Chapter 1

Introduction

1.1 Introduction

Wireless networks can be broadly classified into two categories: centralized and de-centralized. A canonical example of a centralized network is a cellular network, where all operations are controlled by basestations, for example, when should each user transmit or receive, thereby avoiding simultaneous transmission (interference) by closely located nodes. Prominent examples of de-centralized or *ad hoc* networks include sensor or military networks. Sensor network is deployed in a large physical area to either monitor physical parameters, such as temperature, rainfall, and animal census, or intrusion detection. In a military network, a large number of disparate military equipment, e.g., battle tanks, helicopters, ground forces, is connected in a decentralized manner to form a robust and high throughput network. Ad hoc networks are attractive because of their scalability, self-configurability, robustness, etc.

Vehicular network is a more modern example of an ad hoc wireless network, where a large number of sensors are deployed on the highways as well as mounted on vehicles that are used for traffic management, congestion control, and quick accident information exchange. Many other applications of ad hoc wireless networks are also envisaged such as deploying large number of sensors in large building for helping fire fighters in case of fire emergency and in case of earthquakes.

The key feature that distinguishes centralized and ad hoc wireless networks is interference. With centralized control, interference can be avoided in contrast to ad hoc networks, where there is no mechanism of inhibiting multiple transmitters from being active simultaneously. Thus, ad hoc networks give rise to complicated signal interaction at all receiver nodes. As compared to additive noise, interference is structured, and treating interference as noise is known to be sub-optimal. Thus, performance analysis of ad hoc wireless networks is far more complicated than centralized wireless networks.

In this book, we are interested in studying the physical layer issues of ad hoc wireless networks, such as finding the limits on the reliable rate of information transfer and ensuring connectivity among all nodes of the network. Traditionally, the Shannon capacity has been used to characterize the reliable rate of information transfer in communication systems. In a wireless network, however, finding the Shannon capacity is challenging and has remained unsolved. The main impediment in finding the Shannon capacity of wireless networks is the complicated nature of interference created by multiple simultaneously active transmitters at each other's receivers and network topology that directly influences the signal interaction.

To get some meaningful insights to the fundamental limits of throughput in wireless networks, alternate notions of capacity have been introduced and analyzed, such as transmission [1] and throughput/transport [2] capacity, which are defined by relaxing the reliability constraints compared to the Shannon capacity.

One key relaxation/assumption made for the purposes of analyzing these new capacity metrics is that the nodes of the network are assumed to be distributed uniformly at random in the area of interest, called the *random* wireless networks. The random node location assumption allows the use of tools from stochastic geometry and percolation theory for theoretical capacity analysis. In Chapter 2, we argue that random node location assumption is not very limiting for a practical ad hoc network.

Major focus of this book is on finding the transmission and the throughput capacity of random wireless networks. Through the transmission capacity formulation, we also quantify the effects of using multiple antennas at each node, using two-way communication between source and destination, effect of ARQ protocol, and using “smart” scheduling protocols in the random wireless networks. From here on in this book, when we say wireless network, we mean a random wireless network unless specified differently.

A necessary condition for finding the maximum rate of transmission or throughput between a pair of nodes in a wireless network is to ensure that they are connected to each other or have a connected path between each other, under a suitable definition of connection. Since any source can have an arbitrary choice for its destination, essentially, we need network wide connectivity, that is, each node pair should be reachable from every other node via connected paths. This condition is simply called as *connectivity* of the wireless network. Connectivity in a wireless network depends on the density of nodes, radio (transmission) range of any node, topology of the network, connection model between nodes, etc. In this book, we present relevant results from the percolation theory and then describe their application in finding the network parameters that ensure connectivity in wireless networks. Using percolation theory, we also study the size of the largest connected component in wireless networks and find conditions when the size of the largest connected component is a non vanishing fraction of the total number of nodes, which implies approximate connectivity.

The book is divided into two parts, first part exclusively deals with a single-hop model for wireless networks, where each source has a destination at a fixed distance from it and transmits its information directly to its destination without the help of any other node in the network. We define the notion of transmission capacity for the single-hop model and derive it for single antenna nodes, multiple antenna nodes, with scheduling protocols, and under two-way communication scenarios. The first part of the book also includes the performance analysis of cellular wireless network techniques using tools from stochastic geometry that are developed in the earlier chapters of the first part.

In the second part, we deal with the more relevant model of multi-hop communication for a wireless network and define two notions of capacity, namely the delay normalized transmission capacity and the throughput capacity and present their analysis. In addition, in the second part, we also study the connectivity and percolation properties of a multi-hop wireless network under the signal-to-noise-plus-interfence ratio (SINR) model.

This chapter sets up the background for studying wireless networks from a physical layer point of view. We begin by describing the basics of point-to-point communication, where a single transmitter is interested in communicating with a single receiver. To keep the discussion general, we consider the case when each node is equipped with multiple antennas. We first discuss the role of multiple antennas in improving the error-probability performance as a function of number of transmit and receive antennas with the optimal maximum likelihood (ML) decoder. We then state some difficulties in using the optimal maximum likelihood decoder, such as an exponential complexity, and present the more popular sub-optimal decoders such as zero-forcing decoder that have linear decoding complexity. We also discuss the error rate performance degradation while using the sub-optimal zero-forcing decoder.

Next, we define the notion of Shannon capacity and present results on the Shannon capacity of the point-to-point communication channel with multi-antenna equipped nodes. We show that the Shannon capacity scales linearly with the minimum of the number of transmit and receive antennas. We next present the outage formulation for characterizing capacity (called outage capacity) in non-ergodic channels, for which the Shannon capacity is zero. The non-ergodic channel is of interest since the popular slow-fading channel model of wireless signal propagation, where channel coefficients remain constant for sufficient amount of time, falls in the class of non-ergodic channels. The outage formulation also helps in defining the transmission capacity of wireless networks.

Next, we describe the received signal model at any node of a wireless network, where multiple transmitters are active at the same time. Using examples of some basic building blocks of a wireless network, we discuss some of the difficulties in finding the Shannon capacity of a wireless network. We then motivate the definitions of alternate capacity metrics, such as transmission capacity and throughput capacity, which are defined under a relaxed reliability constraint compared to the Shannon capacity.

We end this chapter by presenting some details on studying connectivity in wireless networks under various link connection models.

1.2 Point-to-Point Wireless Signal Propagation Model

Consider a wireless communication channel between a single transmitter T_0 equipped with N_t antennas and a single receiver R_0 with N_r antennas. Let the distance between T_0 and R_0 be d, then the received signal at R_0 at time t is given by

$$\mathbf{y}[t] = d^{-\alpha/2}\sqrt{\frac{P}{N_t}}\sum_{m=0}^{M-1}\mathbf{H}_m\mathbf{x}[t-m] + \mathbf{w}[t], \tag{1.1}$$

where M is the number of distinct multiple fading paths between the transmitter and the receiver, $d^{-\alpha/2}$ is the distance-based path-loss function, α is the path-loss exponent that is typically in the range $(2,4)$, $\mathbf{H}_t \in \mathbb{C}^{N_r\times N_t}$ is the channel coefficient matrix at time t between the transmitter and the receiver, where $\mathbf{H}_t(i,j)$ is the channel coefficient between the $i\text{th}$ receive and $j\text{th}$ transmit antenna. The $N_t \times 1$ transmit signal vector at time t is $\mathbf{x}[t]$ with unit power constraint, $\mathbb{E}\{\mathbf{x}[t]^\dagger\mathbf{x}[t]\} = 1$, P is the average transmitted power, and $\mathbf{w}[t]$ is additive white Gaussian noise vector with entries that are independent and $\mathcal{CN}(0,1)$ distributed.

Assumption 1.2.1 *Throughout this book, we will use the simple distance-based path-loss function of $d^{-\alpha/2}$ that is valid in far-field, however, has a singularity in the near-field at $d = 0$.*

Assumption 1.2.2 *We will also always assume a flat fading channel, that is, no multi-path $\mathbf{H}_t = \mathbf{0}$ for $t > 0$, for which the signal model* (1.1) *simplifies to*

$$\mathbf{y} = d^{-\alpha/2}\sqrt{\frac{P}{N_t}}\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1.2}$$

where the entries of $\mathbf{H}$ *are assumed to be independent and* $\mathcal{CN}(0,1)$ *distributed to model a rich scattering channel (Rayleigh fading). We also assume throughout this book that matrix* $\mathbf{H}$ *is perfectly known at the receiver.*

To decode the transmit signal $\mathbf{x}$, the optimal decoder is the maximum a posteriori (MAP) decoder that declares that signal to be transmitted which is the most likely signal $\mathbf{x}$ given the knowledge of $\mathbf{y}$. Assuming an uniform distribution over the input signals, MAP decoding is equivalent to ML decoding, where the decoded codeword maximizes the likelihood of $\mathbf{y}$ given $\mathbf{x}$. Mathematically, ML decoding solves the following optimization problem.

$$\max_{\mathbf{x}} \mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{H}).$$

For the signal model (1.2), since each entry of $\mathbf{w}$ is independent and $\mathcal{CN}(0,1)$ distributed,

$$\mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{H}) = \frac{1}{\pi} \exp^{-\left(\mathbf{y}-d^{-\alpha/2}\sqrt{\frac{P}{N_t}}\mathbf{H}\mathbf{x}\right)\left(\mathbf{y}-d^{-\alpha/2}\sqrt{\frac{P}{N_t}}\mathbf{H}\mathbf{x}\right)^{\dagger}},$$

which can be simplified to conclude that the ML decoder decodes vector $\mathbf{x}$ that solves

$$\max_{\mathbf{x}} \mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{H}) = \min_{\mathbf{x}} ||\mathbf{y} - d^{-\alpha/2}\sqrt{\frac{P}{N_t}}\mathbf{H}\mathbf{x}||^2. \tag{1.3}$$

Thus, the ML decoder decodes $\mathbf{x}$, which is the closest codeword to the received signal $\mathbf{y}$ in terms of the Euclidean distance. With ML decoding, all the components of vector $\mathbf{x}$ are decoded jointly, thereby making the complexity exponential in the size of $\mathbf{x}$ which is N_t.

Assuming that the channel matrix $\mathbf{H}$ remains constant for $T \geq N_t$ time slots, and if the transmitter codes across T time slots to send codeword $\mathbf{X}_i = [\mathbf{x}_i[1] \dots \mathbf{x}_i[T]]$, the probability of decoding the codeword matrix $\mathbf{X}_j = [\mathbf{x}_j[1] \dots \mathbf{x}_j[T]]$ instead of $\mathbf{X}_i$ with an ML decoder is [3]

$$\mathbb{P}(\mathbf{X}_i \to \mathbf{X}_j) \leq \left(\prod_{k=1}^{\mathsf{div}} \sigma_k \left(\mathbf{X}_i - \mathbf{X}_j\right)\right)^{-N_r} P^{-\mathsf{div} N_r}, \tag{1.4}$$

where

$$\mathsf{div} = \min_{\mathbf{X}_i \neq \mathbf{X}_j} \{\text{rank}(\mathbf{X}_i - \mathbf{X}_j)(\mathbf{X}_i - \mathbf{X}_j)^{\dagger}\} \tag{1.5}$$

and $\sigma_k \left(\mathbf{X}_i - \mathbf{X}_j\right)$ are the non-zero eigenvalues of $(\mathbf{X}_i - \mathbf{X}_j)(\mathbf{X}_i - \mathbf{X}_j)^{\dagger}$. Thus, to minimize the pairwise error probability (1.4), one has to maximize the minimum of the rank of the difference of any two codeword matrices $\mathbf{X}_i$ and $\mathbf{X}_j$ (1.5). Clearly, with $T \geq N_t$, the maximum value of div is N_t (since $\mathbf{X}_i \in \mathbb{C}^{N_t \times T}, \ \forall\ i$) and for achieving $\mathsf{div} = N_t$, the codewords $\mathbf{X}_i$'s should be coded in space and time; hence the codebook consisting of codewords $\mathbf{X}_i$'s is called a space–time block code (STBC). STBCs with $\mathsf{div} = N_t$ are called full-diversity achieving STBCs, and their error probability is proportional to $P^{-N_t N_r}$. Thus, with multiple transmit and receive antennas, the reliability of a wireless channel can be improved exponentially with the increasing transmission power.

Even though ML decoding provides with the best error probability performance, its decoding complexity is very high because of the joint decoding of all elements of transmitted vector $\mathbf{x}$. Several simple decoders with reduced decoding complexity are also known in literature, for example, minimum mean square error (MMSE) decoder and zero forcing (ZF) decoder. ZF decoder is specially attractive for its simple decoding rule and incurs linear decoding complexity in N_t (the number of elements of $\mathbf{x}$). We describe the ZF decoder in brief and present its error probability performance. We will use the ZF decoder in Chapter 3 to analyze the effects of using multiple antennas in a wireless network.

With ZF decoder, to decode stream $\mathbf{x}(\ell)$ of the transmitted vector

$$\mathbf{x} = [\mathbf{x}(1), \dots, \mathbf{x}(N_t)]^T,$$

the received signal (1.2) is multiplied with a vector $\mathbf{q}_\ell^\dagger \in \mathbb{C}^{N_r \times 1}$, which belongs to the null space of the columns $\mathbf{H}(j), j = 1, \dots, \ell-1, \ell+1, \dots, N_t$ of the channel matrix $\mathbf{H}$, to cancel the inter-stream interference from all other streams $\mathbf{x}(j), j = 1, \dots, \ell-1, \ell+1, \dots, N_t$. With this operation, from (1.2), the resulting signal can be written as

$$\mathbf{y}(\ell) = d^{-\alpha/2}\sqrt{\frac{P}{N_t}}\mathbf{q}_\ell^\dagger \mathbf{H}(\ell)\mathbf{x}(\ell) + \mathbf{q}_\ell^\dagger \mathbf{w}, \tag{1.6}$$

$\forall \ \ell = 1, \dots, N_t$, where there is no inter-stream interference from $\mathbf{x}(j), j = 1, \dots, \ell-1, \ell+1, \dots, N_t$. Thus, with a ZF decoder, each of the N_t data streams of $\mathbf{x}$ can be decoded independently of each other using (1.6), thereby incurring linear decoding complexity compared to the exponential decoding complexity of the ML decoder. This sub-optimal receiver, however, has poor error probability performance because of correlating the noise components in y_ℓ for different $\ell = 1, \dots, N_t$, and the error probability is proportional to $P^{N_r - N_t + 1}$ [4], instead of $P^{-N_t N_r}$ with the ML decoder.

We next discuss the alternative use of multiple antennas in improving the capacity of the point-to-point communication channel. We first define the concept of Shannon capacity, a measure of reliable throughput and show that Shannon capacity increases linearly with the minimum of the number of transmit and receive antennas.

1.3 Shannon Capacity

Definition 1.3.1 *The Shannon capacity C for a communication channel is defined as the largest quantity such that for any rate $R < C$, reliable communication is possible. By reliable communication, we mean that the probability of error can be driven down to zero with increasing block length. Conversely, if the rate of transmission $R \geq C$, the probability of error is lower bounded by a constant.*

Definition 1.3.2 *Let $x[n]$ and $y[n]$ be the input and output of a channel at time n, respectively, then a channel is called a discrete memoryless channel (DMC), if given the most recent input, the output is independent of all previous inputs and outputs, that is,*

$$\mathbb{P}(y[n] \mid x[1], \dots, x[n], y[1], \dots, y[n-1]) = \mathbb{P}(y[n] \mid x[n])$$

for $n = 1, 2, 3, \dots$. Thus, in a DMC, given the input at time n, the output at time n is independent of all the past inputs and outputs.

C. E. Shannon, in his 1948 seminal paper [5], proved that the capacity of a DMC defined by $\mathbb{P}(\mathbf{y}|\mathbf{x})$, with input $\mathbf{x} = [x[1], \dots, x[n]]$ and output $\mathbf{y} = [y[1], \dots, y[n]]$ is given by

$$C = \max_{\mathbb{P}(\mathbf{x})} \mathsf{I}(\mathbf{x}; \mathbf{y}), \tag{1.7}$$

where $\mathsf{I}(\mathbf{x}; \mathbf{y})$ is the mutual information between $\mathbf{x}$ and $\mathbf{y}$ [17]. This result is popularly known as Shannon's channel coding theorem.

Specializing this result for the multiple antenna channel (1.2), when $\mathbf{H}$ is known at the receiver, we have that $\mathsf{I}(\mathbf{x};\mathbf{y}|\mathbf{H}) = \mathbb{E}_{\mathbf{H}}\left\{\log\det\left(\mathbf{I}_{N_r} + \frac{P}{N_t}\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right)\right\}$, and the Shannon capacity of the multiple antenna channel is

$$C = \max_{tr(\mathbf{Q})\le N_t} \mathbb{E}_{\mathbf{H}}\left\{\log\det\left(\mathbf{I}_{N_r} + \frac{P}{N_t}\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right)\right\}, \tag{1.8}$$

where $\mathbf{Q} = \mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\}$ is the covariance matrix of the input signal $\mathbf{x}$. The optimization in (1.8) depends on whether the channel coefficient matrix $\mathbf{H}$ is known at the transmitter (referred to as CSIT) or not (called CSIR). With CSIT, the Shannon capacity [7] is

$$C = \mathbb{E}_{\mathbf{H}}\left\{\sum_{k=1}^{\min\{N_t,N_r\}} \log\left(\xi\sigma_k\left(\mathbf{H}\right)\right)^+\right\},$$

where ξ is the Lagrange multiplier satisfying the power constraint

$$\sum_k (\xi - \sigma_k\left(\mathbf{X}_i - \mathbf{X}_j\right)^{-1})^+ = P,$$

and $\sigma_k\left(\mathbf{H}\right)$ is the kth eigenvalue of $\mathbf{H}\mathbf{H}^\dagger$ indexed in the decreasing order.

On the other hand, with CSIR, when transmitter has no information about $\mathbf{H}$, the Shannon capacity [7] is

$$C = \mathbb{E}_{\mathbf{H}}\left\{\log\det\left(\mathbf{I}_{N_r} + \left(\frac{P}{N_t}\right)\mathbf{H}\mathbf{H}^\dagger\right)\right\}.$$

Thus for large signal power P, with CSIT or CSIR, by using multiple antennas at both the transmitter and the receiver, the channel capacity grows linearly with $\min\{N_t, N_r\}$. The $\min\{N_t, N_r\}$ factor is generally referred to as *spatial degrees of freedom*.

Next, we look at an alternate notion of capacity that is useful for non-ergodic channels for which Shannon capacity is zero.

1.4 Outage Capacity

The Shannon capacity formulation is useful for an ergodic multiple antenna fading channel, where in either each time slot or after a block of T time slots, an independent channel realization of $\mathbf{H}$ is drawn from a given distribution. T is generally referred to as the coherence time of the wireless channel. An ergodic model is valid for fast-fading case, where the fading channel coefficients change fast and the communication duration is long enough to get averaging over multiple independent blocks. Another model of interest is the non-ergodic or the slow-fading channel model, where the channel coefficients vary very slowly. To be specific, with the slow-fading model, it is assumed that at the start of the transmission, an independent realization of the channel matrix is drawn from any given distribution, but then is held fixed for the total communication duration. This model is well suited for low mobility wireless channels requiring short duration communication, where the coherence time is large enough compared to the total transmission time.

It is easy to see that the Shannon capacity of any non-ergodic channel is zero, because with increasing block length no averaging is available, and the error probability is lower bounded by a constant for any non-zero rate of transmission. To have a meaningful definition of capacity for the non-ergodic channel, concept of outage capacity was introduced in [7], which is described as follows. Let B bits/sec/Hz be the desired rate of communication. Then channel outage at rate B is defined to be the event that the mutual information is less than B, and the outage probability is defined as

$$P_{\text{out}}(B) = \mathbb{P}(\mathsf{I}(\mathbf{x};\mathbf{y}) < B)).$$

The outage capacity $C_{\text{out}}(\epsilon)$ is defined to be the maximum rate of transmission B for which the outage probability is below a certain threshold ϵ, that is,

$$C_{\text{out}}(\epsilon) := \max_{P_{\text{out}}(B)\leq\epsilon} B.$$

The outage capacity can be interpreted as the maximum possible rate for which there exists a code whose probability of error can be made arbitrarily small for all but a set of $\mathbf{H}$, whose total probability is less than ϵ. Thus, in essence, outage capacity is the maximum rate which is guaranteed with success probability of at least $(1-\epsilon)$.

The outage capacity formulation naturally extends to a wireless network and will be used to define a throughput metric for a wireless network called the transmission capacity in Chapter 2.

For the multiple antenna channel (1.2), with an ML decoder, the outage probability can be simplified to obtain

$$P_{\text{out}}(B) = \inf_{\mathbf{Q},\mathbf{Q}\geq 0, tr(\mathbf{Q})\leq N_t} \mathbb{P}\left(\log\det\left(\mathbf{I}_{N_r} + \frac{P}{N_t}\mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}\right) < B\right),$$

where $\mathbf{Q}$ is the covariance matrix of the transmitted vector $\mathbf{x}$.

For the most popular Rayleigh channel fading model, where each entry of the channel matrix $\mathbf{H}$ is i.i.d. $\mathcal{CN}(0,1)$ distributed, the distribution of the maximum mutual information expression $\log\det(\mathbf{I}+\mathbf{H}\mathbf{H}^{\dagger})$ is unknown. Consequently, finding the outage capacity of the multiple antenna channel has remained unsolved. The mutual information expression can be significantly simplified if instead of an ML decoder, we use a ZF decoder, where different data streams sent by the transmitter are decoupled before decoding. From [4] for (1.6), with N_t independent data streams, and assuming that each data stream is required to have rate B, and outage probability constraint ϵ, the outage capacity of a $N_t \times N_r,\ N_r \geq N_t$ multiple antenna channel with ZF decoder is

$$C_{\text{out}}^{ZF}(\epsilon) = \max_{\mathbb{P}(\log(1+|g|^2)<B)\leq\epsilon} N_t B, \tag{1.9}$$

where $|g|^2$ is the signal power after zero forcing other $N_t - 1$ signal components and hence $|g|^2 \sim \chi^2(2(N_r - N_t + 1))$. Thus, the outage capacity of the multiple antenna channel with ZF decoder can be found by using the CDF of a χ^2 distributed random variable with $N_r - N_t + 1$ degrees of freedom.

After discussing the point-to-point communication scenario, we next look at the signal interactions in a wireless network, which is of primary interest in this book.

1.5 Wireless Network Signal Model

Consider a wireless network with K nodes, where the nth node's location is denoted by T_n. We assume that each node has N antennas for transmission and reception. The received signal at the

mth node is

$$\mathbf{y}_m = d_{mm}^{-\alpha}\mathbf{H}_{mm}\mathbf{x}_m + \sum_{k=1,k\neq m}^{K} \mathbf{1}_k d_{km}^{-\alpha}\mathbf{H}_{km}\mathbf{x}_k + \mathbf{w}_m, \tag{1.10}$$

where $\mathbf{x}_m$ is the transmitted signal from transmitter T_m, d_{km} and $\mathbf{H}_{km} \in \mathbb{C}^{N\times N}$ are the distance and channel coefficient matrix between the kth transmitter and the mth receiver node, respectively, $\mathbf{1}_k$ is the indicator function that represents whether kth node is active/transmitting, and $\mathbf{w}_m$ is the AWGN vector. We will assume throughout this book that each entry of $\mathbf{H}_{km}$ is i.i.d. and $\mathcal{CN}(0,1)$ distributed to model a Rayleigh fading channel. Scheduling policy of transmitter k defines the indicator function $\mathbf{1}_k$ that critically determines the network performance, since it controls the amount of interference seen at any receiver.

Remark 1.5.1 *In a wireless network, the signal transmitted by the mth node ($\mathbf{x}_m$) could be its own signal or a signal that is being forwarded by it to facilitate communication between some other source–destination pair, in which case $\mathbf{x}_m$ is function of the received signal in previous time slots.*

In a wireless network, there are various source–destination configurations possible, for example, a single node might be interested in communicating with a single node (unicast), few nodes (broadcast), or all nodes (multicast), or two different nodes might be interested in communicating with the same node, or a relay might be helping a single source–destination pair communicate. A wireless network can essentially be broken down into four building block channels that are listed as follows:

- Interference channel: A canonical example of an interference channel is where there are two source–destination pairs that are interested in receiving their own information and do not care about the other pair's data.
- Relay channel: In its simplest form, in a relay channel, a single node (designated relay) helps a single source–destination pair communicate. In more complex form, multiple relays can help multiple source–destination pairs to communicate.
- Broadcast channel: The simplest broadcast channel is where a single source wants to communicate with two destinations, where the information content for the two destinations has both common and private components. Extensions to multiple destinations are also possible.
- Multiple access channel: A multiple access channel is where multiple sources want to communicate with a single destination at the same time.

1.5.1 Information Theoretic Limits of Wireless Networks

From an information theoretic point of view, one of the basic questions is to find the limit on reliable rate of information transfer (Shannon capacity) in a wireless network. In comparison to point-to-point communication, where the Shannon capacity is a scalar quantity, the Shannon capacity of a wireless network is a region spanned by rate tuples corresponding to various source–destination pairs that can be simultaneously supported, such that the error probability can be made arbitrarily small for large block lengths.

Clearly, finding the Shannon capacity of the four basic building block channels discussed above is a prerequisite for finding the Shannon capacity of a wireless network. Unfortunately, the Shannon capacity of the relay channel and the interference channel is unknown, precluding the possibility of finding the Shannon capacity of a wireless network. Today, one of the biggest challenges (some people call it the "holy grail") in information theory is to find the Shannon capacity of a wireless network. A simple upper bound on the Shannon capacity of the wireless network can be found using the Fano's inequality, called the cut-set bound; however, there is no known strategy with achievable rates close to the upper bound.

Primary reason for the intractability of the Shannon capacity of the wireless network is the strict reliability constraint that requires the error probability to be arbitrarily small for large block lengths and complicated signal interaction resulting in interference, which is hard to charcterize. In practice, however, if the SINR seen at the receiver is above a threshold, communication can be deemed successful with sufficient reliability. This SINR model of successful transmission gives rise to the concept of *transmission capacity* [1] and *throughput capacity* [2] that were introduced to understand the fundamental limits on the overall throughout of the wireless network as a function of the number of nodes.

Transmission capacity definition uses the concept of outage probability as a reliability metric. Assuming that all the source–destination pairs are at a fixed distance from each other, the transmission capacity is defined to be the maximum density of nodes per unit area such that the outage probability at each node is below a threshold for a fixed rate of transmission by each node. In essence, given a quality of service QoS constraint (rate of transmission and outage probability), transmission capacity counts the maximum number of concurrently allowed transmissions in a given area. In Chapter 2, we discuss the concept of transmission capacity in detail and derive it using tools from stochastic geometry. We also quantify the effects of multiple antenna nodes, interference cancelation, spectrum sharing on the transmission capacity and bi-directional communication in Chapters 3 and 4. We also highlight the use of stochastic geometric tools to analyze some important performance measures in cellular wireless network in Chapter 5, which are hard to find otherwise.

The alternate notion of capacity, called the per-node throughput capacity for a random wireless network with density n is defined to be $\mathtt{t}(n)$ bits/sec/Hz, if there is a spatial and temporal scheduling strategy, such that each node can send $\mathtt{t}(n)$ bits/sec/Hz on average to its randomly chosen destination with high probability. The network wide throughput capacity is obtained by multiplying the density of nodes n with $\mathtt{t}(n)$. In Chapter 9, we will discuss the concept of throughput capacity and derive the seminal result of [2], which showed that the network-wide throughput capacity of a random wireless network with density n, scales as order $\sqrt{n}$ under the SINR model. The order $\sqrt{n}$ scaling is specific to the SINR model and is not an information theoretic limit. We next derive an information theoretic upper bound of order $n \log n$ on the throughput capacity and then show that a hierarchical cooperation strategy can achieve a throughput capacity of order n by using multi-antenna transmission using distributed antennas of different nodes in Chapter 9.

Both the throughput and transmission capacity yield the same scaling with respect to the number of nodes of the wireless network. Because of the use of the outage probability framework, however, quantifying the effects of advanced physical layer techniques, such as equipping nodes with multiple antennas, using successive interference cancelation, and ARQ, on the transmission capacity is easier than on throughput capacity.

An important distinction between the transmission and throughput capacity is in the averaging of successful transmission event. The transmission capacity uses an outage probability constraint

$\mathbb{P}(\mathsf{SINR} < \beta) \leq \epsilon$ and counts the number of successful transmissions satisfying the outage probability constraint. In contrast, in the throughput capacity definition, we are counting the number of nodes that can simultaneously transmit so that the SINR (realization and not the probability) for each pair of transmissions is above a threshold.

1.6 Connectivity in Wireless Networks

Global network connectivity is more or less a prerequisite for ensuring efficient operation of a wireless network. For example, in a military network, where for obvious reasons, it is imperative to have an active communication link between any pair of nodes. Similarly, in a sensor network, fusion node needs to have a path from each of nodes for data collection and processing. Even though connectivity is desirable, ensuring it is quite challenging, since even a single isolated node breaks the connectivity of the network.

In addition to connectivity, efficient routing protocols that are robust to node failures/outages, intelligent network management tools for transmission scheduling, smart application layer protocols are equally important for a smooth operation of a wireless network. In this book, however, we will restrict ourselves to studying the physical layer properties of the wireless network such as connectivity and refer the reader to [8] for the discussion on higher layer issues such as routing, link management, and scheduling etc.

Connectivity in a wireless network is defined for a variety of link connection models, for example, the disc model, and the SINR model. With the disc connection model, two nodes within a fixed distance are assumed to be connected. The motivation behind this model comes from the radio range of each node—the distance to which each node's signal can be received with sufficient strength. The disc connection model, however, assumes that simultaneously active transmitters do not interfere with each which is an idealization. A more realistic connection model is the SINR model that allows multiple nodes to transmit at the same, where a link between two nodes exists if the SINR between them is above a threshold. SINR model is far more complicated than the disc model, since it gives rise to directed links in contrast to the disc model, and the existence of a link between any pair of nodes depends on the formation of all other links.

Under any connection model, a wireless network can be naturally thought of a graph, where an edge in the graph corresponds to a link in the wireless network. Because of this association, graph theoretic tools, namely, percolation theory is used to study the connectivity properties of the graph. In particular, questions like: what is the minimum radio range required to ensure connectivity in a large wireless network and when does a connected component of unbounded size exists as a function of the density of the nodes are answered using percolation theory. The event of formation of an unbounded component is generally referred to as *percolation* and from a wireless network perspective, percolation guarantees long-range communication possibility. Percolation theory is not only useful for studying connectivity properties, but as we will see in Chapter 9, it is also useful in deriving the throughput capacity of a wireless network.

In Chapter 7, we give a brief introduction to the basics of percolation theory that are required for deriving the results presented in this book. In particular, we describe the basic ideas behind main results in discrete percolation theory over square lattice and hexagonal face lattice, and study some properties of the continuum percolation. In Chapters 7 and 8, we discuss in detail the connectivity properties of the wireless network under the disc model and the SINR connection models, respectively. For the disc model, we derive the critical radio range required for connectivity as a function of the number of nodes. For the SINR model, we show that if nodes use multiple