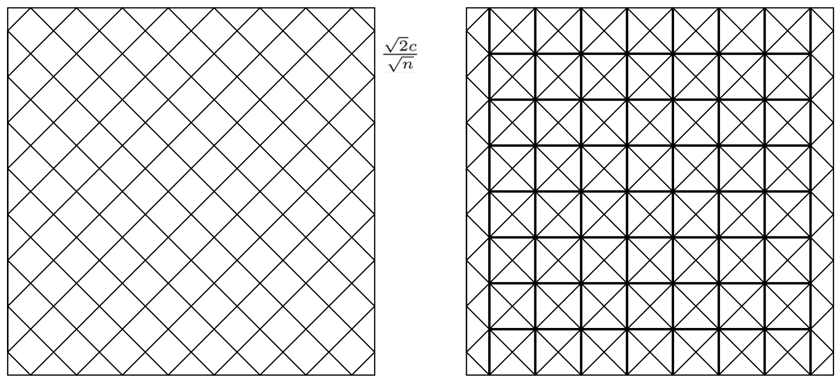

Fig. 7.1 Construction of the bond percolation model. A small square in the left figure is said to be *open* if it contains at least one node, meaning the corresponding edge of the square lattice in the right figure is open. (Based on Figure 2 of [9].)

A square is called *open* if it contains at least one node[2]; otherwise, it is called *closed*. Note that the small squares are open or closed independent of each other.

Now connect one diagonal of each such small square in the way depicted in Figure 7.1; thus a square lattice of size $\sqrt{2}c/\sqrt{n}$ emerges. The edges in the lattice can be open, with probability p, or closed with probability $1 - p$, independently from each other – the *bond percolation model* in percolation theory [10]. A path consisting of edges, is called an *open path* if all the edges are open. In this subsection we call such small horizontal squares h-squares for brevity.

We consider the number of open paths that go from the left side of the square lattice to the right side. Slice the unit square into horizontal rectangles of width $\frac{\sqrt{2}c}{\sqrt{n}} \log \frac{\sqrt{n}}{\sqrt{2}c}$; see Figure 7.2. Note there are $\log m \times m$ h-squares in each rectangle with $m := \frac{\sqrt{n}}{\sqrt{2}c}$. The following theorem shows that one can find many open paths going from left to right in each rectangle.

[2] "Open" and "closed" are used here according to the convention in percolation theory in the opposite sense that they are used in electrical circuits. "Open" means there exists a "channel" such that "fluid", in this case packets, can flow across; while "closed" means that packets cannot flow.

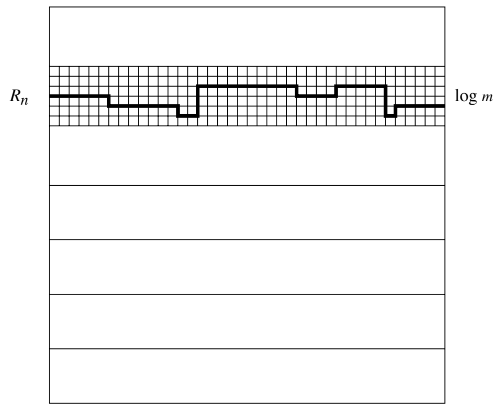

Fig. 7.2 The unit square is divided into $m/\log m$ horizontal rectangles of width $\log m/m$ each. Thus each rectangle contains $\log m \times m$ h-squares. (Based on Figure 4 of [9].)

Theorem 7.2. For c large enough, there exists a positive constant $\gamma = \gamma(c)$ such that there are $\gamma \log m = \gamma \log(\frac{\sqrt{n}}{\sqrt{2}c})$ disjoint open paths in each rectangle crossing from left to right, with probability approaching one as n goes to infinity.

To prove this result, the following results from percolation theory are needed (see Eqs (1.15), (6.41) and Theorem 2.45 in [10] for reference).

Lemma 7.3. Let Z_n be a square lattice with $n \times n$ h-squares and $0 < p \leq 1$. The probability $P_p(0 \leftrightarrow \partial Z_n)$ that there exists an open path from the center 0 of Z_n to its boundary ∂Z_n, is bounded by

$$P_p(0 \leftrightarrow \partial Z_n) \leq \frac{4}{3}(3p)^n.$$

Lemma 7.4. Let R_n be a rectangle embedded in the square lattice. Let $A_{n,1}$ be the event that there exists an open path between the left and right side of R_n, and $A_{n,b}$ the event that there are b edge-disjoint such crossings. We have

$$1 - P_p(A_{n,b}) \leq (\frac{p}{p-p'})^b \, (1 - P_{p'}(A_{n,1})),$$

for all $0 \leq p' < p \leq 1$.

Proof of Theorem 7.2: We focus on one particular rectangle R_n. Consider a new bond percolation model on R_n such that the "open" probability for each edge is p', a parameter to be determined later. Denote this model as R'_n, and define its "dual graph", Q'_n, of R'_n as follows. Put a vertex in the center of each h-square of R_n and connect the neighboring vertices; see Figure 7.3. An edge in Q'_n is *open* if and only if it crosses an open edge in S'_n. Now denote by $A_{n,1}$ the event that there is an *open* path in R'_n that crosses from left to right, and by $B_{n,1}$ the event that there is a *closed* path in Q'_n that crosses from bottom to top. We have $A_{n,1} \cap B_{n,1} = \phi$, because occurrence of both means a crossing between an open edge in R'_n and a closed edge in Q'_n, which is a contradiction to the construction of Q'_n. Furthermore, we have $P_{p'}(A_{n,1}) + P_{p'}(B_{n,1}) = 1$ because whenever $A_{n,1}$ does not occur $B_{n,1}$ occurs.

Index the bottom-line vertices of Q'_n by i, $1 \leq i \leq m$, and denote by $\bar{B}_{n,1}(i)$ the existence of an *open* path in Q'_n that crosses it from the bottom to the top and starts from vertex i. Since there are $\log m$ vertices in each column of Q'_n, we have for all q and i,

$$P_q(\bar{B}_{n,1}(i)) \leq P_q(0 \leftrightarrow \partial S_{\log m}).$$

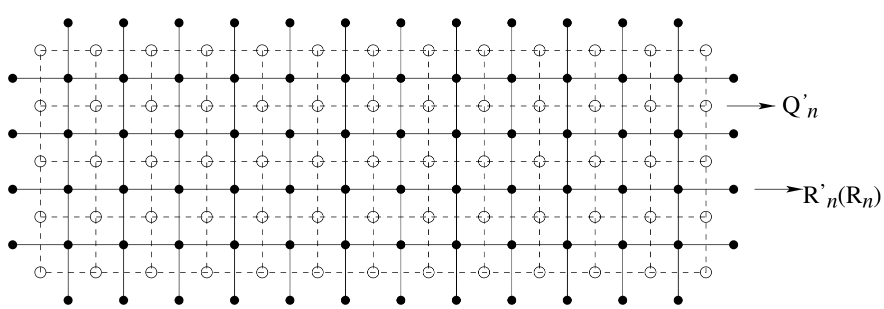

Fig. 7.3 A picture of R'_n, shown in solid lines, and its dual graph Q'_n, shown in dotted lines. (Based on Figure 5 of [9].) Note that no open path traversing from left to right in R'_n means the existence of a closed path in Q'_n from bottom to top.

Therefore, because edges in Q'_n and R'_n are closed with probability $1 - p'$,

$$\begin{aligned} P_{p'}(B_{n,1}) &\leq \sum_{i=1}^{m} P_{1-p'}(\bar{B}_{n,1}(i)) \\ &\leq \sum_{i=1}^{m} P_{1-p'}(0 \leftrightarrow \partial S_{\log m}) \\ &\leq \frac{4m}{3}(3(1-p'))^{\log m}, \end{aligned}$$

where the first inequality is a union bound, and the last inequality is by Lemma 7.3.

Now we apply Lemma 7.4 to lower bound the probability that there are $\gamma \log m$ disjoint paths in rectangle R_n crossing from left to right. Actually, we have

$$1 - P_p(A_{n,\gamma \log m}) \leq (\frac{p}{p-p'})^{\gamma \log m} P_{p'}(B_{n,1}), \tag{7.2}$$

for all $p' < p$. By choosing $p' = 2p - 1$, we get $p/(p-p') = e^{c^2} - 1 < e^{c^2}$ and $1 - p' = 2(1-p) = 2e^{-c^2}$. Substituting into (7.2) we get

$$\begin{aligned} 1 - P_p(A_{n,\gamma \log m}) &\leq (e^{c^2})^{\gamma \log m} P_{p'}(B_{n,1}) \\ &\leq m^{\gamma c^2} \frac{4m}{3} (6e^{-c^2})^{\log m} \\ &= \frac{4}{3} m^{(\gamma - 1)c^2 + \log 6 + 1}. \end{aligned}$$

Thus the probability of finding $\gamma \log m$ disjoint paths in a particular rectangle is bounded above as

$$P_p(A_{n,\gamma \log m}) \geq 1 - \frac{4}{3} m^{(\gamma-1)c^2 + \log 6 + 1}.$$

Since there are $m/\log m$ such rectangles and they are independent of each other, the probability of having $\gamma \log m$ disjoint paths in every rectangle is

$$\begin{aligned} (P_p(A_{n,\gamma \log m}))^{m/\log m} &\geq (1 - \frac{4}{3} m^{(\gamma-1)c^2 + \log 6 + 1})^{\frac{m}{\log m}} \\ &= \exp\left(\frac{m}{\log m} \log(1 - \frac{4}{3} m^{(\gamma-1)c^2 + \log 6 + 1})\right). \end{aligned}$$

Notice if $(\gamma - 1)c^2 + \log 6 + 1 \leq -1$, then the above expression goes to one as m goes to infinity. It is easy to check that $c^2 > \log 6 + 2$ and $\gamma \leq 1 - \frac{\log 6 + 2}{c^2}$ suffices. □

Thus for each of the $\frac{\sqrt{n}}{\sqrt{2}c} / \log \frac{\sqrt{n}}{\sqrt{2}c}$ horizontal rectangles we find $\gamma \log \frac{\sqrt{n}}{\sqrt{2}c}$ disjoint paths that cross the unit square from left to right. Similarly, we can find $\gamma \log \frac{\sqrt{n}}{\sqrt{2}c}$ disjoint vertical paths that cross the unit square from bottom to top, for each of the $\frac{\sqrt{n}}{\sqrt{2}c} / \log \frac{\sqrt{n}}{\sqrt{2}c}$ vertical rectangles. These paths form a *highway system* for the routing we will present below. Note that all this happens with probability going to one as n goes to infinity.

B. The Achievable Throughput on the Highways

Recall each edge along a highway path means the corresponding 45°-angled square is not empty. Pick one node as the designated sender and receiver; it is the intended receiver from the previous hop, and the sender for the next hop with an intended receiver within range $2\sqrt{2}c/\sqrt{n}$. Let every such designated node use a common power $P > 0$. For each hop along every such path to have a constant rate, we can use the idea of "non-interfering groups" from Sections 5 and 6. That is, we use time-division multiplexing among different but finite non-interfering groups. Suppose there are M^2 such groups as shown in Figure 7.4; then when one node is transmitting, the received signal

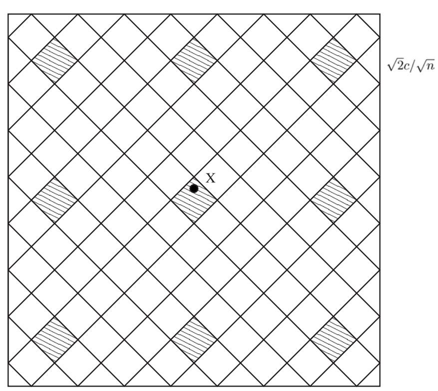

Fig. 7.4 A non-interfering group when $M = 3$, with X being an active transmitter. Also shown is the first layer of interfering 45°-angled squares to X.

power is at least $P/(2\sqrt{2}c/\sqrt{n})^\alpha$. Moreover, the interference experienced by an active node can be upper bounded as follows. Assume node X is such an active node; see Figure 7.4. The possible interfering 45°-angled squares are those in the same non-interfering group as X's. They are located on a series of squares centering at the 45°-angled square containing X; Figure 7.4 shows the first layer of interfering squares. Clearly there are (at most) 8 interfering 45°-angled squares on the first layer; the minimum distance between X and an interfering node within them is at least $(M-1)c/\sqrt{n}$. It is also easy to see that there are (at most) $8k$ interfering 45°-angled squares on the k-th layer, with distance at least $(kM-1)c/\sqrt{n}$ from X, for all $k>0$. Thus the received SINR at X satisfies the following:

$$
\begin{aligned}
SINR &= \frac{\frac{P_i}{|X_i - X_{R(i)}|^\alpha}}{N + \sum_{\substack{k\in\mathcal{T}\\ k\neq i}} \frac{P_k}{|X_k - X_{R(i)}|^\alpha}} \\
&\geq \frac{\frac{P}{(2\sqrt{2}s_n)^\alpha}}{N + \sum_{k=1}^{\infty} 8k \frac{P}{(kMs_n - s_n)^\alpha}} \\
&= \frac{\frac{P}{(2\sqrt{2})^\alpha}}{Ns_n^\alpha + \frac{8P}{M^\alpha}\sum_{k=1}^{\infty}\frac{k}{(k-\frac{1}{M})^\alpha}},
\end{aligned}
$$

where $s_n := \sqrt{2}c/\sqrt{n}$. Similar to the derivation in (6.2), it is easy to show there exists $c_\alpha > 0$ such that, whenever $M>1$ and $\alpha>2$, $\sum_{k=1}^{\infty}\frac{k}{(k-\frac{1}{M})^\alpha} < c_\alpha$. Thus,

$$
\begin{aligned}
SINR > \frac{\frac{P}{(2\sqrt{2})^\alpha}}{Ns_n^\alpha + \frac{8P}{M^\alpha}c_\alpha} &= \frac{\frac{1}{(2\sqrt{2})^\alpha}}{\frac{8}{M^\alpha}c_\alpha}(1+o(1)) \\
&= \frac{(M/2\sqrt{2})^\alpha}{8c_\alpha}(1+o(1)). \qquad (7.3)
\end{aligned}
$$

So in order to ensure $SINR \geq \beta$, one only needs $M = 2\sqrt{2}(8c_\alpha\beta)^{1/\alpha}$ for n large.

In summary, we have shown the following theorem.

Theorem 7.5. There exists a time-sharing scheme such that each hop on a highway path can transmit once in every M^2 time slots, thus achieving constant rate along each path, where M is a constant depending only on β and α.

7.4 Routing in three phases

Time slots can be divided into three parts for three phases of transmission: Draining phase, highway phase, and delivery phase, as designated in [9].

A. Draining phase

In this phase, data packets from a 45°-angled square are sent onto the highway system in one hop. We call each 45°-angled square a "source" square, and each such square is regarded as one unit for the draining phase. Data packets from a source square are sent to the nearest designated node of a horizontal path in the same horizontal rectangle. Recall that there are $\frac{\sqrt{n}}{\sqrt{2}c}/\log\frac{\sqrt{n}}{\sqrt{2}c}$ horizontal rectangles each with $\gamma\log\frac{\sqrt{n}}{\sqrt{2}c}$ disjoint paths. So the distance of the first hop of a packet is at most $\frac{c\sqrt{2}\log(\sqrt{n}/\sqrt{2}c)}{\sqrt{n}}$.

For the first hop to be successful, one needs to schedule the transmissions with intended range $\frac{c\sqrt{2}\log(\sqrt{n}/\sqrt{2}c)}{\sqrt{n}}$. This can be done in a similar way as we did for the highway system in the last subsection. According to Theorem 7.5, and the analysis for (7.3), we know each source square can achieve a rate of $\frac{W}{M^2}$ bits per second with $M = 2\sqrt{2}(8c_\alpha\beta)^{1/\alpha}\cdot\big(c\sqrt{2}\log(\sqrt{n}/\sqrt{2}c)\big)$.

It is easy to show by the union bound that there are at most $c^2\log n$ nodes in each 45°-angled square, with probability going to one as n goes to infinity. Thus with probability going to one, each node can achieve a rate of $\frac{W}{M^2c^2\log n} = \Theta(\frac{W}{(\log n)^3})$ bits per second for the draining phase.

B. Highway phase

Once a packet is in the highway system, it will be routed first horizontally then vertically towards its destination. The vertical highway path the packet switches to is the vertical path that is nearest to the destination node's 45°-angled square in the same vertical rectangle.

Now we calculate the load of each horizontal highway path. Actually, since each 45°-angled square contains no more than $c^2 \log n$ nodes, each horizontal rectangle contains no more than

$$N_1 := \frac{\frac{\sqrt{2}c}{\sqrt{n}} \log \frac{\sqrt{n}}{\sqrt{2}c}}{c^2/n} \cdot c^2 \log n = \sqrt{2}c \cdot \log \frac{\sqrt{n}}{\sqrt{2}c} \cdot \log n \cdot \sqrt{n} = \Theta(\sqrt{n})$$

nodes. So, supposing each node generates data at rate $\lambda(n)$, then each highway path needs to convey no more than $\lambda(n)N_1$ bits per second, which is of order $\Theta(1)$ if $\lambda = c_1/\sqrt{n}$. Since each highway path can accommodate $1/M^2$ rate, by selecting c_1 sufficiently small, no highway path will be overloaded.

C. Delivery phase

The delivery phase is exactly symmetric to the Draining phase.

Thus we have shown a scheme that achieves $\Theta(W/\sqrt{n})$ bits per second for every node. □

Remark 7.6. As shown in the scheme, a common power level can be employed at every node. To achieve $\Theta(W/\sqrt{n})$ bits per second for each node, a three stage transport scheme, with different transmission ranges in each stage, is used. We also note that, similarly to the discussion at the end of Section 4.2, $\Theta(W/\sqrt{n})$ bits per second is also achievable for networks under the Generalized Physical Model.

7.5 Notes

This section is based on [9].

8

An Information Theory for Transport Capacity: The High Attenuation Regime

In Sections 2–7, we have presented results addressing the performance bounds for wireless networks where the criterion for success of a transmission was modeled either by the Protocol Model, specifying a guard zone around a receiver or an interference footprint around a transmitter, or by the Physical Model, requiring a minimum SINR for decoding a signal. Once transmissions among nearby nodes are established, then nodes in the network can communicate with each other by routing packets in a multi-hopped fashion. All this is very much modeled on current technology.

Yet, wireless communication, especially in a network environment, due to its special nature, allows many more possibilities such as relaying by amplifying-and-forwarding rather than decoding-and-forwarding, multiple-access communication, and broadcast, just to name a few. Some of these strategies are motivated by their utility in some problems in multi-user information theory [6]. In fact the design space of strategies is infinite dimensional. Thus the question arises as to what would be the performance limits for wireless networks if any causal strategy is allowed. Is it possible that the scaling laws are different from those of the Protocol and Physical Models?

We address this fundamental question from this section onwards. We will consider scaling laws of transport capacity from an information theoretic perspective. As will be seen in the sequel since linear scaling of the transport capacity is easy to achieve, the challenging problem is to establish the upper bound on the transport capacity, which is the focus of this section.

In this section, we first study models of how wireless signals attenuate with distance. We suppose that a signal arriving at a distance r from the source is attenuated by a factor $Ge^{-\gamma r}/r^{\delta}$. Then we define the transport capacity from an information theoretic point of view. Finally the main results on the scaling behavior of transport capacity are given for networks that are spread out, i.e., there is a minimum separation distance between nodes.

The results to follow show that the scaling behavior of transport capacity fundamentally depends on how fast signal attenuates with distance, i.e., how large γ and δ are.

We consider in this section the case when $\gamma > 0$ or δ is large – which we call the *high attenuation regime*. One can show that then (i) the transport capacity is upper bounded by a multiple of the total power if the network is under a total power constraint; and (ii) the transport capacity scales linearly in n, the network size, if instead the network is under an individual power constraint. A construction to show linear growth in transport capacity is shown for networks on a grid.

The result in (i) shows that each bit-meter of transport requires a minimum positive energy expenditure, thus establishing a fundamental energy cost for information transport. One can draw an analogy with more classical results in single-link information theory, where the work of Shannon shows that there is a linear bound on the throughput in terms of the received signal power. The result above shows that there is a similar result for *wireless networks* in the high attenuation regime if one replaces received signal power by *total transmission power* used over the entire network, and simultaneous replaces throughput by the metric of transport capacity.

The linear scaling of the transport capacity result in (ii) when the pathloss exponent is large, agrees with the square-root scaling law $O(\sqrt{An})$ obtained earlier under the non-information theoretic approach. This is because at least $\Theta(n)$ square-meters are needed for a network of n nodes to satisfy the minimum positive distance separation constraint, and the transport capacity is $\Theta(\sqrt{An})$ for previous models. This implies that the multi-hop mode of operation is indeed an order-optimal architecture whenever it realizes this order, as it does in the best case. This provides a quantitative basis for the choice of the multi-hop mode of information transfer.

The assumption that any two nodes are separated by a minimum positive distance reflects physically realistic situations where nodes are separated by at least some fraction of a wavelength used for communication.

This section is based on [26]. Some recent improvements will be briefly discussed in the last subsection, Section 8.6.

The *low attenuation regime*, when $\gamma = 0$ and δ is small, will be discussed in the next section, where we show that the transport capacity can scale super-linearly in n, or may even not be bounded by the total power. In this case, other architectures for information transfer, involving say multi-user receivers can provide superior scaling laws in comparison to the multi-hop mode, as is shown by examples. However, the results require an unrealistically low path loss exponent, and further research is needed for realistic values.

8.1 Model and definitions

This subsection presents the network model, signal propagation model, and an information theoretic definition of transport capacity.

A. Wireless Network

Suppose there are n nodes denoted by $i \in \mathbb{N} := \{1, 2, \cdots, n\}$. The distance between two nodes i, j is denoted as d_{ij}. Nodes in the network are assumed to be separated by a distance of at least $d_{min} > 0$, i.e., $d_{ij} \geq d_{min}$ for all $i \neq j$. Note that this implies that as n increases the network domain must keep growing at least linearly in the number of nodes.

We suppose that transmissions happen in discrete time. At time instants $t = 1, 2, \cdots$, each node $i \in \mathbb{N}$ transmits a signal $X_i(t)$.

After attenuation due to distance, the received signal $Y_j(t)$ at node j is

$$Y_j(t) = \sum_{i \neq j} \frac{Ge^{-\gamma d_{ij}}}{d_{ij}^{\delta}} X_i(t) + Z_j(t),$$

where $\{Z_j(t)\}$ is iid noise with Gaussian distribution of zero mean and variance σ^2. $G > 0$ is a constant gain. The parameter $\delta > 0$ is called the *path loss exponent*[1], while $\gamma \geq 0$ is called the *absorption constant*. A positive γ generally prevails except for a vacuum environment [8].

Each node $i \in \mathbb{N}$ has a power P_i for transmission, meaning $\sum_{t=1}^{T} X_i^2(t) \leq TP_i$, where T is the transmission horizon. In this section we consider two constraints on power consumption:
Total Power Constraint: $\sum_i^n P_i \leq P_{total}$ that can be shared by all the nodes as they please; or
Individual Power Constraint: $P_i \leq P_{ind}$, $\forall i \in \mathbb{N}$.

We consider two kinds of networks:

- *Linear Networks*: n nodes are located on a straight line; and
- *Planar Networks*: n nodes are located on a plane.

The reason to consider line networks is because the analysis is easier to understand, and they can be used to model networks such as those formed by cars in a highway.

Two special networks of particular interests are the *regular linear network*, in which nodes are equally spaced apart on a straight line, and the *regular planar network*, in which nodes are located on a $\sqrt{n} \times \sqrt{n}$ integer lattice grid.

B. Definitions of Feasible Rates and Transport Capacity

For a wireless network of n nodes, we now define the notion of a *feasible information rate vector* $(R_{ij}, i, j \in N)$ as is standard in information theory [6]. Note that R_{ij} could be zero. For notation convenience we define $R_{ii} = 0$, $\forall i \in \mathbb{N}$.

[1] Compared with the power loss exponent α used in Sections 4, 6 and 7, a path loss exponent δ corresponds to $\alpha = 2\delta$.

Definition 8.1. *A* $(R_{ij}, i, j \in \mathbb{N}; T)$ *code with error probability* $P_e^{(T)}$ consists of the following.

- *Message Selection*: With W_{ij} denoting the message to be sent from node i to node j, we assume that all the messages $\{W_{ij}\}$ are independent, and uniformly distributed over their respective ranges $\{1, 2, \ldots, 2^{TR_{ij}}\}$.
- *Signaling Scheme*: The symbol $X_i(t)$, for $t \geq 1$, that node i sends out at time t, can depend in any causal manner on its own outgoing messages $W_i := (W_{ij}, j \in \mathbb{N})$, as well as the values of its past received symbols $Y_i^t := (Y_i(1), Y_i(2), \cdots, Y_i(t))$ for $t \geq 1$; define $Y_i^0 := \phi$. That is, we allow a set of encoding functions $f_{i,t}$ such that, for all $i \in \mathbb{N}, t \geq 1$, $X_i(t) = f_{i,t} = f_{i,t}(Y_i^{t-1}, W_i)$.
- *Decoders*: For each node j, there is a decoding function $\widehat{W}_{ij}(Y_j^T, W_j)$, $i \in \mathbb{N}$, by which it tries to recover the message intended for it from node i.
  The Error Probability is defined as $P_e^{(T)} := \text{Prob}((\hat{W}_{ij}, i, j \in \mathbb{N}) \neq (W_{ij}, i, j \in \mathbb{N}))$.

Definition 8.2. *Feasible rate vector with power constraint*: A rate vector $(R_{ij}, i, j \in \mathbb{N})$ is said to be *feasible with total power constraint* P_{total}, if there exists a sequence of $((R_{ij}, i, j \in \mathbb{N}), T)$ codes such that $1/T \sum_{t=1}^{T} \sum_{i=1}^{n} X_i(t)^2 \leq P_{total}$, *a.s.*, with $P_e^{(T)} \to 0$, as $T \to \infty$. If instead the power constraint is such that $1/T \sum_{t=1}^{T} X_i(t)^2 \leq P_{ind}$, $\forall i \in \mathbb{N}$, *a.s.*, as $T \to \infty$, the rate vector is said to be *feasible with individual power constraint* P_{ind}.

We now introduce the information theoretic definition of transport capacity.

Definition 8.3. An n-node network's transport capacity is defined as

$$C_T(n) := \sup_{(R_{ij}, i, j \in \mathbb{N}) \text{ feasible}} \sum_{ij} R_{ij} \cdot d_{ij},$$

where d_{ij} is the distance between nodes i and j.

8.2 Main results in high attenuation regime

We present results showing that, when $\gamma > 0$ or δ is large, then the transport capacity is upper bounded by a multiple of the total power. This implies that it scales linearly in n when the nodes are subject to individual power constraints.

First we consider planar networks.

A. Planar Networks

For the case of total power constraint, we have the following bound on transport capacity.

Theorem 8.4. If $\gamma > 0$ or $\delta > 3$, then for any planar network we have $C_T(n) \leq \frac{c_1(\gamma,\delta,d_{min})G^2}{\sigma^2} \cdot P_{total}$, where

$$c_1(\gamma,\delta,d_{min}) := \begin{cases} \dfrac{2^{2\delta+7}\log e}{\gamma^2 {d_{min}}^{2\delta+1}} \dfrac{e^{-\gamma d_{min}/2}(2 - e^{-\gamma d_{min}/2})}{(1 - e^{-\gamma d_{min}/2})} & \text{if } \gamma > 0, \\ \dfrac{2^{2\delta+5}(3\delta - 8)\log e}{(\delta-2)^2(\delta-3){d_{min}}^{2\delta-1}} & \text{if } \gamma = 0 \text{ and } \delta > 3. \end{cases} \tag{8.1}$$

It follows immediately from the above theorem that the transport capacity under the individual power constraint cannot grow faster than linear in n.

Theorem 8.5. If $\gamma > 0$ or $\delta > 3$, we have $C_T(n) \leq \frac{c_1(\gamma,\delta,d_{min})G^2 P_{ind}}{\sigma^2} \cdot n$ for any planar network.

Theorem 8.5 gives an upper bound for the transport capacity. On the other hand, one can show the linear growth is indeed achievable for regular networks.

Theorem 8.6. Suppose $\gamma > 0$ or $\delta > 1$, and each node is subject to an individual power constraint P_{ind}. Then a regular planar network of n nodes (distributed on a $\sqrt{n} \times \sqrt{n}$ grid) can achieve $C_T(n) \geq \mathcal{S}\left(\frac{e^{-2\gamma}G^2 P_{ind}}{c_3(\gamma,\delta)P_{ind}+\sigma^2}\right) \cdot n$, where $\mathcal{S}(\cdot)$ is the Shannon function

$\frac{1}{2}\log(1+x)$, and

$$c_3(\gamma,\delta) := \begin{cases} \dfrac{4(1+4\gamma)e^{-2\gamma} - 4e^{-4\gamma}}{2\gamma(1-e^{-2\gamma})} & \text{if } \gamma > 0, \\ \dfrac{16\delta^2 + (2\pi - 16)\delta - \pi}{(\delta-1)(2\delta-1)} & \text{if } \gamma = 0 \text{ and } \delta > 1. \end{cases} \tag{8.2}$$

B. Linear Networks

We have the following analogous results for linear networks.

Theorem 8.7. If $\gamma > 0$ or $\delta > 2$, and the network is subject to a total power constraint, then $C_T(n) \le \frac{c_2(\gamma,\delta,d_{min})G^2}{\sigma^2} \cdot P_{total}$, where

$$c_2(\gamma,\delta,d_{min}) := \begin{cases} \dfrac{2e^{-2\gamma d_{min}}\log e}{(1-e^{-\gamma d_{min}})^2(1-e^{-2\gamma d_{min}}){d_{min}}^{2\delta-1}} & \text{if } \gamma > 0, \\ \dfrac{2\delta(\delta^2-\delta-1)\log e}{(\delta-1)^2(\delta-2){d_{min}}^{2\delta-1}}, & \text{if } \gamma = 0 \text{ and } \delta > 2. \end{cases} \tag{8.3}$$

Theorem 8.8. If either $\gamma > 0$ or $\delta > 2$, and the network is subject to an individual power constraint, we have $C_T(n) \le \frac{c_2(\gamma,\delta,d_{min})G^2 P_{ind}}{\sigma^2} \cdot n$.

In the following subsections we will present the key ideas of the proofs for the results presented in this subsection. The interested reader is referred to [26] for complete proofs.

8.3 A max-flow-min-cut lemma relating rates with received power

This subsection presents a lemma that bounds the information flow from one set of nodes S to the rest of the network $\mathbb{N}\backslash S$ by a function of the received power in $\mathbb{N}\backslash S$ that is received from signals originating in S.

Lemma 8.9. [*Max-flow min-cut bound*] Suppose receptions in a wireless communication network are modeled by $Y_j(t) = \sum_{i \ne j} \alpha_{ij} X_i(t) + Z_j(t), j \in \mathbb{N}$, where: (i) $\{\alpha_{ij}\}$ is a sequence of *known* *deterministic* numbers. (ii) $\{Z_j(t)\}$ is an i.i.d. Gaussian noise process independent of the signal process $\{X_i(t)\}$, and $E|Z_j(t)|^2 = \sigma^2$.

If $\{((R_{ij}, i, j \in \mathbb{N}); T, P_e^{(T)}), T \geq 1\}$ is a sequence of codebooks, then for any subset S of $\mathbb{N}$, we have

$$R_{SD} \leq \frac{1}{T} + R_{SD} P_e^{(T)} + \frac{1}{T} \sum_{t=1}^{T} \sum_{j \in \mathbb{N} \setminus S} \frac{1}{2} \log(1 + \frac{E|\sum_{i \in S} \alpha_{ij} X_i(t)|^2}{\sigma^2}), \tag{8.4}$$

where $\mathbb{N} \setminus S$ denotes those nodes in $\mathbb{N}$ but not in S, $R_{SD} := \sum_{i \in S, j \in \mathbb{N} \setminus S} R_{ij}$, and $P_e^{(T)}$ is the probability of decoding error.

Proof. Let $D := \mathbb{N} \setminus S$ be the set of destination nodes, and let $W_{ij} := \{1, 2, \ldots, 2^{TR_{ij}}\}$ denote the message set from node i to j. We use the following notation:

$$\begin{aligned} U_j(t) &:= \sum_{i \in S} \alpha_{ij} X_i(t), \quad j \in D; \\ V_j(t) &:= U_j(t) + Z_j(t); \\ W_{SD} &:= \{W_{ij} : i \in S, j \in D\}; \\ W_D &:= \{W_{ij} : i \in D, j \in \mathbb{N}\}; \\ W_i &:= \{W_{ij} : j \in \mathbb{N}\}. \end{aligned}$$

Let $V_D(t) := \{V_j(t) : j \in D\}$, $V_D^t := \{V_D(k) : k = 1, \ldots, t\}$, and similarly for Y, U, Z.

First we want to show that $W_{SD} \to \{V_D^T, W_D\} \to \{Y_D^T, W_D\}$ forms a Markov chain. This can be done by showing that Y_D^T is a deterministic function of (V_D^T, W_D). Actually, this can be seen by noticing that for any $j \in D$, $2 \leq t \leq T$,

$$\begin{aligned} Y_j(t) &= V_j(t) + \sum_{i \in D, i \neq j} \alpha_{ij} X_i(t) \\ &= V_j(t) + \sum_{i \in D, i \neq j} \alpha_{ij} f_{i,t}(Y_i^{t-1}, W_i), \quad \text{and} \\ Y_j(1) &= V_j(1) + \sum_{i \in D, i \neq j} \alpha_{ij} f_{i,1}(W_i). \end{aligned}$$

Now by Fano's Lemma and the property of a Markov chain, we have

$$H(W_{SD}|V_D^T, W_D) \le H(W_{SD}|Y_D^T, W_D) \le 1 + TR_{SD}P_e^{(T)}.$$

Thus,

$$\begin{aligned}
&TR_{SD}\\
&\quad = H(W_{SD}) = I(W_{SD}; V_D^T, W_D) + H(W_{SD}|V_D^T, W_D)\\
&\quad \le I(W_{SD}; V_D^T, W_D) + 1 + TR_{SD}P_e^{(T)}\\
&\quad = I(W_{SD}; W_D) + I(W_{SD}; V_D^T|W_D) + 1 + TR_{SD}P_e^{(T)}\\
&\quad = 0 + h(V_D^T|W_D) - h(V_D^T|W_{SD}, W_D) + 1 + TR_{SD}P_e^{(T)}\\
&\quad \le h(V_D^T) - h(V_D^T|W_{SD}, W_D) + 1 + TR_{SD}P_e^{(T)},
\end{aligned}$$

with

$$\begin{aligned}
&h(V_D^T|W_{SD}, W_D)\\
&\quad = \sum_{t=1}^{T} h(V_D(t)|V_D(1), \ldots, V_D(t-1), W_{SD}, W_D)\\
&\quad \ge \sum_{t=1}^{T} h(V_D(t)|V_D(1), \ldots, V_D(t-1), \Gamma_S(t), W_{SD}, W_D)\\
&\quad = \sum_{t=1}^{T} h(V_D(t)|\Gamma_S(t)) \ge \sum_{t=1}^{T} h(V_D(t)|U_D(t)).
\end{aligned}$$

Hence,

$$\begin{aligned}
&TR_{SD}\\
&\le h(V_D^T) - \sum_{t=1}^{T} h(V_D(t)|U_D(t)) + 1 + TR_{SD}P_e^{(T)}\\
&= h(V_D^T) - \sum_{t=1}^{T}\sum_{j\in D} h(Z_j(t)) + 1 + TR_{SD}P_e^{(T)}\\
&\le \sum_{t=1}^{T}\sum_{j\in D} h(V_j(t)) - \sum_{t=1}^{T}\sum_{j\in D} \frac{1}{2}\log(\pi e\sigma^2) + 1 + TR_{SD}P_e^{(T)}
\end{aligned}$$

$$\leq \sum_{t=1}^{T} \sum_{j \in D} (\frac{1}{2} \log(\pi e(E|U_j(t)|^2 + \sigma^2)) - \frac{1}{2} \log(\pi e \sigma^2)) + 1 + TR_{SD} P_e^{(T)}$$

$$= \sum_{t=1}^{T} \sum_{j \in D} \frac{1}{2} \log(1 + \frac{E|U_j(t)|^2}{\sigma^2}) + 1 + TR_{SD} P_e^{(T)},$$

where the last inequality comes from the fact that Gaussian distribution maximizes entropy for given covariance; see Lemma 2 in [23]. □

Given that the above lemma and noticing $\log(1+x) \leq x \log e$, $\forall x > 0$, the following corollary is immediate.

Corollary 8.10. If $(R_{ij}, i, j \in \mathbb{N})$ is a feasible rate vector, then for any subset $S \in \mathbb{N}$,

$$R_{SD} = \sum_{i \in S, j \in D} R_{ij} \leq \frac{\log e}{2\sigma^2} \cdot \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i \in S} E \left(\sum_{j \in D} \alpha_{ij} X_j(t) \right)^2,$$

where $D := \mathbb{N} \backslash S$.

8.4 Upper bounds for linear and planar networks

In this subsection we present the proof of Theorem 8.7, i.e., that the transport capacity is bounded by a multiple of the total power for linear networks, whenever $\gamma > 0$ or $\delta > 2$. Since the idea to prove Theorem 8.5, the planar case, proceeds similarly, we refer the reader to [26] for it. Moreover, we only show the proof for the case when $\gamma = 0$. For $\gamma > 0$, the proof is similar.

Proof of Theorem 8.7: Let $a_i d_{min}$ denote the coordinate of the node i, and let R_S denote the sum of the rates of the source destination pairs with source nodes in set S, and ending in the rest of the network, $\mathbb{N} \backslash S$. Applying Corollary 8.10 to the subsets:

$$\begin{aligned} \mathcal{N}_q^- &= \{i \in \mathbb{N} : a_i \leq q\}, \text{ and} \\ \mathcal{N}_q^+ &= \{i \in \mathbb{N} : a_i > q\}, \forall q \in \mathbb{Z} := \{\text{integer}\}, \end{aligned} \tag{8.5}$$

we get for any q,

$$\frac{2\sigma^2}{\log e} \cdot R_{\mathcal{N}_q^-} \le \text{liminf}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i\in\mathcal{N}_q^-} E\left(\sum_{j\notin\mathcal{N}_q^-} \frac{X_j(t)}{(a_j - a_i)^\delta d_{min}^\delta} \right)^2 \quad (8.6)$$

$$\frac{2\sigma^2}{\log e} \cdot R_{\mathcal{N}_q^+} \le \text{liminf}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i\in\mathcal{N}_q^+} E\left(\sum_{j\notin\mathcal{N}_q^+} \frac{X_j(t)}{(a_i - a_j)^\delta d_{min}^\delta} \right)^2 \quad (8.7)$$

Suppose a rate R_{ij} is supported over a distance d_{ij} towards the $+\infty$ direction. Consider the sequence $\{R_{\mathcal{N}_q^-}\}$, for q from $-\infty$ to $+\infty$. By definition, the rate R_{ij} must be counted at least $\lfloor d_{ij}/d_{min} \rfloor$ times among the sequence. By symmetry, a left travelling rate must be counted in $\{R_{\mathcal{N}_q^+}\}$ similarly. Thus we get

$$\begin{aligned}\sum_{i,j\in\mathbb{N}} R_{ij} \cdot d_{ij} &\le 2d_{min} \sum_{i,j\in\mathbb{N}} R_{ij} \cdot \lfloor d_{ij}/d_{min} \rfloor \\ &\le 2d_{min} \sum_{q=-\infty}^{+\infty} R_{\mathcal{N}_q^-} + 2d_{min} \sum_{q=-\infty}^{+\infty} R_{\mathcal{N}_q^+}. \end{aligned} \quad (8.8)$$

Now we show that

$$\sum_{q=-\infty}^{+\infty} R_{\mathcal{N}_q^-} \le \frac{c_2(\gamma, \delta, d_{min})}{4 d_{min} \sigma^2} P_{total}, \quad (8.9)$$

and by (8.6), we will then only need to show that

$$\begin{aligned}&\tfrac{1}{T} \textstyle\sum_{t=1}^{T} \sum_{q=-\infty}^{+\infty} \sum_{i\in\mathcal{N}_q^-} E\left(\sum_{j\notin\mathcal{N}_q^-} \frac{X_j(t)}{(a_j - a_i)^\delta} \right)^2 \\ &\qquad \le \frac{c_2(\gamma, \delta, d_{min}) d_{min}{}^{2\delta-1}}{2\log e} P_{total}. \end{aligned} \quad (8.10)$$

The intuition is that the received power is no more than the transmitted power.

Since $X_j(t)$ satisfies

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{j\in\mathcal{N}} X_j^2(t) \le P_{total}, \quad \text{a.s.,} \quad (8.11)$$

we only need to prove that for any t,

$$\sum_{q=-\infty}^{+\infty} \sum_{i\in\mathcal{N}_q^-} \left(\sum_{j\notin\mathcal{N}_q^-} \frac{X_j(t)}{(a_j - a_i)^\delta}\right)^2 \leq \frac{c_2(\gamma,\delta,d_{min}) {d_{min}}^{2\delta-1}}{2\log e} \sum_{i\in\mathcal{N}} X_i^2(t). \tag{8.12}$$

First, we observe that the left hand side (L.H.S.) of (8.12) is a summation of infinite terms of the form $\beta_{jk}X_j(t)X_k(t)$. If every $X_j(t)X_k(t)$ is replaced with the larger value $\frac{1}{2}(X_j^2(t) + X_k^2(t))$, it is easy to see that

$$\text{L.H.S. of (8.12)} \leq \sum_{k\in\mathcal{N}} \left(\sum_{q=-\infty}^{\lceil a_k\rceil-1} \sum_{i\in\mathcal{N}_q^-} \sum_{j\notin\mathcal{N}_q^-} \frac{1}{(a_j - a_i)^\delta (a_k - a_i)^\delta} \right) X_k^2(t).$$

This would imply (8.12) if for any $k \in \mathcal{N}$,

$$\sum_{q=-\infty}^{\lceil a_k\rceil-1} \sum_{i\in\mathcal{N}_q^-} \sum_{j\notin\mathcal{N}_q^-} \frac{1}{(a_j - a_i)^\delta (a_k - a_i)^\delta} \leq \frac{c_2(\gamma,\delta,d_{min}) {d_{min}}^{2\delta-1}}{2\log e}. \tag{8.13}$$

This can be established as follows.

Let $\underline{a}_q := \min_{j\notin\mathcal{N}_q^-} a_j$, and note that $\min_{i\neq j} |a_i - a_j| \geq 1$. Then we have

$$\begin{aligned}
\text{L.H.S. of (8.13)} &\leq \sum_{q=-\infty}^{\lceil a_k\rceil-1} \sum_{i\in\mathcal{N}_q^-} \sum_{l=0}^{\infty} \frac{1}{(l + \underline{a}_q - a_i)^\delta} \frac{1}{(a_k - a_i)^\delta} \\
&\leq \sum_{q=-\infty}^{\lceil a_k\rceil-1} \sum_{i\in\mathcal{N}_q^-} \frac{\delta - 1 + \underline{a}_q - a_i}{(\delta-1)(\underline{a}_q - a_i)^\delta} \frac{1}{(a_k - a_i)^\delta} \\
&= \sum_{\{i:\, a_i < a_k\}} \sum_{q=\lceil a_i\rceil}^{\lceil a_k\rceil-1} \left[\frac{1}{(\underline{a}_q - a_i)^\delta} + \frac{1}{(\delta-1)(\underline{a}_q - a_i)^{\delta-1}} \right]
\end{aligned}$$

$$
\begin{aligned}
&\times \frac{1}{(a_k - a_i)^\delta} \\
&\le \sum_{\{i:\, a_i < a_k\}} \left[\sum_{l=1}^{\infty} \frac{1}{l^\delta} + \frac{1}{\delta - 1} \sum_{l=1}^{\infty} \frac{1}{l^{\delta-1}} \right] \frac{1}{(a_k - a_i)^\delta} \\
&\le \frac{\delta^3 - \delta^2 - \delta}{(\delta - 1)^2(\delta - 2)} = \frac{c_2(\gamma, \delta, d_{min}) {d_{min}}^{2\delta - 1}}{2 \log e}, \qquad (8.14)
\end{aligned}
$$

where we have used the fact that for any real $a > 0$ and $\beta > 1$,

$$
\sum_{l=0}^{+\infty} \frac{1}{(l+a)^\beta} \le \frac{1}{a^\beta} + \int_0^\infty \frac{1}{(a+x)^\beta} dx \le \frac{\beta - 1 + a}{(\beta - 1)a^\beta}. \qquad (8.15)
$$

Hence (8.12) is established, and thus (8.9) follows. Similarly, we have

$$
\sum_{q=-\infty}^{+\infty} R_{\mathcal{N}_q^+} \le \frac{c_2(\gamma, \delta, d_{min})}{4 d_{min} \sigma^2} P_{total}. \qquad (8.16)
$$

Finally, combining (8.8), (8.9) and (8.16), the proof for Theorem 8.7 is completed. □

8.5 Achieving linear growth using regular networks

In this subsection, we show that when $\delta > 1$, a regular network under individual power constraint can achieve $\Theta(n)$ transport capacity. This will prove Theorem 8.6. We only consider the case where $\gamma = 0$ and $\delta > 1$.

Consider a regular planar network with n nodes such that two neighboring nodes are one meter apart from each other. Each node $i \in \mathbb{N}$ is a source, and it selects one of its 4 nearest neighbors as its only destination node.

Now each node independently generates its codebook according to a Gaussian distribution with variance $P = P_{ind} - \epsilon$, where $\epsilon > 0$.

After the block of T transmissions, every destination node decodes its source's signal, treating all the other transmissions as Gaussian noise. Thus, if we can show $\forall j \in \mathbb{N}$,

$$
c_3(\gamma, \delta) P \ge \sum_{i \in \mathbb{N}, i \ne j, i \text{ is not } j\text{'s source}} \frac{P}{d_{ij}^{2\delta}}, \qquad (8.17)
$$

i.e., $c_3(\gamma,\delta)P$ is an upper bound for interference, then any rate R_0 satisfying the following is achievable for every source–destination pair:

$$R_0 < \frac{1}{2}\log\left(1+\frac{P}{c_3(\gamma,\delta)P+\sigma^2}\right).$$

This bound can be shown as follows.

$$\begin{aligned}
&\sum_{i\in\mathbb{N},i\neq j,i\text{ is not }j\text{'s source}} \frac{P}{d_{ij}^{2\delta}} \\
&\leq 4\times\left(2\sum_{i=1}^{\infty}\frac{1}{i^{2\delta}}+\int_1^{\infty}\int_0^{\infty}\frac{1}{(x^2+y^2)^{\delta}}dxdy\right)P \\
&\leq 4\times\left(2\cdot\frac{2\delta}{2\delta-1}+\frac{\pi}{4\delta-4}\right)P \\
&\leq \frac{16\delta^2+(2\pi-16)\delta-\pi}{(\delta-1)(2\delta-1)}P \\
&\leq c_3(\gamma,\delta)P.
\end{aligned}$$

8.6 Notes

This section is based on [26]. In [27] it has recently been shown that, for the transport capacity to be upper bounded by a multiple of the total power, and therefore to be linear in the network size under individual power constraint, $\delta > 3/2$ suffices for linear networks, while $\delta > 5/2$ suffices for planar networks. The effect of random phases of channels is also discussed in the same paper. The linear growth rate of transport capacity for an improved bound on δ, and a *generalized* transport capacity are also examined in [4].

9

An Information Theory for Transport Capacity: The Low Attenuation Regime

This section continues with the treatment of wireless network information theory from the last section. We now address the scaling behavior of transport capacity when the attenuation is slow – *the low attenuation regime*.

In this low attenuation regime, by a *coherent relaying and interference subtraction* (CRIS) strategy, super-linear growth in transport capacity can be achieved for networks under the individual power constraint. The ratio of transport capacity to total power can also be unbounded. This result shows that there can be a fundamental connection between the attenuation properties of the medium and the amount of information that can be transported. In particular the architecture for information transport in wireless networks with very low attenuation can be quite different from what is adequate when the attenuation is high. At present however, the results presented below are only for unrealistically low attenuations, and further research is need to bridge the gap with the high attenuation case.

The same model and definitions as in last section are adopted here. In particular, we still consider both linear and planar networks with a minimum separation distance d_{min} among nodes. The signal is assumed to attenuate with distance by a factor of G/r^δ.

We first present the main results. Then we present the coherent relaying and interference subtraction (CRIS) scheme. The application of this scheme to achieve super-linear growth in transport capacity will be shown in Section 9.3. We present the major steps of the proofs; for complete proofs we refer the reader to [26].

9.1 Main results in low attenuation regime

When there is no absorption ($\gamma = 0$) and δ is small, one finds results in great contrast to the high attenuation regime. In some cases, the transport capacity cannot be upper bounded by a finite multiple of the total power. In some other cases, the transport capacity even scale super-linearly in n under the individual power constraint.

These results are attained by a strategy of coherent relaying with interference subtraction (CRIS). For a source–destination pair (s,d), the n nodes are divided into a sequence of groups, $G_0 \to G_1 \to \cdots \to G_m$, such that $G_0 = \{s\}$ and $G_m = \{d\}$. Groups with higher numbers will be said to be more "downstream," though they are not necessarily closer to the destination. Nodes in group l will dedicate a portion of their powers to coherently transmit for the benefit of nodes in their downstream nodes. During decoding, each node first subtracts from its received signal the already known portion of the superposed component coming from downstream nodes, and then decodes the new information.

Theorem 9.1. Suppose $\gamma = 0$, i.e., there is no absorption. Then, for planar networks under the total power constraint P_{total}:

i) If $\delta < 3/2$, then any arbitrarily large transport capacity can be supported in a regular planar network with large enough n, using the CRIS strategy.
ii) If $\delta < 1$, then CRIS can support a fixed rate $R_{\min} > 0$ for any single source–destination pair in any regular planar network, irrespective of the distance between them.

The following is the corresponding result for regular linear networks.

Theorem 9.2. When $\gamma = 0$, for linear networks under the total power constraint P_{total}:

i) If $\delta < 1$, then any arbitrarily large transport capacity can be supported in a regular linear network with large enough n using the CRIS strategy.
ii) If $\delta < 1/2$, then CRIS can support a fixed rate $R_{\min} > 0$ for any single source–destination pair in any regular planar network, irrespective of the distance between them.

For the scaling law of transport capacity under the individual power constraint we have:

Theorem 9.3. When $\gamma = 0, 1/2 < \delta < 1$, and each node is subject to a individual power constraint, a super-linear $\Theta(n^{\theta})$ bit-meters/second scaling of the transport capacity is feasible for any $\theta \in (1, 1/\delta)$, for some linear networks.

9.1.1 Achievable rate by CRIS

The results in the low attenuation regime are based on the idea of coherent multistage relaying with interference subtraction. This strategy is of independent interest and we will present in Section 9.2. The achieved rate is described in the following setup.

Theorem 9.4. Consider the following *Gaussian Relay Network*: There are $M + 1$ nodes $\{0, 1, \cdots, M\}$ with node 0 as the source node, and node M as its end destination node. Nodes $\{1, 2, \cdots, M - 1\}$ are there to help the source node 0 to send information to the destination node M. Node 0 can only send, and node M can only listen. At each node j, the received signal is $Y_j(t) = \sum_{i \neq j} \alpha_{ij} X_i(t) + N_j(t)$, with $N_j(t)$ being an independent Normal$(0, \sigma^2)$ random variable. Then any rate satisfying the following is achievable by CRIS from node 0 to M:

$$R < \min_{1 \leq j \leq M} \mathcal{S}\left(\frac{1}{\sigma^2} \sum_{k=1}^{j} \left(\sum_{i=0}^{k-1} \alpha_{ij} \sqrt{P_{ik}}\right)^2\right), \tag{9.1}$$

where $\{P_{ik}\}$ is any allocation of powers that satisfies $P_{ik} \geq 0$ and $\sum_{k=i+1}^{M} P_{ik} \leq P_i$.

A simple generalization using groups of nodes to function as a coherent relay gives the following result.

Theorem 9.5. Consider the following generalized Gaussian multiple relay network. There are $M+1$ *groups* of nodes sequentially denoted by $\mathcal{N}_0, \mathcal{N}_1, \ldots, \mathcal{N}_M$ with $\mathcal{N}_0 = \{0\}$ as the source, $\mathcal{N}_M = \{M\}$ as the destination, and the other $M-1$ groups as $M-1$ stages of relay. Suppose there are n_i nodes in Group $\mathcal{N}_i$, $i \in \{0,1,\ldots,M\}$, and the power constraint for each node in $\mathcal{N}_i$ is $P_i/n_i \geq 0$. Then any rate R satisfying the following is achievable from node 0 to M:

$$R < \min_{1 \leq j \leq M} S\left(\frac{1}{\sigma^2}\sum_{k=1}^{j}\left(\sum_{i=0}^{k-1}\alpha_{\mathcal{N}_i\mathcal{N}_j}\sqrt{P_{ik}/n_i}\cdot n_i\right)^2\right), \tag{9.2}$$

where $P_{ik} \geq 0$ satisfies $\sum_{k=i+1}^{M} P_{ik} \leq P_i$, and $\alpha_{\mathcal{N}_i\mathcal{N}_j} := \min\{\alpha_{k\ell} : k \in \mathcal{N}_i, \ell \in \mathcal{N}_j\}$, $i,j \in \{0,1,\ldots,M\}$.

9.2 Method of coherent relaying with interference subtraction

In this subsection we present the strategy used in achieving the single source–destination rate in Theorem 9.4. Let us denote by $\mathcal{S}(x) := 1/2\log(1+x)$, Shannon's function.

We consider $M+1$ nodes with source node 0, relay nodes $\{1,2,\cdots,M-1\}$, and destination node M, each with a power budget P_i, $0 \leq i \leq M-1$. The source node can only transmit, and the destination node can only listen. The received signal at node j at time t is $Y_j(t) = \sum_{i \neq j}\alpha_{ij}X_i(t) + Z_j(t)$, with $\{Z_j(t)\}$ being a sequence of iid, zero mean, Gaussian noise with variance $\sigma^2 > 0$.

The basic idea behind CRIS is that each node divides its energy into several parts used for helping different downstream nodes in forwarding messages to M. Different nodes helping the same downstream node will combine their energy coherently. A node while decoding can subtract from its received signal the part emitted by downstream nodes, since it has already successfully decoded any messages that they know. (Note that the channel is assumed to be perfectly known.)

First we illustrate the strategy for a network of three nodes as in Figure 9.1. Then we will formally present the general strategy. The interested reader is referred to [26] for the proofs.

Consider the network in Figure 9.1 consisting of source s, relay r and destination d. Let α_{sr}, α_{sd} and α_{rd} denote the corresponding signal attenuation factors. The whole transmission time is divided equally into blocks of the same size, say T_0. In each block (except the first and the last), node s divides its power P_s into two parts, θP_s and $(1-\theta)P_s$ with $0 \leq \theta \leq 1$. These are used for different purposes: (i) The part θP_s is used to "inform" relay r so that it can help coherently in the next block. By Shannon's formula any rate R satisfying

$$R < \mathcal{S}\left(\frac{\alpha_{sr}^2 \theta P_s}{\sigma^2}\right), \tag{9.3}$$

is achievable for this task. (ii) The part $(1-\theta)P_s$ is used to "collaborate" with relay r, which will transmit coherently with s to send a signal to receiver d, using its full power P_r. Note that the collaboration information has already arrived at r at the end of the previous block.

At the end of this block, what node d receives is the addition of three components: (a) the signal due to the coherent cooperation between s and r with power $\left(\alpha_{sd}\sqrt{(1-\theta)P_s} + \alpha_{rd}\sqrt{P_r}\right)^2$; (b) the "bonus" signal sent by s intended mainly for r for preparing the next-block cooperation, with power $\alpha_{sd}^2 \theta P_s$; and (c) the noise with power σ^2.

Now, the decoding procedure for node d is as follows. At the end of each block, it decodes based on signals from both this block and the previous one. The information bearing parts for this decoding are: the

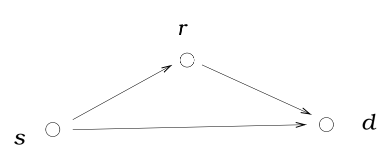

Fig. 9.1 The single relay channel.

first part (a) of this block, and the second part (b) of the previous block. Note that they both represent the same information. Note also that the first part in the previous block can be removed before decoding, since it becomes known after the decoding at the end of the previous block. The following rate is thus achievable:

$$
\begin{aligned}
R &< \mathcal{S}\left(\frac{\left(\alpha_{sd}\sqrt{(1-\theta)P_s}+\alpha_{rd}\sqrt{P_r}\right)^2}{\alpha_{sd}^2\theta P_s+\sigma^2}\right)+\mathcal{S}\left(\frac{\alpha_{sd}^2\theta P_s}{\sigma^2}\right) \\
&= \mathcal{S}\left(\frac{\alpha_{sd}^2 P_s+\alpha_{rd}^2 P_r+2\alpha_{sd}\alpha_{rd}\sqrt{(1-\theta)P_sP_r}}{\sigma^2}\right).
\end{aligned}
$$

Together with the constraint (9.3), this leads to the following end-to-end achievable rate

$$
\begin{aligned}
R < \max_{0\le\theta\le1}\min\Bigg\{&\mathcal{S}\left(\frac{\alpha_{sr}^2\theta P_s}{\sigma^2}\right), \\
&\mathcal{S}\left(\frac{\alpha_{sd}^2 P_s+\alpha_{rd}^2 P_r+2\alpha_{sd}\alpha_{rd}\sqrt{(1-\theta)P_sP_r}}{\sigma^2}\right)\Bigg\}. \qquad (9.4)
\end{aligned}
$$

The scheme here is different from the scheme in [7]. Here messages are not partitioned into cells. For decoding, the destination waits until it has received all the relevant signals, and it then determines the message directly once and for all.

After this illustration of the basic idea for a three-node relay network, we now present the scheme for a network of $M+1$ nodes.

The CRIS Scheme

The transmission consists B blocks, each of T slots. A sequence of $B-M+1$ indices, $w_b \in \{1,\ldots,2^{TR}\}, b=1,2,\ldots,B-M+1$ will be sent over in a total of TB slots. (Note that as $B\to\infty$, the rate $TR(B-M+1)/TB$ is arbitrarily close to R for any T.)

9.2.1 Generation of codebooks

Randomly generate M^2 matrices $\mathcal{X}_k(b_0)$, for $k=1,\ldots,M$, and $b_0=1,\ldots,M$, each of size $2^{TR}\times T$, with every element independently chosen with Gaussian distribution $N(0,1-\varepsilon_1)$. These are the codebooks.

In the same block, different nodes use independent codewords from different codebooks. Since it takes M blocks to transmit one complete message, every node is assigned M independent codebooks. The total number of codebooks needed is M^2 since there are M nodes. The M^2 matrices are revealed to all the $M+1$ nodes. Let $\mathcal{X}_k(b) := \mathcal{X}_k(b \bmod M), b = 1, 2, \ldots, B$. Denote by $x_k(b, w)$ the w-th row of the matrix $\mathcal{X}_k(b)$, for $w \in \{1, \ldots, 2^{TR}\}$. It denotes the w-th codeword.

9.2.2 Encoding

At the beginning of each block $b \in \{1, \ldots, B\}$, every node $i \in \{0, \ldots, M-1\}$ has estimates (see the sequel) $\widehat{w}_{b-k+1,i}$ of w_{b-k+1}, $k \geq i+1$ (with $\widehat{w}_{b-k+1,0} = w_{b-k+1}$) and sends the following vector of length T in the block:

$$\vec{X}_i(b) := \sum_{k=i+1}^{M} \sqrt{P_{ik}} x_k(b, \widehat{w}_{b-k+1,i}).$$

We set

$$\widehat{w}_{b_1,i} := w_{b_1} := 0 \text{ for any } b_1 \leq 0, \text{ and } x_k(b, 0) := 0. \tag{9.5}$$

Every node $k \in \{1, \ldots, M\}$ thus receives the vector:

$$\begin{aligned} \vec{Y}_k(b) &= \sum_{\substack{0 \leq i \leq M-1 \\ i \neq k}} \alpha_{ik} \vec{X}_i(b) + \vec{Z}_k(b) \\ &= \sum_{\substack{0 \leq i \leq M-1 \\ i \neq k}} \sum_{l=i+1}^{M} \alpha_{ik} \sqrt{P_{il}} x_l(b, \widehat{w}_{b-l+1,i}) + \vec{Z}_k(b) \\ &= \left(\sum_{l=1}^{k} \sum_{i=0}^{l-1} + \sum_{l=k+1}^{M} \sum_{\substack{0 \leq i \leq l-1 \\ i \neq k}} \right) \alpha_{ik} \sqrt{P_{il}} x_l(b, \widehat{w}_{b-l+1,i}) + \vec{Z}_k(b). \end{aligned}$$

Let

$$\widehat{\vec{Y}}_k(b) := \vec{Y}_k(b) - \sum_{l=k+1}^{M} \sum_{\substack{0 \leq i \leq l-1 \\ i \neq k}} \alpha_{ik} \sqrt{P_{il}}\, x_l(b, \widehat{w}_{b-l+1,k}) \tag{9.6}$$

serve as an estimate by node k of $\sum_{l=1}^{k} \sum_{i=0}^{l-1} \alpha_{ik} \sqrt{P_{il}} x_l(b, \widehat{w}_{b-l+1,i})$.

9.2.3 Decoding

We use joint-typical set decoding. The following definitions are standard.

Definition 9.6. The set $A_\epsilon^{(T)}$ of jointly typical sequences $\{(x^T, y^T)\}$ with respect to the joint density function $f(x,y)$ is the set of T-sequences with empirical entropies ϵ-close to the true entropies, i.e.,

$$\begin{aligned} A_\epsilon^{(T)} = \Big\{ & (x^T, y^T) \in \mathbb{R}^T \times \mathbb{R}^T : \\ & \left| -\frac{1}{T} \log f(x^T) - h(X) \right| < \epsilon, \\ & \left| -\frac{1}{T} \log f(y^T) - h(Y) \right| < \epsilon, \\ & \left| -\frac{1}{T} \log f(x^T, y^T) - h(X,Y) \right| < \epsilon \Big\}, \end{aligned}$$

where $f(x^T, y^T) = \prod_{i=1}^{T} f(x_i, y_i)$.

Definition 9.7. $A_\epsilon^{(T)}(P,N)$ denotes the set $A_\epsilon^{(T)}$ with respect to the joint density function

$$f(x,y) = g_P(x) g_N(y-x) = \tfrac{1}{\sqrt{2\pi P}} \exp\left(-\tfrac{x^2}{2P}\right) \cdot \tfrac{1}{\sqrt{2\pi N}} \exp\left(-\tfrac{(y-x)^2}{2N}\right).$$

Given these definitions, we now introduce the decoding procedure.

At the *end* of each block $b \in \{1, \ldots, B\}$, every node $k \in \{1, \ldots, M\}$ (for $b - k + 1 \geq 1$) declares $\widehat{w}_{b-k+1,k} = w$, if w is the unique value in $\{1, \ldots, 2^{TR}\}$ such that in *all* the blocks $b - j, j = 0, 1, \ldots, k-1$:

$$\begin{aligned} & \Bigg\{ \sum_{i=0}^{k-j-1} \alpha_{ik} \sqrt{P_{i,k-j}}\, x_{k-j}(b-j, w),\ \widehat{\vec{Y}}_k(b-j) \\ & \quad - \sum_{l=k-j+1}^{k} \sum_{i=0}^{l-1} \alpha_{ik} \sqrt{P_{il}}\, x_l(b-j, \widehat{w}_{b-j-l+1,k}) \Bigg\} \\ & \in A_\epsilon^{(T)}(\bar{P}_{k,j}, N_{k,j}), \end{aligned}$$

where

$$\bar{P}_{k,j} := \left(\sum_{i=0}^{k-j-1} \alpha_{ik} \sqrt{P_{i,k-j}} \right)^2 (1 - \varepsilon_1),$$

$$N_{k,j} := \sum_{l=1}^{k-j-1} \left(\sum_{i=0}^{l-1} \alpha_{ik} \sqrt{P_{il}} \right)^2 (1 - \varepsilon_1) + \sigma^2.$$

Otherwise, if an unique w as above does not exist, an error is declared and $\widehat{w}_{b-k+1,k}$ is set to 0.

9.3 Achieving super-linear growth in low attenuation regime

In this subsection we show that one can apply the CRIS strategy, presented in the previous subsection, to achieve super-linear growth of transport capacity when attenuation is slow, i.e., Theorems 9.1 and 9.2. The key steps for proving Theorem 9.1 are presented. The interested reader is refereed to [26] for the complete proofs.

Recall the following Theorem 9.1.

Theorem 9.1 *Suppose $\gamma = 0$. Then, for planar networks under the total power constraint P_{total}:*

i) If $\delta < 3/2$, then any arbitrarily large transport capacity can be supported in a regular planar network with large enough n, using the CRIS strategy.

ii) If $\delta < 1$, then CRIS can support a fixed rate $R_{\min} > 0$ for any single source–destination pair in any regular planar network, irrespective of the distance between them.

Proof. Consider one source–destination pair where the source node s is located at $(0,0)$ and the destination node d is located at $(m^q, 0)$, with q a positive integer to be determined.

We need the cooperation of $m-1$ groups of relay nodes: Group $\mathcal{N}_i$ consists of n_i nodes in a neighborhood of the node $(i^q, 0)$, for $i = 1, \ldots, m-1$, with $\mathcal{N}_0 = \{s\}$, $n_0 = 1$. Let the n_i nodes in Group $\mathcal{N}_i$ use same power $P_{ik} = \frac{P}{(k-i)^\lambda k^\mu}$, for $0 \le i < k \le n$, where $\lambda > 1, \mu > 1$ are two constants to be determined later, and

$$P := \frac{(\lambda-1)(\mu-1)}{\lambda\mu} P_{total}.$$

It is easy to verify that the total power constraint is satisfied.

Now by Theorem 9.5 (pp. 230), the following rate is achievable

$$R < \min_{1\le j\le r} \mathcal{S}\left(\frac{1}{\sigma^2}\sum_{k=1}^{j}\left(\sum_{i=0}^{k-1}\frac{\sqrt{P_{ik}/n_i}\cdot n_i}{r^{\delta}_{\mathcal{N}_i\mathcal{N}_j}}\right)^2\right), \tag{9.7}$$

where $r_{\mathcal{N}_i\mathcal{N}_j}$ is the maximum distance between any node in Group $\mathcal{N}_i$ and any node in Group $\mathcal{N}_j$.

For any $i = 1,2,\ldots,m-1$, let Group $\mathcal{N}_i$ be the set of nodes located at: $\{(u,v) : i^q \le u \le i^q + i^{q-1} - 1, -i^{q-1} \le v \le i^{q-1}\}$. It is easy to check that these groups are disjoint from each other and $n_i > i^{2(q-1)}$. Furthermore, for any $0 \le i < j < m$, $d_{ij} < j^q - i^q + i^{q-1} + j^{q-1} + j^{q-1} < 3j^q$. Hence by (9.7), the following rate is achievable:

$$R < \min_{1\le j\le r} \mathcal{S}\left(\frac{1}{\sigma^2}\sum_{k=1}^{j}\left(\sum_{i=0}^{k-1}\frac{\sqrt{P_{ik}}\cdot i^{q-1}}{3^{\delta}j^{q\delta}}\right)^2\right). \tag{9.8}$$

For $1 + 2q - \lambda - \mu > 0$, one can show (details in [26]) that the right hand side of (9.8) is lower bounded by $\Omega(j^{1+2q-\lambda-\mu-2q\delta})$.

Now we proceed with two cases.

Case 1. $\delta < 1$.

Choose q such that $1 + 2q - \lambda - \mu - 2q\delta > 0$. Then there exists $\underline{P} > 0$ such that for any j,

$$\sum_{k=1}^{j}\left(\sum_{i=0}^{k-1}\frac{\sqrt{P}\cdot i^{q-1}}{(k-i)^{\lambda/2}k^{\mu/2}3^{\delta}j^{q\delta}}\right)^2 \ge \underline{P}.$$

Then by (9.8), for any m, any $R < \mathcal{S}(\frac{\underline{P}}{\sigma^2})$ is achievable. Without loss of generality, this means that any $R < S(\frac{\underline{P}}{\sigma^2})$ is achievable with power constraint P_{total} for any single source–destination pair. Furthermore, $R \cdot m^q$ is an achievable network transport with power constraint P_{total}, which tends to infinity as $m \to \infty$.

Case 2. $1 \le \delta < \frac{3}{2}$.

In this case, choose q such that $1 + 3q - \lambda - \mu - 2q\delta > 0$. Then we have

$$m^q \min_{1 \le j \le m} \mathcal{S}\left(\frac{1}{\sigma^2} \sum_{k=1}^{j} \left(\sum_{i=0}^{k-1} \frac{\sqrt{P} \cdot i^{q-1}}{(k-i)^{\lambda/2} k^{\mu/2} 3^{\delta} j^{q\delta}} \right)^2 \right)$$
$$= \Omega(m^{1+3q-\lambda-\mu-2q\delta}) \to \infty, \quad \text{as } m \to \infty.$$

This means that an arbitrarily large transport capacity is achievable with a fixed total power constraint P_{total}. □

9.4 Notes

This section is based on [26].

10

The Transport Capacity of Wireless Networks with Fading

In Sections 8 and 9 we have studied the transport capacity of wireless networks from an information theoretic point of view, addressing the fundamental scaling behavior of wireless networks, and attempting to elucidate the order optimal architectures.

However, a very important issue in practical wireless networks, and unmodeled in these sections, is the presence of fading [22]. Fading is caused by the fact that wireless signals, in the form of electromagnetic waves, travel to a receiver along different paths because of scattering, shadowing, reflections, etc., [22, 5]. A fundamental question therefore is: How does fading influence the performance of wireless networks? This is the question addressed in this section.

We begin with a short discussion of multi-path fading in wireless communications. A baseband model incorporating fading and the attenuation due to distance will be presented, which can be specialized to various common fading models in the literature. The model in Sections 8 and 9 will be seen to be a special case.

Then the main results on the scaling behavior of transport capacity in high attenuation regime are presented. Specifically, if the fading is "power bounded," then even if the channel state information (CSI) is known before transmission to every node, the transport capacity is shown to be still upper bounded by a linear function of the network size. On the other hand, if the fading is independent from time to time, then even if the CSI is unknown to every node, a constructive scheme is provided to achieve linear growth of the transport capacity. Interestingly, peaky signalling and random time-scheduling emerge as strategies of interest to combat unknown fading. These results suggest that the scaling laws and order optimal architectures are not affected by the presence or absence of fading, be it slow or fast, and be it frequency flat or frequency selective.

10.1 Network under fading: model and definitions

We begin with a short discussion on fading channels, and then present the model under fading.

A. Fading Channels

In wireless communications, due to the physical environment, for example walls and trees, the electromagnetic waves travel to receivers along a multitude of paths. Along each path, the signal could encounter reflection, delay and path loss, which vary with time.

Fading can be characterized into four different types by comparing the frequency bandwidth used and how fast the environment changes.

First one can compare the signal bandwidth W with the channel coherence bandwidth B_{coh}. Intuitively, B_{coh} characterizes how far apart two different sinusoids with frequencies f_1 and f_2 have to be to face significantly different fading gains. If the bandwidth W of the signal is much smaller than the channel coherence bandwidth B_{coh}, i.e., $W \ll B_{coh}$, then the channel gain is almost the same for different frequency components in the signal. In other words, the receiver cannot distinguish the paths, and thus the channel only has a multiplicative effect on the signal. This situation is called frequency non-selective or *flat fading*. If on the other hand $W \geq B_{coh}$, different frequencies fade differently. The receiver can now get several resolvable paths, and such a channel is called *frequency selective*. Note that in this situation, intersymbol interference is introduced.

One can also characterize a fading channel by comparing the time duration T_s of a signal symbol with the channel coherence time T_{coh}. A channel is called *slow fading* if $T_s \ll T_{coh}$, meaning the channel gain is constant during transmission. If otherwise $T_s \geq T_{coh}$, the channel gain is varying during transmission, and this is called *fast fading*.

A common discrete model for a point-to-point fading channel is the tapped-delay baseband model, in which the received signal $Y(t)$ is given by

$$Y(t) = \sum_{l=0}^{L-1} H_l(t) X(t-l), \quad t = 1, 2, \cdots, \tag{10.1}$$

where L is the number of paths and $H_l(t)$ is the path gain for the l-th path.

The subject of fading channels has been an active research field for decades; the above characterization can be found in various references, for example in [22, 5, 25].

B. Network Model under Fading ($L = \infty$)

Based on (10.1) and the model in Section 8, we consider the following model for wireless networks in this section.

Consider a network consisting of n nodes in $\mathbb{N} := \{1, 2, \ldots, n\}$, located on the plane. The base-band model for the communications among them is described by the following equation:

$$\begin{aligned} Y_j(t) = \sum_{i \neq j} \frac{G e^{-\gamma d_{ij}}}{d_{ij}^{\delta}} \left(\sum_{l=0}^{\infty} H_{ijl}(t) \cdot X_i(t - \tau_{ij} - l) \right) \\ + Z_j(t), \quad t \geq 1, j \in \mathbb{N}. \end{aligned} \tag{10.2}$$

We assume that:

(1) δ, γ, G, d_{ij} and τ_{ij} are *deterministic real* variables known to all the nodes:

- δ and γ are the path-loss exponent and absorption constant of the attenuation, respectively. Throughout this section we assume that $\gamma > 0$ or ($\gamma = 0$, $\delta > 3$). $G > 0$ is just a constant gain.

- d_{ij} is the distance between node i and j, with $d_{ij} \geq d_{min} > 0$.
- $\tau_{ij} := \lfloor \frac{d_{ij}}{r_0} \rfloor$ is the propagation delay for signals from i to j, where r_0 is the distance that a signal travels in one time slot. (Later on we will see that the results do not really depend on the precise value of τ_{ij}.)

(2) $\{H_{ijl}(t)\}$, $\{Z_j(t)\}$, $\{Y_j(t)\}$ and $\{X_i(t)\}$ are *complex* variables.

- $\{H_{ijl}(t) : t \geq 1, i, j \in \mathbb{N}\}$ is the random fading process.
- $\{Z_j(t) : j \in \mathbb{N}, t \geq 1\}$ are i.i.d. complex circular Gaussian[1] noises with variance $E|Z_j(t)|^2 = \sigma^2 > 0$, independent of the fading process $\{H_{ijl}(t)\}$, and are not observable to the users.
- $\{X_i(t)\}$ is the complex base-band signal sequence node i transmits, and $\{Y_j(t)\}$ is the complex base-band signal sequence node j receives.

(3) Each node is subject to an individual power constraint P. Since the channel has multiple paths with delays, we need to model what may have been transmitted before time 0, the time when the useful transmissions begin. We simply suppose that the signals prior to $t = 0$ are independent of all else, unknown to the nodes, and satisfy

$$|X_i(t)|^2 \leq \bar{P} < \infty, \quad \forall t \leq 0, i \in \mathbb{N}. \tag{10.3}$$

With the same set of assumptions, one can specialize the model to the case when L is finite.

C. Network Model under Fading ($L < \infty$)

When there are no more than L paths for every channel, the base-band model for the communications in the network is described by the

[1] A complex random variable Z is circular Gaussian if it can be represented as $Z = Z_1 + \nu Z_2$ where Z_1 and Z_2 are two i.i.d. (real) Gaussian random variables, and ν is the square root of -1.

following equation:

$$Y_j(t) = \sum_{i \neq j} \frac{Ge^{-\gamma d_{ij}}}{d_{ij}^{\delta}} \left(\sum_{l=0}^{L-1} H_{ijl}(t) \cdot X_i(t - \tau_{ij} - l) \right) + Z_j(t), \quad t \geq 1, j \in \mathbb{N}. \tag{10.4}$$

Remark 10.1. We have presented the basic models and assumptions above. Yet more assumptions for the fading process $\{H_{ijl}(t)\}$ are needed. For example, one needs to specify whether a sender/receiver node knows $\{H_{ijl}(t)\}$ before transmission, and how fast it changes with time, etc. These further assumptions will be explicitly made clear for each result we will present below.

Remark 10.2. Clearly the models for $L = \infty$ and $L < \infty$ include all four fading patterns discussed at the beginning of this subsection.

Remark 10.3. The definition of transport capacity is the same as in Section 8.

10.2 Main results

We determine both an upper bound and a lower bound for the transport capacity of networks in high attenuation regime for the case $\delta > 3$. Such results also apply if $\gamma > 0$, irrespective of $\delta \geq 0$. It can be shown that, under certain general assumptions on the power of fading, even if every node knows the fading process non-causally before transmission, the transport capacity is upper bounded by $c_1 n$ for a constant $c_1 < \infty$. On the other hand, Theorem 10.9 shows that even when fading is independent from time to time and "power" bounded, $c_2 n$ bit-meters/second is achievable for a large class of networks, for some $c_2 > 0$.

It will be clear that, if $L < \infty$ and $\{H_{ijl}(t)\}$ is a set of independent random variables with variances less than a constant $\sigma_H^2 > 0$, and their distribution functions are continuous, then both requirements on fading are satisfied. That is, for such a class of independent fading networks, the transport capacity is indeed of order $\Theta(n)$.

A. Upper bound when $L = \infty$.

We establish an upper bound on the transport capacity under the following assumption on the fading process, for the case $L = \infty$.

Definition 10.4. A fading process is called *power limited* if there exist positive constants $\nu \in (0,1)$ and $\bar{H} > 0$, such that for any $i, j \in \mathbb{N}$, the fading process satisfies

$$\limsup_{T} \frac{1}{T} \sum_{t=1}^{T} \sum_{l=0}^{\infty} \nu^{-l} |H_{ijl}(t)|^2 \leq \bar{H}, \text{a.s.} \tag{10.5}$$

Theorem 10.5. Suppose the fading process is power limited, then even if the channel state information (CSI) $\{H_{ijl}(t)\}$ is known non-causally to all transmitters and receivers, the transport capacity is bounded as

$$C_T(n) \leq c_1 \cdot n, \quad \text{for all } n,$$

where

$$c_1 := \begin{cases} \frac{8G\sqrt{P\bar{H}}\log e}{\sigma\sqrt{1-\nu}(d_{min}/\sqrt{2})^{\delta-1}} \left(\frac{2(\delta-2)}{\delta-3} + \frac{\delta-1}{\delta-2} \right), \\ \qquad\qquad \text{if } \gamma = 0, \delta > 3; \\ \frac{48G\sqrt{P\bar{H}}\log e \cdot d_{min}^{1-\delta}}{\sigma\sqrt{1-\nu}} \frac{e^{-(\sqrt{2}/2)\gamma d_{min}}}{\left(1-e^{-(\sqrt{2}/4)\gamma d_{min}}\right)^4}, \\ \qquad\qquad \text{if } \gamma > 0, \delta \geq 0. \end{cases}$$

B. Lower Bound when $L = \infty$.

We need to establish a feasible lower bound on the transport capacity under the following assumptions on the fading process and node distributions, for the case $L = \infty$.

Definition 10.6. A fading process will be called *independent* if the following is true.

- $\{\mathcal{H}_j(t),\, t \geq 1\}$ is a sequence of independent random vectors, where $\mathcal{H}_j(t) := (H_{ijl}(t), i \in \mathbb{N}, l \geq 0)$ for all $j \in \mathbb{N}$. That is, the fading parameters are independent from time to time.
- There exist $a > 0$ and $p^* > 1/2$ such that, for all $i, j \in \mathbb{N}$, $t \geq 1$, $\Pr(|H_{ij0}(t)| \geq a) \geq p^*$.
- There exist $\nu \in (0,1)$ and $\tilde{H} > 0$ such that, for all $i, j \in \mathbb{N}$,

$$\sum_{l=0}^{\infty} \nu^{-l} E|H_{ijl}(t)|^2 \leq \tilde{H}. \tag{10.6}$$

Remark 10.7. The following are two special cases of independent fading processes: i) $\{H_{ijl}(t)\}$ is a set of constants with $|H_{ijl}(t)| \geq \epsilon > 0$, $\forall i, j, l, t$; ii) $\{H_{ijl}(t)\}$ is a set of non-zero iid random variables with continuous distribution and $E|H_{ijl}(t)|^2 < \infty$.

Definition 10.8. A network is called *nearly regular* if there exists $\zeta \geq 1$ such that for every node j there exists another node $\hat{j}$ with $d_{j,\hat{j}} \leq \zeta d_{min}$, i.e., every node can find a nearby neighbor; see Figure 10.1.

Now we have the following result.

Theorem 10.9. Suppose the fading process is independent and the network is nearly regular. Then, even if the CSI is *unknown* to transmitters and receivers, for any $\bar{p} \in (1/2, p^*)$ there exists a constant $c_2 > 0$, such that for any network,

$$C_T(n) \geq c_2 \cdot n, \quad \text{for all } n,$$

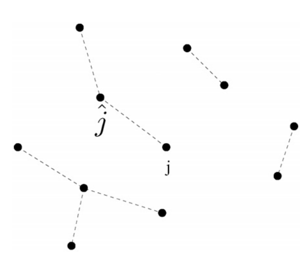

Fig. 10.1 A nearly regular network.

where

$$c_2 := \begin{cases} \min\left\{\frac{1}{2}, \frac{PG^2(\zeta d_{min})^{-2\delta}a^2\cdot(p^*-\bar{p})}{\frac{16\tilde{H}PG^2d_{min}^{-2\delta}}{1-\nu}(1+2^{2+\delta}(\frac{2\delta-1}{\delta-1}+\frac{2\delta}{2\delta-1}))+4\delta^2}\right\} \\ \qquad\cdot(1-H(\bar{p}))\cdot d_{min}, \quad \text{if } \delta>3, \gamma=0; \\ \min\left\{\frac{1}{2}, \frac{PG^2(\zeta d_{min})^{-2\delta}e^{-2\gamma\zeta d_{min}}a^2\cdot(p^*-\bar{p})}{\frac{16\tilde{H}PG^2d_{min}^{-2\delta}}{1-\nu}\left(e^{-2\gamma d_{min}}+\frac{12e^{-\sqrt{2}\gamma d_{min}}}{(1-e^{-\sqrt{2}\gamma d_{min}})^2}\right)+4\delta^2}\right\} \\ \qquad\cdot(1-H(\bar{p}))\cdot d_{min}, \quad \text{if } \delta\geq 0, \gamma>0, \end{cases}$$

with $H(\bar{p}) := -\bar{p}\log(\bar{p}) - (1-\bar{p})\log(1-\bar{p})$.

C. Results when $L<\infty$

The model in (10.4) is common in the literature. It is a special case of the model for $L=\infty$. Since there are only a finite number of paths for each channel, the requirements on fading for similar results as in Theorems 10.5 and 10.9 can be simplified. The following results are immediate.

Theorem 10.10. If $\limsup_T \frac{1}{T}\sum_{t=1}^{T}|H_{ijl}(t)|^2 \leq \bar{H}_0 < \infty$, a.s., $\forall i,j,l$, then there exists a positive constant c_1 such that $C_T(n)\leq c_1\cdot n$, for all n, even if the channel state information $\{H_{ijl}(t)\}$ is known non-causally to all transmitters and receivers.

Theorem 10.11. If $\{H_{ijl}(t)\}$ are independent continuous random variables with $E|H_{ijl}(t)|^2 \leq \tilde{H}_0 < \infty$, and $Pr(|H_{ijl}(t)|\geq a)\geq p^*$ for some $a>0$, $p^*>\frac{1}{2}$, $\forall i,j,l,t$, then there exists a positive constant c_2 such that $C_T(n)\geq c_2\cdot n$, for all nearly regular networks of n nodes, even if the channel state information $\{H_{ijl}(t)\}$ is unknown to any transmitters and receivers.

10.3 Upper bound

This subsection presents the proof for the upper bound (Theorem 10.5). A max-flow-min-cut lemma connecting rate to power for general networks will be presented first, which is similar to the one in Section 8.3.

The lemma will then be adapted to the networks satisfying the model presented earlier. Finally, in Section 10.3.3, the upper bound will be proved by a random cut-set technique.

10.3.1 Max-flow min-cut revisited

As in Section 8.3 for networks without fading, we have the following lemma.

Lemma 10.12. (Max-flow min-cut bound) Suppose receptions in a wireless communication network are modeled as

$$Y_j(t) = \sum_{i \neq j} \sum_{l=0}^{\infty} A_{ijl}(t) X_i(t - \tau_{ij} - l) + Z_j(t), j \in \mathbb{N},$$

where: (i) The τ_{ij}'s are deterministic non-negative integers. (ii) $\{A_{ijl}(t)\}$ is a sequence of *known deterministic* complex numbers. (iii) $\{Z_j(t)\}$ is the i.i.d. circular Gaussian noise process independent of the signal process $\{X_i(t)\}$, and $E|Z_j(t)|^2 = \sigma^2$.

Then for any subset S of N, any feasible rate vector $\{R_{ij},\, i, j \in \mathbb{N}\}$ satisfies

$$R_{SD} \leq \frac{1}{T} + R_{SD} P_e^{(T)} + \frac{1}{T} \sum_{t=1}^{T} \sum_{j \in \mathbb{N} \setminus S} \log \left(1 + \frac{E|\sum_{i \in S} \sum_{l=0}^{\infty} A_{ijl}(t) X_i(t - \tau_{ij} - l)|^2}{\sigma^2} \right),$$

where $\mathbb{N} \setminus S$ denotes those nodes in $\mathbb{N}$ but not in S, $R_{SD} := \sum_{i \in S, j \in \mathbb{N} \setminus S} R_{ij}$, and $P_e^{(T)}$ is the probability of the decoding error.

Note here that a time-varying fading process is allowed. The proof is similar to that of Lemma 8.9; see [28].

Applying Lemma 10.12 to the wireless networks under consideration (10.2), one can get the following corollary.

Corollary 10.13. If the fading process in a wireless network modeled in (10.2) is power limited (Definition 10.4), and the realization of the fading process is *known beforehand* to all nodes, then for any subset

S of $\mathbb{N}$, and feasible rate vector $\{R_{ij}, i, j \in \mathbb{N}\}$, $R_{SD} := \sum_{i \in S, j \in \mathbb{N} \setminus S} R_{ij}$ satisfies

$$R_{SD} \leq \log e \cdot \frac{2\beta\sqrt{P\bar{H}}}{\sigma\sqrt{1-\nu}} \cdot \sum_{i \in S, j \in \mathbb{N} \setminus S} d_{ij}^{-\delta} e^{-\gamma d_{ij}}. \tag{10.7}$$

Proof. We show the proof for the case when there are only a finite number of paths, i.e., $H_{ijl}(t) = 0$, for all $l \geq L$. The proof for the general case can be found in [28].

We denote $\alpha_{ij} := d_{ij}^{-\delta} e^{-\gamma d_{ij}}$. Let $D = \mathbb{N} \setminus S$. Since the rate vector is feasible, we know $P_e^{(T)} \to 0$ as $T \to \infty$. So by Lemma 10.12, we have

$$R_{SD} \leq \frac{1}{T} \sum_{t=1}^{T} \sum_{j \in D} \log \left(1 + \frac{G^2}{\sigma^2} E \left| \sum_{i \in S} \sum_{l=0}^{L-1} \alpha_{ij}^{-\delta} H_{ijl}(t) X_{i,T}(t - \tau_{ij} - l) \right|^2 \right) + o(1).$$

Furthermore,

$$\begin{aligned}
&\frac{G^2}{\sigma^2} E \left| \sum_{i \in S} \sum_{l=0}^{L-1} \alpha_{ij}^{-\delta} H_{ijl}(t) X_{i,T}(t - \tau_{ij} - l) \right|^2 \\
&= \frac{G^2}{\sigma^2} E \left| \sum_{i \in S} \sum_{l=0}^{L-1} \left(\alpha_{ij}^{-\delta/2} H_{ijl}(t) \nu^{-l/2} \right) \cdot \left(\alpha_{ij}^{-\delta/2} X_{i,T}(t - \tau_{ij} - l) \nu^{l/2} \right) \right|^2 \\
&\leq \frac{G^2}{\sigma^2} E \left(\sum_{i \in S} \sum_{l=0}^{L-1} \left| \alpha_{ij}^{-\delta/2} H_{ijl}(t) \nu^{-l/2} \right|^2 \right) \\
&\qquad \times \left(\sum_{i \in S} \sum_{l=0}^{L-1} \left| \alpha_{ij}^{-\delta/2} X_{i,T}(t - \tau_{ij} - l) \nu^{l/2} \right|^2 \right) \\
&= \frac{G^2}{\sigma^2} \left(\sum_{i \in S} \sum_{l=0}^{L-1} \alpha_{ij}^{-\delta} \cdot |H_{ijl}(t)|^2 \cdot \nu^{-l} \mu \right) \\
&\qquad \times \left(\sum_{i \in S} \sum_{l=0}^{L-1} \alpha_{ij}^{-\delta} \cdot E \left| X_{i,T}(t - \tau_{ij} - l) \right|^2 \cdot \nu^{l} \mu^{-1} \right),
\end{aligned}$$

where the inequality comes from the Cauchy-Schwarz inequality, and the last equality follows from the fact that we know $H_{ijl}(t)$ beforehand, with μ a constant to be determined later.

It is easy to verify that for any non-negative x and y, $\log(1+xy) \leq (x+y)\log e$. Hence

$$\begin{aligned}
&R_{SD} - o(1) \\
&\quad \leq \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in D}\log\left(1 + \left(\sum_{i\in S}\sum_{l=0}^{L-1}\frac{\alpha_{ij}^{-\delta}G\mu}{\sigma}|H_{ijl}(t)|^2\nu^{-l}\right)\right. \\
&\qquad \left.\times\left(\sum_{i\in S}\sum_{l=0}^{L-1}\frac{\alpha_{ij}^{-\delta}G\mu^{-1}}{\sigma}E|X_{i,T}(t-\tau_{ij}-l)|^2\nu^{l}\right)\right) \\
&\quad \leq \log e\cdot\frac{1}{T}\sum_{t=1}^{T}\sum_{j\in D}\sum_{i\in S}\sum_{l=0}^{L-1}\left(\frac{\alpha_{ij}^{-\delta}G\mu}{\sigma}|H_{ijl}(t)|^2\nu^{-l}\right. \\
&\qquad \left.+\frac{\alpha_{ij}^{-\delta}G\mu^{-1}}{\sigma}E|X_{i,T}(t-\tau_{ij}-l)|^2\nu^{l}\right) \\
&\quad = \log e\sum_{i\in S, j\in D}\frac{\alpha_{ij}^{-\delta}G}{\sigma}\left(\mu\cdot\frac{1}{T}\sum_{t=1}^{T}\sum_{l=0}^{L-1}|H_{ijl}(t)|^2\nu^{-l}\right. \\
&\qquad \left.+\mu^{-1}\cdot\frac{1}{T}\sum_{t=1}^{T}\sum_{l=0}^{L-1}E|X_{i,T}(t-\tau_{ij}-l)|^2\nu^{l}\right). \qquad (10.8)
\end{aligned}$$

Because of the power-boundedness of the fading (10.5), (10.3), and the individual power constraint P on the signals, we have

$$\begin{aligned}
&\text{R.H.S. of (10.8)} \\
&\quad \leq \log e\sum_{i\in S, j\in D}\frac{\alpha_{ij}^{-\delta}G}{\sigma}\left(\mu\bar{H} + o(1) + \mu^{-1}\right. \\
&\qquad \left.\times\sum_{l=0}^{L-1}E\left(\frac{1}{T}\sum_{t=1}^{T}|X_{i,T}(t-\tau_{ij}-l)|^2\right)\nu^{l}\right) \\
&\quad \leq \log e\sum_{i\in S, j\in D}\frac{\alpha_{ij}^{-\delta}G}{\sigma}\left(\mu\bar{H} + o(1) + \mu^{-1}\right.
\end{aligned}$$

$$\times \sum_{l=0}^{L-1} \left(P + \frac{\tau_{ij} + l}{T} \bar{P} \right) \nu^l \Bigg)$$

$$= \log e \sum_{i \in S, j \in D} \frac{\alpha_{ij}^{-\delta} G}{\sigma} \Bigg(\mu \bar{H} + o(1) + \mu^{-1} \frac{P}{1-\nu}$$

$$+ \mu^{-1} \frac{\tau_{ij}}{T} \frac{\bar{P}}{1-\nu} + \mu^{-1} \frac{\bar{P}}{T} \sum_{l=0}^{L-1} l \nu^l \Bigg)$$

$$= \log e \sum_{i \in S, j \in D} \frac{\alpha_{ij}^{-\delta} G}{\sigma} \Bigg(\mu \bar{H} + o(1) + \mu^{-1} \frac{P}{1-\nu}$$

$$+ \frac{\mu^{-1}}{T} \left(\frac{\tau_{ij} \bar{P}}{1-\nu} + \bar{P} \frac{\nu}{(1-\nu)^2} \right) \Bigg).$$

Now, letting $T \to \infty$, we get

$$R_{SD} \leq \log e \sum_{i \in S, j \in D} \frac{\alpha_{ij}^{-\delta} G}{\sigma} \left(\mu \bar{H} + \mu^{-1} \frac{P}{1-\nu} \right).$$

The result then follows by setting $\mu = \sqrt{\frac{P}{\bar{H}(1-\nu)}}$. □

10.3.2 Connecting rate vector to distance using random cut-sets

We now exhibit a natural relationship between cut-sets and the distance-rate product. It allows one to easily convert results for rate vectors across cut-sets, a staple feature in network information theory, to results on the transport capacity.

Lemma 10.14. If for a set of numbers $\{a_{ij}, i, j \in \mathbb{N}\}$, $\sum_{i,j} a_{ij} I_{[i \in S, j \in \mathbb{N} \setminus S]} \geq 0$ holds for every subset S of N, then $\sum_{i,j} a_{ij} d_{ij} \geq 0$.

Proof. By the symmetry of the condition, for any subset S of N,

$$\sum_{i=1}^{n} \sum_{j \neq i} a_{ij} (I_{[i \in S, j \notin S]} + I_{[i \notin S, j \in S]}) \geq 0. \tag{10.9}$$

Now we will construct a random event to select a subset S of $\mathbb{N}$, such that that probability that node i will be separated from node j is proportional to d_{ij}.

Cover the n-node network with a big disk C. First pick a diameter uniformly, then pick a point on that diameter uniformly. Then pick a straight line to cut C, passing through the point and being perpendicular to the diameter. Now we pick equally likely one of the two sets of nodes separated by the line as S.

It is easy to show that

$$\text{Prob}\{ \text{ Line } F \text{ separates points } i \text{ and } j\} = \frac{d_{ij}}{\pi r},$$

where r is the radius of the disk.

Taking the expectation on both sides of (10.9), we get the result. □

10.3.3 Proof of upper bound

We only show the proof of Theorem 10.5 for the case when $\gamma = 0$, $\delta > 3$. The general case can be shown similarly.

The following lemma is needed to bound $\sum_{j \neq i} d_{i,j}^{-\delta'}$ for $\delta' > 2$.

Lemma 10.15. For any $i_0 \in \mathbb{N}$ and $\delta' > 2$,

$$\sum_{j \neq i_0} d_{i_0,j}^{-\delta'} \leq \frac{4}{(d_{min}/\sqrt{2})^{\delta'}} \left(\frac{2(\delta' - 1)}{\delta' - 2} + \frac{\delta'}{\delta' - 1} \right).$$

Proof. The reason that the sum is finite is because there is a minimum separation between any pair of nodes. This means the maximum number of nodes that can be packed within a disk centered at i_0 is proportional to its area. Since $d^{-\delta'}$ decreases rapidly enough with distance, the sum is upper bounded by a finite number. The complete proof can be found in [28]. □

Now we prove Theorem 10.5.

Proof of Theorem 10.5: for the case $\gamma = 0$, $\delta > 3$: By Corollary 10.13, we know that for any subset S of $\mathbb{N}$, $\sum_{i,j} R_{ij} I_{[i \in S, j \in \mathbb{N} \setminus S]} \leq$

$\log e \cdot \frac{2\beta\sqrt{P\bar{H}}}{\sigma\sqrt{1-\alpha}} \cdot \sum_{i,j} d_{ij}^{-\delta} I_{[i\in S, j\in \mathbb{N}\setminus S]}$. Hence by Lemma 10.14,

$$\sum_{ij} R_{ij} d_{ij} \le \log e \cdot \frac{2\beta\sqrt{P\bar{H}}}{\sigma\sqrt{1-\nu}} \cdot \sum_{ij} d_{ij}^{1-\delta}.$$

Applying Lemma 10.15 we get the desired result. □

10.4 Achieving linear growth

In this subsection we cover the main ideas behind the proof of linear growth of transport capacity for nearly regular planar networks (Figure 10.1) under independent fading.

For a given nearly regular network, we will let each node $j \in \mathbb{N}$ only transmit to its nearest neighbor $\hat{j}$, which is no more than ζd_{min} away.

The whole idea of the construction is to focus on the channel between each individual transmitter-receiver pair. The channel will be translated into a binary symmetric channel (BSC) with a crossover probability strictly less than half, in every slot that the source node is transmitting. Building on top of this BSC one can certainly achieve a strictly positive rate. Thus Theorem 10.9 is proven since the information travels at least a distance d_{min} for each pair.

Suppose node 2 is node 1's chosen destination. Let us focus on pair (1,2). We want to "build" a BSC between them whenever node 1 is transmitting. Three facts can cause channel degradation in addition to the additive circular Gaussian noise: (i) *Channel variation*: This is a random variable that could be very large or very small; (ii) *Inter-channel interference*: The simultaneous transmissions can cause interference; and (iii) *Inter-symbol interference*, since the fading is frequency selective.

Since the CSI is not known, to combat these one needs to exploit the statistical properties of the independent fading process (Definition 10.6). From that definition we notice the following facts:

- With probability $p_1 > 1/2$, the amplitude of the first tap gain, $|H_{120}(t)|$, will be larger than a positive constant a. On the other hand, it will be smaller than a positive constant

M with probability larger than $p_2 > 1/2$, since its variance is bounded above;
- The signal attenuates significantly with distance according to $r^{-\delta}e^{-\gamma r}$. This will help to eliminate interference caused by other transmissions;
- Tap gains $|H_{12l}(t)|$ decrease significantly as l gets large. This will help to eliminate inter-symbol interference.

Taking into account the above considerations, the transmissions are constructed as follows.

(1) *On-off coding*: Every sender node generates a random binary (0 or 1) codebook sequence. The codebooks are independent of each other.
(2) *Random time-sharing before transmission*: Each sender only selects a small portion of the communication horizon to be its duty slots for transmission. By doing so it can save power, avoid inter-symbol interference, and decrease the interference to others at the same time. A node's duty slots are determined before transmission by a random time-sharing mechanism that is known to all nodes.
(3) *Peaky signaling and randomizing phase*: During a node's duty slots, a code symbol will be sent out after first multiplying by a high gain and a random phase. The purpose of the random phase is to make interference look like random noise.
(4) *Threshold decoding*: At the receiver side, it will set a threshold $M > 0$ and compare the received signal strength with it. It declares a "1" if the threshold is exceeded; otherwise it declares a "0". Thus the channel is turned into a binary-input-binary-output channel.

Now we present the above steps in more detail.
We begin by fixing a number $\bar{p} \in (1/2, p^*)$. For $\epsilon > 0$ sufficiently small, we introduce the following quantities for brevity:

$$P_\epsilon := P - \epsilon,$$

$$\lambda := \begin{cases} \frac{4}{\left(d_{min}/\sqrt{2}\right)^{2\delta}}\left(\frac{2\delta-1}{\delta-1}+\frac{2\delta}{2\delta-1}\right), & \text{if } \delta > 3, \gamma = 0; \\ \frac{12 d_{min}^{-2\delta} e^{-\sqrt{2}\gamma d_{min}}}{(1-e^{-\sqrt{2}\gamma d_{min}})^2}, & \text{if } \delta \geq 0, \gamma > 0, \end{cases} \tag{10.10}$$

$$\theta := \min\Bigg\{ 1/2, \tag{10.11}$$

$$\frac{(P-\epsilon)\beta^2(\zeta d_{min})^{-2\delta} e^{-2\gamma\zeta d_{min}} a^2 \cdot (p^* - \bar{p})}{\frac{16\tilde{H}(P-\epsilon)\beta^2}{1-\alpha}\left(d_{min}^{-2\delta} e^{-2\gamma d_{min}} + \lambda\right) + 4\delta^2}\Bigg\},$$

$$\theta_\epsilon := \theta - \epsilon.$$

For any $\epsilon > 0$ making the above quantities positive, and $\epsilon_1 \in (0, \bar{p} - 1/2)$, we will show that the information rate $\theta_\epsilon(1 - H(\bar{p} - \epsilon_1) - \epsilon_1)$ is achievable for every node pair j and $\hat{j}$, simultaneously.

From now on, $\bar{p}$, ϵ and ϵ_1 are all fixed.

10.4.1 Random coding

Each node is given $\theta_\epsilon T$ slots to transmit during a communication horizon T.

For a given rate $R = 1 - H(\bar{p} - \epsilon_1) - \epsilon_1$, the n nodes generate their codebooks individually, independently of each other. Node j generates a $2^{\theta_\epsilon TR} \times \theta_\epsilon T$ random matrix with entries being i.i.d. binary valued r.v.'s with distribution $p(x)$ such that $\Pr(X = 0) = 1/2$, and $\Pr(X = 1) = 1/2$. The w^{th} codeword is the w^{th} row of this matrix. The codebook of node j is denoted as $C_j := \{X_w^j = (X_{w,1}^j, X_{w,2}^j, \ldots, X_{w,\theta_\epsilon T}^j) : w = 1, 2, \ldots, 2^{\theta_\epsilon TR}\}$, and it is revealed to the intended receiver node $\hat{j}$.

10.4.2 Time scheduling of transmissions

If all the nodes transmit in the same time slots, then each receiver will face strong interference from nearby nodes. So we make nodes transmit in a timeshared fashion. Specifically, for any given large $T > 0$, each node $j \in \mathbb{N}$ only transmits at a set of pre-selected increasing time-slots t_k^j, $k = 1, 2, \ldots, \theta_\epsilon T$. This set is called the set of *duty slots of node* j. The corresponding (intended) receiver $\hat{j}$ will decode based only on the signals it receives at time slots t_k^j, $k = 1, \ldots, \theta_\epsilon T$. (Note that, without loss of generality, we assume here that $\tau_{j\hat{j}} = 0$ for all $j \in \mathbb{N}$).

Now nodes "bid" for their duty slots in the following way, before transmission.

Every node $j \in \mathbb{N}$ independently generates a sequence of i.i.d. Bernoulli r.v.'s $B_j(t)$, $t \geq 1$, with $\Pr\{B_j(t) = 0\} = 1 - 2\theta$, and $\Pr\{B_j(t) = 1\} = 2\theta$. Define $B_j(t) := 0$ for $t \leq 0$. Let $\hat{t}_j(k)$ be the time slot in which node j gets the k^{th} 1 in its sequence.

According to the transmission scheme we will present later, a codeword will be multiplied by a gain $\sqrt{P_\epsilon/\theta}$ and a random phase before being sent out. Based on this, and by a careful analysis [28], one can show that, on average, if node j takes $\{\hat{t}_j(k)\}$ as its duty slots, the interference caused by simultaneous transmissions from other nodes is upper bounded by $\frac{2P_\epsilon\beta^2}{1-\alpha}(d_{min}^{-2\delta}e^{-2\gamma d_{min}} + \lambda), \forall j \in \mathbb{N}, k = 1, \ldots, \theta_\epsilon T$.

Based on this observation, one can prove the following lemma on the existence of a low interference schedule.

Lemma 10.16. (Bounded interference) Let the indicator function $b_i(t)$, $i \in \mathbb{N}$, be defined as follows:

$$b_i(t) = \begin{cases} 1, & \text{if slot } t \text{ is in node } i\text{'s duty slot;} \\ 0, & \text{if otherwise.} \end{cases} \tag{10.12}$$

Then for all T sufficiently large, there exists a set of natural numbers $\{t_k^j, k = 1, \ldots, \theta_\epsilon T; j \in \mathbb{N}\}$ such that if we let node j's duty slots be this set, then

$$\begin{aligned} & d_{min}^{-2\delta}e^{-2\gamma d_{min}}\beta^2 \sum_{l=1}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} b_j(t_k^j - l) \\ & \quad + \sum_{i \neq j, \hat{j}} d_{i\hat{j}}^{-2\delta} e^{-2\gamma r_{i\hat{j}}} \beta^2 \sum_{l=0}^{\infty} \alpha^l \frac{P_\epsilon}{\theta} b_i(t_k^j - \tau_{i\hat{j}} - l) \\ & \leq \frac{4P_\epsilon\beta^2}{1-\alpha}(d_{min}^{-2\delta}e^{-2\gamma d_{min}} + \lambda), \quad \forall j \in \mathbb{N}, k = 1, \ldots, \theta_\epsilon T, \end{aligned}$$

where λ is defined in (10.10).

10.4.3 The transmission schedule and random phases

Each node $j \in \mathbb{N}$ chooses a message W_j uniformly from $\{1, 2, \ldots, 2^{\theta_\epsilon TR}\}$. During j's duty slot t_k^j $(k = 1, 2, \ldots, \theta_\epsilon T)$, it first generates a random phase, $\exp(\vartheta\phi_j(t_k^j))$, where ϑ is the square root of -1, and $\phi_j(t_k^j) \sim U[0, 2\pi)$. Then it transmits $X_{w_j,k}^j \cdot \exp(\vartheta\phi_j(t_k^j))$. The random phases $\exp(\vartheta\phi_j(t))$'s are introduced just to help the decoding by eliminating the possible correlations among signals and fading processes. The receivers need not know their exact values.

10.4.4 Decoding by thresholding

Upon receiving the complex baseband signal sequence $\{Y_{\hat{j}}(t_k^j), 1 \le k \le \theta_\epsilon T\}$, node $\hat{j}$ first passes the sequence through a simple thresholding filter, as follows:

$$Y_k^{\hat{j}} = \begin{cases} 1, \text{ if } \left|Y_{\hat{j}}(t_k^j)\right|^2 \ge M; \\ 0, \text{ otherwise,} \end{cases} \quad k = 1, 2, \ldots, \theta_\epsilon T,$$

where

$$M := \left(\frac{4\tilde{H} P_\epsilon \beta^2}{1-\alpha} (d_{min}^{-2\delta} e^{-2\gamma d_{min}} + \lambda) + \sigma^2 \right) \Big/ (p^* - \bar{p}). \quad (10.13)$$

This is shown in Figure 10.2.

Then, node $\hat{j}$ declares that the index $\hat{W}_j$ was sent if

$$\frac{1}{\theta_\epsilon T} \sum_{k=1}^{\theta_\epsilon T} I_{[X_{\hat{W}_j,k}^j / \sqrt{P_\epsilon/\theta} = Y_k^{\hat{j}}]} \ge \bar{p} - \epsilon_1, \quad (10.14)$$

and there is no other codeword X_w^j that satisfies this same condition.

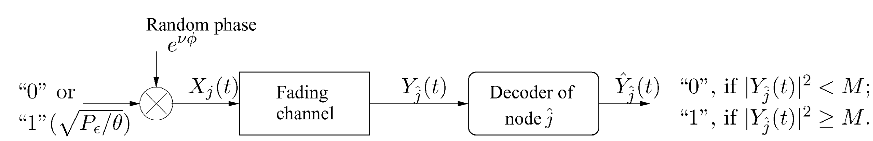

Fig. 10.2 The communication system.

If no such $\hat{W}_j$ exists, or if there is more than one such, then an error is declared.

The analysis of the probability of decoding error can be found in [28].

10.5 Notes

The section is based on [28].

11

MIMO Techniques for Wireless Networks

In Sections 8–10 we characterized the fundamental scaling behavior of the performance of wireless ad hoc networks, from an information theoretic point of view. In this chapter, still addressing similar questions, we show how multi-input multi-output (MIMO) techniques and analysis of random matrices, can be used to derive bounds on performance of wireless networks.

Using MIMO techniques, we first present an upper bound, in terms of individual powers and distances between nodes, on the transport capacity.

As an example of its usage, in Section 11.2, we address the information flow in a special class of random wireless networks, where n nodes in the left half of the domain want to transmit information to n other nodes in the right half. An upper bound is derived using MIMO techniques, showing that the sum rate cannot grow faster than a sub-linear function in n.

11.1 An upper bound on transport capacity using MIMO techniques

We begin with a short discussion on MIMO systems, and then apply the techniques to derive an upper bound on the transport capacity.

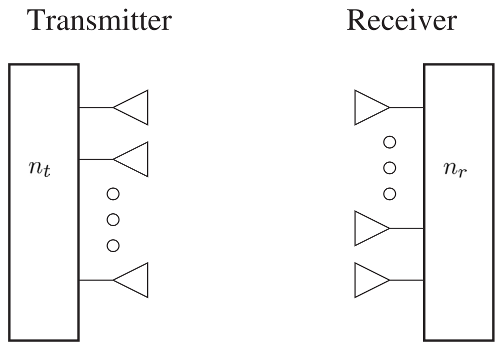

Fig. 11.1 An $n_t \times n_r$ MIMO system with n_t transmit antennas and n_r receive antennas.

A. MIMO systems

Recently as a result of developments in antenna and digital signal processing technology, multiple antennas on wireless devices have become feasible. Compared to the one antenna case, MIMO systems enjoy more reliable communications (called diversity gain) and sometimes much higher data rate (called multiplexing gain). This is due to the fact that signals can be collected through different paths with different path gains; see [23, 24]. The study of MIMO systems has been for some time now, e.g. [23, 24] and the references therein.

A typical setup for a MIMO system with n_t transmit antennas and n_r receive antennas is shown in Figure 11.1. If one denotes by h_{ij} the channel gain from the i-th transmit antenna, $1 \leq i \leq n_t$, to the j-th receive antenna, $1 \leq j \leq n_r$, then assuming additive Gaussian noise the system can be described by the following equation.

$$Y = H_{n_r \times n_t} X + Z,$$

where $H_{n_r \times n_t}$ is a $n_r \times n_t$ matrix with entries $\{h_{ij}\}$, Z is Gaussian noise with power σ^2, and X and Y are two vectors denoting the transmitted signal and received signal, respectively. In addition, the signal vector X is subject to a power constraint: $E||X_i||^2 \leq P_i$, for $1 \leq i \leq n_t$.[1]

It can be shown that the maximum information rate is achieved by using Gaussian random code books, i.e., with X is a vector of Gaussian

[1] Sometimes the signal power constraint is on the total power consumption: $E||X||^2 \leq P$.

random variables; see e.g., [23, 24]. Let us denote the covariance matrix of X by K_x. If H is only known to the receiver, then the maximum rate achievable is

$$\max_{K_x \succeq 0, (K_x)_{ii} \leq P_i} E[\log \det(I + \frac{1}{\sigma^2} H K_x H^\dagger)],$$

where $K_x \succeq 0$ denotes that K_x is non-negative definite. However if H is known to both the transmitter and the receiver, the maximum rate is

$$E\left[\max_{K_x \succeq 0, (K_x)_{ii} \leq P_i} \log \det(I + \frac{1}{\sigma^2} H K_x H^\dagger)\right].$$

Note that now K_x is a function of H.

B. Upper-bounding the transport capacity

Consider an n-node network, $\mathbb{N} := \{X_1, X_2, \cdots, X_n\}$, on the plane. The (discrete time) communications are in an iid flat fading environment. Specifically, the received signal at node j at time t is

$$Y_j(t) = \sum_{i \neq j} \frac{H_{ij}(t)}{d_{ij}^\delta} X_i(t) + Z_j(t), \tag{11.1}$$

where $X_i(t)$ is the signal transmitted by node X_i at time t, d_{ij} is the distance between nodes X_i and X_j, and $\{Z_j(t), \forall j, t\}$ is iid circular Gaussian noise with variance σ^2. $\{H_{ij}(t), t \geq 0\}$ is a stationary and ergodic stochastic process, with the marginal probability distribution symmetric with respect to the origin, and independent for each pair of nodes (X_i, X_j). For simplicity we assume that $E|H_{ij}(t)|^2 = 1$ for all i, j, t. Furthermore, we suppose that each node X_i is subject to an individual power constraint P_i.

The following is an upper bound on the transport capacity.

Theorem 11.1. If the channel state information (CSI) $\{H_{ij}(t), \forall i, j, t\}$ is known, then the transport capacity is upper bounded by

$$\frac{1}{\sigma^2} \sum_{i,k,j=1}^{n} \frac{\sqrt{P_i P_k} \min(d_{ij}, d_{kj})}{d_{ij}^\delta d_{kj}^\delta}. \tag{11.2}$$

If instead the CSI is unknown, then the transport capacity is upper bounded by

$$\frac{1}{\sigma^2}\sum_{i,j=1}^{n}\frac{P_i d_{ij}}{d_{ij}^{2\delta}}. \tag{11.3}$$

Proof. We begin with the case where no CSI is available. Consider a subset of nodes $S \in \mathbb{N}$, and define $D := \mathbb{N}\backslash S$. We now bound the information flow from S to D by MIMO techniques.

Assume that nodes in S can collaborate as a *super-node* with $|S|$ transmitting antennas, as do the nodes in D with $|D|$ receiving antennas. Assume further that there is a genie that provides all CSI to *only* the nodes in D. These assumptions can only result in larger rates. Now the system is equivalent to a MIMO communication system with receiver-only CSI, with $|S|$ transmitting antennas and $|D|$ receiving antennas. Therefore, the information rate from S to D is upper bounded by the capacity of such a MIMO channel as follows:

$$\sum_{i\in S, j\in D} R_{ij} \le \max_{K_x\succeq 0,(K_x)_{ii}\le P_i} E[\log\det(I + \frac{1}{\sigma^2}HK_xH^{\dagger})] \tag{11.4}$$

$$:= \max_{K_x\succeq 0,(K_x)_{ii}\le P_i} R(K_x), \tag{11.5}$$

where K_x is a $|S|\times|S|$ covariance matrix, $K_x \succeq 0$ denotes that K_x is non-negative definite, and H is a matrix with entries of the form $H_{ij}(t)/d_{ij}^{\delta}$. By the assumption on the fading process $\{H_{ij}(t)\}$, one can show that the optimal covariance matrix K_x, in order to maximize the right hand side of (11.4), is diagonal (see [14] for details). Using the fact that the determinant of a Hermitian matrix is at most as large as the product of its diagonal entries (the Hadamard inequality), the right hand side of (11.4) is upper bounded by:

$$\max_{K_x\succeq 0,(K_x)_{ii}\le P_i} R(K_x) \le E\left[\sum_{j\in D}\log\left(1+\frac{1}{\sigma^2}\sum_{i\in S}P_i|H_{ij}|^2 d_{ij}^{-2\delta}\right)\right]$$

$$\le \sum_{j\in D}\log\left(1+\frac{1}{\sigma^2}\sum_{i\in S}P_i/d_{ij}^{2\delta}\right) \tag{11.6}$$

$$\leq \frac{1}{\sigma^2} \sum_{j \in D} \sum_{i \in S} \frac{P_i}{d_{ij}^{2\delta}}, \tag{11.7}$$

where (11.6) follows from Jensen's inequality, and (11.7) because $\log(1+x) \leq x$.

Combining (11.4) and (11.7), we get

$$\sum_{i \in S, j \in D} R_{ij} \leq \frac{1}{\sigma^2} \sum_{j \in D} \sum_{i \in S} \frac{P_i}{d_{ij}^{2\delta}}.$$

Applying Lemma 10.14 of Section 10.3.2 we thus get

$$\sum_{ij} R_{ij} d_{ij} \leq \frac{1}{\sigma^2} \sum_{i,j=1}^{n} \frac{P_i d_{ij}}{d_{ij}^{2\delta}}.$$

This shows (11.3). The bound in (11.2) can be shown similarly; see [14]. □

Remark 11.2. Theorem 11.1 together with further arguments can be used to show that the transport capacity of wireless networks under flat fast fading, and with $d_{ij} \geq d_{min} > 0$, is upper bounded by a linear function in n.

11.2 The sum rate across a cut-set for a class of random wireless networks

In this subsection, we consider the scaling behavior of the sum of the rates from one side of a wireless network to the other side.

Random Network: We consider the following random network consisting of n nodes, $\{(X_i^-, X_i^+), i = 1, \cdots, n\}$, distributed in the domain $\Omega_n := [-\sqrt{n}, \sqrt{n}] \times [0, \sqrt{n}]$. Node X_i^- wishes to send information to node X_i^+, for all i; see Figure 11.2. The location of X_i^- is selected uniformly from domain $\Omega_n^- := [-\sqrt{n}, 0] \times [0, \sqrt{n}]$, while X_i^+ is uniformly selected from $\Omega_n^+ := [0, \sqrt{n}] \times [0, \sqrt{n}]$, independently. All the pairs are independent from each other.

As in Sections 8–10, we assume that a signal attenuates with distance d according to the function

$$g(d) = \frac{e^{-\gamma d}}{d^\delta},$$

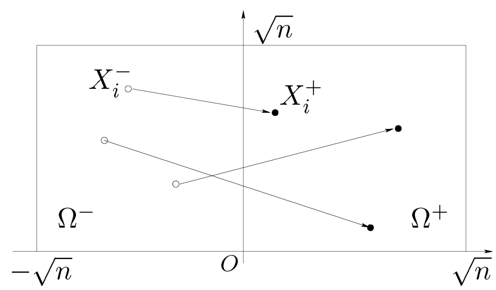

Fig. 11.2 A random network with 3 sender nodes and 3 receiver nodes.

where δ is the path-loss exponent,and γ is the absorption constant. We assume that there is no fading, and that each node is subject to an individual power constraint P. Furthermore, the received signal at each node is corrupted by an additive iid zero mean Gaussian noise with variance σ^2.

We are interested in the sum rate of the n pairs.

Theorem 11.3. If $\gamma > 0$, then the sum rate is upper bounded by $K_1\sqrt{n}(\log n)^2$, almost surely as $n \to \infty$, where K_1 is a positive constant.

Theorem 11.4. If $\alpha > 1$ and $\gamma = 0$, then the sum rate is upper bounded by $K_2 n\left(\frac{\log n}{n^{\frac{1}{2}-\frac{1}{\alpha}}} + \frac{1}{\sqrt{n}}\right)$, almost surely as $n \to \infty$, where K_2 is a positive constant.

We only present the key steps in proving Theorem 11.3; see [17] for details.

Proof of Theorem 11.3: Similar to Section 11.1, we assume that the nodes in Ω^- can collaborate as one super-node with n transmitting antennas, while nodes in Ω^+ can collaborate as one super-node with n receiving antennas. Furthermore, we allow the transmitting antennas to optimally re-allocate their powers while preserving the sum constraint at each super-node. That is, instead of individual power constraints, we impose the following cumulative constraint:

$$\sum_{i=1}^{n} E|X_i^-|^2 \leq nP.$$

Now the communication has been transformed into a MIMO channel with Gaussian noise. If we assume the received signal at node X_i^+ is Y_i^+, for all i, then the MIMO channel can be expressed as

$$Y_j^+ = \sum_{i=1}^{n} H_{ij} X_i^- + Z_j, \forall j,$$

where $\{Z_j\}$ is iid zero mean Gaussian noise with variance σ^2, and $H_{ij} = e^{-\gamma d_{ij}} d_{ij}^{-\delta}$, with $d_{ij} = |X_i^- - X_j^+|$. The sum rate R_{sum} is now upper bounded by the channel capacity. Using the information theoretic result for MIMO channels [23], the channel capacity is achieved when $(X_1^-, \cdots, X_n^-)$ is a jointly Gaussian vector with some covariance matrix Q. So we have

$$\begin{aligned} R_{sum} &\leq \max_{p_X : \sum_i E|X_i^-|^2 \leq nP} I(X_1^-, \cdots, X_n^-; Y_1^+, \cdots, Y_n^+) \\ &\leq \max_{Q \succeq 0 : Tr(Q) \leq nP} \log \det(I + HQH^\dagger). \end{aligned}$$

By a unitary transformation of the matrix H (see [23]), one obtains that

$$R_{sum} \leq \max_{p_i \geq 0 : \sum_{i=1}^{n} p_i \leq nP} \sum_{i=1}^{n} \log(1 + p_i \lambda_i^2) \leq \sum_{i=1}^{n} \log(1 + nP\lambda_i^2) =: B_n,$$

where $\{\lambda_i\}$ are the singular values of the matrix H, in decreasing order.

Noting that $\{\lambda_i^2\}$ are the eigenvalues of matrix $HH^\dagger$, we can bound B_n by the following majorization argument (see [18], pp.218, Theorem 9.B.1). We know that the eigenvalues $\{\lambda_i^2\}$ majorize the diagonal elements of $HH^\dagger$, i.e.,

$$\begin{aligned} &\sum_{i=1}^{l} \lambda_i^2 \geq \sum_{i=1}^{l} (HH^\dagger)_{ii}, \quad \forall 1 \leq l \leq n-1, \text{ and} \\ &\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} (HH^\dagger)_{ii}. \end{aligned}$$

On the other hand, by Proposition 3.C.1 of [18], pp. 54, we know that the function $(x_1, x_2, \cdots, x_n) \mapsto \sum_{i=1}^{n} \log(1 + nPx_i)$ is Schur-concave. That is, it is a function $f\colon \mathbb{R} \mapsto \mathbb{R}$ such that $f(x_1, \cdots, x_n) \leq f(y_1, \cdots, y_n)$ whenever $(x_1, \cdots, x_n)$ majorizes $(y_1, \cdots, y_n)$ in the sense above ([18], Definition 3.A.1, pp. 54). We therefore have

$$B_n \leq \sum_{i=1}^{n} \log(1 + nP(HH^{\dagger})_{ii}).$$

Moreover, since $e^{-\gamma d} d^{-\delta}$ is decreasing in d, we have

$$(HH^{\dagger})_{ii} = \sum_{l=1}^{n} |H_{il}|^2 = \sum_{l=1}^{n} (g(|X_i^- - X_l^+|))^2 \leq n(g(|\tilde{x}_i|))^2,$$

where $\tilde{x}_i$ is the x-coordinate of node X_i^-. So finally we obtain

$$R_{sum} \leq \sum_{i=1}^{n} \log\left(1 + n^2 P \frac{e^{-2\gamma|\tilde{x}_i|}}{|\tilde{x}_i|^{2\delta}}\right).$$

Notice that this is a sum of n iid random variables, with $|\tilde{x}_i|$ being uniformly distributed in $[-\sqrt{n}, 0]$. Through further analysis based on this fact (see [17]), this sum can be shown to be upper bounded by $K_1 \sqrt{n} (\log n)^2$, almost surely as $n \to \infty$, where K_1 is a positive constant. □

11.3 Notes

This section is based on [14] and [17].

12

Concluding Remarks

With the advent of wireless networks and sensor networks, there has naturally been much interest in quantifying what they can provide in the way of information transfer. Also, there is great interest in determining what the appropriate architectures are for operating them. This text presents one theory to address these questions. Starting with models which are salient to current technology and the current proposals to operate them in a multi-hop manner, we have shown how one can get insights by studying their behavior in terms of the number of nodes in the network. The performance measure of transport capacity measures the distance hauling capacity of wireless networks. We have provided sharp order estimates of the transport capacity in the best case. For random networks, we have studied the common throughput capacity that can be supported for all the nodes. The constructive procedure for obtaining the sharp lower bound gives insight into an order optimal architecture for wireless networks operating under a multi-hop strategy. Such results, it is hoped, can enable one to understand the forest of wireless networks as well as the role of protocols and trees in it.

However, the models and the mode of operation studied above beg the question of whether better, perhaps much better, information hauling capacity can be achieved by resorting to other more powerful techniques that the wireless medium can conceivably support. To address this, one is forced to turn to the fundamental formulation proposed by Shannon for the study of communication. However, in the case of networks, progress in information theory has been stymied for several decades by the intractability of obtaining exact results even for apparently simple networks such as the three-node relay channel or the two-by-two cross interference channel.

One needs to somehow overcome this barrier. For this purpose, we have introduced an enriched model of wireless networks in which spatial locations of nodes, distances between nodes, and the attenuation function for signals as a function of distance is more explicitly accounted for. What emerges from the resulting study is an interesting connection between physical properties of the medium such as the attenuation of signals with distance, and properties in the large such as the network-wide quantity of transport capacity.

There is an interesting dichotomy between the high attenuation regime and the low attenuation regime. In the former case, there is a fundamental energy cost in joules to be expended by the network in transmission energy for information transfer of one bit-meter. This leads to a scaling law that is of the same order as is achievable through multi-hop transport. This shows that the multi-hop mode of information transfer is in fact order-optimal in load-balanced scenarios. Such an architectural result provides some macroscopic strategic insight into design strategies for wireless networks where attention is often focused on more tactical considerations involving protocol improvements.

In the low attenuation regime, it is interesting that other strategies emerge as of interest which could even challenge entrenched notions such as spatial reuse of spectrum. Currently, these results are for unrealistically low attenuations, lower even than the inverse square law case.

An interesting open problem is that of bridging the gap between these two cases studied. Also of interest is obtaining more explicit topology and demand-dependent information on wireless network capacity than the notions of transport capacity and throughput capacity can provide. Further, the preconstants in bounds and order results need to be sharpened.

Clearly, much remains to be done.

Acknowledgements

This material is based upon work partially supported by NSF under Contracts NSF CNS 05-19535, ANI 02-21357 and CCR-0325716, AFOSR under Contract No. F49620-02-1-0217, DARPA/AFOSR under Contract No. F49620-02-1-0325, USARO under Contract No. DAAD19-01010-465, and DARPA under Contact No. N00014-0-1-1-0576.

References

[1] A. Agarwal and P. R. Kumar, "Improved capacity bounds for wireless networks," *Wireless communications and mobile computing*, vol. 4, pp. 251–261, 2004.

[2] A. Agarwal and P. R. Kumar, "Capacity bounds for ad hoc and hybrid wireless networks," *ACM SIGCOMM Computer Communications Reveiew*, vol. 34, no. 3, pp. 71–81, July 2004.

[3] D. Aguayo, J. Bicket, S. Biswas, D. S. J. D. Couto, and R. Morris, *MIT roofnet implementation*, http://www.pdos.lcs.mit.edu/roofnet/design/, August 2003.

[4] S. H. A. Ahmad, A. Jovicic, and P. Viswanath, "Upper bounds to generalized transport capacity of wireless networks," to appear in *IEEE Transactions on Information Theory*, June 2006.

[5] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[6] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, Inc., 1991.

[7] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, pp. 572–584, 1979.

[8] M. Franceschetti, J. Bruck, and L. Schulman, "Microcellular systems, random walks and wave propagation," in *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 5, pp. 1304–1317, May 2004.

[9] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, "Closing the gap in the capacity of random wireless networks," in ISIT 2004, p. 438.

[10] G. R. Grimmett, *Percolation.* Springer, 1999.

[11] 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.), pp. 547–560, Boston, MA: Birkhauser, March 1998. ISBN 0-8176-4078-9.

[12] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, pp. 388–404, March 2000.

[13] P. Johansson, M. Kazantzidis, R. Kapoor, and M. Gerla, "Bluetooth: An enabler for personal area networking," *IEEE Network*, vol. 15, pp. 28–37, September/October 2001.

[14] A. Jovicic, P. Viswanath, and S. R. Kulkarni, "Upper bounds to transport capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2555–2565, November 2004.

[15] E.-S. Jung and N. Vaidya, "A power control MAC protocol for ad hoc networks," in *ACM International Conference on Mobile Computing and Networking (MobiCom)*, September 2002.

[16] V. Kawadia and P. R. Kumar, "Principles and protocols for power control in ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 5, pp. 76–88, January 2005.

[17] O. Leveque and I. E. Telatar, "Information-theoretic upper bounds on the capacity of large extended ad hoc wireless networks," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 858–865, March 2005.

[18] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications,* Series in Mathematics in Science and Engineering. vol. 143, New York-London: Academic, 1979.

[19] J. P. Monks, V. Bharghavan, and W.-M. Hwu, "A power controlled multiple access protocol for wireless packet networks," in *Proceedings of INFOCOM*, pp. 219–228, 2001.

[20] S. Narayanaswamy, V. Kawadia, R. S. Sreenivas, and P. R. Kumar, "Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the COMPOW protocol," *Proceedings of European Wireless 2002. Next Generation Wireless Networks: Technologies, Protocols, Services and Applications*, pp. 156–162, Florence, Italy, February 25–28 2002.

[21] M. D. Penrose, "The longest edge of the random minimal spanning tree," *The Annals of Applied Probability*, vol. 7, no. 2, pp. 340–361, 1997.

[22] J. G. Proakis, *Digital Communications*. McGraw-Hill, fourth ed., 2001.

[23] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, November/December 1999.

[24] D. Tse and P. Viswanath, *Fundamental of wireless communications*. Cambridge UK: Cambridge University Press, 2005.

[25] V. V. Veeravalli, in Wideband Multi-Antenna Wireless Channels: Statistical Modeling, Analysis, and Simulation, IEEE VTC 2001 Tutorial, October 2001. Tutorial #7.

[26] L.-L. Xie and P. R. Kumar, “A network information theory for wireless communication: Scaling laws and optimal operation,” *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 748–767, February 2004.

[27] L.-L. Xie and P. R. Kumar, “On the path-loss attenuation regime for positive cost and linear scaling of transport capacity in wireless networks,” submited to *IEEE Transactions on Information Theory*, February 2005.

[28] F. Xue and P. R. Kumar, “The transport capacity of wireless networks under fading channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 834–847, March 2005.

[29] Bluetooth SIG Inc., “Specification of the Bluetooth system: Core,” http://www.bluetooth.org/, 2001.

[30] IEEE 802.11 Standard, “IEEE Local Area Network Working Group,” IEEE Standard for Information Technology.

[31] IEEE P802.16-REVd/D5-2004, “Part 16: Air interface for fixed broadband wireless access systems,” May 13, 2004.

[32] Intel White Paper, “IEEE802.16 and WiMAX: broadband wireless access for everyone,” July 2003.

[33] LAN-MAN Standard Committee of the IEEE Computer Society, *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)*, IEEE, 2003.

[34] ZigBee Alliance, http://www.caba.org/standard/zigbee.html.