The Dissertation Committee for Xiangying Yang
certifies that this is the approved version of the following dissertation:

CDMA Ad Hoc Networks: Design and Performance Tradeoffs

Committee:

_____
Gustavo de Veciana, Supervisor

_____
Jeffrey G. Andrews

_____
Ross Baldick

_____
Sanjay Shakkottai

_____
Harrick M. Vin

CDMA Ad Hoc Networks: Design and Performance Tradeoffs

by

Xiangying Yang, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2005

Acknowledgments

The University of Texas at Austin
August 2005

CDMA Ad Hoc Networks: Design and Performance Tradeoffs

Publication No. _____

Xiangying Yang, Ph.D.
The University of Texas at Austin, 2005

Supervisor: Gustavo de Veciana

This dissertation proposes new principles for designing and performance evaluation for spread spectrum based ad hoc networks. We first highlight the advantages of spread spectrum, in the form of Code Division Multiple Access (CDMA), in handling quality of service (QoS) requirements, enhancing energy efficiency, and enabling spatial multiplexing of bursty traffic. Then, based on stochastic geometric models and simulation, we show the ALOHA-like random channel access and 802.11-like simple contention and handshaking based schemes are poor at achieving good capacity or efficient spatial reuse, especially under bursty and heavy load. We show that this is because the closest interferers severely penalize the performance of the network, particularly for a direct sequence CDMA (DS-CDMA) system. Therefore, it is necessary to reconsider system design for spread spectrum ad hoc networks. To this end, we consider improving system performance at different network layers. At the physical layer, we first propose to use interference cancelation techniques, in particular, successive interference cancelation (SIC), at receivers to handle strong nearby interferers. Our analysis not only shows the significant improvement on capacity from SIC but also indicates that just canceling a few nearest interferers will provide most of the performance gain. Therefore, SIC is particularly suitable for DS-CDMA ad hoc networks to enhance capacity, incurring only a small amount of extra complexity. In addition, at the MAC layer, we show how idealized contention resolution among randomly distributed nodes results in clustering of successful transmitters and receivers, in turn leading to efficient spatial reuse. This motivates explicitly inducing clustering among contending nodes to achieve even better spatial reuse. We propose two distributed mechanisms to realize such clustering and show substantial capacity gains over simple random access/ALOHA-like and even RTS/CTS based protocols – on the order of 100-700%. We examine under what regimes such gains can be achieved, and how clustering and contention

vi

resolution mechanisms should be optimized to do so. We further extend our MAC design for inducing clustered contention in ad hoc networks to support hop-by-hop relaying on different spatial scales. By allowing nodes to relay beyond the set of nearest neighbors using varying transmission ranges (scales), one can reduce the number of hops between a source and destination so as to meet end-to-end delay requirements. To that end we propose a multi-scale MAC clustering and power control mechanism to support transmissions with different ranges while achieving high spatial reuse. The considerations, analysis and simulations included in this thesis suggest that the principle of inducing spatial clustering in contention has substantial promise towards achieving high spatial reuse, QoS, and energy efficiency in spread spectrum ad hoc networks.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Wireless ad hoc networks are wireless communication systems with little or no infrastructure support. The applications for wireless ad hoc networks have expanded in recent years to include numerous military applications and the emerging wireless sensor networks, as well as many other interesting and commercially viable applications including mesh backhaul for wireless broadband and range extension for cell-based networks, e.g. [25, 7]. Given reliable and efficient ad hoc networks, in addition to existing Internet applications, promising future application paradigms such as Ubiquitous Computing and Grid Computing are likely to thrive. Despite the high level of interest and commercial potential, fundamental ad hoc network design principles are still not well understood, in part, due to difficulties in devising acceptable models capturing the salient characteristics of such system.

Current research efforts on wireless ad hoc networks involve new system designs at all layers. Different physical layer choices have been proposed for use in ad hoc networks in addition to the classic narrow-band communications, including spread spectrum or Code Division Multiple Access (CDMA), directional antenna (beam forming) and multiple-in multiple-out antenna (MIMO). Different physical layer choices will dramatically impact MAC layer design and even routing protocols due to their different characteristics, and more importantly lead to significant differences in network performance in terms of capacity, QoS support, etc., and type of applications that can be supported.

In this dissertation, we motivate the use of a CDMA physical layer in ad hoc networks and highlight benefits in terms of performance. To facilitate our analysis and performance evaluation, we narrow the scope of our work as follows.

- We assume a simple channel model with only path-loss attenuation and ignore shadowing and fast fading. In particular, if the transmitted power is ρ and the path-loss exponent is $\alpha > 2$ then the received power at a distance $d > 1$ from the transmitter is $\rho d^{-\alpha}$. While channel fluctuations are critical in wireless communications, these usually can be mitigated by a fast closed-loop power control and advanced receiver

design leading at worst to a graceful degradation of system performance.

- Devices are assumed to be equipped with an omnidirectional antenna which can be used in one of the following operation modes: transmit, receiver, idle or sleep (power saving mode). This is a worst case scenario assuming only the simplest/cheapest hardware is available at each node. More complicated hardware will enable a larger design space and thus generally improved performance. In addition, we assume each device is only equipped with a single transceiver. This limitation is natural given cost and size considerations. The implication is that a typical device can only operate in one of these states at any given time.

- Node mobility is not considered at the timescale of packet transmissions that we consider, typically measured in *msec*. This is much smaller than the timescale of common human or vehicle mobility, typically measured in *sec*. In addition, many applications like residential mesh networks may not be subject to mobility at all.

- Throughout this dissertation, we consider dense networks subject to random traffic patterns, i.e., sources/destinations and routing strategies are randomly generated/selected such that transmissions can be modeled as random in the network. We do not assume particular routing strategies or protocols in this dissertation, although routing strategies are implicitly reflected on the selection of transmission ranges.

Organization of this dissertation. In Chapter 2, we briefly review the key design challenges, fundamental limitations of classic system design - in particular we focus on MAC designs for a narrow-band ad hoc network. In Chapter 3, we motivate the use of a CDMA physical layer in ad hoc networks to enable the realization of tradeoffs among multiple key performance metrics. In Chapter 4, we take an analytical approach to evaluate the performance of existing designs of CDMA ad hoc networks and show that the performance and efficiency, particularly in terms of spatial reuse, of previous designs may not be satisfactory in a heavily loaded dense ad hoc network. This motivates the need for more efficient, robust system designs for such networks. In Chapter 5 a physical layer approach based on advanced receivers, capable of canceling interference, is proposed so as to significantly improve capacity. In Chapter 6, we propose a novel MAC design principle and practical protocol design based on inducing spatial clustering of contentions in DS-CDMA ad hoc networks to further enhance the spatial reuse.

Chapter 2

Basic Understanding of Ad Hoc Networks

In this chapter, we begin by discussing the fundamental problem in ad hoc networks - interference among transmissions - and how it limits capacity. We focus on the case where a narrow-band network is used.

2.1 Interference in ad hoc networks

Dense wireless networks are by nature interference-limited, which means that increasing the transmit power of all nodes in the network simultaneously will not substantially increase the overall throughput of the network. Ad hoc networks pose a particularly challenging interference environment because of the lack of agreed-upon centralized receivers (as in a cellular network) means that one of the major challenges is to ensure successful transmissions given the ad hoc nature of interferers. Therefore, one can expect that for such distributed networks, since scheduling transmissions is a challenging task and capacity is severely constrained by interference, capacity will be one of the key performance metrics one should consider in the network design.

2.1.1 Terms related to interference

We first review a few useful terms that will be used in this dissertation. They are all related to interference in wireless communications but not limited to wireless ad hoc networks.

The near far problem. As shown in Fig. 2.1, because the signal power attenuates with distance, receiver *B* will receive a much stronger transmission power from nearby transmitter *C* than those that are far away, e.g. transmitter *A*. Thus intuitively, severe path loss

or large differences in transmission ranges lead to what is called the near-far problem. In cellular systems, such discrepancies are taken care of by performing power control between the base station and mobile devices so as to equalize the receive power at the base station. However in an ad hoc network, this problem is much harder to solve since transmitters do not share a centralized receiver as they do in a cellular network. For example, consider the case that transmitter C transmits to some node but not A. C may perform power control but only with respect to its intended receiver instead A and thus C may still severely interfere with transmission $A \rightarrow B$.
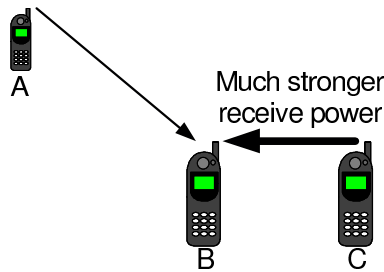
Figure 2.1: Near far problem. Since C is much close to B than A, due to path loss, , the receive power from C is much stronger than that from A at receiver B.

Capture. The near far problem is not always deleterious. In Fig. 2.1, if C indeed intends to transmit to B, B may be able to successfully receive from C even under A's interference. This is called 'capture', referring to successful reception of the stronger signal even when it 'collides' with one or more weaker signals.

The hidden terminal problem. The hidden terminal problem is related to the near far problem and is exhibited on the left of Fig. 2.2. When A transmits to B, C may not hear this transmission since C is far away from A. If C chooses to transmit to D, C may severely interfere with receiver B. As shown on the right of Fig. 2.2, one possible approach to dealing with this problem is to use request-to-send(RTS) and clear-to-send(CTS) handshaking before actual transmissions. A sends a RTS message to C and notifies C of the intended transmission. If C successfully receives the RTS message, it will reply back with a CTS message. The CTS message not only acknowledges that A can proceed with its transmission, but also informs nearby potential interferers about the ongoing transmission. Potential interferers who also intend to transmit, e.g., C, thus will back off for at least the transmission time to avoid collision. The hidden terminal problem can be difficult to tackle when nodes are using unequal transmission power levels, where RTS and CTS signaling may not fully eliminate hidden nodes. This will be discussed in more detail in Section 2.1.2.
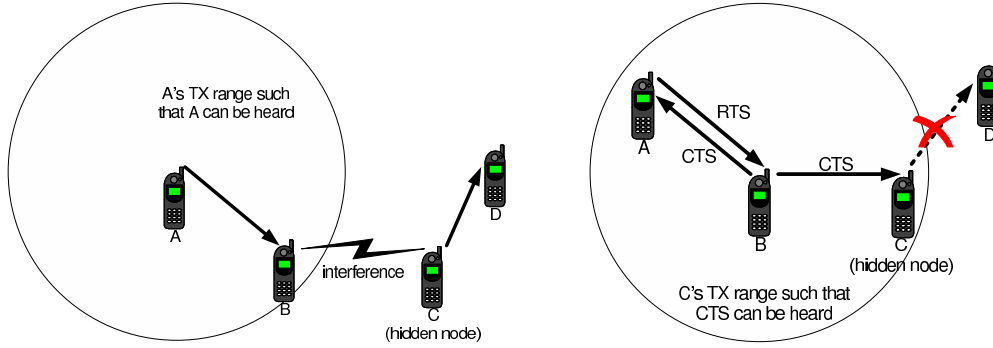
Figure 2.2: Hidden terminal problem. On the left, *C* may interfere with transmission $A \rightarrow B$ since it is not aware of this ongoing transmission. On the right, assuming roughly equal transmission power in the network, the hidden terminal problem can be solved by RTS/CTS handshaking, which allow *C* to be aware of transmission $A \rightarrow B$ and choose to back off.

2.1.2   Modeling interference and outage

In the sequel, we consider two models for successful transmissions in an ad hoc network. Following [16], they are referred to as the *physical model* and *protocol model*. The major difference between these two models is the way they capture the interference seen by a typical receiver.

Physical model

As shown in Fig. 2.3, in the physical model, we consider a typical receiver located at origin *O*, with an associated transmitter at a distance *d* away. Assume that a set of concurrent transmit nodes (including nodes relaying packets), i.e., interferers, are located at $\phi = \{x_i\}$ on the plane. Nodes are interchangeably referred to/by their locations. Let $|x_i|$ denote the distance from node $i \in \phi$ to the origin. Transmitter *i* uses a fixed power level ρ_i. We capture the spatial attenuation of signal power using a basic path loss model where if a transmitter uses a power level ρ the receive power at a distance *d* is given by $\rho \times d^{-\alpha}$. The path loss exponent α is typically assumed to be between 3 and 5.

The ambient noise density is denoted by N_o. Thus the total ambient noise power is $N_o B \equiv \eta$, where *B* is the total available bandwidth. Note that in a dense network with relatively heavy traffic, the network is interference limited and one may ignore the ambient noise in this scenario.

The typical receiver at *O*, sees the degraded powers from *all* other concurrent transmitters as interference. An outage happens when the SINR at the receiver falls below a

certain threshold β, resulting in an unsuccessful transmission, i.e.,

$$\frac{\rho d^{-\alpha}}{\eta + \sum_{x_i \in \phi} \rho_i |x_i|^{-\alpha}} \leq \beta. \tag{2.1}$$
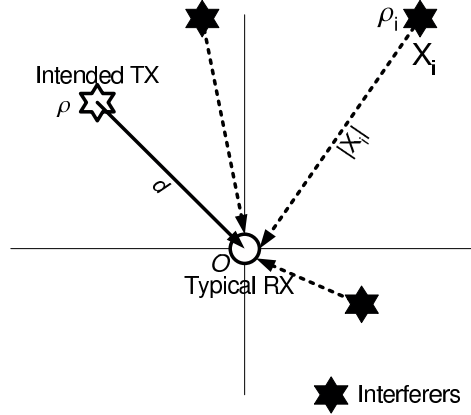
Figure 2.3: In the physical model, a typical receiver at O is interfered by all concurrent transmitters.

Note that ϕ only includes interfering nodes and β is determined by the desired data rate and modulation scheme that are selected. When there are a large number of nodes in space, calculating the exact interference term $\sum_{x_i \in \phi} \rho_i |x_i|^{-\alpha}$ in (2.1) may incur difficulty in analysis. There are, however, more intuitive ways to understand interference and outage. We will discuss an approximation to the physical model, called the protocol model, in the sequel.

Protocol model

Let us consider the rough geometry of transmission and interference ranges. Note that the aggregate interference is a sum of non-i.i.d. random variables. Interferers in the 'near' field of a typical receiver contribute very strong interference. 'Far' field interferers may still potentially contribute enough aggregate interference to cause an outage although each transmitter only contributes a small amount due to path loss. To quantify near/far fields, we define the *critical interference range r* to be the largest distance from a receiver at which a *single* interferer could be located and cause an outage, i.e., the largest r such that

$$\frac{\rho d^{-\alpha}}{\rho r^{-\alpha}} \leq \beta \quad \Longrightarrow \quad r = (\beta)^{\frac{1}{\alpha}} d. \tag{2.2}$$

Note that in a narrow-band system, usually $\beta > 1$ and consequently $r > d$. Thus to ensure successful transmission, at least a disc of radius r around each successful receiver should

contain no transmitter - this is shown in Fig. 2.4. Note that the suppression range of radius r is only a *necessary* condition for successful reception because far field interferers may still contribute an aggregate interference high enough to cause an outage. However as discussed in the sequel, under some typical conditions, a suppression range of radius r is indeed close to a sufficient condition. One can also enforce a margin Δ over r such that this model better approximates the previous physical model.
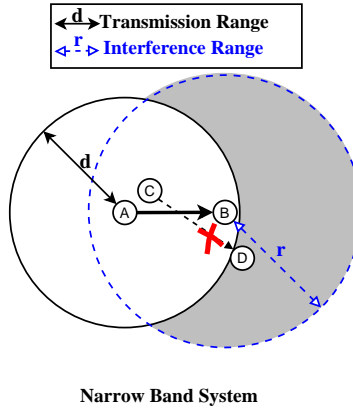
Figure 2.4: In a narrow-band system, ensure transmission $A \rightarrow B$ to be successful, no interfering transmission is allowed in the neighborhood of receiver B within distance r.

Thus in the protocol model, a transmission is considered unsuccessful, i.e., an outage occurs, if there is at least one interferer within the distance $r + \Delta$, $\Delta \geq 0$ of the intended receiver.

Accuracy of the protocol model

There are several aspects that will impact the accuracy of the protocol model. A closer look at the protocol model leads to a better understanding of the nature of interference in an ad hoc network.

Path loss. Path loss dictates how severe the near-far problem will be. Nearby interferers become more dominant in terms of their contribution to the overall interference power with larger path loss. Therefore, one can expect the protocol model to be more accurate in a system where the path loss exponent is large, say larger than 3. Indeed as we shall see later in Chapter 4, in a network with nodes having random locations, roughly a fraction of $\frac{\alpha - 1}{\alpha}$ outage events are caused by nearest interferers and are thus captured by the protocol model.

Node distribution. In an environment with severe path loss, usually nearby interferers make the dominant contribution to the interference power. However, this is based on the

assumption that the interferers are randomly located and the spatial distribution of interferers is fairly homogeneous. Without any of these conditions, the protocol model is unlikely to be accurate. For example, one can consider an extreme (but maybe rare in reality) case where many interferers are outside but very close to the interference range of a receiver and thus it is still sufficient to cause an outage.

Transmission power. When nodes transmit with heterogeneous power levels, the protocol model may no longer be useful. As shown in Fig. 2.5 a transmission with low power has no interferers in its 'near-field'. However a 'far-field' transmission, using high transmission power, may still severely interfere with low-power transmissions, i.e., the definition of near-field/far-field is directly affected by transmission powers, e.g., see Eq. (2.2), and thus we do not have unified notions or abstractions that make sense at the network level when heterogeneous transmission powers/ranges are being used. This type of hidden node problem arises when nodes are using aggressive power control, and as mentioned previously, poses a big challenge for system design, in particular at MAC layer.
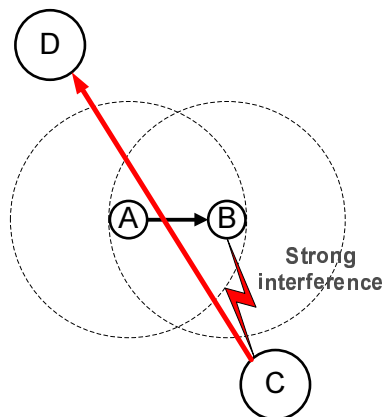
Figure 2.5: A hidden terminal problem when aggressive power control exists. Even all transmissions perform RTS/CTS handshaking before transmitting actual data, long range transmissions with strong power may still severely interfere low power transmissions, who have very small transmission range such that others may not be aware of their RTS/CTS.

Outage regime. Finally, even if the above conditions hold, such that using the protocol model is appropriate, the accuracy still depends on the outage regime of interest. Generally, in a low outage regime, the protocol model is fairly accurate. In a high outage regime, with the aggregated far-field interference being closer to sufficiently causing outages, the protocol model becomes less accurate because it totally ignores far field interference. Note that the outage regime's impact may not be as critical as other aspects mentioned above.

Therefore, one must choose a proper interference model for different scenarios. Although the protocol model is usually easier to analyze, the context will determine its

accuracy.

2.2 Fundamental limits on the capacity of ad hoc wireless networks

2.2.1 Previous results on wireless network capacity

Given their interference limited nature, ad hoc networks are fundamentally capacity limited. There have been some notable recent results on ad hoc network capacity [16, 47, 15, 11]. Gupta and Kumar [16], for example, established that the *transport capacity* of an ad hoc network, defined as the number of bit-meters carried over a given time interval for a network of nodes occupying a unit area, is $O(\sqrt{\lambda})$, where λ is the density of transmitting nodes. Their "physical" model takes the form of an SINR requirement, i.e., the ratio of signal power over interference plus noise power must exceed some threshold, with powers measured at the receiver. However, their analysis focuses on a deterministic SINR model, employs a deterministic channel access scheme and, thereby, precludes the occurrence of outages. By contrast, in order to accurately model the behavior of a distributed ad hoc network at the physical and MAC layer, our model includes a stochastic SINR requirement coupled with random channel access. That is, under a random distribution of transmitters, the probability of the SINR ratio being inadequate for successful reception must be below some constant ε, which we call the outage constraint.

Taking this a step further, more recent work [55, 21, 54] has shown that the scaling of transport capacity depends on the amount of attenuation in the channel. Roughly speaking, in the low attenuation regime with no channel absorption and small path loss, the transport capacity can be unbounded even under a fixed power constraint, by using coherent relaying and interference subtraction. On the other hand, in the high attenuation regime with channel absorption or high path loss, the transport capacity is bounded by the total available power and thus scales as $\Theta(n)$ when the n nodes in the network are individually power constrained, regardless of channel fading [56]. In particular, the low path loss versus high path loss regimes are classified as follows. [55] first identifies the different scaling of transport capacity in different channel attenuation regimes. Consider the channel attenuation model of $\frac{e^{-\gamma r}}{r^{\alpha}}$, where r denotes a distance, γ is the absorption exponent and α is the path loss exponent. For the 2-D network case, if $\gamma > 0$ OR $\alpha > 3$, the capacity is bounded by total transmission power and thus linearly in the total number of nodes when individual node is power constrained; if $\gamma = 0$ and $\alpha < 3/2$, the transport capacity may possibly be unbounded even with fixed transmission power. [56] shows when $\gamma > 0$ or $\alpha > 3$, (roughly) the same scaling property for transport capacity holds even under general models of fading channels. [21] improves the above work by extending the feasible attenuation regime,

9

where the transport capacity scales linearly in total number of nodes, to $\alpha > 5/2$. [54] further improves the linear scaling regime to $\alpha > 2$, under the condition that attenuation phases between node pairs are independently and uniformly distributed. Otherwise, when the phases are arbitrary, the linear scaling regime of capacity remains to be $\alpha > 5/2$.

This dissertation mostly works with the high path loss regime. Because we assume transmissions without joint decoding, we will usually let $\alpha > 2$ capture our high path loss regime. Our results will also show that the capacity of CDMA ad hoc networks is sensitive to the channel path loss, and in fact as we shall show later in Chapter 3 the different types of CDMA behave quite differently.

2.2.2 Interpreting the scaling property of ad hoc network capacity

One of the basic insights provided by recent work on the capacity scaling of ad hoc networks is that it is maximized when traffic is relayed along nearest neighbor paths to a destination. Indeed it turns out to be better to maximize the density of concurrent transmissions in order to achieve a maximum amount of forward progress, i.e., bits-meter/sec. The intuitive reason for this is that transmissions take up space, i.e., they cause a zone around the transmitter where other nodes cannot successfully decode from a different transmitter. Since the number of nodes affected is quadratic with the transmission radius d, i.e., about πd^2, and progress of information is linear in d, smaller transmission ranges will increase the overall network throughput.

The result of [16], in addition, indicates the fundamental limit on transport capacity will not change by using multiple channels, e.g., as long as the total available bandwidth B is fixed, TDMA, FDMA and CDMA optimally all achieve the same theoretical capacity.

Given the above result and intuition, a natural question to ask is whether the use of a different physical layer, say CDMA, in ad hoc networks is worthwhile? We will answer this in the next chapter.

Chapter 3

Why Use a CDMA Physical Layer in Ad Hoc Networks?

In this chapter, we motivate the use of a CDMA physical layer in ad hoc networks. We show the promise of a CDMA physical layer in terms of achieving flexible tradeoffs among different performance metrics, which a simple narrow-band physical layer may not be able to achieve.

3.1 Other important performance metrics for ad hoc networks

In practice, maximizing capacity is but one of a myriad of possible design goals, depending on the physical constraints and intended applications. At least two other objectives are critical in ad hoc network applications: quality of service (QoS) and energy efficiency. For example, in applications supporting delay sensitive traffic like voice, one should ensure certain upper bounds on end-to-end delay; in a network consisting of battery operated devices, one should balance the traffic load and conserve energy. In order to design such networks one must be able to appreciate tradeoffs among the various figures of merit. One should not, for example, consider system capacity without tieing it to energy efficiency, or, for some applications, consider 'capacity' without an understanding of the quality of service, e.g., delays, an end-to-end transmission will incur. One can visually understand these tradeoffs as shown in Fig. 3.1, in which the performance tradeoff is also bounded by the design complexity at different network layers. Thus network designers will need to in particular consider tradeoffs among (but not limited to) the following key performance metrics:

- Capacity,

- End-to-end delay,

11
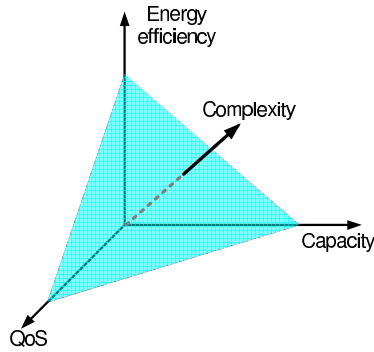
Figure 3.1: Performance tradeoffs among capacity, energy efficiency, QoS and system complexity.

- Energy efficiency,

- Load balancing,

- Design complexity.

We will discuss some of these key performance metrics in the sequel and expose the interaction among them. In particular, although nearest neighbor routing maximizes the transport capacity, long range relaying may still be appealing to enhance other key performance metrics.

QoS in terms of end-to-end delay. In practice when nearest neighbor routing is used in order to maximize capacity, a packet may need to be relayed by a large number of nodes prior to reaching its destination. Each intermediate node would typically incur a delay, depending on the MAC protocol's contention overheads, making it difficult to meet end-to-end delay requirements. For example, as shown in Fig. 3.2, assume each hop roughly incurs a delay τ due to queueing and MAC contention, using longer hops can significantly reduce end-to-end delay. In addition, [37] shows that nearest neighbor routing will likely cause bottlenecks in the network topologies where a few critical nodes have high degree of connectivity. Hence traffic with competing routes going through these bottlenecks will incur large queueing delays due to the congestion at these nodes.

Energy efficiency. When using nearest neighbor routing, intermediate nodes would typically be switching among transmit, receive and idle modes, further decreasing the amount of time they can spend in the sleep mode. Depending on the actual energy characteristics of the nodes the first three modes can be fairly energy hungry[42][43], see Table. 3.1. An alternative would be to permit nodes to use longer transmission distances and relaying
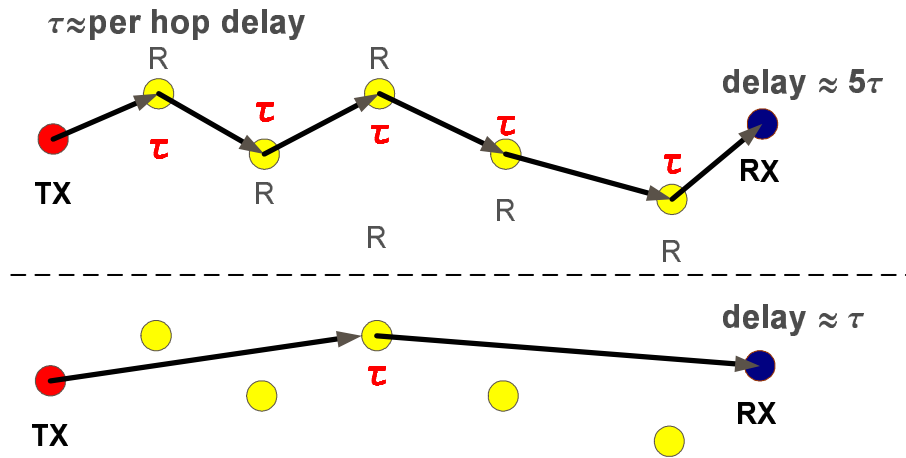
Figure 3.2: Using longer hops reduces end-to-end delay.

through nodes that are in a larger 'neighborhood' as shown in Fig. 3.3. This would allow for paths with fewer hops between the source and destination, possibly enabling a larger number of nodes to dwell in the energy efficient sleep mode. Suppose the typical energy consumption for TX, RX, Idle and Sleep modes are 10(long range transmission)/1(short range transmission), 1, 1 and 0, respectively. Consider a packet is relayed from the source to the destination. Using nearest neighbor routing, it takes 6 transmissions and all 7 nodes are actively operating which lead to an energy consumption of $6 \times 7 = 42$ units. Using long range transmission, even each transmission uses 10 times the energy compared to the case of nearest neighbor routing, one can similarly calculate the energy consumption is only $2 \times 10 + 4 = 24$ units.

Route efficiency and reliability   In an ad hoc network, route reliability and maintenance is critical because if a route is broken, route discovery can be expensive in terms of energy consumption and delay. Many factors can lead to poor route reliability as shown in Fig. 3.4. Nodes might move fast and be out of reach. Nodes may be off due to insufficient battery power. Even if a node is stable in terms of mobility and energy, it may be congested when carrying a heavy load of traffic, leading to excessive delay. With long range transmission, more nodes are available within longer transmission range as candidates for relaying. Thus routing protocols may exploit this to achieve better load balancing and route reliability, by choosing nodes with light traffic load, sufficient battery energy and slow/no mobility.

Table 3.1: Energy consumption of wireless transceivers and devices

Device	TX	RX	Idle	Sleep
Lucent WaveLan Turbo 11Mb card	285mA	185mA	-	9mA
RoamAbout 915 MHz DS/ISA	600mA	300mA	-	36mA
RoamAbout 2.4 GHz DS/ISA	365mA	315mA	-	30mA
2.7V GSM RF Transceiver	31mA	42mA	-	1μA
Nokia C020/C021 wireless LAN card	1.7W	1.3W	-	0.1/0.2W
AT&T WaveLan 915MHz	-	-	1318.9mW	177.3mW
AT&T WaveLan 2.4GHz	-	-	1148.6.9mW	143.3mW
Metricom Ricochet wireless modem	-	-	346.9/431.0mW	93.5mW
Apple Newton PDA	-	-	1187.8mW	164.2mW
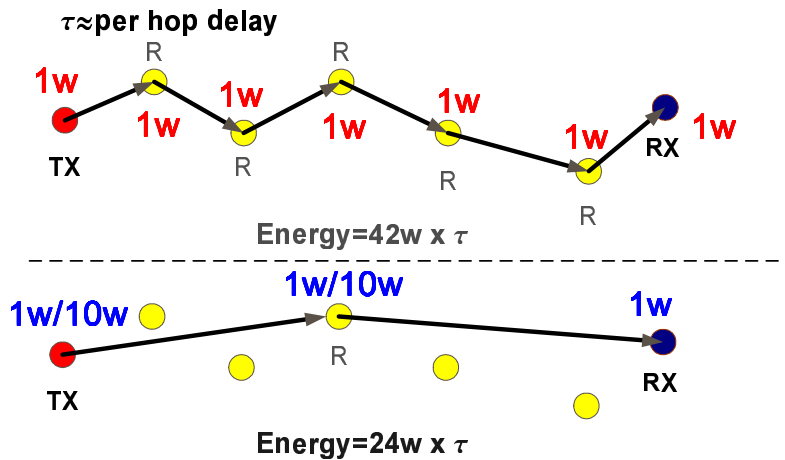
Figure 3.3: Using longer hops can reduce overall energy consumption by allowing more nodes to stay in the energy efficient Sleep mode.
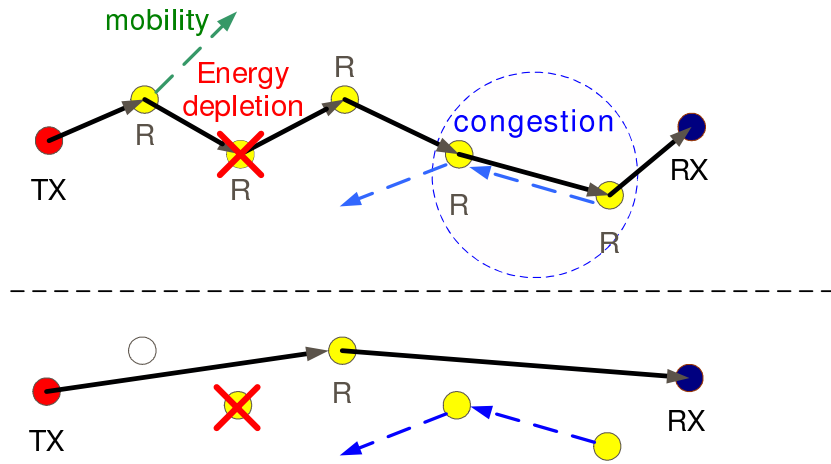
Figure 3.4: Using longer hops helps to maintain route reliability by skipping nodes that are not suitable for relaying traffic.

3.1.1   Why narrow-band is inflexible for realizing performance tradeoffs?

As can be seen in a dense network, for several of the key metrics, it may be worthwhile to abandon the mandate of nearest-neighbor routing, even though this might maximize network capacity. However, long range relaying may not directly lead to efficient tradeoffs among different performance metrics. Long range relaying might require higher transmission power levels for successful reception, and increase interference, possibly compromising network capacity, since more nodes are affected by the associated higher power levels – this is particularly the case in a *dense* narrow-band network. When traffic is bursty, additional contention is required to serialize transmissions, e.g., back-offs and retransmissions, which might lead once again to poor delay performance and energy waste. A narrow-band physical layer thus appears to be poor at allowing tradeoffs between capacity and the other performance metrics.

An approach to achieving a high degree of spatial reuse which enables concurrent overlapping of transmissions and flexible resource allocation among traffic is thus desirable. We see this as one of the key motivations for using a CDMA physical layer in ad hoc networks.

3.2   Why use CDMA in ad hoc networks?

3.2.1   Brief overview of CDMA technology in cellular networks

Spread spectrum in the form of CDMA has proven very robust in interference-limited cellular networks. The central tenant of CDMA is that since dense networks are interference-

limited, designing for time or frequency orthogonality (as in TDMA or FDMA) is not appropriate, since other-cell interference and other imperfections compromise the orthogonality anyway. On the other hand, CDMA tolerates all sources of interference within reasonable bounds by leveraging the high spreading gain. Due to its robustness, achieved capacity, and other implementation and political factors, CDMA is currently the underlying physical layer technology for all three of the important third generation cellular standards: CDMA2000, WCDMA, and TD-SCDMA. Based on this success, it is natural to seriously consider the viability of CDMA for emerging mesh networks.

CDMA techniques have historically been divided into two very different types of modulation: frequency hopping and direct sequence.

In frequency hopped CDMA (FH-CDMA), the total bandwidth B is divided into m orthogonal[1] frequency bands of bandwidth B/m. At each time instant, the transmitter chooses one of the m bands with equal probability, based on a pseudorandom code sequence that is also known to the receiver. Assuming the transmitter and receiver are synchronized, they both hop in unison and are able to successfully communicate. If there are other users in the network, there will be occasional collisions when two transmitters pick the same frequency band, but by coding over time, it may be possible to recover from a moderate number of collisions. Examples of well-known systems that use frequency hopping include Bluetooth, which has 80 frequency bands of 1 MHz width ($m = 80$, $B = 80$ MHz), and a hop interval of 625 microseconds, and GSM, which has a variable number of possible frequency bands of width $B/m = 200$ KHz, and a hop interval of 4.617 msec.

Direct sequence CDMA (DS-CDMA), which is often referred to simply as CDMA, also involves synchronized pseudorandom codes, but in this case the code sequence of bandwidth B is multiplied with the user's data sequence of bandwidth B/m, creating a transmitted sequence of bandwidth B. Since the transmitted signal is "spread" to a bandwidth m times larger than the original transmit sequence, m is called the spreading factor or processing gain. By correlating the same code with the received signal, the desired signal is converted back to a narrow-band signal (i.e. bandwidth B/m) while the noise and interference remain wideband (i.e. bandwidth B) and hence is attenuated by a factor of approximately m at detection. The previously mentioned third generation cellular standards all employ a version of DS-CDMA for multiple access.

Despite their obvious large differences, DS-CDMA and FH-CDMA have very similar properties when used in cellular systems and achieve comparable SINR's for the same system load [39]. Both offer effective "*interference averaging*" capability, i.e., the ability of a receiver to decode a signal in the presence of a substantial number of *concurrent* transmissions, so that a system can be designed for the average, rather than worst case, interference. For example, frequency hopping in GSM allows users near the cell boundaries to avoid

---

[1]In practice, the frequency bands are not strictly orthogonal.

16

the situation where the nearest interfering base station is always transmitting in the same frequency channel as its desired base station.

3.2.2 Benefits in CDMA on ad hoc networking

One can certainly expect similar benefits when using CDMA in an ad hoc network. However, in contrast with the context of cellular systems, FH-CDMA and DS-CDMA have very different characteristics when used in an ad hoc network.

3.2.3 Change of considerations related to interference

In a narrow-band network, for a transmission to be successful, an interference range r should be clear of interferers, see the left panel of Fig. 3.5, in order to allow an acceptable received SINR at the receiver[2]. Since the desired received energy should generally be higher than the received interference, most work on ad hoc networks has assumed that r is larger than the transmission range d, which thus excludes the possibility of spatially overlapped transmissions.

     In contrast, for a CDMA network this restriction is relaxed and spatially overlapped transmissions are allowed, albeit quite differently for FH and DS. For FH-CDMA, although the interference range does not change, as long as there are no transmissions in the same frequency band, multiple transmissions in the same vicinity will be successful. For DS-CDMA, receivers use spreading gain to reduce the required receive SINR threshold for successful transmission by m, which allows the interference range r defined in (2.2) to be reduced in proportion to $m^{1/\alpha}$. To see this, we redefine the *interference range r* in a DS-CDMA ad hoc network to be the largest distance from a receiver at which a *single* interferer could be located and still cause an outage even after being mitigated by a spread gain m, i.e., the largest r such that

$$\frac{\rho d^{-\alpha}}{\rho r^{-\alpha}} \leq \frac{\beta}{m} \implies r = (\frac{\beta}{m})^{\frac{1}{\alpha}} d. \tag{3.1}$$

Hence, for large enough m, the interference range r is less than the transmission range d, which will allow spatially overlapped transmissions, an example of which using DS-CDMA is shown on the right panel of Fig. 3.5.

     The ability of the CDMA physical layer to handle concurrent transmissions that may be spatially overlapping is a fundamental difference between narrow-band and spread spectrum ad hoc networks. Although the degree of such dense spatial reuse is limited by the spreading factor and these concurrent transmissions are not likely to be numerous enough to compensate for the bandwidth efficiency loss due to spreading, as we shall see later, the

---

[2]The same interference range r around transmitters are for successfully receiving ACK from receivers
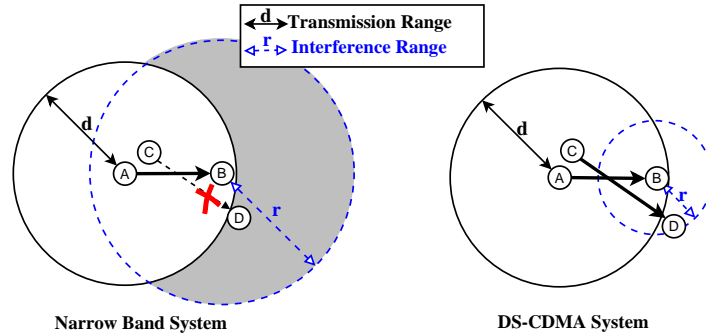
Figure 3.5: The transmission range and interference range for an idealized narrow-band and spread spectrum system. On the left, since $A \rightarrow B$ requires no concurrent transmission in the critical interference range r around a B, $C \rightarrow D$ is not allowed in a narrow-band system, which however may be allowed in the CDMA system shown on the right.

possibility of spatially overlapped transmissions at the physical layer will allow for different design paradigms at the higher network layers, which may afford flexibility in trading off between the different metrics discussed in Section 3.1.

3.2.4 Flexible and efficient tradeoffs among performance metrics.

In a regime where (some) nodes use longer transmission ranges, with a view on meeting QoS service requirements and/or better energy efficiency, a CDMA based physical layer has some key advantages – some of which are akin to those already exploited in cellular networks; some of which are thanks to the efficient spatial reuse, i.e., overlapping transmissions, enabled by the interference averaging ability:

'Spatial multiplexing'. Because receivers can tolerate fluctuations in interference, the network can statistically multiplex concurrent overlapping bursty traffic, e.g., on/off voice streams. A general example of such 'spatial multiplexing' is shown in Fig. 3.6, where applications sharing the network may have different QoS requirements and possibly require different relay scales, e.g., delay-sensitive applications may prefer to use long relay distances and high transmission power to achieve reduced end-to-end delays while best effort traffic can use shorter relay distances and low transmission power allowing for enhancing overall throughput. Such multiscale spatial multiplexing achieves efficient spatial/spectrum reuse and flexible resource allocations, even with heterogeneous flows.

Power control. The ability to average interference, can also help in managing spatial inhomogeneities in nodes' locations, and permit distributed power control to conserve energy with only graceful degradation on performance – e.g., when the power levels among various transmit nodes are not optimally selected due to the lack of central management/global information in an ad hoc network.

18

**Transmission range for
relaying best effort pkts**

**Relay path for a delay
sensitive stream**

**Transmission range for relaying
delay sensitive stream pkts**

**Relay path for a
best effort stream**
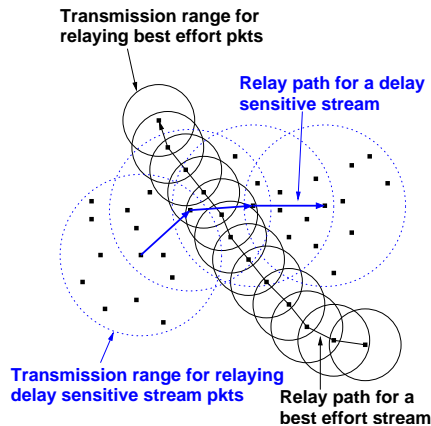
Figure 3.6: An example of 'spatial multiplexing', i.e., heterogenous traffic with different QoS requirements, using different transmit ranges, and co-existing on an ad hoc network. Note that adjacent hops can not be active at the same time because we assume nodes can only either transmit or receive.

Robustness. Critical problems such as hidden and exposed terminals seen in narrow-band systems are no longer as prominent in spread spectrum based ad hoc networks, while the 'near-far' problem can be mitigated via spreading gain and proper power control. These advantages may potentially simplify MAC design and operation.

While CDMA provides a richer design space, one may ask whether this will fundamentally change the paradigms that network designer should follow, in particular at the MAC layer? If so, what would be the desired design principles in CDMA ad hoc networks? We will answer these questions in the rest of this dissertation.

3.3 Appendix: Code Assignment

Code assignment is not the topic of this dissertation. However, because we consider a network based on CDMA, we must understand how to assign codes efficiently in a distributed fashion. In a CDMA system, a node needs to know which code to use in transmitting or receiving a particular packet. Code assignment schemes are required and can generally be categorized as follows:

- Receiver-based code assignment (RCA). Each node is assigned a receiving code such that no two logical neighbors of any node will have the same code. All neighbor nodes send packets to a node using its code as the address. The benefit is RCA only incurs a smaller overhead for decoding at receivers.

- Transmitter-based code assignment (TCA). All neighbor nodes of a given node have different codes for transmitting so that no two neighbor nodes can cause primary conflicts. Although TCA requires extra overheads at receivers to search through known neighbors for proper decoding, it may be useful when broadcast/multicast are desirable.

- Pairwise code assignment (PCA). Another approach is to assign codes to a Tx-Rx pair (edge) such that no two adjacent edges in the logical topology have the same code. While PCA may be complicated to implement, it uses the limited code space efficiently.

- Hybrid scheme Any combination of the above. For example, the authors in [41] proposed two hybrid schemes: the common-transmitter- based protocol and the receiver/ransmitter based protocol. In the first protocol, the fields in the packet header that contain the source and destination addresses are spread using a common code, while the rest of the packet is spread using the transmitter's code. An idle terminal constantly monitors the common code. Upon recognizing its address in the destination field, the listening terminal switches to the code of the transmitting node to receive the rest of the packet. The receiver/transmitter-based works similarly, but with the common code replaced with the receiver's code.

In this dissertation, we assume either RCA or receiver/transmitter-based hybrid assignments because we only consider point-to-point transmission and want to reduce system overheads. This code assignment problem is trivial if the network size is small. Assigning a unique code to each node is inefficient when the network size grows. In addition the number of codes is usually constrained by the available spectrum and designated spreading factor, therefore, spatial reuse of codes becomes increasingly important when we extend CDMA to a large scale network.

In the past a number of code assignment protocols have been proposed with the constraint that all neighbors of a node have different codes. The relationship between CDMA code assignment problems and graph coloring is shown in [18] and a distributed two-phase algorithm which realizes a pairwise code assignment is proposed. In [31] two frequency channels, one for data and one for control (i.e., FDMA-like partitioning) are used. A common spreading code is used by all nodes over the control channel, while several terminal-specific codes can be used over the data channel. These existing efforts show that a proper code assignment can indeed be realized in a distributed way.

Chapter 4

Performance Analysis of Existing CDMA Ad Hoc Network Designs

In this chapter, we will systematically analyze the performance of some existing popular MAC designs for CDMA ad hoc networks. we argue that while FH-CDMA works well under simple MAC scheduling schemes, most existing system designs may not be able to leverage the full potential of a DS-CDMA physical layer for high level of spatial reuse or that they can only achieve this through centralized management. Therefore, we motivate the need for reconsidering the design principles to achieve better performance in DS-CDMA ad hoc networks.

4.1 Existing work on CDMA ad hoc networks

There has been a history and recently a resurgence of interest in using spread spectrum in ad hoc networks [32][33][40][19][13][38], suggesting some of the advantages mentioned in Chapter. 3. For example, early work by Pursley and Taipale [33] studied error probabilities for spread spectrum ad hoc networks and found that frequency hopping was generally preferable to direct sequence due to the near-far effect, a result that will be revisited later but from a network capacity perspective. [40] focused on choosing the optimal transmission range to optimize successful relay progress per time slot, which yields the same result that maximizing capacity requires nearest neighbor routing. Both [13] [38] initiated protocol designs for spread spectrum ad hoc networks.

MAC protocol design for spread spectrum ad hoc networks has been studied in recent work. In [38] a random access/ALOHA-like protocol, which is simple to implement but inefficient from a spatial reuse perspective, is proposed. This will be shown later in this dissertation. In [10] joint power control and scheduling are considered so as to achieve optimal spatial reuse but assuming a centralized scheduler to solve a global optimization

problem. Such centralized implementation may not be feasible in a distributed system. MAC designs using local signaling among nodes [12][20][30], e.g., exchanging RTS/CTS messages over a common code/frequency control channel, realize transmitter-receiver handshaking and may improve the performance over random access schemes. However, they usually assume signaling itself in the control channel is contention free, which may not be true in a practical system.

In the sequel, we will analyze the performance of these different classes of MAC protocols. Note that existing MAC design approaches usually only resort to simulation results for performance evaluation and thus offer little insight on which design principle is preferable and why that would be the case. Instead, we will take a stochastic geometric approach to systematically analyze the performance of MAC protocols, in particular focusing on spatial reuse and illustrate their difference in terms of performance.

4.2   ALOHA-like random channel access

In this section we consider a model for ALOHA-like random channel access which is analytically tractable. The model assumes no carrier sensing or handshaking and is akin to those in [38][40][3]. The basic premise is that spatially distributed nodes randomly (or based on a predefined pseudo-random sequence) choose to be transmitters with certain probability and do so without contending/signaling with each other. The lack of a coordination phase among concurrent transmitter-receivers reduces overheads but increases the likelihood of failed transmissions. We consider both FH-CDMA and DS-CDMA ad hoc networks and analyze their spatial reuse performance and impacts of system parameters on the performance. In particular, we consider two scenarios. In the first, transmitters use a fixed transmit power level, while in the second, transmitters use variable power levels depending on different transmit distances to their associated receivers.

4.2.1   Capacity under outage constraint - no power control case.

We begin by introducing, and then elaborating on, a simple stochastic geometric model for transmitters and receivers in an ad hoc network. The simplicity is key to allowing tractable analysis, yet the salient characteristics are still captured. We assume that a set of transmit nodes (including nodes relaying packets) are spatially distributed according to a homogenous Poisson point process $\Pi = \{X_i\}$ with intensity λ [45] on the plane. In addition, for the FH-CDMA case we assume each transmitter chooses its sub-channel independently. We let Π_M denote the set of transmitters which select sub-channel M, for $M = 1, \ldots, m$ and they are independently sampled at random. Because of the independent sampling assumption, each process Π_M is a homogeneous Poisson point process with intensity $\frac{\lambda}{m}$.

For simplicity, we initially assume that all transmitters use the same transmission power, ρ, and all transmission distances are over the same distance d, i.e., each transmitter is assumed to be sending to a receiver, which is modeled as being at a random location a distance d away. In addition, we will assume receive nodes, are always available at these randomly selected locations. These assumptions will be relaxed in the subsequent model. Our initial focus and model are geared at investigating the scenario where the transmit range d exceeds the typical nearest neighbor distance. The model captures a homogenous offered load where packets are typically relayed along hops with a transmission range d, leading to a homogenous distribution of transmitters. We shall further assume that transmissions are synchronous, or at least approximately so. As discussed in [44] an approximate synchronization provides significant advantages, which we revisit in details later. See also [38][36] for representative protocols based on synchronized contention.

In Chapter. 3 we highlighted the ability of FH and DS-CDMA physical layers to mitigate interference, through hopping and spreading respectively. In particular, FH-CDMA divides the available bandwidth, B, into m sub-channels, each with bandwidth $\frac{B}{m}$. A receiver attempting to decode a signal from a transmitter on sub-channel $M \in \{1, \ldots, m\}$ only sees interference from other simultaneous transmissions on that sub-channel. By contrast in a DS-CDMA network, a receiver sees interference from all transmitters, it uses the spreading factor to reduce the minimum SINR required for successful reception. If the nominal SINR requirement for FH-CDMA is β, then DS-CDMA has a reduced SINR requirement of $\frac{\beta}{m}$, assuming a rather conservative PN code cross-correlation [14][1].

To evaluate the spatial reuse performance, we first introduce a useful notion termed the *optimal contention density*, which corresponds to the maximum spatial density of nodes that can contend for the channel subject to a constraint on the typical outage probability. We then define our metric for ad hoc network spatial reuse, termed *transmission capacity*, to be the density of successful transmissions resulting from an optimal contention density of transmitters, i.e., the optimal contention density thinned by the probability of success. Note that in the low outage regime, which is of interest in practice, transmission capacity is equal to the maximum density of successful transmissions subject to the outage constraint because the capacity is more sensitive to the contention density than the success probability, which is close to one. As we will discuss shortly, the transmission capacity is a natural outgrowth of previous results on ad hoc network capacity, allowing consideration of a random distribution of nodes *and* a constraint on outage (or equivalently, success) probability in random channel access. A principal benefit of this approach will be the derivation of simple expressions for upper and lower bounds on transmission capacity, that clearly demonstrate the dependence of ad hoc network capacity on key system design parameters.

---

[1]Some in the literature use $\frac{1}{3m}$

To evaluate the outage probability we condition on a typical receiver[2] at the origin O. The outage constraint on λ corresponds to ensuring that the probability that the receive SINR is below the appropriate threshold, is less than ε. For the above FH-CDMA and DS-CDMA models, these are given by:

$$FH \qquad \mathbb{P}\left(\frac{\rho r^{-\alpha}}{\eta + \sum_{X_i \in \Pi_M} \rho |X_i|^{-\alpha}} \leq \beta\right) \leq \varepsilon, \qquad (4.1)$$

$$DS \qquad \mathbb{P}\left(\frac{\rho r^{-\alpha}}{m\eta + \sum_{X_i \in \Pi} \rho |X_i|^{-\alpha}} \leq \frac{\beta}{m}\right) \leq \varepsilon. \qquad (4.2)$$

The interference terms, e.g., $\sum_{X_i \in \Pi} \rho |X_i|^{-\alpha}$ defined in (4.2), are a special type of Poisson shot-noise process [29] and does not generally have a closed-form distribution. This makes the outage probability difficult to compute. Therefore in our analysis we will resort to obtaining careful bounds. Note the special case where $\alpha = 4$ does have a closed-form solution, see [40], which we will use later.

We let λ^ε denote the optimal contention density, i.e., the maximum density λ for Π such that the outage probability at a typical receiver is less than ε, where $\varepsilon \in (0,1)$. We will obtain upper and lower bounds on λ^ε denoted by $\lambda_u^{\varepsilon,DS}$, $\lambda_l^{\varepsilon,DS}$ and $\lambda_u^{\varepsilon,FH}$, $\lambda_l^{\varepsilon,FH}$ for the DS-CDMA and FH-CDMA cases respectively. Thus the bounds on the transmission capacity $c_u^{\varepsilon,DS}$, $c_l^{\varepsilon,DS}$ and $c_u^{\varepsilon,FH}$, $c_l^{\varepsilon,FH}$ correspond to bounds on the optimal contention density multiplied by $1 - \varepsilon$. These bounds and the transmission capacity ratio of FH-CDMA over DS-CDMA are given in the following theorem.

Theorem 4.2.1. *Let $\varepsilon \in (0,1)$, $\kappa = \frac{d^{-\alpha}}{\beta} - \frac{\eta}{\rho}$, and $h(\alpha) = \frac{\alpha-1}{\alpha}$. The lower and upper bounds on transmission capacity subject to the outage constraint ε for FH-CDMA when transmitters employ a fixed transmission power ρ for receivers that are a fixed distance r away are:*

$$c_l^{\varepsilon,FH} = \lambda_l^{\varepsilon,FH}(1-\varepsilon) \geq h(\alpha)\frac{m}{\pi}\kappa^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2)$$

$$c_u^{\varepsilon,FH} = \lambda_u^{\varepsilon,FH}(1-\varepsilon) = \frac{m}{\pi}\kappa^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2),$$

as $\varepsilon \to 0$.

---

[2]To be rigorous, we first consider a typical transmitter giving what is known as the Palm distribution for transmitters on the plane [45]. By Slivnyak's Theorem [45] this conditional distribution is also a homogenous Poisson point process with intensity λ with an additional point (the typical transmitter) at the origin. Now shifting this entire point process so that the receiver associated with the desired transmitter lies at the origin, we have that conditional distribution of interferers (excluding the transmitter of interest) is a homogenous Poisson point process with intensity λ. We indeed denote this shifted point process of *interferer* locations by $\Pi = \{X_i, i \in \mathbb{N}\}$. Similar argument applies to Π_M associated with sub-channel M for the FH-CDMA case.

The lower and upper bounds on transmission capacity subject to the outage constraint ε for DS-CDMA when transmitters employ a fixed transmission power ρ for receivers that are a fixed distance r away are:

$$c_l^{\varepsilon,DS} = \lambda_l^{\varepsilon,DS}(1-\varepsilon) \;\; \geq \;\; h(\alpha)\frac{1}{\pi}\left(m\kappa\right)^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2)$$

$$c_u^{\varepsilon,DS} = \lambda_u^{\varepsilon,DS}(1-\varepsilon) \;\; = \;\; \frac{1}{\pi}\left(m\kappa\right)^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2)$$

as $\varepsilon \to 0$.

Proof. Let us briefly consider how these results are obtained. The outage constraints in (4.1) and (4.2) can be rewritten as

$$FH \qquad \mathbb{P}^0\left(\sum_{i\in\Pi_M} |X_i|^{-\alpha} \geq \kappa\right) \leq \varepsilon, \qquad\qquad (4.3)$$

$$DS \qquad \mathbb{P}^0\left(\sum_{i\in\Pi} |X_i|^{-\alpha} \geq m\kappa\right) \leq \varepsilon. \qquad\qquad (4.4)$$

These are complex functions of the contention density λ so we will resort to obtaining careful bounds.

Consider the DS-CDMA case. Specifically, according to (4.4), the outage event is given by

$$E(\lambda) = \left\{\sum_{i\in\Pi} |X_i|^{-\alpha} \geq m\kappa\right\}.$$

We will define events $E_u(\lambda,s)$ and $E_l(\lambda,s)$ such that $E_u(\lambda,s) \subset E(\lambda) \subseteq E_l(\lambda,s)$ and the probabilities of all events $E(\lambda)$, $E_l(\lambda,s)$ and $E_u(\lambda,s)$ increase in λ. Here s is a parameter that will be discussed in the sequel. Therefore if we solve for the largest possible λ such that $\mathbb{P}^0(E_u(\lambda,s)) \leq \varepsilon$, we obtain an upper bound $\lambda_u^{\varepsilon,DS}$ on the optimal contention density, i.e., if $\lambda > \lambda_u^{\varepsilon,DS}$ the outage probability must exceed ε. If we solve for the smallest λ such that $\mathbb{P}^0(E_l(\lambda,s)) \geq \varepsilon$, we obtain a lower bound on the optimal contention density $\lambda_l^{\varepsilon,DS}$, i.e., if $\lambda < \lambda_l^{\varepsilon,DS}$ the outage probability must not exceed ε.

In order to define $E_l(\lambda,s)$ and $E_u(\lambda,s)$, we consider the overall interference a receiver sees from both the 'near field' and 'far field'. As shown in Fig 4.1, the near and far fields are the regions inside and outside of a circle of radius s around the typical receiver at the origin, denoted $b(O,s)$ and $\bar{b}(O,s)$ respectively. The radius s is selected to be small enough such that one or more nodes within distance s would cause an outage. According to (4.4) this means $s^{-\alpha} \geq m\kappa$, which limits $s \leq (m\kappa)^{-\frac{1}{\alpha}}$. In the sequel we shall consider optimizing over s subject to this constraint so as to get the tightest bounds. The rational for separating near and far field's interference is that near-field nodes contribute a major part of the interference.

The outage events associated with near and far field's interference is defined as follows.
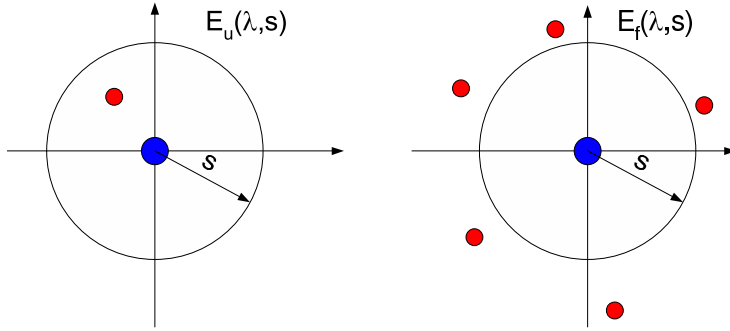
Figure 4.1: On the left, the outage events $E_u(\lambda, s)$ corresponds to an outage caused by near field interfering transmitters, i.e., one or more interferers within distance $s \leq (m\kappa)^{-\frac{2}{\alpha}}$. On the right, the outage event $E_{f(\lambda,s)}$ corresponds to an outage caused far field interfering transmitters, i.e., the aggregate interference level is beyond some threshold such that $\sum_{i \in \Pi \cap \bar{b}(O,s)} |X_i|^{-\alpha} \geq m\kappa$. Note that for DS-CDMA, s is less than the transmission range but for FH-CDMA, s can be larger.

Definition 4.2.1.

$$
\begin{aligned}
E_u(\lambda, s) &= \left\{ \Pi \cap b(O, s) \neq \emptyset \right\}, \\
E_f(\lambda, s) &= \left\{ \sum_{i \in \Pi \cap \bar{b}(O,s)} |X_i|^{-\alpha} \geq m\kappa \right\}, \\
E_l(\lambda) &= E_u(\lambda, s) \cup E_f(\lambda, s),
\end{aligned}
$$

where $E_u(\lambda, s)$ and $E_f(\lambda, s)$ correspond, respectively, to the event where there is one or more nodes within a distance s of the origin, and to the event where there is sufficient interference from *outside* of the same distance s to cause an outage.

It is fairly straightforward to show that these events satisfy the following relationships (see Appendix):

i) $E_u(\lambda, s) \subset E(\lambda)$ for all $s \leq (m\kappa)^{-\frac{1}{\alpha}}$.

ii) $E(\lambda) \subseteq E_l(\lambda, s)$ for all $s \leq (m\kappa)^{-\frac{1}{\alpha}}$.

iii) $E_u(\lambda, s)$ and $E_f(\lambda, s)$ are independent events and thus

$$
\begin{aligned}
\mathbb{P}^0(E_l(\lambda, s)) &= \mathbb{P}^0(E_u(\lambda, s)) + \mathbb{P}^0(E_f(\lambda, s)) \\
&\quad - \mathbb{P}^0(E_u(\lambda, s))\mathbb{P}^0(E_f(\lambda, s)).
\end{aligned}
$$

Note that the probabilities of the events $E(\lambda)$, $E_u(\lambda, s)$, $E_f(\lambda, s)$ and $E_l(\lambda)$ defined above, all increase in λ. Also recall that the optimal contention density is the largest λ such that $\mathbb{P}^0(E(\lambda)) < \varepsilon$. For any given s, one can thus obtain parameterized lower and upper bounds $\lambda_l^{\varepsilon,DS}(s)$ and $\lambda_u^{\varepsilon,DS}(s)$ on the optimal contention density and as discussed previously can optimize over all feasible s to obtain the tightest bounds.

To obtain the upper bound on the optimal contention density, we calculate the probability of event $E_u(\lambda, s)$, which is equal to the void probability of a homogeneous Poisson point process with intensity λ [45], i.e.,

$$\mathbb{P}^0(E_u(\lambda, s)) = 1 - \mathbb{P}(\Pi \cap b(0, s) = \emptyset) = 1 - e^{-\lambda \pi s^2}. \tag{4.5}$$

By setting the above equation equal to ε, one can solve for the upper bound $\lambda_u^{\varepsilon,DS}(s)$ for the optimal contention density given the cutoff distance s and the tightest $\lambda_u^{\varepsilon,DS}$ upper bound can then be obtained by further optimizing over $s \in [0, (m\kappa)^{-\frac{1}{\alpha}}]$.

For the lower bound on the optimal contention density, since $E(\lambda) \subseteq E_l(\lambda, s)$ and $\mathbb{P}^0(E_l(\lambda, s)) = \mathbb{P}^0(E_u(\lambda, s)) + \mathbb{P}^0(E_f(\lambda, s)) - \mathbb{P}^0(E_u(\lambda, s))\mathbb{P}^0(E_f(\lambda, s))$, one can write the sufficient condition for the outage constraint $\mathbb{P}^0(E(\lambda)) < \varepsilon$ as

$$\mathbb{P}^0(E_u(\lambda, s)) < \varepsilon_1, \ \ \mathbb{P}^0(E_f(\lambda, s)) < \varepsilon_2, \tag{4.6}$$

where constants ε_1 and ε_2 are such that $\varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2 = \varepsilon$.

Each condition in (4.6) gives to a lower bound, $\lambda_u^{\varepsilon_1,DS}(s)$ and $\lambda_f^{\varepsilon_2,DS}(s)$ respectively, for some s. The derivation for $\lambda_u^{\varepsilon_1,DS}(s)$ is the same as $\lambda_u^{\varepsilon,DS}(s)$ by substituting ε with ε_1. However estimating $\lambda_f^{\varepsilon_2,DS}(s)$ is not straightforward. We shall do this using Chebychev's inequality to obtain lower bound on $\lambda_f^{\varepsilon_2,DS}(s)$. Note that the tightness of the lower bound also depends on both the choices of $(\varepsilon_1, \varepsilon_2)$ and s. Finally we use both $\lambda_u^{\varepsilon_1,DS}(s)$ and $\lambda_f^{\varepsilon_2,DS}(s)$ and optimize over s such that $s \leq (m\kappa)^{-\frac{1}{\alpha}}$ and $(\varepsilon_1, \varepsilon_2)$ such that $\varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2 = \varepsilon$ to derive the tightest lower bound for the optimal contention density:

$$\lambda_l^{\varepsilon,DS} = \max_{s,(\varepsilon_1,\varepsilon_2)} \left(\min(\lambda_u^{\varepsilon_1,DS}(s), \lambda_f^{\varepsilon_2,DS}(s)) \right).$$

Considering equations (4.3)(4.4), it is clear that the exact same analysis for FH-CDMA holds provided we replace λ with λ/m and $m\kappa$ with κ. A detailed proof is included in the Appendix. $\square$

There are a number of observations that can be made from inspecting the upper and lower bounds derived in Theorem 1. We will save the bulk of such discussion until Section 4.2.3 when we will examine some plots of these expressions to enhance intuition. Before continuing, however, we would like to note some similarities between the preceding bounds on transmission capacity and other possible metrics of ad hoc network capacity, in particular transport capacity and network sum capacity.

27

In Theorem 1, we rediscover the scaling property between capacity and transmission range d to be $\Theta(d^{-2})$, the same as the result of [3]. One can understand this scaling as packing as many concurrent transmissions spatially as possible, with each occupying an area $\Theta(d^2)$. Spreading factor m allows for certain relaxations on the overlapping among concurrent transmissions, which is prohibited in narrow band systems, but eventually only provides a constant gain on the number of concurrent transmissions that can be scheduled. Thus qualitatively, the scaling of transmission capacity in d is still $\Theta(d^{-2})$. Relating this to transport capacity, the transport capacity (for a fixed bandwidth of 1 Hz and a fixed area of $1\ m^2$) of an ad hoc network is essentially the maximum number of legal transmissions (i.e. n), multiplied by the transmission range d, multiplied by the achieved spectral efficiency of each transmission, b. In the best case, this was shown to scale as $\Theta(\sqrt{n})$ in [16]. In this paper, the number of legal transmissions (i.e. have received SINR above β) is a stochastic measure λ which has been shown in Theorem 1 to be inversely proportional to d^2. Hence, it can be readily seen that our comparable metric for the transport capacity is $\lambda \cdot d \to \Theta(\sqrt{n})$ since $\lambda \propto n$. Thus, transmission capacity recovers the basic scaling result of [16], but also includes a stochastic notion of successful transmissions and allows the computation bounds for non-asymptotic n.

It is also natural to wonder how the transmission capacity or transport capacity might relate to measurable network throughput, for example the sum data rate over all the nodes. With a target SINR of β it is well known that the achievable spectral efficiency of each legal transmission is $b \geq \log_2(1 + \beta/\Gamma)$, where Γ is "the Gap", the penalty of a particular modulation and coding technique and error probability specification relative to Shannon capacity, which has $\Gamma \to 1$. The inequality is due to the fact that the legal transmissions have a received SINR of *at least* β. The latter observation makes the typical relation of $b \leq c = \log_2(1 + \beta)$ inaccurate. The sum capacity of the ad hoc network per unit area and bandwidth, called the area spectral efficiency, can then be lower bounded by the transmission capacity multiplied by b, computed as just described. We note that it will not be simple to attain an exact expression for sum capacity due to the stochastic nature of the actual received SINR for each node γ_i, which leads each node i to have a unique achievable data rate $b_i \leq \log_2(1 + \gamma_i)$. If we define $\bar{b} = E_\Pi[b_i]$, then the average area spectral efficiency will be $\lambda \cdot \bar{b}$. In summary, there is a direct relation between transmission capacity and traditional information theoretic measures of network capacity, but the former is more direct.

Theorem 4.2.1 shows that there is a capacity improvement of FH-CDMA over DS-CDMA on the order of $\Theta(m^{1-\frac{2}{\alpha}})$, which can be quite large as the channel attenuation increases, i.e. $\alpha > 2$, or spread gain increases. The intuition is that FH-CDMA is less sensitive to the near-far problem than DS-CDMA. In addition, we observe a capacity gain – the capacity improves linearly for FH-CDMA and sub-linearly for DS-CDMA in the spreading gain m (when $\alpha > 2$). Hence, if the traffic in an ad hoc network does not require a very

high data rate, e.g., voice traffic, it may be desirable to use a high spreading gain in order to achieve robust interference tolerance, low delay, or high transmission capacity.

For both models, the capacity under the outage constraint indeed scales approximately linearly in the constraint constant ε for small ε. Therefore the outage constraint significantly limits the transmission capacity when such networks use a random channel access strategy.

4.2.2 Capacity under outage constraint - power control case.

In our second model we remove the assumption that all transmitters use the same transmission power and have associated receivers at the same distance. In real ad hoc networks transmission relay distances will be variable as will interference power levels. This suggests transmitters should use power control since if the signal power is too high it may cause unnecessary interference and if it is too low the signal may not be successfully received. Finding a system-wide optimal set of transmission power levels is the subject of recent work [10], which only offers a centralized solution for global power control and scheduling. In this work we take a simple distributed approach of assuming that transmitters choose their transmission power as a function of their distance from their intended receiver but independently of the interference level at the receiver. We call this *pairwise power control* since each transmitter and receiver pair determine the transmission power independently of other pairs. Specifically, the transmitter chooses its transmission power such that the signal power at the receiver will be some fixed level ρ. Thus if a transmitter and receiver are separated by a distance d then the transmitter will employ a transmission power ρd^{α} so that the received signal power is ρ. We make no particular assumption on the value of ρ, other than $\rho > \eta\beta$, which is required to keep the received signal power above the noise floor.

Devices are assumed to have a maximum transmission power of ρ_{max}. Solving $\rho d^{\alpha} \leq \rho_{max}$ for d gives a maximum transmission distance of $d_{max} = \left(\frac{\rho_{max}}{\rho}\right)^{\frac{1}{\alpha}}$. We assume that transmission distances are i.i.d. with (continuous) CDF $F_D(d)$ and PDF $f_D(d)$ and that the interval $[1, d_{max}]$ is the support of this distribution.

Formally, our second model consists of a *marked* homogeneous Poisson point process[3] $\Phi = \{(X_i, D_i)\}$ where the points $\{X_i\}$ again denote the locations of interfering transmitters and the marks $\{D_i\}$ denote the distance between transmitter i and its intended receiver. We assume the marks are independent and identically distributed with CDF $F_D(d)$, and that the marks are also independent of the transmitter locations. We use $|X_i|$ to denote the distance from node i to the origin. Similar to the first model, we define the sampled sub-process Φ_M

---

[3]Similar to the first model, we evaluate the outage probability using the Palm distribution for the marked point process Φ, which indeed is a shifted version of Φ with a typical receiver placed at the origin O.

as a homogeneous marked Poisson point processes consisting of all interfering transmitters in Φ that are on sub-channel m, for $M = 1, \ldots, m$.

To evaluate the interference a typical receiver sees, we define the function $l(r, d)$ as giving the signal power level at a distance r from the transmitter when the transmitter's intended recipient is at a distance d. Thus, $l(r, d) = \rho \left(\frac{d}{r} \right)^{\alpha}$. Note in particular that $l(d, d) = \rho$, i.e., at the distance of the intended receiver the signal power is the desired level. Note that the transmission power is $l(1, d) = \rho d^{\alpha}$.

The appropriate outage constraints on λ are now given by:

$$FH \qquad \mathbb{P}\left(\frac{\rho}{\eta + \sum_{(X_i, D_i) \in \Phi_m} l(|X_i|, D_i)} \leq \beta \right) \leq \varepsilon, \qquad (4.7)$$

$$DS \qquad \mathbb{P}\left(\frac{\rho}{m\eta + \sum_{(X_i, D_i) \in \Phi} l(|X_i|, D_i)} \leq \frac{\beta}{m} \right) \leq \varepsilon. \qquad (4.8)$$

We can use these to obtain upper and lower bounds for the second model. These bounds on the transmission capacity of FH-CDMA and DS-CDMA, are given in the following theorem.

Theorem 4.2.2. *Let* $\varepsilon \in (0, 1)$, $\delta = \frac{1}{\beta} - \frac{\eta}{\rho}$, *and let* $g(\alpha) = \frac{1}{2} \left((\alpha + 1)(\alpha - 1) \right)^{\frac{1}{\alpha}}$. *Suppose* $F_D(d) = \frac{d^2 - 1}{d_{max}^2 - 1}$ *for* $0 \leq d \leq d_{max}$, *i.e., intended receivers are such that the transmission distances are uniformly selected in range* $[1, d_{max}]$. *The lower and upper bounds on transmission capacity for FH-CDMA when transmitters employ pairwise power control are given by:*

$$c_l^{\varepsilon, FH} = \lambda_l^{\varepsilon, FH}(1 - \varepsilon) \geq g(\alpha) \frac{m \delta^{\frac{2}{\alpha}}}{\pi} \left(\frac{d_{max}^2 - 1}{d_{max}^{2\alpha+2} - 1} \right)^{\frac{1}{\alpha}} \varepsilon + \Theta(\varepsilon^2)$$

$$c_u^{\varepsilon, FH} = \lambda_u^{\varepsilon, FH}(1 - \varepsilon) = \frac{4}{\pi} \frac{m \delta^{\frac{2}{\alpha}} (d_{max}^2 - 1)}{d_{max}^4} \varepsilon + \Theta(\varepsilon^2)$$

as $\varepsilon \to 0$.

The lower and upper bounds on transmission capacity for DS-CDMA when transmitters employ pairwise power control are given by:

$$c_l^{\varepsilon, DS} = \lambda_l^{\varepsilon, DS}(1 - \varepsilon) \geq g(\alpha) \frac{(m\delta)^{\frac{2}{\alpha}}}{\pi} \left(\frac{d_{max}^2 - 1}{d_{max}^{2\alpha+2} - 1} \right)^{\frac{1}{\alpha}} \varepsilon + \Theta(\varepsilon^2)$$

$$c_u^{\varepsilon, DS} = \lambda_u^{\varepsilon, DS}(1 - \varepsilon) = \frac{4}{\pi} \frac{(m\delta)^{\frac{2}{\alpha}} (d_{max}^2 - 1)}{d_{max}^4} \varepsilon + \Theta(\varepsilon^2)$$

as $\varepsilon \to 0$.

30

Proof. The proof of Theorem 4.2.2 is similar to that of Theorem 4.2.1. Briefly, the outage constraints (4.7) and (4.8) can be written as

$$FH \qquad \mathbb{P}^0\Big(\sum_{(X_i,D_i)\in\Phi_m}\Big(\frac{|X_i|}{D_i}\Big)^{-\alpha}\geq\delta\Big)\leq\varepsilon, \qquad (4.9)$$

$$DS \qquad \mathbb{P}^0\Big(\sum_{(X_i,D_i)\in\Phi}\Big(\frac{|X_i|}{D_i}\Big)^{-\alpha}\geq m\delta\Big)\leq\varepsilon. \qquad (4.10)$$

Consider the FH-CDMA case. Let $m\in\{1,\ldots,m\}$ denote a particular sub-channel used in FH-CDMA and let us define the following events:

Definition 4.2.2.

$$\begin{aligned}
E(\lambda) &= \Big\{\sum_{(X_i,D_i)\in\Phi_m}\Big(\frac{D_i}{|X_i|}\Big)^{\alpha}>\delta\Big\}, \\
E_u(\lambda,s) &= \Big\{\Phi_m\cap(b(O,s)\times[s\delta^{\frac{1}{\alpha}},d_{max}))\neq\emptyset\Big\}, \\
E_{l_1}(\lambda,s) &= \Big\{\Phi_m\cap(b(O,s)\times\mathbb{R}^+)\neq\emptyset\Big\}, \\
E_{l_2}(\lambda,s) &= \Big\{\sum_{(X_i,D_i)\in\Phi_m}\Big(\frac{D_i}{|X_i|}\Big)^{\alpha}\mathbf{1}(X_i\in\bar{b}(O,s))>\delta\Big\}, \\
E_l(\lambda) &= E_{l_1}(\lambda,s)\cup E_{l_2}(\lambda,s).
\end{aligned}$$

Here the event $E(\lambda)$ corresponds to an outage event. The event $E_u(\lambda,s)$ consists of all outcomes where there are one or more transmitters within s of the origin with transmission distances exceeding $s\delta^{\frac{1}{\alpha}}$. This threshold is the smallest transmission distance such that a single transmitter in $b(O,s)$ transmits to a receiver at this distance will cause an outage at the origin. The event $E_{l_1}(\lambda,s)$ consists of all outcomes with one or more transmitters in $b(O,s)$. Note however that not all outcomes in $E_{l_1}(\lambda,s)$ will cause an outage. Finally, the event $E_{l_2}(\lambda,s)$ consists of all outcomes where the aggregate interference power at the origin caused by the transmitters outside $b(O,s)$ suffices to cause an outage at the origin.

Applying a similar approach as in our first model based on the events defined above, we can prove he result in Theorem 4.2.2. For detailed proof, please see the Appendix. □

Theorem 4.2.2 shows how the transmission capacity scales in the fundamental system parameters, e.g., transmission distance, spreading factor and outage constraint. We see that the scaling is the same as Theorem 4.2.1 for both DS-CDMA and FH-CDMA. Power control in cellular networks solves the "near-far" problem by equalizing receiving powers at the central base station. The pairwise power control scheme in ad hoc networks cannot fully solve the "near-far" problem since transmitters have different intended receivers,

Table 4.1: Simulation Parameters (unless otherwise noted)

Symbol	Description	Value
m	Spreading factor	16
d	Transmission radius	10m
ε	Target outage probability	0.1
β	Target *SINR*	$3 = 4.77$dB
ρ	Transmit Power	1
α	Path loss exponent	3

but it offers a simple and distributed means by which to mitigate the interference across concurrent transmissions.

4.2.3 Numerical Results and Interpretations

The derived transmission capacity results are evaluated in this section for some typical parameters in order to show how ad hoc network capacity can be expected to scale with path loss and spreading, and to compare frequency hopping and direct sequence spread spectrum. Additionally, a simulated ad hoc network where nodes are spatially distributed according to a Poisson point process is used to show how the derived bounds perform relative to simulated performance. The simulations are carried out using the parameter values enumerated in Table 4.1. All simulation results shown in the plots are confidence intervals, although the intervals are too small to be visually distinguished from a point. Note that we only show the numerical results for our first model with fixed transmission power and relay distance because the numerical results of the second model is basically the same as the first one. For reference, we also include Table 4.1 that shows analytically how transmission capacity scales for FH, DS, and narrowband ($m = 1$) modulation, and it can be noted that they only differ in terms of their scaling with regards to m.

Outage probability vs. transmission density

The first investigation is to study the outage probability $p_o(\lambda) = \mathbb{P}^0(E(\lambda))$ versus the transmission density λ. The outage lower bound $\mathbb{P}^0(E_u(\lambda, s))$ is given by (4.14). Let $Y(\lambda, s) = \sum_{i \in \Pi \cap \bar{b}(O,s)} |X_i|^{-\alpha}$ denote the normalized far field interference from transmitters outside the circle $b(O, s)$, as shown in Fig 4.1. Let $y = \kappa$ for FH-CDMA and $y = m\kappa$ for DS-CDMA denote the normalized SINR requirement as expressed in Equation (4.3) and (4.4) respectively. We obtain an upper bound of the outage probability by applying the same technique
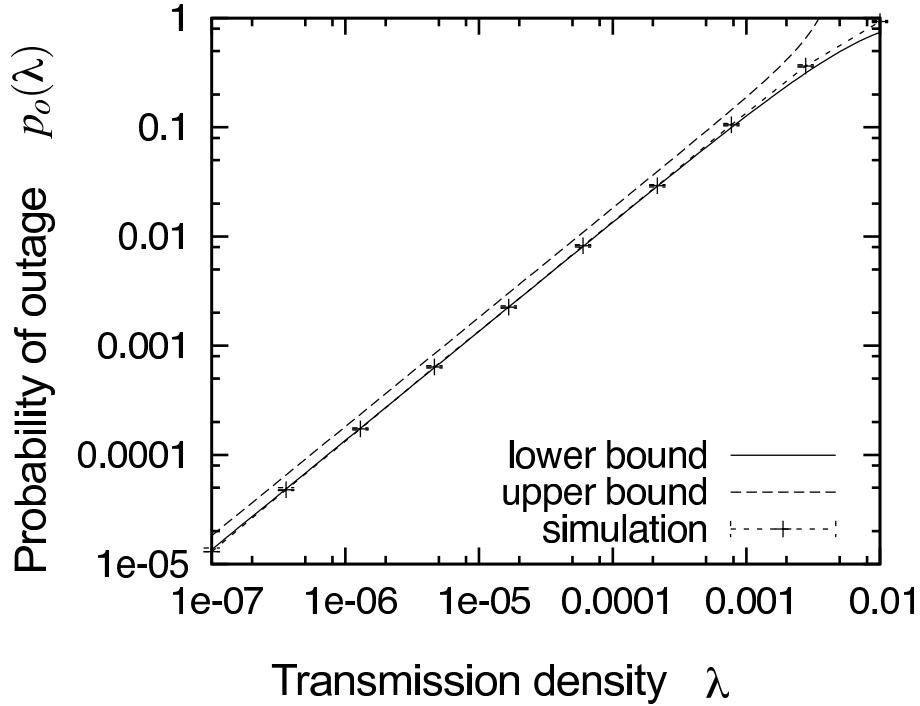
Figure 4.2: Numerical and simulation results for the probability of outage $p_o(\lambda)$ versus the transmission density λ. The numerical bounds are the upper and lower bounds on $p_o(\lambda)$. The simulation results (with confidence intervals) are seen to fall between the lower and upper bounds.

used in the proof of the lower bound in Theorem 4.2.1:

$$
\begin{aligned}
\mathbb{P}^0(E(\lambda,s)) &= \mathbb{P}^0(E_u(\lambda,s) \cup E_f(\lambda,s)) \\
&\leq \mathbb{P}^0(E_u(\lambda,s)) + \mathbb{P}^0(E_f(\lambda,s)) \\
&= \mathbb{P}^0(E_u(\lambda,s)) + \mathbb{P}^0(Y(\lambda,s) > y),
\end{aligned}
$$

where $\mathbb{P}^0(E_u(\lambda,s))$ is the same as above and $\mathbb{P}^0(Y(\lambda,s) > y)$ can be upper bounded using the Chebychev inequality.

Figure 4.2 plots numerical and simulation results of $p_o(\lambda)$ vs. λ; the simulated outage probability falls between the lower and upper bounds as predicted. The plot illustrates that the lower bound is reasonably tight with respect to the simulated performance, and that as expected from our analytical expressions in Theorem 4.2.1, outage probability increases about linearly in the transmission density in the low outage regime.

Transmission capacity vs. path loss exponent

Figure 4.3 shows the transmission capacity $c^\varepsilon = (1-\varepsilon)\lambda^\varepsilon$ versus the path loss exponent α for both FH-CDMA and DS-CDMA systems. The bounds given in Theorem 4.2.1 are plotted along with the simulation results, and as expected, the simulated transmission density falls between the lower and upper bounds. The plot illustrates that the upper bound is fairly tight. Recall that the lower and upper bounds are given up to an asymptotic order $\Theta(\varepsilon^2)$. Here these expressions are plotted assuming that this constant is zero. So, the tightness of the upper bound is due partly to the neglect of this small term, and also due to the fact the close-in interfering nodes – which are what determine the upper bound – appear to dominate the transmission capacity.

As can be seen in this plot, frequency hopping is increasingly superior to direct sequence as the path loss becomes worse. The interpretation for this is that as the path loss worsens, dramatically more power is needed to reach the desired transmitter, so dramatically more interference is caused to neighbors. FH-CDMA systems typically avoid the interference by hopping with an occasional collision, so their performance is improved by this effect since now the aggregate interference from far away nodes in the utilized frequency slot is decreased. On the other hand, DS-CDMA receivers, which must suppress with interference from other transmitters that are closer to it than the desired transmitter, are at a distinct disadvantage since the desired power decreases more quickly than the interference power of close-in nodes as the path loss exponent increases.

The decay in transmission capacity as $\alpha \to 5.5$ is a consequence of the SNR (absent any interference) being below the SINR requirement: solving $\frac{\rho r^{-\alpha}}{\eta + 0} = \beta$ for α yields $\alpha = 5.5$ for the parameters given in Table 4.1. This is the value of α such that even absent any interference, the SNR ratio at the receiver is below the SINR requirement β. In other words, the received power is very close to or below the noise floor.

Transmission capacity vs. spreading factor

The final investigation studies the transmission capacity c^ε versus the spreading factor m for both FH-CDMA and DS-CDMA systems. Figure 4.4 plots the numerical and simulation results, with the simulated random network falling between the lower and upper bounds as predicted, with again the upper bound relatively accurately approximating the actual transmission capacity. Note that we plot $\frac{c^\varepsilon}{m}$ versus m. The transmission capacity is normalized by m to account for the fact that increasing m requires a commensurate increase in bandwidth. Thus c^ε/m is a rough measure of the spectral efficiency.

The key insight from this plot is that FH-CDMA capacity is unaffected by the spreading gain, whereas DS-CDMA capacity grows steadily worse as more spreading is employed. The interpretation is that the amount of interference that can be suppressed with
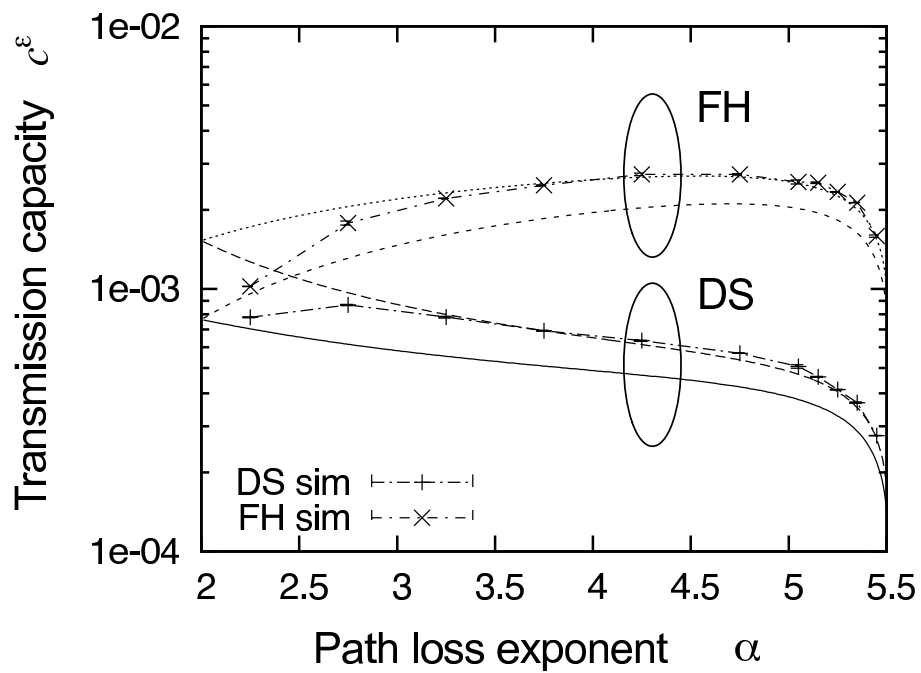
Figure 4.3: Numerical and simulation results for the transmission capacity c^ε versus the path loss exponent α. The upper bound appears relatively tight relative to the simulation results. The decay in transmission capacity as $\alpha \to 5.5$ is a consequence of the received power approaching the ambient noise floor.

the spreading factor does not compensate for the fact the bandwidth had to be increased (or the data rate decreased) by a factor of *m*.
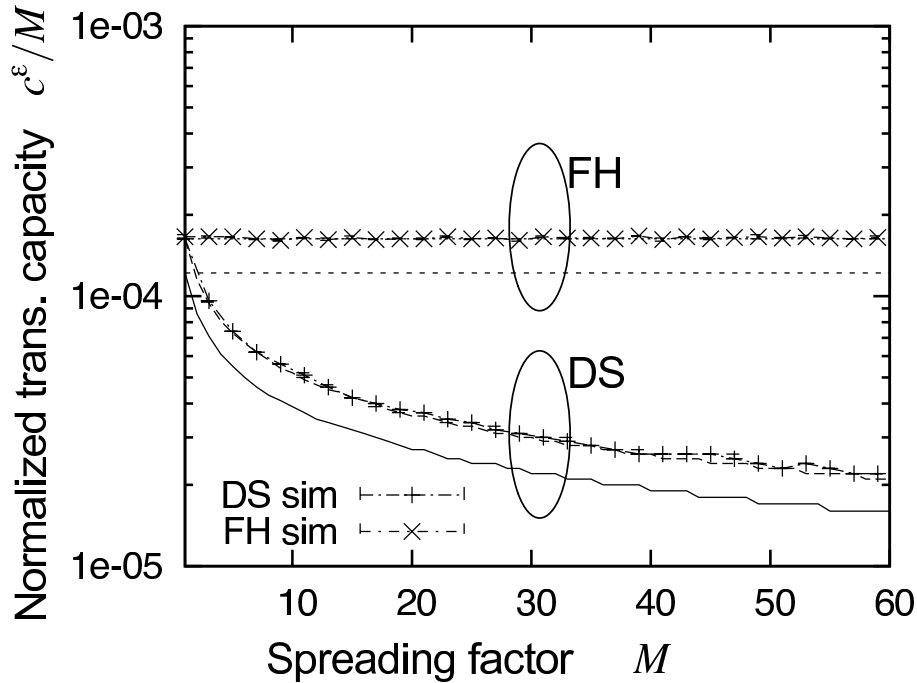
Figure 4.4: Numerical and simulation results for the transmission capacity c^ε/m versus the spreading factor *m*. Frequency hopping's advantage over direct sequence is increased as *m* increases.

4.2.4 Frequency Hopping vs. Direct Sequence vs. Narrowband

A recurring theme throughout the discussion thus far has been the apparent superiority of FH-CDMA to DS-CDMA, whenever the path loss exponent $\alpha > 2$, as can be directly observed from the bounds in Theorems 1 and 2. In this section, we discuss the significance and meaning behind this result. First, we note that although in some environments (notably indoor or urban canyons) the path loss exponent is sometimes modeled to be less than 2 due to reflections, these environments usually have significant shadowing and fading, bringing the "effective" path loss exponent to much greater than 2 if all these effects were lumped into just the path loss model. So, in general it is reasonable to assume that the effective $\alpha > 2$, and the common assumption in terrestrial environments of $\alpha = 4$ results in both FH-CDMA and narrowband (NB) having a higher normalized transmission capacity than

36

Table 4.2: Transmission capacity scalings

	FH-CDMA	DS-CDMA	Narrowband
Spreading factor (m)	m	$m^{\frac{2}{\alpha}}$	1
Target SINR (β)	$\frac{1}{\beta^{\frac{2}{\alpha}}}$	$\frac{1}{\beta^{\frac{2}{\alpha}}}$	$\frac{1}{\beta^{\frac{2}{\alpha}}}$
Outage Constraint (ε)	ε	ε	ε
Transmission range (d)	d^{-2}	d^{-2}	d^{-2}

DS-CDMA by a factor of $\sqrt{m}$.

Those experienced with CDMA may recall that FH and DS perform identically or within a small constant of each other assuming perfect power control [39, 50] in a cellular environment, with no dependence on the spreading factor. So why is FH-CDMA better than DS-CDMA by such a wide margin in ad hoc networks? The reason is that "perfect" power control is impossible to achieve in an ad hoc network due to the random locations of the transmitters and receivers, so the near-far problem is impossible to reconcile with power control. In addition, we have assumed DS-CDMA uses non-orthogonal codes generated from P/N sequence while FH-CDMA is using orthogonal FDMA channels. As a result, it is better to *avoid* interference than to attempt to *suppress* interference. As a simple example, consider three transmit-receive pairs shown in Fig. 4.5. In DS-CDMA, transmission B will continually overwhelm the receivers for transmissions A and C unless an enormous spreading factor is employed. By contrast, in FH-CDMA this situation is only a problem when transmission B is in the same frequency slot as A or C, each of which occurs only with probability $1/m$.

Other readers may have noted that throughout this paper we have assumed a CDMA matched filter (MF) receiver, which is known to be highly suboptimal in the multiuser CDMA environment, particularly when receive powers are widely varied. We have made the MF assumption since this is still the predominant CDMA receiver used in practice, and also the easiest to analyze, since the interference suppression is simply $1/m$ (or similar, depending on the exact codes used). However, there is no question that interference-aware CDMA receivers will, at least in theory, significantly outperform the MF. An enormous number of such receivers have been proposed over the past 20 years, ranging from the maximum likelihood detector (best performance, highest complexity) to linear multiuser detectors (lowest complexity, but questionable robustness), as described in [49, 52, 1] and the references therein. The transmission capacity framework can be adapted to analyze the improvement resulting from such receivers, and we will visit this in Section 5 by considering successive interference cancellation (SIC), with the conclusion that ideal SIC has large gains over DS and even FH, but more realistic SIC has far more modest gains, and might not exceed the capacity of FH in many cases. We specifically would like to caution readers that idealistic assumptions like perfect interference cancellation/suppression are especially

dubious in ad hoc networks, since even a small fraction of residual interference from nearby nodes can constitute a very large amount of interference in the absence of centralized power control or sophisticated MAC scheduling.
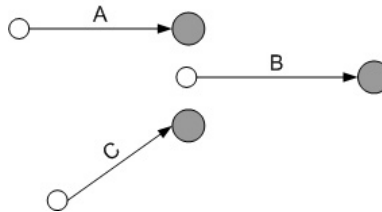
Figure 4.5: Illustration of the near-far problem in ad hoc networks. Transmission B destroys reception of A and C unless an enormous spreading factor is used, or the interference is avoided altogether by frequency hopping or scheduling.

4.2.5 Maximizing transmission capacity

The above analysis is based on the physical model. One can only obtain bounds for the capacity because the close form distribution of the aggregate interference is unknown. In addition, the bounds are only interesting in the low outage regime with $\varepsilon \ll 1$. One might ask, without the outage constraint, what is the spatial intensity of contenders that maximizes the intensity of successful transmissions given a fixed transmission distance d. Specifically will a lack of an outage constraint change the scaling properties of the transmission capacity? Let us use the protocol model in a DS-CDMA network to study the role of contention density for transmission capacity for a wider range of outage regimes. The following analysis can be generalized to a narrow-band network or a particular sub-channel in a FH-CDMA network, by setting $m = 1$.

Dumbbell-like protocol model: modeling concurrent transmissions and spatial reuse in spread spectrum ad hoc networks.

Consider a spatially distributed set of transmitters. We assume that a transmitting node suppresses all receive nodes within a critical range r – see (3.1). Indeed the high interference from such nodes would preclude successful reception. Similarly, a receiver must have no interfering transmitters within a distance r of itself. As discussed earlier, these are only necessary conditions nevertheless good approximations to ensure successful transmissions. Equivalently, as shown in the left panel in Fig. 4.6, each transmission corresponds to a dumbbell with disks of radius $r/2$ at the transmitter and receiver, connected by a bar

of length d. A successful transmission is modeled by a dumbbell without prohibited over-laps, i.e., no transmit disc overlaps with a receive disc on either end.[4] Thus, among the three contending transmissions on the left panel in Fig. 4.6, only two transmissions can be successful. This is shown on the right panel in Fig. 4.6. Spatial reuse corresponds to realizing a high spatial density of dumbbells subject to *at least* satisfying the rules on overlaps. We shall use this dumbbell model to illustrate contention and later clustering phenomena among transmissions. Subsequent MAC designs will not be based on this dumbbell model and our simulations will factor the actual interference seen by receiver, i.e., both near and far field contributions.
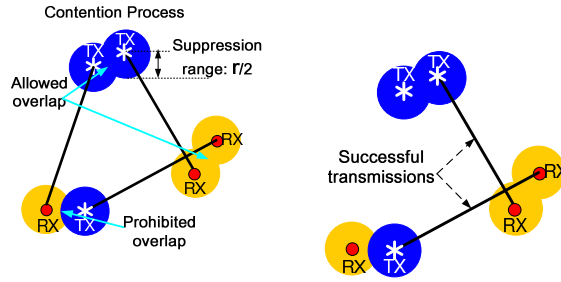
Figure 4.6: On the left contentions among three concurrent transmissions. On the right, after contentions, only those two transmissions whose receivers do not have prohibited overlap survive.

We can show that near-field interference as captured by nodes within a distance r from a receiver, see (3.1), provides a reasonable approximate abstraction for the relevant source of interference, particularly when α is large. To see this, let $\mathbf{B}(x, r)$ denote a ball centered at x with a radius r. We let the event $\mathbf{E_1}$ denote the occurrence that *at least* one interferer is within $\mathbf{B}(O, r)$ which in turn would necessarily cause an outage for a receiver at the origin. It follows that the outage probability for a typical receiver $p_o(\lambda, d)$ is such that

$$p_o(\lambda, d) \geq \mathbb{P}(\mathbf{E_1}), \tag{4.11}$$

For a Poisson point process with intensity λ, the probability of $\mathbf{E_1}$ is given by

$$p_o(\lambda, d) \geq \mathbb{P}(\mathbf{E_1}) = 1 - e^{-\lambda \pi r^2} = 1 - e^{-\lambda \pi (\frac{\beta}{m})^{\frac{2}{\alpha}} d^2}.$$

As shown in Fig. 4.7, in the low outage regime, the outage lower bound provided by $\mathbb{P}(\mathbf{E_1})$ is very accurate. Indeed, one can analytically show this bound has a slope of $\pi d^2 \lambda (\frac{\beta}{m})^{\frac{2}{\alpha}}$ at the point of zero outage, which equals to the slope of exact outage probability solution when $\alpha = 4$ [40]. The slope, i.e., the linearization of the outage probability/lower bound is also

---

[4]Note that, if two such disks overlap, then the associated nodes are within r of each other.

shown in Fig. 4.7, which turns out to be an accurate approximation even in the relatively high outage regime. The point of this analysis is to support the intuitive abstraction for the interference and outage in a network where nodes randomly distributed: nearby interferers within interference range *r* contribute most outage and thus only considering near-field interference, is reasonably accurate. This abstraction allows a simple consideration on how contention among transmitters-receivers occurs.
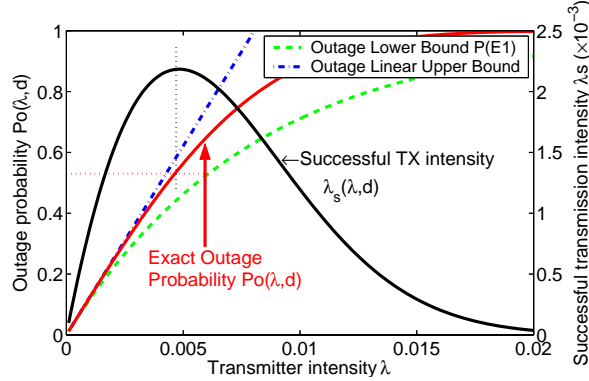
Figure 4.7: Effective capacity λ_s is maximized when $p_o \approx 0.5$. Also shown are our outage lower bound and its linear approximation, which actually serves as a tight outage upper bound. All bounds/approximations are close to the exact analytical result in the low outage regime for $p_o < 0.5$.

Optimal contention density without outage constraint

Fig. 4.7 also exhibits the tradeoff between the outage probability $p_o(\lambda, d)$, the intensity of contenders λ and the intensity of successful transmissions $\lambda_s(\lambda, d) = \lambda(1 - p_o(\lambda, d))$. The figure shows numerical results for $\lambda_s(\lambda, s)$ using the exact analysis of outage probability when $\alpha = 4$ in [40], and the approximate linear upper bound discussed earlier, i.e.,

$$\lambda_s(\lambda, d) \geq \lambda \left(1 - \pi d^2 \lambda (\frac{\beta}{m})^{\frac{2}{\alpha}} \right). \tag{4.12}$$

One can in principle determine the λ which maximizes $\lambda_s(\lambda, d)$, say λ^*, which is marked by the cross in Fig. 4.7 and according to (4.12) is roughly given by

$$\lambda^* = \frac{1}{2\pi d^2}(\frac{m}{\beta})^{\frac{2}{\alpha}} \text{ and } \lambda_s(\lambda^*, d) = \frac{1}{4\pi d^2}(\frac{m}{\beta})^{\frac{2}{\alpha}}. \tag{4.13}$$

Note from Fig. 4.7 to achieve a maximal capacity, one incurs a high outage probability, roughly 0.5. This observation, also holds for the exact analysis in [3], wherein transmitters use a exponentially distributed transmission power and outage probability at the optimal

contender intensity maximizing the capacity is roughly $1 - e^{-1} \approx 0.63$. The key observation here is that maximizing capacity based on a random access MAC will require a high density of transmitters resulting in high spatial reuse but also a high likelihood of outage. Outage in turn may be associated with poor energy efficiency, causing retransmissions and thus increased loads, or congestion collapse, and further compromise packet delay or service quality. In addition, high outage regime does not fundamentally change the scaling property of transmission capacity of CDMA ad hoc networks. One can show that FH-CDMA still has a capacity gain over DS-CDMA in an ad hoc network, via simple exchange of variables in the above analysis. The capacity gain of FH-CDMA over DS-CDMA remains the same in the high outage regime, i.e., the switch of outage regime does not change the basic scaling properties of the capacity in a random channel access MAC protocol.

4.2.6 Summary of ALOHA-like random channel access.

For a FH-CDMA ad hoc network, the reduced contention intensity on each sub-channel suggests that simple contention resolution schemes can be used, e.g., ALOHA-like random channel access, as they generally work well under light load. Indeed by employing a simple random channel access MAC we show that FH-CDMA, due to its interference avoidance feature, achieves a much higher spatial reuse than DS-CDMA ($\alpha > 2$). As the capacity scales linearly in m, i.e., the overall occupied bandwidth of the system, the spatial reuse is considered to be efficient. In a DS-CDMA ad hoc network, since interference from near by interferers can not be fully mitigated (with realistic m), under a random channel access, DS-CDMA achieves worse capacity than FH-CDMA.

Finally we stress that MAC protocols based purely on spatial random access generally only work well under light load. Their performance is not robust in heavy or bursty load as shown in Fig. 4.7, for which, at their best, these protocols can only achieve a moderate spatial reuse but will likely incur a high outage probability and thus poor energy efficiency.

4.2.7 Simple contention resolution and handshaking reduces data collisions but does not increase spatial reuse

The random access MAC model considered above is quite crude, in the sense that transmitter nodes send data without prior signaling and thus some transmissions are lost due to outages. More sophisticated protocols introduce carrier sensing and contention/signaling phases prior to data transmission [12][20][44]. Only the 'winners' of the contention/signaling process subsequently transmit. The goal of contention is to eliminate/defer certain contending transmitters, which will cause excessive interference and likely outage to other receivers if all transmit at the same time. The goal of signaling is to do hand-shaking between a transmitter and receiver to ensure that the intended receiver is indeed available. For example,

41

typical signaling mechanisms use three way hand-shaking. A transmitter sends a RTS mes-
sage to its intended receiver. A receiver which successfully gets a RTS message replies back
with a CTS message, which is hopefully in turn received by the transmitter and suppresses
other neighboring transmitters with high probability. After a successful hand-shake, the
data is transmitted. With such handshaking in place one can ensure receivers are available,
and possibly dramatically reduce the likelihood of outage during actual data transmissions.
Signaling messages can be quite small relative to data packets and thus such mechanisms
are worthwhile to reduce outages on data transmission and save on energy even if this is at
the cost of failures in signaling during the contention/hand-shaking process. An analysis of
the capacity in this case need only factor the density of successful RTS/CTS exchanges one
can achieve, with the outage probability during data transmissions being negligible. Thus,
assuming CTS signals are always successful, the spatial density of successful transmitters
is roughly the same to the random channel access case, i.e., such signaling scheme may not
help much to improve spatial reuse.

More improvement on spatial reuse can be realized by proper back-off strategies,
e.g., by keeping NAV state in 802.11b protocol, which are effective under light loads. Under
heavy or bursty traffic, however, even with back-off, the performance degrades to that of
random channel access[26][22].

In summary, simple variations of RTS/CTS or carrier sensing based narrow-band
MAC protocols fail to exploit the much reduced interference range when using DS-CDMA,
see Fig. 3.5, they are too conservative in contention resolution, generally leading to a poor
spatial reuse. Therefore, there is a need to revisit MAC designs for DS-CDMA for better
spatial reuse performance.

4.3   Idealized scheduling and centralized contention resolution for DS-CDMA

In this section, we study the performance of idealized/centralized MAC scheduling algo-
rithms. These algorithms generally provide much better performance than distributed con-
tention resolution schemes, thanks to available global information and centralized manage-
ment. Although these algorithms are generally not practical to implement, they shed light
on designing efficient MAC for ad hoc networks and the desired performance that practical
designs should aim at.

4.3.1   Joint scheduling and power control

Recently, [10] proposed a joint scheduling and power control algorithm for DS-CDMA ad
hoc networks. As shown in Fig. 4.8, this algorithm has two stages. In stage 1, it thins

the set of intended transmissions by deferring those transmissions that can not be fully handled by power control, i.e., either they severely interfere nearby transmissions or they themselves are severely interfered by their neighbors. In Stage 2, it performs an optimization on the transmit power levels for the remaining set of transmissions. This algorithm repeats the thinning operation in Stage 1 until the Stage 2 optimization is feasible. [10] showed that given the set of intended transmissions, this algorithm achieves close-to optimal performance in terms of the number concurrent successful transmissions and power consumption.
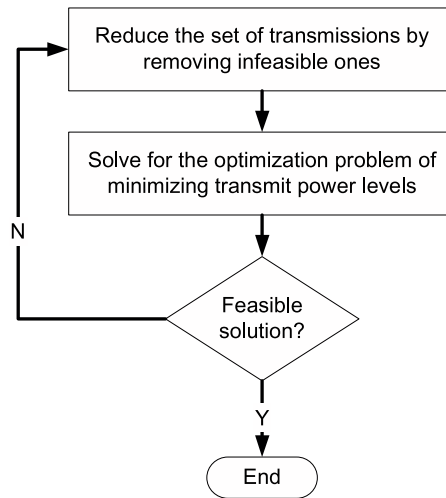
Figure 4.8: Diagram (simplified) of the joint scheduling and power control algorithm in [10].

4.3.2   Centralized contention resolution algorithms

Intuitively MAC contention resolution schemes 'remove' transmissions with prohibited overlaps. For example, consider the realization of the contenders on the top left panel of Fig. 4.9. In this case two receivers have a prohibited overlap and only one transmission will be successfully scheduled by the RTS/CTS handshaking scheme discussed earlier – see the middle figure on the top of Fig. 4.9. Yet a sophisticated contention resolution mechanism could achieve a better spatial reuse. For example, an *idealized* contention resolution process might allow at least one of a set of prohibited overlaps to survive – i.e., remove dumbbells with prohibited overlaps one at a time until no such overlaps were left. A possible result with two successful transmissions of such an idealized scheme is shown on the top right panel of Fig. 4.9. We want to compare the performance of centralized scheduling schemes with ALOHA-like random channel access schemes. Therefore, we simplify the power control part by employing the same fixed transmission power/distance model we used before,

43
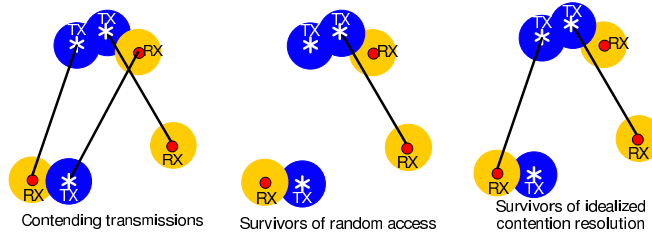
Figure 4.9: On the left panel, an initial contending pattern of 3 transmissions. On the middle panel, only one successful transmissions under random access protocol. On the right panel, two successful transmissions under idealized contention resolution scheme.

and only focus on contention resolution algorithms. Let us consider the following two schemes.

Centralized greedy contention resolution. We consider the centralized greedy contention resolution to approximate the joint scheduling and power control algorithm proposed in [10]. Given a set of contenders, this scheme iteratively examines the subset of remaining transmissions and removes one transmission pair at a time based on which is currently seeing the worst SINR on either its transmitter or receiver side. Contention resolution finishes when all surviving transmissions have sufficient SINR at both receivers and transmitters such that signaling and data transmissions are guaranteed to be successful. Clearly though impractical, such a scheme is close to optimal.

Centralized random algorithm The centralized greedy algorithm discussed above is very computationally expensive even if centralized management is available. For purposes of performance comparison, we consider another centralized schemed – centralized random algorithm. Given a set of contenders, this scheme iteratively examines the subset of remaining transmission pairs and *randomly* removes one transmission pair with an insufficient SINR on either its transmitter or receiver side. Contention resolution finishes when all surviving transmissions have sufficient SINR at both receivers and transmitters such that signaling and data transmissions are guaranteed to be successful.

A simulation of these two idealized schemes is shown in Fig. 4.11. For comparison, we also plot random channel access in Fig. 4.10, with one allowing all transmissions to contend and the other only allowing a density of λ^* defined in (4.13) to contend. For all schemes we discussed, given an initial set of contending transmissions, we plot a subset of successful transmissions. We discover that successful transmitters and receivers tend to be clustered, which is more obvious when the spatial reuse level is high, e.g., see the centralized greedy algorithm. Note that centralized schemes, in particular the centralized greedy algorithm, are generally much better in terms of spatial reuse than simply removing

all transmitter-receiver pairs with prohibited overlaps as the random channel access does, yet it would not be straightforward to implement.[5]
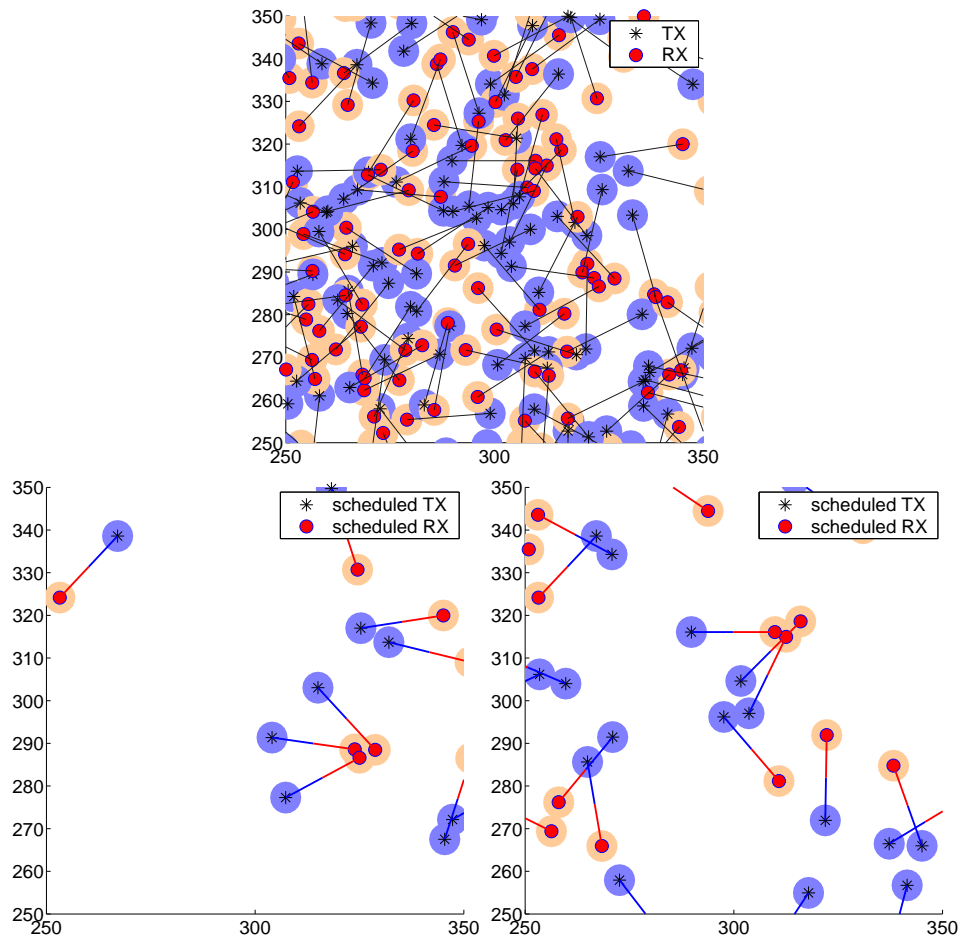
Figure 4.10: On the top panel a realization of contending transmissions in our simulation; on the bottom left panel, the surviving transmissions after random channel access; and on the bottom right panel, the surviving transmissions after thinning the contenders via random channel access, according to the optima contention density in Theorem. 4.13.

---

[5]Implementing such contention resolution in a distributed system requires substantial signaling. Note that signaling also involves contention. Thus achieving successful signaling in turn causes significant overheads per transmission.
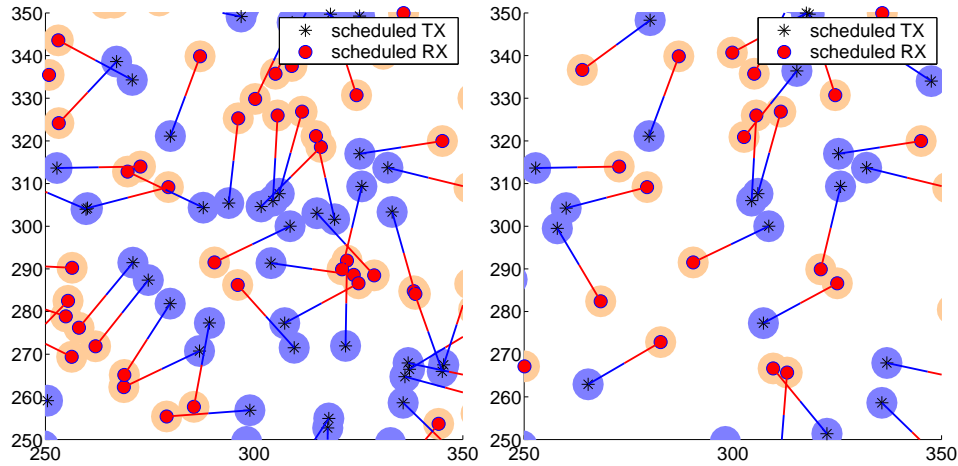
Figure 4.11: On the left panel transmissions surviving an *Centralized Greedy* contention resolution of prohibited overlaps; on the right panel transmissions surviving an *Centralized Random* contention resolution of prohibited overlaps

4.4 Summary

Through theoretical analysis, we show that FH-CDMA works well in traditional narrow-band MAC protocols such as random channel access, since by its nature, FH-CDMA is equivalent to overlaying several orthogonal narrow-band channels. In addition, the actual contention level is low for each individual channel, for which simple contention schemes are generally sufficient to provide good performance.

On the other hand, DS-CDMA performs poorly when used with traditional MAC protocols compared to FH-CDMA. The rest of this dissertation will then focus on improving system designs of DS-CDMA ad hoc networks, addressing specific problems that lead to poor performance of DS-CDMA, and show significant performance gain can be achieved.

One of the factors that contributes to the poor performance of DS-CDMA is the prominent near-far problem in ad hoc networks, which severely penalizes DS-CDMA when quasi-orthogonal codes are used. This motivates us to consider possible ways to inhibit strong interference at receivers, without compromising system capacity or overheads. One of the promising approaches is to use advanced receivers that are capable of canceling interference, at the cost of using additional signal processing resources. We will consider this in Chapter. 5.

Another factor that leads to poor performance of DS-CDMA is that most existing MAC designs do not efficiently leverage the full potential of DS-CDMA. Our simulations systematically exhibited two key aspects of contention-based MAC that have perhaps not been fully appreciated to date. First, as shown in our earlier analysis to achieve a high

46

density of successful transmissions, i.e., a dense packing of dumbbells, one needs to have a high density of contenders. As a result, a significant number of transmitters will need to defer due to contention resolution or see a high outage probability under random access MAC protocol–roughly 50%; this will not lead to the desired high degree of spatial reuse either. The second observation, is that successful transmitters and receivers are *clustered*, in particular when efficient spatial reuse is achieved – see right panel in Fig. 4.9 and simulations in Fig. 4.10, 4.11. This is a unique property of DS-CDMA based ad hoc networks where receivers are capable of interference averaging and thus can tolerate certain level of neighboring interference. This suggests that by explicitly inducing spatial clustering in contention mechanisms, one might further improve spatial reuse. We consider this novel MAC design principle in Chapter. 6.

4.5 Appendix: Proof of Theorem 4.2.1 and Theorem 4.2.2

4.5.1 Proof of Theorem 4.2.1

We first present the following Lemma that will be used in proving Theorem 4.2.1.

Lemma 4.5.1. *Assume transmitters' locations are modeled by a homogeneous Poisson point process Π with intensity λ and transmissions are using fixed transmission power to a distance d. Let $Y(\lambda,s) = \sum_{i \in \Pi \cap \bar{b}(O,s)} |X_i|^{-\alpha}$ denote the normalized far field interference from transmitters outside the circle $b(O,s)$. The mean and variance of $Y(\lambda,s)$ are*

$$
\begin{aligned}
\mathbb{E}[Y(\lambda,s)] &= \frac{2\pi}{\alpha-2}s^{2-\alpha}\lambda = \mu(s)\lambda \\
Var\big(Y(\lambda,s)\big) &= \frac{\pi}{\alpha-1}s^{2(1-\alpha)}\lambda = \sigma^2(s)\lambda.
\end{aligned}
$$

Proof of Lemma 4.5.1 Recall we assume $\alpha > 2$. We compute the mean and variance using

47

Campbell's Theorem [45] as follows:

$$\mathbb{E}\left[\sum_{i\in\Pi\cap\bar{b}(O,s)}|X_i|^{-\alpha}\right] = 2\pi\lambda\int_s^\infty r^{-\alpha}r\,dr$$

$$= \frac{2\pi}{2-\alpha}\lambda r^{2-\alpha}\Big|_s^\infty$$

$$= \frac{2\pi}{2-\alpha}\lambda(0-s^{2-\alpha})$$

$$= \frac{2\pi}{\alpha-2}s^{2-\alpha}\lambda.$$

$$Var\left(\sum_{i\in\Pi\cap\bar{b}(O,s)}|X_i|^{-\alpha}\right) = 2\pi\lambda\int_s^\infty r^{-2\alpha}r\,dr$$

$$= \frac{2\pi}{2(1-\alpha)}\lambda r^{2(1-\alpha)}\Big|_s^\infty$$

$$= \frac{\pi}{1-\alpha}\lambda(0-s^{2(1-\alpha)})$$

$$= \frac{\pi}{\alpha-1}s^{2(1-\alpha)}\lambda.$$

$$\blacksquare$$

Proof of Theorem 4.2.1. We first consider the DS-CDMA case. According to Definition 4.2.1, $E_u(\lambda,s)$ consists of all outcomes where there are one or more transmitters within distance s of the origin O, and the event $E_f(\lambda,s)$ consists of all outcomes where the set of transmitters outside the ball $b(O,s)$ generate enough interference power to cause an outage at the origin. Then one can show the following properties.

Property 4.5.1. *The events in Definition 4.2.1 satisfy the following properties:*

a) *For* $s \le (M\kappa)^{-\frac{1}{\alpha}}$, $E_u(\lambda,s) \subset E(\lambda) \subseteq (E_u(\lambda,s) \cup E_f(\lambda,s)) = E_l(\lambda,s)$, $\forall\ 0 < s < (M\kappa)^{-\frac{2}{\alpha}}$.

b) *Each of* $\mathbb{P}^0(E(\lambda))$, $\mathbb{P}^0(E_u(\lambda,s))$, *and* $\mathbb{P}^0(E_f(\lambda,s))$ *increases in* λ *for fixed s.*

c) $\mathbb{P}^0(E_u(\lambda,s))$ *is increasing in s, while* $\mathbb{P}^0(E_f(\lambda,s))$ *is decreasing in s for fixed* λ.

d) $E_u(\lambda,s)$ *and* $E_f(\lambda,s)$ *are independent events.*

To prove Property *a*, consider the following facts: (*i*) If $s \le s_{max} = (M\kappa)^{-\frac{1}{\alpha}}$ then even if there is only one node in $b(O,s)$, and even if that one node is as far away from the origin as possible, i.e., $|X_i| = s$, then the normalized interference generated by that node, $s^{-\alpha} \ge \left((M\kappa)^{-\frac{1}{\alpha}}\right)^{-\alpha} = M\kappa$, is still sufficient to cause an outage. This proves the statement.

(*ii*) Consider an outage outcome $\omega \in E(\lambda)$. Suppose $\omega \in E(\lambda)$ and $\omega \notin E_u(\lambda, s)$. Then ω constitutes an outage but there are no nodes in $b(O, s)$. Then clearly the interference is caused by nodes in $\bar{b}(O, s)$, which means $\omega \in E_f(\lambda, s)$. Suppose $\omega \in E(\lambda)$ and $\omega \notin E_f(\lambda, s)$. Then ω constitutes an outage but the external interference generated by nodes in $\bar{b}(O, s)$ is insufficient to cause outage. Then this means there are one or more transmitters in $b(O, s)$, which means $\omega \in E_u(\lambda, s)$. Thus, $\omega \in E(\lambda)$ implies either $\omega \in E_u(\lambda, s)$ or $\omega \in E_f(\lambda, s)$, which is equivalent to saying $\omega \in (E_u(\lambda, s) \cup E_f(\lambda, s))$.

Property *b* and *c* are straightforward by considering the number of interfering transmitters considered in each event.

To prove Property *d*, recall that $E_u(\lambda, s)$ only depends upon the space $b(O, s)$ and $E_f(\lambda, s)$ only depends upon the space $\bar{b}(O, s)$. The independence property of the Poisson point process states that the number of points $N(A)$ and $N(B)$ in disjoint regions A and B are independent random variables, hence $E_u(\lambda, s), E_f(\lambda, s)$ are independent events.

The upper bound λ_u^ε. We need to calculate $\mathbb{P}^0(E_u(\lambda, s))$. Consider the probability that there are no transmitters in $b(O, s)$, which is simply the void probability for $b(O, s)$ [45]. For a Poisson point process in the plane with intensity λ, the void probability for a set $A \subset \mathbb{R}^2$ is $e^{-\lambda \nu(A)}$, where $\nu(\cdot)$ denotes the Lebesgue measure of the set contained in the argument. Thus,

$$\mathbb{P}^0(E_u(\lambda, s)) = 1 - \mathbb{P}^0(\Pi \cap b(O, s) = \emptyset) = 1 - e^{-\lambda \pi s^2}, \tag{4.14}$$

for $s \in (0, (M\kappa)^{-\frac{2}{\alpha}})$.

Now given the outage constraint $\mathbb{P}^0(E_u(\lambda, s)) \leq \varepsilon$, we obtain the parameterized upper bound

$$\lambda_u^\varepsilon(s) = -\frac{1}{\pi} s^{-2} \ln(1 - \varepsilon),$$

for all $s \in (0, (M\kappa)^{-\frac{2}{\alpha}})$ and all $\varepsilon \in (0, 1)$. We further optimize this bound over s and find that the tightest (smallest) upper bound is obtained by choosing the largest possible $s = (M\kappa)^{-\frac{1}{\alpha}}$. Thus, the final upper bound on the optimal contention density is

$$\begin{aligned}
\lambda_u^\varepsilon &= -\frac{1}{\pi}(M\kappa)^{\frac{2}{\alpha}} \ln(1 - \varepsilon) \\
&= \frac{1}{\pi}(M\kappa)^{\frac{2}{\alpha}} \varepsilon + \Theta(\varepsilon^2),
\end{aligned}$$

for all $\varepsilon \in (0, 1)$.

The lower bound λ_l^ε. We need to calculate $\mathbb{P}^0(E_l(\lambda,s))$. Because of Property d, we have

$$
\begin{aligned}
\mathbb{P}^0(E_l(\lambda,s)) &= \mathbb{P}^0(E_u(\lambda,s) \cup E_f(\lambda,s)) \\
&= \mathbb{P}^0(E_u(\lambda,s)) + \mathbb{P}^0(E_f(\lambda,s)) \\
&\quad - \mathbb{P}^0(E_u(\lambda,s))\mathbb{P}^0(E_f(\lambda,s)).
\end{aligned}
$$

The outage constraint $\mathbb{P}^0(E(\lambda)) < \varepsilon$ then can be rewritten as

$$
\mathbb{P}^0(E_u(\lambda)) < \varepsilon_1, \quad \mathbb{P}^0(E_f(\lambda)) \le \varepsilon_2, \tag{4.15}
$$

for some ε_1 and ε_2 such that $\varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2 = \varepsilon$.

Given the constants ε_1 and ε_2, define the contention density imposed by conditions in (4.15) to be

$$
\begin{aligned}
\lambda_u^{\varepsilon_1}(s) &= \max\{\lambda \,|\, \mathbb{P}^0(E_u(\lambda,s)) \le \varepsilon_1\}, \\
\lambda_f^{\varepsilon_2}(s) &= \max\{\lambda \,|\, \mathbb{P}^0(E_f(\lambda,s)) \le \varepsilon_2\}.
\end{aligned}
$$

Thus the lower bound of the optimal contention density can be derived as

$$
\lambda_l^\varepsilon = \max_{s,(\varepsilon_1,\varepsilon_2)} \left\{ \min\{\lambda_u^{\varepsilon_1}(s), \lambda_f^{\varepsilon_2}(s)\} \right\}, \tag{4.16}
$$

for $s \in [0, (M\kappa)^{-\frac{2}{\alpha}}]$ and $(\varepsilon_1,\varepsilon_2)$ satisfying $\varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2 = \varepsilon$. To obtain the lower bound we will consider $\lambda_u^{\varepsilon_1}(s)$ and $\lambda_f^{\varepsilon_2}(s)$ separately first. And then choose to maximize the minimum of the two for all feasible s and choices of $(\varepsilon_1,\varepsilon_2)$ pairs. It can be seen that if we increase $\mathbb{P}^0(E_u(\lambda,s))$, $\mathbb{P}^0(E_f(\lambda,s))$ decreases, i.e., changing s or $(\varepsilon_1,\varepsilon_2)$ must increase one of $\lambda_u^{\varepsilon_1}(s)$ and $\lambda_f^{\varepsilon_2}(s)$ but decrease the other one at the same time. Thus according to (4.16), a necessary condition for the optimized lower bound is that the choice of s and $(\varepsilon_1,\varepsilon_2)$ is such that $\lambda_u^{\varepsilon_1}(s) = \lambda_f^{\varepsilon_2}(s)$ if this is feasible.

Based on our calculation of the upper bound, we have that $\lambda_u^{\varepsilon_1}(s) = -\frac{1}{\pi}s^{-2}\ln(1 - \varepsilon_1) = \frac{1}{\pi}s^{-2}\varepsilon_1 + \Theta(\varepsilon_1^2)$.

$\lambda_f^{\varepsilon_1}(s)$ is not straightforward to calculate. Thus we resort to finding its lower bound using Chebychev's inequality. To this end, we use our estimates for the mean and variance of the (normalized) far field interference Y defined earlier in Lemma 4.5.1. The lower bound

on $\lambda_f^{\varepsilon_1}(s)$ is obtained as follows:

$$
\begin{aligned}
\lambda_f^{\varepsilon_2}(s) &= \max\left\{\lambda \,\Big|\, \mathbb{P}^0\big(E_f(\lambda,s)\big) \le \varepsilon_2\right\} \\
&= \max\left\{\lambda \,\Big|\, \mathbb{P}^0\Big(\sum_{i \in \Pi \cap \bar{b}(O,s)} |X_i|^{-\alpha} \ge M\kappa\Big) \le \varepsilon_2\right\} \\
&= \max\left\{\lambda \,\Big|\, \mathbb{P}^0\big(Y(\lambda,s) \ge M\kappa\big) \le \varepsilon_2\right\} \\
&\ge \max\left\{\lambda \,\Big|\, \mathbb{P}^0\Big(\frac{|Y(\lambda,s)-\mu(s)\lambda|}{M\kappa-\mu(s)\lambda} \ge 1\Big) \le \varepsilon_2\right\} \\
&\ge \max\left\{\lambda \,\Big|\, \frac{\sigma^2(s)\lambda}{(M\kappa-\mu(s)\lambda)^2} \le \varepsilon_2\right\} \\
&= \underline{\lambda}_f^{\varepsilon_2}(s).
\end{aligned}
$$

The first three equalities are by definition, the second inequality is Chebychev bound. Clearly $\underline{\lambda}_f^{\varepsilon_2}(s)$ is achieved when $\frac{\sigma^2(s)\lambda}{(M\kappa-\mu(s)\lambda)^2} = \varepsilon_2$. By solving this equation and keeping the dominant term under the condition that ε is small, we obtain

$$
\begin{aligned}
\underline{\lambda}_f^{\varepsilon_2}(s) &= \frac{(M\kappa)^2}{\sigma^2(s)}\varepsilon_2 + \Theta(\varepsilon_2^2) \\
&= \frac{(\alpha-1)(M\kappa)^2}{\pi}s^{2(\alpha-1)}\varepsilon_2 + \Theta(\varepsilon_2^2)
\end{aligned}
$$

We now move on optimizing the parameterized bound over s and $(\varepsilon_1,\varepsilon_2)$. In the calculation of $\mathbb{P}^0(E_u)$ and $\mathbb{P}^0(E_u)$, only Chebychev's inequality introduces estimation error. If we minimize this error, we optimize our lower bound. It can be seen that the estimation error decreases in s. Thus we need to choose the largest possible $s = (M\kappa)^{-\frac{1}{\alpha}}$. In addition, as mentioned before, we need to choose ε_1 and ε_2 properly in order to equalize $\lambda_u^{\varepsilon_1}(s)$ and $\underline{\lambda}_f^{\varepsilon_2}(s)$ to obtain the tightest lower bound. Therefore, given $s = (M\kappa)^{-\frac{1}{\alpha}}$, we have

$$
\begin{aligned}
\lambda_u^{\varepsilon_1}(s)\Big|_{s=(M\kappa)^{-\frac{1}{\alpha}}} &= \frac{1}{\pi}(M\kappa)^{\frac{2}{\alpha}}\varepsilon_1 + \Theta(\varepsilon_1^2), \\
\underline{\lambda}_f^{\varepsilon_2}(s)\Big|_{s=(M\kappa)^{-\frac{1}{\alpha}}} &= \frac{(\alpha-1)(M\kappa)^{\frac{2}{\alpha}}}{\pi}\varepsilon_2 + \Theta(\varepsilon_2^2)
\end{aligned}
$$

In order to equalize $\lambda_u^{\varepsilon_1}(s)$ and $\underline{\lambda}_f^{\varepsilon_2}(s)$, we must have $\varepsilon_2 = \frac{\alpha-\sqrt{\alpha^2-4(\alpha-1)\varepsilon}}{2(\alpha-1)} = \frac{1}{\alpha}\varepsilon + \Theta(\varepsilon^2)$ and $\varepsilon_1 = (\alpha-1)\varepsilon_2 = \frac{\alpha-1}{\alpha}\varepsilon + \Theta(\varepsilon^2)$.

Thus our lower bound on the optimal contention density is

$$
\lambda_l^{\varepsilon} = h(\alpha)(M\kappa)^{\frac{2}{\alpha}} + \Theta(\varepsilon^2), \tag{4.17}
$$

where $h(\alpha) = \frac{\alpha-1}{\alpha}$.

Looking at equations (4.3) and (4.4), it's clear that the exact same analysis for FH-CDMA holds provided we replace $M\lambda$ with λ and $M\kappa$ with κ, i.e., if

$$Mh(\alpha)\frac{1}{\pi}(M\kappa)^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2) \leq M\lambda^{\varepsilon,DS} \leq M\frac{1}{\pi}(M\kappa)^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2)$$

holds for DS, then,

$$h(\alpha)\frac{1}{\pi}M\kappa^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2) \leq \lambda^{\varepsilon,FH} \leq \frac{1}{\pi}M\kappa^{\frac{2}{\alpha}}\varepsilon + \Theta(\varepsilon^2)$$

holds for FH. ∎

4.5.2 Proof of Theorem. 4.2.2

Before proving Theorem 4.2.2, we first show the following Lemma that will be used in the proof later.

Lemma 4.5.2. *In the pairwise power control model, assume uniformly distributed transmission distance D_i with $F_D(d) = \frac{d^2-1}{d_{max}^2-1}$ for $d \in [1, d_{max}]$. Let $Y'(\lambda, s) = \sum_{i \in \Phi_m \cap \bar{b}(O,s)} \left(\frac{D_i}{|X_i|}\right)^\alpha$ denote the normalized far field interference. The mean and variance of the random variable $Y'(\lambda, s)$ are*

$$\mathbb{E}[Y'(\lambda, s)] = \frac{4\pi(d_{max}^{\alpha+2}-1)s^{2-\alpha}}{M(\alpha-2)(\alpha+2)(d_{max}^2-1)}\lambda = \mu(s)\lambda,$$

$$Var(Y'(\lambda, s)) = \frac{\pi(d_{max}^{2\alpha+2}-1)s^{2(1-\alpha)}}{M(\alpha-1)(\alpha+1)(d_{max}^2-1)}\lambda = \sigma^2(s)\lambda.$$

Proof of Lemma 4.5.2 Recall we assume $\alpha > 2$. We compute the mean and variance using

Campbell's Theorem as follows:

$$
\begin{aligned}
\mathbb{E}[Y'(\lambda,s)] &= 2\pi\frac{\lambda}{M}\int_s^\infty r^{-\alpha}r\,dr\int_1^{d_{max}} x^\alpha dF_D(x) \\
&= 2\pi\frac{\lambda}{M}\int_s^\infty r^{1-\alpha}dr\int_1^{d_{max}}\frac{2x^{1+\alpha}}{d_{max}^2-1}dx \\
&= 2\pi\frac{\lambda}{M}\frac{s^{2-\alpha}}{\alpha-2}\frac{2}{d_{max}^2-1}\frac{d_{max}^{2+\alpha}-1}{2+\alpha} \\
&= \frac{4\pi(d_{max}^{\alpha+2}-1)s^{2-\alpha}}{M(\alpha-2)(\alpha+2)(d_{max}^2-1)}\lambda \\
Var\big(Y'(\lambda,s)\big) &= 2\pi\frac{\lambda}{M}\int_s^\infty r^{-2\alpha}r\,dr\int_1^{d_{max}} x^{2\alpha}dF_D(x) \\
&= 2\pi\frac{\lambda}{M}\frac{s^{2-2\alpha}}{2(\alpha-1)}\frac{2}{d_{max}^2-1}\frac{d_{max}^{2+2\alpha}-1}{2(\alpha+1)} \\
&= \frac{\pi(d_{max}^{2\alpha+2}-1)s^{2(1-\alpha)}}{M(\alpha-1)(\alpha+1)(d_{max}^2-1)}\lambda.
\end{aligned}
$$

∎

Proof of Theorem 4.2.2. We first prove the FH-CDMA case. Consider a particular sub-channel $m \in \{1,\ldots,M\}$ used in FH-CDMA. According to Definition 4.2.2, the event $E(\lambda)$ consists of all outage outcomes. The event $E_u(\lambda,s)$ consists of all outcomes where there are one or more transmitters within s of the origin with transmission distances exceeding $s\delta^{\frac{1}{\alpha}}$. This threshold is the smallest transmission distance such that even one transmitter in $b(O,s)$ with such a transmission distance will cause an outage at the origin. The event $E_{l_1}(\lambda,s)$ consists of all outcomes with one or more transmitters in $b(O,s)$; but note that not all outcomes in $E_{l_1}(\lambda,s)$ will cause an outage. Finally, the event $E_{l_2}(\lambda,s)$ consists of all outcomes where the interference power at the origin caused by all the transmitters outside $b(O,s)$ is adequate to cause an outage at the origin. Note that the events in Definition 4.2.2 have similar properties as the Properties $a-d$ mentioned in the proof of Theorem 4.2.1. In particular for $s \leq d_{max}\delta^{-\frac{1}{\alpha}}$, $E_u(\lambda,s) \subset E(\lambda) \subseteq E_l(\lambda,s) = E_{l_1}(\lambda,s) \cup E_{l_2}(\lambda,s)$ and $E_{l_1}(\lambda,s)$ and $E_{l_2}(\lambda,s)$ are two independent events.

The upper bound λ_u^ε. To obtain the upper bound, we need to calculate $\mathbb{P}^0\big(E_u(\lambda,s)\big)$. Since the marks are independent of the point locations, the points with certain marks form a thinned process, where the thinning is proportional to the mark probability, i.e., the intensity of the thinned process is $\frac{\lambda}{M}\bar{F}_D\big(s\delta^{\frac{1}{\alpha}}\big)$. Thus the probability of event $E_u(\lambda,s)$ is given by

$$
\mathbb{P}^0\big(E_u(\lambda,s)\big) = 1 - \exp\Big\{-\frac{\lambda}{M}\bar{F}_D\big(s\delta^{\frac{1}{\alpha}}\big)\pi s^2\Big\},
$$

for $s \in (0, d_{max}\delta^{-\frac{1}{\alpha}})$. We can derive the parameterized upper bound $\lambda_u^{\varepsilon}(s)$

$$\lambda_u^{\varepsilon}(s) = -\frac{M\ln(1-\varepsilon)}{\pi\bar{F}_D(s\delta^{\frac{1}{\alpha}})s^2} = \frac{M\varepsilon}{\pi\bar{F}_D(s\delta^{\frac{1}{\alpha}})s^2} + \Theta(\varepsilon^2),$$

for all $\varepsilon \in (0,1)$.

To get the tightest upper bound, we need to optimize this result over feasible s. Note that $\mathbb{P}^0(E_u(\lambda,s))$ is not monotone increasing in s as in the case of model 1. In particular, $\mathbb{P}^0(E_u(\lambda,s))$ is increasing in s for s small, and decreasing in s as s approaches $d_{max}\delta^{-\frac{1}{\alpha}}$. Intuitively, for s small we can accept any mark but the circle $b(O,s)$ is small, while for s large the circle $b(O,s)$ is large enough but only the largest marks may be admitted. Thus, $\lambda_u^{\varepsilon}(s)$ is also not monotone in s: $\lambda_u^{\varepsilon}(s)$ is large for s small and s near s_{max}.

Consider the example with randomly selected S-D pairs in space with distance uniform in $[1, d_{max}]$, i.e., $F_D(d) = \frac{d^2-1}{d_{max}^2-1}$. By taking derivative of $\lambda_u^{\varepsilon}(s)$ with respect to s and setting it to 0, we can solve for the minimizer $s_u^* = \frac{d_{max}}{\sqrt{2}\delta^{\frac{1}{\alpha}}}$, which give us the tightest upper bound

$$\lambda_u^{\varepsilon} = \frac{4M\delta^{\frac{2}{\alpha}}(d_{max}^2-1)}{\pi d_{max}^4}\varepsilon + \Theta(\varepsilon^2),$$

for all $\varepsilon \in (0,1)$.

The lower bound λ_l^{ε}. Same as the calculation of $\lambda_l^{\varepsilon_u}(s)$ in Theorem 4.2.1, we have

$$\lambda_{l_1}^{\varepsilon_1}(s) = -\frac{M}{\pi}s^{-2}\ln\left(1-\varepsilon_1\right)$$

for all $s > 0$ and $\varepsilon \in (0,1)$.

We also need to bound $\mathbb{P}^0(E_{l_2}(\lambda,s))$ and thus obtain lower bound on $\lambda_{l_2}^{\varepsilon_1}(s)$. This is done in a similar way to what we did for model 1. We apply Chebychev's inequality, for which Lemma 4.5.2 provides the mean and variance for the normalized far field interference Y', and we obtain the lower bound on $\lambda_{l_2}^{\varepsilon_2}(s)$ as follows:

$$\begin{aligned}
\lambda_{l_2}^{\varepsilon_2}(s) &\geq \frac{(\alpha-1)(\alpha+1)M\delta^2 s^{2(\alpha-1)}(d_{max}^2-1)}{\pi(d_{max}^{2\alpha+2}-1)}\varepsilon_2 + \Theta(\varepsilon_2^2) \\
&= \underline{\lambda}_{l_2}^{\varepsilon_2}(s).
\end{aligned}$$

To obtain λ_l^{ε}, we again need to optimize $\lambda_l^{\varepsilon} = \max_{s>0,(\varepsilon_1,\varepsilon_2)}\left\{\min\left\{\lambda_{l_1}^{\varepsilon_1}(s), \underline{\lambda}_{l_2}^{\varepsilon_2}(s)\right\}\right\}$ over s and $(\varepsilon_1,\varepsilon_2)$. As previously noted, in this model, $\mathbb{P}^0(E_{l_1}(s))$ is no longer the exact outage probability cause by near field interference, but only a conservative estimation. One really have to optimize s and $(\varepsilon_1,\varepsilon_2)$ concurrently, which is very complicated. For simplicity we instead choose to set $\varepsilon_1 = \varepsilon_2 = 1 - \sqrt{1-\varepsilon} = \frac{\varepsilon}{2} + \Theta(\varepsilon^2)$ and then only optimize over

s to equalize $\lambda_{l_1}^{\varepsilon_1}(s)$ and $\lambda_{l_2}^{\varepsilon_2}(s)$, for which we obtain

$$\lambda_l^{\varepsilon} \geq g(\alpha)\frac{M\delta^{\frac{2}{\alpha}}}{\pi}\left(\frac{d_{max}^2-1}{d_{max}^{2\alpha+2}-1}\right)^{\frac{1}{\alpha}}\varepsilon+\Theta(\varepsilon^2),$$

where $g(\alpha) = \frac{1}{2}\left((\alpha-1)(\alpha+1)\right)^{\frac{1}{\alpha}}$.

Looking at equations (4.9) and (4.10), it's clear that the exact same analysis for DS-CDMA holds provided we replace λ with $M\lambda$ and δ with $M\delta$, i.e., if

$$\lambda_l^{\varepsilon,FH} \geq g(\alpha)\frac{M\delta^{\frac{2}{\alpha}}}{\pi}\left(\frac{d_{max}^2-1}{d_{max}^{2\alpha+2}-1}\right)^{\frac{1}{\alpha}}\varepsilon+\Theta(\varepsilon^2)$$

$$\lambda_u^{\varepsilon,FH} = \frac{4M\delta^{\frac{2}{\alpha}}(d_{max}^2-1)}{\pi d_{max}^4}\varepsilon+\Theta(\varepsilon^2)$$

holds for FH, then,

$$\lambda_l^{\varepsilon,DS} \geq g(\alpha)\frac{M^{\frac{2}{\alpha}}\delta^{\frac{2}{\alpha}}}{\pi}\left(\frac{d_{max}^2-1}{d_{max}^{2\alpha+2}-1}\right)^{\frac{1}{\alpha}}\varepsilon+\Theta(\varepsilon^2)$$

$$\lambda_u^{\varepsilon,DS} = \frac{4}{\pi}\frac{M^{\frac{2}{\alpha}}\delta^{\frac{2}{\alpha}}(d_{max}^2-1)}{d_{max}^4}\varepsilon+\Theta(\varepsilon^2)$$

holds for DS. ∎

55

Chapter 5

Employing Successive Interference Cancelation at Receivers

5.1   General idea of interference cancelation

Since ad hoc networks are interference-limited, interference suppression naturally increases the communication quality and allows for denser networks. Intuitively, strong interference is worse than weak interference, but when interference-aware receivers are employed, strong interference may actually be preferable and result in an increase in capacity [6, 8]. Since these early works, an entire research field has emerged on multiuser receivers particularly for CDMA systems [49]. But until recently, for a variety of reasons, these techniques have not been widely deployed [1]. Employing advanced signal processing will become increasingly attractive with the progression of Moore's Law, and it is a natural extension of the simple matched filter DS-CDMA receiver that has been assumed up until this point in the dissertation.

In contrast to the matched-filter receiver that treats the signals of other users as background noise, successive interference cancelation (SIC) is a nonlinear type of multiuser receiver in which users are decoded one after another, with the interference from the prior user subtracted before proceeding on to the next user. This simple technique has been proven to achieve the Shannon capacity region boundaries for both the downlink [9] and uplink cellular scenarios [35].

A simple SIC receiver is shown in Fig. 5.1. SIC is especially promising for ad hoc networks since it is well-suited to asynchronous signals of unequal powers [2] [51]. Note that since SIC cancels interference iteratively, the cancelation order is critical to the performance of SIC. A receiver should cancel interference from interferers in the order of decreasing interference power levels, in order to obtain the best SINR after SIC. Note that cancelation error in terms of residual interference power after cancelation is inevitable
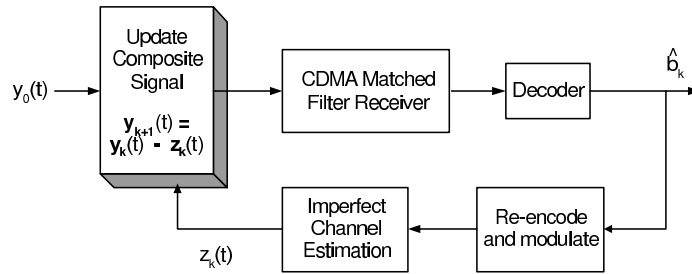
Figure 5.1: Successive interference cancelation

mostly due to inaccurate estimation of the phase and amplitude of interference instead of wrong estimation of bit decision in previous decoded/canceled users because bit error rate (BER) is generally assumed to be low. Recent research shows the typical cancelation error is roughly 10%[17].

Employing SIC at receivers incurs an extra cost in hardware as well as network management. Fortunately, hardware for SIC in DS-CDMA is relatively simple with much smaller complexity than other interference cancelation implementations and does not increase with the number of interferers a receiver chooses to cancel. To perform cancelation, receivers also need to acquire codes used by interferers. Thus each node need maintain certain state information about neighbors, i.e., their code and relative distance corresponding to potential interference power level. Practical SIC performance will rely on the accuracy of such state information.

5.2 DS-CDMA with SIC under random channel contention

5.2.1 Modeling SIC performance.

Successive interference cancellation allows users to be decoded one at a time, and then subtracted out from the composite received signal in order to improve the performance of subsequently decoded users. In practice, this corresponds to decoding the strongest user first, since it will experience the best SINR and hence be the most accurately decoded, which is a prerequisite for accurate interference cancellation. More generally, by similar reasoning, users should be decoded in order of their received powers [51, 53], even though this is not always the preferred order from an information theoretic viewpoint [48]. In an ad hoc network with a path loss channel model, this corresponds to canceling the interference from nodes closer to the receiver than the desired transmitter.

An accurate characterization of the performance gains due to SIC should be based on a plausible interference cancellation scenario, otherwise the results can in fact be quite optimistic and misleading. Particularly, an accurate model should capture that a SIC-

57

equipped receiver is able to reduce the interference power of up to K nearest interfering nodes by a factor ζ (i.e. residual interference power of ζ), assuming these nodes are closer than our desired transmitter, since it has been shown that a particular (unequal) distribution of received powers is needed for SIC systems to perform well, especially when the interference cancellation is imperfect [5, 2]. However, it is difficult to work with this exact model in a mathematical framework, since it requires a characterization of the joint distribution of the distances of the K nodes nearest to the origin, see [46].

Instead of pursuing this exact approach, we utilize a closely-related SIC model that is more amenable to analysis. In particular, define the *cancellation radius*, denoted r_{sic}, such that the receiver is capable of reducing the interference power by ζ of any and all transmitters located within distance r_{sic} of it. As shown in Fig. 5.2, this approximation corresponds to a modified version of the stochastic geometric model we used in Section. 4, in which interferers within the circle of radius r_{sic} only impose a fraction, i.e., ζ, of the interference power. The cancellation radius is chosen so that there are K interfering nodes falling within the radius *on average*. Since the average number of points in a Poisson point process of intensity λ falling in a circle of radius r is $\lambda \pi r^2$, we find $r_{\text{sic}} = \sqrt{\frac{K}{\pi \lambda}}$. It is normally only feasible to cancel the interference from those nodes whose interference power measured at the receiver exceeds the signal power. Thus we add the requirement that $r_{\text{sic}} \leq d$, where d is the distance from a transmitter to its intended receiver, i.e., by requiring the cancellation radius not exceed the signal transmission radius we are ensuring the interference power of cancelled nodes exceeds the signal power.

Definition 5.2.1. A (K, ζ) *SIC receiver* operating in a network with a transmission density of λ is capable of reducing the interference power by a factor $1 - \zeta$ for all interfering nodes within distance $r_{\text{sic}} = \min(d, \sqrt{\frac{K}{\pi \lambda}})$.

Note that $r_{\text{sic}} = d$ for $\lambda \leq \lambda_r = \frac{K}{\pi} d^{-2}$, and is decreasing in λ for $\lambda > \lambda_r$. Put simply, for low densities all the nodes closer than the desired transmitter are cancellable. For higher densities we can only cancel the closest K nodes (on average), which are inside $r_{\text{sic}} \leq d$.

Now, let $b(O, r) = \{x : |x| \leq r\}$ be the ball of radius r centered at the origin, and let $\bar{b}(O, r) = \mathbb{R}^2 \setminus b(O, r)$. The appropriate modification to Definition 4.2.1 that allows for SIC is as follows.

Definition 5.2.2. The *optimal contention density* for a network of (K, ζ) capable receivers, $\lambda^{\varepsilon, \text{sic}}$, is the maximum spatial density of nodes that can contend for the channel subject to the constraint that the typical outage probability is less than ε for some $\varepsilon \in (0, 1)$:

$$\lambda^{\varepsilon, \text{sic}} = \sup\left\{\lambda : \mathbb{P}^0\left(\frac{\rho r^{-\alpha}}{Y(\lambda)} \leq \frac{\beta}{m}\right) \leq \varepsilon\right\}, \tag{5.1}$$
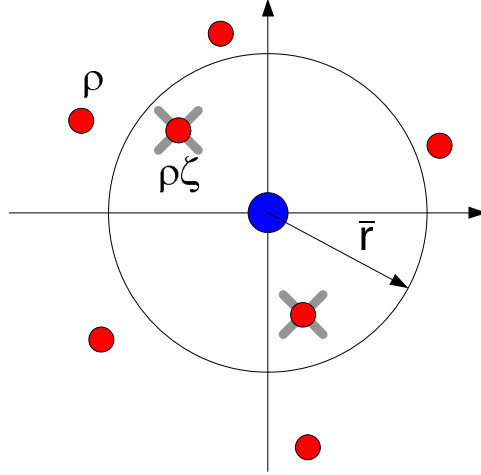
Figure 5.2: Stochastic geometric model for non-perfect successive interference cancelation, in which the interference from interferes inside $B(O, \bar{r})$ is effectively thinned by a factor of ζ.

where

$$Y(\lambda) = \zeta \sum_{i \in \Pi \cap b(O, r_{\text{sic}})} \rho |X_i|^{-\alpha} + \sum_{i \in \Pi \cap \bar{b}(O, r_{\text{sic}})} \rho |X_i|^{-\alpha}. \tag{5.2}$$

As in Chapter 4, we let the transmission capacity is thus given by $c^{\varepsilon, \text{sic}} = \lambda^{\varepsilon, \text{sic}}(1 - \varepsilon)$.

The first term is the partially cancelled normalized aggregate interference seen at the receiver at the origin from all nodes lying within the cancellation radius; the second term is the uncancelled interference from nodes lying outside that set.

We consider three different types of CDMA receivers, where ζ is the fractional interference left after performing interference cancelation:

1. Conventional DS-CDMA, $\zeta = 1$, which corresponds to the case of Theorem 4.2.1.

2. DS-CDMA with perfect interference cancelation of the strongest K users, i.e. $(K, \zeta = 0)$.

3. DS-CDMA with imperfect cancelation of the strongest K users, (K, ζ).

5.2.2 Capacity analysis of DS-CDMA ad hoc network employing SIC

Recall that the function $h(\alpha) = 1 - \frac{1}{\alpha}$ and $\kappa = \frac{d^{-\alpha}}{\beta}$[1]. Define the normalized SINR threshold $\bar{\beta} = \beta/m$. Note that $0.5 < h(\alpha) < 0.8$ for $2 \le \alpha \le 6$, which constitutes the usual accepted

---

[1]Here we have ignored the ambient noise for the simplicity of results.

range for the path loss exponent [34].

Theorem 5.2.1. *Under the simplified stochastic geometric model, in which we assume a receiver can cancel all interferers within a cancelation radius* $r_{\text{sic}} = \sqrt{\frac{K}{\pi\lambda}}$, *the upper and lower bounds for the transmission capacity for the following four SIC cases:*

1. *Conventional DS-CDMA without SIC (nsic), i.e.,* $\zeta = 1$,

2. *DS-CDMA with perfect interference cancelation of the strongest K users (psic), i.e.* $(K, \zeta = 0)$.

3. *and DS-CDMA with imperfect cancelation of the strongest K users (sic),* (K, ζ).

are given by

    Conventional CDMA, $(K = \infty, \zeta = 1)$:

$$c_{l,C}^{\varepsilon,\text{nsic}} = \lambda_{l,C}^{\varepsilon,\text{nsic}}(1-\varepsilon) = h(\alpha)\frac{1}{\pi}(m\kappa)^{\frac{2}{\alpha}}\varepsilon + O(\varepsilon^2) \tag{5.3}$$

$$c_u^{\varepsilon,\text{nsic}} = \lambda_u^{\varepsilon,\text{nsic}}(1-\varepsilon) = \frac{1}{\pi}(m\kappa)^{\frac{2}{\alpha}}\varepsilon + O(\varepsilon^2) \tag{5.4}$$

    Perfect Interference Cancelation, $(K, \zeta = 0)$:

$$c_{l,M}^{\varepsilon,\text{psic}} = \lambda_{l,M}^{\varepsilon,\text{psic}}(1-\varepsilon) = \frac{(\alpha-2)}{2\pi\bar{\beta}^{1-\frac{2}{\alpha}}}(m\kappa)^{\frac{2}{\alpha}}\varepsilon + O(\varepsilon^2), \ \ \varepsilon \leq \frac{2K}{\alpha-2}\bar{\beta} \tag{5.5}$$

$$c_u^{\varepsilon,\text{psic}} = \lambda_u^{\varepsilon,\text{psic}}(1-\varepsilon) = (1+\frac{K}{\varepsilon})\frac{1}{\pi}(m\kappa)^{\frac{2}{\alpha}}\varepsilon + O(\varepsilon^2). \tag{5.6}$$

    Partial Cancelation of Nearby Nodes (K, ζ):

$$c_{l,M}^{\varepsilon,\text{sic}} = \lambda_{l,M}^{\varepsilon,\text{sic}}(1-\varepsilon) = \sup_{(\varepsilon_1,\varepsilon_2):\varepsilon_1+\varepsilon_2=\varepsilon}\left[\frac{\alpha-2}{2}\frac{\bar{\beta}^{\frac{2}{\alpha}}}{(1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}\pi}(m\kappa)^{\frac{2}{\alpha}}\varepsilon_1, \zeta^{-\frac{2}{\alpha}}\frac{1}{\pi}(m\kappa)^{\frac{2}{\alpha}}\varepsilon_2\right] + O(\varepsilon^2),$$

$$\varepsilon < [(1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}]\frac{2K}{\alpha-2} \tag{5.7}$$

$$c_u^{\varepsilon,\text{sic}} = \lambda_u^{\varepsilon,\text{sic}}(1-\varepsilon) = \zeta^{-\frac{2}{\alpha}}\frac{1}{\pi}(m\kappa)^{\frac{2}{\alpha}}\varepsilon + O(\varepsilon^2), \ \ \varepsilon < 1 - \exp(-K\zeta^{\frac{2}{\alpha}}) \tag{5.8}$$

Proof. Similar to that of Theorem. 4.2.1. See the Appendix for details. □

    Note that for perfect SIC and imperfect SIC, we only list the regimes that are most likely to hold given our choice of parameters. In particular, relatively small outage constraints with $\varepsilon < 0.1$ and small SIC error with $\zeta < 0.1$ are considered. In addition, for these two cases, we only list the results derived from Markov inequality (subscripted by M) for

simplicity of the expressions. If we use Chebychev inequality (subscripted by C) in cases of psic and sic, as we do in the cases of conventional CDMA without SIC, we will obtain tighter bounds, which however do not have clean close-form expressions for the SIC case. See Appendix for the complete results.

Although the idealized perfect SIC case is not very realistic, it is useful for understanding the factors that may impact the performance of SIC. Comparing the perfect SIC case with conventional CDMA, we obtain the following observations.

Capacity improvement from perfect SIC. Consider the upper bound for example. Compared with no SIC, perfect SIC improves the capacity upper bound by

$$\frac{\lambda_u^{\varepsilon,\text{psic}}}{\lambda_u^{\varepsilon,\text{nsic}}} \approx \frac{K}{\varepsilon} + 1 \quad \text{when} \quad \varepsilon \ll 1, \tag{5.9}$$

which is linear in K. This implies that in the scenario where SIC is close to perfect, for the maximum performance, as many users should be cancelled as possible. On the other hand, for the lower bound based on Markov's inequality, similar to Theorem 4.2.1, we can obtain the lower bound for conventional CDMA as follows.

$$c_{l,M}^{\varepsilon,\text{nsic}} = \lambda_{l,M}^{\varepsilon,\text{nsic}}(1-\varepsilon) = \left(1 - \frac{2}{\alpha}\right)\frac{\varepsilon}{\pi}(m\kappa)^{\frac{2}{\alpha}} + O(\varepsilon^2). \tag{5.10}$$

Compared with no SIC, for $\varepsilon < \varepsilon_{c,M}^{\text{psic}}$, perfect SIC improves the Markov lower bound by a factor

$$\frac{\lambda_{l,M}^{\varepsilon,\text{psic}}}{\lambda_{l,M}^{\varepsilon,\text{nsic}}} = \frac{\alpha}{2}\left(\frac{m}{\beta}\right)^{1-\frac{2}{\alpha}}, \tag{5.11}$$

which is independent of K. This indicates that the transmission capacity is sensitive in path loss in the regime of low path loss, where the capacity observes the lower bound closely since far-field interference becomes significant. A similar conclusion can be reached based on the lower bound derived from Chebychev inequality, see Appendix.

Tightness of the bounds. As will be further demonstrated in the numerical results section, the performance improvement due to SIC under our model is relatively insensitive to K for the small ε outage regime. Recall our SIC model: up to K interfering nodes within transmission distance r_{sic} are cancelled. For small ε the transmission capacities supporting that QoS are such that it is unlikely for more than 1 interfering node to be within r_{sic} in the first place. Thus increasing K may not have any significant effect. Of course in regimes with high transmission densities having a higher K will be of great value, but this regime will not support a small ε QoS level. The bounds for perfect SIC are very loose for small ε. It is straightforward to see that

$$\lambda_u^{\varepsilon,\text{psic}} = \frac{\varepsilon + K}{\pi(\bar{\beta}^{\frac{1}{\alpha}}d)^2} + O(\varepsilon^2), \tag{5.12}$$

61

and corresponding bound ratios:

$$\frac{\lambda_{l,M}^{\varepsilon,\text{psic}}}{\lambda_{u}^{\varepsilon,\text{psic}}} = \frac{\alpha-2}{2}\bar{\beta}^{\frac{2}{\alpha}-1}\left(\frac{\varepsilon}{\varepsilon+K}\right),$$

This ratio is arbitrarily close to 0 as $\varepsilon \to 0$. The poor bound ratio is a consequence of the fact that there is no known "upper bound event" that provides a sufficient condition for outage and also results in a tight bound. The imperfect SIC model does not suffer from this problem.

Next we consider the imperfect SIC with partial cancellation error. Again we focus on the practical capacity gain obtained via imperfect SIC and the tightness of our capacity bounds.

Capacity gain from partial interference cancellation. The performance improvement due to SIC is very sensitive to the cancellation effectiveness parameter ζ as $\zeta \to 0$, especially for small ε. Looking at the upper bound, for example, we see that for small ε:

$$\frac{d}{d\zeta}\lambda_{u}^{\varepsilon,\text{sic}} \propto -\zeta^{-(1+\frac{2}{\alpha})}, \tag{5.13}$$

which means

$$\lim_{\zeta\to 0}\frac{d}{d\zeta}\lambda_{u}^{\varepsilon,\text{sic}} = \infty. \tag{5.14}$$

Thus our model suggests that technology improvements which improve cancellation effectiveness may yield large increases in the transmission capacity.

Tightness of the bounds. For small ε it is straightforward to show that the bounds are reasonably tight. In particular,

$$\lambda_{l,M}^{\varepsilon,\text{sic}} = \frac{(\alpha-2)\bar{\beta}^{\frac{2}{\alpha}}}{2(1-\zeta)\bar{\beta}+(2\zeta+\alpha-2)\bar{\beta}^{\frac{2}{\alpha}}}\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}+O(\varepsilon^2),$$

with corresponding bound ratios of

$$\frac{\lambda_{l,M}^{\varepsilon,\text{sic}}}{\lambda_{u}^{\varepsilon,\text{sic}}} = \frac{(\alpha-2)\left(\bar{\beta}\zeta\right)^{\frac{2}{\alpha}}}{2(1-\zeta)\bar{\beta}+(2\zeta+\alpha-2)\bar{\beta}^{\frac{2}{\alpha}}}+O(\varepsilon^2).$$

Note that these bound ratios are 0 for $\zeta=0$, consistent with the poor bound ratios for perfect SIC, but $(\alpha-2)/\alpha$ and $(\alpha-1)/\alpha$ respectively for $\zeta=1$, consistent with the ratios obtained for no SIC.

5.2.3 Performance evaluation.

In order to see how the transmission capacity varies across the different techniques (FH, DS, DS with SIC) and also as parameters vary, four plots are presented, Fig. 5.3-5.6. For all cases, we use the Markov lower bounds for comparison. Note that the accuracy of the Markov lower bounds is poor when the second order variability of far field interferers becomes significant in contribution to the outage probability, e.g., when the spreading factor m is small or path loss exponent α is small.

Unless otherwise noted, in these plots the received signal-to-thermal-noise ratio is $SNR = 20$dB, the communication distance is $d = 10$m, the path loss exponent is $\alpha = 4$, the fractional cancelation error is $\zeta = 0.1$, the outage probability is $\varepsilon = 0.01$, and the number of cancelable users is $K = 5$.

Our key results can be summarized as:

- Overall transmission capacity decreases as the spreading factor m increases (since the available bandwidth decreases by m) for all techniques except frequency hopping.

- Perfect SIC increases capacity by approximately an order of magnitude.

- It appears as though much of the capacity of perfect SIC may be attainable, because it is only necessary to cancel a few strong nodes to get a large portion of the capacity gain from SIC.

- Capacity decreases in severe propagation environments (large α), except for FH.

It should be mentioned that although the capacity of DS-CDMA ad hoc networks decay as the spreading increases, this does not mean that the best solution is to avoid spreading. Indeed, straightforward capacity analysis of DS-CDMA cellular systems reached similar conclusions, with implementation details such as frequency reuse and voice activity eventually giving CDMA an advantage in cellular [14]. When considering delay constraints, robustness, security (anti-jamming and low probability of detection/intercept), and the need for strong error-correction codes, we suspect that spread spectrum systems will prove very attractive for wireless ad hoc networks.

As shown in Fig. 5.3, since all DS-CDMA systems lose capacity as the spreading factor increases, the spreading factor should be chosen to be as small as possible while still allowing interference averaging. With SIC, such loss of spectrum efficiency when increasing m is significantly mitigated in a DS-CDMA system.

As shown in Fig. 5.4, direct sequence capacity reduces as the path loss becomes worse because much higher transmit power levels must be used, which further cripples nearby nodes. In contrast, for a FH-CDMA system, nodes only have a $1/m$ chance of colliding with nearby nodes, and meanwhile, the farther nodes now cause less interference
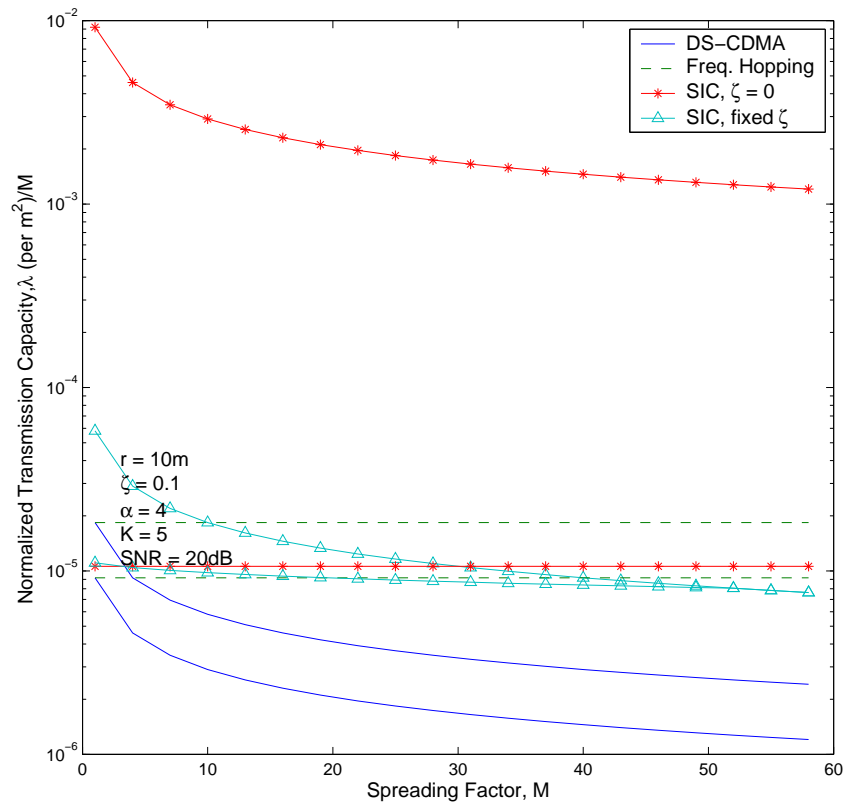
Figure 5.3: Normalized transmission capacity vs. spreading factor.
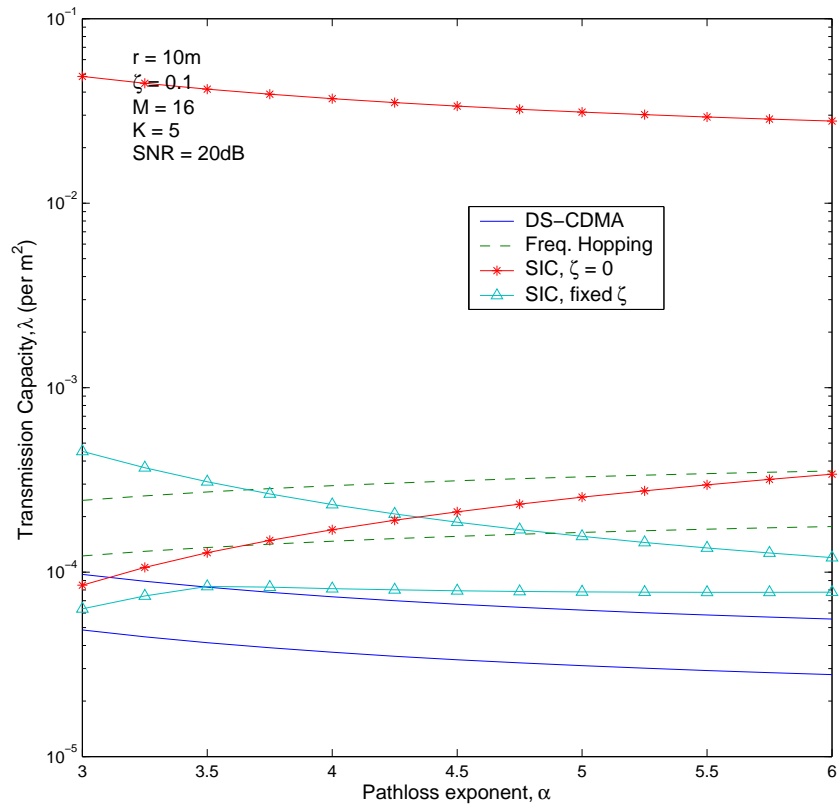
Figure 5.4: Transmission capacity vs. path loss exponent.

than before, which explains why the capacity of a FH-CDMA system improves in α. Again, SIC in DS-CDMA helps to tolerate the negative impact on capacity due to the path loss. Note that with SIC, the upper bound and the lower bound behave differently in the path loss exponent α. One can expect the capacity observes the lower bound closely in the low path loss regime, in which far field interference is significant and captured by the lower bound. The capacity observes the upper bound closely in the high path loss regime, in which far field interference is negligible and thus only considering nearest interferers, as the derivation for the upper bound does, is quite accurate.
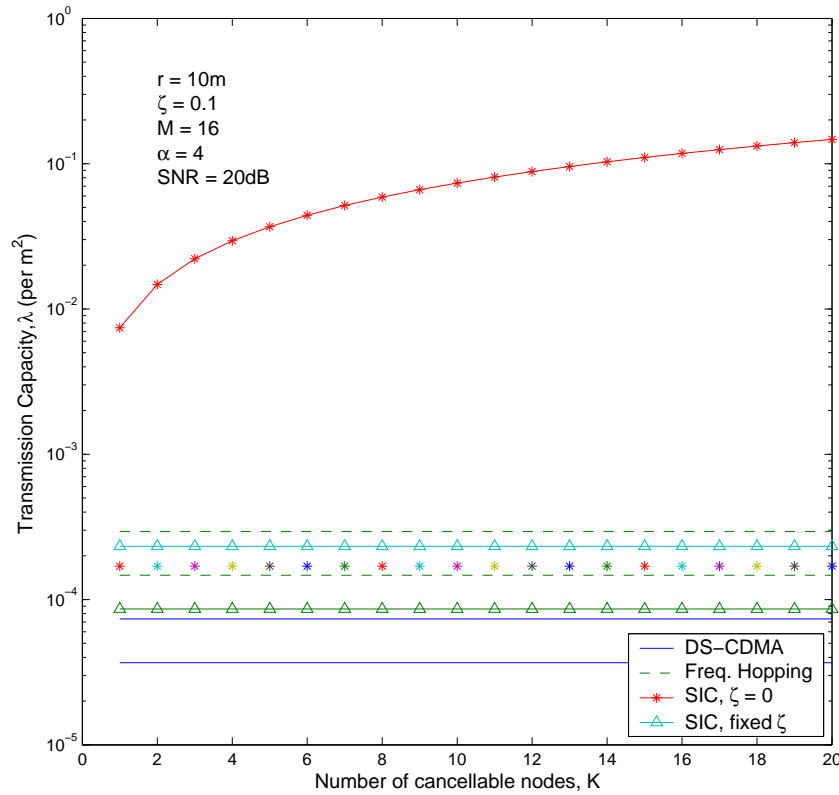
Figure 5.5: Transmission capacity vs. number of cancelable users.

As shown in Fig. 5.5, it is obvious that only a few interferers need to be canceled to get the most of the gain from SIC. This is because these are the closest interferers who cause most of the interference. This outcome has favorable implications for receiver complexity and latency. This observation remains true unless the SIC error is extremely small, for which canceling more interferers will have significant capacity improvement.

As shown in Fig. 5.6, the capacity degrades gradually as interference cancelation error increases and converges to conventional CDMA case as $\zeta \to 1$. When ζ is very small,
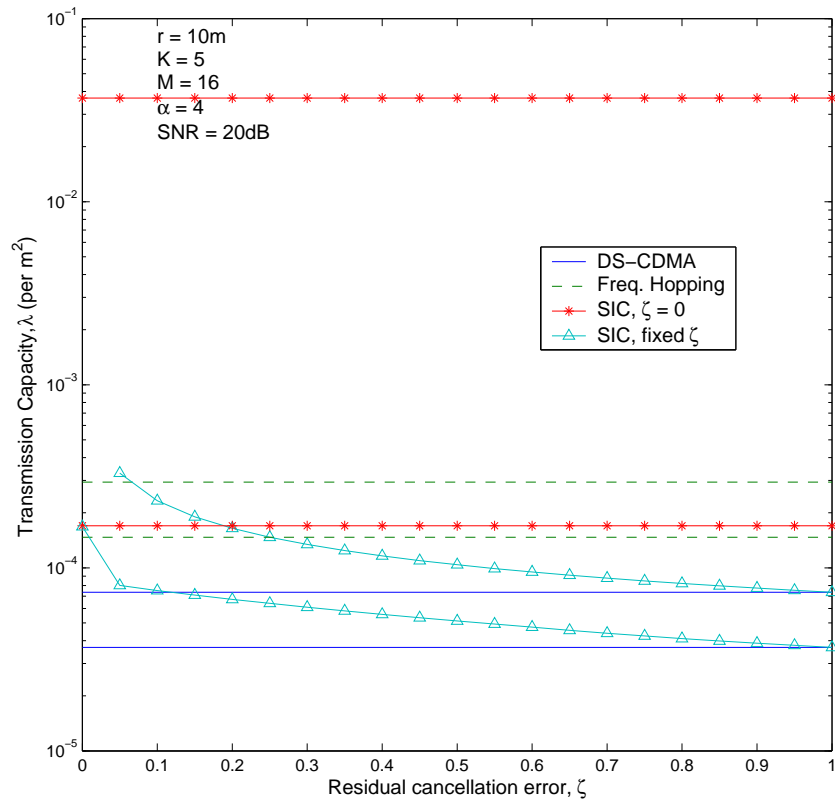
Figure 5.6: Transmission capacity vs. residual cancelation error.

DS-CDMA with SIC can outperform FH-CDMA.

5.3 Summary of SIC in CDMA ad hoc networks

SIC significantly improves the capacity of DS-CDMA systems, particularly under a simple random channel access MAC protocol. The major reason is given such a MAC protocol, a few nearby interferers contribute most of the interference, which is well suited for canceling via SIC. Even though SIC potentially causes delay due to iterative signal processing, a receiver need only cancel a few strong interferers and thus the delay factor will be insignificant when compared with the much reduced end-to-end delay, thanks to the higher capacity achieved and enabling of long range transmission. In addition, technology advances enable low power computation and signal processing, which warrant the energy cost for SIC to significantly reduce the energy waste due to outage and retransmissions. Therefore, by using SIC, we introduce extra system design complexity, but improve all three major performance metrics: capacity, delay and energy efficiency.

5.4 Appendix: The Complete Results and Proofs of Theorem 5.2.1

5.4.1 The Complete Results of Theorem 5.2.1 - Perfect SIC Case

For the convenience of the reader, we first summarize here most of the notation used in the proof of Theorem 5.2.1.

$a \vee b$	$\max\{a,b\}$
$a \wedge b$	$\min\{a,b\}$
$x \in A \setminus B$	$x \in A, x \notin B$
$b(O,r)$	$\{x : |x| \leq r\}$, i.e. a ball of radius r centered at origin
$a(O,r_1,r_2)$	$\{x : r_1 \leq |x| \leq r_2\}$, i.e. an annulus between r_1 and r_2

ρ	transmit power
d	transmit distance
β	target (required) SINR after despreading
$\bar{\beta}$	$\bar{\beta} \doteq \frac{\beta}{m}$
ε	required outage probability, i.e. $P^0[SINR \leq \beta] \leq \varepsilon$
α	path loss exponent
y	$y \doteq \frac{d^{-\alpha}}{\beta}$
r_s	$r_s \doteq y^{-\frac{1}{\alpha}} = \beta^{\frac{1}{\alpha}} d$ is the "splitting radius"
K	maximum *and* expected number of cancelled users
r_{sic}	radius around receiver that includes, on average, K interferers
$\zeta \in (0,1)$	residual interference after cancellation
psic	specifies perfect SIC, i.e $\zeta \to 1$
nsic	specifies no SIC, i.e $\zeta \to 0$

$\Pi(\lambda)$	A Poisson Point Process Π with density λ
λ_u	Upper bound on λ, i.e. $\lambda \geq \lambda_u \Rightarrow P^0[SINR \leq \beta] \geq \varepsilon$
λ_l	Lower bound on λ, i.e. $\lambda \leq \lambda_l \Rightarrow P^0[SINR \leq \beta] \leq \varepsilon$
λ_C	A lower bound attained with Chebychev inequality
λ_M	A lower bound attained with Markov inequality
λ^{ε}	specifies that the resulting density is for some ε
λ_r	λ such that K users are $\in b(0,r_{\text{tx}}$ on average.
λ_s	λ such that K users are $\in b(0,y^{-\frac{1}{\alpha}}) = b(0,r_s)$

The major result is a set of closed form expressions for lower and upper bounds on the transmission capacity in the perfect SIC case.

Theorem 5.4.1. As $\varepsilon \in (0,1) \to 0$, the lower and upper bounds on the transmission capacity when receivers are equipped with perfect SIC ($\zeta = 0$) are:

$$c_l^{\varepsilon,\text{psic}} = (1-\varepsilon)\lambda_l^{\varepsilon,\text{psic}}, \qquad c_u^{\varepsilon,\text{psic}} = (1-\varepsilon)\lambda_u^{\varepsilon,\text{psic}}. \tag{5.15}$$

The upper bound on the optimal contention density is:

$$\lambda_u^{\varepsilon,\text{psic}} = \frac{-\ln(1-\varepsilon) + K}{\pi r_s^2}. \tag{5.16}$$

The Markov (M) lower bound on the optimal contention density is:

$$\lambda_{l,M}^{\varepsilon,\text{psic}} = \begin{cases} \left(\frac{(\alpha-2)\varepsilon}{2}\right)\frac{\bar\beta^{\frac{2}{\alpha}-1}}{\pi r_s^2} & \varepsilon \le \frac{2K}{\alpha-2}\bar\beta \\[2mm] \left(\frac{(\alpha-2)\varepsilon}{2}\right)^{\frac{2}{\alpha}}\frac{K^{1-\frac{2}{\alpha}}}{\pi r_s^2} & \text{else} \\[2mm] \left(1-\frac{2}{\alpha}\right)\frac{\varepsilon+K}{\pi r_s^2} & \varepsilon \ge \frac{2K}{\alpha-2} \end{cases} \tag{5.17}$$

The Chebychev (C) lower bound on the optimal contention density is:

$$\lambda_{l,C}^{\varepsilon,\text{psic}} \ge \sup_{(\varepsilon_u,\varepsilon_f):\varepsilon_u+\varepsilon_f=\varepsilon}\left\{\lambda_u^{\varepsilon_u,\text{psic}} \wedge \lambda_{f,C}^{\varepsilon_f,\text{psic}}\right\}, \tag{5.18}$$

where

$$\lambda_{f,C}^{\varepsilon,\text{psic}} = \begin{cases} \frac{\alpha-2}{2\bar\beta}\frac{1}{\pi}d^{-2} + \frac{(\alpha-2)^2}{8(\alpha-1)}\frac{1}{\pi}d^{-2}\frac{1}{\varepsilon}\left(1-\sqrt{1+\frac{8(\alpha-1)}{(\alpha-2)\bar\beta}\varepsilon}\right) \\[3mm] \dfrac{y^{\frac{2}{\bar\alpha}}}{\left(\frac{2}{\alpha-2}K^{1-\frac{\alpha}{2}}\pi^{\frac{\alpha}{2}} + K^{\frac{1-\alpha}{2}}\frac{\pi^{\frac{\alpha}{2}}}{\sqrt{\alpha-1}\sqrt{\varepsilon}}\right)^{\frac{2}{\bar\alpha}}} \\[3mm] \frac{\alpha-2}{2}\frac{1}{\pi}y^{\frac{2}{\alpha}} + \frac{(\alpha-2)^2}{8(\alpha-1)}\frac{1}{\pi}y^{\frac{2}{\alpha}}\frac{1}{\varepsilon}\left(1-\sqrt{1+\frac{8(\alpha-1)}{\alpha-2}\varepsilon}\right) \end{cases} \tag{5.19}$$

where the three expressions (a), (b), (c) hold for the regimes

$$(a) \qquad \left(\lambda_r < \lambda_M^{\text{psic}} \text{ and } \varepsilon \le \varepsilon_{c,C}^{\text{psic}}\right) \text{ or } \lambda_r > \lambda_M^{\text{psic}}$$

$$(b) \qquad \left(\lambda_s < \lambda_M^{\text{psic}} \text{ and } \varepsilon_{c,C}^{\text{psic}} < \varepsilon \le \varepsilon_{k,C}^{\text{psic}}\right)$$

$$\text{or } \left(\lambda_r \le \lambda_M^{\text{psic}} < \lambda_s \text{ and } \varepsilon > \varepsilon_{c,C}^{\text{psic}}\right)$$

$$(c) \qquad \lambda_s < \lambda_M^{\text{psic}} \text{ and } \varepsilon > \varepsilon_{k,C}^{\text{psic}}$$

respectively. The constants are given by:

$$\lambda_M^{\text{psic}} = \begin{cases} \frac{\alpha-2}{2\pi}d^{-2}\bar\beta^{-1}, & \alpha-2-2\bar\beta K \le 0 \\[2mm] \frac{1}{\pi}\left(\frac{\alpha-2}{2}\right)^{\frac{2}{\alpha}}K^{1-\frac{2}{\alpha}}y^{\frac{2}{\alpha}}, & \text{else} \\[2mm] \frac{\alpha-2}{2\pi}d^{-2}\bar\beta^{-\frac{2}{\alpha}}, & \alpha-2-2\bar\beta Kr^\alpha > 0 \end{cases} \tag{5.20}$$

and

$$\lambda_r = \frac{K}{\pi}d^{-2}, \quad \lambda_s = \frac{K}{\pi}y^{\frac{2}{\alpha}} \tag{5.21}$$

and

$$\varepsilon_{c,C}^{\text{psic}} = \frac{K(\alpha-2)^2\bar{\beta}^2}{(\alpha-1)(\alpha-2-2\bar{\beta}K)^2}$$

$$\varepsilon_{k,C}^{\text{psic}} = \frac{K(\alpha-2)^2}{(\alpha-1)(\alpha-2-2K)^2}.$$

Proof. The idea behind the proof is to identify necessary and sufficient conditions for outage, calculate or bound the probabilities of the corresponding events, and then determine the spatial densities such that the probabilities of the necessary and sufficient events equal the specified QoS parameter ε. The sufficient condition event we employ is the set of realizations of the point process Π with one or more interfering nodes close enough to the receiver so that one such node alone is capable of causing outage. The necessary condition is more complex: if we have an outage it can be due to either a few nodes near the receiver or the combination of a large number of far away nodes. Let

$$F_u^{\text{psic}}(\lambda), \quad F^{\text{psic}}(\lambda), \quad F_l^{\text{psic}}(\lambda) \tag{5.22}$$

be events parameterized by the spatial density λ so that

$$F_u^{\text{psic}}(\lambda) \subset F^{\text{psic}}(\lambda) \subset F_l^{\text{psic}}(\lambda). \tag{5.23}$$

The probability of all three events will be nondecreasing in λ. The event $F_u^{\text{psic}}(\lambda)$ is the sufficient event, $F^{\text{psic}}(\lambda)$ is the outage event, and $F_l^{\text{psic}}(\lambda)$ is the necessary event. As illustrated in Figure 5.7, lower and upper bounds on the optimal contention density are obtainable from the probabilities of the necessary and sufficient events, provided we can solve

$$\lambda_u^{\varepsilon,\text{psic}} = \{\lambda : \mathbb{P}^0(F_u^{\text{psic}}(\lambda)) = \varepsilon\},$$

$$\lambda_l^{\varepsilon,\text{psic}} = \{\lambda : \mathbb{P}^0(F_l^{\text{psic}}(\lambda)) = \varepsilon\}$$

for λ. Since these equations are in general not solvable for λ, we define several different events that will help us attain bounds on λ.

Definition 5.4.1.

$$F^{\text{psic}}(\lambda) = \left\{Y(\lambda) > y\right\},$$

$$F_u^{\text{psic}}(\lambda) = \left\{\Pi \cap a(O, r_{\text{sic}}, r_s) \neq \emptyset\right\}$$

$$F_f^{\text{psic}}(\lambda) = \left\{Y(\lambda, r_s) > y\right\}$$

71
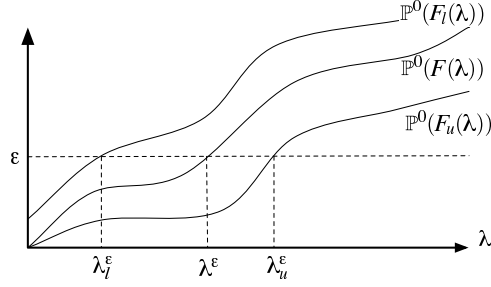
Figure 5.7: Illustration of the technique to find lower and upper bounds on the contention density: $\lambda_l^\varepsilon \leq \lambda^\varepsilon \leq \lambda_u^\varepsilon$ through the use of necessary and sufficient events for outage.

where $Y(\lambda)$ is given by Definition 5.2.2 with $\zeta = 0$ and

$$Y(\lambda, r_s) = \sum_{i \in \Pi \cap \bar{b}(O, r_{\text{sic}} \vee r_s)} |X_i|^{-\alpha} \tag{5.24}$$

is the normalized aggregate interference by all nodes outside of the radius $r_{\text{sic}} \vee r_s$.

In words, $F^{\text{psic}}(\lambda)$ is the outage event, $F_u^{\text{psic}}(\lambda)$ is the event that one or more nodes lie in the annulus with radii r_{sic} and r_s, and $F_f^{\text{psic}}(\lambda)$ is the event that the aggregate interference generated by nodes outside the radius $r_{\text{sic}} \vee r_s$ is sufficient to cause an outage. It is straightforward to establish that

$$F_u^{\text{psic}}(\lambda) \subset F^{\text{psic}}(\lambda) \subset F_l^{\text{psic}}(\lambda) = F_u^{\text{psic}}(\lambda) \cup F_f^{\text{psic}}(\lambda). \tag{5.25}$$

It is helpful to think of $r_s \doteq y^{-\frac{1}{\alpha}}$ as the radius splitting the "near-field" interference, $b(O, r_s)$, from the "far-field" interference, $\mathbb{R}^2 \setminus b(O, r_s)$. A similar approach is employed in the proof of Theorem 4.2.1 for the proof of the transmission capacity without SIC, but with the additional degree of freedom that the near/far field boundary was optimized over all s. It is shown that $s = r_s$ is the optimal splitting radius. A similar optimization could be performed here but the analysis becomes much more complex and the tractability of the model is lost. For that reason we use a fixed near/far field splitting radius r_s throughout this paper, where r_s is the maximum radius such that a single node at that distance from the receiver can by itself cause an outage at the receiver. Clearly

$$\mathbb{P}^0(F_u^{\text{psic}}(\lambda)) \leq \mathbb{P}^0(F^{\text{psic}}(\lambda)) \leq \mathbb{P}^0(F_l^{\text{psic}}(\lambda)), \tag{5.26}$$

and

$$\mathbb{P}^0(F_l^{\text{psic}}(\lambda)) \leq \mathbb{P}^0(F_u^{\text{psic}}(\lambda)) + \mathbb{P}^0(F_f^{\text{psic}}(\lambda)). \tag{5.27}$$

Define two spatial transmission density thresholds:

$$\lambda_r = \frac{K}{\pi d^2}, \quad \lambda_s = \frac{K}{\pi r_s^2}. \tag{5.28}$$

72

These correspond to the densities where there are on average K users inside of a radius d and r_s, respectively.

It is straightforward to see then that

$$r_{\text{sic}} = \begin{cases} d, & \lambda \leq \lambda_r \\ \sqrt{\frac{K}{\pi\lambda}}, & \text{else} \end{cases} \tag{5.29}$$

and that

$$r_{\text{sic}} \begin{cases} \geq r_s, & \lambda \leq \lambda_s \\ \leq r_s, & \text{else} \end{cases}. \tag{5.30}$$

Upper Bound. We begin by finding the upper bound; this requires solving $\mathbb{P}^0(F_u^{\text{psic}}) = \varepsilon$ for λ. The probability of one or more nodes lying in the annulus with radii r_{sic} and r_s is simply one minus the void probability for the set:

$$\mathbb{P}^0(F_u^{\text{psic}}(\lambda)) = \begin{cases} 1 - \exp\{-\lambda\pi(r_s^2 - r_{\text{sic}}^2)\}, & \lambda > \lambda_s \\ 0, & \text{else} \end{cases}. \tag{5.31}$$

It is evident here why the upper bound is so weak for the perfect SIC case: the upper bound event is zero for all $\lambda \leq \lambda_s$. Unfortunately there is no other easily computable sufficient event available. Note that the map $\lambda \to \mathbb{P}^0(F_u^{\text{psic}}(\lambda))$ is onto $[0,1)$ and monotone increasing in λ; hence a unique inverse exists for all $\varepsilon > 0$. Setting this expression equal to ε and solving for λ yields:

$$\lambda_u^{\varepsilon,\text{psic}} = \frac{1}{\pi r_s^2}\left(-\ln(1-\varepsilon) + K\right) \geq \left(1 + \frac{K}{\varepsilon}\right)\frac{1}{\pi r_s^2}\varepsilon + O(\varepsilon^2). \tag{5.32}$$

∎

Lower Bounds. We turn now to the lower bound. The lower bound event $F_l^{\text{psic}}(\lambda, s)$ is the union of two events, $F_u^{\text{psic}}(\lambda)$ and $F_f^{\text{psic}}(\lambda)$; the probability of both events is increasing in λ. Moreover, a consequence of the assumption that the node positions form a Poisson process is that the two events are independent seeing as they concern disjoint regions of $\mathbb{R}^2$. Fix ε and consider some pair $(\varepsilon_u, \varepsilon_f)$ such that $\varepsilon_u + \varepsilon_f = \varepsilon$. If we can identify a pair $(\lambda_u^{\varepsilon_u,\text{psic}}, \lambda_f^{\varepsilon_f,\text{psic}})$ satisfying:

$$\mathbb{P}^0(F_u(\lambda_u^{\varepsilon_u,\text{psic}})) \leq \varepsilon_u, \ \mathbb{P}^0(F_f(\lambda_f^{\varepsilon_f,\text{psic}})) \leq \varepsilon_f \tag{5.33}$$

then

$$\mathbb{P}^0(F_l(\lambda_u^{\varepsilon_u,\text{psic}} \wedge \lambda_f^{\varepsilon_f,\text{psic}})) \leq \varepsilon_u + \varepsilon_f = \varepsilon. \tag{5.34}$$

Thus $\lambda_u^{\varepsilon_u,\text{psic}} \wedge \lambda_f^{\varepsilon_f,\text{psic}}$ is a valid lower bound since choosing $\lambda < \lambda_u^{\varepsilon_u,\text{psic}} \wedge \lambda_f^{\varepsilon_f,\text{psic}}$ ensures the outage probability is less than ε. This argument holds for all partitions $(\varepsilon_u, \varepsilon_f)$ summing

73

to ε. The greatest lower bound is obtained by maximizing $\lambda_u^{\varepsilon_u,\text{psic}} \wedge \lambda_f^{\varepsilon_f,\text{psic}}$ over all feasible partitions:

$$\lambda_l^{\varepsilon,\text{psic}} = \sup_{(\varepsilon_u,\varepsilon_f)\,:\,\varepsilon_u+\varepsilon_f=\varepsilon} \left\{ \lambda_u^{\varepsilon_u,\text{psic}} \wedge \lambda_f^{\varepsilon_f,\text{psic}} \right\}. \tag{5.35}$$

Note that the minimum of two functions is maximized by minimizing the distance between them. If that minimum distance is zero then the minimum of the two functions is their value at the point of intersection.

The probability $\mathbb{P}^0(F_f^{\text{psic}}(\lambda)) = \mathbb{P}^0(Y(\lambda,r_s) > y)$ cannot be computed exactly; it must be bounded. We'll obtain two bounds via the Markov and Chebychev inequalities. The former is weaker but the resulting equations are simpler, the latter is stronger but the equations are more complex. To compute the Markov bound we need $\mathbb{E}^0[Y(\lambda,r_s)]$; to compute the Chebychev bound we need $\mathbb{E}^0[Y(\lambda,r_s)]$ and $Var(Y(\lambda,r_s))$. Both are obtainable via Campbell's Theorem, see the proof of Theorem 4.2.1, which gives that $E[\sum_{x\in\Pi} f(x)] = \int_{\mathbb{R}^2} \lambda f(x)\nu_d(dx)$ where $\nu_d(\cdot)$ is the Lebesgue measure (area). We begin with the Markov bound:

$$\mathbb{P}^0(F_f^{\text{psic}}(\lambda)) \leq \frac{\mathbb{E}^0[Y(\lambda,r_s)]}{y}. \tag{5.36}$$

It is straightforward to compute:

$$\frac{\mathbb{E}^0[Y(\lambda,r_s)]}{y} = \begin{cases} \frac{2\pi}{\alpha-2}\bar{\beta}d^2\lambda, & \lambda \leq \lambda_c \\ \frac{2}{\alpha-2}K^{1-\frac{\alpha}{2}}\frac{1}{y}(\pi\lambda)^{\frac{\alpha}{2}}, & \lambda_c < \lambda \leq \lambda_k \\ \frac{2\pi}{\alpha-2}r_s^2\lambda, & \lambda > \lambda_k \end{cases}. \tag{5.37}$$

The value of the bound at the critical points is

$$\begin{aligned} \varepsilon_{c,M}^{\text{psic}} &= \frac{\mathbb{E}^0[Y(\lambda_c,r_s)]}{y} = \bar{\beta}\frac{2K}{\alpha-2} \\ \varepsilon_{k,M}^{\text{psic}} &= \frac{\mathbb{E}^0[Y(\lambda_k,r_s)]}{y} = \frac{2K}{\alpha-2}. \end{aligned}$$

This function is monotone increasing in λ and is onto $[0,1]$; hence the inverse function exists. Setting the expression equal to ε and solving for λ gives:

$$\lambda_{f,M}^{\varepsilon,\text{psic}} = \begin{cases} \left(\frac{\alpha-2}{2}\right)\bar{\beta}^{\frac{2}{\alpha}-1}\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}, & \varepsilon \leq \varepsilon_{c,M}^{\text{psic}} \\ \left(\frac{\alpha-2}{2}\right)^{\frac{2}{\alpha}}\left(\frac{K}{\varepsilon}\right)^{1-\frac{2}{\alpha}}\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}, & \varepsilon_{c,M}^{\text{psic}} < \varepsilon \leq \varepsilon_{k,M}^{\text{psic}} \\ \left(\frac{\alpha-2}{2}\right)\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}, & \varepsilon \geq \varepsilon_{k,M}^{\text{psic}} \end{cases} \tag{5.38}$$

It remains to maximize $\lambda_{f,M}^{\varepsilon_f,\text{psic}} \wedge \lambda_u^{\varepsilon_u,\text{psic}}$ over all $(\varepsilon_u,\varepsilon_f)$ such that $\varepsilon_u + \varepsilon_f = \varepsilon$. Note that $\lambda_u^{\varepsilon_u,\text{psic}} \geq \lambda_s$ for all ε while $\lambda_{f,M}^{\varepsilon_f,\text{psic}} \leq \lambda_s$ for all $\varepsilon \leq \varepsilon_{k,M}^{\text{psic}}$. Thus the optimum splitting pair is

$\varepsilon_f = \varepsilon$ and $\varepsilon_u = 0$, and the corresponding minimum of the two functions is $\lambda_{f,M}^{\varepsilon,\text{psic}}$. Note that $\varepsilon_{k,M}^{\text{psic}} \geq 1$ for all $K \geq 2$ and all $\alpha \leq 4$. We now find the optimal splitting pair when $\varepsilon > \varepsilon_{k,M}^{\text{psic}}$. Note that $\lambda_u^{\varepsilon_u,\text{psic}}$ is non-linear in ε and hence finding the point of intersection with $\lambda_{f,M}^{\varepsilon_f,\text{psic}}$ is complicated. We find a lower bound on $\lambda_u^{\varepsilon_u,\text{psic}}$ by linearizing around $\varepsilon = 0$; this leaves us with the problem of maximizing the minimum of

$$\lambda_u^{\varepsilon_u,\text{psic}} \geq \frac{K}{\pi(\bar{\beta}^{\frac{1}{\alpha}}d)^2} + \frac{\varepsilon_u}{\pi(\bar{\beta}^{\frac{1}{\alpha}}d)^2}$$

$$\lambda_{f,M}^{\varepsilon_f,\text{psic}} = \left(\frac{\alpha-2}{2}\right)\frac{\varepsilon_f}{\pi(\bar{\beta}^{\frac{1}{\alpha}}d)^2}, \ \varepsilon \geq \varepsilon_{k,M}^{\text{psic}}.$$

It is straightforward to establish the optimal splitting pair for the linearized $\lambda_u^{\varepsilon_u,\text{psic}}$ is:

$$\varepsilon_u = \left(\left(1 - \tfrac{2}{\alpha}\right)\varepsilon - \tfrac{2}{\alpha}K\right) \vee 0, \quad \varepsilon_f = \left(\tfrac{2}{\alpha}\varepsilon + \tfrac{2}{\alpha}K\right) \wedge 1, \tag{5.39}$$

where the two functions share a common value at this point of

$$\lambda_{l,M}^{\varepsilon,\text{psic}} = \left(1 - \frac{2}{\alpha}\right)\frac{\varepsilon + K}{\pi(\bar{\beta}^{\frac{1}{\alpha}}d)^2}, \ \varepsilon > \varepsilon_{k,M}^{\text{psic}}. \tag{5.40}$$

■

We now consider the Chebychev lower bound. The variance of the far-field interference is again found by Campbell's Theorem:

$$Var(Y(\lambda, r_s)) = \begin{cases} \frac{\pi}{\alpha-1}d^{2(1-\alpha)}\lambda, & \lambda \leq \lambda_r \\ \frac{1}{\alpha-1}K^{1-\alpha}(\pi\lambda)^{\alpha}, & \lambda_r < \lambda \leq \lambda_s \\ \frac{\pi}{\alpha-1}y^{2(1-\frac{1}{\alpha})}\lambda, & \lambda > \lambda_s \end{cases} \tag{5.41}$$

Chebychev's inequality yields: for $y > \mathbb{E}^0[Y(\lambda, r_s)]$:

$$\mathbb{P}^0(Y(\lambda, r_s) > y) \leq \frac{Var(Y(\lambda, r_s))}{(y - \mathbb{E}^0[Y(\lambda, r_s)])^2}. \tag{5.42}$$

Substituting the expressions for the mean and variance:

$$\frac{Var(Y(\lambda, r_s))}{(y - \mathbb{E}^0[Y(\lambda, r_s)])^2} = \begin{cases} \dfrac{\frac{\pi}{\alpha-1}d^{2(1-\alpha)}\lambda}{\left(y - \frac{2\pi}{\alpha-2}d^{2-\alpha}\lambda\right)^2} \\[2em] \dfrac{\frac{1}{\alpha-1}K^{1-\alpha}(\pi\lambda)^{\alpha}}{\left(y - \frac{2}{\alpha-2}K^{1-\frac{\alpha}{2}}(\pi\lambda)^{\frac{\alpha}{2}}\right)^2} \\[2em] \dfrac{\frac{\pi}{\alpha-1}y^{2(1-\frac{1}{\alpha})}\lambda}{\left(y - \frac{2\pi}{\alpha-2}y^{1-\frac{2}{\alpha}}\lambda\right)^2} \end{cases} \tag{5.43}$$

where the three expressions above hold for

$$\lambda \leq \lambda_r, \quad \lambda_r < \lambda \leq \lambda_s, \quad \lambda > \lambda_s. \tag{5.44}$$

75

The value of the above bound at the critical points λ_r, λ_s is

$$
\begin{aligned}
\varepsilon_{c,C}^{\text{psic}} &= \frac{K\bar{\beta}^2(\alpha-2)^2}{(\alpha-1)(\alpha-2-2\bar{\beta}K)^2} \\
\varepsilon_{k,C}^{\text{psic}} &= \frac{K(\alpha-2)^2}{(\alpha-1)(\alpha-2-2K)^2}.
\end{aligned}
$$

Note that the first and third expressions in the Chebychev bound have the form

$$
\frac{\sigma^2\lambda}{(y-\mu\lambda)^2}, \tag{5.45}
$$

for constants μ and λ independent of ε. Setting this equation equal to ε and solving for λ yields, for $y > \mu\lambda$:

$$
\frac{y}{\mu} + \frac{\sigma^2}{2\mu^2\varepsilon}\left(1 - \sqrt{1 + \frac{4\mu y}{\sigma^2}\varepsilon}\right) = \frac{y^2}{\sigma^2}\varepsilon + O(\varepsilon^2). \tag{5.46}
$$

Setting the three expressions in the Chebychev bound equal to ε and solving for λ gives:

$$
\lambda_{f,C}^{\varepsilon,\text{psic}} = \begin{cases}
\frac{\alpha-2}{2\bar{\beta}}\frac{1}{\pi}d^{-2} + \frac{(\alpha-2)^2}{8(\alpha-1)}\frac{1}{\pi}d^{-2}\frac{1}{\varepsilon}\left(1 - \sqrt{1 + \frac{8(\alpha-1)}{(\alpha-2)\bar{\beta}}\varepsilon}\right) \\[2mm]
\dfrac{y^{\frac{2}{\alpha}}}{\left(\frac{2}{\alpha-2}K^{1-\frac{\alpha}{2}}\pi^{\frac{\alpha}{2}} + \frac{K^{\frac{1-\alpha}{2}}\pi^{\frac{\alpha}{2}}}{\sqrt{\alpha-1}\sqrt{\varepsilon}}\right)^{\frac{2}{\alpha}}} \\[3mm]
\frac{\alpha-2}{2}\frac{1}{\pi}y^{\frac{2}{\alpha}} + \frac{(\alpha-2)^2}{8(\alpha-1)}\frac{1}{\pi}y^{\frac{2}{\alpha}}\frac{1}{\varepsilon}\left(1 - \sqrt{1 + \frac{8(\alpha-1)}{\alpha-2}\varepsilon}\right)
\end{cases} \tag{5.47}
$$

The condition that $y > \mathbb{E}^0[Y(\lambda, r_s)]$ can be expressed as $\lambda < \lambda_M^{\text{psic}}$ where λ_M^{psic} is the unique density such that $\mathbb{E}^0[Y(\lambda_M^{\text{psic}}, r_s)] = y$. Straightforward algebra yields:

$$
\lambda_M^{\text{psic}} = \begin{cases}
\frac{\alpha-2}{2\pi}d^{-2}\bar{\beta}^{-1}, & \alpha-2-2\bar{\beta}K \leq 0 \\
\frac{1}{\pi}\left(\frac{\alpha-2}{2}\right)^{\frac{2}{\alpha}}K^{1-\frac{2}{\alpha}}y^{\frac{2}{\alpha}}, & \text{else} \\
\frac{\alpha-2}{2\pi}d^{-2}\bar{\beta}^{-\frac{2}{\alpha}}, & \alpha-2-2\bar{\beta}Kr^\alpha > 0
\end{cases} \tag{5.48}
$$

Note that the Chebychev bound is monotone increasing for $\lambda < \lambda_M^{\text{psic}}$, monotone decreasing for $\lambda > \lambda_M^{\text{psic}}$ and has a singularity at λ_M^{psic}. Inverting the bound requires a careful analysis of when each of the three conditions occur:

$$
\lambda_r \leq \lambda_s \leq \lambda_M^{\text{psic}}, \quad \lambda_r \leq \lambda_M^{\text{psic}} \leq \lambda_s, \quad \lambda_M^{\text{psic}} \leq \lambda_r \leq \lambda_s. \tag{5.49}
$$

Looking at Figure 5.8, it is apparent that the appropriate expression for the inverse depends on both ε's position relative to $\varepsilon_{c,C}^{\text{psic}}$ and $\varepsilon_{k,C}^{\text{psic}}$ as well as λ_M^{psic}'s position relative to λ_r and λ_s.
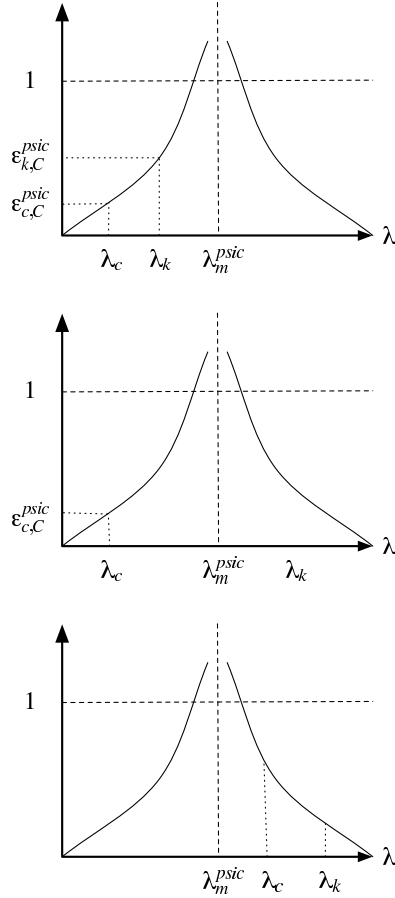
Figure 5.8: Three possibilities for the three transmission densities: $\lambda_r, \lambda_s, \lambda_M^{\text{psic}}$. The Chebychev bound is a convex increasing function for $\lambda < \lambda_M^{\text{psic}}$; inversion of the function requires a careful analysis of the cases when each of the three inverse expressions for the bound are appropriate.

In particular, the three expressions above hold for

$$(a) \qquad \left(\lambda_r < \lambda_M^{\text{psic}} \text{ and } \varepsilon \leq \varepsilon_{c,C}^{\text{psic}} \right) \text{ or } \lambda_r > \lambda_M^{\text{psic}}$$

$$(b) \qquad \left(\lambda_s < \lambda_M^{\text{psic}} \text{ and } \varepsilon_{c,C}^{\text{psic}} < \varepsilon \leq \varepsilon_{k,C}^{\text{psic}} \right)$$

$$\text{or } \left(\lambda_r \leq \lambda_M^{\text{psic}} < \lambda_s \text{ and } \varepsilon > \varepsilon_{c,C}^{\text{psic}} \right)$$

$$(c) \qquad \lambda_s < \lambda_M^{\text{psic}} \text{ and } \varepsilon > \varepsilon_{k,C}^{\text{psic}}$$

respectively. ∎

It remains to maximize $\lambda_{f,M}^{\varepsilon_f, \text{psic}} \wedge \lambda_u^{\varepsilon_u, \text{psic}}$ over all $(\varepsilon_u, \varepsilon_f)$ such that $\varepsilon_u + \varepsilon_f = \varepsilon$. Note that although the expressions are quite messy the actual algorithm to find the optimal pair

is quite simple: find the splitting $(\varepsilon_u, \varepsilon_f)$ summing to ε that minimizes the distance

$$\left| \lambda_{f,M}^{\varepsilon_f, \text{psic}} - \lambda_u^{\varepsilon_u, \text{psic}} \right|, \tag{5.50}$$

where the optimal splitting pair is found by finding the intersection of the two functions for the cases when the minimum distance is zero. For those cases where the two functions do not intersect then the optimal pair is trivial: the smaller function gets ε and the larger function gets 0. This can be easily done on a computer; we will study this algorithm in the numerical results section. $\qquad\square$

Note that we have shown the Markov lower bound for the perfect case has a capacity improvement that is independent of K. Now consider the Chebychev lower bound and we can show a similar result. For small ε the capacity can be shown to be

$$\lambda_{l,C}^{\varepsilon, \text{psic}} = (\alpha - 1) \frac{\varepsilon}{\pi(\bar{\beta}d)^2} + O(\varepsilon^2), \tag{5.51}$$

with a corresponding performance improvement of

$$\frac{\lambda_{l,C}^{\varepsilon, \text{psic}}}{\lambda_{l,C}^{\varepsilon, \text{nsic}}} = \alpha \bar{\beta}^{2(\frac{1}{\alpha} - 1)}, \tag{5.52}$$

which again is independent of K, giving similar intuition and scaling to that of the Markov lower bound.

5.4.2 Appendix: Complete Results of Theorem 5.2.1 - Imperfect SIC Case.

The major result is again a set of expressions for lower and upper bounds on the transmission capacity. The expressions for the lower bound are given in terms of an easy optimization problem the solution of which is trivial for a computer.

Theorem 5.4.2. Let $\varepsilon \in (0,1)$ and $y = \frac{d^{-\alpha}}{\bar{\beta}}$. As $\varepsilon \to 0$, the lower and upper bounds on the transmission capacity when receivers are equipped with imperfect SIC ($\zeta \in (0,1)$) are:

$$c_l^{\varepsilon, \text{sic}} = (1 - \varepsilon)\lambda_l^{\varepsilon, \text{sic}}, \quad c_u^{\varepsilon, \text{sic}} = (1 - \varepsilon)\lambda_u^{\varepsilon, \text{sic}}. \tag{5.53}$$

The upper bound on the optimal contention density is:

$$\lambda_u^{\varepsilon, \text{sic}} = \begin{cases} \dfrac{1}{\zeta^{\frac{2}{\alpha}}} \dfrac{-\ln(1-\varepsilon)}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} \\[2.5ex] \dfrac{K}{1+\zeta^{\frac{2}{\alpha}}} \dfrac{1}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} + \dfrac{1}{1+\zeta^{\frac{2}{\alpha}}} \dfrac{-\ln(1-\varepsilon)}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} \\[2.5ex] \dfrac{-\ln(1-\varepsilon)}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} \end{cases} \tag{5.54}$$

where the three expressions hold for the regimes

$$\varepsilon \le \varepsilon_{k,u}^{\text{sic}}, \quad \varepsilon_{k,u}^{\text{sic}} < \varepsilon \le \varepsilon_{l,u}^{\text{sic}}, \quad \varepsilon \ge \varepsilon_{l,u}^{\text{sic}}. \tag{5.55}$$

The constants are given by:

$$\varepsilon_{k,u}^{\text{sic}} = 1 - \exp\left\{-K\zeta^{\frac{2}{\alpha}}\right\}, \quad \varepsilon_{l,u}^{\text{sic}} = 1 - \exp\left\{-K\zeta^{-\frac{2}{\alpha}}\right\}. \tag{5.56}$$

The Markov (M) lower bound on the optimal contention density is:

$$\lambda_{l,M}^{\varepsilon,\text{sic}} \ge \sup_{(\varepsilon_u,\varepsilon_f):\varepsilon_u+\varepsilon_f=\varepsilon} \left\{\lambda_u^{\varepsilon_u,\text{sic}} \wedge \lambda_{f,M}^{\varepsilon_f,\text{sic}}\right\}, \tag{5.57}$$

where

$$\lambda_{f,M}^{\varepsilon,\text{sic}} = \begin{cases} \dfrac{\alpha-2}{2} \dfrac{\bar{\beta}^{\frac{2}{\alpha}}}{(1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}} \dfrac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}, & \varepsilon \le \varepsilon_{c,M}^{\text{sic}} \\ \textit{see below}, & \varepsilon_{c,M}^{\text{sic}} \le \varepsilon < \varepsilon_{k,M}^{\text{sic}} \\ \dfrac{\alpha-2}{2} \dfrac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}, & \varepsilon > \varepsilon_{k,M}^{\text{sic}} \end{cases} \tag{5.58}$$

and $\lambda_{f,M}^{\varepsilon,\text{sic}}$ for $\varepsilon_{c,M}^{\text{sic}} \le \varepsilon < \varepsilon_{k,M}^{\text{sic}}$ is the unique solution to the equation:

$$\frac{2\pi\lambda}{(\alpha-2)y}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}} + \zeta y^{1-\frac{2}{\alpha}}\right] = \varepsilon. \tag{5.59}$$

The constants are given by:

$$\varepsilon_{c,M}^{\text{sic}} = \left[(1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}\right]\frac{2K}{\alpha-2}, \quad \varepsilon_{k,M}^{\text{sic}} = \frac{2K}{\alpha-2} \tag{5.60}$$

The Chebychev (C) lower bound on the optimal contention density is:

$$\lambda_{l,C}^{\varepsilon,\text{sic}} \ge \sup_{(\varepsilon_u,\varepsilon_f):\varepsilon_u+\varepsilon_f=\varepsilon} \left\{\lambda_u^{\varepsilon_u,\text{sic}} \wedge \lambda_{f,C}^{\varepsilon_f,\text{sic}}\right\}, \tag{5.61}$$

where

$$\lambda_{f,C}^{\varepsilon,\text{sic}} = \begin{cases} \dfrac{\alpha-1}{(1-\zeta)\bar{\beta}^{2(1-\frac{1}{\alpha})}+\zeta} \dfrac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} + O(\varepsilon^2) \\ \textit{see below} \\ (\alpha-1)\dfrac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} + O(\varepsilon^2) \end{cases} \tag{5.62}$$

For $\varepsilon_{c,C}^{\text{sic}} < \varepsilon \le \varepsilon_{k,C}^{\text{sic}}$ $\lambda_{f,C}^{\varepsilon,\text{sic}}$ is the unique solution to the equation:

$$\frac{\frac{\pi}{\alpha-1}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\alpha} + \zeta y^{2(1-\frac{1}{\alpha})}\right]\lambda}{\left(y - \frac{2\pi}{\alpha-2}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}} + \zeta y^{1-\frac{2}{\alpha}}\right]\lambda\right)^2} = \varepsilon. \tag{5.63}$$

79

These three expressions (a), (b), (c) hold for the regimes

$$(a) \quad \left(\lambda_r < \lambda_M^{\text{sic}} \text{ and } \varepsilon \le \varepsilon_{c,C}^{\text{sic}}\right) \text{ or } \lambda_r > \lambda_M^{\text{sic}}$$

$$(b) \quad \left(\lambda_s < \lambda_M^{\text{sic}} \text{ and } \varepsilon_{c,C}^{\text{sic}} < \varepsilon \le \varepsilon_{k,C}^{\text{sic}}\right)$$

$$\text{or } \left(\lambda_r \le \lambda_M^{\text{sic}} < \lambda_s \text{ and } \varepsilon > \varepsilon_{c,C}^{\text{sic}}\right)$$

$$(c) \quad \lambda_s < \lambda_M^{\text{sic}} \text{ and } \varepsilon > \varepsilon_{k,C}^{\text{sic}}$$

respectively. The constants are given by:

$$\lambda_M^{\text{sic}} = \begin{cases} \frac{\alpha-2}{2\pi d^2\left((1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}\right)}, & \alpha-2-2K\left((1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}\right) < 0 \\ \textit{see below}, & \textit{else} \\ \frac{\alpha-2}{2\pi d^2\bar{\beta}^{\frac{2}{\alpha}}}, & \alpha-2-2K > 0 \end{cases} \tag{5.64}$$

where λ_M^{sic} is given by the unique solution of the equation

$$\frac{2\pi\lambda}{\alpha-2}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}}+\zeta y^{1-\frac{2}{\alpha}}\right] = y \tag{5.65}$$

when $\alpha-2-2K\left((1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}\right) \ge 0$ and $\alpha-2-2K \le 0$, and

$$\lambda_r = \frac{K}{\pi}d^{-2}, \quad \lambda_s = \frac{K}{\pi}y^{\frac{2}{\alpha}} \tag{5.66}$$

and

$$\varepsilon_{c,C}^{\text{sic}} = \frac{K(\alpha-2)^2\left((1-\zeta)\bar{\beta}^2+\zeta\bar{\beta}^{\frac{2}{\alpha}}\right)}{(\alpha-1)\left((\alpha-2)-2K\left((1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}\right)\right)^2}$$

$$\varepsilon_{k,C}^{\text{sic}} = \frac{K(\alpha-2)^2}{(\alpha-1)(\alpha-2-2K)^2}.$$

Proof. It was earlier shown that $r = r_s$ is a critical radius in the sense that it is the maximum distance a single interfering node can be from a receiver and still generate sufficient interference to cause an outage at that receiver. When the interference is partially cancelled through the use of imperfect SIC with parameter $\zeta \in (0,1)$ then the corresponding critical radius is $\zeta^{\frac{1}{\alpha}}r_s$: if the partially cancelled node is any further away it cannot by itself cause an outage. With this in mind, define the following three spatial density thresholds and compute the corresponding value of r_{sic}:

$$\lambda_r = \frac{K}{\pi}d^{-2}, \quad \lambda_s = \frac{K}{\pi}y^{\frac{2}{\alpha}}, \quad \lambda_l = \frac{K}{\pi}\zeta^{-\frac{2}{\alpha}}y^{\frac{2}{\alpha}}$$
$$r_{\text{sic}} = d, \quad r_{\text{sic}} = r_s, \quad r_{\text{sic}} = \zeta^{\frac{1}{\alpha}}r_s. \tag{5.67}$$

Looking at Figure 5.9, for each possible spatial density of transmissions λ we can identify the distances from the receiver where an interfering node at that distance will by itself generate sufficient interference to cause an outage. For example, for $\lambda \in (\lambda_s, \lambda_l)$ nodes at distances in the range $(0, \zeta^{\frac{1}{\alpha}} r_s)$ generate interference that is partially cancelled but even so sufficient to cause an outage, and nodes at distances in the range $(\sqrt{\frac{K}{\lambda}}, r_s)$ are outside the cancellation radius but are nonetheless sufficiently close to cause an outage.
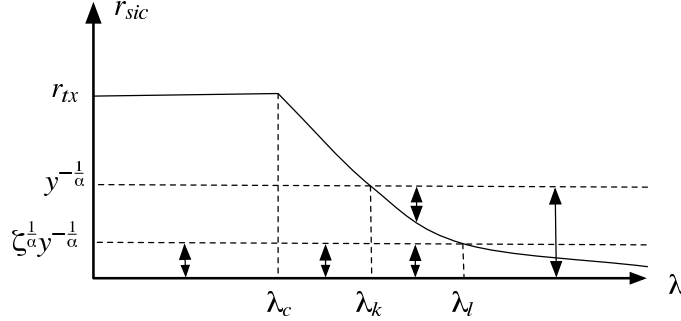
Figure 5.9: The cancellation radius r_{sic} versus the spatial transmission density λ. The near/far field separation radius is r_s, this is also the farthest distance that an uncanceled node can be from the receiver and still cause an outage. The farthest distance that a canceled node can be from the receiver and still cause an outage is $\zeta^{\frac{1}{\alpha}} r_s$. The arrows denote the annular regions around the receiver where a single node could cause outage provided that node is in the near field.

The upper bound event corresponding to Figure 5.9 is:

$$
F_u^{\text{sic}}(\lambda) = \begin{cases} \{\Pi \cap b(O, \zeta^{\frac{1}{\alpha}} r_s)\} \\ \{\Pi \cap \left(b(O, \zeta^{\frac{1}{\alpha}} r_s) \cup a(O, \sqrt{\frac{K}{\pi\lambda}}, r_s)\right)\} \\ \{\Pi \cap b(O, r_s)\} \end{cases} \tag{5.68}
$$

where the three cases apply for the intervals

$$
\lambda \leq \lambda_s, \quad \lambda_s < \lambda \leq \lambda_l, \quad \lambda > \lambda_l \tag{5.69}
$$

respectively. The probability of the event is:

$$
\mathbb{P}^0(F_u^{\text{sic}}(\lambda)) = \begin{cases} 1 - \exp\{-\lambda\pi\zeta^{\frac{2}{\alpha}} r_s^2\} \\ 1 - \exp\{-\lambda\pi(1 + \zeta^{\frac{2}{\alpha}}) r_s^2 + K\} \\ 1 - \exp\{-\lambda\pi r_s^2\} \end{cases} \tag{5.70}
$$

The bound evaluated at the critical points λ_s, λ_l gives:

$$
\begin{aligned}
\varepsilon_{k,u}^{\text{sic}} &= \mathbb{P}^0(F_u^{\text{sic}}(\lambda_s)) = 1 - \exp\{-K\zeta^{\frac{2}{\alpha}}\} \\
\varepsilon_{l,u}^{\text{sic}} &= \mathbb{P}^0(F_u^{\text{sic}}(\lambda_l)) = 1 - \exp\{-K\zeta^{-\frac{2}{\alpha}}\}
\end{aligned}
$$

81

The map $\lambda \to \mathbb{P}^0(F_u^{\mathrm{sic}}(\lambda))$ is onto $[0,1)$ and monotone increasing in λ; hence a unique inverse exists for all $\varepsilon > 0$. Setting this expression equal to ε and solving for λ yields:

$$
\lambda_u^{\varepsilon,\mathrm{sic}} = \begin{cases}
\zeta^{-\frac{2}{\alpha}}\frac{1}{\pi}y^{\frac{2}{\alpha}}\left(-\ln(1-\varepsilon)\right), & \varepsilon \leq \varepsilon_{k,u}^{\mathrm{sic}} \\
\frac{1}{1+\zeta^{\frac{2}{\alpha}}}\frac{1}{\pi}y^{\frac{2}{\alpha}}\left(-\ln(1-\varepsilon)+K\right), & \varepsilon_{k,u}^{\mathrm{sic}} < \varepsilon \leq \varepsilon_{l,u}^{\mathrm{sic}} \\
\frac{1}{\pi}y^{\frac{2}{\alpha}}\left(-\ln(1-\varepsilon)\right), & \varepsilon > \varepsilon_{l,u}^{\mathrm{sic}}
\end{cases}
\tag{5.71}
$$

We turn now to the lower bound. Define the following events:

Definition 5.4.2.

$$
\begin{aligned}
F^{\mathrm{sic}}(\lambda) &= \left\{Y(\lambda) > y\right\}, \\
F_f^{\mathrm{sic}}(\lambda) &= \left\{Y(\lambda, r_s, \zeta) > y\right\}
\end{aligned}
$$

where $Y(\lambda)$ is given by Definition 5.2.2 and

$$
\begin{aligned}
Y(\lambda, r_s, \zeta) &= \sum_{\Pi \cap a(O, r_s, r_{\mathrm{sic}})} \zeta |X_i|^{-\alpha} \\
&+ \sum_{\Pi \cap \bar{b}(O, r_{\mathrm{sic}} \vee r_s)} |X_i|^{-\alpha}
\end{aligned}
$$

is the normalized aggregate interference generated by all partially cancelled nodes in the annulus $a(O, r_s, r_{\mathrm{sic}})$ plus the interference generated by the uncanceled nodes outside the radius $r_{\mathrm{sic}} \vee r_s$.

We first compute the Markov bound on $\mathbb{P}^0(F_f^{\mathrm{sic}}(\lambda))$: $\mathbb{E}^0[Y(\lambda, r_s, \zeta)]$

$$
= \begin{cases}
2\pi\lambda\left[\zeta \int_{r_s}^d r^{-\alpha}rdr + \int_d^\infty r^{-\alpha}rdr\right] \\
2\pi\lambda\left[\zeta \int_{r_s}^{\sqrt{\frac{K}{\pi\lambda}}} r^{-\alpha}rdr + \int_{\sqrt{\frac{K}{\pi\lambda}}}^\infty r^{-\alpha}rdr\right] \\
2\pi\lambda \int_{r_s}^\infty r^{-\alpha}rdr
\end{cases}
$$

$$
= \begin{cases}
\frac{2\pi\lambda}{\alpha-2}\left[(1-\zeta)d^{2-\alpha} + \zeta y^{1-\frac{2}{\alpha}}\right] \\
\frac{2\pi\lambda}{\alpha-2}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}} + \zeta y^{1-\frac{2}{\alpha}}\right] \\
\frac{2\pi\lambda}{\alpha-2}y^{1-\frac{2}{\alpha}}
\end{cases}
$$

where the three expressions hold for the intervals

$$
\lambda \leq \lambda_r, \quad \lambda_r < \lambda \leq \lambda_s, \quad \lambda > \lambda_s
\tag{5.72}
$$

respectively. The Markov bound is:

$$
\mathbb{E}^0[Y(\lambda, r_s, \zeta)]/y =
\begin{cases}
\frac{2\pi\lambda}{(\alpha-2)y}\left[(1-\zeta)d^{2-\alpha}+\zeta y^{1-\frac{2}{\alpha}}\right] \\[2mm]
\frac{2\pi\lambda}{(\alpha-2)y}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}}+\zeta y^{1-\frac{2}{\alpha}}\right] \\[2mm]
\frac{2\pi\lambda}{\alpha-2}r_s^2
\end{cases}
\tag{5.73}
$$

The value of the bound at the critical points is

$$
\begin{aligned}
\varepsilon_{c,M}^{\text{sic}} &= \frac{\mathbb{E}[Y(\lambda_r, r_s, \zeta)]}{y} = \left[(1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}\right]\frac{2K}{\alpha-2} \\[2mm]
\varepsilon_{k,M}^{\text{sic}} &= \frac{\mathbb{E}[Y(\lambda_s, r_s, \zeta)]}{y} = \frac{2K}{\alpha-2}
\end{aligned}
$$

This function is monotone increasing in λ and is onto $[0,1]$; hence the inverse function exists. Setting the expression equal to ε and solving for λ gives:

$$
\lambda_{f,M}^{\varepsilon,\text{sic}} =
\begin{cases}
\frac{\alpha-2}{2}\frac{\bar{\beta}^{\frac{2}{\alpha}}}{(1-\zeta)\bar{\beta}+\zeta\bar{\beta}^{\frac{2}{\alpha}}}\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}, & \varepsilon \le \varepsilon_{c,M} \\[2mm]
\text{see below}, & \varepsilon_{c,M} \le \varepsilon < \varepsilon_{k,M} \\[2mm]
\frac{\alpha-2}{2}\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2}, & \varepsilon > \varepsilon_{k,M}
\end{cases}
\tag{5.74}
$$

It is not possible to obtain a closed form expression for $\lambda_{f,M}^{\varepsilon,\text{sic}}$ for $\varepsilon_{c,M}^{\text{sic}} \le \varepsilon < \varepsilon_{k,M}^{\text{sic}}$; it is the unique solution to the equation:

$$
\frac{2\pi\lambda}{(\alpha-2)y}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}}+\zeta y^{1-\frac{2}{\alpha}}\right] = \varepsilon.
\tag{5.75}
$$

The same comments made in the proof of Theorem 4.2.2 regarding finding the optimal splitting pair to maximize $\lambda_{f,M}^{\varepsilon_f,\text{sic}} \wedge \lambda_u^{\varepsilon_u,\text{sic}}$ over all $(\varepsilon_u, \varepsilon_f)$ such that $\varepsilon_u + \varepsilon_f = \varepsilon$ hold here.

We now consider the Chebychev lower bound. The variance of the far-field interference is again found by Campbell's Theorem: $Var(Y(\lambda, r_s, \zeta))$

$$
\begin{aligned}
&=
\begin{cases}
2\pi\lambda\left[\zeta\int_{r_s}^{d}r^{-2\alpha}rdr+\int_{d}^{\infty}r^{-2\alpha}rdr\right] \\[2mm]
2\pi\lambda\left[\zeta\int_{r_s}^{\sqrt{\frac{K}{\pi\lambda}}}r^{-2\alpha}rdr+\int_{\sqrt{\frac{K}{\pi\lambda}}}^{\infty}r^{-2\alpha}rdr\right] \\[2mm]
2\pi\lambda\int_{r_s}^{\infty}r^{-2\alpha}rdr
\end{cases} \\[4mm]
&=
\begin{cases}
\frac{\pi\lambda}{\alpha-1}\left[(1-\zeta)d^{2-2\alpha}+\zeta y^{2(1-\frac{1}{\alpha})}\right] \\[2mm]
\frac{\pi\lambda}{\alpha-1}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\alpha}+\zeta y^{2(1-\frac{1}{\alpha})}\right] \\[2mm]
\frac{\pi\lambda}{\alpha-1}y^{2(1-\frac{1}{\alpha})}
\end{cases}
\end{aligned}
$$

where the three expressions hold for the intervals

$$
\lambda \le \lambda_r, \quad \lambda_r < \lambda \le \lambda_s, \quad \lambda > \lambda_s
\tag{5.76}
$$

83

respectively. The Chebychev bound is: for $y > \mathbb{E}^0[Y(\lambda, r_s, \zeta)]$:

$$\mathbb{P}^0(F_f^{\text{sic}}(\lambda)) \leq \frac{Var(Y(\lambda, r_s, \zeta))}{\left(y - \mathbb{E}^0[Y(\lambda, r_s, \zeta)]\right)^2}$$

$$= \begin{cases} \dfrac{\frac{\pi}{\alpha-1}\left[(1-\zeta)d^{2-2\alpha}+\zeta y^{2(1-\frac{1}{\alpha})}\right]\lambda}{\left(y - \frac{2\pi}{\alpha-2}\left[(1-\zeta)d^{2-\alpha}+\zeta y^{1-\frac{2}{\alpha}}\right]\lambda\right)^2} \\[20pt] \dfrac{\frac{\pi}{\alpha-1}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\alpha}+\zeta y^{2(1-\frac{1}{\alpha})}\right]\lambda}{\left(y - \frac{2\pi}{\alpha-2}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}}+\zeta y^{1-\frac{2}{\alpha}}\right]\lambda\right)^2} \\[20pt] \dfrac{\frac{\pi}{\alpha-1}y^{2(1-\frac{1}{\alpha})}\lambda}{\left(y - \frac{2\pi}{\alpha-2}y^{1-\frac{2}{\alpha}}\lambda\right)^2} \end{cases}$$

where the three expressions hold for the same three intervals as above. The value of the bound at the critical points is

$$\varepsilon_{c,C}^{\text{sic}} = \frac{Var(Y(\lambda_r, r_s, \zeta))}{\left(y - \mathbb{E}^0[Y(\lambda_r, r_s, \zeta)]\right)^2}$$

$$= \frac{K(\alpha-2)^2\left((1-\zeta)\bar{\beta}^2 + \zeta\bar{\beta}^{\frac{2}{\alpha}}\right)}{(\alpha-1)\left((\alpha-2) - 2K\left((1-\zeta)\bar{\beta} + \zeta\bar{\beta}^{\frac{2}{\alpha}}\right)\right)^2}$$

$$\varepsilon_{k,C}^{\text{sic}} = \frac{Var(Y(\lambda_s, r_s, \zeta))}{\left(y - \mathbb{E}^0[Y(\lambda_s, r_s, \zeta)]\right)^2}$$

$$= \frac{K(\alpha-2)^2}{(\alpha-1)(\alpha-2-2K)^2}.$$

Setting the three expressions in the Chebychev bound equal to ε and solving for λ yields rather unwieldy expressions. The first and third cases are in the form of (5.45); for simplicity we employ the solution given in the right side of (5.46) which is obtained by linearizing in ε around $\varepsilon = 0$. Applying this result to the first and third cases above we find:

$$\lambda_{f,C}^{\varepsilon,\text{sic}} = \begin{cases} \frac{\alpha-1}{(1-\zeta)\bar{\beta}^{2(1-\frac{1}{\alpha})}+\zeta}\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} + O(\varepsilon^2) \\[12pt] \text{see below} \\[8pt] (\alpha-1)\frac{\varepsilon}{\pi\left(\bar{\beta}^{\frac{1}{\alpha}}d\right)^2} + O(\varepsilon^2) \end{cases} \tag{5.77}$$

The second case, as was also true for the Markov bound, cannot be put in closed form. It is given by the unique solution of the equation:

$$\frac{\frac{\pi}{\alpha-1}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\alpha}+\zeta y^{2(1-\frac{1}{\alpha})}\right]\lambda}{\left(y - \frac{2\pi}{\alpha-2}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}}+\zeta y^{1-\frac{2}{\alpha}}\right]\lambda\right)^2} = \varepsilon. \tag{5.78}$$

84

As was discussed in the proof of Theorem 4.2.2, the condition that $y > \mathbb{E}^0[Y(\lambda, r_s, \zeta)]$ can be expressed as $\lambda < \lambda_M^{\text{sic}}$ where λ_M^{sic} is the unique density such that $\mathbb{E}^0[Y(\lambda_M^{\text{sic}}, r_s, \zeta)] = y$. Straightforward algebra yields:

$$\lambda_M^{\text{sic}} = \begin{cases} \frac{\alpha-2}{2\pi d^2 \left((1-\zeta)\bar{\beta} + \zeta \bar{\beta}^{\frac{2}{\alpha}}\right)}, & \alpha - 2 - 2K\left((1-\zeta)\bar{\beta} + \zeta\bar{\beta}^{\frac{2}{\alpha}}\right) < 0 \\ \text{see below,} & \text{else} \\ \frac{\alpha-2}{2\pi d^2 \bar{\beta}^{\frac{2}{\alpha}}}, & \alpha - 2 - 2K > 0 \end{cases} \qquad (5.79)$$

where λ_M^{sic} is given by the unique solution of the equation

$$\frac{2\pi\lambda}{\alpha-2}\left[(1-\zeta)\left(\frac{K}{\pi\lambda}\right)^{1-\frac{\alpha}{2}} + \zeta y^{1-\frac{2}{\alpha}}\right] = y \qquad (5.80)$$

when $\alpha - 2 - 2K\left((1-\zeta)\bar{\beta} + \zeta\bar{\beta}^{\frac{2}{\alpha}}\right) \geq 0$ and $\alpha - 2 - 2K \leq 0$. Again referring to Figure 5.8, it is apparent that the appropriate expression for the inverse depends on both ε's position relative to $\varepsilon_{c,C}^{\text{sic}}$ and $\varepsilon_{k,C}^{\text{sic}}$ as well as λ_M^{sic}'s position relative to λ_r and λ_s. In particular, the three expressions above hold for

$$(a) \qquad \left(\lambda_r < \lambda_M^{\text{sic}} \text{ and } \varepsilon \leq \varepsilon_{c,C}^{\text{sic}}\right) \text{ or } \lambda_r > \lambda_M^{\text{sic}}$$

$$(b) \qquad \left(\lambda_s < \lambda_M^{\text{sic}} \text{ and } \varepsilon_{c,C}^{\text{sic}} < \varepsilon \leq \varepsilon_{k,C}^{\text{sic}}\right)$$

$$\text{or } \left(\lambda_r \leq \lambda_M^{\text{sic}} < \lambda_s \text{ and } \varepsilon > \varepsilon_{c,C}^{\text{sic}}\right)$$

$$(c) \qquad \lambda_s < \lambda_M^{\text{sic}} \text{ and } \varepsilon > \varepsilon_{k,C}^{\text{sic}}$$

respectively.

It remains to maximize $\lambda_{f,C}^{\varepsilon_f,\text{sic}} \wedge \lambda_u^{\varepsilon_u,\text{sic}}$ over all $(\varepsilon_u, \varepsilon_f)$ such that $\varepsilon_u + \varepsilon_f = \varepsilon$. The same comments apply here that were made for selecting the optimal splitting pair for the Markov bound: finding the optimal pair is trivial for a computer, whereas the corresponding expressions for the optimal are both messy and don't necessarily provide any insight.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

It is relevant to check whether these bounds are correct by setting parameter ζ to be boundary values, i.e., 0 and 1, and comparing with the cases of perfect SIC and conventional CDMA without SIC. First Consider the upper bound. Note that

$$\lim_{\zeta \to 1} \varepsilon_{k,u}^{\text{sic}} = \lim_{\zeta \to 1} \varepsilon_{l,u}^{\text{sic}} = 1 - e^{-K}, \qquad (5.81)$$

and that $\lim_{\zeta \to 1} \lambda_u^{\varepsilon,\text{sic}} = \lambda_u^{\varepsilon,\text{nsic}}$ as expected. Next note that

$$\lim_{\zeta \to 0} \varepsilon_{k,u}^{\text{sic}} = 0, \quad \lim_{\zeta \to 0} \varepsilon_{l,u}^{\text{sic}} = 1, \qquad (5.82)$$

and that $\lim_{\zeta \to 0} \lambda_u^{\varepsilon,\mathrm{sic}} = \lambda_u^{\varepsilon,\mathrm{psic}}$ as expected.

Now consider the Markov and Chebychev lower bounds. When $\zeta = 0$ it is straightforward to see the expressions for $\lambda_{l,M}^{\varepsilon,\mathrm{sic}}, \lambda_{l,C}^{\varepsilon,\mathrm{sic}}$ will reduce to those of $\lambda_{l,M}^{\varepsilon,\mathrm{psic}}, \lambda_{l,C}^{\varepsilon,\mathrm{psic}}$ after appropriate choice of the optimal splitting pair $(\varepsilon_u, \varepsilon_f)$. Similarly, when $\zeta = 1$ it is straightforward to see the expressions for $\lambda_{l,M}^{\varepsilon,\mathrm{sic}}, \lambda_{l,C}^{\varepsilon,\mathrm{sic}}$ will reduce to those of $\lambda_{l,M}^{\varepsilon,\mathrm{nsic}}, \lambda_{l,C}^{\varepsilon,\mathrm{nsic}}$ after appropriate choice of the optimal splitting pair $(\varepsilon_u, \varepsilon_f)$.

Therefore, the bounds for the case of imperfect SIC case with parameter (K, ζ) become the bounds for the other three cases under the appropriate substitutions. Hence, the imperfect SIC case can be considered to be a generalized result for the transmission capacity of DS-CDMA ad hoc networks.

Chapter 6

Induce Spatial Clustering in MAC Contention

One can consider inducing spatial clustering of transmitters in many ways. To intuitively show the benefit of inducing clustering we first consider an idealized deterministic placement. Then we propose two approaches. The first assumes that nodes contend synchronously and are aware of their locations. These capabilities are used to directly generate a spatial clustering of transmitters. The second assumes nodes are able to monitor interference levels to roughly infer relative locations of nodes and use signaling stages to achieve clustering of transmissions. We use these two representative distributed mechanisms to exhibit the benefits of *inducing spatial clustering* in a practical system.

'Clustering' in ad hoc networks is discussed in [28][27][23][4][57]. The scope of the clustering approach in these papers is quite different from ours in that they deal with realizing a virtual hierarchical clustered structure *above MAC layer*, which enables efficient routing or enables nodes to do power control within a local cluster based on existing non-homogeneities in the spatial distribution of nodes or network connectivity. In addition, the clusters they created are meant to be maintained for relatively long time scales for routing purposes. One of the novel ideas in our work is an exploration of intentionally *inducing* clustering in MAC contention processes so as to enhance the capacity of spread spectrum-based ad hoc networks. Our method of inducing clusters naturally structures clusters during each packet transmission slot, i.e., a much smaller time scale. To our knowledge there is no previous work suggesting the benefits of clustered contention processes *at the MAC layer* of spread spectrum ad hoc networks nor how they might be optimized to realize its substantial promise towards achieving high spatial reuse, QoS and energy efficiency.

6.1 Idealized deterministic clustering.

We begin by considering an idealized deterministic clustering of transmitters and receivers. The pioneering work of [24] showed that the optimal spatial reuse can be achieved by placing transmitters/receivers on a regular grid. We shall extend their result to a spread spectrum ad hoc network where optimal spatial reuse is achieved by clustering.

Assume we are free to select *both* the locations and states, e.g., transmitter or receiver, of nodes. Specifically, as shown on the left in Fig. 6.1, we assume tight clusters of transmitters and receivers are placed at the centers of cells in a regular square grid of size d^2 according to a checkerboard pattern. Each transmitter is assumed to transmit to a distinct receiver in one of the four neighboring cells and so has a fixed transmission range d. The number of nodes n within each cluster will be determined to ensure all transmissions are successful. Let π_n^d denote a set of locations in $\mathbb{R}^2$, corresponding to n-clustered transmitters in this checkerboard configuration, and where the origin corresponds to the center of a receive cluster cell. As seen earlier, in order for a transmitter at X_0 to successfully send to a receiver a distance d located at origin one must have that

$$\frac{\rho d^{-\alpha}}{\sum_{X_i \in \pi_n^d \setminus \{X_0\}} \rho |X_i|^{-\alpha}} \geq \frac{\beta}{m}$$

Fact 6.1.1 states this constraint in terms of a maximum allowable number of transmit nodes per cluster.

Fact 6.1.1. *Under the clustered grid placement of transmit nodes π_n^d, and $\alpha > 2$, a maximum number of nodes per cluster of*

$$\lfloor \frac{m/\beta + 1}{k(\alpha)} \rfloor, \text{ where } k(\alpha) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{4}{((2i)^2 + (2j+1)^2)^{\alpha}}$$

can be placed while ensuring no outage. This gives a density of successful transmissions of
$\lambda_s = \frac{1}{2d^2} \lfloor \frac{m/\beta+1}{k(\alpha)} \rfloor.\square$

This fact is shown through a brute force calculation of the aggregate interference offered by transmitter clusters in π_n^d at various ranges from the origin, which equals $nk(\alpha)d^{-\alpha}\rho$. Since $k(\alpha) \approx 4$ this suggests we need only consider the interference due to the $4n$ nodes which are a distance d from the origin. Thus in the Section 6.2 we will also focus on interference from the nearest clusters.

Comparing the best achievable spatial density of successful transmissions for randomly distributed MAC versus our ideally clustered grid, i.e., (4.13) to Fact 6.1.1, we have an approximate gain factor of $\frac{\pi}{2} \left(\frac{m}{\beta} \right)^{1-\frac{2}{\alpha}}$, e.g., when $\alpha = 4$, $m = 512$ and $\beta = 10$dB a 10-fold increase in capacity.
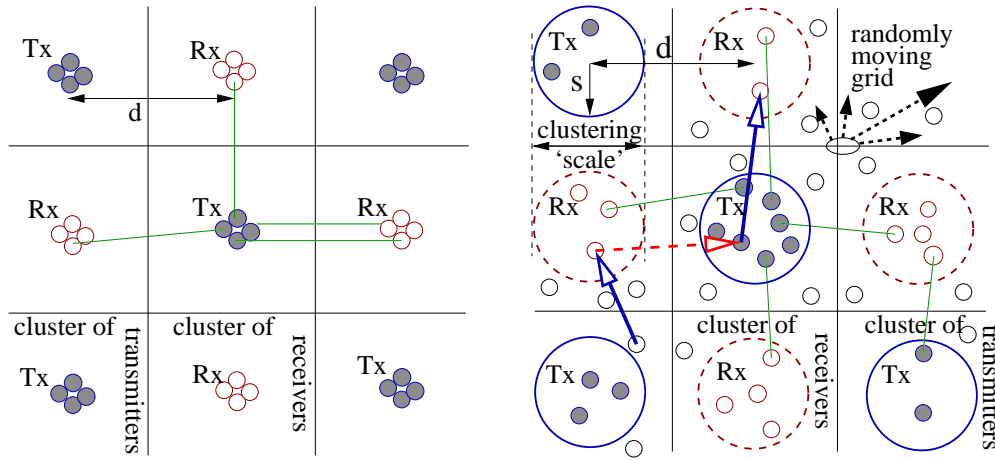
Figure 6.1: On the left an idealized deterministic placement. On the right a clustering of randomly located nodes in a random virtual grid with an example of typical routing patterns marked with hollow-headed arrows.

6.2 Inducing clustering based on a virtual grid.

In practice, one can not choose the placement of transmitters and receivers, i.e., it is unrealistic to expect that one can arrange for such a checkerboard clustered pattern. Yet for a homogeneous distribution of transmitters wishing to send a distance of roughly d one can approximate this pattern. Suppose, for example, that nodes are location aware and can determine their location relative to a known *virtual grid* of span d whose location evolves in a 'random' but known manner with time. This can be achieved with GPS-capable synchronized nodes, that share a randomization seed driving the evolution of the grid – [38][36] have used this shared seed idea to allow nodes to infer other nodes' states. Given this information and an a priori convention, a node can determine if it lies within a current transmitter/receiver cluster. As shown on the right of Fig. 6.1, we assume for now that nodes within a transmitter cluster transmit/relay to receivers in a neighboring receiver cluster.

Furthermore we let the parameter s determine the spatial scale of clustering and thus proximity of clustered nodes. Note that nodes that do not fall in either a transmitter or a receiver cluster region can defer, e.g., enter the sleep mode, unless they are sources or destinations.[1] 'Random moving' of the grid happens in relatively large time scale and is mainly for balancing long-term energy consumption. If s is too small, each cluster will contain but a few transmitters and we may under-utilize the available capacity. If it is too large, there may be too many transmitters and/or interference variability (due to increased

---

[1]This requires a routing protocol to give special considerations to the first and last hops, e.g., the first hop of the typical route on the right of Fig. 6.1, which should also require a proper power control.

proximity), resulting in outages at receivers. So one may consider what is a good choice for s depending on the intensity λ of the Poisson point process Π of active nodes.

Let us first evaluate the outage probability of a receiver at the center of a receiver cluster and use this as an approximate estimate for the outage probability of a typical receiver. As for the deterministic placement, we will focus on the nearest 4 transmit clusters as the source of interference. Using Campbell's theorem [45], we can evaluate the mean and variance of the interference as follows.

Fact 6.2.1. *Let Y denote the aggregate interference power level from the four transmit clusters closest to origin, i.e.,*

$$Y = \sum_{X_i \in \Pi} \mathbf{1}(X_i \cap A(s) \neq \emptyset) \times \rho|X_i|^{-\alpha}$$

where $A(s) = \bigcup_{i=-1,1} \bigcup_{j=-1,1} \mathbf{B}((i \times d, j \times d), s)$ is the union of these transmit cluster discs of radius s which are closest to the origin. Then

$$E[Y] = \lambda \int_{A(s)} \rho|x|^{-\alpha}dx, \quad Var(Y) = \lambda \int_{A(s)} \rho^2|x|^{-2\alpha}dx.$$

$\square$

Assuming Y is approximately Gaussian the outage probability for a receiver located at the origin is given by

$$p_o(\lambda, d, s) = \mathbb{P}(\frac{\rho d^{-\alpha}}{Y} < \frac{\beta}{m}) \approx \frac{1}{2} - \Phi\left(\frac{\frac{m}{\beta} - E[Y]}{\sqrt{Var(Y)}}\right). \tag{6.1}$$

Suppose in fact each transmitter finds a distinct receiver and thus there are no collisions due to concurrent transmissions to a receiver. Then for a fixed λ and d one can consider optimizing the cluster scale s so as to maximize the mean number of successful transmitters per cluster. Let n^* denote maximal mean number of successful transmissions per cluster, i.e., given by

$$n^*(\lambda, d) = \max_s \{\lambda \pi s^2 (1 - p_o(\lambda, d, s)) \mid s > 0\}, \tag{6.2}$$

and s^* denote the optimal s maximizing (6.2). In order to compare with the previous result in Theorem 4.2.1, we define the transmission capacity to be the density of successful transmissions, which in the case of square grid, is $\frac{n^*(\lambda, d)}{2d^2}$. We consider two regimes corresponding to different node density.

6.2.1 Regime 1: High node density.

In this regime, $\lambda \pi d^2$ is larger than $\frac{m}{\beta}$, resulting in an optimal clustering scale $s^* \ll d$, i.e., nodes are clustered closely around the center of each cell. This is akin to the deterministic

placement considered earlier. However each cluster will have a random, Poisson distributed, number of nodes with mean $\lambda\pi s^{*2}$. In this regime Fact 6.2.1 gives

$$E[Y] = 4\lambda\pi s^{*2}\rho d^{-\alpha} \ \text{ and } \ \text{Var}(Y) = \sigma_Y^2 = 4\lambda\pi s^{*2}\rho^2 d^{-2\alpha}.$$

We can in turn estimate the outage probability using (6.1). Note this optimization problem of (6.2) depends only on $z = \lambda\pi s^2$, the mean number of contending nodes per cluster.

As shown on the left panel of Fig. 6.2, in regime 1, the capacity achieved is close to the case of idealized deterministic placement. As shown on the right panel in Fig. 6.2, in this high density regime, we again obtain a transmission capacity that grows roughly linearly in $\frac{m}{\beta}$ as in the case of idealized deterministic placement, leading to a significant improvement over random access/ALOHA protocols, on the order of $\left(\frac{m}{\beta}\right)^{1-\frac{2}{\alpha}}$ or around 700% when $\frac{m}{\beta} = 50$, $\alpha = 4$.

6.2.2   Regime 2: Low node density.

In general if the intensity of active nodes is not high, the optimal choice of s^* will become comparable to d, i.e., one needs to increase the cluster scale so that a sufficient number of nodes can be scheduled. In this case the statistics of the interference are affected by both the variability of the number of nodes per cluster, and the increased interference variability due to their larger set of possible location within a cluster and relative to the origin. As shown on the left of Fig. 6.2 in the low density regime, if the spatial intensity λ is too small and there is not a sufficient number of nodes inside each cluster of size s, this negatively impacts the capacity. As λ grows larger, the capacity improves but the improvement eventually is marginal. Even in this regime, the achieved capacity is still significantly larger than random access/ALOHA. As shown on the right of Fig. 6.2, the capacity improves in $\frac{m}{\beta}$ sub-linearly but closer to linearly when λd^2 is larger.

6.2.3   Summary - virtual grid mechanisms.

Note that (not shown in Fig. 6.2) for both high density and low density regimes, the outage probability is significantly lower at the operating point achieving the highest transmission capacity, e.g., only about 5% for $\frac{m}{\beta} = 30$ compared with the outage probability of 50% in random access/ALOHA-like protocols, and this improves in m. When $\alpha = 2$, the capacity gain becomes marginal and independent of m. However, the benefit of low outage remains significant and improves with m.

An ad hoc network may have a non homogenous spatial density of nodes or traffic. Thus it is desirable to let the clustering scale adapt to such inhomogeneities. A straightforward approach would be to modify the our virtual grid mechanism such that each cell i
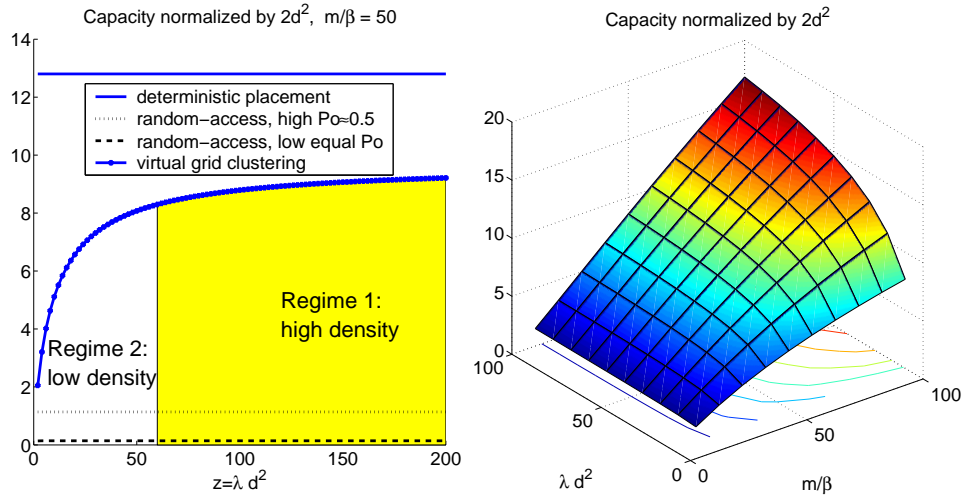
Figure 6.2: On the left, a comparison of capacities of different schemes. Here 'random access with low equal p_o' is obtained by setting the outage constraint ε in Theorem. 4.2.1 equal to the outage probability in the virtual-grid clustering protocol. On the right, the max capacity improves in both λd^2 and $\frac{m}{\beta}$.

has different cluster scale s_i. In this case the choice of s_i might be selected based on our analysis, i.e., $s_i^2 \approx \frac{1}{\pi \lambda_i} c(\frac{m}{\beta})$, where λ_i is the local node intensity in virtual cell i.

The virtual grid approach achieves a good tradeoff between capacity and energy efficiency. For example, nodes that are not covered by the transmit or receive cluster areas, can put themselves to sleep, until the grid moves. Furthermore, the overheads are low because nodes can infer locality of traffic and thus contention or signaling are not required, except that collisions where two or more transmitters send to the same receiver must be avoided. Collisions should be unlikely however, since the corresponding routing protocol should take advantage of long relay distance to achieve load balancing. The achievable capacity is close to that achieved by an ideal deterministic placement. Note that relatively low spatial intensity compared with spatial scale d may negatively impact the overall capacity because it prevents this mechanism from effectively inducing clustering.

6.3 Inducing clustering via multistage contention.

As seen in Section 4, protocols for ad hoc networks based on distributed contention resolution and handshaking induce weak clustering among scheduled nodes, see Fig. 4.10. Since this is particularly beneficial to enhance the capacity of spread-spectrum based ad hoc networks, one might ask how these mechanisms might be optimized to achieve clustering.

6.3.1 Spatial Packing versus Thinning

Contention based MAC protocol designs for narrow-band ad hoc networks mostly focus on the concept of 'thinning', i.e., given a set of contenders, use MAC layer contention resolution to reduce the number of contenders until (almost) all survivors can realize successful transmissions. Most existing MAC designs For spread spectrum ad hoc networks are also rooted in design concepts for narrow-band systems which do not fully leverage a CDMA PHY layer's capabilities. A common 'thinning' approach, e.g., IEEE 802.11b ad hoc mode, uses RTS and CTS signaling messages. We here first revisit the example in Fig. 4.9 from a protocol point of view. A transmitter intending to transmit sends an RTS to its receiver; the receiver, upon successfully receiving the RTS message, sends back a CTS message to confirm a successful handshake. As shown on the top of Fig.6.3, suppose three intended transmissions contend simultaneously but will interfere with each other, in particular, C interferes B and A interferes F. After contention, only the transmission from C→D can succeed in handshaking and proceed with data transmission, while B and F are not able to successfully receive RTS messages due to the interference from C and A respectively.

Instead, consider the 'packing' approach as shown at the bottom of Fig.6.3. Conceptually, we start by scheduling only a subset of the transmissions and then check whether it is possible to schedule (pack) more transmissions incrementally given the previously scheduled transmissions. In this example, if we start with the transmission E→F, then A→B should not be scheduled since A will interfere with receiver F. However it is possible to schedule the transmission C→D successfully without (severely) interfering with the transmission E→F. From this simple example, it is straightforward to observe the advantage of packing over thinning: with packing we might schedule two concurrent transmissions while with thinning we can at most schedule one transmission.

We can conceptually view 'thinning' and 'packing', as shown in Fig. 6.4, as a series of functions defined on the set of all contenders S.

- <u>Thinning</u> can be represented by functions $f_2, f_3, \ldots$ such that $S_{i+1} = f_{i+1}(S_i)$, where S_i is the set of surviving contenders at the beginning of Stage i and Stage1 includes all the contenders, i.e., $S = S_1$. If thinning ends at Stage n, the set of scheduled contenders is S_{n+1}. A MAC based on thinning will be designed such that the transmissions in S_{n+1} will be successful with high probability.

- <u>Packing</u> can be represented by functions $f_2', f_3', \ldots$ such that $S_{i+1} = f_{i+1}'(S_1^*, \ldots, S_i^*)$, where S_i is the set of contenders for Stage i and $S_i^* \subset S_i$ is the set of surviving contenders of Stage i. If packing ends at Stage n, the set of scheduled contenders is $S_1^* \cup S_2^* \ldots \cup S_n^*$, which again should correspond to a set of concurrent transmissions with high probability to be successful.

Below we consider a representative approach of spatial packing, starting with the
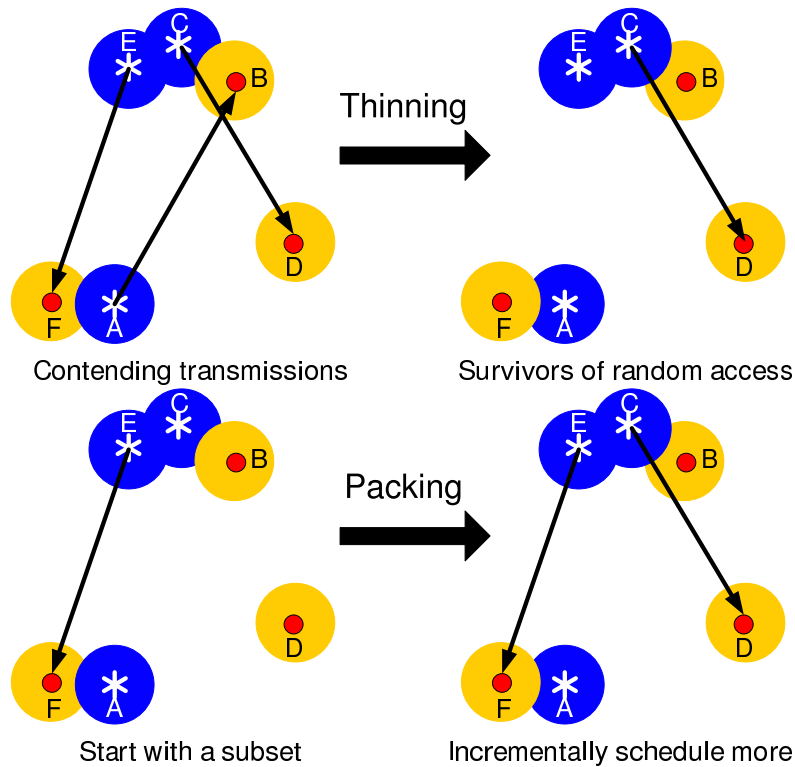
Figure 6.3: On the top, an example of thinning contenders with only one surviving transmission. On the bottom, an example of a packing of contenders with two surviving transmissions.

case where nodes use a common transmission distance/power and then connecting to the case where they use heterogeneous transmission distance/power. In particular we consider the following design criteria:

1. Each node has a single transceiver with which it can either transmit or receive at any point in time.

2. A receiver will not decode signaling messages that are not intended for it. [2].

3. Given these constraints, our design relies on sensing the signaling power during synchronous contention. Indeed, if contention and data transmission happen in fixed slots, nodes can therefore choose to contend or back off based only on measured

---

[2]Signaling may either use a separate common control code channel, or use the same code channel as data transmissions. A common signaling channel requires extra hardware complexity for nodes to consistently receive signaling messages and in addition the (narrowband) signaling channel may be congested. If signaling uses the same code that data transmission chooses, to receive other nodes' signaling messages, each node needs to search through the code space, which incurs significant overhead.
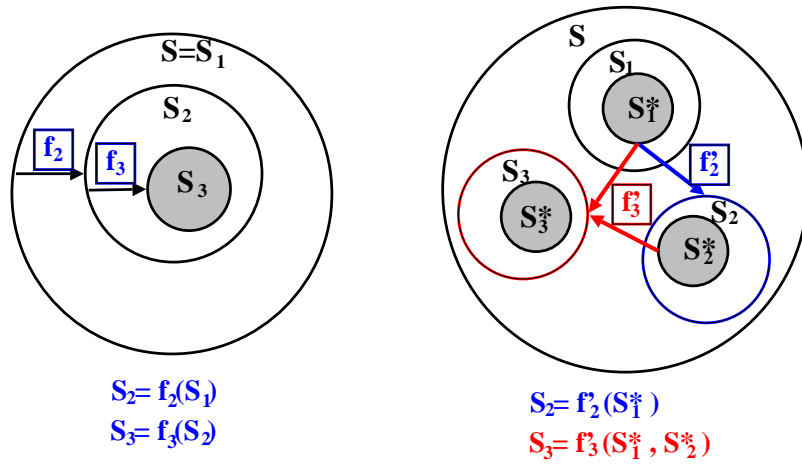
Figure 6.4: On the left, an abstract representation for thinning and on the right for packing.

RTS and CTS power levels, and need not actually receive these signaling messages in order to know how long a transmission will take.

We will show the spatial packing approach can efficiently induce clustering among transmissions and thus achieve high spatial reuse.

6.3.2 A multistage contention protocol.

Let us consider inducing clustering through a modified synchronous multi-stage RTS/CTS mechanism. Consider a two-stage example with the timing diagram shown in Fig. 6.5.

Stage 1 handshaking: In Stage 1 a subset of transmitters perform the three-way handshaking with their intended receivers, i.e., RTS, CTS, followed by an additional 'confirmation' RTS message. Only transmitter-receiver pairs who successfully exchange the three messages survive the first stage. These survivor pairs serve as 'seeds' for clusters in the subsequent handshaking stage(s).

Stage 1 monitoring: During Stage 1 contention, potential transmitters and receivers[3] not participating in the first stage handshaking process synchronously monitor interference levels, for which they can indeed distinguish RTS and CTS time slots. Doing so permits them to evaluate their proximity to surviving Stage 1 transmitters and receivers.

Stage 2 handshaking: In Stage 2, transmitters that sensed a 'strong' (see below) CTS signal in Stage 1 do not participate in Stage 2, i.e., are suppressed since they would likely interfere with the a successful Stage 1 receiver. Similarly a Stage 2 receiver which successfully

---

[3]Those who will not be active at this cycle do not need to monitor, which is more efficient than [30] in which all nodes have to do consistent monitoring.

receives an RTS from a transmitter, will only send back a CTS, if during Stage 1 it did no sense a 'strong' confirmation RTS signal. Thus the role of the Stage 1 'confirmation' RTS is to signal receivers in the Stage 2 that they will be interfered with and thus to suppress their CTS.

This process can be carried out through multiple stages to achieve a higher level of spatial reuse. Under the condition that *all transmissions use equal transmission power and relay distance*, contention might be performed in different ways, e.g., as shown in Fig. 6.6 survivors of Stage 1, might also concurrently participate in Stage 2 with 'virtual' RTS/CTS exchange, permitting actual Stage 2 contenders to estimate aggregate interference, rather than simply local interactions. We call this 'virtual' signaling because they will exchange RTS and CTS regardless the result of Stage 2 contention, e.g., a CTS is sent even the RTS is not received. This approach eliminates the need for the confirmation RTS slot and reduces the contention overheads.
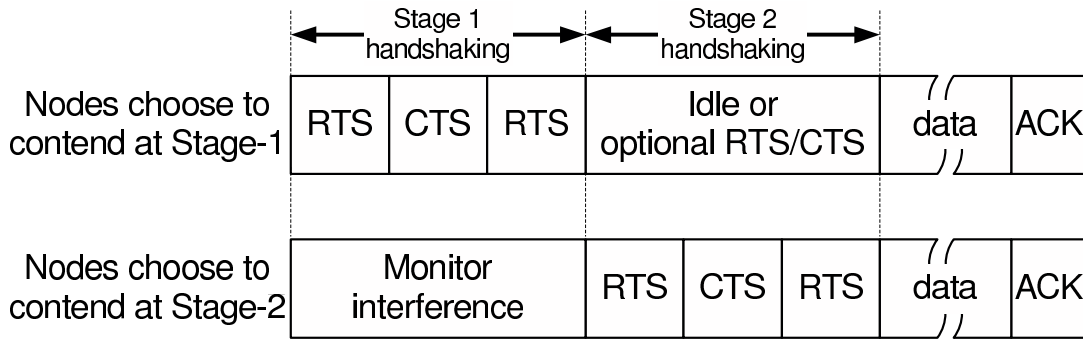

Figure 6.5: Timing diagrams of a two-stage contention MAC with the top for Stage 1 transmitter/receiver and the bottom for Stage 2 transmitter/receiver.

We need to formally define thresholds which are used to decide when signals should be deemed strong enough to result in suppression. The critical range analysis, see (3.1), suggests that a single interferer will cause outage for a transmitter-receiver pair using transmit power level ρ over a transmission range d, if the interference, as seen at the receiver exceeds $\frac{\rho d^{-\alpha} m}{\beta}$. To tolerate measurement uncertainty in the interference, we introduce a backoff factor c, where $0 < c \leq 1$ and thus a signal will be deemed strong if it exceeds $c \times \frac{\rho d^{-\alpha} m}{\beta}$. Note that c should be close to 1 otherwise we may be too conservative in utilizing available capacity. For purposes of visualizing this with 'dumbbells' and analytically studying clustering phenomenon later, we calculate the clearance range r_c around transmitters and receivers to be $r_c = c^{-\alpha} \left(\frac{\beta}{m} \right)^{-\alpha} d$.

The assumption of equal transmission power is important for the protocol designs in Fig. 6.5 and 6.6. Our multistage contention protocol allows transmitters and receivers to
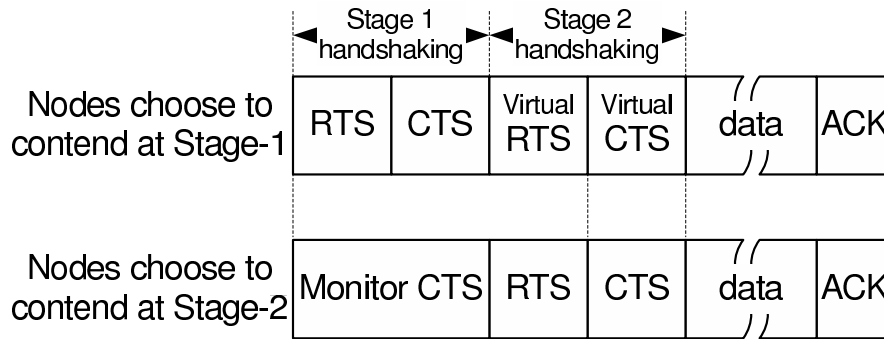
Figure 6.6: Timing diagrams of a two-stage contention MAC with the top for Stage 1 transmitter/receiver and the bottom for Stage 2 transmitter/receiver. Stage 1 successful contenders also participate in Stage 2 contention, which eliminates the need for confirmation RTS slots.

contend in the same fashion in terms of transmission power and back-off threshold. Given a clustering pattern of successful transmissions, this symmetry in contention induce clustering among not only receivers but transmitters, which can be observed on the right panel of Fig. 6.7. Thus transmissions in both directions, i.e., RTS/DATA and CTS/ACK, will succeed. Thus we can allow Stage 1 contenders to participate in Stage 2 contention and we need only a single ACK slot for all transmissions contending at different stages. We will see in the sequel that when transmission power levels are heterogeneous, extra considerations are needed for the protocol design.

As exhibited by the simulation in Fig. 6.7, nodes that survive Stage 1 only achieve 'weak' clustering while Stage 2 survivors are dense and clustered. One may naturally ask:

- How does Stage 1 (simple contention) realize initial weak clustering and interact with subsequent stage(s) to realize these gains?

- What is the capacity gain provided by multistage contention process and how should it be optimized?

In the sequel, we will use our dumbbell model to analyze the multistage contention process and answer these questions.

6.3.3   How multistage contention leads to clustering?

The basic intuition of inducing clustering with multistage contention is to let initial contenders in Stage 1 generate a spatial reuse pattern which serves as a seed for subsequent contenders to further enhance the clustering pattern. The back-off strategy based on RTS/CTS power level is visualized in Fig. 6.8 with our dumbbell model, i.e., subsequent transmitters
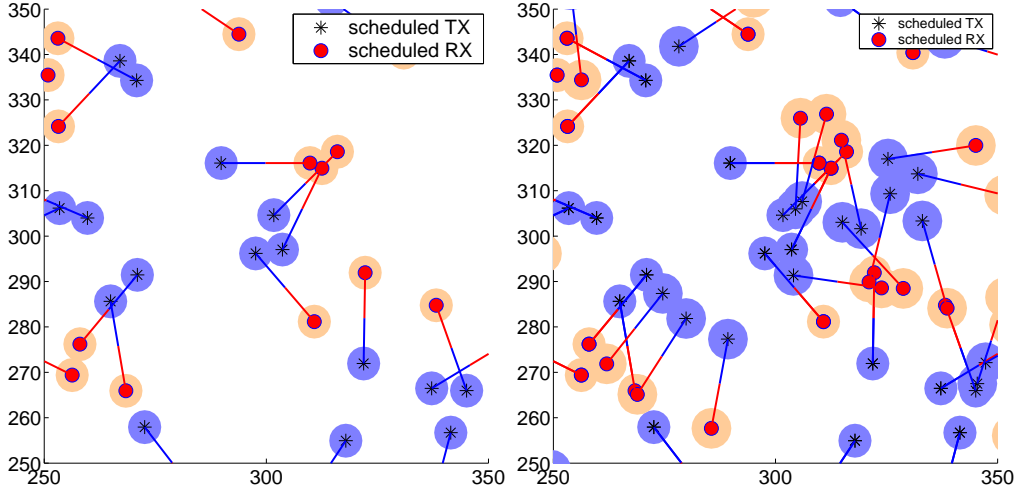
Figure 6.7: On the left, contention result of successful transmitter-receiver pairs, which serve as cluster 'seeds' for Stage 2. On the right the contention result after Stage 2, in which transmitters/receivers are indeed closely clustered with stage-1 transmitters/receivers and this significantly increases the overall clustering level.

or receivers will back-off if they have prohibited overlaps, e.g., E is too close to B and senses a strong CTS. On the other hand, subsequent transmitters/receivers clustering with existing transmitters/receivers, e.g., C is close to A, are likely to be able to contend and succeed. Thus with multistage contention, we spatially 'pack' subsequent transmissions by clustering them with existing transmissions scheduled in previous stages.

'Weak' clustering from Stage 1. To quantify the clustering effect after Stage 1, we will consider given a successful receiver at the origin, what is the intensity of other successful receivers around the origin after Stage 1. 'Clustering' means the intensity of successful receivers should be higher close to the origin. Let $\lambda^{(1)}$ and $\lambda^{(2)}$ denote the intensity of contending transmitters in Stage 1 and 2 respectively. Consider a receiver that succeeds Stage 1 and suppose it is located at the origin O. Since it was successful during Stage 1 it must have cleared a disc of radius r_c of transmitters around it. Now conditioning on this receiver contending transmitters *outside* the disc are still homogenously distributed with intensity $\lambda^{(1)}$. Specifically let us evaluate the intensity of successful receivers within the ball $B(O, r_c)$ centered at the origin. Fact 6.3.1 summarizes the results in this regard.

First denote $a(x)$ to be the area of $B(x, r_c) \backslash B(o, r_c)$ as shown on the left of Fig. 6.9, which is the interference region of a receiver located at x, uncovered by the interference
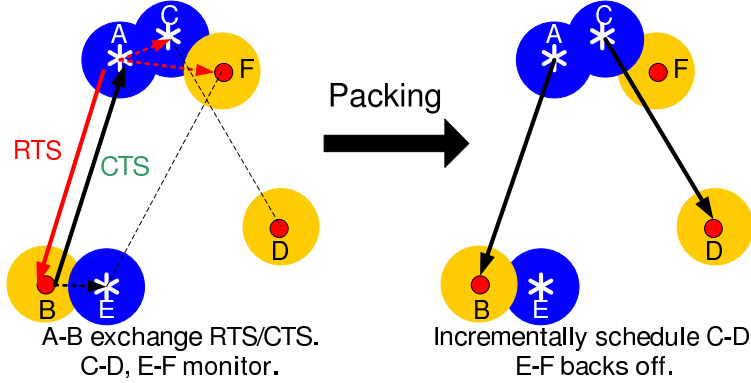
Figure 6.8: Multistage contention achieves clustering by spatial packing.

region of a given successful receiver at O. One can calculate

$$
\begin{aligned}
a(x) &= x^2\Big[\omega\sin\theta+\omega^2(\pi-\theta)+(\omega^2+1-2\omega\cos\theta)\times \\
&\quad\Big(\pi-\arccos\frac{1-\omega\cos\theta}{\sqrt{\omega^2+1-2\omega\cos\theta}}\Big)\Big]-\pi r_c^2,
\end{aligned}
$$

with $\omega=r_c/x$ and $\cos\theta=\frac{x}{2r_c}$.

Fact 6.3.1. *Consider Stage 1 contention of the multi-stage mechanism described above. Conditioning on a successful receiver at the origin O, the intensity of other successful receivers $\lambda_s^{(1)}(x)$ within the disc $B(O,r_c)$ at a distance x from the origin is roughly given (upper-bounded) by $\lambda_s^{(1)}(x)=\lambda^{(1)}(1-p_s(x))$, where $p_s(x)=1-e^{-\lambda^{(1)}a(x)}$ is the probability that a receiver a distance x, $0<x<r_c$, from the origin, is suppressed by one or more Stage 1 transmitters within the area $a(x)$.*

Proof. The above result is similar to that used to compute the outage lower bounds. Consider a receiver located a distance x from the origin. The receiver will survive Stage 1, if it has no transmitters within a ball of radius r_c of itself. As shown in Fig. 6.9 part of this ball has already been cleared of transmitters, since a successful Stage 1 receiver lies at the origin. Thus our candidate receiver will be successful if there are no transmitters within the region $B((x,0),r_c)\setminus B(O,r_c)$ whose area is given by $a(x)$, with the probability of this occurring given by 1 minus the probability that a homogenous Poisson point process places no points in a region of area $a(x)$. It then follows that the intensity of Stage 1 receivers within $B(O,r_c)$ which are sent RTS's by transmitters outside this disc is also homogenous and has intensity $\lambda^{(1)}$ as long as $r_c<\frac{d}{2}$. The clustering of successful receivers is because the success probability of a receiver $1-p_s(x)$ decreases sharply with distance to the origin, as shown on the right of Fig. 6.9. $\qquad\square$
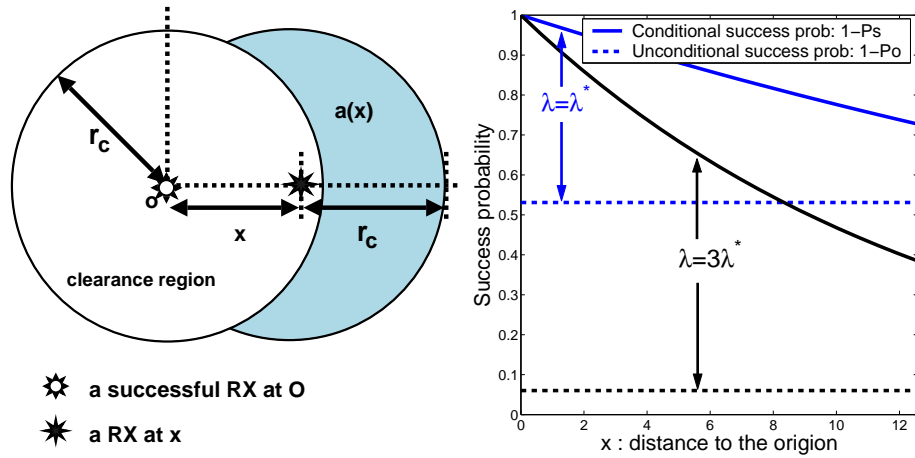
Figure 6.9: On the left the area $a(x)$ for obtaining outage lower bound conditioning on a successful receiver at O. On the right given the intensity λ of contending transmitters, the upper bounds for $1 - p_s(x)$, the success probability of a receiver at distance x conditioning on a successful receiver at O, and the success probability $1 - p_o(\lambda, d)$ without conditioning. For $p_o(\lambda, d)$ and λ^* see (4.11)(4.13).

Fact 6.3.1 shows that *given* a successful receiver at the origin, the probability $p_s(x)$ that another Stage 1 receiver at distance x away from it is successful decreases quickly with distance. Furthermore, the *conditional* intensity of other *successful* receivers within $B(O, r_c)$ forms a non-homogenous Poisson process with intensity $\lambda_s^{(1)}(x)$ – the graph on the right in Fig. 6.9 exhibits this decay in intensity as a function of x. This decay is more significant when the intensity is higher. Note that the above analysis was carried out by conditioning on a random event - 'a successful receiver at O', with probability $1 - p_o$ decreasing in $\lambda^{(1)}$, and thus the clustering in Stage 1 does not really increase significantly in $\lambda^{(1)}$.

Optimizing the additional clustering realized by Stage 2. An intuitive explanation for the substantial clustering after Stage 2 is that survivors of Stage 1 serve as seeds to Stage 2 by creating areas around transmitters and receivers where they suppress receivers and transmitters respectively, in order to avoid close-by interference and further enhance clustering in subsequent stages. In fact these results point to the robustness of the proposed two stage signaling. Because the first stage 'seeds' the Stage 2 clusters, the second stage is able to handle fairly high density of contenders achieving a high success rate. Overall, unlike Stage 1, the system capacity is not as sensitive to the intensity of Stage 2 contenders. Furthermore only transmitters which know they will not severely interfere a Stage 1 receiver would in fact attempt Stage 2 signaling.

Fact 6.3.1 can be used to approximately study the additional clustering induced

by Stage 2. After Stage 1 transmitter-receiver pairs are already scheduled in space, and will suppress nearby Stage 2 receivers and transmitters respectively. Thus, the process of Stage 2 transmitters which contend depends on the process of successful Stage 1 receivers and thus is no longer Poisson. Nevertheless successful Stage 2 receivers will most likely cluster around Stage 1 receivers. Considering such a receiver at the origin and *assuming* that the Stage 2 transmitters which actually contend correspond roughly to a homogenous Poisson process with intensity $\lambda^{(2)}$ outside $B(o, r_c)$, we can reuse the results in Fact 6.3.1. Specifically the expected number of successful Stage 2 receivers clustered within the ball $B(O, r_c)$ of a Stage 1 receiver is approximately

$$\int_o^{r_c} \lambda^{(2)} e^{-\lambda^{(2)} a(x)} 2\pi x dx. \tag{6.3}$$

To maximize (6.3), i.e., maximize spatial reuse, we solve for the optimal $\lambda^{(2)} \approx \frac{1.75}{2r_c^{2\alpha}}$, which is about $1.75\pi c^2$ times of the optimal contending intensity for Stage 1, see (4.13). Subsequently, the mean number of successful Stage 2 receivers around a successful Stage 1 receiver is roughly 0.93, i.e., we get roughly a 93% improvement on the capacity at the second stage and the expected number of successful receivers per cluster to be roughly 1.93. Our simulation results in Fig. 6.7 match this analysis quite well.

6.3.4 Handling multi-class or non-homogenous traffic and node distributions with multi-scale contention and clustering.

Why multi-scale contention with heterogeneous power is challenging?

A realistic network may support transmissions with different relay distances for at least two reasons. First, spatial intensity of nodes may be heterogenous and nodes may need to use different distances to maintain connectivity. Second, applications sharing the network may have different QoS requirements and possibly require different relay strategies, e.g., relaying on different spatial scales, see the discussion of 'spatial multiplexing' and Fig. 3.6 in Chapter 3. Third, the network may consist of heterogeneous devices with different transceiver capabilities. In this case nodes should use power control to choose transmit power levels corresponding to the desired relay distances. Note such power control geared at achieving performance or service differentiation is different from close-loop feedback power control compensating channel variations. It is initiated by users or applications. Moreover, such power control is not needed to ensure successful transmissions, which in fact has been sufficed to a MAC scheduling.

Inducing spatial clustering to achieve high spatial reuse faces additional difficulties in this context. First, monitoring nodes can not correctly infer what are the interference regions of contending nodes in the previous stage(s) given a mixture of heterogeneous transmissions with different power levels. This suggests transmissions with similar power
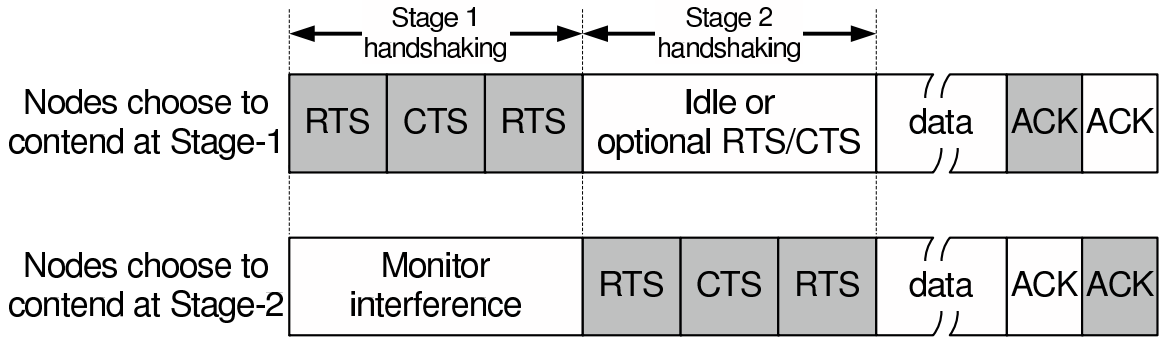
Figure 6.10: Multiscale multistage contention protocol requires separate ACK slots for each contention stage.

levels should contend together. Second, contention at different power level does not fully solve the hidden node problem. As previously discussed in Fig. 2.5, transmissions with low power and short range, e.g. $A \rightarrow B$, perform RTS/CTS handshaking before transmitting actual data, long range transmissions with strong power like $C \rightarrow D$ may severely interfere low power transmissions because C may not sense RTS/CTS from A or B. This suggests transmissions with higher power should contend before transmissions with lower power.

A multi-class multi-stage contention protocol.

One approach to deal with this problem is via a multi-class and multi-stage contention process. The basic idea is to allow transmissions with higher transmission power to perform handshaking first so as to enable transmitters and receivers in subsequent stages to detect their RTS/CTS and correctly estimate interference regions. Specifically, consider a network where nodes use one of k possible relay distances $d_i, i = 1, \ldots k$ satisfying $d_1 > d_2 \ldots > d_k$ each with an associated transmit power level $\rho_i^t, i = 1 \ldots k$. Suppose these power levels are known, and the associated ranges are such that typically the receive power are the same, e.g., $\rho_i^t = \rho d_i^\alpha$. Note that this is an idealized model and we will discuss imperfect power control vs. relay distance later in Section 6.4. In the sequel we refer to nodes which choose relay distance d_i to be of class i. Our new contention protocol is a variation on the multi-stage RTS/CTS/RTS process considered earlier. As shown in Fig. 6.10, we assume that only class i nodes perform handshaking at Stage i based on monitoring interference levels for stages $1, \ldots i - 1$ and thus inferring whether they will interfere with, or be interfered by, contenders in previous stages, by taking into account predefined power levels used at each stage.

The intuition for this choice is that by allowing only a given class to contend at a particular stage, nodes monitoring the process can obtain reasonable estimates of the

102

proximity of contenders based on a priori knowledge of their transmitting power levels and the received interference levels. Furthermore the ordering in which classes contend (based on transmission ranges) is enforced because if $i < j$ then the packing achieved by class i is likely to be less dense than that of class j. This ensures that large range transmissions are effectively packed prior to committing to short range ones. This is akin to packing large 'objects' first and then squeezing the smaller ones as appropriate within the gaps. This ordering also ensures a contender will hear the signaling of all relevant contenders with higher transmission power in previous stages, which solves the hidden terminal problem even under heterogeneous transmission power levels.

This approach, achieves a multi-scale clustering and high spatial reuse of successful transmissions. Fig. 6.11, exhibits a realization of a two-class scenario. Transmissions with long relay distances and larger interference regions (dumbbell size) are scheduled in Stage 1, while transmissions with shorter relay distance are scheduled in Stage 2. As can be seen, in addition to clusters of receivers for both classes of transmitters, additional fine scale clustering of short range transmission fill the void area remaining after Stage 1.
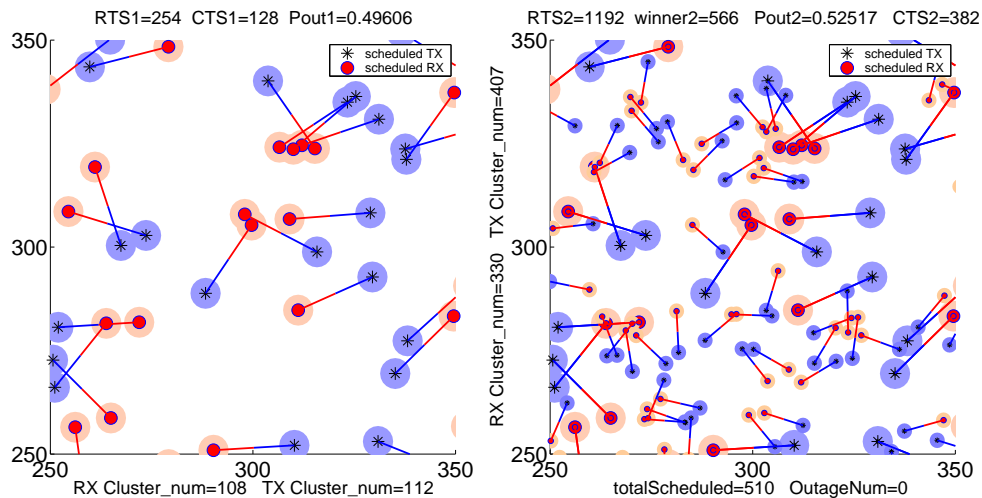

Figure 6.11: On the left, the resulting transmitter-receiver pairs of a multi-stage multi-class contention protocol's Stage 1 contention among nodes relaying delay sensitive traffic, i.e., using longer relay distances and thus having larger large interference ranges. On the right the resulting transmitter-receiver pairs for Stage 1 and 2 for a multi-class multi-stage contention protocol. Note how the shorter range transmissions cluster around Stage 1 receivers as well as independently in voids left during Stage 1.

By contrast with the case where nodes use homogeneous transmission power levels considered in last section, we can no longer allow all nodes to contend together as we did in Fig. 6.6. This is because the desired multiscale clustering pattern for nodes with

heterogeneous power levels is biased toward inducing receiver clusters. By clustering receivers along with proper power control corresponding to different transmission distances, i.e., $\rho^t = \rho \times d^\alpha$, receive signal power and interference power across nodes will be similar. Thus receivers will achieve similar SINR, as shown on the left of Fig. 6.12, and all successfully receive, assuming a fixed data rate for all nodes. However, for CTS/ACK signaling messages in the reverse direction, concurrent transmissions with power control lead to uneven SINR at the destinations, with nodes close by the receiver cluster overwhelmed by high power interference and suffering a very poor SINR, as shown on the right of Fig. 6.12. For this reason, we need separate ACK slots for each class.

The above problem can also negatively impact performance when there is certain heterogeneity among transmission power and distance within the same class, e.g., nodes in class i may use any power from 1mw to 10mw with corresponding transmission range 10m to 18m, assuming path loss exponent $\alpha = 4$. We can solve this problem by sending CTS/ACK at some maximal power level, e.g., 10mw for class i. Thus CTS/ACK will be received by transmitters with roughly equal SINR and succeed with high probability even if heterogeneity among transmission power exists. Indeed, here we mimic the power control strategy used by cellular networks with receiver clusters corresponding to base stations, transmitters corresponding to mobile terminals, RTS/DATA corresponding to reverse link using variable power and CTS/ACK corresponding to forward link using equal power. Simulations show that this leads to a 10% performance gain for the parameters mentioned above.

6.3.5 Performance evaluation of multi-stage contention and clustering MAC

We shall compare the performance of multistage contention (later referred as Packing) and random channel access or contention (later referred as Thinning) with two centralized schemes: centralized greedy algorithm (Centralized$_{greedy}$) and centralized random (Centralized$_{random}$) algorithm studied in Chapter. 4. In particular the Packing scheme has three stages, with the first two stages being identical to the two-stage version discussed in Section 6.3.2 and the last stage consists of retries by those who fail in the previous two stages.

We fix the path loss exponent to be 4 and assume all transmissions are of the same distance. If not specifically mentioned, the spreading factor is 512 and the SINR threshold required for successful transmission after de-spreading is 10dB. We fix the number of nodes in a rectangle area and randomize their locations for each round, for which different MAC schemes are applied to the same realization of nodes. Each performance point is an average of ten rounds. We also only consider nodes falling inside some margin to eliminate edge effects.

Our first simulation examines the spatial reuse achieved by different schemes given
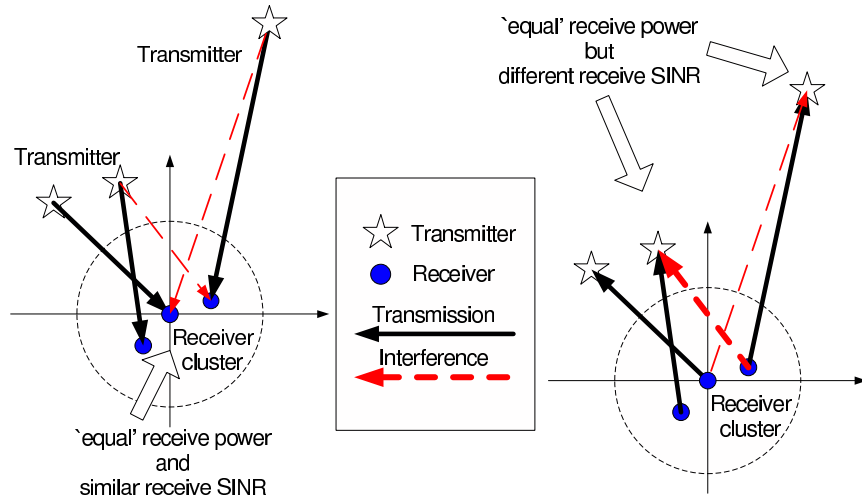
Figure 6.12: A typical multi-scale clustering pattern, in which receivers are likely clustered but transmitters may not. On the left, when receiving RTS/DATA, all receivers achieve similar receive signal power and SINR with power control. On the right, when receiving CTS/ACK, transmitters achieve uneven SINR if power control is in place.

the same set of intended transmissions. We also vary the density of contending transmissions and show how spatial reuse scales with the contention intensity. As shown in Fig. 6.13, we plot the number of successful transmissions achieved by Centralized$_{greedy}$, Centralized$_{rand}$ and Thinning. For Packing, we plot the *overall* successful transmissions achieved at the end of each stage. As expected, Centralized$_{greedy}$ has the best performance and Thinning has the worst performance. The performance of Packing is slightly lower than Centralized$_{greedy}$ but remains better than Centralized$_{rand}$, which is very impressive for Packing since Centralized$_{rand}$ is centralized. In this simulation, Packing is properly configured by choosing the right contention intensities at each stage, in particular, we let the Stage2 contention intensity be approximately twice of that in Stage1 according to the result from Section. 6.3.3. The performance of Packing almost remains increasing in the range of contention intensities tested. However, as discussed before, Thinning's performance is sensitive the contention intensity and there is some optimal contention intensity for Thinning to achieve the best performance, e.g., in Fig. 6.13 this happens when the normalized contention intensity is roughly 4. Finally, Stage2 achieves most of the performance gain, which indicates that our multistage protocol can be implemented with only two or three stages without compromising potential performance.

Our second simulation tests the robustness of Packing by intentionally assigning suboptimal contention intensities at each stage. In particular, we let the contention intensity at Stage1 be twice that of Stage2, i.e., we have too high contention intensity initially and in-

sufficient contention intensity at Stage2. As shown in Fig. 6.14, the performance of Packing is only slightly worse than the previous simulation when parameters are optimally chosen and remains increasing or flat throughout the range of intensities tested. It is only when the overall intensity is extremely high, that its performance starts decreasing as Thinning. We can also observe that Stage3 contributes more significantly to the overall spatial reuse when the Packing is not optimally configured. Therefore, Packing is quite robust in performance thanks to a multistage implementation.

Finally we examine the scaling of spatial reuse in spreading factor m for different schemes. Recall that in Section. 6.1 we show that optimal scheme can achieve a spatial reuse linear in m and in Chapter. 4 we show that thinning can only achieve one which is sub-linear in m, i.e., $\Theta(m^{\frac{2}{\alpha}})$, for both low-outage and high-outage regimes. As shown in Fig. 6.15, Centralized$_{greedy}$ and Packing are both efficient because not only their spatial reuse scales roughly linearly in m but also much faster than Centralized$_{rand}$ and Thinning, whose spatial reuse scales only sub-linearly and relatively slow in m. Therefore, Packing is well suited as the choice for spread spectrum ad hoc networks.
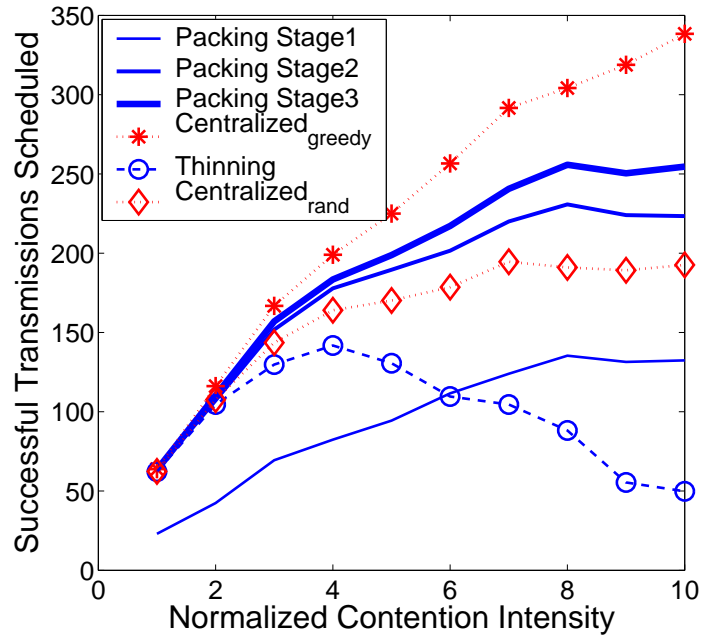

Figure 6.13: Performance of the Packing approach surpasses Centralized$_{rand}$ when contention intensity at each stage is optimally chosen.
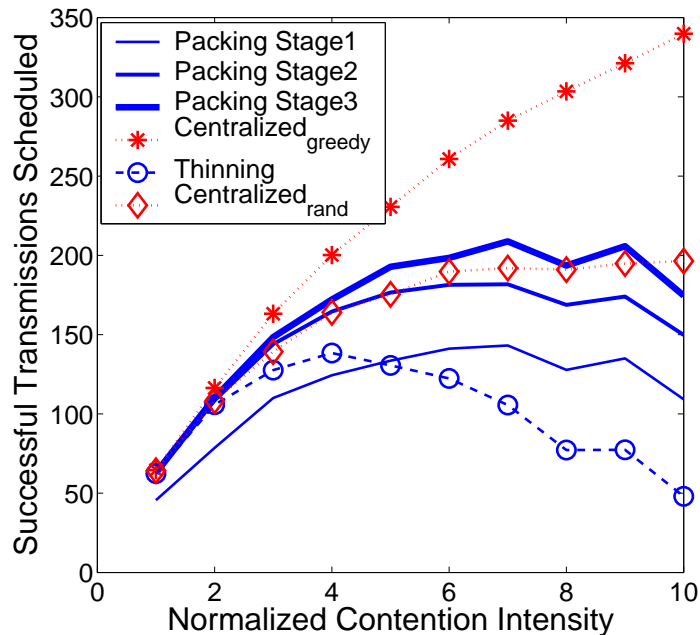
Figure 6.14: Performance of Packing is robust when contention intensity at each stage is not optimally chosen.

6.3.6 Summary - multi-stage contention and clustering.

First, when the contention parameters are properly configured, multistage contention can achieve close-to optimal capacity with 2 or 3 stages. Second, the associated overhead for each successful transmission, is fairly close to the simple RTS/CTS mechanism with additional overhead to monitor local interference levels. Finally the previous clustering and passive monitoring of the contention process benefit the next round of contention reducing the sequential failure rate of contention significantly, e.g., by 65% relative to homogenous Poisson point process in our simulations. For example, assuming nodes switch between transmit and receive modes, successful transmitters in a given round tend to cluster, making themselves likely to be successful in the next round as receivers.

6.4 Practical design and implementation considerations

The previous section suggests a mechanism for inducing spatial clustering on multiple scales has some significant benefits in terms of meeting QoS requirements and enabling power savings while achieving efficient spatial reuse. This section gives some thoughts on design and implementation considerations. Due to space constraints, we will only briefly introduce some of the problems.
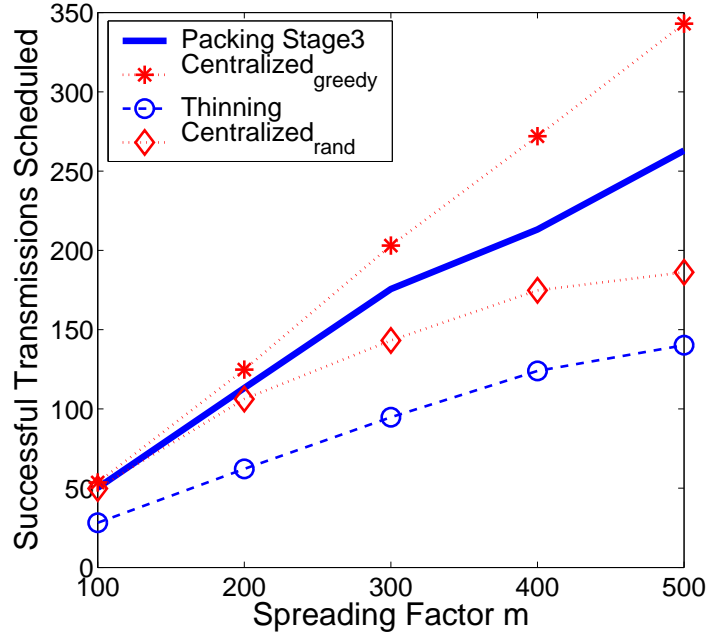
Figure 6.15: The scaling of performance of different approaches in spreading factor *m*.

Synchronization. Throughout this dissertation, we have assumed a slotted system with synchronous contention and transmission. We believe synchronization is critical for efficient MAC scheduling in ad hoc networks. First, data transmission and ACK are well protected after handshaking, which eliminates the need for maintaining states, e.g., NAVs in 802.11 and [30]. Second, synchronous contention provides better priority access and thus better QoS support than asynchronous implementations[44]. Although synchronization incurs extra overheads, such as inter-frame spacing, similar MAC inefficiency also exist for asynchronous contention resolution in which the required carrier sensing usually causes conservative back-off both spatially and temporally. Therefore benefits of synchronization will warrant these overheads.

Spreading factor. Our clustering approach requires a system with a large spreading factor, i.e., $\frac{\beta}{m} \ll 1$. The performance gains of clustering are more significant when the spreading factor *m* is large. Note that we conservatively assume correlation among quasi-orthogonal codes is $\frac{1}{m}$ while some in the literature use $\frac{1}{3m}$, which would allow clustering to work well even for systems with relatively moderate spreading.

Transmission power and range. To be effective, our multi-stage contention MAC requires pre-defined power levels. This is actually a realistic model since real devices typically only do discrete power control. Yet the transmission range can continuously change rather than fixed per class as assumed in this dissertation. This leads to a variable interfer-

ence range r even at the same transmitting power level. Thus a back-off criterion should always conservatively assume monitored transmissions are of the maximal transmission range allowed by the power levels associated with each class. This causes a penalty in the capacity, but only marginal in particular when path loss is large.

Node spatial distribution. We have assumed homogeneous Poisson point process throughout this dissertation. What if the spatial density is not homogeneous? For the multi-stage contention, existing clustering patterns may indeed help the performance and this scheme is very robust for inducing clustering among contenders for different scenarios. By contrast, the virtual-grid scheme, works well in a 'dense' network, i.e., when there are many nodes within each virtual cell of size d^2. In a 'dense' network, our virtual-grid scheme is robust to fluctuations of spatial density.

Routing protocol. The importance of routing has been mentioned in previous items. In particular, the challenge in the new design paradigm proposed in this dissertation is to enable long relay distance. Maintaining the information of all the nodes in a large neighborhood may not be feasible. Thus the actual design may try to discover preliminary routes consisting small hops and associated power budget. Then a source node may recompute a new route by skipping some relay points on preliminary routes. At the same time, using spread spectrum results in a reduction in the bandwidth per channel, while enabling multiple quasi orthogonal channels. Thus a routing algorithm for such networks should carefully spatially balance traffic loads across available nodes (channels) to achieve efficient operation.

Chapter 7

Conclusion

In ad hoc network design, one must consider tradeoffs among different key performance metrics such as capacity, QoS, energy efficiency and system complexity, etc. A CDMA physical layer allows for flexible tradeoffs among different performance metrics. However, to achieve these benefits, network designs at different layers need to be changed accordingly to fully leverage the capability of CDMA. While simple designs akin to those used in narrow-band systems work well for FH-CDMA systems, they fail to achieve good performance in DS-CDMA systems, in particular in a dense network with heavy load. We therefore propose practical design approaches that can be applied to improve network capacity without compromising or even enhancing other performance metrics across different network layers, e.g., SIC at the physical layer, inducing clustering at the MAC layer, and virtual grid clustering at the network layer using GPS aided routing. We show significant improvements on performance of DS-CDMA systems can be achieved while maintaining distributed management and low complexity in the network.

Bibliography

[1] J. G. Andrews. Interference cancellation for cellular systems: A contemporary overview. *IEEE Communication Magazine*, 12(2):19–29, Apr. 2005.

[2] J. G. Andrews and T. Meng. Optimum power control for successive interference cancellation with imperfect channel estimation. *IEEE Trans. Wireless Communications*, 2(2):375–83, Mar. 2003.

[3] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler. A spatial reuse aloha mac protocol for multihop wireless mobile networks. In *Proc. Annual ALLERTON Conference on Communication.*, 2003.

[4] S. Bandyopadhyay and E. Coyle. An energy efficient hierarchical clustering algorithm for wireless sensor networks. In *Proc. IEEE INFOCOM*.

[5] R. M. Buehrer. Equal BER performance in linear successive interference cancellation for CDMA systems. 49(7):1250–58, July 2001.

[6] A. B. Carleial. A case where interference does not reduce capacity. *IEEE Trans. Inform. Theory*, 21(5):569–70, Sept. 1975.

[7] S. Cass. Viva mesh vegas: The gambling capital antes up for a new mobile broadband technology. *IEEE Spectrum*, 2005.

[8] T. Cover. Broadcast channels. *IEEE Trans. Inform. Theory*, 18(1):2–14, 1972.

[9] T. Cover and J. Thomas. *Elements of information theory*. John Wiley & Sons, INC, 1991.

[10] T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad-hoc networks. In *Proc. IEEE INFOCOM*, June 2002.

[11] A. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proc. IEEE INFOCOM*, March 2004.

[12] J. Garcia-Luna-Aceves and J. Raju. Distributed assignment of codes for multihop packet-radio networks. In *Proc. IEEE MILCOM*, volume 1, 1997.

[13] S. Ghez, S. Verdú, and S. Schwartz. Stability properties of slotted ALOHA with multi-packet reception capability. *IEEE Transactions on Automatic Control*, 33(7):640–649, July 1988.

[14] K. S. Gilhousen and et. al. On the capacity of a cellular CDMA system. *IEEE Trans. Vehicular Technology*, 40(2):303–12, May 1991.

[15] M. Grossglauser and D. Tse. Mobility increases the capacity of ad-hoc wireless networks. *IEEE Transactions on Networking*, 10(4):477–486, August 2002.

[16] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[17] A. Hasan and J. Andrews. Cancellation Error Statistics in a CDMA System using Successive Interference Cancellation. In *Proc. Allerton Conf. on Communications, Cont., and Comp*, Aug 2003.

[18] L. Hu. Distributed Code Assignments for CDMA Packet Radio Networks. *IEEE/ACM Trans. on Networking*, 1(6):668–77, Dec. 1993.

[19] J. Hui. Throughput analysis for code division multiple accessing of the spread spectrum channel. SAC-2(4):482–486, July 1984.

[20] M. Joa-Ng and I. Lu. Spread spectrum medium access protocol with collision avoidance in mobile ad-hoc wireless network. In *Proc. IEEE INFOCOM*, March 1999.

[21] A. Jovicic, P. Viswanath, and S. Kulkarni. Upper boundsto transport capacity of wireless networks. *IEEE Trans. Inform. Theory*, 50.

[22] R. Jurdak, C. Lopes, and P. Baldi. A survey, classification and comparative analysis of medium access control protocols for ad hoc networks. *IEEE Communications Surveys and Tutorials.*, 6(1):2–16, First Quarter 2004.

[23] V. Kawadia and P. Kumar. Power control and clustering in ad hoc networks. In *Proc. IEEE INFOCOM*.

[24] L. Kleinrock and J. Silvester. Spatial reuse in multihop packet radio networks. *Proceedings of the IEEE*, 751(1):156–167, January 1987.

[25] E. Knightly. Wireless community mesh networks - hype or the next big frontier? In *ACM MobiCom Panel*, Philadelphia, PA, 2004.

[26] S. Kumar, V. S. Raghavan, and J. Deng. Medium access control protocols for ad-hoc wireless networks: A survey. *Elsevier Ad Hoc Networks Journal*, to appear.

[27] T. Kwon and M. Gerla. Clustering with power control. In *Proceedings of IEEE MILCOM*, 1999.

[28] C. Lin and M. Gerla. Adaptive clustering for mobile wireless networks. *IEEE Journal on Selected Areas in Communications*, 15(7):1265–1275, September 1997.

[29] S. Lowen and M. Teich. Power law shot noise. *IEEE Trans. Inform. Theory*, 36:1302–1318, November 1990.

[30] A. Muqattash and M. Krunz. CDMA-based MAC protocol for wireless ad hoc networks. In *Proc. ACM MOBIHOC*, June 2003.

[31] A. Muqattash, M. Krunz, and W. E. Ryan. Solving the Near Far Problem in CDMA-based Ad Hoc Networks. *Ad Hoc Networks Journal*, 1(4):435–53, Nov. 2003.

[32] M. Pursley. The role of spread spectrum in packet radio networks. *Proceedings of the IEEE*, 75(1):116–134, 1987.

[33] M. Pursley and D. Taipale. Error probabilities for spread-spectrum packet radio with convolutional codes and viterbi decoding. 35(1):1–12, Jan. 1987.

[34] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice-Hall, Upper Saddle River, New Jersey, second edition, 2002.

[35] B. Rimoldi and R. Urbanke. A rate splitting approach to the gaussian multiple access channel. *IEEE Trans. Inform. Theory*, 42(2):364–75, Mar. 1996.

[36] R. Rozovsky and P. Kumar. Seedex: A mac protocol for ad hoc networks. In *Proc. ACM MOBIHOC*, October 2001.

[37] T. Shepard. Decentralized channel management in scalable multihop spread-spectrum packet radio networks. Technical Report MIT/LCS/TR-670, MIT, 1995.

[38] T. Shepard. A channel access scheme for large dense packet radio networks. In *Proc. ACM SIGCOMM*, pages 219–230, August 1996.

[39] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt. *Spread Spectrum Communications Handbook*. Mc-Graw Hill, revised edition edition, 1994.

[40] E. Sousa and J. Silvester. Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network. *IEEE Journal on Selected Areas in Communications*, 8(5):762–771, June 1990.

[41] E. S. Sousa and J. Silvester. Spreading Code Protocols for Distributed Spread-Spectrum Packet Radio Networks. *IEEE Trans. Communications*, 36(34):272–81, Mar. 1988.

[42] M. Stemm and R. H. Katz. Measuring and reducing energy consumption of network interfaces in hand–held devices. *IEICE Trans. on Commun. – Special Issue on Mobile Computing*, E80–B(8):1290–1302, 1997.

[43] J. Stine. Energy conserving protocols for wireless data networks. Technical Report Ph.D. Dissertation, University of Texas at Austin, 2001.

[44] J. Stine and G. de Veciana. A paradigm for quality of service in wireless ad hoc networks using synchronous signaling and node states. *IEEE JSAC Special Issue on Quality of Service Delivery in Variable Topology Networks*, 22(7):1301–1321, September 2004.

[45] D. Stoyan, W. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications, 2nd Edition*. John Wiley and Sons, 1996.

[46] D. Stoyan and H. Stoyan. *Fractals, Random Shapes, and Point Fields*. John Wiley and Sons, 1994.

[47] S. Toumpis and A. J. Goldsmith. Capacity regions for wireless ad hoc networks. *IEEE Trans. Wireless Communications*, 24(5):736–48, May 2003.

[48] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, To Appear.

[49] S. Verdu. *Multiuser Detection*. Cambridge, Cambridge, UK, 1998.

[50] A. J. Viterbi. Spread-spectrum communications – myths and realities. In *IEEE Communication Magazine*, volume 4, May 1979.

[51] A. J. Viterbi. Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels. *IEEE J. Selected Areas Commun.*, 8:641–9, May 1990.

[52] X. Wang and H. Poor. *Wireless Communication Systems: Advanced Techniques for Signal Reception*. Prentice-Hall, 2003.

[53] D. Warrier and U. Madhow. On the capacity of cellular CDMA with successive decoding and controlled power disparities. In *Proc., IEEE Veh. Technology Conf.*, pages 1873–7, May 1998.

[54] 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. *submitted to IEEE Transactions on Information Theory*.

[55] L.-L. Xie and P. R. Kumar. A network information theory for wireless communication: Scaling laws and optimal operation. *IEEE Trans. Inform. Theory*, 50(5):748–767, May 2004.

[56] F. Xue, L.-L. Xie, and P. R. Kumar. The transport capacity of wireless networks over fading channels. *IEEE Trans. Inform. Theory*, 51(3):834–847, 2005.

[57] O. Younis and S. Fahmy. Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach. In *Proc. IEEE INFOCOM*.

Vita

Xiangying Yang was born in SiChuan Province China, on November 18, 1975, the son of Chunzheng Yang and Zhuozi Gao. After completing his work at Tanjin the 1st High School, he entered Tsinghua University in Beijing, China, where he received his B.S. in electrical engineering, in July 1998. He continued his graduate study in the Unite States since 1998. He received M.S. and Ph.D. both in electrical and computer engineering (ECE) from University of Texas at Austin in 2000 and 2005. He is a recipient of Texas Telecommunications Engineering Consortium (TxTEC) Fellowship and a member of Tau Beta Pi. His research interests include web caching, peer-to-peer applications and cross-layer design of wireless ad hoc networks. He is currently working in the Broadband Wireless Division at Intel Corp, Portland OR.

Permanent Address: 288 Nanjing Rd
China Institute of Hematology Science
Tianjin 300020
China

This dissertation was typeset with LaTeX 2_ε by the author.