# Schemes and Architectures for Wireless Ad Hoc Networks and Cooperative Communication

## THÈSE N° 3760 (2007)

PRÉSENTÉE À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
Institut de systèmes de communication
Section systèmes de communication

## ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

Par


**Tarik Tabet**

Ingénieur en systèmes de communication Diplômé de l'EPFL,
DEA Signal et Communication de l'Université de Nice Sophia-Antipolis,
de nationalité marocaine


acceptée sur proposition du jury:

Prof. Raymond Knopp, directeur de thèse
Pr. Bixio Rimoldi, codirecteur de thèse
Prof. Emre Telatar, président du jury
Pr. Giuseppe Caire, rapporteur
Pr. Mérouane Debbah, rapporteur
Pr. Suhas Diggavi, rapporteur

Lausanne, EPFL Mars 2007

# Schemes and Architectures for Wireless Ad Hoc Networks and Cooperative Communication

Tarik Tabet

# Abstract

The focus of this thesis is on the study of decentralized wireless multi-hop networks. We are particularly interested in establishing bounds on the traffic-carrying capabilities of wireless ad hoc networks and conditions on the scalability of such networks with node mobility. This theoretical investigation brings forward challenges on the design of such networks. This leads to a second part of this thesis that considers the feasibility and the design of physical layer architectures and schemes for decentralized wireless multi-hop networks.

In the first part of this thesis, bounds on the capacity of wireless ad hoc networks with two types of non-uniform traffic patterns are established. We focus on the impact of traffic patterns where local communications predominate and show the improvement in terms of per user-capacity over ad hoc networks with unbounded average communication distances. We then study the capacity of hybrid wireless networks, where long-distance relaying is performed by a fixed overlay network of base-stations. We investigate the scaling of capacity versus the number of nodes and the density of base-stations in the area of the network. It is shown that the gain in performance is mainly due to the reduction in the mean number of hops from source to destination.

Then, we investigate the impact of mobility on the ad hoc network capacity. We propose a set of necessary and sufficient conditions under which the long-term averaged throughput in an ad hoc network can remain constant as the number of nodes increases. The main idea is to use a *connectivity graph*, that does not represent the actual physical network, but rather the available communication resources. This graph also allows to translate the problem of maximizing the throughput in ad hoc networks to the multi-commodity flow problem and directly apply related results.

In constrast to these macroscopic studies, in the second part we focus

on a microscopic analysis of ad hoc wireless networks. We are interested in characterizing the performance of decentralized multiple-access and retransmission schemes for multi-hop wireless networks with the goal of drawing conclusions on cross-layer design. We investigate different transmission strategies in order to assess the tradeoff between spatial density of communications and the range of each transmission. We present tools for characterizing the spatial throughput as a function of topological parameters (e.g node population density) and system parameters (propagation, bandwith etc). The results of this work also show that coding and retransmissions provide means of reliable communication coupled with a completely decentralized multiple-access strategy.

Finally, an efficient protocol for the delay-limited fading Automatic Retransmission reQuest (ARQ) single relay channel is considered for cooperative communications. The proposed protocol exploits two kinds of diversity: (i) space diversity available through the cooperative (relay) terminal, which retransmits the source's signals, (ii) ARQ diversity obtained by leveraging the retransmission delay to enhance the reliability. The performance characterization is in terms of the achievable diversity, multiplexing gain and delay tradeoff for a high signal-to-noise ratio (SNR) regime. Then, by letting the source's power level vary over the retransmission rounds, we show the benefits of power control on the diversity.

## Keywords

*Foundations of sensor and ad hoc wireless networks, traffic patterns, connectivity graph, mobility, multi-commodity flow, ARQ, incremental redundancy, multi-hop communications, physical layer issues in ad hoc networks, multiple access techniques, cross-layer design, Poisson Point process, multi-hop routing, cooperative diversity, relay channel, diversity-multiplexing-delay tradeoff, power control.*

# Résumé

Le but de cette thèse est l'étude des réseaux sans fil décentralisés multi-sauts. Nous nous intéressons en particulier à l'établissement des limites sur les possibilités de trafic des réseaux Ad Hoc sans fil, et des conditions sur le comportement et l'échelonnage de tels réseaux avec la mobilité des noeuds. Cette recherche théorique présente des défis quant à la conception de tels réseaux. Ceci nous mène à la deuxième partie de cette thèse qui aborde la faisabilité et la conception des architectures et des arrangements de couche physique des réseaux sans fil décentralisés multi-sauts.

Dans la première partie de cette thèse, les limites sur la capacité des réseaux Ad Hoc sans fil avec deux types de modèles de trafic non-uniformes sont établies. Nous nous concentrons sur l'impact des modèles de trafic, là où la communication locale prédomine et montrons l'amélioration en termes de capacité par utilisateur sur les réseaux Ad Hoc avec des distances de communication moyennes illimitées. Ensuite, Nous étudions la capacité de réseaux hybrides sans fil, là où le relayage est effectué par un réseau fixe de recouvrement des stations de base. Nous étudions l'échelonnage de la capacité en fonction du nombre de noeuds et de la densité des stations de base. Il est démontré que le gain dans la performance est principalement dû à la réduction du nombre moyen de sauts de la source à la destination.

Ensuite, nous étudions l'impact de la mobilité sur la capacité du réseau Ad Hoc. Nous proposons un ensemble de conditions nécessaires et suffisantes, sous lesquelles le débit dans un réseau Ad Hoc, peut rester constant à mesure que le nombre de noeuds augmente. L'idée principale est d'employer un *graphe de connectivité*, qui ne représente pas le réseau physique réel, mais plutôt les ressources disponibles de communication. Ce graphe permet également de transformer le problème consistant à maximiser le débit dans les réseaux Ad Hoc en problème polyvalent de multi-flots, et d'appliquer directement des résultats relatifs.

En contraste avec ces études macroscopiques, nous nous concentrons dans la deuxième partie sur une analyse microscopique des réseaux Ad Hoc sans fil. Nous caractérisons la performance des schémas multi-accès décentralisés et de retransmission pour les réseaux sans fil de multi-sauts, avec l'objectif de tirer des conclusions sur la conception multi-couches. Nous étudions différentes stratégies de retransmission afin d'évaluer l'interaction entre la densité spatiale de communication et la portée de chaque transmission. Nous présentons des outils pour caractériser le débit spatiale en fonction de paramètres topologiques (par exemple densité de population par noeud) et de paramètres de système (propagation, bande de fréquence etc). Les résultats de ce travail montrent également que le codage et les retransmissions fournissent des moyens de communication fiable couplés avec une stratégie de multiple-accès complètement décentralisée.

Enfin, un protocole efficace pour le canal relais semi-duplex ARQ est envisagé pour des communications coopératives. Le protocole proposé exploite deux sortes de diversité : (i) diversité spatiale disponible par la borne (de relais) coopérative, qui retransmet les signaux depuis la source, (ii) la diversité ARQ obtenue par les retransmissions ARQ afin d'augmenter la fiabilité. La caractérisation de la performance est en termes de la diversité réalisable, du gain de multiplexage et du délai pour un haut régime de rapport signal/bruit (SNR). Puis, en laissant le niveau de puissance de la source changer pour chaque retransmission, nous montrons les avantages du contrôle de la puissance et son impact sur la diversité.

## Mots-clés

*Fondements des réseaux sans fil ad hoc et sensor, modèles de trafic, graphe de connectivité, mobilité, multi-flots, ARQ, incremental redundancy, communications multi-sauts, couche physique, techniques d'accès multiple, conception multi-couches, processus Point Poisson, routage multi-saut, diversité cooperative, canal relais, diversité-multiplexage-delai compromis, contrôle de puissance.*

# Contents

# List of Figures

# List of Tables

# List of Acronyms

| | |
|---|---|
| ABR | Associativity-Based Routing |
| ACK | Acknowledgment |
| AF | Amplify and Forward |
| ARQ | Automatic Retransmission reQuest |
| BPCU | Bit Per Channel Use |
| BTS | Base-Station |
| CDMA | Code Division Multiple Access |
| CSI | Channel State Information |
| CSMA | Carrier Sense Medium Access |
| CSMA/CA | Carrier Sense Medium Access with Collision Avoidance |
| DDF | Dynamic Decode and Forward |
| EGPRS | Enhanced General Packet Radio Service |
| ESS | Extended Service Set |
| MAC | Medium-Access Control |
| MACA | Multiple Access with Collision Avoidance |
| MACAW | Multiple Access with Collision Avoidance for Wireless |
| MANET | Mobile Ad Hoc Network |
| MGF | Moment Generating Function |
| MIMO | Multiple Input Multiple Output |
| MISO | Multiple Input Single Output |
| ML | Maximum Likelihood |
| NACK | Negative Acknowledgment |
| OFDM | Orthogonal Frequency Division Multiple |
| OFDMA | Orthogonal Frequency Division Multiple Access |
| PHY | Physical Layer |
| RCPC | Rate Compatible Convolutional Codes |
| RF | Radio Frequency |
| S-BTS | Source Base-Station |

| | |
|---|---|
| S-C | Source-Collector |
| S-D | Source-Destination |
| SINR | Signal to Interference Noise Ratio |
| SISO | Single Input Single Output |
| SNR | Signal to Noise Ratio |
| TCP | Transmission Control Protocol |
| TDMA | Time Division Multiple Access |
| TG | Task Group |
| WLAN | Wireless Local Area Network |
| WMN | Wireless Mesh Network |

# Acknowledgments

I would like to thank my supervisors Prof. Raymond Knopp and Prof. Bixio Rimoldi for their guidance an support throughout my Ph.D. studies. I am very thankful to Raymond for the freedom he gave me in pursuing my own course of research, and, at the same time, providing me with strong support and friendship. His creativity, deep knowledge of theoretical and practical aspects of communication systems was a source of continuous learning and of inspiration. I am also very grateful to Bixio, whose door was always open, for teaching me the path one should take during his doctoral studies, asking the right questions and keeping in mind the big picture.

Many thanks to Prof. Christina Fragouli and my colleague and friend Sanket Dusad, with who I enjoyed collaborating during my Ph.D work.

It was a great pleasure and honor for me to have, on my thesis jury, Prof. Giueseppe Caire, Prof. Mérouane Debbah, Prof. Suhas Diggavi and Prof. Emre Telatar. I would like to thank them for accepting to read and review the first version of this thesis.

A team of english speakers, latex specialists and good friends contributed in improving the quality of this dissertation. I very much appreciated the help of Dinkar, Jamal, Marius, Peter, Sanket, Sibi, Shrinivas and Vishwambhar. The help of my friends Djalel, Menouer, Moushab, Naoufel and Osmane in the translation of the abstract was valuable. I also wish to thank all my colleagues from LCM, LTHC, LTHI and LICOS for their friendship and also the past members of LCM for their warm welcome especially Rajesh Manakkal and Nicolae Chiurtu.

My stay in Lausanne during my studies was made very enjoyable by my many friends. In particular special thanks to my arabic friends which I am not going to mention in order not to forget anyone of them. The moments we spent all together will be cherished and memorable.

Un grand merci to Françoise Behn for her excellent administration of the lab. She was always supporting and helping to the point I considered her as my third supervisor.

Last but not least, I would like to thank my family for their love, support and constant encouragement, this thesis is dedicated to them. Cette thèse n'aurait jamais vu le jour sans la confiance et le soutien de mes parents. Je ne saurais comment exprimer ma gratitude à mon père à qui je dois tout dans ma vie. Il a toujours su nourrir et enrichir mes facultés analytiques et mon esprit critique par ses observations si précieuses et par nos *monakachat* interminables. Il a su me donner goût aux sciences et aux mathématiques par des petits problèmes et jeux dignes de l'époque des califes de Bagdad. Tu aurais aimé que je sois Docteur en médecine, je te dédie ce titre de Docteur ès sciences. De même ces quelques lignes ne suffiront pas pour remercier ma mère, elle qui a passé des nuits et des nuits blanches à prier pour moi. Sa foi, sa sagesse, ses remarques pertinentes m'ont toujours porté vers l'avant et m'ont aidé à surmonter les difficultés de tous les jours. Je voudrais également remercier mes deux frères pour avoir toujours cru en moi. Enfin, je remercie ma femme pour ses encouragements et son aide si précieuse et surtout sa patience: je sais que ce n'est pas facile de supporter les changements d'humeur d'un thésard en fin de thèse.

# Chapter 1

# Introduction

## 1.1 Overview

There have been tremendous advances in wireless communications in recent years, including in wireless radios, networks, and mobile devices. The existing network infrastructure of planned cellular networks is being complemented by heterogeneous self-organizing systems with hybrid infrastructure and peer-to-peer communication modes. It aims at overcoming inherent limitations to achieve unprecedented coverage, throughput, flexibility, and cost-efficiency.

Ad hoc wireless networks is currently a hot research area due to the increase of the need for connectivity "anywhere", "anytime" and, in particular, "anyhow" (with and without a fixed infrastructure) [1]. Currently, most wireless networks require base-stations to operate; the most common examples are cellular telephony and IEEE 802.11 networks (infrastructure mode as depicted in Fig.1.1)[2]. While these traditional networks concentrate on single-hop communications (nodes communicating to fixed infrastructure), ad hoc wireless networks (see Fig.1.2) can be described as multi-hop wireless networks, where there is not a central or dominant node. All the nodes are at the same hierarchical level. Wireless ad hoc networks are formed by a set of hosts that communicate with each other over a wireless channel. Each node has the ability to communicate directly with another node in its physical neighborhood. They operate in a self-organized and decentralized manner and message communication takes place via multi-hop spreading. A packet is sent to its target node through a set of intermediate nodes that act as routers. Multi-hopping replaces therefore in theory the need for fixed in-

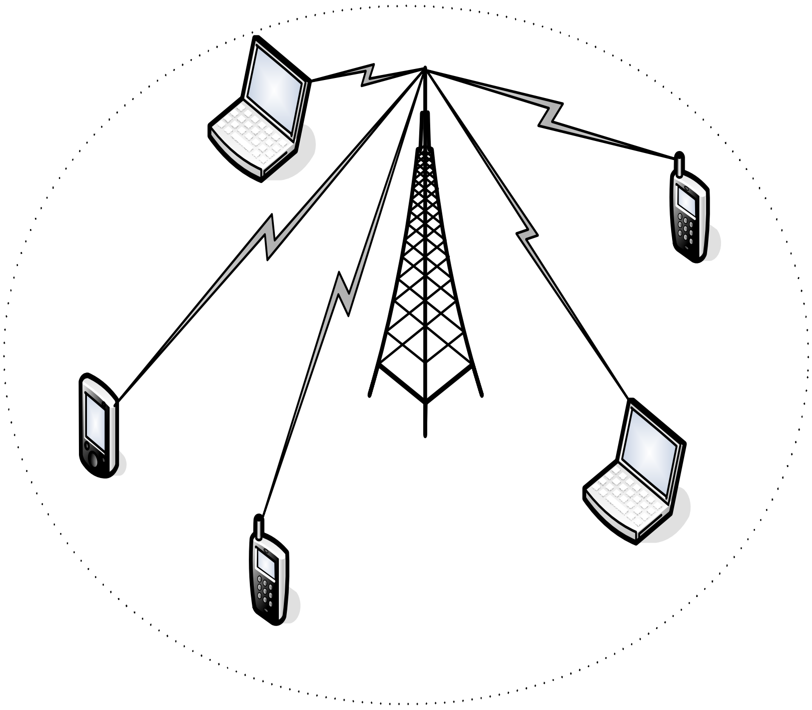

**Figure 1.1:** Infrastructure-based wireless network.

frastructure, since the network is only made of similar devices collaborating to provide connectivity, by acting as terminals and as relays simultaneously, and all the nodes have the same capabilities and the same responsibilities [3].

Ad hoc networks are suited for use in situations where an infrastructure is unavailable or to deploy one is not cost effective. One of the many possible uses of mobile ad hoc networks is in some business environments, where the need for collaborative computing might be more important outside the office environment than inside. An example application could be a business meeting outside the office to brief clients on a given assignment. A mobile ad hoc network (MANET) [4] can also be used to provide crisis management services [57], such as in disaster recovery, where the entire communication infrastructure is destroyed and resorting communication quickly is crucial. By using a mobile ad hoc network, an infrastructure could be set up in hours instead of weeks, as is required in the case of wireline communication.

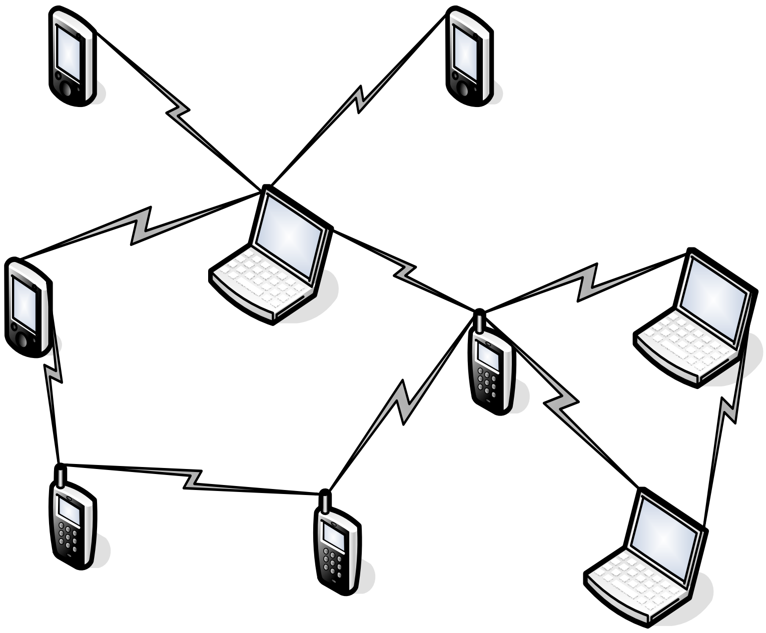

**Figure 1.2:** Ad hoc wireless network.

Another application example of a mobile ad hoc network is Bluetooth [5], which is designed to support a personal area network by eliminating the need of wires between various devices, such as printers and personal digital assistants. Internet Service Providers are also seeking multi-hop networks (wireless mesh networks [7]) solutions to implement public internet access which could simultaneously target the markets of residential, business and travel. The goal is to provide broadband wireless connectivity solutions to both indoor and outdoor communications in urban or rural environments without the need for extremely costly wired network infrastructure. Relevant to the mesh networking paradigm is the extension under development by the 802.11s ESS Mesh Networking Task Group (TG). The scope of this TG is to extend the IEEE 802.11 architecture and protocol for providing the functionality of an extended service set (ESS) mesh. Access points will be capable of establishing wireless links among each other to enable automatic topology learning and dynamic path configuration. Similarly, IEEE 802.16 group is actively working on new specifications for mesh networks and has established sub-working groups to focus on new standards for wireless mesh networks as 802.16g. Particular ad hoc network systems include packet radio networks, sensor networks [6], personal communication systems, establishing communication for emergency or rescue operations, disaster relief efforts and military networks and wireless local area networks.

Robustness is an obvious advantage of wireless ad hoc networks over cellular networks. If a node dissipates its power supply, or malfunctions, the nodes in its vicinity will take over its routing responsibilities. This makes ad hoc networks extremely reliable, as each node is connected to several other nodes. On the other hand, if a base-station becomes inoperable, all the nodes in its cell will lose their connection to the network. A second advantage is the superior resource utilization that ad hoc networks achieve: in cellular networks, users experiencing "poor" wireless link with any base-station are either denied service, or the system consumes a lot of resources (bandwidth and energy) to support their operation. On the other hand, in ad hoc networks, there are many different paths with which a packet can reach its destination. If the channel between two nodes is in a deep fade, the network will work around that link (provided there are other nodes around to handle the traffic). There are many other reasons why the ad hoc topology may be more appropriate than the cellular topology. By overcoming traditional wireless network limitations, ad hoc networks open the door to remarkable new wireless capabilities:

- Instant, automatic formation and evolution of wireless networks. Nodes can automatically join and leave the network anytime.

- Increased coverage and performance. High data throughput requires a high signal-to-noise ratio. However, signals weaken as the transmitter receiver distance increases; meaning lower performance. But in an ad hoc network, full signal strength is restored with each "hop" in the network.

- Lower infrastructure and operational costs. Most ad hoc networks require less back-haul than a traditional wireless network which can greatly reduce deployment and operating expenses.

- Robustness against shadowing. Multi-hopping enables packets to be instantly routed around obstacles and interference.

## 1.2 Ad Hoc Wireless Networks Challenges and Current Research

Although it is quite easy to implement such networks on a small scale, things become much more complicated when the number of nodes is large. Since ad hoc networks of hundreds or thousands of nodes are envisioned for some applications, it is imperative that all algorithms employed by the nodes be distributed. This will greatly improve their usefulness, however it will clearly complicate their design. Moreover, lack of any centralized control and possible node mobility give rise to many issues at the network, medium access, and physical layers, which have no counterparts in the cellular networks or in the wired networks like Internet. Therefore, communication design in an ad hoc wireless network is a very challenging and complicated task. The design of ad hoc wireless networks seems to require novel approaches, since they have peculiar characteristics which differ substantially from those of fixed networks or cellular networks, for which well-established design techniques already exist.

The first issue that comes to mind is a fundamental concept in information theory which is the capacity measured in bit/s [8]. In a network, an extension of this concept has lead to the transport capacity as defined in [9], given by the product between the data-rate (bit/s) and the distance (m) through which the bits can be carried. This approach considers multi-hop communications in order to maximize the transmission of information in the network while not dealing with constraints on the delay, power consumption and the impossibility of knowing the signal to interference noise ratio SINR.

One should think also of connectivity, power control [13, 14] and interference management: is it always possible to find appropriate relays between given source and destination in order to establish the communication? How to control the power at which a signal is transmitted so as to be high enough to reach the intended receivers, while causing minimal interference at other receivers? These problems were analyzed in a seminal work by Gupta and Kumar [10] . The critical transmission range of nodes in a network to ensure that the network is connected with high probability is derived. Later in [12], the connectivity-throughput tradeoffs are explored, and conditions for good performance are explained.

Another issue in ad hoc wireless networks is the design of the routing [16] and the medium access [15] protocols. Due to node mobility and fast fading, topology changes are very frequent in wireless ad hoc networks making the design of such protocols very complex. The choice of the medium access should guarantee a coordination between the nodes so that packets collisions are avoided, while at the same time the channel is utilized efficiently. The use of time sharing or dynamic assignment of frequency bands is too complex. This is in contrast with cellular networks where the base-station allocates channels or time slots to users within a cell. Unlike these *transmission scheduling* policies, where contention is excluded by paying a high price at the spatial reuse, one could think of random multi-access as the current preferred medium access mechanism for ad hoc wireless networks. In these *contention based* protocols, such as Aloha [17], CSMA (Carrier Sense Medium Access) [18], MACA (Multiple Access with Collision Avoidance) [19], MACAW (Multiple Access with Collision Avoidance for Wireless) [20], and IEEE802.11 [2], nodes contend for the channel for each packet they need to transmit. In most protocols specifically designed for use in ad hoc networks, the actual data packet transmission is preceded by an exchange of two small control packets that aim to inform the other nodes that a node is attempting to reserve the channel. Nodes that either have received a control packet or sense the presence of carrier remain silent. This mechanism is typically referred to as Carrier Sense Medium Access with Collision Avoidance (CSMA/CA) [21]. The routing problem is more difficult in wireless networks than in conventional wired networks, like the Internet [22]. In the latter, one strives to find the shortest sequence of segments or links (each segment having two nodes at its end points) connecting the source with the destination. These routing protocols designed for wired networks do not handle topology changes gracefully. A naive solution would be to simply extend the previous protocols, focusing on the design of routing strategies which try to track the evolution of the network's topology. In particular, the solutions proposed in [23] range from proactive routing protocols, where an updated description of the network topology is maintained at each node, to reactive routing protocols which dynamically try to adapt to the changing conditions only if needed. In all of these cases, it often seems that the physical layer is ignored and it is assumed that the physical layer makes each link in the network an error-free connection. Hence, each node should only worry about the forwarding of an incoming packet. These comments emphasize the fact that routing should rely on some measure of the physical constraints, local information, traffic density and interference. The concept of associativity-based routing (ABR) [24] is interesting: it indicates that the route to be preferred should not be the shortest one, but the one passing through the densest area of the network. This should ensure the longest possible route lifetime. The routing protocol for example should play a role in the interference management in the same way TCP (Transmission Control Protocol) is dealing with congestion control. Indeed, delay is caused also in ad hoc wireless networks by interference and fading. A route should be considered as a resource.

## 1.3 Contributions

In the first part of this thesis, we consider large scale networks from a macroscopic view. We address the problem of how throughput in a wireless network scales as the number of users grows. Significant steps toward such a systematic study have been initiated in [9, 26]. In [9], the authors studied a distributed wireless communication model focusing on the case when the nodes are not mobile or equivalently, the desired time scale of communication is assumed to be much faster than the mobility of the nodes. They showed that in a random traffic pattern and random node distribution the total throughput (in bits per second) can grow no faster than $\sqrt{n}/\sqrt{\log n}$, $n$ being the number of nodes in the network, and they gave an achievability result of this growth rate. These results were obtained assuming multi-hop forwarding of packets with single-user decoding and a scheme that uses only nearest-neighbor communications. The latter is justified by the fact that long-range communication between nodes is not preferable since it causes a high level of interference that precludes other nodes from communicating. This nearest-neighbor communication strategy leads to an increase of the number of hops needed to convey a packet from a source to a destination. As a result, most communications have to occur between nearest neighbors, at distances of order $1/\sqrt{n}$, with each packet going through many other nodes (serving as relays) before reaching the destination. The number of hops in a typical route is of order $\sqrt{n}$. Because much of the traffic carried by the nodes are relayed traffic, the actual useful throughput per user pair is small. Nodes acting as relays have to use their radio device not only for transmitting their data, but also the data from other nodes, the channel may thus become very busy, degrading the end-to-end throughput. Therefore relaying is the key reason for capacity reduction in an ad hoc network.

Several works show how these factors affect the capacity of the network and try to find ways to alleviate these effects. In [26], Grossglauser and Tse propose a scheme that takes advantage of the mobility of the nodes. By exploiting node mobility as a type of multi-user diversity, they show that the throughput can increase dramatically when nodes are mobile rather than fixed. Gastpar and Vetterli [31] study the capacity under a different traffic pattern. There is only one active source and destination pair, while all other nodes serve as relay, assisting the transmission between this source-destination pair. The capacity is shown to scale as $O(\log n)$. Our contribution regarding this problematic is as follows:

- In Chapter 2, we continue the investigation along the lines of [9] but show the impact of traffic patterns on the throughput capacity. The study of [9] assume a uniform traffic pattern, where each pair of nodes is equally likely to communicate. The packet path length grows with the physical dimensions of the network leading to a growth in the relay load (since the number of hops to reach the destination increases). In contrast to this analysis, we focus in this work on traffic patterns that allow the per-node capacity to scale well with the size of the network. Our key contribution is that we are able to derive bounds on the per-node capacity of an ad hoc wireless network where local communications predominate by using a simple deterministic scheme. Moreover our results capture the impact of local communications on the increase of throughput independently of the path length distribution. We also investigate a hybrid wireless network, a tradeoff between a purely ad hoc network and a cellular one. In the latter, data is always forwarded through the base-station, whereas in our model of a hybrid network, a cellular mode (data forwarded from source to destination through the base-station) and a pure ad hoc mode (data forwarded from source to destination using multi-hop relaying communications) coexist. The primary interest is to reduce the transmit power of mobile terminals through multi-hop relaying which leads to a decrease in the number of hops from the source to the destination (since traffic is absorbed by the base-station). We assess the tradeoff between the number of base-stations in the network and the increase in the throughput capacity of a hybrid ad hoc network due to the additional infrastructure. Finally, a discussion on the impact of event-driven traffic patterns in conducted. Nodes report to the network collector measurements on a particular event. We show the impact of the density of events on the throughput for a wireless sensor network with $n$ sensing nodes and one collector.

As stated above, the main performance limitation is due to the relay traffic load. In [26], a natural strategy to overcome this limitation with mobility, is for each source node to split its packet stream to as many different nodes as possible. These nodes then serve as mobile relays and whenever they get close to the final destination, they hand the packets off to the final destination. The basic idea is that since there are many different relay nodes, the probability that at least one is close to the destination is significant. On the other hand, each packet goes through at most one relay node and, hence, the throughput can be kept high. Although the basic communication problem is point-to-point, this strategy effectively creates multi-user diversity by distributing packets to many different intermediate nodes. This result assumes a 2-dimensional mobility pattern, where the trajectory of each node is an independent, stationary and ergodic random process with uniform distribution on the unit disk. That is, the mobility pattern is homogeneous with respect to each node, and the sample path of each node covers all the space over time. A natural question to ask is whether this good performance is specific to the particular generous mobility pattern, or whether it can be achieved under more restricted mobility conditions. Our contribution in that sense is as follows:

- In Chapter 3, Our main contribution is a method that allows to check whether, for a given mobility pattern, a constant $\Theta(1)$ throughput per source-destination pair is possible. We introduce the *connectivity graph*, that does not represent the actual physical network, but rather the available communication resources, for a given transmission policy. The connectivity graph offers an abstraction of the communication capabilities of the ad hoc network: we can study the long-term averaged throughput between source-destination pairs in the actual ad hoc network, by examining information flows in the connectivity graph. Thus, by mapping the ad hoc network problem into a graph problem, we establish a bridge between the multi-commodity flow and the ad hoc network problem. This bridge can be used in both directions. We analyzed the properties of the connectivity graph to develop a set of necessary and sufficient conditions under which constant $\Theta(1)$ per-node throughput is possible. We then identified the structural properties these conditions imply for the connectivity graph and how they translate in properties for the underlying mobility pattern.

In the second part of this thesis, we take a microscopic view of ad hoc wireless networks where the design of such networks is emphasized. For example, in the approaches described in Chapter 2 and Chapter 3, the *physical layer PHY* was not considered explicitly, it was taken into account by the abstraction of the problem. Chapter 2 and Chapter 3 presented useful results on the scale order of wireless ad hoc networks, but practical implementation issues are not addressed, e.g. the PHY, the multi-hop routing (node location information), the coordination between nodes (to realize scheduling) etc.

Engineers have historically partitioned the design of communication systems into a stack of protocol layers, each serving a particular purpose. These layers range from the highest-layer (farthest from the wireless channel, most abstracted), the application layer, to the lowest layer (closest to the wireless channel, least abstracted), the physical layer. The tasks allocated to each layer are as follows. The *application layer* generates or handles user signals and conveys them through an interface to the transport layer. The *transport layer* often performs packet sequencing, end-to-end retransmission, and flow control. The *network layer* routes messages through the network over a set of point-to-point links created by the link layer. The *link layer*, and its associated *medium-access control (MAC)* sub- layer, maintains a set of virtual point-to-point communication links built on top of the physical layer. Finally, the physical layer incorporates a majority of the analog circuitry and signal processing and provides for transmission reception and processing of signals over the wireless channel. One can ask whether the particular layers are appropriate and whether there is a more natural set of abstractions that presents opportunities for cross-fertilization of ideas across all the layers.

In this part, we focus on a framework characterized by a bottom-up approach, where the impact of the physical layer (PHY) on media access control (MAC) and routing layers is evaluated. The more common approach is to focus on a particular layer and compare two alternative protocol designs. The main advantage of this approach is architectural flexibility, but this may lead to largely suboptimal network designs especially in wireless networks. As an example, consider the power control algorithm that specifies the power with which signals are transmitted. It is tempting to associate it exclusively with the physical layer. Such an association would be correct in the case of wired networks. In the wireless environment the transmit power of a particular node will affect all other nodes in the network, by changing the levels of interference experienced by these nodes. The transmit power will affect the connectivity of the network, which in turn impacts routing. As a consequence of this interdependency, there is a need for a more complex medium access mechanism. On the one hand, this mechanism should be able to control the amount of interference experienced by receivers. On the other hand, it should exploit spatial reuse and, in certain cases, enforce concurrent transmissions, in order to maximize the performance. The medium access should also influence the physical layer. If the total amount of the interference at a receiver during a reception of a packet is high, the physical layer should decrease the transmission rate to cope with it. On the contrary, if the interference is low, the physical layer should benefit from the conditions and transmit with a high rate. Therefore, in the wireless environment, power control, rate adaptation are not confined to the physical layer, but in reality can affect the operation across the physical, MAC, and routing layers. This fact should not be viewed as a complication, but rather as an opportunity for system level cross-layer design, i.e., a design that spans multiple layers of the protocol stack. Our contribution is as follows:

- In Chapter 4, Our work is motivated by cross-layer mechanisms (PHY/MAC/Routing) aiming at maximizing the spectral efficiency of the network. To this end, we present a cross-layer framework for the design of multi-hop wireless networks. We jointly address the properties of the physical and the data link layer in the design of the media-access control (MAC) protocol and provide conclusions on routing strategies based on physical layer metrics. We assume that nodes access the channel at random and employ simple protocols to retransmit the erroneously received packets. We consider two possible retransmission protocols: the classical reference scheme is *Slotted Aloha* (using the wireless setting as described in [58]). It shows the benefit of coupling channel coding for medium-access and where decoding considers only the most recent received block. The second is *Incremental Redundancy* where decoding takes into account all previously received signal blocks and performs soft combining until decoding is achieved successfully. We compare these strategies to the generalization of the collision channel without feedback or delay constraints [59], where the measure of success of a transmission will depend on the achievable ergodic throughput of the corresponding channel. For this analysis, the nodes are taken to be spatially distributed on the plane according to a homogeneous spatial Poisson process which leads to a new representation of interference and collisions between concurrent transmissions. This random characterization of the network is justified by the fact that the homogeneous Poisson point-process is spatially ergodic and thus the performance quantities considered in this work for particular network realizations (network throughputs and information outage probability), will converge quickly to the average performance of the random network. To derive the throughput, we follow the analysis of Nelson and Kleinrock in [60] where they studied the spatial capacity of a slotted Aloha multi-hop network with capture. The spatial throughput is computed by multiplying the number of the simultaneously successful transmissions and the corresponding average jump ( hop distance or expected forward progress ).

The main contributions of this chapter are

- The study of incremental redundancy as a multiple access technique for ad hoc wireless networks
- The representation of interference and collisions statistics in an exact manner from the homogeneous Poisson point process network model
- A cross-layer framework where multi-hop routing protocols are analyzed and tools for characterizing the spatial throughput (bit-m/dim, related to the transport capacity) from a microscopic point-of-view as a function of topological parameters (e.g node population density) and system parameters (propagation, bandwidth etc).

The physical layer conventionally combats fading with coding, spread-spectrum, and multiple antennas. With the recent interest in ad hoc wireless networks, the notion of cooperative communication has received tremendous attention [86, 87, 88]. It offers the opportunity to develop novel communication techniques based on multi-node cooperation that can perform efficiently even over harsh fading channels. This feature is enhanced by ad hoc environments where nodes are sufficiently far apart so that rich scattering exists and can be exploited via spatial diversity. The latter is implemented by creating multiple paths that carry the same information. This redundancy allows the ultimate receivers to essentially average channel variations resulting from fading, shadowing, and other forms of interference. By contrast, classical network architectures only employ a single path through the network and thus forego these benefits. Our contribution in this direction is as follows:

- In Chapter 5, we present an efficient protocol for the delay-limited fading Automatic ReQuest (ARQ) single relay half-duplex channel. The terminals are constrained to employ half-duplex transmission, i.e. they cannot transmit and receive simultaneously. The source is using an ARQ retransmission protocol to send data to the relay and the destination. When the relay is able to decode, both the relay and the source send the same data to the destination providing additional gains. The source and the relay are allowed to transmit in the same channel using cooperative protocols not relying on orthogonal subspaces, allowing for a more efficient use of resources. The ARQ permits the use of communication over a variable number of blocks depending on the quality of the channel. The proposed protocol exploits two kinds of diversity: (i) space diversity available through the cooperative (relay) terminal, which retransmits the source's signals, (ii) ARQ diversity obtained by leveraging the retransmission delay to enhance the reliability. The performance characterization is in terms of the achievable diversity, multiplexing gain and delay tradeoff for a high signal-to-noise ratio (SNR) regime. Finally, we show the benefits of power control on the diversity by controlling the source's power level over the retransmission rounds.

# Part I

# Chapter 2

# Impact of Traffic Patterns on the Throughput Capacity of Ad Hoc Wireless Networks

## 2.1 Introduction

The study of wireless ad hoc networks has recently received significant attention. A purely ad hoc network is a collection of wireless nodes forming a network without the use of any existing network infrastructure or centralized coordination. In [9], Gupta and Kumar determined the scaling of capacity of these networks under simplified propagation and traffic assumptions. They showed that given $n$ nodes randomly located on a unit disk and a uniform traffic pattern (i.e. that nodes are equally likely to communicate with any other node in the network), the aggregate capacity is of $O(\sqrt{n})$ [1] allowing optimal scheduling and relaying of packets. The nodes are however assumed to be fixed throughout the duration of the communication sessions. Because of their assumptions regarding interference and measure of connectivity, their result is not of an information theoretical nature. Xie and Kumar relaxed these assumptions in [25], and proposed another upper bound on the total rate of communication in the network. Based on the assumption of a minimum distance between nodes and a power loss exponent $\alpha > 6$, it is shown that the transport capacity is asymptotically bounded

[1] We use Knuth's notation: $f(n) = O(g(n))$ if $\limsup_{n\to\infty} \frac{f(n)}{g(n)} < +\infty$; $f(n) = \Omega(g(n)$ if $g(n) = O(f(n)$; $f(n) = \Theta(g(n))$ denotes that $f(n) = O(g(n))$ as well as $g(n) = O(f(n))$.

by the sum of the transmit power of the nodes in the network, mainly for domains of size $\Theta(n)$, transport capacity scales as $O(n)$, leading to a rate of communication which is again sub-linear. More recently, it has been shown that even with information-theoretically optimal coding strategies, the per-user capacity still diminishes to zero [37]. The upper bounds have been derived for $\alpha > d \vee 2(d-2) = \max(d, 2(d-2))$, where $d$ is the dimension of the network, for the uniform traffic pattern and for extended networks (i.e., the number of users per unit area is constant, and increasing the number of users implies an increase in geographical area, which is a scenario studied in [25] but with different assumptions as seen above, namely for $\alpha > 6$ or $\alpha > 2(d+1)$). In [26], the model in [9] was modified to take into account mobility and using only one-hop relaying, an $\Theta(n)$ throughput was obtained for a mobile ad hoc network. Even with limited mobility, i.e., when nodes move on large circles, it was shown in [42] that the throughput of ad hoc wireless networks can be enhanced. In [27, 28], the capacity of a three-dimensional wireless ad hoc network is studied. These results provide expressions for the ad hoc network capacity and determine the scalability of such networks as the number of nodes increases to infinity.

As can be concluded from the studies referenced previously, the performance limitation of an ad hoc network comes first from the long-range peer-to-peer communication (that causes excessive interference) and second the increase in relayed traffic in the case of multi-hop transmissions. let $\overline{L}(n)$ be the mean distance traversed by a packet and $r$ be the common transmission range (which is proportional to transmit power) and each node has a randomly chosen destination to which it wishes to send $\lambda(n)$ bits/s. Then each packet has to take $\frac{\overline{L}(n)}{r}$ hops to reach destination. This creates $\frac{\overline{L}(n)\lambda(n)}{r}$ bits/s of traffic per use for other nodes, and if each link is capable of $W$ bits/s, we should have $\lambda(n) \leq \frac{Wr}{\overline{L}(n)}$. The right-hand side is proportional to the range, so it appears that increasing the range should increase the throughput. But an increased range causes more interference and loss of packets (spatial concurrency, simultaneous transmissions), and too small a range increases relay traffic.

In general, the transmitter-receiver pairs are not arbitrarily close to each other and an important physical insight from [9] is the need for multiple hops to reach the destination. Because the large majority of traffic carried by the nodes is relayed traffic, in [26] each packet is constrained to make at most two hops and transmission is limited to nearest neighbors. But since source and destination are nearest neighbors only for a very small fraction of time, the transmission is spread to a large number of intermediate mobile relay nodes, and whenever they get close to the final destination, they hand the packets off to the final destination. Suppose now that the transmitter-receiver pairs of nodes are close to each other, then reliable communication will cause little interference to the other nodes and the scenario is essentially that of a set of non-interfering point-to-point communication systems, or transmitter-receiver pairs communicating through a small number of hops. The studies cited above assume a uniform traffic pattern, where each pair of nodes is equally likely to communicate, so that packet path length grows with the physical dimensions of the network leading to a growth in the relay load (since the number of hops to reach the destination increases). This assumption may not be true in large networks, where users communicate mostly with physically nearby nodes: users in the same department in a university, the same group in a company, and even in the case of telephony, users communicate mostly with neighbors in the same city (or even district) rather than users in other countries.

In [29], traffic patterns that allow the per node capacity to scale well with the size of the network are discussed. The local traffic pattern is scalable where the expected path length clearly remains constant as the network size (equivalently the number of nodes in the case of a large network) grows. [30] illustrates the impact of an exponentially decaying traffic pattern and the relay load on the throughput in the context of a decentralized system with retransmission protocols. In [31], the capacity under a different traffic pattern is studied. There is only one active source-destination S-D pair, while all other nodes serve as relay, helping the transmission between the source and the destination nodes. The capacity is shown to scale as $\Theta(\log n)$. In this chapter, we continue the investigation along the lines of [9] but show the impact of traffic pattern on the throughput capacity. We are able, by using a simplified deterministic scheme, to derive a lower bound on the per-node capacity of an ad hoc wireless network where local communications predominate.

The pessimistic results of [9] dampened the early enthusiasm for ad-hoc networks which would eliminate the need for infrastructure like base-station (BTS). In this work we also investigate a hybrid wireless network, a tradeoff between a purely ad hoc network and a cellular one. In the latter, data is always forwarded through the base-station, whereas in our model of a hybrid network, a cellular mode (data forwarded from source to destination through the base-station) and a pure ad hoc mode (data forwarded from source to destination using multi-hop relaying communications) coexist. The primary interest is to reduce the transmit power of mobile terminals through multi-hop relaying. In [32], the introduction of a sparse network of base-stations was shown to help improving the network connectivity. In [33], the scaling behavior of the throughput capacity of a hybrid network is studied under two particular routing strategies. It was shown that an effective improvement of a hybrid mode over a pure ad hoc mode is provided only if the number of base-stations scales faster than the square-root of the number of nodes in the network. Here, we assess the tradeoff between the number of base-stations in the network and the increase in the throughput capacity of a hybrid ad hoc network due to the additional infrastructure. We assume that base stations are connected to each other by a wired network, and are regularly placed within the ad hoc network. Terminal nodes are reaching the base-stations through multi-hop communications. On the other hand, the link from the base-stations to the terminals (down-link) can be achieved by single-hop communications, since we assume there is no power constraint for base-stations.

Finally, a discussion on the impact of event-driven traffic patterns in conducted, where we show the impact of the density of events on the throughput for a wireless sensor network with $n$ sensing nodes and one collector.

The outline of the chapter is as follows: In Section 2.2 an overview of the results in [9, 26] is discussed. In Section 2.3 we specify the ad hoc network model and our problem model. In Section 2.4 we describe the constructive communication scheme. Section 2.5 deals with the throughput capacity expressions for both locally predominant communications and hybrid networks. Event-driven traffic patterns are discussed also. Finally, in Section 2.6 we draw some conclusions.

## 2.2 Related Work

The aim of this section is to provide an overview and intuitive insights about the scaling capacity with the number of nodes in an ad hoc network.

### 2.2.1 Gupta Kumar Results

In [9], Gupta and Kumar described how the capacity of an ad hoc wireless network scales with the number of nodes $n$. Two types of capacity metrics were defined: (i) transport capacity (in bit-meter per second) is the number of bits transmitted per unit time, multiplied by the distance over which the bits were transported towards their destinations; (ii) throughput capacity (in bit per second) is defined in the usual manner as the average number of bits transmitted per unit time by every node to its destination. Note that transport capacity is an attribute of the whole network, thus an aggregate metric, whereas throughput capacity concerns the transmission rate per node. Moreover the following two network models were introduced in [9].

#### Arbitrary Networks

In this scenario, $n$ nodes are arbitrarily located on a flat disk of unit area. Each node arbitrarily chooses a destination to send its message at an arbitrary rate, and also arbitrarily choose a transmission range or power level.

Successful reception of a transmission over one hop is modeled in two ways:

- *The Protocol Model*: let $X_i$ denote the location of a node, and suppose node $X_i$ transmits to node $X_j$. This transmission is successfully received by node $X_j$ if:

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

  for any other node $X_k$ simultaneously transmitting over the same frequency. The quantity $\Delta > 0$ models situations where a guard zone is specified by the protocol to prevent a neighboring node from transmitting on the same channel at the same time (see Fig.2.1). It also allows for imprecision in the achieved range of transmissions.

- *The Physical Model*: let $\{X_k; k \in \mathcal{T}\}$ be the subset of nodes simultaneously transmitting on the same frequency. Assume node $X_k$ transmits with power $P_k$ , for $k \in \mathcal{T}$. The transmission from a node $X_i$ , $i \in \mathcal{T}$ is successfully received by node $X_j$ if the inequality:

  $$\frac{P_i |X_i - X_j|^{-\alpha}}{N_0 + \sum_{\substack{k \neq i \\ k \in \mathcal{T}}} P_k |X_k - X_j|^{-\alpha}} \geq \beta$$

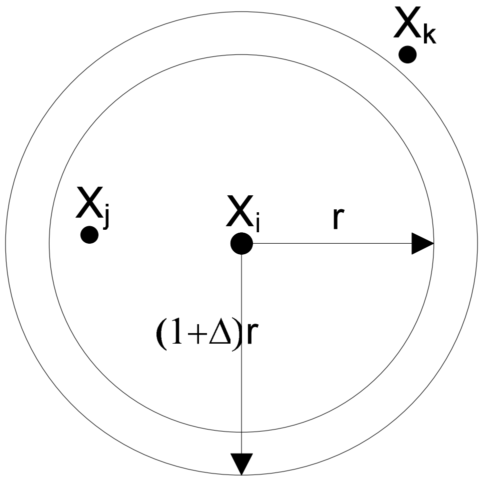


**Figure 2.1:** The Protocol Model.

is satisfied, where $\beta$ is a signal to interference noise ratio (SINR) threshold for successful receptions, $N_0$ is the ambient noise power level, and $\alpha > 2$ indicates the signal power decay with distance $\frac{1}{r^\alpha}$.

For arbitrary networks, node's location and transmission range can be chosen and traffic distributed in a clever way. The transport capacity of such networks, is $\Theta(\sqrt{n})$ bit-meters/s in the protocol model, and $c_3\sqrt{n}$ and $c_4 n^{\frac{\alpha-1}{\alpha}}$ bit-meters/s are respectively lower and upper bounds for the physical model, for some $c_3$ and $c_4$ depending only on $\alpha$ and $\beta$ (specified above). Thus, even under optimal conditions, the throughput is only $\Theta(\frac{1}{\sqrt{n}})$ bits/s for each node for a destination non-vanishingly far away under the protocol model.

### Random Networks

In this scenario, $n$ nodes are uniformly and independently distributed on a planar disk of area $1\,\mathrm{m}^2$. Each node sends data at $\lambda(n)$ bits/s to a randomly chosen destination node. This destination node is picked as follows. A uniformly and independently distributed point is chosen, and the node nearest to this location is chosen as the destination node. Thus, the average separation between source destination S-D pairs is on the order of $1\,\mathrm{m}$.

Similarly to the arbitrary networks case, the physical and the protocol model are adopted as transmission models. The only difference is that a common range $r$ is introduced for all transmissions. The inequality in the protocol model changes to $|X_i - X_j| \leq r$ and $|X_k - X_j| \geq (1 + \Delta)r$, and for the physical model the SINR threshold inequality becomes:

$$\frac{P|X_i - X_j|^{-\alpha}}{N_0 + \sum_{\substack{k \neq i \\ k \in \mathcal{T}}} P|X_k - X_j|^{-\alpha}} \geq \beta$$

A throughput of $\lambda(n)$ bits per second for each node is said to be *feasible* if there is a spatial and temporal scheme for scheduling transmissions, such that every node can send $\lambda(n)$ bits per second on average to its chosen destination node through the intermediate nodes and some buffering strategy in the intermediate nodes. The definition is stated in [9] as follows:

**Definition 2.1.** *The throughput capacity of Random Networks is said to be of order* $\Theta(f(n))$ *bits/s if there are deterministic constants* $c > 0$ *and* $c^{'} < +\infty$ *such that*

$$\lim_{n \to \infty} Prob(\lambda(n) = cf(n)\, is\, feasible) = 1 \tag{2.1}$$

$$\lim_{n \to \infty} Prob(\lambda(n) = c^{'}f(n)\, is\, feasible) = 0 \tag{2.2}$$

Eq.(2.1) is used to prove the lower bound on the throughput whereas Eq.(2.2) is used for the upper bound.

The main result for random networks is as follows: if the nodes are randomly located in the disk, each node uses a fixed range/power and each node sends data to a randomly chosen destination, then each node can obtain a throughput of $\Theta(\frac{1}{\sqrt{n \log n}})$ bits/s in the Protocol Model, while $\frac{c_1}{\sqrt{n \log n}}$ and $\frac{c_2}{\sqrt{n}}$ bits/s are, respectively, lower and upper bounds in the Physical Model for some $c_1$ and $c_2$ depending only on $\alpha$ and $\beta$.

These results demonstrate that as the number of nodes per unit area $n$ increases, the throughput per S-D pair decreases approximately as $1/\sqrt{n}$. This is the best performance achievable even allowing for optimal scheduling, routing, and relaying of packets in the networks and is a somehow pessimistic result on the scalability of such networks. The intuition behind this phenomenon is that a transmission may travel either through a single direct transmission or through multiple hops via relay nodes. As shown in the protocol model, the successful reception of the transmission of a given

S-D pair prohibits simultaneous transmission with in the disk of radius proportional to the transmission distance of the pair: a successful transmission over the range $r$ incurs a cost proportional to $r^2$ by excluding other transmissions in the vicinity of the sender. In order to maximize the transport throughput of a network, i.e., the total number of meters traveled by all bits per time unit, it is therefore beneficial to schedule a large number of short transmissions. The best we can do is to restrict transmissions to neighbors, which are at a typical distance of $1/\sqrt{n}$ (Since the expected distance for each session is $\Theta(1)$, the number of relays a packet has to go through scales as $\sqrt{n}$). The transport capacity is then at most $\sqrt{n}$ bit-m/s . As there are $n$ sessions, it follows that the throughput per session can at best be $O(1/\sqrt{n})$.

### 2.2.2 Impact of the Mobility on the Throughput of Ad Hoc Wireless Networks

The capacity of a mobile ad hoc wireless network was first studied in [26]. The network consists of $n$ nodes all lying on a disk of unit area. The location of the *ith* user at time $t$ is given by $X_i(t)$, which is modeled as a stationary and ergodic process with stationary distribution. Moreover, the trajectories of different users are i.i.d. The intuition of [26] is that any two nodes can be expected to be close to each other from time to time so that we may improve the capacity of the network (the delay tolerance can be usefully exploited in a mobile wireless network). The authors in [26] first show that without relaying, there is no way to achieve a $\Theta(1)$ throughput per S-D pair. The explanation is that the number of simultaneous long-range communications is limited by interference. If transmissions over long distances are allowed, then there are many S-D pairs that are within the range of interference. This limits the number of concurrent transmissions over long distances, and the throughput is then interference limited. On the other hand, if we constrain communications to neighboring nodes, then there is only a small fraction of S-D pairs that are sufficiently close for transmitting a packet. Hence, the throughput is distance limited. To increase the throughput, one needs to find a way to limit the transmission locally, while guaranteeing that there would be enough sender-receiver pairs that have packets to send. The authors in [26] propose to spread the traffic stream between the source and the destination to a large number of intermediate relay nodes. The goal is that in the steady-state, the packet of every source node will be distributed across all the nodes in the network, ensuring that every node will have packets buffered destined to every other node. This ensures that a scheduled sender-receiver pair always has a packet to send, in contrast to the case of direct transmissions. A question that naturally arises is that how many times a packet needs to be relayed. In fact, as the node location processes are independent, stationary, and ergodic, it is sufficient to relay only once. This is because the probability for an arbitrary node to be scheduled to receive a packet from a source node S is equal for all nodes and independent of S. Each packet then makes two hops, one from the source to its randomly chosen relay nodes and one from the relay node to the destination as shown in Fig.2.2.

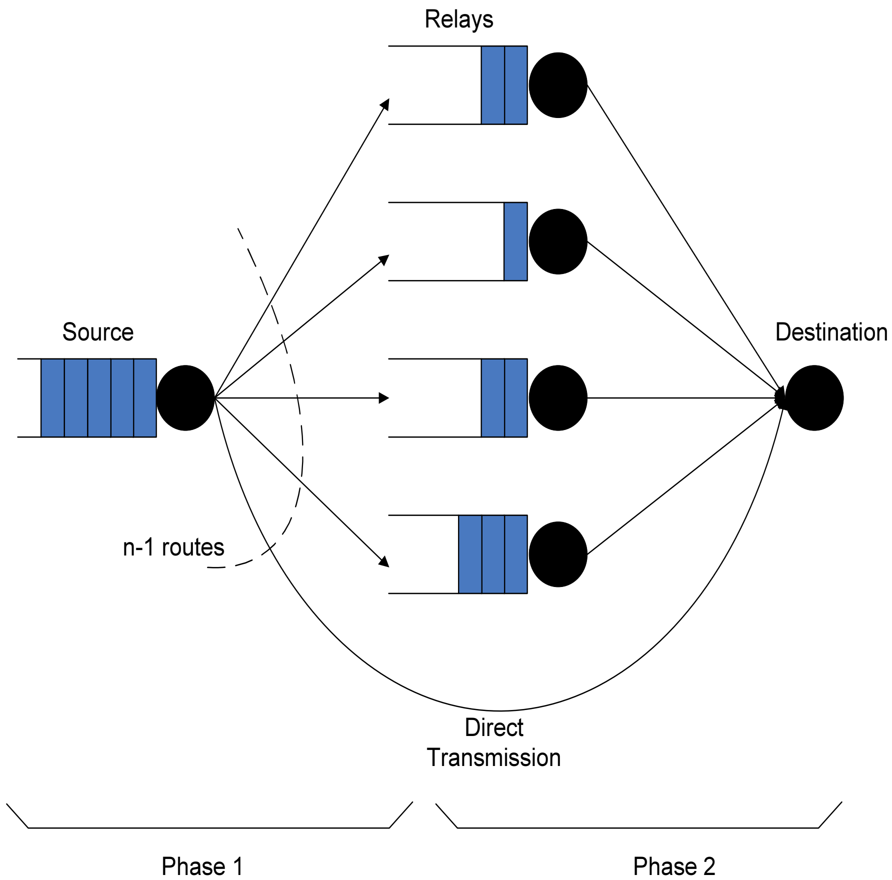


**Figure 2.2:** The two-phase scheduling scheme viewed as a queueing system.

In [26], a scheduling algorithm consisting of two phases is proposed: the scheduling of packet transmissions from sources to relays and the scheduling of transmissions form relays to final destinations. Note that in both phases a transmission from a source directly to a destination is possible. This two-phased algorithm achieves a throughput of $\Theta(1)$ per S-D pair. This result is based on the fact that the probability that two nodes come in proximity and are selected as a S-D pair is $\Theta(1/n)$. For a given S-D pair, there is one direct route and $n-2$ two-hop routes. The throughput of the direct route is $\Theta(1/n)$. Each of the two-hop route can be treated as a single server queue, each with arrival and service rate of $\Theta(1/n)$. The total throughput is $\Theta(1)$ by summing all the throughputs of $n-1$ routes.

The common deduction from the results stated above is that the capacity determining constraint is the number of source-destination paths that pass through a node, i.e., the relaying burden of a node. This is the key observation motivating the rest of this chapter.

## 2.3 Network and Problem Model

We consider a random network where $n$ nodes are distributed uniformly on a two-dimensional area, a square of area $\Theta(n)$ (a similar network model is considered in [34] with unit area). This is a large network where the number of nodes is increasing with the area of the network leading to a fixed density of nodes per area (similar scenario as in [25, 37]), whereas in the Gupta and Kumar model [9], the density of nodes is increasing with the number of nodes as the area is fixed. The large network model is more realistic since one would not expect nodes to get arbitrarily close by letting the number of nodes become very large. We assume that all nodes can act as both transmitters and receivers, and each node wants to communicate with another node chosen randomly and independently among the rest. Therefore there exist $O(n)$ communicating pairs of nodes. Moreover, the nodes are static, relative to the time scale of communication.

In the system we are considering, each node can transmit over a common wireless channel of bandwidth $W$. The propagation model is described by the signal attenuation due to the distance $r$ between the transmitter and the receiver, proportional to $r^{-\alpha}$, where $\alpha$ is the power loss exponent (a positive number typically $\alpha > 2$ which is the usual model outside a small neighborhood of the transmitter). Each node transmits with a common power $P$. Let $\{X_t : t \in \mathcal{T}\}$ be the set of transmitting nodes at a given time, and suppose that node $X_i$ transmits to a node $X_j$, then the signal to interference and noise ratio (SINR) at node $X_j$ is given by:

$$\gamma_{ij} = \frac{P|X_i - X_j|^{-\alpha}}{N_0 + \sum_{\substack{k \neq i \\ k \in \mathcal{T}}} P|X_k - X_j|^{-\alpha}} \tag{2.3}$$

where $N_0$ is the thermal noise at receiver node $j$. We take the transmission rate as Shannon's formula $C_{ij} = W \log_2(1 + \gamma_{ij})$ where single-user decoding is assumed, i.e., each decoder treats the signals from other users as noise, and the single-user decoder for each node has perfect knowledge of the channel gain and the total interference power, i.e., noise and interfering user traffic.

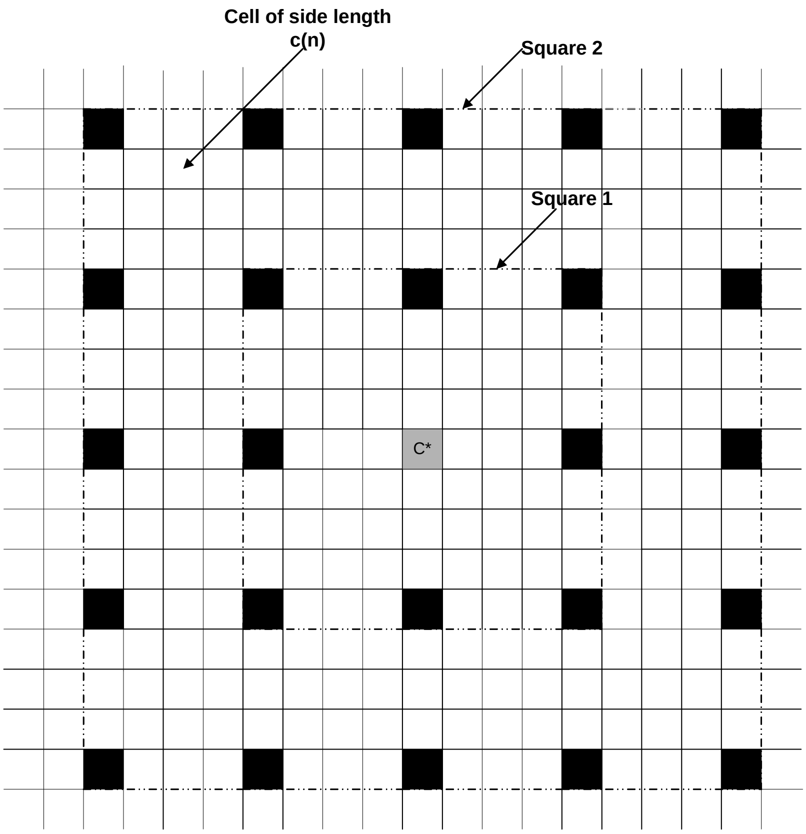


**Figure 2.3:** An example of network area partition with $K = 4$. All nodes in the shaded cells can transmit simultaneously to the eight neighboring cells.

## 2.4 Constructive Communication Scheme

### 2.4.1 The Cell Partition

We partition the area of the network into a set of regular cells, where each cell is a square of side length $c(n)$. We impose that nodes transmit only to nodes in the same cell or in (the eight) adjacent neighboring cells. In this local communication, all cells sufficiently far away can simultaneously transmit with reduced interference. We introduce a parameter $K$ which corresponds to a reuse factor in a cellular system as in [34]. Indeed, all the

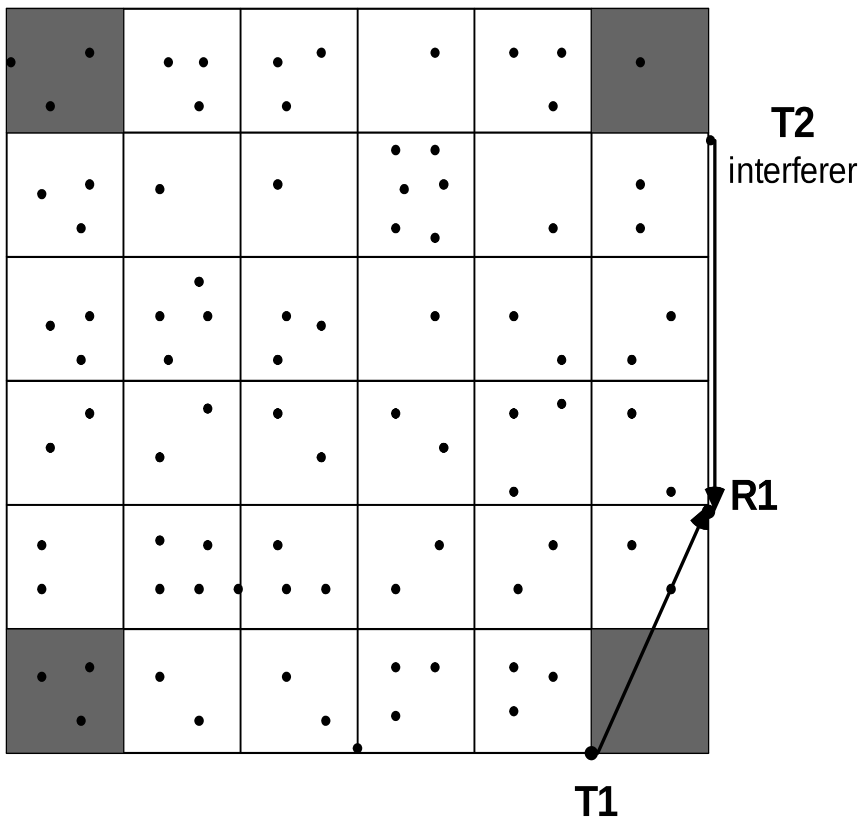


**Figure 2.4:** An example of network area partition with $K = 5$ showing the maximal distance of a single hop (T1-R1) which is always less than the distance interferer-receiver T2-R1.

cells that are a at vertical and horizontal distance of exactly some multiple of $K$, can transmit simultaneously as depicted in Fig.2.3. Then, we choose a finite length time-division scheduling scheme of $K^2$ slots ($K > 2$), in which each cell is assigned one slot to transmit. This scheduling between cells ensures that transmissions from a cell do not interfere with transmissions in simultaneously transmitting cells. Nodes in the same group of nodes transmit with reduced interference, and the distance from an interferer to a receiver is at least $c(n)$. Fig.2.4 shows an example of the cell partition with $K = 5$, where the distance transmitter-receiver is always less than the distance interferer-receiver. When a cell becomes active, packets that are relayed or originated from this cell are scheduled one after the other (one packet by Source-Destination pair).

### 2.4.2 The Routing Strategy

The packet routing is as follows: a packet is relayed from the cell containing the source to the cell containing the destination in a sequence of hops. In each hop, the packet is transferred from one cell to another, in the order in which cells intersect the straight line connecting the source to the destination. This is depicted in Fig.2.5. To make it possible to relay traffic between cells, we need to guarantee that every cell contains at least one node with high probability.

**Lemma 2.1.** *For $c(n) = \sqrt{3 \log n}$, no cell is empty with high probability as $n$ is large.*

*Proof.* For a network of area $\Theta(n)$ where $n$ is the number of nodes in the network, with cells of side length $c(n)$, the probability that a particular cell is empty is bounded by $\left(1 - \frac{c(n)^2}{n}\right)^n$. By using the union bound, we have:

$$\begin{aligned}
\Pr[\textit{at least one cell is empty}] &\leq Q(n)\left(1 - \frac{c(n)^2}{n}\right)^n \\
&= \frac{n}{c(n)^2}\left(1 - \frac{c(n)^2}{n}\right)^n \\
&\overset{(a)}{\leq} \frac{n}{c(n)^2}\exp(-c(n)^2) \qquad (2.4)
\end{aligned}$$

where $Q(n)$ is the number of cells in the network area and it is equal to $\frac{n}{c(n)^2}$, and inequality (a) follows from $1 - x \leq \exp(-x)$. We obtain for $c(n) = \sqrt{3 \log n}$:

$$\begin{aligned}
\Pr[\textit{at least one cell is empty}] &\leq \frac{1}{3n^2 \log n} \\
&\to 0 \qquad (2.5)
\end{aligned}$$

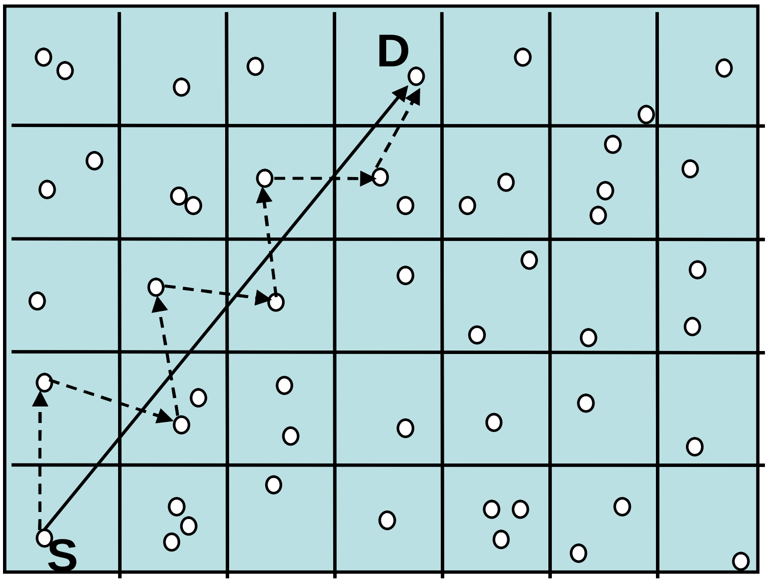


**Figure 2.5:** Routing along straight lines.

and the result is proven for sufficiently large $n$.

**Lemma 2.2.** *If the power loss exponent satisfies the condition $\alpha > 2$, the SINR at node $j$ is lower bounded by a constant, hence the rate transmission of pair $(i, j)$ is asymptotically lower bounded by:*

$$C_{ij}(n) \geq k_1 \tag{2.6}$$

*where $i, j$ belong to adjacent neighboring cells.*

*Proof.* From the routing strategy and the time-division scheme, intuitively such a bound on the SINR exists. To prove this rigorously, we need first to derive a lower bound on the power of the useful signal. Under the routing strategy, each node can transmit only to nodes in the same cell or nodes in neighboring cells. Under this assumption, the maximum distance between a transmitter and a receiver is $\sqrt{5}c(n)$. The power of the useful signal at the receiver is then bounded by:

$$P|X_i - X_j|^{-\alpha} \geq P(15 \log n)^{-\frac{\alpha}{2}} \tag{2.7}$$

We need to bound the interference. This is derived in Appendix 2.A. We obtain:

$$\begin{aligned} I &= \sum_{\substack{k \neq i \\ k \in \mathcal{T}}} P|X_k - X_j|^{-\alpha} \\ &\overset{(\alpha > 2)}{\leq} \frac{cPK(\log n)^{-\frac{\alpha}{2}}}{(K-2)^{\alpha}} \end{aligned} \tag{2.8}$$

where $c$ is a positive number and $\mathcal{T}$ is the set of simultaneously transmitting nodes. By combining Eq.(2.7) with Eq.(2.8), the SINR$(n)$ is lower bounded by SINR$_{\min}(n)$ which is a constant and as $C_{ij}(n) = W \log_2(1+\text{SINR}_{\min}(n))$, we obtain Eq.(2.6).

## 2.5 Throughput Capacity Expressions

### 2.5.1 Local Traffic Pattern

The information packets that are relayed through a particular cell create load for the nodes in the cell, and it is important to compute the maximum number of routes passing through any cell. This helps us estimate how much traffic, apart from its own, each cell has to relay, and the reduction in the node-throughput induced by the relay traffic. We recall that a route is the collection of cells a source will use to forward packets to a destination following the straight line connecting the source to the destination (hence a route is a S-D line).

**Lemma 2.3.** *The number of routes passing through any cell is* $\Theta(\log n)$ *for* $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1)^2$, *whereas it is on the order* $\Theta(\overline{L}(n)\sqrt{\log n})$ *for* $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1)$, *where* $\bar{L}(n)$ *is the average source destination distance. Result that holds with high probability as n is large.*

*Proof.* We first argue intuitively. Let $X$(s-d) be the S-D distance, $\overline{L}(n) = E[X(\text{s-d})]$, the average path length and $n$ being very large. Each S-D line will

[2] $\Theta(g(n)) \leq \Theta(f(n))$ means that $f(n)$ grows faster than $g(n)$

traverse a mean number of $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right)$ cells. Moreover, we consider $\Theta(n)$ pairs S-D. Then, $\Theta\left(n\frac{\overline{L}(n)}{c(n)}\right)$ is the mean number of times all cells are traversed by S-D lines, and we can conclude that $\Theta\left(\frac{n\overline{L}(n)}{c(n)\text{number of cells}}\right)$ is the mean number of routes passing through a cell. For a uniform traffic pattern, the path length is $\Theta(\sqrt{n})$ in a large network of area $\Theta(n)$, and as the number of cells is $\Theta\left(\frac{n}{\log n}\right)$, the mean number of lines passing through a cell is $\Theta\left(\frac{n\overline{L}(n)}{c(n)\text{number of cells}}\right) = \Theta(\sqrt{n\log n})$ as in the Gupta-Kumar model. In [29], traffic patterns that allow the throughput capacity to scale with the network size are discussed. For local traffic patterns, the expected path length (S-D distance) remains constant as the network size grows. Actually, for a local traffic pattern (power decaying law), the path length is $\Theta(1)$. One can notice that for a path length order smaller than a cell side length (i.e $\frac{\overline{L}(n)}{c(n)} = \Theta\left(1/\sqrt{\log n}\right) \leq \Theta(1)$), the source-destination line is completely included in the cell. Even if the line is inside the cell, traffic should be scheduled for this pair S-D, and we should count that at least one line is intersecting the cell. We take it into account by replacing $\frac{\overline{L}(n)}{c(n)}$ by $\left\lceil\frac{\overline{L}(n)}{c(n)}\right\rceil = 1$. The mean number of routes passing through a cell is $\Theta(\log n)$ for $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1)$, whereas it is of the order $\Theta(\overline{L}(n)\sqrt{\log n})$ for $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1)$.

To prove Lemma 2.3 rigorously, we proceed as in [30] to compute the number of routes passing through a cell. Let $Z_i$ be a Bernoulli random variable indicating if the cell is used by the S-D pair $i$ to relay packets to the destination. Then, the number of routes passing through any cell is:

$$L_n = \sum_{i=1}^{n} 1_{\{Z_i=1\}} \tag{2.9}$$

Moreover, we note that:

$$\begin{aligned}
\Pr(Z_i = 1) &= \Theta\left(\frac{\#\text{ of cells traversed by a route}}{\text{total }\#\text{ of cells}}\right) \\
&= \begin{cases} \Theta\left(\frac{\overline{L}(n)c(n)}{n}\right) & \text{if } \Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1) \\ \Theta\left(\frac{c(n)^2}{n}\right) & \text{if } \Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1) \end{cases} \\
&= p_n
\end{aligned} \tag{2.10}$$

We need to bound the actual number of routes going through any cell. Neglecting the edge effects, and using the fact that $L_n$ is a Binomial random variable with parameters $(p_n, n)$ (recall that we consider $n$ S-D pairs), we use a Chernoff bound to obtain, for $\delta > 0$, $t > 0$, $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1)$

$$\begin{aligned} \Pr\left(L_n > \delta \log n\right) &\leq \frac{E[\exp(tL_n)]}{\exp(t\delta \log n)} \\ &= \frac{\left(1 + (e^t - 1)p_n\right)^n}{\exp(t\delta \log n)} \\ &\overset{(a)}{\leq} \exp\left(np_n(e^t - 1) - t\delta \log n\right) \\ &\overset{(2.10)}{=} \exp\left(c(n)^2(e^t - 1) - t\delta \log n\right) \end{aligned} \tag{2.11}$$

where (a) is by using $(1 + x) \leq \exp(x)$. Taking $t = 1$, $\delta = 3e$ in Eq.(2.11), we obtain:

$$\Pr(L_n > 3e \log n) \leq \frac{1}{n^3} \tag{2.12}$$

Similarly for $\delta = \sqrt{3}e$, $t = 1$, $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1)$,

$$\Pr\left(L_n > \sqrt{3}e\overline{L}(n)\sqrt{\log n}\right) \leq \frac{1}{n^3} \tag{2.13}$$

We need to prove that the bounds on $L_n$ in Eq.(2.12), Eq.(2.13) hold for all cells with high probability as $n$ gets large. Let us call $E_i$ the event that the number of lines passing through cell $i$ does not exceed the bounds on $L_n$ in Eq.(2.12) and Eq.(2.13). Then,

$$\begin{aligned} \Pr\Big(\bigcap_{i=1}^{|\mathcal{C}_n|} E_i\Big) &= 1 - \Pr\Big(\bigcup_{i=1}^{|\mathcal{C}_n|} E_i^c\Big) \\ &\overset{(a)}{\geq} 1 - |\mathcal{C}_n| \Pr(E_i^c) \\ &\overset{(b)}{\geq} 1 - n\varepsilon(n) \\ &\to 1 \end{aligned}$$

where $\mathcal{C}_n$ is the set of all cells, (a) is from the union of events bound, (b) is from the fact that there are at most $n$ cells in the network and $\varepsilon(n)$ are the bounds Eq.(2.12), Eq.(2.13). An alternative proof of Eq.(2.14) may be obtained from the Borel-Cantelli Lemma by noticing that $\sum_{n=1}^{\infty} \frac{1}{n^3} < \infty$. We conclude that the number of routes passing through any cell does not exceed almost surely the bounds on $L_n$ in Eq.(2.12), Eq.(2.13). □

We are now ready to state the following result ($c'$, $c''$, $c_1$, $c_2$ are positive constants).

**Theorem 2.1.** *For a large ad hoc network of $n$ nodes, the scheme described above achieves a per-node throughput capacity (with high probability as $n$ gets large):*

$$\begin{aligned}
\lambda(n) &= \begin{cases} c' \frac{1}{\overline{L}(n)c(n)} & \text{if } \Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1) \\ c'' \frac{1}{c(n)^2} & \text{if } \Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1) \end{cases} \qquad (2.14) \\
&= \begin{cases} c_1 \frac{1}{\overline{L}(n)\sqrt{\log n}} & \text{if } \Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1) \\ c_2 \frac{1}{\log n} & \text{if } \Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1) \end{cases}
\end{aligned}$$

*Proof.* We recall that the throughput capacity is computed over all possible time-space scheduling of transmissions and paths. A per node throughput is called *feasible* if there exist satisfying time-space scheduling and routing paths. We denote by $\lambda(n)$ the maximum feasible throughput with high probability as $n$ gets large.

By Lemma 2.2, we guarantee a constant rate to all communications. Lemma 2.3 bounds the number of routes each cell needs to serve. By Lemma 2.1, each cell will contain at least one node to forward the packets of these routes. Due to the time division, each cell will be active every one of $K^2$ slots. Then each path is guaranteed, with high probability as $n$ gets large, a rate of $\frac{k_1}{K^2 L_n}$. Combining these results yields a proof of Theorem 2.1. □

Moreover one can notice that for $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1)$, the maximum power is on the order of the average power, mainly $\Theta(c(n)^\alpha) = \Theta\left((\log n)^{\frac{\alpha}{2}}\right)$; whereas for $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1)$, the maximum power is on the order $\Theta\left((\log n)^{\frac{\alpha}{2}}\right)$ and the average power is $\Theta\left(\overline{L}(n)^\alpha\right)$ showing the benefit of having a very local traffic pattern.

In [9], it was shown that any upper bound on the transport capacity for arbitrary networks under the protocol model is also an upper bound on the transport capacity for random networks under the physical model. For a domain of area $\Theta(n)$ applying the results of [9], the transport capacity is bounded as follows (we scale the upper bound by $\sqrt{\text{area}}$):

$$\lambda(n) n \overline{L}(n) \leq c_3 W n \text{ bit-m/s} \qquad (2.15)$$

where $c_3$ is a positive constant and $\overline{L}(n)$ is the average source destination distance. This leads to an upper bound on the per-node throughput of the order $O\left(\frac{1}{\overline{L}(n)}\right)$. The difference between the lower and upper bound is discussed in the following (Section 2.5.2).

### 2.5.2 Discussion on the Impact of Fading on the Connectivity and the Throughput

As in [9], the lower and upper bound do not coincide under the physical model. This gap between the lower and upper bound is closed in [35]. In [35], the authors note that the requirement of connectivity with high probability in random networks as prescribed in [9], requires higher transmission power at all the nodes, and results in excessive interference. This in turn lowers the throughput of random networks from $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ to $O\left(\frac{1}{\sqrt{n}}\right)$. Motivated by this, the authors in [35] propose using a backbone-based relaying scheme in which instead of ensuring connectivity with probability one, they allow for a small fraction of nodes to be disconnected from the backbone. The nodes in the backbone are densely connected, and can communicate over each hop at a constant rate. Such a backbone traverses up to $\Theta\left(\sqrt{n}\right)$ hops. The nodes that are not a part of the backbone, send their packets to the backbone using single hop communication. The authors then show that the interference caused by these long range transmissions does not impact the traffic carrying capacity of the backbone nodes. Finally, the authors show that the relaying load of the backbone determines the per-node throughput of such a scheme, and this results in an achievable per-node throughput of $\Theta\left(\frac{1}{\sqrt{n}}\right)$. The above approach of allowing a few disconnected nodes in the network has also been used in [36]. Another way to close the gap between the lower and upper bound is by including fading in the model. This idea is motivated by multi-user diversity framework [78]. Indeed, one can decrease the cell size $c(n)$ and hope that the fading will ensure a range of transmission on the order of $O(\sqrt{\log n})$. By decreasing $c(n)$ we relax the connectivity condition and we hope that all the nodes that are out of range can be reached by the increased transmit power due to fading. The decrease of $c(n)$ will lead to an increase in the per-node throughput as derived in Theorem 2.1. We need to ensure that in each cell, only nodes experiencing the best channel conditions (among all neighbors) will transmit to the neighboring cells. The investigation of this idea is left for future research.

### 2.5.3 Hybrid Wireless Networks

A hybrid wireless network is formed by placing a sparse network of base-stations (or access points, gateways) in an ad hoc network. These base-stations are assumed to be connected by a high bandwidth wired network, and only to relay the packets since they do not generate any data traffic themselves. In addition to $n$ nodes randomly located within a square of area $\Theta(n)$, $f(n)$ base-stations are regularly placed within the network area. These base-stations divide the network area in $f(n)$ squares that we call clusters. We have then a collection of $f(n)$ clusters, each of which has a base-station placed in the middle of it as shown in Fig.2.6. As stated be-

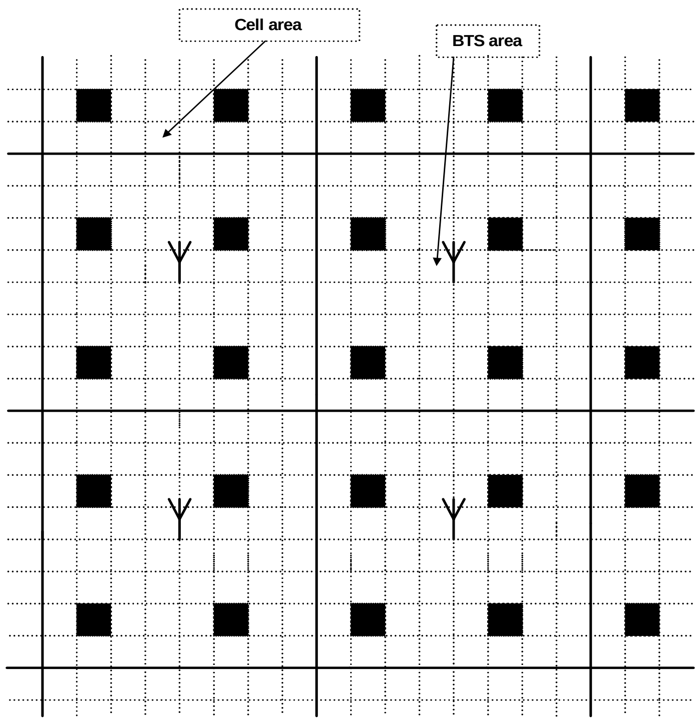


**Figure 2.6:** A hybrid wireless network with base-stations regularly placed in the middle of a BTS area (cluster).

fore, a base-station is never the initiator of a data transmission, but a relay that acts as a gateway between various clusters. Moreover, the infrastructure network is assumed to be an infinite capacity backbone and to have relatively abundant bandwidth and resources. The base stations are not power constrained and have the ability to reach any node within the cluster, whereas the nodes are power limited. A packet that reaches a base-station tunnels through another base station closest to the destination. Because of our subdivision of the network area in $f(n)$ mutually exclusive clusters, each wireless node is close to only one base-station. Within the same cluster, data transmissions are carried out without the use of the base-station. Data are forwarded from the source to the destination in a multi-hop fashion. Transmissions to nodes in other clusters are carried out by routing the data via the infrastructure (base-stations). Data are first transmitted from the source to the closest base-station (the base-station of the cluster) in a multi-hop fashion (i.e. in an ad hoc manner since nodes are power limited); the base-station then transmits the data through the wired infrastructure to the base-station closest to the destination, which finally transmits the data to the destination directly (since the base-station is not power constrained). The transmissions within any mode (ad hoc mode or infrastructure mode) do not interfere. The ad hoc mode and the infrastructure mode go through different sub-channels. Similarly, the infrastructure sub-channel can be divided into up-link and down-link parts. This means that the RF (radio frequency) is built such that an ad hoc transmitter is attached to each base station and that a BTS receiver is attached to each node.

In order to derive the throughput capacity of a hybrid wireless network, we use the deterministic scheme described above and the technical Lemmas derived in Section 2.4. We keep the same cell partition as described in Section 2.4. The network of ad hoc nodes, excluding the base-stations, is required to be connected since it is desirable to have an ad hoc network which can function without any infrastructure. The cell size was determined by the condition that no cell is empty as $n$ gets large, i.e., we have a standalone ad hoc network that can provide connection between any pair of ad hoc nodes without the support of any infrastructure. We do not change the transmission policy (each node in a cell can transmit to a node in the same cell or in the neighboring cells), therefore we do not require that each node is connected with high probability to a base-station. The latter will be reached in a multi-hop fashion (ad hoc mode). On the top of this partition, we add clusters, where each cluster contains a base-station and a number of cells that depends on the number $f(n)$ of base-stations in the network. We assume that $\Theta(f(n)) < \Theta\left(\frac{n}{c(n)^2}\right)$, otherwise we have a purely cellular system where each node can reach a base-station directly since each node will have a base-station within its range and the ad hoc transmissions (relaying done by nodes) are not needed (the distance from the source to the base-station is less than $c(n)$ the range of a node, this is the case for example when a base-station is placed within each cell of size $c(n)^2$). Since we are assuming a frequency division of intra-cell, up-link and down-link data transmissions, there is no interference between the three types of traffic. However, within a sub-channel, interference exists between the same type of traffic. Interference between adjacent clusters may be reduced by employing frequency reuse as in the case of a cellular network. For the ad hoc transmissions, we showed in Section 2.4 a spatial transmission schedule that ensures simultaneous transmissions with reduced interference. Actually, the cells are spatially divided into $K^2$ (a constant number) different groups. Each group is allocated a slot in a round robin fashion, and each cell will be able to transmit once every finite fixed amount of time with reduced interference. Then, Lemma 2.2 is still valid for our analysis of hybrid wireless networks.

From Theorem 2.1, the throughput capacity of a wireless ad hoc network is completely determined by the number of routes each cell needs to serve. Each cell relays the intra-cluster traffic (if the source and the destination are inside the same cluster, data transmissions are done in an ad hoc (multi-hop) fashion) and the traffic to reach the base-station (for a source wanting to communicate with a destination not belonging to the same cluster, data transmissions are sent in a multi-hop fashion to the closest base-station, which then routes data through the infrastructure to the destination). We are assuming now a uniform traffic pattern as in [9]. Sources and the corresponding destinations are randomly and independently placed in the network area. The probability that a node and its corresponding destination are located in the same cluster area is $\Theta\left(\frac{1}{f(n)^2}\right)$ (this happens with high probability as $n$ gets large). We conclude that the number of S-D pairs belonging both to the same cluster area (thus communicating in a pure ad hoc mode) is $\Theta\left(\frac{n}{f(n)^2}\right)$, whereas the number of S-D pairs communicating through the infrastructure is $\Theta\left(\frac{n}{f(n)}\left(1-\frac{1}{f(n)}\right)\right)$. Neglecting the edge effects and bottlenecks around base-stations, and using the results of Lemma 2.3, the number of routes passing through a cell is $\Theta\left(\frac{(\text{\#cells traversed by a route})\ (\text{\#S-D pairs})}{\text{\#total of cells}}\right)$. Assuming that the S-D or the S-BTS mean path is on the order $\Theta\left(\sqrt{\text{cluster area}}\right)=\Theta\left(\sqrt{\frac{n}{f(n)}}\right)$, the number of cells traversed by a route is $\Theta\left(\sqrt{\frac{n}{f(n)}}\frac{1}{c(n)}\right)$ and the total number of cells in a cluster area is $\Theta\left(\frac{n}{f(n)c(n)^2}\right)$. We obtain that the number of routes passing through a cell due to the intra-cluster traffic is $\Theta\left(\frac{c(n)\sqrt{n}}{f(n)^{\frac{3}{2}}}\right)$, whereas the number of routes due to the traffic S-BTS is $\Theta\left(\frac{c(n)\sqrt{n}(f(n)-1)}{f(n)^{\frac{3}{2}}}\right)$. We are now ready to state the following result whose proof stems from the technical Lemmas derived above.

**Theorem 2.2.** *For a hybrid wireless network of $n$ nodes and $f(n)$ base-stations regularly placed within the network area, and under the deterministic scheme and the routing strategy described above, the per-node throughput capacity is:*

$$\lambda(n) = c_4 \frac{f(n)^{\frac{1}{2}}}{c(n)\sqrt{n}} = c_5 \sqrt{\frac{f(n)}{n \log n}} \tag{2.16}$$

*with high probability as $n$ gets large ($c_4$, $c_5$ are positive constant).*

Suppose now that $\Theta f(n) = \Theta\left(\frac{n}{c(n)^2}\right) = \Theta\left(\frac{n}{\log n}\right)$, this corresponds to the case where each cell contains a base-station. Long-distance relaying is performed by the infrastructure (no need for ad hoc mode). Since the number of nodes per area is constant (fixed density), we have $O(\log n)$ nodes per cell. We can schedule each node in the network without any conflict by a schedule of length $3K^2 \log n$, and the per node throughput is of the order $\frac{1}{\log n}$. Eq.(2.16) gives the same result. Remember that in the case of a local traffic pattern where the mean path length is less than $c(n)$, a similar per node throughput was obtained. This is mainly due to the fact that the scenario is essentially a set of point-to-point communication systems, where data transmissions are unlikely to use relaying to reach the destination. The only way to increase the throughput is then to reduce the cell size (in the case of a local traffic pattern).

Moreover for a cellular system with $f(n)$ base-stations $\left(\Theta(f(n)) < \Theta\left(\frac{n}{c(n)^2}\right)\right)$ where each node communicates directly with a base-station (no ad hoc mode), the average power of each node is of the order $O\left(\left(\frac{n}{f(n)}\right)^{\frac{\alpha}{2}}\right)$ (where $\sqrt{\frac{n}{f(n)}}$ is the mean source base-station distance in the setting described above), this is higher than the average power for a hybrid wireless network with the same number of base-stations and under the same setting where the average power of each node is $O\left((\log n)^{\frac{\alpha}{2}}\right)$, showing the benefit of a hybrid wireless network over a purely cellular system.

### 2.5.4 Discussions on Event-Driven Traffic Patterns for Wireless Sensor Networks

The motivation behind this section comes from the fact that many sensor networks are event-driven and have spatially-correlated transmissions. Indeed, most sensor networks can be grouped in subsets since they are deployed in the same geographical area. This is done to increase reliability and fault-tolerance.

Consider a wireless sensor network formed by $n$ nodes and one collector over an area of $\Theta(n)$. This could be easily extended to the case of say $h(n)$ collectors. Indeed, for simplicity we make some assumptions for the particular scenario of interest, but this analysis could be extended to other scenarios. Among these assumptions is the fact that we consider that nodes have an infinite buffer and events happen continuously. No burstiness or periodicity of events is considered for the moment, this leads to a continuous sensing and transmission of nodes. The creation of packets occurs at a regular and deterministic rate. The analysis of the impact of buffer size, queueing delays and times of arrival of events are beyond the scope of this section, but note that this is a direction one should consider for completeness. We assume that not all the nodes that sense an event need to report it. By using the cell partitioning of Section 2.4, only one node in the vicinity of an event in a cell $c^*$ (e.g. an event that happens in a cell $c^*$) is elected to report this event. One can imagine a transmission policy in which nodes cooperate by exchanging informations about a given event and decide which node is to report it. This is motivated by the fact that all the observations made by the nodes in $c^*$ are highly correlated, and in order to decrease traffic and concurrent transmissions only one session from a source node (in the vicinity of a given event) to the collector is allowed. It is then convenient to assume that all the events are distributed uniformly on the area of the network or by a homogeneous Poisson point process, such that for any region $\mathbf{R}$ of area $A(\mathbf{R})$, the number of events in the region has a Poisson distribution with parameter $\sigma A(\mathbf{R})$, i.e.,

$$\Pr[k \text{ in } \mathbf{R}] = e^{-\sigma A(\mathbf{R})} \frac{(\sigma A(\mathbf{R}))^k}{k!} \tag{2.17}$$

The main parameter is the the event density given by,

$$\sigma(n) = \frac{\#\text{of events}}{\text{area}}$$

The number of events is closely linked to the geographical area $\Theta(n)$. Thus, the large network model (fixed density of nodes) is suitable for the state of this application. If we increase the network area, the events to be reported will increase. An example application could be intelligent transportation systems [43], where vehicles detects dangerous situations in traffic. The number of events in this example clearly increases with the area covered by the roads. Moreover, for some applications, we may resort to a non-homogeneous Poisson point process where the density depends on the geographical location, mainly when the event intensity is higher in some regions. This will make the analysis more complex since the traffic and relay load will depend on the geographical region. To simplify the analysis, one can study the worst-case scenario, in which the throughput is determined by considering that all traffic loads are equal to the load generated by the region with the highest density. In the following, the only key parameter that matters is the event density. It will allow the computations of the number of source collector S-C pairs, and the relay traffic and load produced by reporting these events. Finally, we assume that all the events are detected correctly by the nodes. For future directions, one could think of extending this model in order to take into account the probability of correct detection of an event.

Again in order to derive the throughput capacity of an event-driven model for wireless sensor networks, we use the deterministic scheme described above and the technical Lemmas derived in Section 2.4. We keep the same cell partition as described in Section 2.4. The connectivity condition on the cell size is kept the same as above. Remember, the cell size was determined by the condition that no cell is empty as $n$ gets large. We do not change the transmission policy (each node in a cell can transmit to a node in the same cell or in the neighboring cells), and we assume that the traffic overhead between nodes in a given cell induced by the election of a source of event reports is negligible compared to the traffic source collector. One can suppose that a sub-channel is dedicated for intra-cell communications in order to avoid interference. With this model, one can show that the number of source-collector pairs is of the order of the number of events occurring in a network of area $\Theta(n)$, i.e. $O(n\sigma(n))$. The throughput is determined by the relay load, and we need to compute the number of routes going through each cell. Since the collector is placed in the center of the network area, the source collector mean path distance is of the order $\Theta(\sqrt{n})$. Using the results of Lemma 2.3, the number of routes passing through a cell is:

$$\Theta\left(\frac{(\#\text{cells traversed by a route})\ (\#\text{S-C pairs})}{\#\text{total of cells}}\right) = \Theta\left(\sigma(n)\sqrt{n\log n}\right)$$

Thus, the event reporting throughput $\lambda(n)$ achieved by the model and under the deterministic scheme and routing strategy described above is:

$$\lambda(n) = \Omega\left(\frac{1}{\sigma(n)\sqrt{n\log n}}\right) \tag{2.18}$$

This result is quite encouraging as an initial derivation. It means that for an event density of the order of $O\left(\frac{1}{\sqrt{n}}\right)$, a per-event throughput of $\Omega\left(\frac{1}{\sqrt{\log n}}\right)$ is achievable. This is quite realistic, it means that the number of events needed to be reported is of the order of $O(\sqrt{n})$, where $n$ is the number of nodes in a network of area $\Theta(n)$. Similarly, for an event density of the order of $O\left(\frac{1}{\sqrt{n\log n}}\right)$, a per-event throughput of $\Omega(1)$ is achievable. The result in Eq.(2.18) could also be written in a large network of area $\Theta(n)$ as:

$$\lambda(n) = \Omega\left(\frac{\sqrt{n}}{\#\text{of events}\sqrt{\log n}}\right) \tag{2.19}$$

The increase of throughput is due to the reduction of traffic load in the network since we have $O(\sqrt{n})$ communications instead of $O(n)$. This result could even be improved, if we introduce some randomness in the arrival of events, in order to model the rarity of events and the fact that the nodes do not have to transmit all the time.

## 2.6 Conclusions

Following [9], we constructed a scheme that achieves the throughput capacity of a large ad hoc wireless network with high probability as the number of nodes increases. The proofs are made simple and more intuitive (we do not resort to the Vapnik-Chervonenkis Theorem for example), and we were able to study the asymptotic behavior of ad hoc wireless networks under a local traffic pattern as well as hybrid wireless networks.

For a local traffic pattern we showed the impact of the mean S-D distance on the throughput. Moreover there is a limit in the throughput improvement as the mean path length becomes smaller than the cell side length. For a local traffic pattern for which $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) \leq \Theta(1)$, we obtain a per-node throughput larger than $\frac{1}{\log n}$, whereas it is larger than $\frac{1}{\overline{L}(n)\sqrt{\log n}}$ for $\Theta\left(\frac{\overline{L}(n)}{c(n)}\right) > \Theta(1)$. It seems that the way to increase the throughput capacity is by relaxing the connectivity condition, which can be achieved by decreasing $c(n)$, the cell side length.

In this chapter, we have demonstrated also the benefits of using a hybrid wireless network in terms of per-node capacity. The base-stations are regularly placed within the network area, and the analysis is based on the subdivision of the network into $f(n)$ clusters, where $f(n)$ is the number of base-stations in the network. Moreover, the infrastructure network is assumed to have abundant bandwidth and resources. Inside each cluster, the communications are done in a pure ad hoc mode, whereas if the source and the destination do not belong to the same cluster, packets first reach the base-station in a multi-hop fashion and tunnel through the infrastructure to the base-station nearest to the destination. We obtain a per node throughput larger than $\sqrt{\frac{f(n)}{n\log n}}$(e.g. for $f(n) = (\log n)^2$, we obtain $\sqrt{\frac{\log n}{n}}$). The gain in performance is mainly due to the reduction in the mean number of hops from the source to the destination.

Finally, an initial step on the study of event-driven traffic patterns and their impact on the throughput expressions is conducted. The results are promising, and further developments and extensions of the model considered are left for future research.

## Appendix 2.A Bound on the Interference

Let us bound the interference. Consider a particular cell $c^*$. If one node from this cell is transmitting, all other simultaneous transmissions may occur in cells belonging to the same set of cells that are at a vertical and horizontal distance of exactly some multiple of $K$. Actually, the interfering cells are placed along the perimeter of concentric squares, whose center is $c^*$, and each square contains $(2lK+1)^2, l = 1, 2..., S(n)$ cells and $2lK, l = 1, 2..., S(n)$ interfering cells as depicted in Fig.2.3, where $S(n)$ is the number of such concentric squares. For example, take the particular case where $K = 4$, the first concentric square contains 8 interfering cells, whereas the second concentric square contains 16 interfering cells. Each node in the intended cell $c^*$ transmits information packets to nodes in the eight neighboring cells. Then, the distance between these nodes (the possible receivers in the eight adjacent cells) and the interfering ones is at least $l(K-2)c(n), l = 1, 2..., S(n)$. As we are considering a lower bound, we take the worst-case and neglect edge effects. Then, the number of concentric squares (irrespective of the position of the intended cell, since the worst case is when the intended cell is at one corner of the area) is at most $S(n) \leq \left\lceil \frac{\sqrt{\frac{n}{\log n}}}{K} \right\rceil$. We proceed in upper bounding the interference at the receiver:

$$
\begin{aligned}
I &= \sum_{\substack{k \neq i \\ k \in \mathcal{T}}} P|X_k - X_j|^{-\alpha} \\
&\leq \sum_{i=1}^{S(n)} \frac{2PKi}{[i(K-2)c(n)]^\alpha} \\
&= \frac{2PK}{[(K-2)c(n)]^\alpha} \sum_{i=1}^{S(n)} i^{1-\alpha} \\
&\leq \frac{2PK}{[(K-2)c(n)]^\alpha} \left[1 + \int_1^{S(n)} x^{1-\alpha} \mathrm{d}x\right] \\
&= \frac{2PK}{[(K-2)c(n)]^\alpha} \left[1 + \frac{1}{2-\alpha}\left(S(n)^{2-\alpha} - 1\right)\right] \\
&\overset{(\alpha \geq 2)}{=} \frac{2PK}{[(K-2)c(n)]^\alpha}\left(\frac{\alpha-1}{\alpha-2}\right) + S(n)^{2-\alpha} \frac{2PK}{[(K-2)c(n)]^\alpha} \frac{1}{2-\alpha} \\
&\overset{(\alpha > 2)}{\leq} \frac{cPK(\log n)^{-\frac{\alpha}{2}}}{(K-2)^\alpha}
\end{aligned} \tag{2.20}
$$

where $c$ is a positive number. The thermal noise $N_0$ is negligible as $n \to \infty$, and by combining Eq.(2.7) with Eq.(2.20), the $\mathrm{SINR}(n)$ is lower bounded by $\mathrm{SINR}_{\min}(n)$ which is a constant and as $C_{ij}(n) = W \log_2(1 + \mathrm{SINR}_{\min}(n))$, we obtain the result Eq.(2.6).

# Chapter 3

# Connectivity Graph and Conditions for Constant Throughput in Wireless Ad Hoc Networks

## 3.1 Introduction

In the area of wireless networks that operate in ad hoc mode, a question that has recently attracted significant research interest and activity, is the derivation of scaling laws for the capacity. This activity was sparked by the seminal work of [9], who proposed to model wireless ad hoc networks as random geometric graphs and examined the scaling of the long-term averaged throughput with the number of nodes $n$. Their main result is that, given $\Theta(n)$ randomly selected source-destination pairs, the throughput per source-destination pair scales at least as $\Theta\left(\frac{1}{\sqrt{n \log n}}\right)$ and at most as $\Theta\left(\frac{1}{\sqrt{n}}\right)$ for random networks, where nodes are placed uniformly at random on a given network area. The achievability is proved in the asymptotic sense by designing proper routing and transmission mechanisms. An insight from [9] is that, to maximize the throughput, we need to minimize the transmission power (range) of each node, while still keeping the network connected. That is, we need to reduce the interference region of each transmission and schedule as many non-interfering concurrent transmissions as possible.

This work was extended by a number of works that established scaling laws assuming different network models. For example, in the model of [9], the geographical area is fixed and the density of nodes is increasing with the number of nodes (dense network). In [35], techniques from percolation theory are used to show that in dense random networks a $\Theta(1/\sqrt{n})$ throughput is achievable. A different network model assumes that the number of nodes is increasing with the area of the network leading to a fixed density of nodes per area. For this case, in [37] and in [25], information theoretic upper bounds on the rate of communication are derived, as a function of the value of power loss exponents. Other follow-up works include [38, 39, 40, 41]. A result common to the different models is that, for fixed nodes networks, the throughput per source-destination pair vanishes as the number of nodes grows.

On the contrary, if nodes are allowed to move, a constant $\Theta(1)$ throughput can be achieved per source-destination pair even if the number of nodes grows to infinity [26]. This result assumes a 2-dimensional mobility pattern, where the trajectory of each node is an independent, stationary and ergodic random process with uniform distribution on the unit disk. That is, the mobility pattern is homogeneous with respect to each node, and the sample path of each node covers all the space over time. In the analysis, transmission is restricted to the closest nodes, and at each given instance, each link between any two nodes is activated with probability $\Theta(1/n)$. Then, a two-phase scheduling policy is employed. In the first phase, source nodes transmit the packets to the closest receiver (which can be a relay or a destination node) and in the second phase transmitters forward the packets that have as destination their closest receiver. Thus, for any source destination pair, $(n-2)$ relay nodes receive and transmit packets at rate $\Theta(1/n)$ while source nodes also transmit directly to the destination at a rate $\Theta(1/n)$. Note that the flow between each source-destination pair sums up to a fixed rate $\Theta(1)$. It is clear that network capacity can be drastically improved when mobility is effectively exploited.

A natural question to ask is whether this good performance is specific to the particular generous mobility pattern, or whether it can be achieved under more restricted mobility conditions. The work in [42] made progress in answering this question, by demonstrating that the same order of throughput can still be achieved under restricted 1-dimensional mobility. In their mobility pattern, each node is restricted to move on a randomly and independently chosen great circle on the unit sphere. However, the general question still remains, which is, under what conditions on the mobility patterns of the users a throughput of $\Theta(1)$ is achievable.

Recently, another model of restricted mobility has been examined in [51], where nodes are confined to overlapping neighborhoods, and the throughput scaling as a function of the neighborhood size is analyzed. In [54], the impact of the mobility pattern on the relay throughput (i.e. the maximum rate at which a node can relay data from the source to the destination) is studied. It is shown that the relay throughput depends on the node mobility pattern only via its stationary node position distribution and that a node mobility pattern that results in a uniform steady-state distribution for all nodes achieves the lowest relay throughput.

Independently, in computer science, the problem of multi-commodity flow has received significant attention. In this problem, we are given a graph with edges of a fixed capacity, a set of source-destination pairs, each with its own demand, and we are asked to maximize the amount of information flow that can be simultaneously routed for all source-destination pairs. The emphasis of the work in the literature is in deriving min-cut bounds on the achievable rates and characterizing under what conditions these bounds are tight, see for example [44, 45, 46]. Recently, the area of network coding has emerged, where it is demonstrated that by allowing flows to mix, we may achieve significant throughput benefits for the multi-commodity problem over directed graphs [47]. At this time, the prevailing conjecture is that network coding does not offer benefits for the case of the undirected multi-commodity problem [48]. For our results we will not use network coding techniques, i.e., we will assume that nodes can only forward and not combine their incoming information flows.

In this chapter, our interest is in mobile ad hoc and wireless sensor networks. Our main contribution is a method that allows to check whether, for a given mobility pattern, a constant $\Theta(1)$ throughput per source-destination pair is possible. Intuitively, the reason we get a decreased throughput in fixed nodes networks, is that the average number of hops that a packet needs to traverse from a source to a destination scales with $n$ (the number of nodes in the network). On the contrary, in [26, 42], mobility enables a routing strategy where the number of hops is limited to at most two. Our work is motivated by the observation that in a complete graph between each source-destination pair there exists a set of max-flow paths whose length is upper bounded by two. This set consists of $n-2$ paths of length two and one path of length one. A graph is complete if any two nodes are connected with an edge, which intuitively seems to correspond to the fact that in the uniform mobility pattern any two nodes can become neighbors and successfully transmit to each other. In other words, the routing approach in [26, 42] seems to exactly correspond to the routing approach we would use to solve the multi-commodity flow problem on a complete graph of equal capacity edges.

To make this loose connection precise, we first decompose the communication problem into a "transmission policy" and a "scheduling policy". As we argue in Section 3.3, this decomposition does not affect whether constant throughput is achievable asymptotically. We then introduce the *connectivity graph*, that does not represent the actual physical network, but rather the available communication resources, for a given transmission policy. The connectivity graph offers an abstraction of the communication capabilities of the ad hoc network: we can study the long-term averaged throughput between source-destination pairs in the actual ad hoc network, by examining information flows in the connectivity graph. Thus, by mapping the ad hoc network problem into a graph problem, we establish a bridge between the multi-commodity flow and the ad hoc network literature, that can be used in both directions.

The focus of this chapter is in using the properties of the connectivity graph to develop a set of necessary and sufficient conditions under which constant $\Theta(1)$ throughput is possible. We illustrate how these conditions apply to a number of different topologies, including the topologies in [9, 26, 42]. We then try to understand what structural properties these conditions imply for the connectivity graph and how they translate into properties for the underlying mobility pattern. Interestingly, we provide an example where constant throughput may be possible to achieve by mobile nodes that have a restricted number of neighbors. That is, each node may successfully communicate with at most $n^{1/t}$ other nodes, for a finite $t$.

Although we focus on scaling results when the number of nodes of the network increases, the same approach can be used to analyze throughput and design routing over finite networks. In fact, independent of our work [49], the authors in [50] proposed the use of a structure similar to our "connectivity graph" to optimize scheduling in finite size networks.

The chapter is organized as follows. Section 3.2 formally defines our problem, and Section 3.3 introduces the decomposition into a transmission and a routing policy. Section 3.4 introduces the concept of connectivity graph. Section 3.5 presents the proposed set of conditions. Section 3.6 applies these conditions to different topologies. Section 3.7 investigates what structural properties of the connectivity graph are necessary to achieve constant throughput, and Section 3.8 concludes the chapter.

## 3.2 Problem Statement

Closely following the model in [9, 26, 42], we consider a wireless ad hoc network with $n$ nodes, that can act as transmitters and receivers. Nodes may be mobile, according to a given mobility pattern. For example in [9], nodes are fixed, while in [42] nodes move on great circles. We randomly choose $\Theta(n)$ source-destination pairs $\{s_i, t_i\}$. We assume discrete slotted time, where time is divided into equal duration slots. During each time slot, one information unit can be transmitted from each transmitter. We also assume point-to-point communication where, during a successful transmission, a transmitter can convey information to exactly one receiver.

Our main interest is in deriving asymptotic results when the number of nodes $n$ increases. We distinguish between the terms "transmitter-receiver" that may refer to intermediate nodes acting as relays along an information flow path, and the terms "source-destination" that refers to the pair of end nodes wishing to exchange information. Each source has an infinite amount of information to convey to its destination. Each node has possibly infinite buffering capabilities. The source may transmit information to the destination either directly, or by routing the information through a number of intermediate relays before attaining the final destination.

Let $\lambda(n)$ be the long-term averaged throughput we can guarantee to $\Theta(n)$ source-destination pairs $(s_i,\ t_i)$, $i = 1 \ldots \Theta(n)$, measured in information units per time slot. We will use the following definitions.

**Definition 3.1.** *A throughput of $\lambda(n)$ is feasible if there exists a spatial and temporal scheme for scheduling transmissions, such that by operating the network in a multi-hop fashion and buffering at intermediate nodes when awaiting transmission, a source can send $\lambda(n)$ information units per time* *slot on average to its chosen destination node. That is, there is a* $T < \infty$ *such that in every interval of* $[(i-1)T, iT]$ *time slots, every source can send* $T\lambda(n)$ *information units to its corresponding destination.*

The following is a simplified definition of the aggregate throughput [9], which is sufficient for our purposes, since we assume that each transmission conveys exactly one unit of information.

**Definition 3.2.** *At a given time-slot, the aggregate throughput* $\Gamma$ *is equal to the total number of concurrent successful transmissions.*

Whether an attempted transmission is successful depends on factors such as the underlying channel model, and the interference from other transmissions. For example, in [9] the authors considered two models for determining the success of a transmission over the shared wireless channel:

1. The *protocol model*, where a given transmitter-receiver pair successfully communicate if no other node transmits within a disk area centered at the intended receiver. The radius of the disk depends on the distance between the transmitter and the receiver and on a protocol dependent constant;

2. The *physical model*, where a successful transmission occurs when the signal to interference and noise (SINR) ratio at the receiver is above a certain threshold.

Our goal is to determine, whether for a given successful communication model and a given mobility pattern, a throughput of $\lambda(n) = \Theta(1)$ is feasible. We will approach this goal through the following steps.

1. We will decompose the communication problem into a transmission strategy and a scheduling policy. We will argue that this decomposition does not affect whether $\lambda(n) = \Theta(1)$ is feasible (Section 3.3).

2. We will construct a connectivity graph, that embodies a given transmission policy, and thus, transforms our wireless network problem to a multi-commodity problem over a graph (Section 3.4).

3. We will derive necessary and sufficient conditions for $\lambda(n) = \Theta(1)$ using the connectivity graph (Section 3.5).

In fact, these steps define a methodology, which we can follow to determine whether it is possible to achieve constant throughput over an ad hoc network with a given transmission policy and mobility pattern. We will follow this methodology over several examples in Section 3.6.

## 3.3 Decomposition

The described model employs two random processes, one referring to the placement of the nodes at each time slot, and the other to the choice source-destination pairs.

- *Node placement.* This random process determines, for any given node of the network the probability of being at a specific area of the network. In a static network, the nodes are randomly placed on the unit area disk. From then on, for all time slots, we have the same fixed realization of the node placement random process. In a mobile network, we have a mobility model that we assume can be described as a random process. [1]

- *Source - Destination Pairs.* The second random process pairs sources with destinations uniformly at random.

Our results apply only in the case where these two random processes are independent, that is, the source-destination pairs are chosen independently from the node placement.

In this chapter, we will also assume that all nodes have the same transmission range. Commercial ad hoc and sensor networks are typically simple identical devices, subject to power constraints. Thus the assumption that all nodes are identical devices that can transmit at the same range is a typical one in the literature. Moreover, in order to analyze asymptotic results, and since the source-destination pairs are chosen randomly, we will mainly examine mobility models that have some form of uniformity from the nodes point of view. That is, there exists a common (possibly parametrized) process that describes the mobility of all nodes. For such models these is no loss of generality in assuming common transmission range.

[1]Deterministic mobility models can also be included in this category by assigning probability one to a desired pattern.

**Lemma 3.1.** *A necessary condition to achieve throughput* $\lambda(n) = \Theta(1)$ *is that the average (over time) aggregate throughput* $E(\Gamma)$ *satisfies*

$$E(\Gamma) = \Theta(n). \tag{3.1}$$

*Proof.* If $\lambda(n) = \Theta(1)$, then $\Theta(n)$ source-destination pairs achieve an average throughput of $\Theta(1)$ information units per time slot. Thus on the average, at least $\Theta(n)$ information units have to be exchanged per time-slot, i.e., $E(\Gamma) \geq \Theta(n)$. But, since we consider point-to-point communication, and we have at most $n$ nodes transmitting at each given time slot, $\Gamma \leq \Theta(n)$. Thus, we conclude that $E(\Gamma) = \Theta(n)$ information units have to be successfully transmitted on the average per time-slot. □

**Lemma 3.2.** *(From [9]) Given a network with n nodes, under the protocol and the physical model, the transmission policy that maximizes* $\lambda(n)$ *is one where nodes transmit only to their closest neighbors at the smallest range that allows the network to be connected. The typical range is of the order of* $O\left(\frac{1}{\sqrt{n}}\right)$ *in a network of n nodes of unit area.*

*Proof.* A formal proof is provided in [9], here we briefly outline the intuition. Let $\overline{P}$ be the mean distance traversed by an information unit from a source to the destination, and let $r$ be the common transmission range (which is proportional to the transmit power), and each node has a randomly chosen destination to which it wishes to send $\lambda(n)$ bits/s. Then, each information unit reaches the destination after $\frac{\overline{P}}{r}$ hops. This creates $\frac{\overline{P}\lambda(n)}{r}$ units of traffic, and if each link is capable of $W$ information units per time slot, we have that $\lambda(n) \leq \frac{Wr}{\overline{P}}$. The right-hand side is proportional to the transmission range $r$, so it appears that increasing $r$ should increase throughput. However, increasing the range increases the interference, since each transmission causes interference within a disc of area $\pi r^2$. For example, nodes close to the receiver are required to be idle to avoid collision.

Increasing the range $r$ leads to loss (due to interference) quadratic in $r$, while decreasing the range $r$ leads to loss (due to number of hops) linear in $r$. Hence the optimal transmission policy is to reduce the range of transmission $r$ to minimize interference, while still keeping the network connected. □

Results in [41] also prove the optimality, in the scaling law sense, of multi-hop communication strategies for maximizing the transport capacity.

Lemma 3.1 is straightforward, and Lemma 3.2 effectively reviews results from [9] for completeness. Together, they imply that there exists exactly one [2] transmission policy, that may allow to achieve $\Gamma = \Theta(n)$ and thus constant throughput, namely, attempting to transmit only to the closest neighbors. The transmission policy can also be thought as the transmission range and power the nodes are allowed to employ. Thus, we argue that, under our assumptions, we can decompose the communication problem into two parts,

- a transmission policy, and
- a scheduling policy,

and that this decomposition does not affect whether we can achieve constant throughput. For the rest of the chapter, we will assume that the transmission policy and the scheduling policy are optimized independently. The scheduling policy determines how the information is routed from the sources to the destinations, takes into account scheduling of competing nodes, dictates which nodes are allowed to transmit, which packets they will transmit and to whom. More specifically, the scheduling policy determines, at a given time-slot, what information units will be conveyed between the transmitter-receiver pairs, that are allowed to communicate under the transmission policy.

Finally, note that, although this decomposition does not affect whether we can or cannot achieve constant throughput, it does affect other system parameters, such as delay. In this chapter we exclusively focus our attention on long-term averaged throughput.

## 3.4 Connectivity Graph

A transmission policy determines during each time slot a set of transmitters that attempt to transmit to a set of receivers. The attempted transmissions are successful or not, according to the underlying model for successful communication. For example, an attempted transmission might fail due to interference from other transmissions, or because the transmitter happens to be out of range from other nodes. As a result, during each time slot, we get a configuration of simultaneously successful transmissions between

[2] Perhaps more than one, but all are equivalent from the average point of view.

transmitter-receiver pairs, which we call a *communication pattern*. This policy takes into account scheduling of competing nodes and transporting data from the source to the destination across a sequence of links. In the following we do not focus on a particular routing/physical policy, we rather assume that the latter is given and can be optimized depending on the application and can be chosen independently from the transmission policy. For example in [26], $\Theta(n)$ nodes at each time slot attempt to transmit to their nearest neighbor, and the transmission is deemed successful if the signal to noise and interference power is above a threshold. Fig. 3.1 depicts possible such communication patterns during discrete time slots.

Because we are interested in long-term averaged throughput, and we assume that nodes have infinite buffering capabilities, if we were to randomly permute the different time slots (communication patterns) the throughput we could achieve would not be affected. In other words, the time-correlation between communication patterns, and their evolution in time, does not affect the long-term average throughput. In fact, for a given transmission policy, we can construct a single graph, that summarizes all the information we need in order to calculate the achievable throughput.

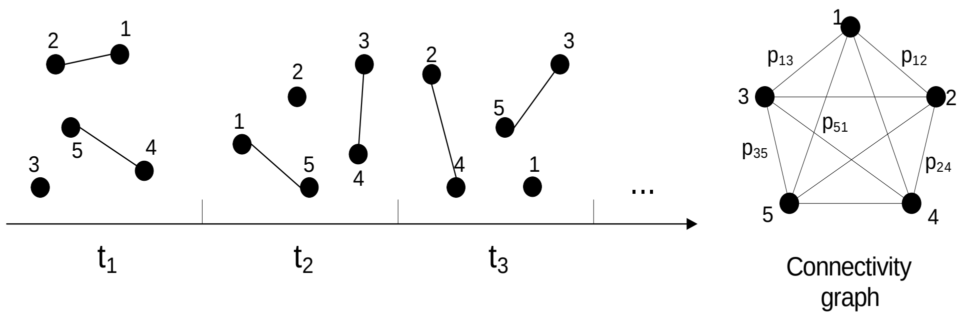


**Figure 3.1:** An ad hoc network with $n = 5$ nodes. During each time slot a different communication pattern occurs.

To do so, we assume that that there exists an ergodic and stationary process which determines at each time slot which nodes can successfully communicate. This random process is a high-level/end-result model that incorporates the mobility model, the physical model of the network, the criterion for successful transmission between two nodes, generally all factors that affect the success of a transmission.

At each time slot, the random process realization corresponds to a communication pattern, i.e., the set of point-to-point simultaneous successful transmissions during that time slot. For example in Fig. 3.1, during each time slot one such communication pattern occurs.

In other words, for every two nodes $i$ and $j$ and at each time slot $t$, we observe the realization of an indicator function $I_t(i, j)$. $I_t(i, j)$ indicates whether the connection (edge) between two nodes exists or not, whether the nodes can directly and successfully transmit to each other.

We construct a *connectivity graph* that connects every two nodes $i$, $j$ with an edge $e = (i, j)$ of capacity $C(e) = p_{ij}$, where

$$p_{ij} = \lim_{n\to\infty} \frac{\sum_{t=1}^{n} I_t(i,j)}{n}.$$

The capacity $p_{ij}$ expresses the fraction of the time slots nodes $i$ and $j$ are "connected", i.e., can directly and successfully transmit to each other. Thus, $p_{ij}$ upper bounds the fraction of information units per time slot that nodes $i$ and $j$ can exchange on the average. Since we assumed an ergodic and stationary random process, and for $n \to \infty$, we will, with probability going to one, observe the same values for the connectivity graph.

The connectivity graph *does not* represent the actual ad hoc network at any given time slot, but rather summarizes the long-term average communication capabilities of the network. We emphasize that this graph does not express either the correlation between successive instantiations of communication patterns, or the correlation among successful transmissions at any communication pattern, but rather offers an abstraction of the communication capabilities of the original ad hoc network, that we can use to study long-term averaged throughput.

For example, in [26] it is proved that for the uniform mobility pattern, at a given time slot, any pair of nodes $i$ and $j$ will be close to each other and will successfully exchange one unit of information with probability $\Theta(1/n)$. That is, over a number of time slots $N >> n$, nodes $i$ and $j$ will exchange on the average $\Theta(\frac{N}{n})$ information units. Thus the connectivity graph associated with this model is a complete graph, where each edge $(i, j)$ has capacity $p_{ij} = \Theta(1/n)$. This graph captures this fact, that node $i$ will be able to successfully send one information unit to every other node $j$, on the average once every $n$ time slots. Thus, node $i$ will be able on the average to directly transmit to node $j$ throughput of $\Theta(1/n)$ information units per time slot.

In practice, given a mobility pattern and transmission policy that can be described through a deterministic or probabilistic model, we can in general easily construct the corresponding connectivity graph, as we will demonstrate for a number of examples in Section 3.6. We can also empirically construct the connectivity graph by actually recording the number of successful transmissions between two pairs of nodes over time.

Note that we constructed the connectivity graph as an undirected graph, by implicitly assuming a symmetrical transmission policy. That is, at a given time-slot, if node $i$ can successfully transmit to node $j$, then node $j$ as well can successfully transmit to node $i$, and we can choose to schedule one of these two transmissions. This is not necessarily true for all transmission policies, and in full generality we would need to create a directed connectivity graph. However, in all examples in the literature, the transmission policies can be considered symmetric. Thus for simplicity in this chapter we will only consider undirected connectivity graphs.

It is easy to see that in the connectivity graph, the capacities associated with the edges adjacent to a node $i$, have to sum to a number smaller or equal to one. This is because we assumed "point-to-point" transmissions and not multicasting/broadcasting over the wireless medium, which implies that a node at each time slot can successfully transmit to at most one receiver.

**Lemma 3.3.** *For any node $i$ in the connectivity graph*

$$\sum_{e:\, adjacent\ to\ node\, i} C(e) \leq 1, \tag{3.2}$$

where $C(e)$ is the capacity of edge $e$. As a result, a source can transmit throughput at most $\Theta(1)$ towards its destination, and consequently the destination cannot possibly receive more, i.e.,

$$\lambda(n) \leq 1. \tag{3.3}$$

## 3.5 Throughput Analysis

In this section, we are going to study whether constant throughput is possible, by examining necessary and sufficient conditions on the associated connectivity graph.

### 3.5.1 Min-cut Bounds

By introducing the connectivity graph, we have effectively translated the problem of scheduling in an ad hoc network, to a multi-commodity flow problem over the connectivity graph. A routing policy in an ad hoc network, can be mapped to routing along paths in the connectivity graph. Thus, we can directly apply min-cut bounds from the multi-commodity flow literature, which we briefly review in the following.

Consider a graph $G = (V, E)$, and $k$ source-destination pairs $(s_i, t_i)$. Let $D_i$ denote the throughput demand between source-destination pair $(s_i, t_i)$. The *sparsest cut* $\mathcal{S}$ of an undirected multi-commodity flow problem is defined to be the minimum over all cuts of the ratio of the capacity of the cut to the demand of the cut [44, 45]. If a cut divides the vertices of the graph into sets $U$ and $V - U$, the capacity of the cut is calculated as the sum of the capacities of the edges that connect $U$ and $V - U$ (edges that have one endpoint in $U$ and the other in $V - U$):

$$C(U, V - U) = \sum_{e \in (U, V-U)} C(e). \tag{3.4}$$

The *demand* is defined as the sum of the demands of pairs $(s_i, t_i)$, where $s_i$ and $t_i$ are on opposite sides of the cut that separates $U$ from $V - U$, and are completely disconnected by the cut:

$$D(U, V - U) = \sum_{(s_i \in U \text{ and } t_i \in V-U) \text{ or } (t_i \in U \text{ and } s_i \in V-U)} D_i. \tag{3.5}$$

Then the sparsest cut equals

$$\mathcal{S} = \min_U \frac{C(U, V - U)}{D(U, V - U)}. \tag{3.6}$$

A multi-commodity flow is *feasible*, if

$$\sum_i f_i(e) \leq C(e), \tag{3.7}$$

where $f_i(e)$ is the flow between $(s_i, t_i)$ that is routed through edge $e$. A feasible flow satisfies the sparsity upper bound.

The sparsity bound generalizes the min-cut max-flow bound by Ford and Fulkerson, that applies for one source-destination pair ($k = 1$). Unlike the min-cut bound however, the sparsity upper bound is not always achievable. Known results include:

- For the uniform multi-commodity problem, where each pair of nodes form a source-destination pair, and where the demand is equal, i.e., $k = \binom{n}{2}$ and $D_i = \lambda(n)$, we can always achieve throughput within a factor of $O(\log n)$ from the sparsity upper bound [44].

- For the same problem over planar graphs, we can achieve the bound within a factor of $O(1)$.

### Application

In our case of interest, $D_i = \lambda(n)$, $k = \Theta(n)$, and we are interested in order arguments. Applying the sparsity bound we get the following lemma:

**Lemma 3.4.** *A necessary condition to achieve throughput $\lambda(n)$ is that, the capacity of any cut $(U, V-U)$ that separates $m$ sources from $m$ destinations, $1 \leq m \leq n$, has to be greater or equal to $\lambda(n)\Theta(m)$, i.e.,*

$$\lambda(n)\Theta(m) \leq \min_U C(U, V-U). \tag{3.8}$$

Thus, a necessary condition to achieve $\lambda(n) = \Theta(1)$, is that the capacity of the min-cut that separates the $m$ sources from $m$ destinations is greater or equal to $\Theta(m)$. We underline that Lemma 3.4 applies not only for $\lambda(n) = \Theta(1)$, but for all $\lambda(n)$ achievable under the assumption that the transmission and scheduling policy are optimized separately. This, as we discussed, is not restrictive for obtaining constant throughput.

### 3.5.2 Aggregate Throughput Bounds

A different type of bound can be derived from the observation that the total resources required from the $k$ source-destination pairs have to be smaller than the total resources available at the network.

For example, if we use a path of length two to route throughput say $\lambda(n)$, we need to use $2\lambda(n)$ of the network capacity resources to deliver this information from the source to the destination. Generally, let $\mathcal{P}_i = \{P_i^1, \cdots, P_i^{m_i}\}$ be a set of $m_i$ paths that convey flow

$$f_i = \sum_{j=1}^{m_i} f_i(P_i^j), \tag{3.9}$$

between the pair $(s_i, t_i)$, where $f_i(P_i^j)$ is the flow along the path $P_i^j$. Then the network resources required by the pair $(s_i, t_i)$ equal

$$R_i = \sum_{j=1}^{m_i} \text{length}(P_i^j) f_i(P_i^j), \tag{3.10}$$

where length$(P_i^j)$ is the length in number of edges that comprise path $P_i^j$. In total, all source-destination pairs, will need resources

$$R = \sum_{i=1}^{k} R_i = \sum_{i=1}^{k} \sum_{j=1}^{m_i} \text{length}(P_i^j) f_i(P_i^j). \tag{3.11}$$

These required resources cannot exceed the available network resources, i.e.,

$$R \leq C = \sum_{e \in E} C(e). \tag{3.12}$$

Eq. (3.12) is a direct application of Eq. (3.7), since we can get the same inequality as

$$\sum_{e \in E} C(e) \geq \sum_{e \in E} \sum_{i} f_i(e) = \sum_{i} \sum_{e \in E} f_i(e) = \sum_{i} \sum_{j} \text{length}(P_i^j) f_i(P_i^j). \tag{3.13}$$

## Application

In our framework, $C = \Theta(E(\Gamma))$, that is, the network resources available per time slot equal the average aggregate throughput $E(\Gamma)$, where averaging is over time. Moreover, we have that $\lambda(n) = \sum_j f_i(P_i^j)$ for all $i$. Let $p_{i,j}$ be the fraction of the flow $f_i$ routed through path $P_i^j$, i.e. ,

$$p_{i,j} = \frac{f_i(P_i^j)}{\lambda(n)}, \quad \sum_{j} p_{i,j} = 1. \tag{3.14}$$

Then Eq. (3.13) can be rewritten as

$$\lambda(n) \sum_{i} \sum_{j=1}^{m_i} p_{i,j} \text{length}(P_i^j) \leq \Theta(E(\Gamma)). \tag{3.15}$$

We can interpret $p_{i,j}$ as the probability that a given information unit sent by source $s_i$ to $t_i$ utilizes path $P_i^j$, and $\sum_{j=1}^{m_i} p_{i,j}\text{length}(P_i^j)$ as the average path length that the information units from source $s_i$ experience.

**Definition 3.3.** *We define the average path length $\bar{P}$, averaged over all source-destination pairs and all paths, to be the average path length that a source would utilize to route information, calculated as*

$$\bar{P} = \frac{\sum_{i=1}^{\Theta(n)} \sum_{j=1}^{m_i} p_{i,j} length(P_i^j)}{n}. \tag{3.16}$$

In Lemma 3.1, we observed that $E(\Gamma) \leq \Theta(n)$. Thus we get

**Lemma 3.5.** *A necessary condition to achieve throughput*

- $\lambda(n)$ *is that*

$$\bar{P} \leq \frac{\Theta(E(\Gamma))}{n\lambda(n)}, \tag{3.17}$$

- $\lambda(n) = \Theta(1)$ *is that*

$$\Gamma = \Theta(n), \text{ and } \bar{P} = \Theta(1). \tag{3.18}$$

This lemma makes more precise the intuitive connection between throughput and average path length discussed in the introduction. Indeed, from $\Gamma = \Theta(n)$ and Eq. (3.17) we get that

$$\lambda(n) \leq \Theta(\frac{1}{\bar{P}}). \tag{3.19}$$

This bound agrees with the result of [52] for local traffic patterns. The authors in [52] show that for such traffic patterns, and fixed networks, the per source-destination throughput scales as $O\left(\frac{1}{\bar{P}\sqrt{\log n}}\right) \leq \lambda(n) \leq O\left(\frac{1}{\bar{P}}\right)$ if $O(\bar{P})$ grows faster than $O(\sqrt{\log n})$, and scales as $O\left(\frac{1}{\log n}\right) \leq \lambda(n) \leq O\left(\frac{1}{\bar{P}}\right)$ if $O(\bar{P})$ grows slower than $O(\sqrt{\log n})$. Eq.(3.19) strengthens our connectivity graph approach to study Ad Hoc wireless networks. Hence, we are able to emphasize the main structural properties of such networks, and their impact on the throughput. Again Eq.(3.19) states that the main decrease in the throughput is due to the relay load which is proportional to the average path length or equivalently the number of hops a packet has to go through from source to destination.

### 3.5.3 Necessary and Sufficient Conditions for $\lambda(n) = \Theta(1)$

In this section we will derive necessary and sufficient conditions for constant throughput, under the assumption of *integral-edge routing*, in the sense that, if flow $f_i$ utilizes edges $e$ that has capacity $C(e)$, then $f_i(e) = \Theta(C(e))$.This assumption is not restrictive, since we can allow parallel edges, and decompose an edge of capacity $C(e)$ to $m$ edges $e_1, \dots e_m$ such that $\sum C(e_i) = C(e)$.

Consider a source-destination pair $(s_i, t_i)$. To achieve throughput $f_i = \lambda(n) = \Theta(1)$, source $s_i$ needs to transmit $\Theta(1)$ units of information towards destination $t_i$. Accordingly, in the connectivity graph, there need to exist a set of paths through which source $s_i$ will route information $\Theta(1)$ towards the destination $t_i$. Requiring additionally that, when the total network resources are shared by all source-destination pairs, the overall flow is feasible, we get the following theorem.

**Theorem 3.1.** *Consider the connectivity graph associated with a given wireless network, and a set of randomly selected source-destination pairs $(s_i, t_i)$. The following conditions are* necessary and sufficient *to achieve $\lambda(n) = \Theta(1)$ throughput.*
*For each $(s_i, t_i)$ there exists a set of $m_i$ paths $\mathcal{P}_i = \{P_i^1, \cdots, P_i^{m_i}\}$ on the connectivity graph such that*

1. *Each source $s_i$ transmits to $t_i$ flow $f_i = \sum_j f_i(P_i^j) = \Theta(1)$ information units per time slot.*

2. *Each edge in the connectivity graph is used a finite number of times by the union of all flows.*

*Proof.* 1. *Sufficiency:* From construction and from condition 1 there exists a set of paths $\mathcal{P}_i = \{P_i^1, \cdots, P_i^{m_i}\}$ through which source $s_i$ transmits $\Theta(1)$ information units per time slot towards destination $t_i$. From the second condition, each path intersects at most a finite number of times with any other path, and thus sharing edges does not affect the order of throughput along the path.

2. *Necessity:* It is straightforward that the first condition is necessary. Assume that there does not exist a set of paths that satisfy the second condition. Consider the edges $e$ where $g(e, n)$ flows intersect, with $\lim_{n\to\infty} g(e, n) = \infty$. From assumption removing these edges will violate the first condition for at least one receiver, say $R_i$. But then $R_i$ will observe throughput of order

$$O(\frac{1}{\min_{e:g(e,n)\to\infty} g(e, n)})$$

that is not constant. Similarly, under our assumption that $f_i(e) = \Theta(C(e))$, we get the contradiction that

$$g(n)\Theta(C(e)) = \sum_e f_i(e) \leq C(e).$$

□

Note that the theorem stipulates the existence of paths that satisfy these conditions: there might also exist paths that do not satisfy the above conditions and still allow a constant throughput. For example, there might exist an edge that is used an infinite number of times, but whose removal (and of the corresponding paths that use it) would not affect the order of the throughput that the receivers experience.

## 3.6 Applications

In this section we are going to apply our proposed methodology to a number of examples, starting with the three cases in the literature we have discussed.

#### Example 1

In the Gupta-Kumar model [9] the nodes are static and transmit within a fixed radius. The transmission policy is such that $\Gamma = \Theta(n)$, and is independent of the scheduling policy. For a given realization of node placement random process, with high probability the associated connectivity graph will have a large fraction of nodes with constant degree $c$.

Consider a $c$-regular graph where the weight associated with each edge is at most $1/c$. For each $(s_i, t_i)$ pair there exist $c$ max-flow paths, each carrying throughput at most $1/c$, of average length $\Theta(\sqrt{n})$. Thus, from Lemma 3.5 we get that

$$\lambda(n) \leq \Theta(\frac{1}{\sqrt{n}}). \tag{3.20}$$

We can get the same result from the sparsity cut bound as follows. The nodes are placed uniformly at random on the area of a unit circle. Take a cut across a diameter of this circle. This cut will separate order $\Theta(n)$ source-destination pairs. However, the capacity of the cut will be of order $\Theta(\sqrt{n})$ (determined by the number of nodes within a small distance from the cut), which leads again to Eq. (3.20).

If we start from a random geometric graph, where nodes transmit at the same range to their closest neighbors, the corresponding connectivity graph will have the structure of a planar graph. For planar graphs, the sparsity bound can be achieved within a constant factor $O(1)$ [46].

### Example 2

In the Grossglausser and Tse model [26], uniform mobility implies that the connectivity graph is a complete graph with uniform weight associated with every edge. That is, each node has degree $n-1$, and the weight associated with each edge is $\Theta(1/n)$ (corresponding to the probability that two nodes are nearest neighbors and the probability that a feasible sender receiver pair is scheduled). This is depicted in Fig.3.2.

We are going to apply the sufficient conditions in Theorem 3.1.

- From each source to each destination node, there exist $n-2$ edge-disjoint paths of length two, and one path of length one. Indeed, since the graph is complete, there exists an edge from node $s_i$ to all other nodes in the graph (including the destination $t_i$) and an edge from $t_i$ to every other node in the graph, which put together form the described max-flow paths. Using these paths, we can route throughput $\Theta(1) = (n-1)\frac{1}{n}$. This shows that Condition 1 is verified.
- We are going to show that if $\Theta(n)$ source-destination pairs share the network and use the paths previously described, each edge is used a finite number of times. Consider edge $(k, l)$ between nodes $k$ and $l$. This edge is going to be used only if node $k$ is a source, or if node $l$ is a destination - so at the most it is going to be used two times. This verifies Condition 2.

### Example 3

In the Diggavi, Grossglauser and Tse model [42], mobility is restricted since the nodes are only allowed to move along great circles in the surface of a sphere, and constant throughput is still achieved by employing the same two-phases policy. In [42], it is argued that the two phases policy achieves a constant throughput mainly for two reasons. First, each node spends the same order amount of time as the nearest neighbor to every other node. This ensures that each source spreads its information units uniformly across

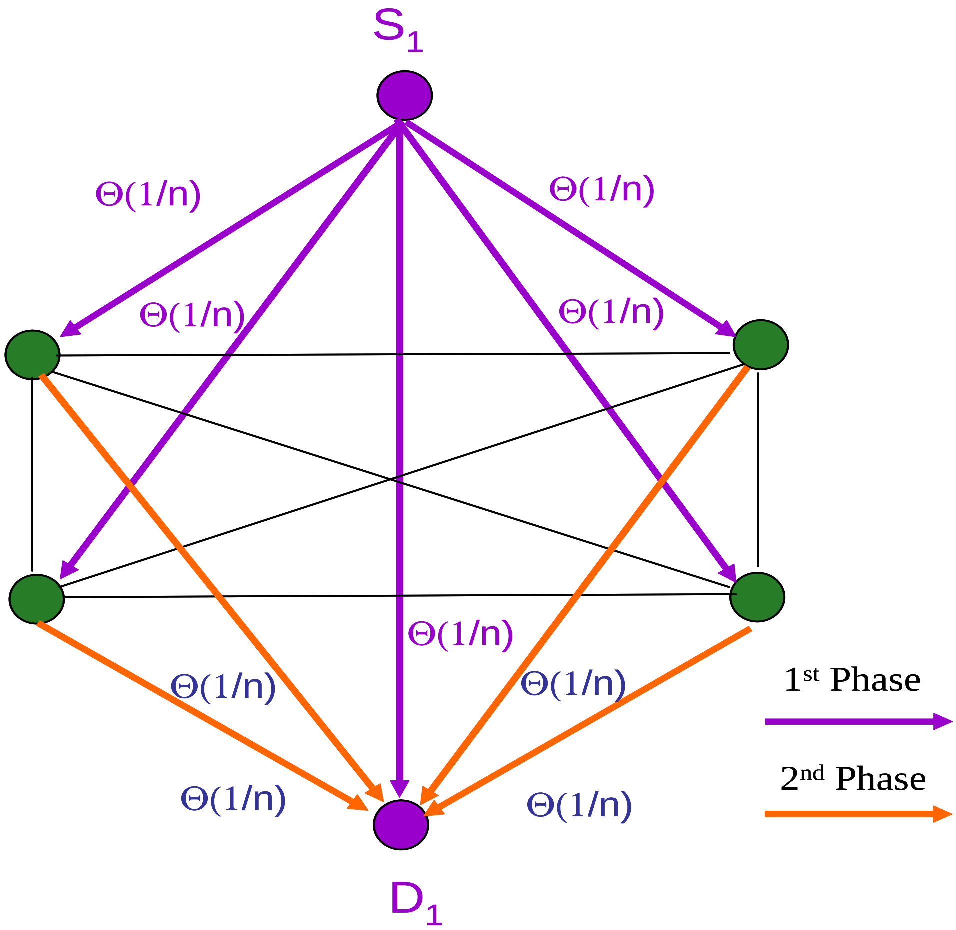


**Figure 3.2:** The connectivity graph is a complete graph in the case of uniform mobility pattern.

all other nodes and the traffic is equally distributed among all relays. Secondly, similarly to [9], communications are constrained to nearest neighbors. Hence, the capture probability is not vanishingly small even in a large system, despite the fact that there are $O(n)$ interfering nodes transmitting simultaneously. These observations correspond to having the connectivity graph be a complete graph, where the degree for every node is of the order $\Theta(n)$ and the weights associated with each edge are $\Theta(1/n)$. Thus, the analysis is the same as in the previous example. Having a connectivity graph to be a complete graph with equal capacity edges is a sufficient, but not necessary condition, for constant throughput. Moreover, the two-phase routing strategy, is tied to the fact that the graph is complete, and thus it is also not necessary for constant throughput, as the next example demonstrates.

#### Example 4

Consider the example of geographic gossip over a sensor network as described in [55]. The $n$ sensors collect a measurement and it is of interest that all nodes can compute the average of all $n$ sensor measurements. Traditional gossip algorithms solve this problem by having each node randomly pick one of their one-hop neighbors and exchange their current values. The pair of nodes compute the pairwise average which becomes the new value for both nodes, and this process is iterated until all nodes converge to the global average. [55] introduces the notion of geographic gossip where geographic routing is used to exchange information with random nodes who are far away in the network. The wireless sensor network is then modeled as a random geometric graph where the transmission range scales as $\Theta(\sqrt{\frac{\log n}{n}})$ and the number of hops between a source and a randomly uniformly distributed destination scales as $\Theta(\sqrt{\frac{n}{\log n}})$. This random geometric graph could be represented by our connectivity graph again.

#### Example 5

Consider the rectangular grid depicted in Fig. 3.3 that has $2d$ lines. Assume that the nodes are uniformly distributed on the lines of the grid. Thus, each line will contain $n/2d$ nodes. Assume that our transmission policy is such that, nodes in the same line or in intersecting line can communicate with each other if they are within a certain range (protocol model), but nodes in parallel lines cannot communicate. Since the min-cut between parallel lines is at most equal to $d$, a necessary condition to have constant throughput is that $n/2 = O(d)$.

Now consider the case where $n = 2d$, that is, each line contains exactly one node. Then the corresponding connectivity graph is a complete bipartite graph, since from assumption nodes in parallel lines do not communicate. It is easy to see that, for a complete bipartite as depicted in Fig. 3.4, if the source-destination pairs belong to parallel lines, there exist $n/2$ non-intersecting paths of length two, while if they belong to intersecting lines, there exist $n/2 - 1$ non-intersecting paths of length three and one path of

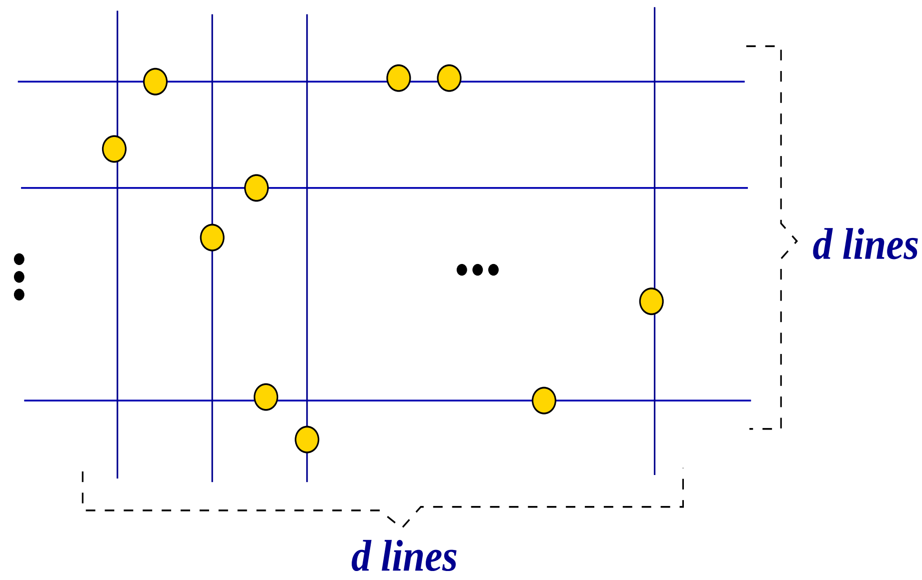


**Figure 3.3:** Mobility on a rectangular grid.

length one. Thus, from Theorem 3.1, if each edge has capacity $\Theta(1/n)$, it is possible to achieve constant throughput.

If nodes are allowed to uniformly randomly move on their line, then ignoring edge-effects, the capacity of each edge is upper bounded by $\Theta(1/n^2)$. However, it is possible to construct a mobility pattern that leads to edge capacities $\Theta(1/n)$. For example, to avoid edge-effects we can assume that the square grid envelopes the surface of a torus. Thus the nodes move on parallel (horizontal and vertical) "circles" instead of lines, i.e.., the end points of the line segments are connected. We can then construct a mobility pattern, where nodes move clockwise on their circle, where at time slots $(i) \mod (d) = k$, the node at the horizontal circle $i$ is within range (and successfully communicates) with the node at the vertical circle $(i + k) \mod (d)$.

### Example 6

This example is a direct application from the multi-commodity flow literature. Consider a connectivity graph, that has constant degree, and is an expander graph [53]. Expander graphs have been the subject of much study in combinatorics and computer science. Broadly speaking, an expander graph

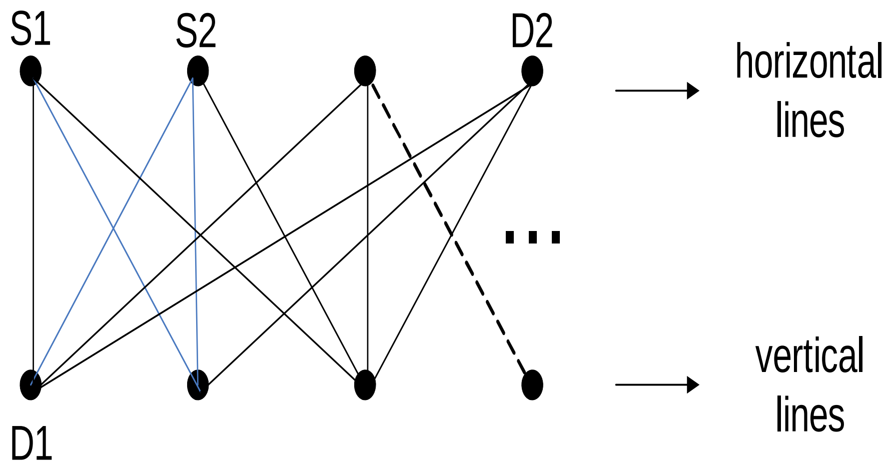


**Figure 3.4:** Bipartite graph.

should be a graph in which every reasonably small set of vertices has many neighbors. Then, the sparsity bound gives that $\lambda(n) \leq \Theta(1)$. However, the best achievable throughput equals $\lambda(n) = \Theta(\frac{1}{\log n})$. This example is used to demonstrate that the sparsity bound cannot always be achieved.

**Example 7**

We consider the example studied in [51] where nodes are placed randomly with $n^{\alpha}, 0 \leq \alpha \leq 1$ overlapping neighborhoods. The $n$ mobile nodes are restricted to move within their assigned neighborhood. In the associated graph of the model described above, each node has degree $n^{1-\alpha}$ and the weight associated with each edge is at most $n^{\alpha-1}$. Moreover, the average path length in this model is $\Theta(n^{\frac{\alpha}{2}})$. Thus, for each $(s_i, t_i)$ pair there exist $n^{1-\alpha}$ max-flow paths, each carrying throughput at most $n^{\alpha-1}$, of average path length $\Theta(n^{\frac{\alpha}{2}})$. From Lemma 3.5 we get that the per source destination throughput is:

$$\lambda(n) = O(n^{-\frac{\alpha}{2}}) \tag{3.21}$$

which corresponds to an aggregate throughput of $O(n^{1-\frac{\alpha}{2}})$. The connectivity graph paradigm takes into account the restricted mobility model in terms of the average path length and the node degree, which motivates the next section.

## 3.7 Degree of Connectivity Graph

In this section, we examine whether there exist fundamental structural properties of the connectivity graph that are necessary in order to achieve constant throughput. We focus on possible restrictions that the degree of the connectivity graph should satisfy. The degree of the connectivity graph corresponds to the number of neighbors a given node may communicate with. We are going to assume that the degree is of the same order for all nodes, that is, all nodes behave in a similar way in terms of mobility.

As we saw in Example 1, if the connectivity graph has a constant degree, that is, every node can communicate with a fixed finite number of neighbors, it is not possible to achieve $\Theta(1)$ throughput. Indeed, as the number of nodes $n$ increases, since the degree is constant, the average path length between a randomly chosen source-destination pair will also increase as a function of $n$, say $P(n)$. But then from Lemma 3.5, $\lambda(n) = O(\frac{1}{P(n)})$.

On the other hand, in Examples $2-4$, where $\Theta(1)$ throughput was possible, the associated connectivity graphs has degree $\Theta(n)$. The question we are looking at is: To achieve throughput $\lambda(n) = \Theta(1)$, is it necessary for the connectivity graph to have degree of order $\Theta(n)$? In other words, to achieve throughput $\lambda(n) = \Theta(1)$ in an ad hoc network with $n$ nodes, is it a necessary condition that the mobility pattern ensures for every node to at some point be able to successfully communicate with $\Theta(n)$ other nodes?

Theorem 3.2 shows that no such conclusion can be derived. There do exist possible connectivity graphs with degree of order $\Theta(n^{\frac{1}{t}})$, for finite $t$, such that the throughput is constant. Such connectivity graphs can be constructed from *de Bruijn* and *Kautz* graphs and their generalizations [56]. The *de Bruijn* graphs $D_B(d,t)$ have $n = d^t$ nodes, and degree at most $2d$. It is a $t$-dimensional graph of $d$ symbols representing overlaps between sequences of symbols (see Fig.3.5). Its $n = d^t$ vertices consists of all possible length-$t$ sequences of the given symbols. In the proof of Theorem 3.2 we show that for each source-destination path there exist max-flow paths of length at most $2t$. For example, if we choose $d = \sqrt{n}$, $t = 2$, and associate weight $\Theta(1/\sqrt{n})$ with each edge, we have a graph that has degree $d = \Theta(\sqrt{n})$, and $\sqrt{n}-1$ paths that carry flow $1/\sqrt{n}$ of length at most 4. Thus the conditions of Theorem 3.1 are satisfied.

**Theorem 3.2.** *Consider the de Bruijn graph $D_B(d,t)$ with $n = d^t$ nodes,*

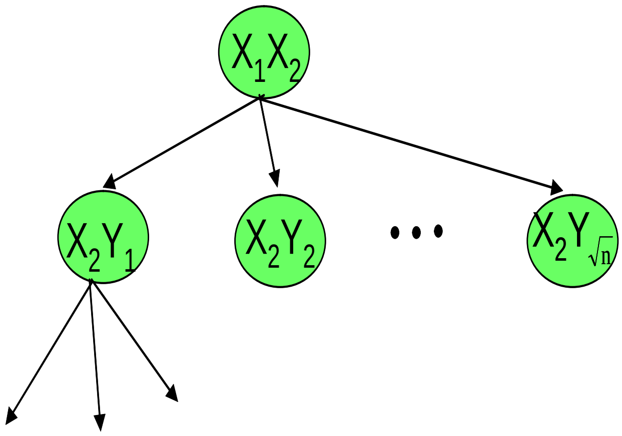


**Figure 3.5:** An example of the De Bruijn graph. Each node can be represented as a 2-dimensional vector over an alphabet of size $\sqrt{n}$.

*and assume that each edge has capacity* $\Theta(1/d)$. *Let* $d = n^{\frac{1}{t}}$ *and* $t$ *be finite. For any* $\Theta(n)$ *randomly chosen source-destination pairs, there exist max-flow paths of length at most* $2t$ *such that* $\lambda(n) = \Theta(\frac{1}{2t})$ *and each edge of the graph is used a finite number of times.*

*Proof.* Consider the de Bruijn graph $D_B(d,t)$ with $n = d^t$ nodes, and associate with each edge capacity $\Theta(1/d)$. We can think of every vertex of the graph as a $t$-dimensional vector over an alphabet of size $d$. Each node $(s_1\ s_2 \ldots s_t)$ connects to the $d$ nodes $(s_2 \ldots s_t\ x)$, for every possible $x$ in the alphabet. For simplicity we start with the case $d = \sqrt{n}$ and $t = 2$. Consider the source node $(s_0\ s_1)$ transmitting to a destination node $(s_2\ s_3)$. To route information we use the following length-four paths, that are also depicted in Fig. 3.6.

1. Source $(s_0\ s_1)$ transmits information $\Theta(1/\sqrt{n})$ to each of the $\sqrt{n}$ neighbors $\{(s_1\ s_0), (s_1\ s_1) \ldots (s_1\ s_d)\}$.
2. Each of the $\sqrt{n}$ nodes $(s_1\ s_i)$ equally distributes the $\Theta(1/\sqrt{n})$ information units it has among its $\sqrt{n}$ neighbors $(s_i\ s_j)$. Thus at the end of this step, every node in the graph has $\Theta(1/n)$ units of the information.
3. All nodes $(s_i\ s_j)$ transmit the $\Theta(1/n)$ information they have to the $\sqrt{n}$ nodes $(s_j\ s_2)$. Thus each $(s_j\ s_2)$ node collects $\Theta(1/\sqrt{n})$ of the information.

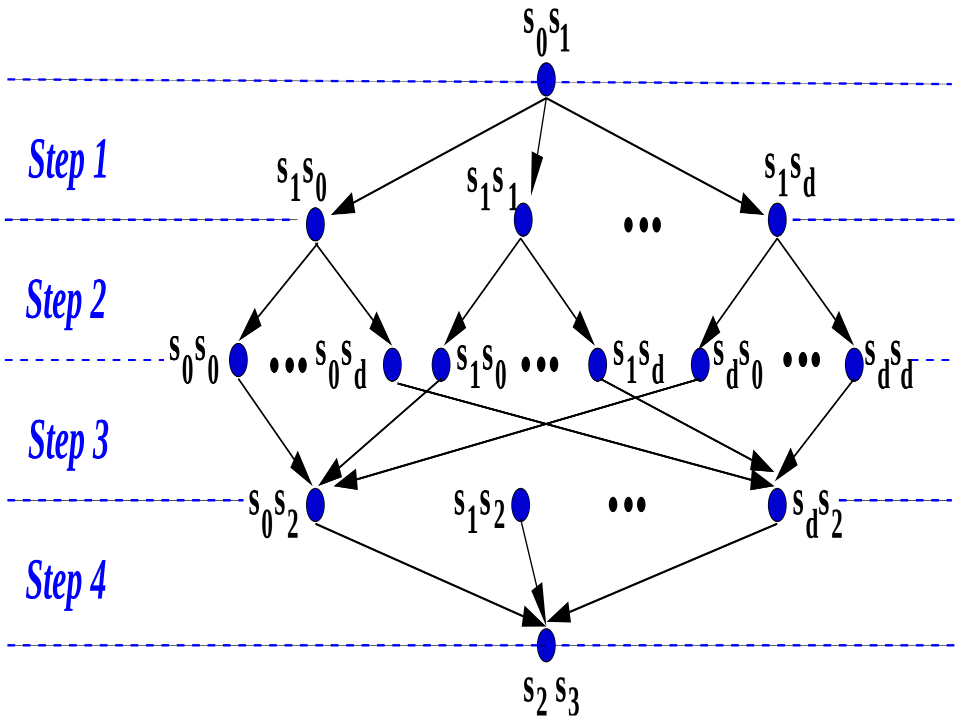


**Figure 3.6:** Proposed set of paths in the de Bruijn graph $D_B(\sqrt{n}, 2)$.

4. Finally, each $(s_j\ s_2)$ node transmits its information to the destination $(s_2\ s_3)$.

We show now that, if all $\Theta(n)$ source-destination pairs use the previously described paths, then each edge of the graph is employed a finite number of times. Let $(ab, bc)$ be an arbitrary edge, we are going to calculate an upper bound on the number of times it is used, by bounding the number of times it may be employed through steps $1-4$.

1. Edge $(ab, bc)$ is used at step 1 only if node $ab$ is a source, and thus at most once.

2. Edge $(ab, bc)$ is used at step 2 if there exists node $xa$ that is a source. In that case it will carry information $\Theta(1/n)$ for the source $xa$. There exist at most $\sqrt{n}$ possible sources $xa$, and the edge $(ab, bc)$ has capacity $1/\sqrt{n}$, thus to accommodate all of them $(ab, bc)$ needs to be used at most once.

3. Edge $(ab, bc)$ is used at step 3 only if there exists a node $cx$ that is

a destination. In that case it will carry information $\Theta(1/n)$ towards the destination $cx$. But there exist at most $\sqrt{n}$ possible destinations $cx$ and thus to accommodate all of them the edge $(ab, bc)$ needs to be used at most once.

4. Finally, edge $(ab, bc)$ is used at step 4 only if node $bc$ is a destination, in which case it needs again to be used once.

Thus in total, each edge needs to be used at most $2t = 4$ times.

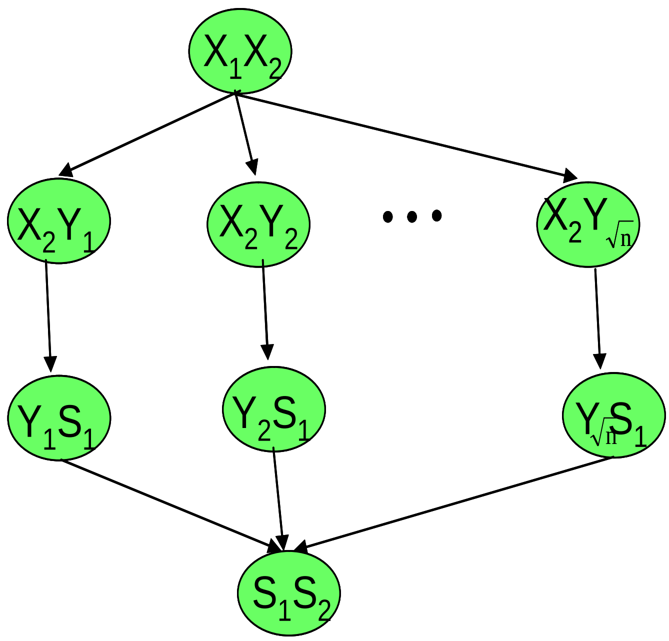


**Figure 3.7:** Between every two nodes, there exist $\sqrt{n}$ edge-disjoint paths of length four.

Generally, it is easy to see that, if we have degree $d = n^{\frac{1}{t}}$, and each edge has capacity $\Theta(\frac{1}{n^{\frac{1}{t}}})$ by following the same approach of equally distributing the information to all neighbors in $t$ steps, until all nodes will have $\Theta(1/n)$ information for each source-destination pair, and then collecting the information towards the receiver in another $t$ steps, we can construct max-flow paths of length $2t$ such that, each edge at every one of the $2t$ steps is used at most once. Indeed, at step $k$, $k = 1 \dots t$, each edge will carry for each source-destination pair a load of at most $\frac{1}{n^{\frac{(k+1)}{t}}}$ and will be used by at most $n^{\frac{k}{t}}$ pairs. The same holds for $k = t + 1 \dots 2t$. Note that the necessary condition in Lemma 3.5 is satisfied: the graph has a total of $n^{1+\frac{1}{t}}$ edges, and the max-flow paths require $\Theta(2tn^{1+\frac{1}{t}})$ edges. □

The proposed max-flow paths used in the proofs are not the only paths that may guarantee constant throughput. Actually, for each source-destination pair there exist max-flow paths of length $t+1$ [56]. However, depending on the random arrangement of pairs, this choice of max-flow paths may lead to congested edges. To divert the traffic and equally distribute it over the network, we may need to use different length paths for each source-destination pair, that makes the proof less straightforward.

From Theorem 3.2 and the previous discussion, we can see that:

**Corollary 3.1.** *It is not possible to achieve constant throughput if the connectivity graph has constant degree. However, we may be able to achieve constant throughput if the connectivity graph has degree $n^{\frac{1}{t}}$, where $t$ is a constant.*

The next question to examine is whether there exists a mobility pattern that corresponds to connectivity graphs of the prescribed degree, mainly what mobility patterns give rise to a connectivity graph of degree $n^{1/t}$. Thus, a future direction of this work is to investigate methods to translate connectivity graphs into mobility patterns. Such a direction will allow the study of mobility patterns that achieve an intermediate throughput, i.e a per node throughput that is not constant and does not depend $\sqrt{n}$. Finally, connectivity graphs with vertices having distinct degrees might be of great interest. As stated in this chapter, all the studies on the mobility in Ad Hoc wireless networks, consider that all nodes of the network have the same mobility model. However, in practice, it might not be the case, and some node might have a restricted mobility pattern whereas others might have a uniform mobility pattern. It would be then interesting to derive the per-node throughput in this case based on the properties one can extract from the connectivity graph.

## 3.8 Conclusions

In this chapter, we proposed to use a connectivity graph to study the long-term averaged throughput of wireless networks operating in an ad hoc mode. Throughput refers to the minimum achievable rate between a source-destination pair for a given routing mechanism and physical model, when the network is shared by $\Theta(n)$ randomly chosen source-destination pairs. The connectivity graph offers an abstraction of the communication capabilities of the network, and forms a natural bridge between the literature in wireless networks and the multi-commodity flow problem. Hence, this graph allows to translate the problem of maximizing the throughput in ad hoc networks to the multi-commodity flow problem and directly apply related results. Using this graph, we proposed a set of necessary and sufficient conditions to achieve constant throughput, and examined structural properties that these conditions imply. We also applied these conditions in a number of configurations in the literature, and demonstrated that they offer an alternative simpler methodology to re-derive these results.

# Part II

# Chapter 4

# Spatial Throughput of Multi-Hop Wireless Networks: Cross-Layer Design

## 4.1 Introduction

The results derived in the work of Gupta and Kumar [9] and in related literature as [26] provide expressions for the ad hoc network capacity and determine the scalability of such networks as the number of nodes increases to infinity.

In contrast to these macroscopic studies, in this work we focus on a microscopic analysis of decentralized ad hoc wireless networks where some form of coordination entity between nodes at the MAC layer (e.g. TDMA-based exclusion techniques, hybrid coordinator in 802.11e) is not desired, for instance due to the need for minimizing protocol overhead or due to node mobility; for example in [30], a different approach illustrates the impact of an exponentially decaying path length traffic pattern and the impact of the relay load on the throughput in the context of a decentralized system with retransmission protocols. An example application for this type of radio interface could be agile and rapidly deployable wireless infrastructure nodes for future public safety [57] (e.g broadband hotspots for emergency/disaster relief) and/or next generation commercial wireless systems. Other application could be rapidly deployable wireless mesh network that is critical to large-scale wireless networks with no pre-existing infrastructure. It en-

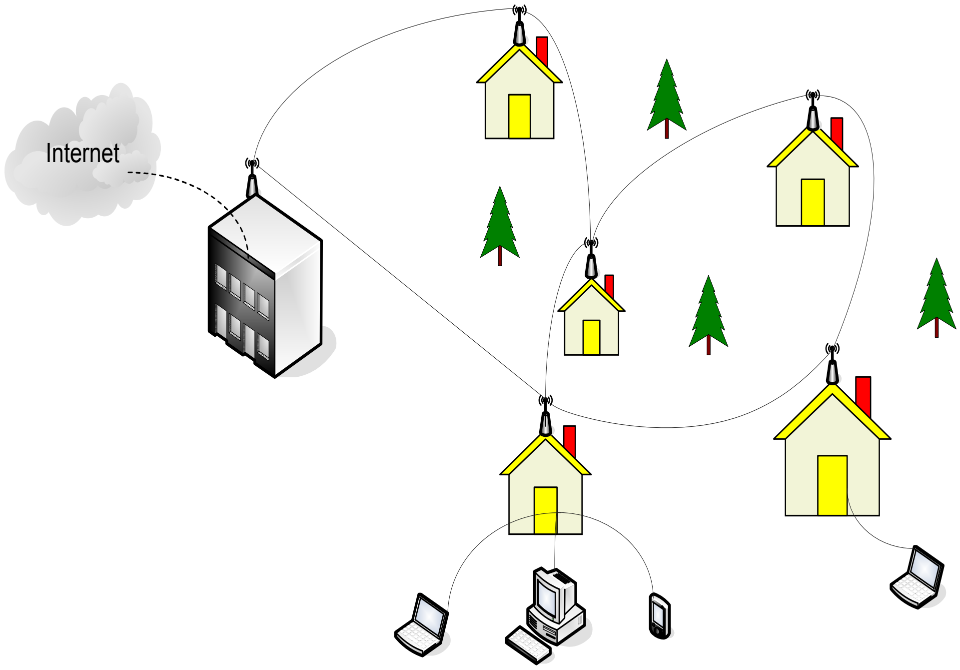


**Figure 4.1:** Example application of Ad Hoc wireless network.

ables quick-and-easy extension of a local area network into a wide area by allowing the geographical area and available bandwidth on the network to scale with the number of participants. Prior efforts on wireless networks, especially multi-hop ad hoc networks, have led to significant research contributions that range from fundamental results on theoretical capacity bounds to various flavors of routing and transport protocols. However, the work is far from enough. The state-of-art is insufficient for deploying sizable wireless mesh networks. Important aspects such as network radio range, network capacity and scalability, manageability, and security remain open problems. Industrial standards groups are also actively working on new specifications for mesh net-working for both short and long range systems. For example, IEEE 802.11s, and IEEE 802.16 have established sub-working groups to focus on new standards for Wireless Mesh Networks (WMN) (see Fig.4.1, Fig.4.2). To address capacity and scaling concerns, researchers are experimenting with systems that use multiple radios, frequency-agile radios, directional and multiple input multiple output (MIMO) antennas. Further, there are renewed interests in carrying out research on MAC protocols and

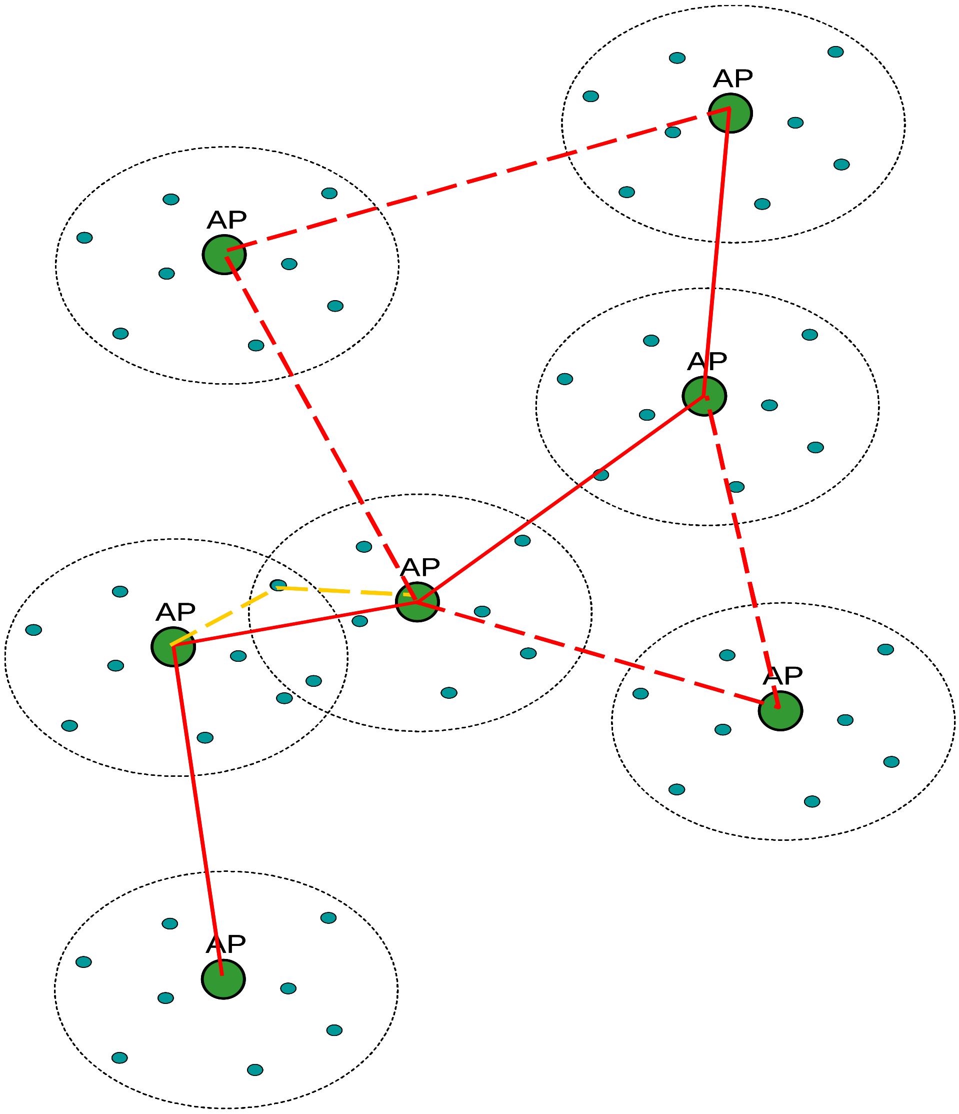


**Figure 4.2:** Extension of WLAN: 802.11s.

cross-layer design that breaks the traditional networking layering norm. It is clear that such novel techniques are interesting from the point of view of increasing system capacity and scalability, extending network range. Our work is also motivated by cross-layer mechanisms (PHY/MAC/Routing) aiming at maximizing the spectral efficiency of the network.

To this end, we present a cross-layer framework for the design of these wireless networks. We jointly address the properties of the physical and the data link layer in the design of the media-access control (MAC) protocol and provide conclusions on routing strategies based on physical layer metrics. We provide a setting to characterize the performance of such networks. We assume that nodes access the channel at random and employ simple protocols to retransmit the erroneously received packets. We consider two possible retransmission protocols: the classical reference scheme is *Slotted Aloha* (using the wireless setting as described in [58]) shows the benefit of coupling channel coding for medium-access and where decoding considers only the most recent received block; the second is *Incremental Redundancy* where decoding takes into account all previously received signal blocks and performs soft combining until decoding is achieved successfully. We compare these strategies to the generalization of the collision channel without feedback or delay constraints [59], where the measure of success of a transmission will be an achievable ergodic throughput of this channel as it will be seen later.

For this analysis, the nodes are taken to be spatially distributed on the plane according to a homogeneous spatial Poisson process which leads to a new representation of interference and collisions between concurrent transmissions. This random characterization of the network is justified by the fact that the homogeneous Poisson point process is spatially ergodic and thus the performance quantities considered in this work for particular network realizations (network throughputs and information outage probability), will converge quickly to the average performance of the random network.

To derive the throughput, we follow the analysis of Nelson and Kleinrock in [60] where they studied the spatial capacity of a slotted Aloha multi-hop network with capture. The spatial throughput is computed in terms of the product of the number of the simultaneously successful transmissions per unit area by the average jump (or expected forward progress) made by each transmission. We carry out its optimization with respect to the channel access probability $p$ as defined in the case of the collision channel without feedback [59]. The relationship between the spatial throughput and the Gupta-Kumar transport capacity is described in [61]. For the purpose of comparison of potential multi-hop routing protocols, we consider three strategies; one that maximizes the expected forward progress (**RS1**) based on long-term averages of signal-to-interference ratios, the second that relays packets to the closest node in range at each hop towards the final destination

(**RS2**) and in the third where the next hop is selected to exploit the best channel and to be the most forward (**RS3**). This last strategy attempts to exploit *instantaneous* channel state information at transmission when choosing candidate routes, rather than relying on average signal-to-interference ratios. All routing strategies assume a-priori information regarding neighbors and destination coordinates. This assumes a form of neighbors discovery and route capabilities knowledge. This is clearly a form of geographic routing strategies [62, 63].

The main contributions of this chapter are:

- The study of incremental redundancy as a multiple access technique for ad hoc wireless networks
- Representation of interference and collisions statistics from the homogeneous Poisson point process network model
- A cross-layer framework where multi-hop routing protocols are analyzed and in particular the channel-driven routing strategy and tools for characterizing the spatial throughput (bit-m/dim, related to the transport capacity) from a microscopic point-of-view as a function of topological parameters (e.g node population density) and system parameters (propagation, bandwith, etc.).

The outline of the chapter is as follows: In Section 4.2, we describe the system model. Section 4.3 deals with the retransmission protocols. In Section 4.4, throughput expressions are derived and we show some numerical results. Finally, in Section 4.5 we draw some conclusions and point out future research directions.

## 4.2 System Model and Setting

### 4.2.1 Network and Propagation Model

We assume that nodes are distributed according to a Poisson point process on the plane with node density $\sigma$. This topology represents an instantaneous snapshot of a mobile network of nodes. Then, for any region $\mathbf{S}$ of area $A(\mathbf{S})$, the number of nodes in the region has a Poisson distribution with parameter $\sigma A(\mathbf{S})$, i.e.,

$$\Pr[k \text{ in } \mathbf{S}] = e^{-\sigma A(\mathbf{S})} \frac{(\sigma A(\mathbf{S}))^k}{k!} \tag{4.1}$$

The propagation model is described by two effects: the signal attenuation due to the distance $r$ between the transmitter and the receiver, proportional to $r^{-\alpha}$, where $\alpha$ is the power loss exponent (positive number) ; and Rayleigh fading that causes random power variations. The envelope of the received signal is therefore, Rayleigh distributed and its power exponentially distributed. The received power $P_R$ from a mobile at distance $r$ is expressed as:

$$P_R = K_{d_0} R_a^2 r^{-\alpha} P = K_{d_0} \gamma r^{-\alpha} P \tag{4.2}$$

where $R_a$ is a Rayleigh distributed random variable (with unit power for simplicity), $\gamma$ is an exponentially distributed random variable (with unit mean) and $P$ is the transmit power. $K_{d_0}$ represents the signal attenuation at a close-in reference distance [66]. This represents a narrowband channel with respect to the coherence bandwidth of the environment.

### 4.2.2 System Model and Setting

In the system we are considering, each node can transmit over a common wireless channel. Apart from the slotted transmission structure where nodes transmit packets within slots of defined duration, nodes are completely uncoordinated (see Fig.4.3). This slotted transmission scheme requires some local frame synchronization method, for instance a form of distributed transmission of pilot signals, or could be based on a common pilot from an external source (e.g. cellular infrastructure, satellite positioning systems, etc.). For the purpose of our analysis, we make the following assumptions:

- An infinite number of packets is available for each source. A packet can be seen as a separate codeword for which transmission is stopped when an acknowledgment of successful decoding is returned by the receiver. Furthermore, we assume that the ACK/NACK feedback signaling channel is error-free and delay-free. The signaling overhead is insignificant with respect to the data channel.

- We suppose single-user decoding where each decoder treats the signals from other users as noise. The single-user decoder for each node has perfect knowledge of the channel gain and the total interference power (i.e. noise and interfering user traffic). This can be achieved in a real system by inserting some pilot symbols.

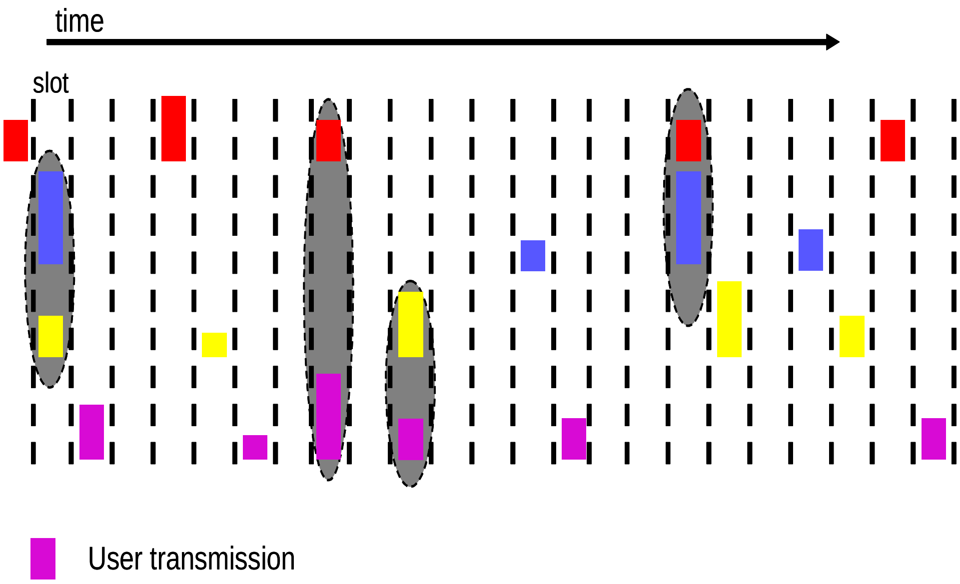


**Figure 4.3:** The slotted transmission and the channel random access. Packets are possibly coded over many slots.

- We assume a block-fading channel model. For the first two routing protocols, we assume that the fading remains constant over the whole slot and is an i.i.d process across successive slots in order to provide diversity against fading. In a real system, this can be achieved via slow frequency hopping across a large system bandwidth, where the number of frequencies is typically larger than the number of retransmission rounds. For the third routing strategy (**RS3**), we will assume a long-term static channel, in the sense that the channel remains constant over all the retransmission rounds of the protocol. In this case, frequency-hopping is not a possibility, since it could imply changing routes within a transmission round. Nevertheless, it is important to note that in this case, diversity against signal fading is achieved by the routing protocol, since the routes are chosen based on the *instantaneous* channel realization. This is clearly a form of *multi-user diversity*. Diversity against interference is achieved in all cases by the retransmission protocol (Aloha or Incremental Redundancy) since in each slot the number of interferers is random due to the random channel-access transmission strategy.

- For each slot, each node transmits a packet with probability $p$ and remains silent with probability $1 - p$ such that transmit and receive nodes have spatial Poisson distributions with average node density $\sigma_t = \sigma p$ and $\sigma_r = \sigma(1 - p)$ respectively. One could incorporate in this model the possibility of having some nodes acting as pure relays. These nodes will always be in receive mode, hence increasing the receive node density. This has practical implications for wireless mesh networks, where some nodes act as gateways or repeaters between different sub-networks or different clusters. Moreover, in a slow frequency-hopping system, $1/p$ could be the number of frequencies (or the number of sub-bands/carriers in an OFDMA Orthogonal Frequency Division Multiple system) when nodes transmit only on a single frequency for any time-slot coupled with a duty cycle to randomize the access to the channel.
- Each node transmits with fixed power $P$.

Moreover, the signal model is given by:

$$y_{j,s} = \sum_{k \in \Gamma} \sqrt{\gamma_{k,j,s} P r_{k,j}^{-\alpha}} x_{k,s} + n_{j,s} \tag{4.3}$$

where the index $s$ denotes the slot, $y_{j,s}$ the received signal at node $j$, $x_{k,s}$ the transmitted signal from node $k$, $\Gamma$ is the set of transmitter nodes and $n_{j,s}$ the background noise at node $j$ of variance $\sigma_n^2 = WN_0$ where $W$ is the available system bandwidth ($N_0$ is the background noise power). This represents a narrowband flat fading channel model and the generalization to frequency selective channels is straightforward but diversity features of the access protocols will be reduced. Moreover, this is a more general multiple-access scenario than the interference model considered in [9, 26], since we are at liberty to optimize the transmission probability $p$ to randomize interference levels, which is important in a microscopic analysis. Note that this creates a random exclusion area around each node.

## 4.3 Retransmission Protocols

All performance metrics considered for throughput analysis are based on information theoretic quantities, i.e, these metrics are function of the instantaneous average mutual information generated by a particular link in a given slot. This represents a modern PHY approach in contrast with previous analysis based on the instantaneous signal to noise interference ratio, and this allows for an analysis covering arbitrary channel coding strategies across several transmission slots.

### 4.3.1 Information Outage Probability

The instantaneous average mutual information for a $(i, j)$ pair of nodes conditioned on the channel gain $\gamma_{i,j,s}$ and the interference power $V$ is (in bit/dim):

$$I_{i,j,s} = I(X_{i,j,s}; Y_{j,s}|\gamma_{i,j,s}, V) = \log_2\left(1 + \frac{\gamma_{i,j,s} P r_{i,j}^{-\alpha}}{\sigma_n^2 + V}\right) \tag{4.4}$$

where $P$ is the transmit power and $r_{i,j} = |X_i - X_j|$where $X_j$ is the position of the receiver, $V$ is defined as the summation of interference power contributions from all interfering transmitters (in the following we will drop the index $s$ standing for the slot):

$$V = \sum_{k \neq i} \gamma_{k,j,s} P r_{k,j}^{-\alpha} \tag{4.5}$$

$P_{out}(r_{i,j})$ is defined as the information outage probability of the channel, or the probability that the mutual information $I_{i,j}$ falls below some fixed spectral efficiency $R$. Expressions of the mutual information necessary for the outage probability evaluation are derived under the assumption that all user signals are Gaussian with flat power spectral density. The Gaussian assumption yields an upper-bound to the minimum achievable outage probability [67][69]. $P_{out}(r_{i,j})$ is given by:

$$\begin{aligned} P_{out}(r_{i,j}) &= \Pr\left(\log_2\left(1 + \frac{\gamma_{i,j} P r_{i,j}^{-\alpha}}{\sigma_n^2 + V}\right) \leq R\right) \\ &= 1 - e^{-\frac{(2^R-1)\sigma_n^2}{P r_{i,j}^{-\alpha}}} \int e^{-v\frac{(2^R-1)}{P r_{i,j}^{-\alpha}}} f_V(v) \mathrm{d}v \end{aligned} \tag{4.6}$$

where $f_V(v)$ is the probability density function of the random variable $V$ and $v$ is the integration variable. Let us define $t = \frac{(2^R-1)}{P r_{i,j}^{-\alpha}}$, and $\int e^{-vt} f_V(v) \mathrm{d}v = \phi_V(t)$ is the moment generating function MGF of $V$. To compute the latter, we follow the procedure in [70] (and references therein) where the problem is different since the fading was not considered. To compute the MGF of $V$, we restrict $V$ to all nodes in a disk $\mathbf{D}_b$ centered at the receiver and having radius $b$, then we let $b \to \infty$. Moreover, given $k$ nodes in a region, $V$ is a sum of independent random variables with uniform distribution and from [77], we obtain:

$$\begin{aligned}\phi_V(t) &= \lim_{b\to\infty}\sum_{k=0}^{\infty}\Pr[k \text{ in } \mathbf{D}_b] E\left[e^{-tV}|k \text{ in } \mathbf{D}_b\right] \\ &= \lim_{b\to\infty}\sum_{k=0}^{\infty}\Pr[k \text{ in } \mathbf{D}_b] E\left[e^{-tP\gamma r^{-\alpha}}\right]^k \end{aligned} \tag{4.7}$$

Defining $g(r) = r^{-\alpha}$ and $\beta = P\gamma$ as an exponential random variable with mean $P$, we obtain:

$$E\left[e^{-tP\gamma r^{-\alpha}}\right] = \int f_r(r)\phi_\beta(tr^{-\alpha})\mathrm{d}r = \int f_r(r)\frac{1}{1+Ptg(r)}\mathrm{d}r \tag{4.8}$$

where $\phi_\beta(tr^{-\alpha}) = \frac{1}{1+Ptg(r)}$ is the MGF of $\beta$ evaluated at $tg(r)$, and $f_r(\cdot)$ is the probability density function of $r$. The probability density function of the distance between a transmitter and a receiver in a disk of radius $b$ has the uniform distribution:

$$f_r(r) = \begin{cases} \frac{2r}{b^2} & \text{if } 0 \le r \le b \\ 0 & \text{elsewhere} \end{cases} \tag{4.9}$$

Thus, Eq.(4.7) becomes (using the fact that all the interfering nodes form a Poisson process with density $\sigma_t$):

$$\begin{aligned}\phi_V(t) &= \lim_{b\to\infty}\sum_{k=0}^{\infty} e^{-\sigma_t\pi b^2}\frac{(\sigma_t\pi b^2)^k}{k!}\left(\int_0^b \frac{2r}{b^2}\frac{1}{1+Ptg(r)}\mathrm{d}r\right)^k \\ &= \lim_{b\to\infty} e^{\sigma_t\pi b^2\left[\int_0^b \frac{2r}{b^2}\frac{1}{1+Ptg(r)}\mathrm{d}r - 1\right]} \\ &= \lim_{b\to\infty} e^{\sigma_t\pi b^2\left[\int_0^b \frac{2r}{b^2}\left(\frac{1}{1+Ptg(r)}-1\right)\mathrm{d}r\right]} \\ &= \exp\left\{-2\sigma_t\pi\int_0^\infty \frac{rPg(r)t}{1+Ptg(r)}\mathrm{d}r\right\} \end{aligned} \tag{4.10}$$

Remember that $t = \frac{(2^R-1)}{Pr_{i,j}^{-\alpha}}$ and with some manipulations, Eq.(4.6) becomes:

$$P_{out}(r_{i,j}) = 1 - e^{\left(-\frac{(2^R-1)\sigma_n^2}{Pr_{i,j}^{-\alpha}}\right)} e^{\frac{-2\sigma_t\pi\Gamma(2/\alpha)\Gamma(1-2/\alpha)r_{i,j}^2(2^R-1)^{2/\alpha}}{\alpha}} \tag{4.11}$$

where $\Gamma(y) = \int_0^\infty x^{y-1}e^{-x}\mathrm{d}x$ is the Gamma function.

### 4.3.2 Slotted Aloha

The Slotted Aloha protocol can provide random multiple access to a common channel with minimal coordination between the channel users. The channel is divided into fixed length slots. This is a simple scheme where the transmitter sends a codeword to the receiver and waits for an acknowledgment from the receiver. A positive acknowledgment (ACK) means the codeword is received successfully, whereas a negative acknowledgment (NACK) means that errors are detected by the receiver. When the transmitter gets a NACK, it will resend the previous codeword to the receiver until it gets an ACK from the receiver. A data packet collision occurs whenever two or more users transmit at about the same time. We are not assuming that all packets involved in a collision are destroyed. Since the nodes will generally be at different distances from the intended receiver, their respective received powers will not be the same. When packets from different nodes collide, it may still be possible to successfully decode the packet with the strongest received signal power, which is known as the"capture effect"[71]. Following the analysis of [58], we define the throughput as:

$$\eta = \frac{R}{\tau} \tag{4.12}$$

where $\tau$ is the mean delay measured in slots for the transmission of an information message and $R$ is the spectral efficiency in bit/dim. In Aloha, the receiver has no memory of the past signals, and the probability of successful decoding after $l$ transmitted slots is given by (in the following, we drop the indices $i, j$ standing for the positions of transmitter receiver for simplicity, we keep only the slot index $s$ in the mutual information):

$$\Pr(I_1 < R, I_2 < R, \cdots, I_l > R) = P_{out}^{l-1}(1 - P_{out}) \tag{4.13}$$

and the mean delay is given by:

$$\tau = \frac{(1 - P_{out}) \sum_{l=1}^{\infty} l P_{out}^{l-1}}{p} = \frac{1}{p(1 - P_{out})} \tag{4.14}$$

Combining Eq.(4.12) and Eq.(4.14) we obtain:

$$\eta = Rp(1 - P_{out}) \tag{4.15}$$

This retransmission protocol could be generalized to $M$-*Slotted Aloha* where the codeword is split in $M$ independent blocks in order to benefit from some diversity. A packet is coded across $M$ slots in order to combat interference, which represents a form of interference diversity. We consider that the $M$ blocks are independent and $M$ is fixed. The mean delay is given by:

$$\begin{aligned}\tau &= \frac{M\sum_{l=0}^{\infty}\Pr\left(\frac{\sum_{s=1}^{M}I_s}{M}<R_i\right)^l}{p_t}\\ &= \frac{M}{\left(1-\Pr\left(\frac{\sum_{s=1}^{M}I_s}{M}<R_i\right)\right)p_t}\end{aligned} \tag{4.16}$$

where the index $s$ stands for the slot sequence.

### 4.3.3 Incremental Redundancy

The basic idea behind incremental redundancy is that the code rate is adjusted by incrementally transmitting redundancy information until decoding is successful. Indeed, if the receiver fails to successfully decode a packet, a NACK is sent to the transmitter. This latter will send additional new redundancy bits which are accumulated and processed by the receiver. As explained in [72], incremental redundancy can be achieved by using rate compatible punctured convolutional codes (RCPC). Transmission starts with the highest rate code of the RCPC code family and additional redundancy bits are sent whenever needed. Incremental redundancy is used in modern cellular radio systems such as EGPRS and evolving 3G standards. In [73, 74] these techniques are studied for single-user fading channels.

To study the achievable rate for incremental redundancy, we consider that node $k$ encodes its message information of $b$ bits each independently of other nodes by using a channel code with code book $\mathcal{C}_k \subset \mathbb{C}^{LN}$ where $N$ is the slot length and $L$ is the accumulate number of slots. For the sake of computing information theoretic quantities, we let $L \to \infty$ $N \to \infty$. Codewords are divided into $L$ sub-blocks of length $N$, and we let $\mathcal{C}_{k,l}$ for $l = 1, \cdots, L$ denote the punctured code of length $lN$ obtained from $\mathcal{C}_k$ by deleting the last $L - l$ sub-blocks. If successful decoding occurs at the $l$-th transmission, the effective coding rate for the current codeword is $R/l$ bit/dim where $R = b/N$. In incremental redundancy, the receiver has memory of the past signals since it accumulates mutual information. Since

$\Pr(I_m^{ir} < R) \leq \Pr(I_n^{ir} < R)$ for $m \leq n$, and

$$Pr(I_l^{ir} < R) = \Pr(\sum_{s=1}^{l} I_s < R)$$

(where the index $s$ stands for the slot sequence), the probability of successful decoding after $l$ transmitted slots is given by:

$$\begin{aligned}\Pr(I_1^{ir} < R, I_2^{ir} < R, \cdots, I_l^{ir} > R) &= \Pr(I_1^{ir} < R, I_2^{ir} < R, \cdots, I_{l-1}^{ir} < R) \\ &\quad - \Pr(I_1^{ir} < R, I_2^{ir} < R, \cdots, I_l^{ir} < R) \\ &= \Pr(I_{l-1}^{ir} < R) - \Pr(I_l^{ir} < R) \quad (4.17)\end{aligned}$$

and the mean delay is given by:

$$\tau = \frac{\sum_{l=0}^{\infty} \Pr(I_l^{ir} < R)}{p} \quad (4.18)$$

The throughput is then:

$$\eta = \frac{Rp}{\sum_{l=0}^{\infty} \Pr(\sum_{s=1}^{l} I_s < R)} \quad (4.19)$$

One can notice that $\Pr(\sum_{s=1}^{l} I_s < R)$ is the cumulative density function of the sum of $l$ i.i.d random variables distributed as $I_s$ and evaluated in $R$. This can be computed numerically by using the characteristic function and discrete Fourier transforms as we have already computed the cumulative density function of $I_s$ in closed form Eq.(4.11).

## 4.4 Throughput expressions

The spatial throughput is expressed as a function of the product of the number of the simultaneously successful transmissions per unit space by the average jump made by each transmission, a result that we maximize with respect to the channel access probability $p$. To calculate the spatial throughput, we introduce the expected forward progress as defined in [75], and in [60] where it is assumed that a packet is randomly relayed to one of the neighboring terminals within a circle of defined radius (constrained range) in a capture environment. The expected forward progress of a packet in the direction of its final destination $F$, is the distance $Z$ between the transmitter and the receiver (an intermediate node) projected onto a line towards the final destination and the transmission to that receiver is successful (note that to make the calculations simple, the forward progress is assumed to be the same for any node on the line perpendicular to the direction of the destination, assumption that is reasonable since the distance $r$ in (Fig.4.4) is much smaller than the source destination distance). In the following we

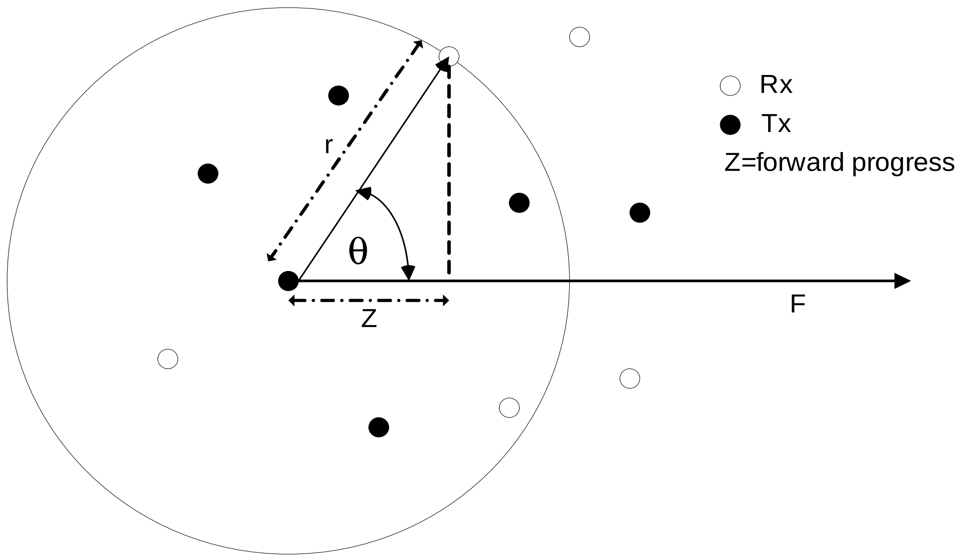


**Figure 4.4:** The forward progress.

present three routing strategies: one that maximizes the expected forward progress by moving the packet to the node most forward towards the final destination; the second moves the packet to the closest node in range and the third one forwards packets to the node most forward and experiencing the best link. Concerning the closest node in range, similar strategy is considered in [76] in the context of mobile info-stations networks, and in [26] where the transmission is spread to a large number of intermediate mobile relay nodes, and whenever they get close to the final destination, they hand the packets off to it, this leads to a transmit range on the order of $O\left(\frac{1}{\sqrt{n}}\right)$, $n$ being the number of nodes in an unit area. In our analysis, we assess the trade-off between the spatial transmission concurrency and the spectral efficiency of the connections. This is explained by realizing that when we decrease the hop-distance, there are less simultaneous transmissions in a given area but we decrease the mutual interference. This leads to an increase in the achievable rate of each pair and consequently of the spectral efficiency of each link, but the potential spatial transmission concurrency is not fully utilized and moreover the number of hops to reach the destination increases. We define the spatial throughput $C$ as the product of the mean total distance traversed in one hop by all transmissions initialized in an unit area $\sigma p E[Z_u]$ (where $E[Z_u]$ is the expected forward progress for strategy $u$) by the bit rate $R$, mainly:

$$C = R\sigma_t E[Z_u] \text{ bit-m/dim} \tag{4.20}$$

The spatial throughput is in function of different cross-layer measures: the PHY target spectral efficiency $R$, the MAC layer parameter (channel access probability) $p$, the routing protocol and the network topology metric $\sigma$, $E[Z_u]$. The spatial throughput is based on outage probabilities to quantify the success of a transmission. This approach takes into account collisions and interferences from simultaneous transmissions in an exact manner. This is in contrast to studies like [60], where a model of restricted reception and transmission domain is considered. The next hop lies in a circle of radius $R$, and only interferers inside this circle are considered.

Also, note that the approach considered in this work at the link model is different from the work of Gupta-Kumar [9]. Indeed in [9], the authors assume a link layer model in which, if the Signal to Interference and Noise Ratio (SINR) at the receiver is greater than a certain threshold $\beta$, then the packet is received successfully by the receiver with probability one. In reality, for a given modulation and coding scheme, as long as there is some noise and interference, i.e., as long as the SINR is finite, there is always a non-zero probability of packet error. While the threshold-based packet reception model used in [9] is a reasonable choice for successful packet reception in a single hop network such as a cellular network, it needs to be refined when applied to a multi-hop network. In the context of an ad hoc network, we know that each packet traverses multiple hops. The links of these hops receive interference from other ongoing transmissions which could potentially corrupt the packet transmission over the given link depending on the current transmissions. In our case, however, each packet transmitted on the link is received successfully with certain probability. This is taken into account by the link outage probability. Another layer is added, mainly when a packet is not successfully received, it is retransmitted over another slot. Moreover the SINR is not deterministic as it depends on factors such as transmitter-receiver separation, fading, power of interference from simultaneous transmissions and noise power. These factors are different for different links along a path. For example, since the nodes are randomly placed, it is possible that for a given hop, the two nodes could be arbitrarily close to each other resulting in a very high SINR for that link. This in turn means that the packet success probability for that link could be arbitrarily close to one. All these parameters (outage probability, SINR) play an important role in tuning the overall network performance. For example, we are at liberty to optimize the modulation and the coding scheme in order to make the error probability of the link as small as possible. Also one can think of a routing strategy that reduces the interference in the network and that makes benefits from the instantaneous SINR value. This will become clearer in the developments below.

Finally, by considering constant of throughput (in contrast to order of throughput), we are able to assess the parameters that affects the spatial throughput of multi-hop networks and the design of such networks. Mainly, the spatial throughput is optimized with respect to the target rate information, the SNR, the channel access probability. The latter depends on the topological parameters as the node density, which is in contrast to the work in [61]. This approach allows a more fine grained adjustment to the PHY/MAC/topology information and routing information especially for the channel driven strategy. Finally, it should be noted that the results on the throughput remain the same when the area of the domain of the network is $A$ rather than normalized unit area. Mainly, the results on the spatial throughput scale by $\sqrt{A}$.

### 4.4.1 Maximal Expected Forward Progress (RS1)

As stated before, the forward progress (the distance traversed in one hop for a successful transmission, which can be seen as a cross-layer measure) is $Z = h(r) = r\cos(\theta)\psi(r)$ (see (Fig.4.4), where $\psi(r)$ is a measure of the success of a transmission, and $\psi(r) = 1 - P_{out}$ for slotted Aloha and $\psi(r) = \frac{1}{\sum_{l=0}^{\infty} \Pr(\sum_{s=1}^{l} I_s < R)}$ for incremental redundancy. To derive the expected maximal forward progress, we compute:

$$\begin{aligned}
&\Pr\left(Z_1 = \max_{j\in\Omega} h(r_j) \leq z\right) \\
&= \lim_{a\to\infty} \sum_{k=0}^{\infty} \Pr\left(\max_{j\in\Omega} h(r_j) \leq z | k \text{ in } \mathbf{D}_{a^{\text{`}}}\right) \Pr\left(k \text{ in } \mathbf{D}_{a^{\text{`}}}\right) \quad (4.21) \\
&= \lim_{a\to\infty} \sum_{k=0}^{\infty} \Pr\left(h(r_j) \leq z\right)^k e^{-\sigma_r \pi a^2/2} \frac{(\sigma_r \pi a^2/2)^k}{k!} \\
&= \lim_{a\to\infty} \exp\left\{-\sigma_r \pi a^2/2 \Pr\left(h(r_j) > z\right)\right\} \quad (4.22)
\end{aligned}$$

where $\Omega$ is the set of all receivers, and since we are looking for a receiver that maximizes the forward progress, we consider a sender-centric transmission model and we restrict to all receivers in a half disk $\mathbf{D}_a$ of radius $a$, the half disk in the direction of the destination, then we let $a \to \infty$, moreover note that we are using $\sigma_r$ the density of receivers, and $\theta$ being uniformly distributed over $[-\pi/2, \pi/2]$. One can write:

$$\begin{aligned}
\Pr\left(h(r_j) > z\right) &= \int_0^a \Pr\left(r\psi(r)\cos(\theta) > z | r\right) f_r(r) \mathrm{d}r \\
&= \frac{4}{\pi a^2} \int_0^a \int_0^{\pi/2} r \mathbf{1}_{\left\{\theta < \arccos\left(\frac{z}{r\psi(r)}\right)\right\}} \mathrm{d}r \mathrm{d}\theta \\
&= \frac{4}{\pi a^2} \int_{\{r\geq 0: z/r\psi(r)<1\}} r \arccos(z/r\psi(r)) \mathrm{d}r \quad (4.23)
\end{aligned}$$

By combining the latter to Eq.(4.21), we obtain:

$$\Pr\left(Z_1 \leq z\right) = e^{-2\sigma_r \left[\int_{\{r\geq 0: z/r\psi(r)<1\}} r \arccos(z/r\psi(r)) \mathrm{d}r\right]} \quad (4.24)$$

which leads to:

$$E[Z_1] = \int_0^1 1 - e^{-2\sigma_r \left[\int_{\{r\geq 0: z/r\psi(r)<1\}} r \arccos(z/r\psi(r)) \mathrm{d}r\right]} \mathrm{d}z \quad (4.25)$$

One can notice that the optimization above (in computing the maximal expected forward progress) is obtained by averaging over all Poisson configurations for the location of interfering nodes and of the receiver node.

### 4.4.2 Expected Forward Progress for the Closest Node in Range Strategy (RS2)

We need first to derive the probability density function of the minimum distance between the transmitter and the receiver among all the receive node distances (in a half disk of radius $a$). By using order statistics (see for example [77]), we have (see Appendix for the proof):

$$f_r^{'}(r) = \sigma_r \pi r e^{-\sigma_r \pi \frac{r^2}{2}}$$

Moreover, the average distance of a node to its closest neighboring node is given by (using again the Gamma function $\Gamma(y) = \int_0^\infty x^{y-1}e^{-x}\mathrm{d}x$) :

$$\begin{aligned} E[r] &= \int_0^\infty \sigma_r \pi r^2 e^{-\sigma_r \pi \frac{r^2}{2}} \\ &= \frac{1}{\sqrt{2\sigma_r}} \end{aligned} \quad (4.26)$$

Remember that the node density is the number of nodes per area, and for an unit area, the average nearest neighbor distance is on the order of $O\left(\frac{1}{\sqrt{n}}\right)$, which is similar to the strategy used in [26] as explained above. For this strategy, we have:

$$\begin{aligned} \Pr(Z_2 \leq z) &= \int_0^\infty \Pr\left(r\psi(r)\cos(\theta) \leq z|r\right) f_r^{'}(r)\mathrm{d}r \quad (4.27) \\ &= \frac{2}{\pi}\int_0^\infty \int_0^{\pi/2} \sigma_r \pi r e^{-\sigma_r \pi \frac{r^2}{2}} \mathbf{1}_{\left\{\theta \geq \arccos\left(\frac{z}{r\psi(r)}\right)\right\}} \mathrm{d}r\mathrm{d}\theta \\ &= 1 - \int_{\{r\geq 0: z/r\psi(r)<1\}} 2r\sigma_r \arccos(z/r\psi(r)) e^{-\sigma_r \pi \frac{r^2}{2}} \mathrm{d}r \end{aligned}$$

The expected forward progress becomes:

$$E[Z_2] = \int_0^1 \int_{\{r\geq 0: z/r\psi(r)<1\}} 2r\sigma_r \arccos\left(\frac{z}{r\psi(r)}\right) e^{-\sigma_r \pi \frac{r^2}{2}} \mathrm{d}r\mathrm{d}z \quad (4.28)$$

### 4.4.3 Channel Driven Maximal Expected Forward Progress (RS3)

A trait of wireless channels is the time variation of the channel strength. As a result, when many users are present, different users will experience peaks in their channel quality at different times. Taking benefit from this channel characteristic is called multi-user diversity and was initially exploited in [78] where a power control scheme was presented for maximizing the uplink information theoretic capacity in a single cell system. It was shown that capacity is maximized by allowing to transmit, at any given time, only the user with the best channel. Moreover, in an ad hoc network context, multi-user diversity was studied in [26] where the transmission strategy benefits from multi-user diversity by distributing packets to many different intermediate nodes which have independent time varying channels to the final destination.

In our setting, since there is no centralized control, multi-user diversity will be exploited in a different manner than [78]. Assume that the instantaneous channel state information of the neighbors is available at the transmitter, so that it could select the next hop the furthest away and with the best channel. Diversity gain arises from the fact that with many neighbors whose channel vary independently, it is likely that there is a receiver with a very good channel (from the transmitter to that node) at any one time. Then, the forward progress is a function of the distance and the channel between the transmitter and the next hop $h(r,\gamma) = r\cos(\theta)\psi(r,\gamma)$, and the outage probability will be conditioned on the instantaneous channel. Moreover, we are assuming that the fading is constant over all the retransmission rounds, and changes independently when the transmission of the current information message is stopped. We can still write Eq.(4.13), Eq.(4.17) and derive throughput formulas by computing the outage probability conditioned on the knowledge of the channel at the transmitter:

$$\begin{aligned} P_{out}(r_{i,j},\gamma_{i,j}) &= \Pr\left(\log_2\left(1+\frac{\gamma_{i,j}Pr_{i,j}^{-\alpha}}{\sigma_n^2+V}\right)\leq R\middle|\gamma_{i,j}\right) \\ &= 1-\Pr\left(0\leq V\leq\frac{\gamma_{i,j}r_{i,j}^{-\alpha}P-(2^R-1)\sigma_n^2}{2^R-1}\middle|\gamma_{i,j}\right) \end{aligned} \tag{4.29}$$

The following result in [79] (*pp.436*) on stable distributions with exponent $\frac{1}{2}$ is useful.

**Lemma 4.1.** *The cumulative distribution function:*

$$G(x) = 2[1-\mathcal{R}(1/\sqrt{x})], \quad x>0 \tag{4.30}$$

*(where $\mathcal{R}$ is the standard normal cumulative distribution function) has the Laplace transform:*

$$\phi(\lambda) = e^{-\sqrt{2\lambda}} \tag{4.31}$$

From Eq.(4.10), the MGF of $V$ is given by:

$$\begin{aligned} \phi_V(t) &= e^{\frac{-2\sigma_t\pi\Gamma(2/\alpha)\Gamma(1-2/\alpha)(Pt)^{2/\alpha}}{\alpha}} \\ &= e^{-\frac{\sigma_t\pi^2\sqrt{Pt}}{2}}, \quad \text{for } \alpha=4 \end{aligned} \tag{4.32}$$

Moreover, from Lemma 4.1, it is straightforward to notice that $\phi^n(\lambda) = \phi(n^2\lambda)$, which is the same as $G^{n\star}(x) = G(n^{-2}x)$[1]. The cumulative distribution of $V$ is then given by:

$$F_V(x) = \text{erfc}\left(\frac{\sigma_t \pi^2 \sqrt{P}}{4\sqrt{x}}\right) \tag{4.33}$$

leading to (note that for $\alpha \neq 4$, $\alpha > 2$, the cdf of $V$ can be computed as infinite series from [79]):

$$P_{out}(r_{i,j}, \gamma_{i,j}) = \text{erf}\left(\frac{\sigma_t \pi^2 \sqrt{2^R - 1}}{4\sqrt{\gamma_{i,j} r_{i,j}^{-\alpha} - (2^R - 1)\sigma_n^2/P}}\right) \tag{4.34}$$

Computing $\Pr\left(Z_3 = \max_{j \in \Omega} h(r_j, \gamma_j) \leq z\right)$ and following the steps of Eq.(4.21), Eq.(4.25), we obtain (where $l(r,\gamma) = \frac{z}{r\psi(r,\gamma)}$):

$$E[Z_3] = 1 - \int_0^1 e^{-2\sigma_r \left[\int\int_{\{r \geq 0 : l(r,\gamma) < 1\}} e^{-\gamma} r \arccos(l(r,\gamma)) \mathrm{d}r \mathrm{d}\gamma\right]} \mathrm{d}z \tag{4.35}$$

### 4.4.4 Numerical Results

The throughput is expressed as a function of different system parameters: the received SNR at one meter $\frac{P}{\sigma_n^2}$, the target information rate $R$ (PHY), the transmit probability $p$ (MAC) and the node density $\sigma$ (topology). The optimal throughput is derived by maximizing over the transmit probability $p$ (where Eq.(4.25), Eq.(4.28), Eq.(4.35) are solved by using numerical integration). In Slotted Aloha, since the throughput for very high target information rate $R$ goes to zero and the throughput is zero for ($R = 0$), there exists an optimal target information rate $R$ given a node density, transmit probability and transmit SNR as shown in Fig.4.5 (actually one can say that the mean delay $\tau$ is growing faster than $R$ which from Eq.(4.12) leads to a zero throughput). Note this is the network optimal target information rate. It implies that most of the nodes will not operate at their own optimal target information. Incremental redundancy is capacity achieving since it benefits from the accumulation of information (this process permits some averaging of the fading and interference affecting the useful signal). *Ergodic* in Fig.4.5 stands for the case where we replace $\psi(r)$ by $E[I_s]$ in Eq.(4.25)

[1]The cdf $G^{n\star}(x)$ has the n-fold convolution of $g$ with itself for pdf, where $g$ is the pdf of $G(x)$

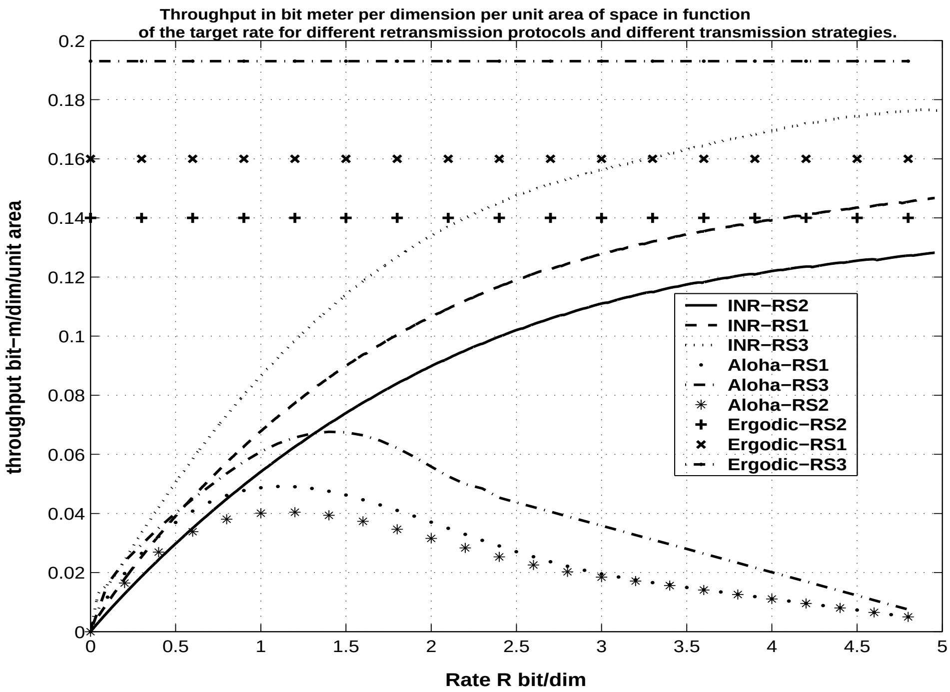


**Figure 4.5:** The Spatial throughput (in bit-meter per dimension for different retransmission protocols and transmissions strategies. Transmit SNR= $5$ dB, node density $\sigma = 1$, power loss exponent $\alpha = 4$.

Eq.(4.28) and $E[I_s|\gamma]$ in Eq.(4.35) where $E[I_s]$ is the ergodic capacity, i.e. the maximum achievable throughput on this channel, without feedback or delay constraints ($I_s$ is defined in Eq.(4.4)). Note that $E[I_s]$ in (**RS1**) and (**RS2**) is in function of $r$, the distance between the transmitter and the receiver. This implies that for a particular transmitter receiver distance, incremental redundancy converges to the same throughput $E[I_s]$ ( remember that incremental redundancy has this flavor " the more we wait the better it is", and thanks to the frequency hopping setting considered, incremental redundancy converges to an ergodic capacity since"all" channel realizations are visited). The spatial capacity is computed then by averaging over all network realizations. In the case of (**RS3**), note in that case the dependance of $E[I_s|\gamma]$ on the channel and the distance between the transmitter and the receiver, it means that for a particular channel realization and particular network realization incremental redundancy converges to $E[I_s|\gamma]$. The spatial throughput is obtained through averaging over channel and network realizations. In contrast to slotted Aloha, and since the optimal spectral efficiency in incremental redundancy is very high (infinite), by choosing a high spectral efficiency, the network will allow the nodes to operate at very good throughputs.

Moreover, we consider a rate adaptive approach where the target rate information is a function of the instantaneous channel state information at the transmitter ( i.e $R = R(\gamma)$). The average throughput is $\int R(\gamma) f_{\gamma}(\gamma) = E[I_s]$. This means that incremental redundancy with very low-rate feedback and without channel state knowledge at the transmitter has the same long-term benefit as a (complex) rate adaptive approach that requires an instantaneous channel state information at the transmitter and feedback.

The closest node in range strategy (in a microscopic analysis) performs worse than the maximal forward progress strategy, this is in contrast to the results stemming from the Gupta Kumar model where communication is limited to nearest neighbors. The maximal expected forward strategy permits the computation of the optimal hop (or relay) distance. The channel driven strategy performs substantially better than the other strategies by exploiting transmissions only to nodes with instantaneously good channels. On each hop, the link capacity is then maximized. By this strategy, we are optimizing the spatial concurrency and the spectral efficiency of each link by exploiting multi-user diversity. This suggests that routing should be based on the instantaneous channel strength of the link, which could require fast route updates (in comparison to existing routing protocols for ad hoc networks) if the channel changes rapidly. In the case of the maximal expected forward progress, routing is based on a *spatial* empirical average of the SINR's at the transmitter among the nodes in its proximity. This is reasonably simple for slowly varying channels and could be included in existing routing protocols.

Finally the impact of node density on the throughput was studied in [11]. It is shown that the the network throughput increases with node density up to a certain density threshold. This is due to the improvement of the node connectivity. However, for a high value of node density, the interferences become too limiting and a decrease of the throughput is noticed. This study is based on percolation theory results for the physical model as described in Chapter 2. It would be interesting to extend such results to our practical scenario.

### Numerical Application for a Practical Scenario

| Transmit power (dBm) | C (bit-m/dim) |
|---|---|
| 5 | 77.7 |
| 10 | 84.1 |
| 15 | 86.4 |
| 20 | 88.0 |
| 25 | 88.2 |
| 30 | 88.6 |
| 35 | 88.7 |

**Table 4.1:** The spatial throughput in bit-m/dim as a function of the transmit power in dBm for incremental redundancy (RS2).

For simulation, we consider the following scenario. The node density is chosen to be $\sigma = 0.1$, which corresponds to $100'000$ nodes for an area of $1\,\text{km}^2$. The nodes transmits over channels of bandwidth $W = 1$ MHz, and the total available bandwidth is of 100 MHz at a frequency carrier of 5.8 GHz. The noise variance is taken to be $-108$ dBm and the signal attenuation at a close-in reference distance is based on measures in [64]. The transmit power varies from 5 dBm to 35 dBm. The network is operating at a rate R=3 bit/dim and we want to assess the variations of the spatial throughput as a function of the transmit power. This approach could not be used for a slotted Aloha system as explained above, where there is an optimal rate R at which the network should operate. Whereas for incremental redundancy, one can trade the throughput for the delay since the higher is R, the higher is the achieved spatial throughput. From Tab. 4.1 and Tab. 4.2, one can see that the throughput saturates quickly with the transmit power. A transmit power of 20 dBm (100 mW) is reasonable. Also, we notice that the optimal channel access probability $p$ does not vary as we increase the transmit power over 20 dBm (note that in general $p$ decreases with increasing transmit power).

| Transmit power (dBm) | C (bit-m/dim) |
|---|---|
| 5 | 100.4 |
| 10 | 108.9 |
| 15 | 112.7 |
| 20 | 114.1 |
| 25 | 114.5 |
| 30 | 114.6 |
| 35 | 114.6 |

**Table 4.2:** The spatial throughput in bit-m/dim as a function of the transmit power in dBm for incremental redundancy (RS3).

### 4.4.5 Implementation

In this section, we are dealing with the implementation issues of the scheme described in this chapter. The main concern is the design of a proper mechanism that ensures the selection of the next hop as described by our routing strategies.

The first assumption stated in this chapter is the ability of the transmitting nodes to discover their routes or their neighborhood. This translates in practice into a routing protocol that relies on geographical locations of nodes in order to compute routes towards destination. For example, this a-priori information regarding neighbor and destination coordinates is used in geographical routing protocols as described in [63, 65]. A simple and efficient way to obtain these geographical locations is to assume that all nodes are equipped with Global Positioning System (GPS) receivers and capabilities. This GPS-based solution is realistic nowadays and easy to implement. One can argue that other location methods can be used for ad hoc wireless networks such as signal strength, angle of arrival and can be suitable for determining local information (i.e., information in the vicinity of the node). However all these methods would be expensive in terms of computations and bandwidth (exchange information overhead). Thus, GPS based location determination method can be an important parameter in reducing information overhead, thus simplifying the distribution of information and limiting infrastructure reliance. Moreover, in case not all the nodes have a GPS receiver, one can design a procedure so that the subset of nodes

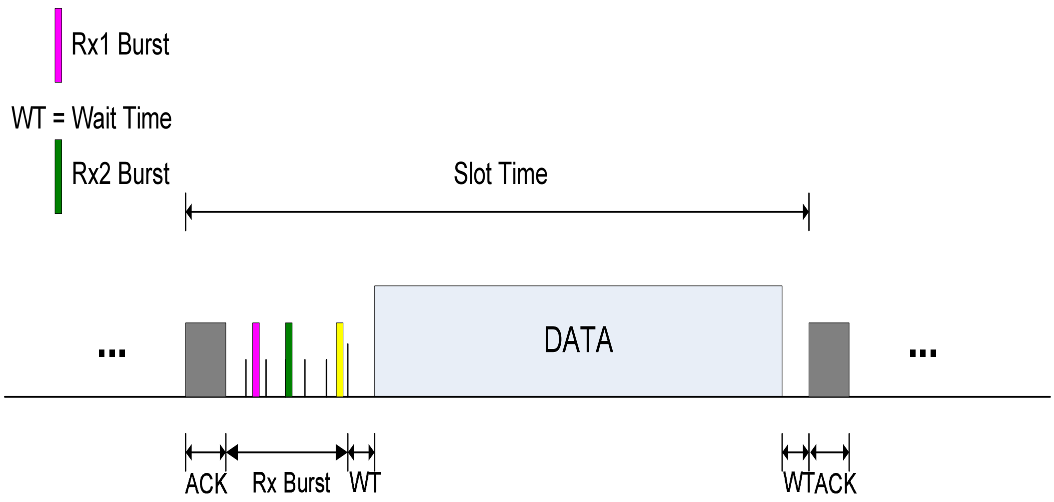


**Figure 4.6:** Slot format for the case where the transmitter selects the next hop.

with GPS capability succeed in supporting the maximum possible number of nodes without GPS capability. As a result local determination is fully enabled.

Once the neighbor and destination coordinates are available at the transmitter, the latter computes the next hop (or the next relay) towards the destination. The computation of the next hop could be either determined by the transmitter (as described in Fig.4.6) or by the receiver at the next hop (as described in Fig.4.7). For the first solution, the transmitter knowing the direction of the final destination and the coordinates of nodes in its vicinity, computes the forward progress (for all strategies described in this chapter) and determines the best node to be the next hop. However this computation is based on the knowledge of the nodes locations, their operating mode (MAC mode, either transmitter or receiver) and an information on the node density. Indeed the transmitter selects the best hop among the receivers. If the station elected as the next hop happens to be in transmit mode, a collision occurs. To cope with this problem, we can include in the spatial throughput derivations this event as a collision with certain probability. Another approach described in Fig.4.6 relies on the signal bursts sent by the receiver nodes. A portion of the slot time is dedicated to these signal bursts. Just before data transmission in the next slot, the nodes that will operate in the receive mode in the next slot will send some burst signals or

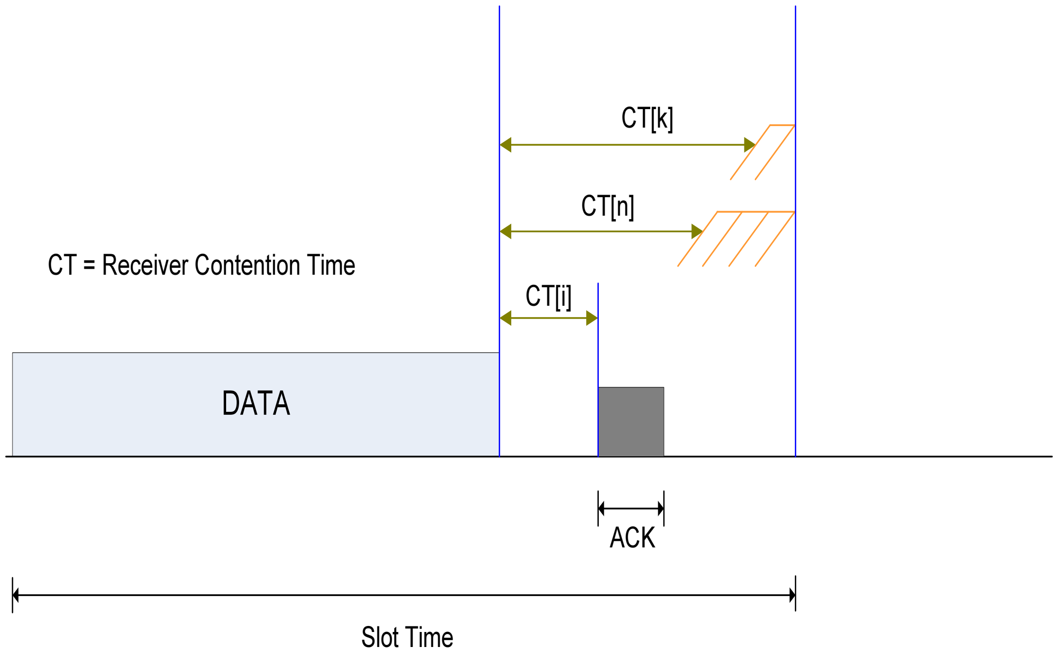


**Figure 4.7:** Slot format for the case where the receiver elects itself as the next hop.

pilots in a broadcast manner. Thus all the transmit nodes in the vicinity will update their receive nodes table in order to keep track of all the potential relay nodes in the next slot. These signal bursts are sent just after the ACK/NACK signaling while all the potential transmitters are listening. In order to avoid collisions between different nodes burst, one can imagine that each receive node has a position within the time allocated for these signals in the slot format. This is similar to a PPM modulation, making all these signals orthogonal at the transmitting nodes, and any other physical layer solution (e.g. OFDM) is also suitable. This scheme has other practical advantages. It ensures a simple and efficient way to estimate the node density at the transmitting nodes, making the computation of the forward progress easy. These received burst signals (or pilots) allow also the computation of a spatial empirical average of the SINR's at the transmitting nodes. This is crucial for the routing strategy (**RS3**).

The second solution, described in Fig.4.7, is a receiver oriented approach. The receiver that realizes the optimal forward progress (depending on the routing strategy used) elects itself as the next hop or the relay for the current transmission. The issue raised by this solution is how a potential relay notifies the other receivers about the forward progress realized for the current transmission. This is important since it prevents collisions between all the ACK signaling from all the receivers (potential relays). The idea one can use to solve this problem is inspired from backoff algorithms and contention windows used in 802.11 protocols. As depicted in Fig.4.7, receiver[$i$] has to wait for a duration of time $CT[i]$ before sending the ACK signaling. This duration is closely linked (for example inversely proportional) to the forward progress realized by the corresponding receiver. Thus the receiver corresponding to the maximal forward progress will be first in accessing the channel and transmitting the ACK signaling. This prevents all the nodes listening to this ACK signaling from sending ACK signals to the corresponding transmitter node and stops the election process of the next hop node.

A more detailed study of such schemes is needed and the impact of these signaling schemes on the overhead and the throughput must be taken into account in a more rigorous manner. This is beyond the scope of this work.

## 4.5 Conclusions and Future Work

We have derived formulas for the spatial throughput for retransmission protocols and transmission strategies for random networks described by a spatial Poisson point process. It is shown that coding and retransmission protocols are a viable and simple solution for providing fully decentralized multiple-access communications in ad hoc wireless networks despite harsh propagation characteristics (interference from nearby competing nodes). Random exclusion and a decentralized protocol allow for the mitigation of the interference coming from other nodes. A routing protocol aiming to maximize the expected forward progress and exploiting multi-user diversity is shown to significantly out-perform other schemes. Future work will focus on more advanced strategies for cooperation, the analysis of multi-user detection techniques, practical coding strategies and distributed synchronization methods.

## Appendix 4.A Proof

The probability distribution function of the distance between a transmitter and a receiver in a disk of radius $b$ is given in Eq.(4.9). By using order statistics, one can write the probability distribution of the minimum distance between a transmitter and a receiver among $k$ transmit-receive nodes distances (in a half disk of radius $b/2$:

$$f_r(r|k \text{ in } \mathbf{D}_{b^{'}}) = \frac{2kr}{b^2}\left(1-\left(\frac{r}{b}\right)^2\right)^{k-1} \tag{4.36}$$

Then, we compute the probability distribution function of the minimum distance between a transmitter and a receiver among all transmit-receive nodes distances (with unconstrained range):

$$\begin{aligned} f_r^{'}(r) &= \sum_{k=1}^{\infty} f_r(r|k \text{ in } \mathbf{D}_{b^{'}})\,\Pr(k \text{ in } \mathbf{D}_{b^{'}}) \\ &= \sum_{k=1}^{\infty} \frac{2kr}{b^2}\left(1-\left(\frac{r}{b}\right)^2\right)^{k-1} \frac{(\sigma_r \pi b^2/2)^k}{k!} e^{-(\sigma_r \pi b^2/2)} \\ &= -e^{-(\sigma_r \pi b^2/2)} \sum_{k=1}^{\infty} \frac{\mathrm{d}\left(1-\left(\frac{r}{b}\right)^2\right)^k}{\mathrm{d}r} \frac{(\sigma_r \pi b^2/2)^k}{k!} \\ &= -e^{-(\sigma_r \pi b^2/2)} \frac{\mathrm{d}\left(e^{\left(1-\left(\frac{r}{b}\right)^2\right)}(\sigma_r \pi b^2/2)\right)}{\mathrm{d}r} \\ &= \sigma_r \pi r e^{-\frac{\sigma_r \pi r^2}{2}} \end{aligned} \tag{4.37}$$

# Chapter 5

# ARQ Based Half-Duplex Cooperative Diversity Protocol

## 5.1 Introduction

The landmark paper of Gupta-Kumar [9] has driven interest in wireless sensor and ad hoc networks. The constraints on the size of the terminals in such ad hoc networks mitigates the presence of multiple antennas and full duplex transmissions. In such a scenario, distributed antennas can be used to provide a mean to combat fading with a similar flavor as that of space diversity. Another application scenario that has great potential is cellular networks. For uplink transmission, from an end user to a base station (access point), a relay can forward the end user message to the base station. The motivation comes from the fact that the end user is close to the cell boundary, and direct transmission requires high power transmission. Moreover the RF technology used is kept simple by using one antenna at the end user preventing the benefit of the promising space-time techniques.

This kind of reliability obtained by the creation of virtual antennas is referred to as cooperative diversity because the terminals share their resources to get the information across to the destination. In this manner, spatial diversity gain can be obtained even when a local antenna array is not available. Cooperative diversity is also very useful when the propagation environment changes slowly as compared to the signaling rate or if the bandwidth of the channel input is less than the coherence bandwidth. In these scenarios, we can not benefit from frequency and time diversity.

Cooperative schemes have attracted significant attention recently, and a variety of cooperation protocols have been studied and analyzed in various papers. The information-theoretic relay channel was first studied by van der Meulen [80], and some of the most important capacity results on relaying were published in [81]. The distinctive property of relay channels in general is that certain terminals, called "relays", receive, process, and re-transmit some information bearing signal(s) of interest in order to improve performance of the system. Cover and El Gamal [81] examine certain non-faded relay channels, developing lower and upper bounds on the channel capacity via random coding and converse arguments, respectively. Generally these lower and upper bounds do not coincide, except in the class of degraded relay channels. The lower bounds on capacity, i.e., achievable rates, are obtained via three structurally different random coding schemes, referred to in [81] as facilitation, cooperation, and a quantization based cooperation scheme. In the facilitation scheme, the relay does not actively help the source, but rather, facilitates the source transmission by inducing as little interference as possible. In the cooperation scheme, the relay fully decodes the source message, and retransmits some information about that signal to the destination. More precisely, the relay encodes the bin index of the previous source message, from a random binning of the source messages as in well-known Slepian-Wolf coding [82]. The source transmits the superposition of a new encoded message and the encoded bin index of the previous message, in a block- Markov fashion. The destination suitably combines the source and relay transmissions, possibly coherently combining the identical bin index transmissions, in order to achieve higher rates than with the direct transmission alone. Of course, full decoding at the relay can, in some circumstances, be a limiting factor; the rates achieved using this form of cooperation are no greater than the capacity of direct transmission from the source to the relay. As one alternative in such circumstances, Cover and El Gamal propose another scheme, in which the relay encodes a quantized version of its received signal. The destination combines information about the relay received signal with its own in order to form a better estimate of the source message. For Gaussian noise channels, the destination can essentially average to two observations of the source message, thereby reducing the noise. Broadly speaking, we can expect cooperation (resp. quantization based cooperation) to be most beneficial when the channel between the source and relay (resp. relay and destination) is particularly good. For intermediate regimes, Cover and El Gamal propose superposition of the two schemes in order to maximize the achievable rates. A comprehensive review of past work on the relay channel and related problems appears in [83], and new information theoretic results can be found in [83, 84, 85] with extensions to multiple relays. The idea of cooperative diversity was pioneered in [86, 87] where the transmitters repeat detected symbols from each other to increase their rate region. Therein, the feasibility of user cooperation in a wireless network is demonstrated by an information theoretic exposition of the gains and a practical CDMA implementation. Taking into account practical constraints such as half-duplex transmission, and channel state information available only at the receiver (preventing from exploiting coherent transmission and combining), low-complexity cooperative diversity protocols are analyzed in [88]. Although previous work focuses primarily on ergodic settings and characterizes performance via Shannon capacity or capacity region, the analysis of these protocols (such as amplify-and-forward, decode-and forward) in [88] is considered in terms of outage and diversity. The proposed schemes in [88] are based on a time division strategy, where the two users rely on the use of orthogonal subspaces to forward each other's signals. In [89], the assumption on orthogonal subspaces is relaxed by allowing the source to continuously transmit over the whole duration of the codeword. The proposed schemes are evaluated in terms of diversity multiplexing tradeoff and extended to the case of multiple relays.

In these studies, there appear to be two general classes of approaches to relay processing. In one class, the relay decodes the source message and retransmits some information about the message. We refer to techniques in this class broadly as *decode-and- forward* schemes. For example, the relay might decode the message and simply repeat the transmission, or it might transmit additional parity bits about the message, as in the cooperation scheme of [81]. In the other class, the relay tries to convey a representation of its received signal to the destination, so that the destination can effectively combine two receive signals and decode the message. For example, the relay might simply amplify its received signal in the case of an *amplify-and-forward*, or it might quantize or rate-distortion code its received signal and encode for transmission to the destination [81].

Recently, the authors of [97] extended the Zheng-Tse formulation [95] and characterized the three dimensional diversity-multiplexing-delay tradeoff in MIMO Automatic Retransmission reQuest (ARQ) channels. They established that delay can be exploited as a potential source for diversity. Thus, retransmission protocol is an appealing scheme to combat fading and its performance has been studied in decentralized ad hoc networks [99]. Inspired by [97], we propose a new scheme for transmission in relay channel utilizing the ARQ to increase the diversity gain. We look at the tradeoff in the high SNR (Signal-to-Noise Ratio) regime and point out the gain achieved by the ARQ.

Following the setup in [88], the terminals are constrained to employ half-duplex transmission, i.e. they cannot transmit and receive simultaneously. The source and the relay are allowed to transmit in the same channel using cooperative protocols not relying on orthogonal subspaces, allowing for a more efficient use of resources. This is in contrast to [88], where the available bandwidth is divided into orthogonal channels allocated to the transmitting terminals. Fig.5.1 and Fig.5.2 depict the difference in the operation mode between protocols relying on orthogonal and non-orthogonal channels.

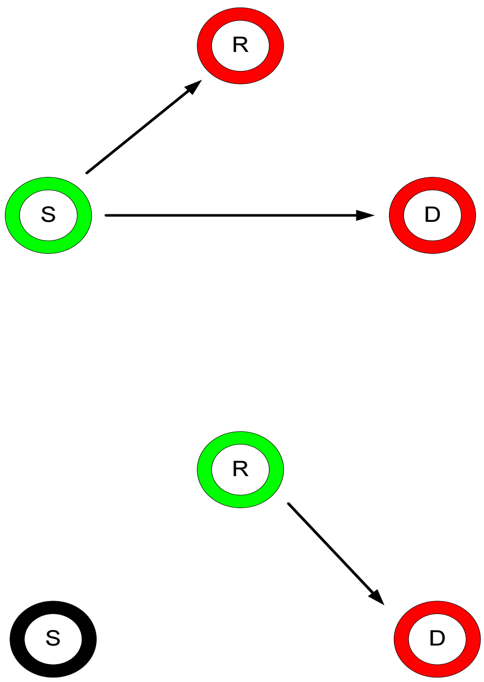


**Figure 5.1:** Orthogonal channels based cooperative protocol.

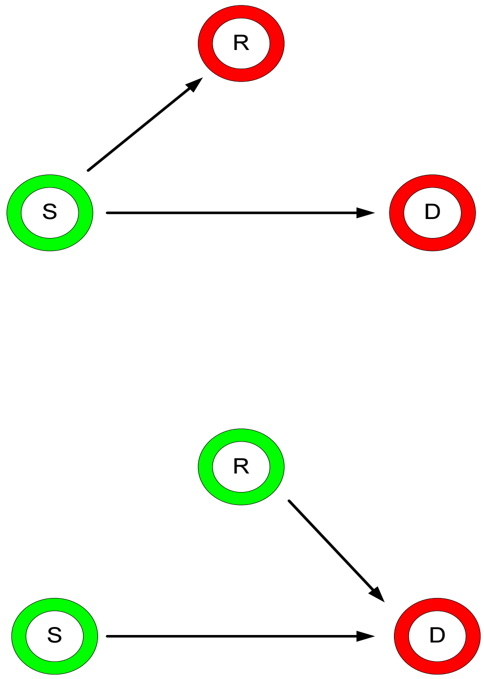


**Figure 5.2:** Non-orthogonal channels based cooperative protocol.

In the dynamic decode and forward scheme proposed in [89] the communication is across one block of fixed length $l$, where $l$ is asymptotically large. In our setting introduced in [91, 90], the ARQ permits the use of communication over a variable number of blocks (henceforth referred to as number of rounds) of fixed length where the number of blocks used depend on the quality of the channel and are upper bounded by a fixed number $L$. If the destination is not able to decode at the end of these $L$ blocks an outage is declared. Similarly, the relay accumulates enough information before it starts cooperating with the source. If the source-relay channel is always in outage, the relay may not be able to decode and thus it will not forward the source message. This scheme is in contrast to the ARQ-DDF (dynamic decode and forward) protocol used in [98] in a multiple access channel with two users, where once a user transmits its message successfully, it can cooperate with the other user on each round, using the DDF strategy. On each round the relay (in this case the user that successfully transmitted its message) will be able to decode the source message in $l^{'} < l$ symbols and will transmit the encoded message using an independent code-book during the rest of the codeword. In our case the number of rounds the relay will collaborate with the source by repeating its signal is random and depends on the quality of the source relay channel. The scheme proposed assesses the role of ARQ temporal diversity (by considering two dynamics of the channel: long-term static and short-term static channel) and cooperative diversity and the results are derived in terms of diversity multiplexing delay tradeoff at high SNR. Finally, a long-term power constraint is assumed in order to highlight the potential gain from a deterministic power control strategy based on [97]. The outline of this chapter is as follows. Section 5.2 contains a summary of the useful results and notations used in the rest of the chapter. We introduce the channel model and the details of the algorithm in Section 5.3. The actual tradeoff for this protocol is analyzed and presented in Section 5.4 for both long-term and short-term quasi-static channels. Section 5.5 proposes a power control scheme for ARQ relay protocol. Finally we summarize and present a few concluding remarks and future directions in Section 5.6.

## 5.2 Background

### 5.2.1 Notation

The symbol $\doteq$ will be used to denote the exponential quality, i.e. $f(\text{SNR}) \doteq \text{SNR}^b$ to denote:

$$\lim_{\text{SNR}\to\infty} \frac{\log f(\text{SNR})}{\log \text{SNR}} = b$$

and similarly for $\dot{\leq}$ and $\dot{\geq}$. $(x)^+$ means $\max(0, x)$. $\mathcal{R}^{n+}$ denotes the set of real $n$-vectors with nonnegative elements, and $\mathcal{A}^+ = \mathcal{A} \cap \mathcal{R}^{n+}$.

### 5.2.2 Capacity-vs.-Outage

The notion of capacity-vs.-outage examines the tradeoff between a fixed rate and the probability that rate is achievable over the composite channel. For example, for a non-ergodic Gaussian fading channel, for a fixed rate $R$ certain channel realizations $h$ will support the rate, i.e., those with $\log\left(1 + |h|^2 P/N_0\right) \geq R$, and other channel realizations will not support the rate, i.e., those with $\log\left(1 + |h|^2 P/N_0\right) < R$. The event $\log\left(1 + |h|^2 P/N_0\right) < R$ is referred to as an outage event, and the probability of this event is referred to as the outage probability of the channel.

The capacity-vs.-outage at outage probability $\alpha$ is defined to be the maximum rate for which the outage probability is less than $\alpha$. Delay-limited capacity is the special case of capacity-vs.-outage corresponding to zero outage. Capacity-vs.-outage was introduced in [68] to examine the performance of certain cellular systems with delay constraints. It is intimately related to the more general and precise $\epsilon$-capacity framework of Verdu and Han [92], and this relationship was solidified in the work of Caire, Taricco, and Biglieri [101]. Both [68], [101] extend the notion to handle block-fading models with delay constraints limiting the number of blocks available for transmission.

For systems with tighter delay constraints, the channel may not exhibit its ergodic nature within a coding interval, so that the Shannon capacity is zero. In such cases, alternative performance metrics such as capacity-vs.-outage/outage probability [68], [93] or delay-limited capacity [94] can be employed.

### 5.2.3 Diversity-Multiplexing Tradeoff (DMT)

The trade-off between diversity and multiplexing was formally defined and studied in the context of point-to-point coherent communications in [95]. A family of codes $\mathcal{C}(\mathrm{SNR})$ of block length $T$, with one code for each $SNR$ level, is said to have a diversity gain of $d$ and spatial multiplexing gain of $r$ if

$$r = \lim_{\mathrm{SNR}\to\infty} \frac{R(\mathrm{SNR})}{\log \mathrm{SNR}} \quad , \quad d = -\lim_{\mathrm{SNR}\to\infty} \frac{\log P_e(\mathrm{SNR})}{\log \mathrm{SNR}}$$

where $R(\mathrm{SNR})$ is the data rate measured in bits per channel use (BPCU) and $P_e(\mathrm{SNR})$ is the average error probability using the maximum likelihood (ML) decoder. For a coherent MIMO channel with $M$ transmit antenna and $N$ receive antenna, and for any multiplexing gain $r \leq \min\{M, N\}$ the optimal diversity gain $d(r)$ is given by the piecewise linear function joining the points $(K, (M-K)(N-K))$ for $K = 0, \ldots, \min\{M, N\}$. $d(r)$ is achieved by the random Gaussian i.i.d code ensemble for all block lengths $T \geq M + N - 1$. This is depicted in Fig.5.3.

Note that the ergodic capacity results suggest that dramatic increases in capacity are possible using multi-antenna systems. For example, for the case of channel state information available to the receiver only, the ergodic capacity increases by $\min\{M, N\}$ b/s/Hz for each additional 3 dB of SNR, in the high SNR regime [96].

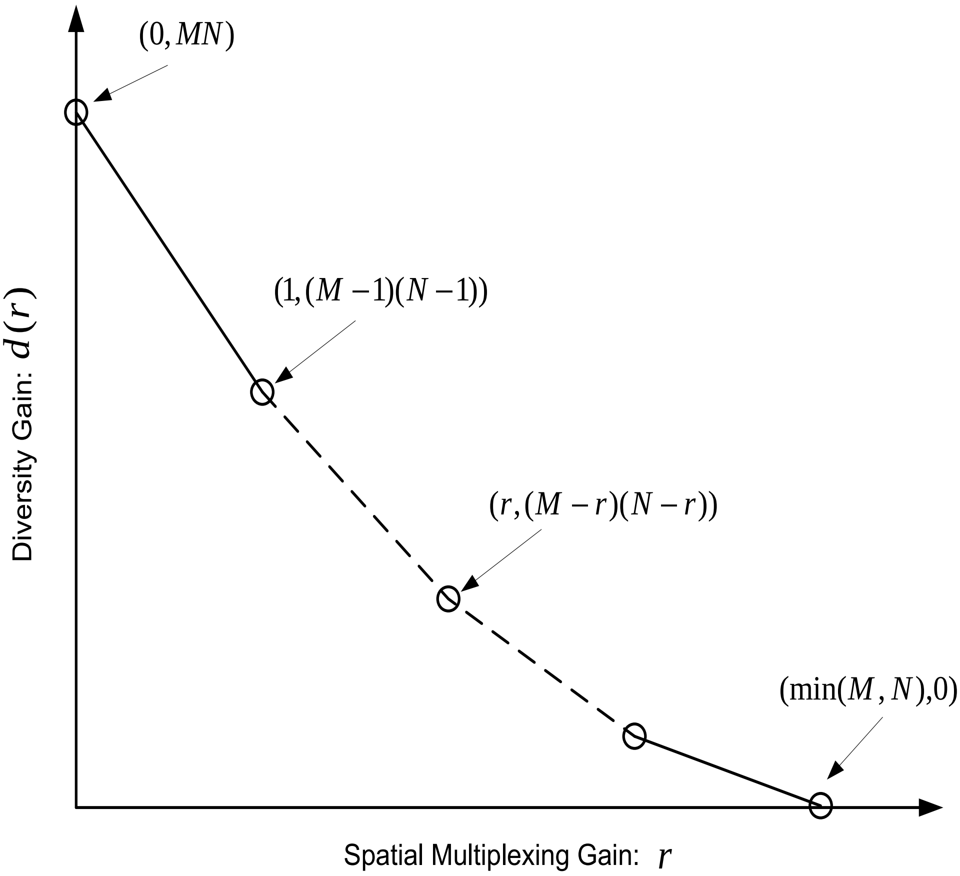


**Figure 5.3:** The optimal diversity multiplexing tradeoff of a MIMO system.

### 5.2.4 Useful Results

Let $h$ be a circularly symmetric complex Gaussian with zero mean and unit variance. $\gamma = |h|^2$ is exponentially distributed with unit mean. Defining $\mu = -\frac{\log \gamma}{\log \text{SNR}}$ we note that $\mu$ is distributed as,

$$f_\mu(\mu) = \log(\text{SNR})\text{SNR}^{-\mu} \exp(-\text{SNR}^{-\mu}) \tag{5.1}$$

which, in the high SNR gives:

$$f_\mu(\mu) \doteq \begin{cases} \text{SNR}^{-\mu} & \text{for } \mu \geq 0 \\ 0 & \text{for } \mu < 0 \end{cases} \tag{5.2}$$

At high SNR, we have $(1 + \text{SNR}\gamma) \doteq \text{SNR}^{(1-\mu)^+}$. Let us define $\mathcal{A}$ as the set describing the outage event. Then, for independent random variables

$\underline{\mu} = [\mu_1, \ldots, \mu_n]$, the outage probability is given by:

$$P_{out} \doteq \int_{\mathcal{A}} f(\underline{\mu}) \mathrm{d}\underline{\mu} \doteq \mathrm{SNR}^{-d} \tag{5.3}$$

where:

$$d = \inf_{(\mu_1, \ldots, \mu_n) \in \mathcal{A}^+} \sum_{j=1}^{n} \mu_j \tag{5.4}$$

From this, it follows that:

$$\Pr(\log(1 + \gamma \mathrm{SNR}) < r \log(\mathrm{SNR})) \doteq \mathrm{SNR}^{-(1-r)} \tag{5.5}$$

and the following results can be obtained:

$$\begin{aligned} &\Pr\left(\sum_{i=1}^{l} \log(1 + \gamma_i \mathrm{SNR}) < r \log(\mathrm{SNR})\right) \\ &\doteq \Pr\left(\sum_{i=1}^{l} (1 - \mu_i)^+ < r\right) \\ &\doteq \int_{\mathcal{A}} \mathrm{SNR}^{-\sum_{i=1}^{l} \mu_i} \mathrm{d}\underline{\mu} \\ &\doteq \mathrm{SNR}^{-d} \end{aligned} \tag{5.6}$$

where the set $\mathcal{A} = \{\underline{\mu} : \sum_{i=1}^{l} (1 - \mu_i)^+ < r\}$ describes the outage event, $\mathcal{A}^+ = \{\underline{\mu} \in \mathcal{R}^{n+} : \sum_{i=1}^{l} (1 - \mu_i)^+ < r\}$ and $d$ is given by:

$$d \doteq \inf_{\underline{\mu} \in \mathcal{A}^+} \sum_{i=1}^{l} \mu_i \doteq l\left(1 - \frac{r}{l}\right) \tag{5.7}$$

Similarly, we have:

$$\begin{aligned} &\Pr\left(l \log(1 + \gamma \mathrm{SNR}) < r \log(\mathrm{SNR})\right) \\ &\doteq \Pr\left(l(1 - \mu)^+ < r\right) \\ &\doteq \int_{\mathcal{A}} \mathrm{SNR}^{\mu} \mathrm{d}\mu \\ &\doteq \mathrm{SNR}^{-d} \end{aligned} \tag{5.8}$$

where the set $\mathcal{A} = \{\mu : l(1 - \mu)^+ < r\}$ describes the outage event, $\mathcal{A}^+ = \{\mu \in \mathcal{R}^{n+} : l(1 - \mu)^+ < r\}$ and $d$ is given by:

$$d \doteq \inf_{\mu \in \mathcal{A}^+} \mu \doteq 1 - \frac{r}{l} \tag{5.9}$$

## 5.3 System Model and Setting

In this work, we consider communication over a relay network with one relay node (R) assisting the transmission of a source(S) destination(D) pair as described in Fig.5.4. Each link has circularly symmetric complex Gaussian zero mean unit variance channel gain $h_{sd}, h_{sr}, h_{rd}$ corresponding to Rayleigh-fading channel, and the channel gains are mutually independent. The additive noises at the relay and the destination are mutually independent circularly symmetric white complex Gaussian. Nodes are operating in half-duplex mode, i.e, a node cannot transmit and receive simultaneously. Moreover, we assume that each decoder has perfect knowledge of the channel gain. Perfect channel state information at the receivers implies that the S-R channel is known to the relay node, while the individual S-D, R-D channels are known to the destination node. The channel state information (CSI) is assumed to be absent at the node which is transmitting. Because of the ARQ protocol, limited feedback is received by the transmitting nodes. Moreover, perfect synchronization is assumed between nodes, which requires some form of distributed pilot signals in practice.

We investigate two scenarios for the channel gains: 1) long-term static channel, where the fading is constant for all the channels over all retransmission (ARQ) rounds, and changes independently when the transmission of the current information message is stopped; 2) short-term static channel where the fading for all the channels is constant over each transmission round (or block) of the ARQ protocol and is an i.i.d process across successive rounds. For the short-term static channel, only an upper-bound on

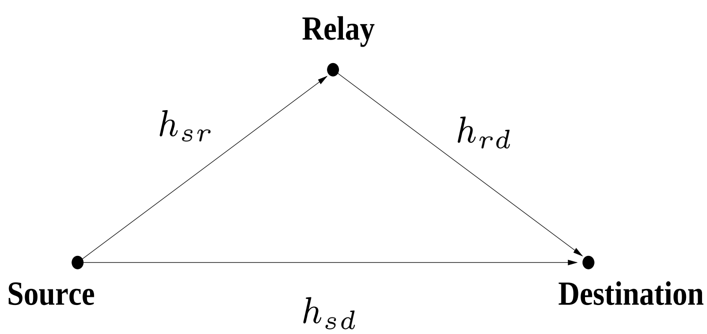


**Figure 5.4:** System Model

the diversity multiplexing delay tradeoff is derived. The ARQ protocol considered in this work is a form of incremental redundancy as studied in [99] [58]. The transmission queue at the source is assumed to be infinite (not concerned by stability issues). The information message of $b$ bits is encoded using a space-time code with code book $\mathcal{C} \subset \mathbb{C}^{2\times LT}$, where $T$ is the number of channel uses taken to transmit one round and $L$ is the maximum number of rounds that can be used to transmit the $b$ information bits. We let $\mathcal{C}_l$ for $l = 1, \cdots, L$ denote the punctured space-time code of length $lT$ obtained from $\mathcal{C}$ by deleting the last $(L-l)T$ columns of the space time code.

The protocol utilizes the ARQ as follows. The destination feeds back a one bit success/failure indication to both the relay and the source. If the relay decodes before the destination then knowing the codebook $\mathcal{C}$ it begins transmitting the second row of the codebook $\mathcal{C}$ to the destination. Thus effectively it becomes a MISO channel increasing the diversity. If the destination decodes before the relay, it just sends the feedback to the source and relay and the source moves on to transmitting the next message. We assume that the relay informs the destination of the starting of its transmission. The source moves on to the next information message in the transmission queue either if $L$ rounds have been exhausted for the message or if the destination sends success feedback. If successful decoding occurs at the $l$-th transmission, the effective coding rate for the current codeword is $R/l$ bit/dim where $R = b/T$. In incremental redundancy, the receiver has memory of the past signals since it accumulates mutual information.

As defined above, the information message is encoded by a space-time

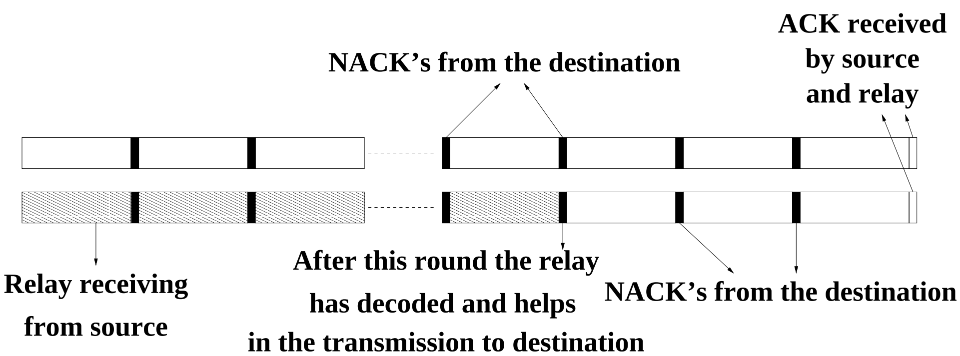


**Figure 5.5:** Message as seen by the destination.

encoder, and mapped in a sequence of $L$ blocks, $\{\mathbf{x}_l \in \mathbb{C}^{2\times T} : l = 1,\ldots,L\}$, and the transmission is as in a MIMO system, where the rows of $\mathbf{x}_l = \begin{bmatrix} \mathbf{x}_{sd;l} & \mathbf{x}_{rd;l} \end{bmatrix}^T$ are transmitted in parallel by the source and the relay. Each symbol of the transmitted codeword has unit power constraint. Let us call $\mathcal{T}_r$ a random variable denoting the block in which the relay was able to decode the source information message. Then, the signal model of our channel is given by:

$$\mathbf{y}_l^d = \sqrt{\frac{\text{SNR}}{2}}\mathbf{h}_l\mathbf{x}_l + \mathbf{n}_l^d \tag{5.10}$$

where $l$ stands for the retransmission round, $\{\mathbf{y}_l^d \in \mathbb{C}^{1\times T}\}$ is the received signal block by the destination, and $\{\mathbf{n}_l^d \in \mathbb{C}^{1\times T}\}$ is the channel noise assumed to be temporally and spatially white with i.i.d entries $\sim \mathcal{N}\mathbb{C}(0,1)$. The channel of the $l$-th round is characterized by the matrix $\{\mathbf{h}_l \in \mathbb{C}^{1\times 2}\}$ as follows:

$$\mathbf{h}_l = \begin{cases} [\ h_{sd} \quad 0\ ] & \text{if } l \in [1, \mathcal{T}_r] \\ [\ h_{sd} \quad h_{rd}\ ] & \text{if } l \in [\mathcal{T}_r + 1, L] \end{cases} \tag{5.11}$$

The received signal at the relay for $l = 1,\ldots,\mathcal{T}_r$ is given by:

$$\mathbf{y}_l^r = \sqrt{\frac{\text{SNR}}{2}} h_{sr;l}\mathbf{x}_{sd;l}^T + \mathbf{n}_l^r \tag{5.12}$$

Note that as $T \to \infty$ using random coding arguments we can find codebooks which are good depending on the instant $\mathcal{T}_r$ at which the relay decodes. It can be shown that by taking the intersection over all the codebooks which are optimal for each $\mathcal{T}_r$ and using random coding arguments we can choose a codebook which is optimal irrespective of the instant $\mathcal{T}_r$ when the relay decodes.

## 5.4 Tradeoff Curves

In this section we derive the tradeoff curves for the case of the long term quasi-static and short-term quasi-static channels. Since we are in the high SNR regime we ignore the factor 2 and use $\text{SNR} \doteq \frac{\text{SNR}}{2}$ for the remaining sections.

### 5.4.1 ARQ Protocol

The ARQ protocol considered in this chapter is based on incremental redundancy as studied in [58]. The destination accumulates information on successive rounds and successful decoding is performed by soft-combining all the received rounds. We define the effective rate in a different manner as follows. Let $\mathcal{T}_d$ be a random variable denoting the stopping time of the transmission of the current message at the destination. Let $\overline{\mathcal{O}_l}$ be the event that the mutual information per channel use at a particular decoder exceeds the transmission rate $R$, i.e, $\overline{\mathcal{O}}_l = \{\sum_{i=1}^{l} I_i > R\}$ for $l = 1, \ldots, L-1$, with $I_i$ being the mutual information of a single ARQ round as defined in Eq.(5.14), Eq.(5.15), Eq.(5.16). Then, we have:

$$\begin{aligned}
\Pr(\mathcal{T}_d = l) &= \Pr(\mathcal{O}_{d,1}, \ldots, \mathcal{O}_{d,l-1}, \overline{\mathcal{O}}_{d,l}) \\
&= \Pr(\mathcal{O}_{d,1}, \ldots, \mathcal{O}_{d,l-1}) - \Pr(\mathcal{O}_{d,1}, \ldots, \mathcal{O}_{d,l}) \\
&= \Pr(\mathcal{O}_{l-1}) - \Pr(\mathcal{O}_l) \qquad (5.13)
\end{aligned}$$

where we used the fact that the random sequence $I_l$ is non-decreasing with probability 1, and $\mathcal{O}_l \subseteq \mathcal{O}_m$ for $l \leq m$ leading to $\Pr(\mathcal{O}_1, \ldots, \mathcal{O}_l) = \Pr(\mathcal{O}_l)$. We have also $\Pr(\mathcal{O}_0) = 1$, and $\Pr(\mathcal{T}_d = L) = \Pr(\mathcal{O}_{d,L-1})$. In our relay channel scenario, the instantaneous mutual informations per channel use for the $j^{th}$ blocks are given by:

$$I_{s;d}^{j} = I^{j}(\mathbf{x}_{sd,j}; \mathbf{y}_j^d | h_{sd,j}) = \log(1 + \mathrm{SNR}\gamma_{sd,j}) \qquad (5.14)$$

$$\begin{aligned}
I_{s,r;d}^{j} &= I(\mathbf{x}_{sd,j}, \mathbf{x}_{rd,j}; \mathbf{y}_j^d | h_{sd,j}, h_{rd,j}) \qquad (5.15) \\
&= \log(1 + \mathrm{SNR}(\gamma_{sd,j} + \gamma_{rd,j})) \\
&= \log(1 + \mathrm{SNR}^{(1-\mu_{sd,j})} + \mathrm{SNR}^{(1-\mu_{rd,j})}) \\
&\doteq \log(\mathrm{SNR}^{(1-\min(\mu_{sd,j}, \mu_{rd,j}))^+})
\end{aligned}$$

$$I_{s;r}^{j} = I(\mathbf{x}_{sr,j}; \mathbf{y}_r^d | h_{sr,j}) = \log(1 + \mathrm{SNR}\gamma_{sr,j}) \qquad (5.16)$$

The throughput of an incremental redundancy ARQ based protocol is determined by the number of rounds needed for successful decoding, and it is defined as $\eta = R/\tau$, where $\tau$ is the average number of rounds needed for successful decoding. The effective multiplexing rate is then defined as,

$$\begin{aligned}
r_e &= \lim_{\mathrm{SNR}\to\infty} \frac{R(\mathrm{SNR})}{\left(\sum_{l=0}^{L-1} \Pr(\mathcal{O}_l)\right) \log \mathrm{SNR}} \\
&= \lim_{\mathrm{SNR}\to\infty} \frac{R(\mathrm{SNR})}{\left(1 + \sum_{l=1}^{L-1} \Pr(\mathcal{O}_l)\right) \log \mathrm{SNR}} \\
&\doteq \frac{r}{\left(1 + \sum_{l=1}^{L-1} \Pr(\mathcal{O}_l)\right)} \qquad (5.17)
\end{aligned}$$

### 5.4.2 Long-Term Static Channel

**Theorem 5.1.** For a long-term static channel the outage probability at the $l^{th}$ round for the proposed protocol is given by,

$$\Pr_{out}(l) \doteq \mathrm{SNR}^{-d^{lt}_{out}(r,l)} \tag{5.18}$$

where,

$$d^{lt}_{out}(r,l) = \begin{cases} (1-r) & \text{for } l=1 \\ (1-\frac{r}{l}) + (1-\frac{r}{l-1}) & \text{for } l \neq 1,3 \\ 2-5r/6 & \text{for } l=3, r<\frac{6}{7} \\ 3-2r & \text{for } l=3, r\geq\frac{6}{7} \end{cases} \tag{5.19}$$

**Proof Outline 1.** We used the fact that $\Pr_{out}(l) = \Pr(\mathcal{O}_l)$. For a long-term static channel, the instantaneous mutual informations per channel use do not vary from one round to another. Denote their common values as $I_{s;d}, I_{s,r;d}$ and $I_{s;r}$. At round $l$, the outage probability for this cooperative channel depends on the fact that the relay was able to decode the message from the source. Suppose that the relay decodes at time $k$ with probability given by (From Eq.(5.13)):

$$\begin{aligned} \Pr(\mathcal{T}_r = k) &= \Pr((k-1)I_{s;r} < r\log(\mathrm{SNR})) - \Pr(kI_{s;r} < r\log(\mathrm{SNR})) \\ &\doteq \mathrm{SNR}^{-(1-r/(k-1))} - \mathrm{SNR}^{-(1-r/k)} \end{aligned} \tag{5.20}$$

If $k < l$, the mutual information is then the sum of the contribution of the source during $k$ rounds (SISO) and the contribution from the source and the relay during $l-k$ rounds (MISO). However if $k \geq l$, the relay has no contribution to the information conveyed to the destination. The outage probability for the ARQ relay long-term static channel is,

$$\begin{aligned}
\mathrm{Pr}_{out}(l) &= \sum_{k=1}^{L} \mathrm{Pr}_{out|\mathcal{T}_r=k}(l)\mathrm{Pr}(\mathcal{T}_r = k) && (5.21)\\
&= \sum_{k=1}^{l-1} \mathrm{Pr}(kI_{s;d} + (l-k)I_{s,r;d} < r\log(\mathrm{SNR}))\mathrm{Pr}(\mathcal{T}_r = k)\\
&\quad + \sum_{k=l}^{L} \mathrm{Pr}(lI_{s;d} < r\log(\mathrm{SNR}))\mathrm{Pr}(\mathcal{T}_r = k)\\
&\doteq \mathrm{SNR}^{-2(1-\frac{r}{l})} \sum_{k=1}^{\lfloor \frac{l}{2} \rfloor} \mathrm{Pr}(\mathcal{T}_r = k)\\
&\quad + \sum_{k=\lfloor \frac{l}{2} \rfloor+1}^{l-1} \mathrm{SNR}^{-(2-\frac{r}{l-k})}\mathrm{Pr}(\mathcal{T}_r = k) + \sum_{k=l}^{L} \mathrm{SNR}^{-(1-\frac{r}{l})}\mathrm{Pr}(\mathcal{T}_r = k)\\
&\doteq \mathrm{SNR}^{-d_{out}^{lt}(r,l)} && (5.22)
\end{aligned}$$

where $d_{out}^{lt}(r,l)$ is as given in Eq. (5.19). We use the following result:

$$\mathrm{Pr}(kI_{s;d} + (l-k)I_{s,r;d} < r\log(\mathrm{SNR})) \doteq \begin{cases} \mathrm{SNR}^{-2(1-r/l)} & \text{for } k \leq \lfloor l/2 \rfloor \\ \mathrm{SNR}^{-(2-r/(l-k))} & \text{for } \lfloor l/2 \rfloor < k \leq l-1 \end{cases} \quad (5.23)$$

*See Appendix 5.A for proofs of Eq.(5.19), Eq.(5.23).* □

### 5.4.3 Short-Term Static Channel

**Theorem 5.2.** For a short-term static channel the outage probability at the $l^{th}$ round for the proposed protocol is given by,

$$\mathrm{Pr}_{out}(l) \doteq \mathrm{SNR}^{-d_{out}^{st}(r,l)} \quad (5.24)$$

where,

$$d_{out}^{st}(r,l) \leq \begin{cases} (1-r) & \text{for } l = 1 \\ l\left(1-\frac{r}{l}\right) + (l-1)\left(1-\frac{r}{(l-1)}\right) & \text{for } l \neq 1 \end{cases} \quad (5.25)$$

**Proof Outline 2.** Unlike in the case of long-term static channel, the instantaneous mutual informations defined above vary from one block to the other. We have:

$$\begin{aligned}
\mathrm{Pr}(\mathcal{T}_r = k) &= \mathrm{Pr}\left(\sum_{i=1}^{k-1} I_{s;r}^{i} < r\log(\mathrm{SNR})\right) - \mathrm{Pr}\left(\sum_{i=1}^{k} I_{s;r}^{i} < r\log(\mathrm{SNR})\right)\\
&\doteq \mathrm{SNR}^{-(k-1)(1-r/(k-1))} - \mathrm{SNR}^{-k(1-r/k)}
\end{aligned}$$

And the outage probability for the ARQ relay long-term static channel is:

$$\begin{aligned}
\Pr_{out}(l) &= \sum_{k=1}^{L} \Pr_{\text{out}|\mathcal{T}_r=k}(l)\Pr(\mathcal{T}_r = k) \\
&\doteq \text{SNR}^{-d^{st}_{out}(r,l)} \\
&= \sum_{k=1}^{l-1} \Pr\left(\sum_{i=1}^{k} I^i_{s;d} + \sum_{i=k+1}^{l} I^i_{s,r;d} < r\log(\text{SNR})\right)\Pr(\mathcal{T}_r = k) \\
&\quad + \sum_{k=l}^{L} \Pr\left(\sum_{i=1}^{l} I^i_{s;d} < r\log(\text{SNR})\right)\Pr(\mathcal{T}_r = k) \\
&\dot{\geq} \sum_{k=l}^{L} \text{SNR}^{-l(1-r/l)}\Pr(\mathcal{T}_r = k)
\end{aligned} \tag{5.26}$$

where $d^{st}_{out}(r,l)$ is as given in Eq. (5.25). □

$d^{st}_{out}(r,l)$ corresponds to an upper-bound on the diversity multiplexing delay tradeoff. For a $2 \times 1$ MISO ARQ systems, the optimal diversity multiplexing delay tradeoff is $2l(1 - r/l)$. One can write:

$$l(1-r/l) \leq d^{st}_{out}(r,l) \leq l\left(1 - \frac{r}{l}\right)+(l-1)\left(1 - \frac{r}{(l-1)}\right) \leq 2l(1-r/l) \quad \text{for } l \neq 1 \tag{5.27}$$

$l(1-r/l)$ corresponds to the case when the source relay channel is physically degraded version of the source destination channel. In this case, the diversity multiplexing delay tradeoff is that of a SISO ARQ system.

### 5.4.4 Diversity Multiplexing Delay Tradeoff

**Theorem 5.3.** *The optimal diversity-multiplexing-delay tradeoff for the ARQ relay channel for the long-term static and short-term static relay channel is,*

$$d^{lt}(r_e, L) = d^{lt}_{out}(r_e, L) = d^{lt}_{out}(r, L)$$
$$d^{st}(r_e, L) = d^{st}_{out}(r_e, L) = d^{st}_{out}(r, L)$$

*subject to the constraint that $TL \geq 2$ for the long-term static channel and $T \geq 2$ for the short-term static channel, $0 \leq r_e < 1$.*

*See Appendix 5.B for proof.*

Note that the way we have defined the effective rate earlier Eq.(5.17) and from the expressions above for both the short-term and long-term static channel, it follows that:

$$r_e \doteq \frac{r}{\left(1+\sum_{l=1}^{L-1} \mathrm{SNR}^{-d_{out}(r,l)}\right)} \implies r_e \doteq r \tag{5.28}$$

The first phenomenon one can notice is that by increasing the value of the retransmission rounds, $L$, the diversity-multiplexing tradeoff curve for the long-term static channel flattens out as in Fig.5.6. Consider the tradeoff curve in Eq.(5.25) for the short-term static channel. Since the channel fades independently to a new realization in each round, transmission in each new round gives additional diversity which explains the multiplicative $L$ and $L-1$ factors in the diversity expression. Note that the factor is $(L-1)$ (both in the multiplication and the division) in the second term as the relay has to wait for at least one round before it can start transmitting to the destination. The reason this multiplicative factor does not show up in the case of the long-term quasi-static channel is that the channel is constant over all ARQ rounds and there is no time diversity benefit. But still there is a gain in the diversity because of the relay to destination channel and because of the ARQ protocol (the factor $r/l$). Moreover, as one can note from Appendix 5.B, the multiplexing gain is determined by the rate of the first block $r_e \doteq r$, which means that most packets are decoded correctly within the first round, and ARQ rounds are used to correct the remaining error events increasing the diversity order without loss in the transmission rate, thus the diversity order is determined by the rate of the code of the combined packets. Another fact established by the ARQ protocol for the half duplex relay channel, is that the time sharing factor (the fraction of time the relay spends in receive mode or transmit mode) chosen such that it minimizes the outage probability is accounted for automatically by incremental redundancy with ACK/NACK bit feedback and by adapting to the channel conditions.

## 5.5 Power Control

We notice that $d_{out}^{lt}(0,l) = 2$ for all $l \neq 1$. Thus the long-term static channel is limiting the performance at low multiplexing-gains, which motivates the use of the power control.

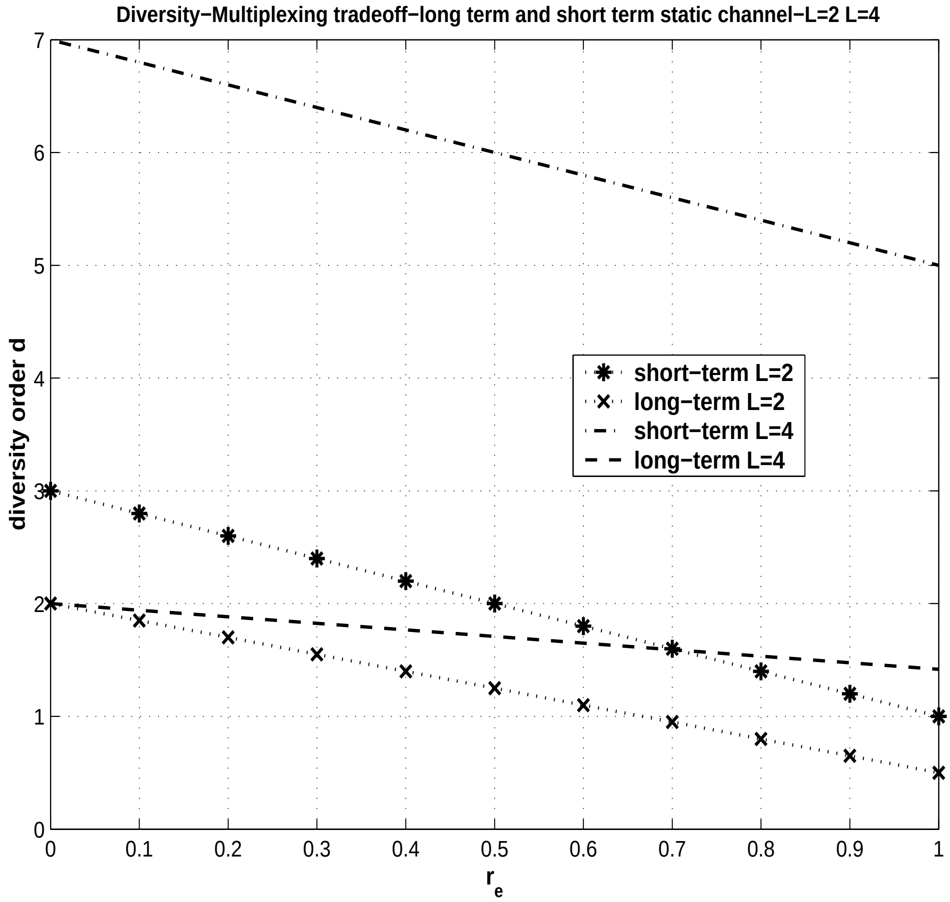


**Figure 5.6:** The diversity-multiplexing tradeoff for different values of the maximum number of ARQ rounds for the short-term (upper-bound) and long term static channel.

Power control was recently applied to the cooperative relay channels. In [100], it was shown that by exploiting the channel state information at the transmitter and an adapted power control algorithm, the outage can be substantially lowered leading to an increase in diversity. The power control algorithm is based on the elegant technique presented in [101]. When channel state information is available at both the source and the relay, power control is an alternative to some coding and transmission techniques. For example, phase correction of the source and relay signals such that these signals coherently combine at the destination, or use of correlation between source and relay signals to maximize the rate at each channel state. Note that all these techniques occur at constant power transmission. In [102], the authors demonstrated that if the entire network state is used to determine the instantaneous transmitter power, only one bit feedback suffices to double the diversity order of the amplify and forward (AF) cooperative channel. Inspired by [97], and noticing that in long-term static channels the ARQ diversity is limited at low multiplexing gains, we construct a power control algorithm for this ARQ single relay channel. For simplicity, we consider a power control in which the relay is restricted to use a constant power in each round, but the source has the ability to vary its power to meet a long-term average power constraint.

Let $P_l = \mathrm{SNR}^{p_l}$ be the power allocated per channel use for the $l^{th}$ round. The power constraint for the long-term static channel is $\frac{\sum_{l=1}^{L} P_l \mathrm{Pr}_{out}(l-1)}{\sum_{l=0}^{L-1} \mathrm{Pr}_{out}(l)} \leq 1$ (where the denominator is the expected number of rounds needed for successful decoding at the receiver). It is straightforward to show that $P_l \leq \frac{L}{\mathrm{Pr}_{out}(l-1)}$ and $p_l \leq d_{out}(r, l-1)$ where $d_{out}(r, l-1)$ is the SNR exponent of the $l-1$-th round outage probability for the ARQ relay channel. The power control policy is optimal when $P_l = \mathrm{SNR}^{d(r,l-1)}$, with $P_1 = 0$ (for ease of notation the index *out* is omitted). Then, Eq.(5.14), Eq.(5.15), Eq.(5.16) become (for long-term static channels): $I_{1,pc}^{j} = \log(1 + \mathrm{SNR}^{1-\mu_{sd}+d(r,j-1)})$, $I_{2,pc}^{j} \doteq \log(\mathrm{SNR}^{(1-\min(\mu_{sd}-d(r,j-1),\mu_{rd}))^+})$ and $I_{3,pc}^{j} = I(\mathbf{x}_{sr,j}; \mathbf{y}_r^d | h_{sr,j}) = \log(1 + \mathrm{SNR}^{1-\mu_{sr}+d(r,j-1)})$. We define for convenience,

$$
\begin{aligned}
q_k &= \Pr(\sum_{i=1}^{k} I_{1,pc}^{i} + \sum_{i=k+1}^{l} I_{2,pc}^{i} < r\log(\mathrm{SNR})) \qquad (5.29)\\
&\doteq \Pr\Big(\sum_{i=1}^{k}(1-\mu_{sd}+d(r,i-1))^{+} + \\
&\quad \sum_{i=k+1}^{l}(1-\min(\mu_{sd}-d(r,i-1);\mu_{rd}))^{+} < r\Big)
\end{aligned}
$$

The outage probability for the ARQ relay long-term static channel with power control is:

$$\begin{aligned}
\Pr_{out}(l) &= \sum_{k=1}^{l-1} q_k \Pr(\mathcal{T}_r = k) + \sum_{k=l}^{L} \Pr\left(\sum_{i=1}^{l} I_{1,pc}^{i} < r\log(\mathrm{SNR})\right)\Pr(\mathcal{T}_r = k) \\
&\doteq \sum_{k=1}^{l-1} q_k \Pr(\mathcal{T}_r = k) \\
&\quad + \sum_{k=l}^{L} \Pr\Bigg(\sum_{i=1}^{l}(1-\mu_{sd}+d(r,i-1))^{+} < r\Bigg)\Pr(\mathcal{T}_r = k) \\
&\dot{\geq} \sum_{k=l}^{L} \Pr\Bigg(\sum_{i=1}^{l}(1-\mu_{sd}+d(r,i-1))^{+} < r\Bigg)\Pr(\mathcal{T}_r = k) \qquad (5.30)
\end{aligned}$$

Now note that,

$$\begin{aligned}
&\Pr\Bigg(\sum_{i=1}^{l}(1-\mu_{sd}+d(r,i-1))^{+} < r\Bigg) \\
&= \Pr\Bigg(\Bigg[\max_{t=1,\dots,l}\sum_{i=1}^{t} d(r,l-i)+t(1-\mu_{sd})\Bigg]^{+} < r\Bigg) \\
&\geq \Pr\left(\left(\sum_{i=1}^{l-1} d(r,i)+l\right)\left[1-\frac{\mu_{sd}}{d(r,l-1)+1}\right]^{+} < r\right) \\
&\doteq \mathrm{SNR}^{-d(r,l)} \qquad (5.31)
\end{aligned}$$

where one can easily show that $\left(\sum_{i=1}^{l-1} d(r,i)+l\right)\left[1-\frac{\mu_{sd}}{d(r,l-1)+1}\right]^{+}$ is, for all $\mu_{sd}$, strictly above $\left[\max_{t=1,\dots,l}\sum_{i=1}^{t} d(r,l-i)+t(1-\mu_{sd})\right]^{+}$, leading to:

$$d(r,l) = \left(1-\frac{r}{\sum_{i=1}^{l-1} d(r,i)+l}\right)(1+d(r,l-1)) \qquad (5.32)$$

From Eq.(5.32) we note that $d(r,l)$ does not depend on $k$. Also $\Pr(\mathcal{T}_r = k)$ decreases as $k$ increases. Combining these two facts and from Eq.(5.30) we see that by using the fact that $r_e \doteq r$,

$$\begin{aligned}
\mathrm{d}_{out,pc}(r_e,l) &\dot{\leq} \left(1-\frac{r_e}{\sum_{i=1}^{l-1} d_{out,pc}(r_e,i)+l}\right)(1+d_{out,pc}(r_e,l-1)) \\
&\quad + \left(1-\frac{r_e}{l-1}\right) \qquad (5.33)
\end{aligned}$$

In the following, the diversity-multiplexing delay tradeoff is computed through Monte-Carlo simulations. The diversity gain in a particular round

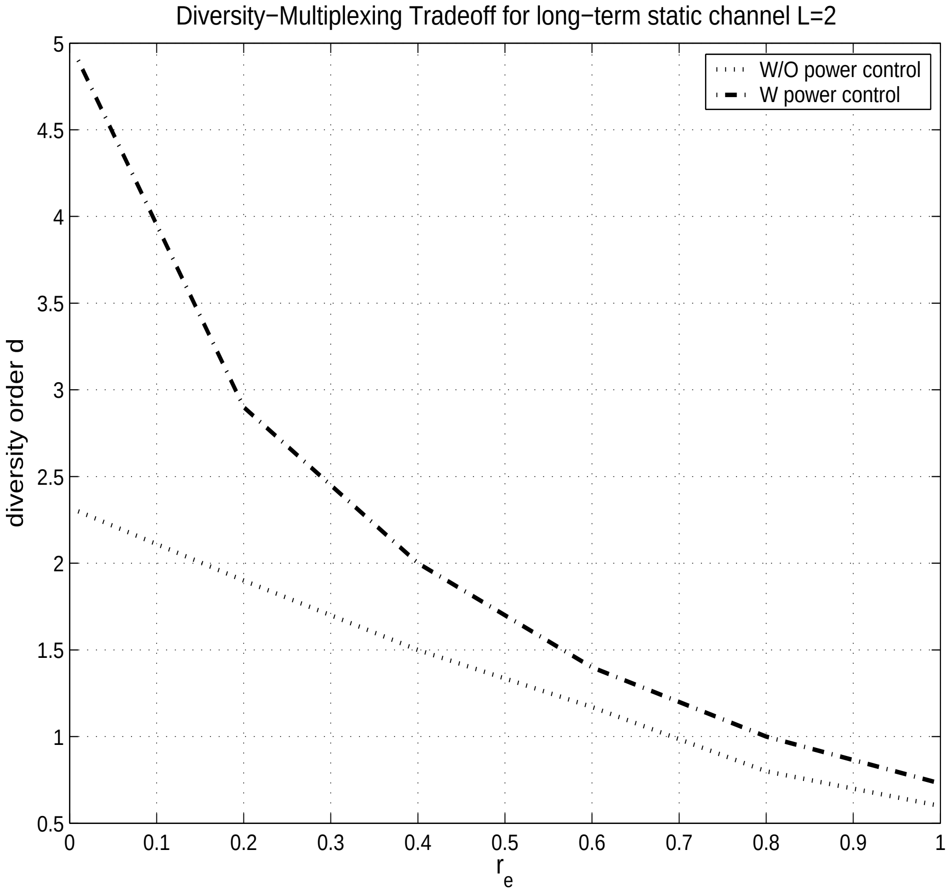


**Figure 5.7:** The diversity-multiplexing tradeoff for $L = 2$ for the long-term static channel with and without power control

and hence the power allocated can be numerically computed in a recursive manner. The diversity gain obtained using power control is significant compared to the constant power case especially at low multiplexing gains as shown in Fig.5.7. Moreover, one can notice that the proposed power control is deterministic in the sense that it does not depend on the knowledge of the channel state but requires only the knowledge of the outage probabilities which can be estimated.

## 5.6 Conclusion

From Fig.5.6, Fig.5.7, it can be seen that a significant gain in diversity is obtained by the proposed protocol. This is also evident from the outage probability expressions for short-term and long-term static channels. Power control is seen to be beneficial at all multiplexing gains by increasing the diversity order. An extension of this work would be to consider the impact of multiple antennas (in particular two antennas considering the practical implications) at the receiver (base station), where the source and the relay collaborate to reach the destination. Another avenue would be to investigate the extension of these schemes to the case of multiple relays relaying the information for a single source destination pair. This protocol can then also be applied to ad-hoc TDMA wireless networks where in each slot all the remaining nodes in the network act as relays for a particular source destination pair.

## Appendix 5.A Proof of the Tradeoff Curve for Long-Term Static Channel

The proof of Eq.(5.23) is based on Fig.5.8, Fig.5.9, Fig.5.10.

$$
\begin{aligned}
&\Pr(kI_{s;d} + (l-k)I_{s,r;d} < r\log(\mathrm{SNR})) \\
&\quad = \Pr(k\log(1+\mathrm{SNR}\gamma_{sd}) \qquad (5.34)\\
&\qquad +(l-k)\log(1+\mathrm{SNR}^{(1-\mu_{sd})}+\mathrm{SNR}^{(1-\mu_{rd})}) < r\log(\mathrm{SNR})) \\
&\quad \doteq \Pr(k\log(\mathrm{SNR}^{(1-\mu_{sd})^+}) \\
&\qquad +(l-k)\log(\mathrm{SNR}^{(1-\min(\mu_{sd},\mu_{rd}))^+}) < r\log(\mathrm{SNR})) \\
&\quad \doteq \Pr(k(1-\mu_{sd})^+ + (l-k)(1-\min(\mu_{sd},\mu_{rd}))^+ < r) \qquad (5.35)\\
&\quad \overset{(a)}{\doteq} \mathrm{SNR}^{-\inf_{\underline{\mu}\in\mathcal{A}^+}(\mu_{sd}+\mu_{rd})}
\end{aligned}
$$

where (a) is from the results in Section (5.2.4) and $\mathcal{A}^+ = \{\underline{\mu} \in \mathcal{R}^{2+} : k(1-\mu_{sd})^+ + (l-k)(1-\min(\mu_{sd},\mu_{rd}))^+ < r)\}$. To solve Eq.(5.34), for the case where $\min(\mu_{sd},\mu_{rd}) = \mu_{sd}$, we obtain that

$$
\inf_{\underline{\mu}\in\mathcal{A}^+} \mu_{sd}+\mu_{rd} = 1-r/l+1-r/l = 2(1-r/l)
$$

as depicted in Fig.5.8. For the case where $\min(\mu_{sd},\mu_{rd}) = \mu_{rd}$, the solution to Eq.(5.34) depends on $k$ (the slot where the relay decodes). For $k \le \lfloor l/2 \rfloor$, $\inf_{\underline{\mu}\in\mathcal{A}^+} \mu_{sd}+\mu_{rd}$ corresponds to $\mu_{sd} = \mu_{rd} = 1-r/l$ as depicted in Fig.5.9 and for $\lfloor l/2 \rfloor < k \le l-1$, we obtain $\mu_{sd} = 1$ and $\mu_{rd} = 1-r/(l-k)$ as depicted in Fig.5.10 which concludes the proof of Eq.(5.23). To derive $d_{out}^{lt}$ we proceed by finding the dominant term at high SNR.

$$
\begin{aligned}
\Pr_{out}(l) &= \sum_{k=1}^{L} \Pr_{out|\mathcal{T}_r=k}(l)\Pr(\mathcal{T}_r=k) \\
&\doteq \mathrm{SNR}^{-2(1-r/l)}\sum_{k=1}^{\lfloor l/2\rfloor}\Pr(\mathcal{T}_r=k) + \sum_{k=\lfloor l/2\rfloor+1}^{l-1}\mathrm{SNR}^{-(2-r/(l-k))}\Pr(\mathcal{T}_r=k) \\
&\quad + \sum_{k=l}^{L}\mathrm{SNR}^{-(1-r/l)}\Pr(\mathcal{T}_r=k) \\
&\doteq \mathrm{SNR}^{-2(1-r/l)} + \underbrace{\sum_{k=\lfloor l/2\rfloor+1}^{l-1}\mathrm{SNR}^{-(2-r/(l-k))}\Pr(\mathcal{T}_r=k)}_{\zeta} \\
&\quad + \underbrace{\mathrm{SNR}^{-(1-r/l)-(1-r/(l-1))}}_{\xi} \qquad (5.36)
\end{aligned}
$$

$$\sum_{k=\lfloor l/2\rfloor+1}^{l-1} \mathrm{SNR}^{-(2-\frac{r}{l-k})}\Pr(\mathcal{T}_r = k) \doteq \begin{cases} 0 & \text{for } l \in \{1,2\} \\ \mathrm{SNR}^{-(3-2r)} & \text{for } l = 3 \\ \mathrm{SNR}^{-(2-r)-(1-\frac{r}{l-2})} & \text{for } l > 3 \end{cases} \quad (5.37)$$

*Proof.* To show Eq.(5.37), let us compare the SNR exponent of the summand of $\zeta$ at $k$ and $k+1$. In other words we need to show that the SNR exponent is decreasing with $k$, and the dominant term of $\zeta$ is the last summand ($k = l-1$). We want to show that $(2-\frac{r}{l-k}+1-\frac{r}{k-1})-(2-\frac{r}{l-k-1}+1-\frac{r}{k}) > 0$. This is equivalent to:

$$\begin{aligned} (2-\frac{r}{l-k}+1-\frac{r}{k-1})-(2-\frac{r}{l-k-1}+1-\frac{r}{k}) &> 0 \\ \frac{1}{(l-k-1)(l-k)}-\frac{1}{k(k-1)} &> 0 \\ l(l-1) &< 2k(l-1) \\ l &< 2k \qquad (5.38) \end{aligned}$$

where we used the fact that $r$ is positive and that $l \neq 1$. Eq.(5.38) is always true since $\lfloor l/2\rfloor + 1 \leq k \leq l-1$, which concludes the proof of Eq.(5.37).

The dominant term of the outage probability is the last term of Eq.(5.36) $\xi$ except for $l = 3$ where it depends on the value of $r$. Indeed for $l = 3$ and $r < 6/7$ the dominant term of the outage probability is the last term of Eq.(5.36) $\xi$ where for $l = 3$, $r \geq 6/7$ the dominant term is the second one in Eq.(5.36) $\zeta$ as given in Eq.(5.37).

*Proof.* Again we need to compare $\xi = 1 - r/l + 1 - r/(l-1)$ and $\zeta = 2 - r + 1 - r/(l-2)$ for $l \neq 1$ and $l \neq 2$.

$$\begin{aligned} 3-r-r/(l-2)-(2-r/l-r/(l-1)) &> 0 \\ \Leftrightarrow r &< \frac{1}{1+\frac{1}{l-2}-\frac{1}{l}-\frac{1}{l-1}} \qquad (5.39) \end{aligned}$$

It is easy to check that $\frac{1}{1+\frac{1}{l-2}-\frac{1}{l}-\frac{1}{l-1}} > 1$ except for $l = 3$ and that $\xi$ is the dominant term except for $l = 3$. For $l = 3$, by replacing $l$ by 3 in Eq.(5.39) we obtain the desired result.

## Appendix 5.B Proof of the Diversity Multiplexing Delay Tradeoff

In the following, all the computations are done for the long-term static channel, the results are easily extended to the short-term static channel. For simplicity of notations in the proofs, we will use the following channel model:

$$\underline{\mathbf{y}}_l^d = \underline{\mathbf{h}}_l \underline{\mathbf{x}} + \underline{\mathbf{n}}_l^d \tag{5.40}$$

where $\underline{\mathbf{y}}_l^d \in \mathcal{C}^{Tl}$ represents the signal received over all transmitted block from 1 to $l$, $\underline{\mathbf{n}}_l^d$ the noise over the $l$ rounds, $\underline{\mathbf{x}} = (\mathbf{x}_1^T, \ldots, \mathbf{x}_L^T)^T$ where $\mathbf{x}_l^T$ is defined in Eq.(5.10) and $\underline{\mathbf{h}}_L = \sqrt{\text{SNR}}(\mathbf{h}_1, \ldots, \mathbf{h}_L)$, $\mathbf{h}_l$ as defined in Eq.(5.11). $\underline{\mathbf{h}}_l$ is obtained from $\underline{\mathbf{h}}_L$ by replacing the last $2T(L-l)$ rows by zero, which corresponds to the fact that the blocks $(\mathbf{x}_{l+1}, \ldots, \mathbf{x}_L)$ have not been transmitted yet and they appear multiplied by a zero channel matrix.

We wish first to show that $d_{out}$ as defined in Eq.(5.19) is an upper bound to the SNR exponent of the ARQ relay system. Consider a system with codebook $\mathcal{C}(\text{SNR})$, first-block rate $r\log(\text{SNR})$ and some decoding rule $\phi = (\phi_1, \ldots, \phi_L)$. We call $k$ the time slot at which the relay decodes and $\overline{\mathcal{E}_k}$ the event that the relay decodes correctly, $\mathcal{E}_k$ the event for the incorrect decoding. Then, the probability of error conditioned on the vector channel: $\underline{\mathbf{h}} = [\underline{h}_{sd}, \underline{h}_{rd}, \underline{h}_{sr}]$, that the relay decodes at slot $k$, and on the particular decoder and codebook is:

$$\begin{aligned} P_{e,k}(\text{SNR}|\underline{\mathbf{h}}, \mathcal{C}(\text{SNR}), \phi) &= \Pr(e, \mathcal{E}_k) + \Pr(e, \overline{\mathcal{E}_k}) \\ &\geq \Pr(e|\overline{\mathcal{E}}_k)\Pr(\overline{\mathcal{E}}_k) \\ &\geq \Pr(e|\overline{\mathcal{E}}_k) \end{aligned} \tag{5.41}$$

From the channel coding theorem, one can show that $\Pr(\overline{\mathcal{E}}_k) \geq 1 - \epsilon$.

Let $E_l$ be the event that the decoding outcome at the destination is not correct with $l$ received rounds and $\mathcal{O}_l$ the event that the destination sends a NACK at round $l$ and $\overline{\mathcal{O}}_l$ being the event for sending an ACK. Based on the results of [97], one can define the probability of error as (knowing that $w$ is the transmitted message, $\hat{w} = \phi_l(\underline{\mathbf{y}}_l)$ and $\phi_l(\underline{\mathbf{y}}_l) = 0$ means that an error is detected and a NACK is sent back to the transmitter):

$$
\begin{aligned}
&\Pr_{e,k}(\text{SNR}|\underline{\mathbf{h}}, \mathcal{C}(\text{SNR}), \phi) && (5.42)\\
\geq\ & \sum_{l=1}^{L-1} \Pr(E_l, \mathcal{O}_1, \ldots, \mathcal{O}_{l-1}, \overline{\mathcal{O}}_l | \overline{\mathcal{E}}_k) + \Pr(E_L, \mathcal{O}_1, \ldots, \mathcal{O}_{L-1} | \overline{\mathcal{E}}_k) && (5.43)\\
=\ & \sum_{l=1}^{L-1} \Pr(\{\phi_l(\underline{\mathbf{y}}_1) = 0\}, \ldots, \{\phi_l(\underline{\mathbf{y}}_{l-1}) = 0\}, \bigcup_{\substack{\hat{w}\neq w\\ \hat{w}>0}} \{\phi_l(\underline{\mathbf{y}}_l) = \hat{w}\} | \underline{\mathbf{h}}, \overline{\mathcal{E}}_k)\\
& + \Pr(\{\phi_l(\underline{\mathbf{y}}_1) = 0\}, \ldots, \{\phi_l(\underline{\mathbf{y}}_{L-1}) = 0\}, \bigcup_{\hat{w}\neq w} \{\phi_l(\underline{\mathbf{y}}_L) = \hat{w}\} | \underline{\mathbf{h}}, \overline{\mathcal{E}}_k)
\end{aligned}
$$

Clearly $\Pr(E_l, \mathcal{O}_1, \ldots, \mathcal{O}_{l-1}, \overline{\mathcal{O}}_l)$ is the probability of undetected error. Indeed if an event $\mathcal{E}$ is detected a NACK is sent and a retransmission occurs, then this event does not count in the error event. $\Pr(E_L, \mathcal{O}_1, \ldots, \mathcal{O}_{L-1})$ is the probability of decoding error with $L$ (maximum number) received blocks (at round $L$, the error is due to the undetected error or the failure of decoding which explains the $\bigcup_{\hat{w}}$ in Eq.(5.42)).

The probability of error (computed by considering a decoder working on each round) is lower-bounded by the probability of error of the optimal maximum likelihood (ML) decoder $\phi_{ml}$ that operates on the whole received signal vector $\underline{\mathbf{y}} = \underline{\mathbf{y}}_L$.

By using the Fano's inequality ($T$ is the number of channel uses per round):

$$
\begin{aligned}
\Pr_{e,k}(\text{SNR}|\underline{\mathbf{h}}, \mathcal{C}(\text{SNR}), \phi) &\geq \Pr_{e,k}(\text{SNR}|\underline{\mathbf{h}}, \mathcal{C}(\text{SNR}), \phi_{ml})\\
&\geq 1 - \frac{I(\underline{\mathbf{x}}; \underline{\mathbf{y}}|\underline{\mathbf{h}}, k)}{Tr \log \text{SNR}} - \frac{1}{Tr \log \text{SNR}}
\end{aligned}
$$

This leads to (based on the results in [95]):

$$
P_{e,k} \dot{\geq} \Pr(I(\underline{\mathbf{x}}; \underline{\mathbf{y}}|\underline{\mathbf{h}}, k) \leq Tr \log \text{SNR}) \tag{5.44}
$$

In our case, we obtain $P_e = E_k[P_{e,k}] \dot{\geq} SNR^{-d_{out}(r,L)}$. This is because the mutual information we use in the Fano's inequality is exactly the mutual information knowing that the relay cooperate at a particular round $k$ and then we average on $k$ which corresponds exactly to the outage probability computed in Eq.(5.21). Since $r_e \leq r$ and $d^{lt}$ is a decreasing function in $r$, we have $d^{lt}(r_e, L) \leq d_{out}(r, L) \leq d_{out}(r_e, L)$.

The achievability of the exponent upper-bound is shown based on a bounded distance decoder [97]. Let $\mathcal{C}(\text{SNR})$ denote a random code generated with i.i.d $\sim \mathcal{N}_c(0, 1)$ components, block length $LT$ and rate $r \log SNR$.

Remember that in our case the source and the relay cooperate fully (once the latter has decoded) and we assume that both nodes are using the same codebook. The probability of error is given by (recalling that $\overline{\mathcal{E}}_k$ the event that the relay decodes correctly):

$$\begin{aligned} P_{e,k}(\mathrm{SNR}) &= \Pr(e, \mathcal{E}_k) + \Pr(e, \overline{\mathcal{E}_k}) \\ &\leq \Pr(e|\overline{\mathcal{E}}_k)\Pr(\overline{\mathcal{E}}_k) \end{aligned} \tag{5.45}$$

where we used the fact that $\Pr(\mathcal{E}_k) \leq \epsilon$ (based on results in [95]), $\Pr(\overline{\mathcal{E}}_k) \leq 1 - \epsilon$ and $\Pr(e|\mathcal{E}_k) \leq 1$.

We define the following bounded distance decoder $\phi$ at each round $l \leq L - 1$ and the signal model at round $l$ is given by Eq.(5.40), and $\underline{\mathbf{x}}(w)$ takes into account the signature of the source and the relay, $w$ is the word message transmitted:

- $\phi_l(\underline{\mathbf{y}}_l) = \hat{w}$ if the channel is not in outage and the codeword $\underline{\hat{\mathbf{x}}}$ corresponding to $\hat{w}$ is the unique codeword in $\mathcal{C}(\mathrm{SNR})$ such that $|\underline{\mathbf{y}}_l - \underline{\mathbf{h}\hat{\mathbf{x}}}| \leq Tl(1+\delta)$
- $\phi_l(\underline{\mathbf{y}}_l) = 0$ in any other case.
- At round $L$, the decoder outputs the index of the minimum distance codeword, i.e, $\phi_L(\underline{\mathbf{y}}_L) = \phi_{ml}(\underline{\mathbf{y}}_L)$

Let us first bound the probability of undetected error $\Pr(E_l, \mathcal{O}_1, \ldots, \mathcal{O}_{l-1}, \overline{\mathcal{O}}_l|\overline{\mathcal{E}}_k) \leq \Pr_k(E_l, \overline{\mathcal{O}}_l)$. An error is undetected if the unique codeword $\underline{\hat{\mathbf{x}}}$ such that $|\underline{\mathbf{y}}_l - \underline{\mathbf{h}\hat{\mathbf{x}}}| \leq Tl(1+\delta)$ does not correspond to the transmitted message $w$. This means that if we draw a sphere centered around the true codeword corresponding to the message transmitted $w$ of radius $Tl(1+\delta)$, an undetected error occurs if the received signal $\underline{\mathbf{y}}_l$ belongs to the other sphere centered around other codewords. This event is included in the event that the received signal belongs to the region corresponding to the complement of the sphere corresponding to the true message transmitted; which is the event that the magnitude of the noise is bigger than the radius of the sphere.

$$\begin{aligned} \Pr_k(E_l, \overline{\mathcal{O}}_l) &\leq \Pr(|\underline{\mathbf{n}}_l|^2 \geq Tl(1+\delta)) \\ &\leq (1+\delta)^{Tl(1+\delta)} \exp(-Tl\delta) \end{aligned} \tag{5.46}$$

Using the Chernoff bound and for some $\beta > 0$, and by letting $\delta = \beta \log \mathrm{SNR}$. This leads to $\Pr_k(E_l, \overline{\mathcal{O}}_l) \dot{\leq} \mathrm{SNR}^{-Tl\beta}$.

Assuming ML decoder, the probability of error at round $L$ is $\Pr(E_L|\overline{\mathcal{E}}_k) = \Pr_k(E_L)$, and using the results in [95], one can show that $\Pr(E_L)\doteq \text{SNR}^{-d^{lt}_{out}(r,L)}$ for $LT \geq 2$ (corresponding to the case where the relay decodes in the first round). Using the following

$$\Pr_{e,k}(\text{SNR}) \leq \sum_{l=1}^{L-1} \Pr(E_l, \overline{\mathcal{O}}_l|\overline{\mathcal{E}}_k) + \Pr(E_L|\overline{\mathcal{E}}_k)$$

we obtain:

$$\begin{aligned} P_e(\text{SNR}) &\dot{\leq} E_k[\sum_{l=1}^{L-1} \Pr_k(E_l, \overline{\mathcal{O}}_l)] + \text{SNR}^{-d^{lt}_{out}(r,L)} \\ &\dot{\leq} \text{SNR}^{-T\beta} + \text{SNR}^{-d^{lt}_{out}(r,L)} \end{aligned}$$

By choosing $LT \geq 2$ and a large $\beta$, one can ensure that $T\beta \geq d^{lt}_{out}(r,L)$ leading to $P_e \dot{\leq} \text{SNR}^{-d^{lt}_{out}(r,L)}$.

The next step is to prove that the effective multiplexing rate $r_e \doteq r$. Note that the effective multiplexing rate is given by:

$$r_e \doteq \frac{r}{1 + \sum_{l=1}^{L-1} \Pr(\mathcal{O}_l)}$$

The condition $r_e \doteq r$ translates to the fact that $\Pr(\mathcal{O}_l)$ are $o(1)$. Let us look at the region formed at the channel output by all possible received vectors $\underline{\mathbf{y}}_l$ and channel matrices $\underline{\mathbf{h}}_l$. We define $\mathcal{A}(\text{SNR}, l)$ as the outage space, $\mathcal{R}_\emptyset$ as the region of channel outputs not included in any sphere of radius $Tl(1+\delta)$ and centered around the codewords, and finally $\mathcal{R}$ is the region of channel outputs included in more than one of such spheres. $\mathcal{R}$ is partitioned into $\mathcal{R}_w$ (the region centered around the true codeword $\underline{\mathbf{x}}$ corresponding to the transmitted message $w$), and $\mathcal{R}_{\overline{w}}$.

$$\begin{aligned} \Pr(\mathcal{O}_l) &= \Pr(\mathcal{A}(\text{SNR}, l) \cup \mathcal{R}_\emptyset \cup \mathcal{R}) & (5.47) \\ &\leq \Pr(\mathcal{A}(\text{SNR}, l)) + \Pr(\overline{\mathcal{A}}(\text{SNR}, l) \cap (\mathcal{R}_\emptyset \cup \mathcal{R}_{\overline{w}})) \\ &+ \Pr(\overline{\mathcal{A}}(\text{SNR}, l) \cap \mathcal{R}_w) & (5.48) \end{aligned}$$

The event $\overline{\mathcal{A}}(\text{SNR}, l) \cap (\mathcal{R}_\emptyset \cup \mathcal{R}_{\overline{w}})$ corresponds to the union of the region of channel outputs not included in any sphere and the intersection of all spheres excluding the one corresponding to the transmitted codeword. One can show that:

$$\Pr(\overline{\mathcal{A}}(\text{SNR}, l) \cap (\mathcal{R}_\emptyset \cup \mathcal{R}_{\overline{w}})) \leq \Pr(|\underline{\mathbf{n}}_l|^2 \geq Tl(1+\delta))$$

We then have:

$$\Pr(\mathcal{O}_l)\dot{\leq}\text{SNR}^{-d_{out}^{lt}(r,l)} + \text{SNR}^{-T\beta} + \Pr(\overline{\mathcal{A}}(\text{SNR}, l) \cap \mathcal{R}_w) \tag{5.49}$$

Noting that $\mathcal{R}_{\overline{w}}$ is the region of the sphere centered around the true codeword and included in other spheres:

$$\mathcal{R}_{\overline{w}} : \bigcup_{\substack{\hat{w}\neq w\\ \hat{w}>0}} \{|\underline{\mathbf{y}}_l - \underline{\mathbf{h}}_l\underline{\mathbf{x}}| \leq Tl(1+\delta), |\underline{\mathbf{y}}_l - \underline{\mathbf{h}}_l\underline{\hat{\mathbf{x}}}| \leq Tl(1+\delta)\} \tag{5.50}$$

This leads to:

$$\begin{aligned}&\Pr(\overline{\mathcal{A}}(\text{SNR}, l) \cap \mathcal{R}_w) \\ &= \Pr(\overline{\mathcal{A}}(\text{SNR}, l), \bigcup_{\substack{\hat{w}\neq w\\ \hat{w}>0}} |\underline{\mathbf{y}}_l - \underline{\mathbf{h}}_l\underline{\hat{\mathbf{x}}}| \leq Tl(1+\delta), |\underline{\mathbf{n}}_l|^2 \leq Tl(1+\delta))\end{aligned} \tag{5.51}$$

Let $a = \underline{\mathbf{h}}_l(\underline{\mathbf{x}} - \underline{\hat{\mathbf{x}}})$, $b = \underline{\mathbf{n}}_l$ and $\Delta = Tl(1+\delta)$, we have:

$$\begin{aligned}\{|a+b|^2 \leq \Delta, |b|^2 \leq \Delta\} &= \{|a+b|^2 \leq \Delta, |b|^2 \leq \Delta, |a|^2 \leq 4\Delta\} \\ &\quad \cup\{|a+b|^2 \leq \Delta, |b|^2 \leq \Delta, |a|^2 > 4\Delta\} \\ &= \{|a+b|^2 \leq \Delta, |b|^2 \leq \Delta, |a|^2 \leq 4\Delta\} \\ &\subseteq \{|a|^2 \leq 4\Delta\}\end{aligned} \tag{5.52}$$

since the event $\{|a+b|^2 \leq \Delta, |b|^2 \leq \Delta, |a|^2 > 4\Delta\}$ is empty.

We obtain:

$$\begin{aligned}&\Pr(\bigcup_{\substack{\hat{w}\neq w\\ \hat{w}>0}} |\underline{\mathbf{y}}_l - \underline{\mathbf{h}}_l\underline{\hat{\mathbf{x}}}| \leq Tl(1+\delta), |\underline{\mathbf{n}}_l|^2 \leq Tl(1+\delta)) \\ &\leq \sum_{\substack{\hat{w}\neq w\\ \hat{w}>0}} \Pr(|\frac{\underline{\mathbf{h}}_l(\underline{\mathbf{x}} - \underline{\hat{\mathbf{x}}})}{2}|^2 \leq Tl(1+\delta))\end{aligned} \tag{5.53}$$

This leads to an upper-bound on the pairwise error probability summed over all distinct messages pairs and conditioned with respect to the channel. Results from [95] yield that for $Tl \geq 2$:

$$\Pr(\overline{\mathcal{A}}(\text{SNR}, l) \cap \mathcal{R}_w)\dot{\leq}\text{SNR}^{-d_{out}^{lt}(r,l)} \tag{5.54}$$

where $d_{out}^{lt}(r,l)$ is the maximum possible SNR exponent for codes with length $lT$ and rate $r/l$. Finally, one can state:

$$\Pr(\mathcal{O}_l)\dot{\leq}\text{SNR}^{-d_{out}^{lt}(r,l)}, \quad \text{for } 0 < l < L \tag{5.55}$$

and recalling the definition of the effective multiplexing rate, we obtain that $r_e \doteq r$.

Using random coding arguments, one can find codebooks which are good depending on the instant $\mathcal{T}_r$ at which the relay decodes. But we need codebooks which are simultaneously optimal irrespective of the instant $\mathcal{T}_r$ when the relay decodes. In the same spirit, we have to show that not only all the exponents (Eq.(5.55)) can be achieved by averaging over the code ensemble, but that there exist codes that achieve them simultaneously. The existence of codes which simultaneously achieve the error probability exponents follows from expurgation (along the lines of lemma 11 in [97]).

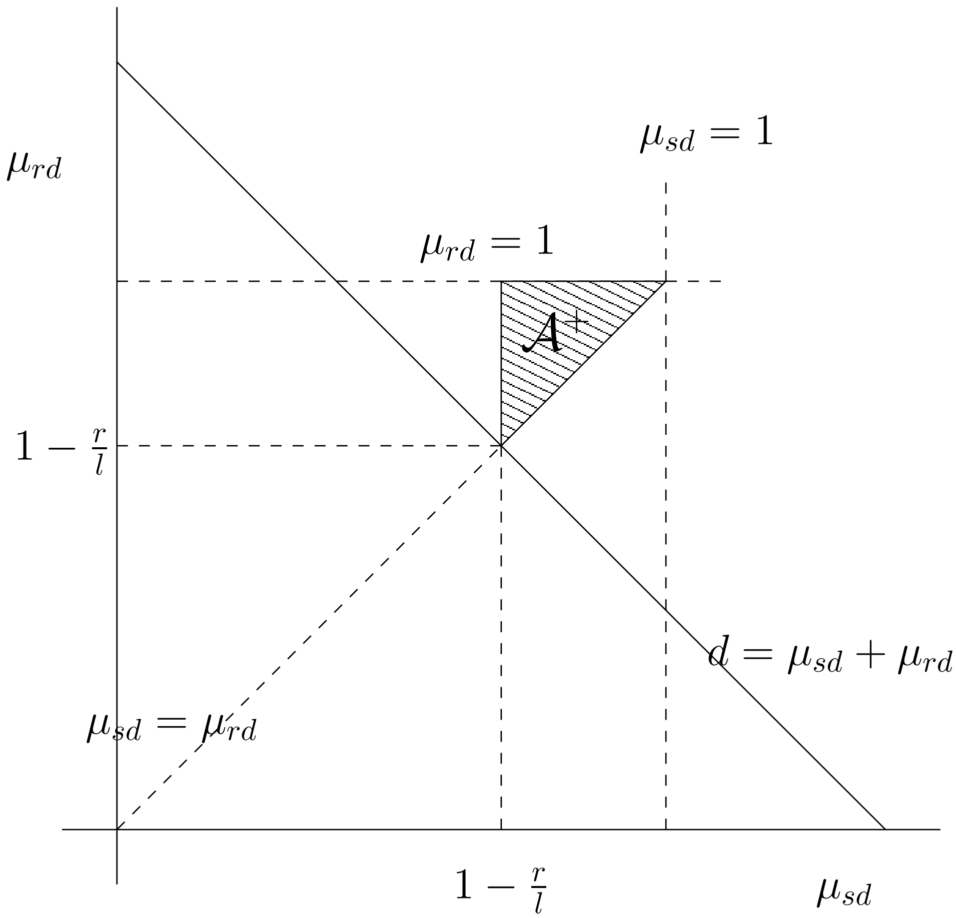


**Figure 5.8:** The region $\mathcal{A}^+$ for the long-term static channel where $\min(\mu_{sd}, \mu_{rd}) = \mu_{sd}$.

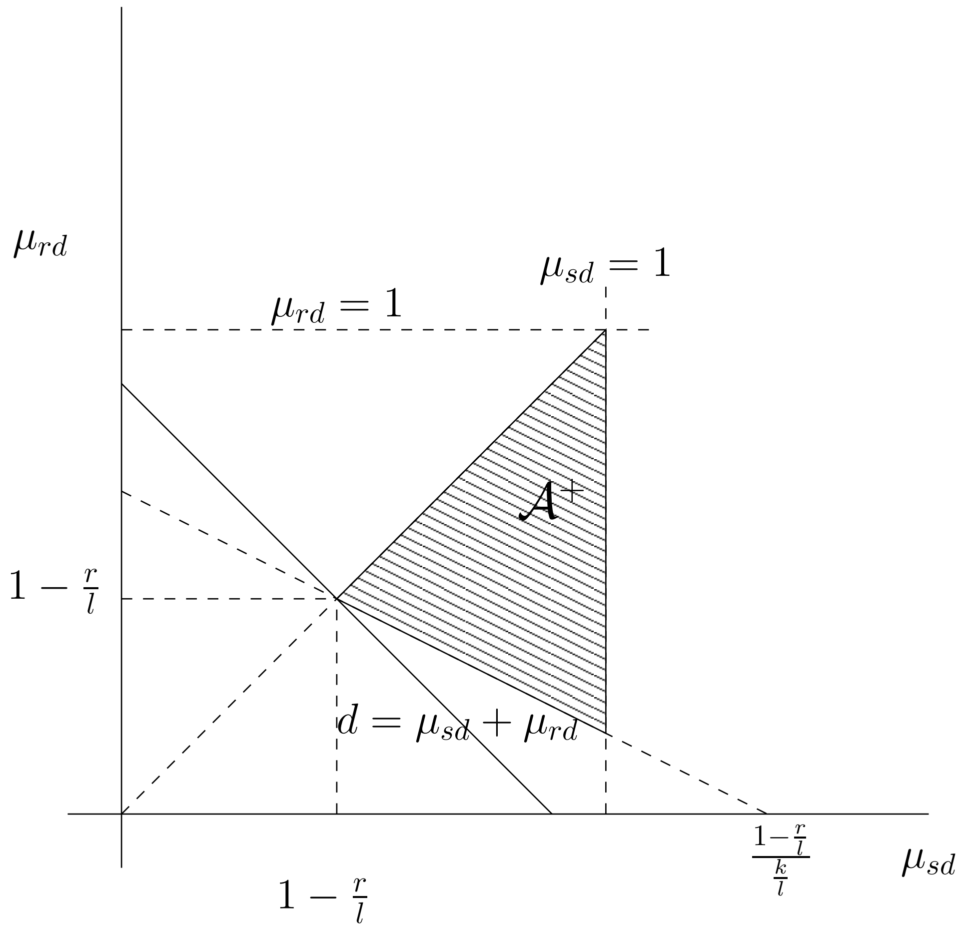


**Figure 5.9:** The region $\mathcal{A}^+$ for the long-term static channel where $\min(\mu_{sd}, \mu_{rd}) = \mu_{rd}$ and $k \leq \lfloor l/2 \rfloor$.

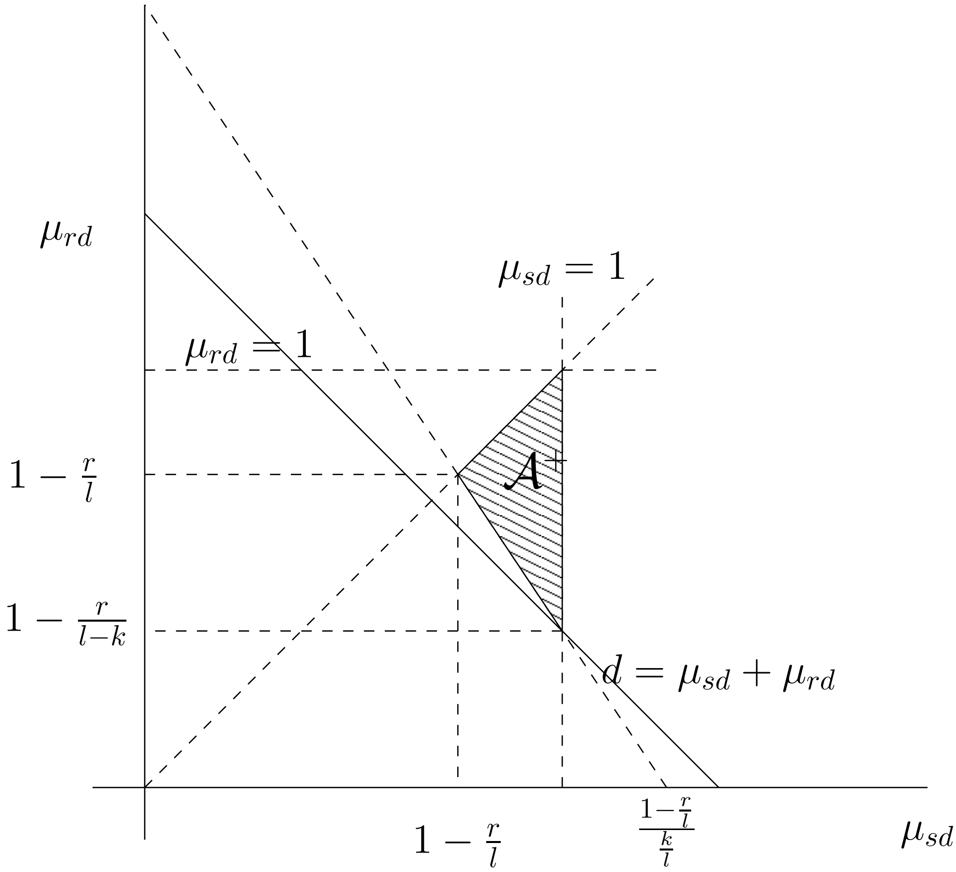


**Figure 5.10:** The region $\mathcal{A}^+$ for the long-term static channel where $\min(\mu_{sd}, \mu_{rd}) = \mu_{rd}$ and $\lfloor l/2 \rfloor < k \leq l - 1$.

# Chapter 6

# Conclusions

In this dissertation, we study issues related to schemes and architectures for wireless ad hoc networks. We create a framework for the establishment of the fundamental properties and limitations of such networks. This framework also highlights cross-layer design and cooperative communications as good design principles. In doing so, the performance evaluation of multi-hop wireless networks is emphasized.

In Chapter 2, we adopt an approach similar to the one developed in [9]. The main idea is to consider random topologies, to allow the number of nodes to go to infinity, and to compute the performance asymptotically. This approach is then applied in a simplified framework. This allows us to study different scenarios while keeping our proofs tractable. We construct a scheme to establish the asymptotic behavior of ad hoc wireless networks under non-uniform traffic patterns. This is motivated by the observation that the performance limitation of an ad hoc network comes first from the long-range peer-to-peer communication (that causes excessive interference) and second from the increase in relayed traffic in the case of multi-hop transmissions. For local traffic patterns, where the nodes communicate mostly with neighbors, we study the effect of the mean source destination distance on the throughput. The per-node throughput is improved compared to ad hoc wireless networks with unbounded average communication distance. Moreover, we show that there is a limit in the throughput improvement due to the connectivity condition. We address also the benefits of using a hybrid wireless network with respect to the per-node capacity. The base-stations are regularly placed within the network area and long-range communications are performed by this overlay network. We assess the tradeoff between the number of base-stations in the network and the increase in the throughput capacity of a hybrid ad hoc network due to the additional infrastructure. The gain in performance is mainly due to the reduction in the mean number of hops from source to destination. Finally, initial results on the state of applications for event-driven traffic patterns for wireless sensor networks are derived. We investigate event-driven traffic patterns where nodes report to the network collector measurements on a particular event. The impact of the event density on the throughput expressions is studied. The increase of throughput is due to the reduction of traffic load in the network since the number of communications is reduced (equivalent to the number of source destination pair in a ad hoc network). Then, we highlight possible extensions, models and scenarios for the event-driven traffic pattern case.

The above capacity results suggest certain guidelines for designing wireless ad hoc networks. Small networks scenarios seem to be a viable solution. Therefore, large wireless networks are feasible where nodes communicate mainly with nearby nodes. Such a scenario is envisaged in a number of emerging applications, like Bluetooth, HomeRF etc. Our model of hybrid wireless network could find application in mesh networks. This increases the coverage of actual wireless networks while keeping the transmit power at a reasonable level.

A future line of research would be to incorporate queueing delays in the formulation. Note that, in all our schemes, we assume that the creation of packets occurs at a regular and deterministic rate. If we assume that the arrival times are random, queuing delays are introduced. We also assume that the buffer size are infinite and nodes always have a packet to transmit, which is not the case in some scenarios of event-driven traffic pattern model. Another possible extension, would be to quantify the reduction in throughput due to communication overhead. Indeed, nodes inside the same cell (a partition of the network area as defined in Chapter 2) need to coordinate in order to schedule consecutive transmissions and to find informations on routes from the source to destination.

In Chapter 3, we propose the use of a connectivity graph, to study the long-term averaged throughput of ad hoc wireless networks with mobility. The connectivity graph offers an abstraction of the communication capabilities of the ad-hoc network, and forms a natural bridge between the literature in ah hoc wireless networks and the multicommodity flow problem. Using this graph, we propose a set of necessary and sufficient conditions to achieve constant throughput, and examine structural properties that these conditions imply. We also apply these conditions in a number of configurations in the literature, and demonstrate that they offer an alternative simpler methodology to re-derive these results.

Future work includes mapping the conditions derived from the connectivity graph on the mobility pattern. One could extend these results by linking the design of routing algorithms to the connectivity graphs. The design of such routing algorithms in wireless networks reduces to finding max-flow paths on a connectivity graph. Finally from a practical point of view, one could develop tools from linear programming to analyze examples of finite size.

In the second part of this dissertation, we focus on architectural design concepts and schemes for wireless ad hoc networks. While the first part of this dissertation deals with limit results in the number of nodes by considering an abstraction of network properties, the second part addresses design and performance evaluation of ad hoc networks. This is done by addressing properties crossing all the layers and some cooperation schemes.

In Chapter 4, we characterize the performance of decentralized multiple-access and retransmission schemes for multi-hop wireless networks. Our cross-layer framework jointly addresses the properties of the physical layer and the data link layer in the design of the media-access control (MAC) protocol and provide conclusions on routing strategies based on physical layer metrics. We investigate different transmission strategies in order to assess the tradeoff between spatial density of communications and the range of each transmission. We derive formulas for the spatial throughput for retransmission protocols and transmission strategies for random networks described by a spatial Poisson point process. It is shown that coding and retransmission protocols are a viable and simple solution for providing fully decentralized multiple-access communications in ad hoc wireless networks despite harsh propagation characteristics (interference from nearby competing nodes). Random exclusion and a decentralized protocol allow for the mitigation of the interference coming from other nodes. A routing protocol aiming to maximize the expected forward progress and exploiting multi-user diversity is shown to significantly out-perform other schemes. Finally we address some implementation issues and provide numerical results for a practical scenario.

A possible extension of this work is the impact of the exchange information and the need of coordination between nodes on the throughput. The implementation of such schemes and protocols requires an overhead of control information that can reduce the throughput of the network. Moreover, in this work nodes are assumed to be synchronized since a slotted transmission mode is considered. Hence, the study of distributed synchronization methods is a future direction. Future work will focus on more advanced cooperation strategies such as clustering, where for each group of nodes, a cluster head is elected to assign opportunity slot to different transmitters. This is a cluster oriented scheduling strategy where the cluster head needs some local information about the nodes in its vicinity. Practical coding strategies for incremental redundancy is also a possible future direction. Additionally, it would be interesting to study the impact of node density on the forward progress based on results from percolation theory as derived in [12].

Finally, in our analysis, we only consider single-user decoder. Nodes treat the interference from other simultaneous transmitters as noise. A possible future line of research would be to investigate the analysis of multi-user detection techniques and the impact on the outage probabilities and the forward progress.

In Chapter 5, we examine cooperative diversity techniques where a terminal relays signals of a source to create a virtual antenna array. Benefits are gained by exploiting spatial diversity in channel to combat multipath fading. Moreover in our scenario, the relay represents an additional resource that is freely utilized by the source terminal. We present an ARQ based protocol for the fading relay channel where diversity is exploited through the cooperative relay terminal and through time diversity from the ARQ protocol. We impose a half-duplex constraint on the transmission mode, but we assume that the relay and the source can transmit simultaneously (non-orthogonal transmission). Indeed, when the relay is able to decode, both the relay and the source send the same data to the destination providing additional gains. The performance characterization of this scheme is in terms of the achievable diversity, multiplexing gain and delay tradeoff for a high signal-to-noise ratio (SNR) regime. Then we construct a power control algorithm for this ARQ single relay channel. For simplicity, we consider a power control in which the relay is restricted to use a constant power in each round, but the source has the ability to vary its power to meet a long term average power constraint. The diversity gain obtained using power control is significant compared to the constant power case especially at low multiplexing gains. Moreover, the proposed power control is deterministic in the sense that it does not depend on the knowledge of the channel state but requires only the knowledge of the outage probabilities, which can be estimated, and the limited feedback from the destination.

An extension of this work would be the impact of multiple antennas (in particular two antennas considering the practical implications) at the receiver (base-station), where the source and the relay collaborate to reach the destination. Another avenue would be to investigate the extension of these schemes to the case of multiple relays relaying the information for a single source destination pair. This protocol can then also be applied to ad-hoc TDMA wireless networks where in each slot all the remaining nodes in the network act as relays for a particular source destination pair. Another generalization could be the broadcast network where relays also have information to receive.

A future line of research would be to incorporate the design of practical coding and decoding algorithms in this framework. Indeed, throughout this chapter, we have employed random coding arguments to evaluate performance of ARQ decode-and-forward scheme. It would be of great value to map the performance advantages of our scheme at high SNR to advantages in terms of bit error rate for different code designs. The investigation of rateless codes [103] as an alternative implementation of incremental redundancy could be a promising solution. More generally, designing effective algorithms, evaluating performance, and selecting codes are necessary for practical implementation of cooperative diversity.

# Bibliography

[1] C.-K. Toh, "Ad Hoc Mobile Wireless Networks: Protocols and Systems", *1st ed. Englewood Cliffs, NJ: Prentice Hall*, December 2001.

[2] "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications", *IEEE Standard 802.11-1997*, IEEE Computer Society LAN MAN Standards Committee, 1997.

[3] Z. J. Haas, M. Gerla, D. B. Johnson, C. E. Perkins, M. B. Pursley, M. Steenstrup, C.-K. Toh, *Eds., IEEE J. Select. Areas Commun., Special Issue on Wireless Ad Hoc Networks*, Volume 17, no. 8, August 1999.

[4] D. K. Kim,"A New Mobile Environment: Mobile Ad Hoc Networks (MANET)," *IEEE Vehic. Tech. Soc. News*, August 2003, pp. 29-35.

[5] J. C. Haartsen, "The Bluetooth radio system," *IEEE Personal Communications*, Volume 7, no. 1, pp. 28-36, February 2000.

[6] G. J. Pottie, "Wireless Sensor Networks", *in Proceedings IEEE Information Theory Workshop*, Volume 1, Killarney, Kerry, Ireland, June 1998, pp. 139-140.

[7] I.F.X.W. Akyildiz, "A survey on wireless mesh networks", *IEEE Communications Magazine*, Volume: 43, Issue: 9, pp. 23-30, September 2005.

[8] C. E. Shannon, "A mathematical theory of communication", *Bell System Technical Journal*, Volume 27, pp. 379-423 and 623-656, July and October, 1948.

[9] P. Gupta, P.R. Kumar, "The Capacity of Wireless Networks " *IEEE Transactions on Information Theory*, Volume: 46 Issue: 2, pp: 388 -404, Mars 2000.

[10] P. Gupta, P. R. Kumar, "Critical Power for Asymptotic Connectivity in Wireless Networks," pp. 547-566, *in Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming.* Edited by W.M. McEneany, G. Yin, and Q. Zhang, Birkhauser, Boston, 1998. ISBN 0-8176-4078-9.

[11] O. Dousse, "Asymptotic properties of wireless multi-hop networks", Thèse EPFL, no 3310, 2005.

[12] O. Dousse, P. Thiran,"Connectivity vs capacity in dense ad-hoc networks", *Proceedings INFOCOM*, Hong Kong, 2004.

[13] S. Ulukus, R. Yates, "Stochastic power control for cellular radio systems," *IEEE Transactions on Communication*, Volume 46, no. 6, pp. 784-798, 1998.

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

[15] R. Gallager, "A perspective on multiaccess channels," *IEEE Transactions on Information Theory (Special Issue on Random Access Communications)*, Volume IT-31, no. 2, pp. 124-142, March 1985.

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

[17] N. Abramson, "The ALOHA system: Another alternative for computer communications", *in Proceedings AFIPS Fall Joint Computer Conference*, Volume 37, Mountvale, NJ, November 1970, pp. 281-285.

[18] L. Kleinrock, F. A. Tobagi, "Packet switching in radio channels: Part I Carrier Sense Multiple Access models and their throughput-delay characteristics", *IEEE Transactions Commun.*, Volume 23, no. 12, pp. 1400-1416, December 1975.

[19] P. Karn, "MACA - A new channel access method for packet radio," *in Proceedings of the 9th Computer Networking Conference*, September 1990, pp. 134-140.

[20] V. Bharghavan, A. Demers, S. Shenkar, L. Zhang, "MACAW: a media access protocol for wireless LANs," *in Proceedings of the SIGCOMM '94 Conference on Communications Architectures, Protocols and Applications*, August 1994, pp. 212-225.

[21] M. Nor, J. Semarak,"Performance of CSMA-CA MAC protocol for distributed radio local area networks," *Proceedings IEEE PIMRC '95* Toronto, Canada, pp. 912-916, September 27-29, 1995.

[22] D. P. Bertsekas, R. Gallager, "Data Networks", *2nd ed. Englewood Cliffs, NJ: Prentice Hall*, December 1991.

[23] S.Basagni, M. Conti, S.Giordano, I .Stojmeovic, "Mobile Ad Hoc Networking", *Willey Inter Science IEEE Press2004*, ISBN 0-471373133.

[24] C.K. Toh, "Associativity-Based Routing for Ad-Hoc Mobile Networks". *Wireless Personal Commun., 4: 1-36*, 1997.

[25] L. Xie, P.R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation." *IEEE Transactions on Information Theory, Volume 50, no. 5, pp. 748-767*, May 2004.

[26] M. Grossglauser, D. N. .C Tse, "Mobility Increases The Capacity of Ad Hoc Wireless Networks " *IEEE/ACM Transactions on Networking*, Volume: 10 Issue: 4, pp: 477-486, August 2002.

[27] P. Gupta, P.R. Kumar, "Internets In The Sky: The Capacity of Three Dimensional Wireless Networks," Communications in Information and Systems, Volume 1, issue 1, pp. 33-49, January 2001.

[28] S. Toumpis, A.J. Goldsmith, "Ad Hoc Network Capacity", *in Proceedings Asilomar Conference of Signals, Systems and Computers*, 2000.

[29] J.Li, C.Blake, D.S.J. De Couto, H.I. Lee, R. Morris,"Capacity of Ad Hoc Wireless Networks", ACM SIGMOBILE, pp. 61-69, July 2001, Rome, Italy.

[30] T. Tabet, R. Knopp, "Throughput Analysis For Decentralized Slotted Regular Wireless Networks", *in Proceedings Asilomar Conference on Signals, Systems, and Communications*, November 2003.

[31] M. Gastpar, M. Vetterli, "On the capacity of wireless networks: The relay case", *In Proc IEEE Infocom 2002*, June 2002.

[32] O. Dousse, P. Thiran, M. Hasler,"Connectivity in Ad-Hoc and Hybrid Networks", *in Proceedings IEEE Infocom 2002*, pp. 1079-1088, June 2002.

[33] B. Liu, Z. Liu, D. Towsley,"On The Capacity of Hybrid Wireless Networks", *in Proceedings IEEE Infocom 2003*, 2003.

[34] S.R. Kulkarni, P. Viswanath, "A Deterministic Approach to Throughput Scaling in Wireless Networks", *to appear in IEEE Transactions on Information Theory*, June 2004.

[35] M. Franceschetti, O. Dousse, D. Tse, P. Thiran, "Closing the gap in the capacity of random wireless networks", in *Proceedings of International Symposium of Information Theory (ISIT)*, Chicago, Illinois, July 2004.

[36] O. Dousse, M. Franceschetti, P. Thiran,"Information theoretic bounds on the throughput scaling of wireless relay networks", *IEEE INFOCOM*, Miami, March 2005.

[37] O. Lévêque, E. Telatar, "Information theoretic upper bounds on the capacity of large extended ad-hoc wireless networks", *IEEE Transactions on Information Theory*, Volume: 51 Issue 3, pp: 858–865, March 2005.

[38] S. Toumpis, A. Goldsmith, "Capacity bounds for large wireless networks under fading and node mobility", *in Proceedings of Allerton Annual Conference on Communications, Control, and Computing*, Monticello, IL, 2003.

[39] A. Agarwal, P. Kumar, "Capacity bounds for ad-hoc and hybrid wireless networks", *ACM SIGCOM Computer Communication Review*, Volume: 34, No 3, pp: 71-81, July 2004.

[40] M. Franceschetti, O. Dousse, D. Tse, P. Thiran, "On the throughput capacity of random wireless networks", *under revision.*

[41] A. Jovisic, P. Viswanath, S. Kulkarni, "Upper bounds to transport capacity of wireless networks", *IEEE Transactions Information Theory*, Volume: 50, No 11, pp: 2555-2565, November 2004.

[42] S. Diggavi, M. Grossglauser, D. Tse, "Even one-dimensional mobility increases capacity of wireless ad-hoc networks", *in Proceedings IEEE International Symposium on Information Theory (ISIT)* , Lausanne, June 2002.

[43] H. Sawant, T. Jindong, Y. Qingyan, W. Qizhi, "Using bluetooth and sensor networks for intelligent transportation systems", *in Proceedings of 7th international IEEE conference on intelligent transportation systems*, Washington DC, 2004.

[44] T. Leighton, S. Rao, "Multicommodity max-flow min-cut theorems and their use in designining approximation algorithms", *Journal of the ACM*, Volume: 46, Issue: 6, pp: 787-832, November 1999.

[45] D. B. Shmoys, "Cut problems and their application to divide-and-conquer", *Approximation algorithms for NP-hard problems*, pp. 192-235, 1996.

[46] P. N. Klein, S. Plotkin, S. Rao, "Excluded minors, network decomposition, and multicommodity flow", *in Proceedings of 25th ACM Symposium on Theory of Computing (1993)*, pp: 682-690, 1993.

[47] N. Harvey, R. D. Kleinberg, A. Lehman, "Comparing network coding with multicommodity flow for the k-pairs communication problem", *MIT report, submitted.*

[48] K. Jain. P. Chou, V. Vazirani, R. Yeung, "On the capacity of multiple unicast sessions in undirected graphs", *in Proceedings of International Symposium on Information Theory, Adelaide September 2005.*

[49] C. Fragouli, T. Tabet, C. Derdiyok, "On the throughput of ad hoc networks using connectivity graphs", *in Proceedings of Allerton Annual Conference on Communications, Control, and Computing*, Monticello, IL, October 2004.

[50] Y. Wu, P. A. Chou, Q. Zhang, K. Jain, W. Zhu, S. Y. Kung, "Network planning in ad-hoc wireless networks: a cross layer approach", *IEEE Journal on Selected Areas on Communications*, Volume: 23, No 1, Jan. 2005.

[51] A. Lozano, S. Kulkarni, P. Viswanath, “Throughput scaling in wireless networks with restricted mobility”, *IEEE International Symposium on Information Theory (ISIT)* , Chicago, June 2004.

[52] T. Tabet, R. Knopp, “Bounds on the throughput capacity of wireless ad-hoc network with non-uniform traffic”, *in Proceedings of First Annual IEEE Communications Society Conference on Sensor and Ad-Hoc Communications and Networks, IEEE SECON*, October 2004.

[53] N. Alon, “Eigenvalues and Expanders”, *manuscript*, 1985.

[54] A. Al Hanbali, R. Groenevelt, P. Nain, E. Altman,“Impact of Mobility on the Performance of Relaying in Ad-Hoc Networks”, *Proceedings of IEEE Infocom*, Barcelona 2006.

[55] A. G. Dimakis, A. D. Sarwate, M. J. Wainwright, “Geographic Gossip: Efficient Aggregation for Sensor Networks”, *Proceedings of IPSN*, 2006

[56] J. Bang-Jensen, G. Gutin, “Digraphs: Theory, Algorithms and Applications”, *Springer-Verlag 2002.*

[57] R. Knopp, N. Nikaein, C. Bonnet, H. Aiache, V. Conan, S. Masson, G. Guibe, C. Le Martret, “Overview of the WIDENS Architecture, a Wireless Ad Hoc Network for Public Safety”, *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks SECON*, Santa Clara, USA, 2004.

[58] G. Caire, D. Tuninetti, “The throughput of hybrid-ARQ protocols for the Gaussian collision channel ” *IEEE Transactions on Information Theory*, Volume: 47 Issue: 5, pp: 1971 -1988, July 2001.

[59] J. Massey, P. Mathys “The collision channel without feedback”, *IEEE Transactions Info. Theory*, Volume 31, pp. 192-204, March 1985.

[60] R. Nelson, L. Kleinrock, “The Spatial Capacity of a Slotted ALOHA Multihop Packet Radio Network with Capture” *IEEE Transactions on Communications*, Volume: COM-32, No. 6, June 1984.

[61] F. Baccelli, B. Blaszczyszyn, P. Muhlethaler, “A Spatial Reuse Aloha MAC Protocol for Multihop Wireless Mobile Networks”, *in Proceedings of Annual Conf. on Communication, Allerton*, September, 2003.

[62] Y.-B. KO, N. H. Vaidya "Location-Aided Routing in mobile Ad hoc networks", *In Proceedings ACM/IEEE Mobicom*, pp 66-75, October 1998.

[63] B. Karp, "Geographic Routing for Wireless Networks". *PhD thesis*, Harvard University, 2000

[64] T. Schwengler, M. Gilbert, "Propagation models at 5.8 GHz-path loss and building penetration", *Radio and Wireless Conference*, RAWCON 2000, pp: 119-124, 10-13 September 2000.

[65] M. Barbeau, E. Kranakis, D. Krizanc, P. Morin, "Improving Distance Based Geographic Location Techniques in Sensor Networks", *in Proceedings of ADHOC-NOW 04, Springer Verlag, LNCS*, pp. 197-210, Volume 3158, 2004.

[66] T.S. Rappaport, "Wireless Communications: Principles and Practice", *Series: Prentice Hall Communications Engineering*, December 2001

[67] S.Shamai and A.Wyner, "Information theoretic considerations for symmetric cellular, multiple-access fading channels - Part I," *IEEE Transactions on Information Theory,* Volume 43, No. 6, pp. 1877-1894, November 1997.

[68] L. H. Ozarow, S. Shamai, A. D. Wyner, "Information theoretic considerations for cellular mobile radio". *IEEE Transactions Veh. Technol.*, 43(5) pp:359-378, May 1994.

[69] G. Caire, R. Knopp, P. Humblet, "System capacity of F-TDMA cellular systems " *IEEE Transactions on Communications* Volume: 46 Issue:12, pp: 1649 -1661, December 1998.

[70] E. S. Sousa, "Performance of a Spread Spectrum Packet Radio Network Link in a Poisson Field of Interferers" *IEEE Transactions on Information Theory*, Volume 38, No. 6, November 1992.

[71] I. M. I. Habbab, M. Kavehrad, C. -E. W. Sundberg, "ALOHA with capture over slow and fast fading radio channels with coding and diversity ", *IEEE Journal on Selected Areas in Communications*, Volume: 7 Issue: 1, pp: 79 -88, January 1989.

[72] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications " *IEEE Transactions on Communications*, Volume: 36 Issue: 4, pp: 389 -400, Apr. 1988.

[73] C.-F. Leanderson, G. Caire, "The performance of incremental redundancy schemes based on convolutional codes in the block-fading Gaussian collision channel", *IEEE Transactions on Wireless Communications*, Volume : 3, No.: 3, pp: 843-854, 2004.

[74] S. Sesia, G. Caire, G. Vivier, "Incremental redundancy Hybrid ARQ schemes based on low-density parity-check codes", *IEEE Transactions on Communications*, Volume 52, No. 8, pp. 1311-1321, 2004.

[75] L. Kleinrock, J. Silvester, "Optimum transmission radii for packet radio networks or why six is a magic number", *in Conf. Rec., Nat. Telecommun. Conf.*, pp. 431-435, December 1978.

[76] W.H. Yuen, R.D. Yates, "Optimum Transmit Range and Capacity of Mobile Infostation Networks", *in IEEE GLOBECOM 2003*, San Francisco, December 2003.

[77] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 1984.

[78] R. Knopp and P.A. Humblet, "Information Capacity and Power Control in Single-Cell Multiuser Communications", *Proceedings IEEE ICC '95,* Seattle, WA, June 1995.

[79] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume II.* Wiley, 1966.

[80] E.C. van der Meulen, "Three-terminal communication channels", *Advanced Applied Probability*, Volume 3, pp. 120-154, 1971.

[81] T.M. Cover, A.A. El Gamal, "Capacity theorems for the relay channel", *IEEE Transactions Information Theory*, Volume 25, pp. 572-584, September 1979.

[82] Thomas M. Cover and Joy A. Thomas. "Elements of Information Theory". *John Wiley & Sons, Inc.*, New York, 1991.

[83] G. Kramer, M. Gastpar, P. Gupta, "Cooperative strategies and capacity theorems for relay networks", *IEEE Transactions on Information Theory*, 51(9):3037-3063, September 2005.

[84] Anders Host-Madsen, "Capacity Bounds for cooperative diversity", *Transactions Information theory*, Volume: 52, Issue: 4, page(s): 1522-1544, April 2006.

[85] M. Khojastepour, A. Sabharwal, B. Aazhang, "On the Capacity of 'Cheap' Relay Networks", *in Proceedings of Conference on Information Sciences and Systems*, April 2003.

[86] A. Sendonaris, E. Erkip, B. Aazhang, "User Cooperation Diversity. Part I. System Description," *IEEE Transactions On Communications,*, pp. 1927-1938, Volume 51, Issue: 11, November 2003.

[87] A. Sendonaris, E. Erkip, B. Aazhang, "User Cooperation Diversity– Part II: Implementation Aspects and Performance Analysis", *IEEE Transactions on Communications*, pp. 1939-1948, Volume 51, Issue:11, November 2003.

[88] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Transactions Information Theory*, Volume 50, no. 12, pp. 3062-3080, December 2004.

[89] K. Azarian, H. El Gamal, P. Schniter,"On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels", *IEEE Transactions on Information Theory*, Volume: 51, Issue: 12, page(s): 4152- 4172, December 2005.

[90] T.Tabet, S. Dusad, R. Knopp, "ARQ Based Half-Duplex Cooperative Diversity Protocol", *in Proceedings of Canadian Workshop on Information Theory*, Montreal, June 2005.

[91] T.Tabet, S. Dusad, R. Knopp, "Achievable Diversity-Multiplexing-Delay Tradeoff in Half-Duplex ARQ Relay Channels", *in Proceedings of IEEE International Symposium on Information Theory* Adelaide, Australia, September 2005.

[92] S. Verdú, T. S. Han, "A general formula for channel capacity". *IEEE Transactions Information Theory*, 40(5):1147-1157, July 1994.

[93] E. Biglieri, J. Proakis, S. Shamai, "Fading channels: Information theoretic and communications aspects". *IEEE Transactions Information Theory*, 44(6):2619-2692, October 1998.

[94] S. V. Hanly, D N.C. Tse, "Multiaccess fading channels — Part II: Delaylimited capacities". *IEEE Transactions Information Theory*, 44(7):2816-2831, November 1998.

[95] L. Zheng and D. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels", *IEEE Transactions on Information Theory*, Volume 49, pp. 1073-1096, May 2003.

[96] I.E. Telatar, "Capacity of multi-antenna Gaussian channels", *European Transactions on Telecomm.*, 10(6):585-596, November 1999.

[97] H. El Gamal, G. Caire, M. O. Damen, "The MIMO ARQ channel: Diversity-multiplexing-delay tradeoff", *IEEE Transactions on Information Theory*, Volume 52, Issue 8, pp. 3601-3621, August 2006.

[98] Y. Nam, K. Azarian, H. El Gamal , "Cooperation Through ARQ", *in Proceedings of IEEE 6th Workshop on Signal Processing Advances in Wireless Communications (SPAWC05)*, June 5-8 2005, New York, pp. 1023-1027.

[99] T. Tabet and R. Knopp, "Spatial throughput of multi-hop wireless networks under different retransmission protocols", *in Proceedings Allerton Conf. on Communication, Control, and Computing,* Monticello, IL, October 2004

[100] N. Ahmed, M. Khojastepour, B. Aazhang, "Outage Minimization and Optimal Power Control for the Fading Relay Channel", *in Proceedings of IEEE Information Theory Workshop, Houston*, *Submitted May 2004.*

[101] G. Caire, G. Taricco, and E. Biglieri, "Optimum Power Control Over Fading Channels," *IEEE Transactions Information Theory*, Volume 45, no. 5, pp. 1468-1489, July 1999.

[102] N. Ahmed, M. Khojastepour, A. Sabharwal, B. Aazhang, "On power control with finite rate feedback for cooperative relay networks", *in Proceedings of International Symposium on Information Theory and its Applications*, *Submitted, March 2004*.

[103] A. Shokrollahi, "Raptor Codes", *IEEE Transactions on Information Theory*, Volume 52, Issue 6, pp. 2551-2567. June 2006.

# Curriculum Vitae

## Tarik Tabet

Born the $4^{\text{th}}$ of October, 1978

**Home Address**
Avenue des Collèges 39,
1009 PULLY, Switzerland

**Phone** : +41 78 611 9565
**e-mail** : tarik.tabet@gmail.com
http://lcmwww.epfl.ch/people/∼tabet.htm

---

### EDUCATION

2002-present **EPFL, School of Computer and Communication Sciences**
PhD Student/Research Assistant, in the Mobile Communications Laboratory (LCM).
PhD thesis entitled *"Schemes and Architectures for Wireless Ad Hoc Networks and Cooperative Communications "*, under the supervision of Prof. R. Knopp and Prof. B. Rimoldi.

1996-2001 **EPFL, Communication Systems Department**
Diploma Engineer degree.
Diploma grade 5.5/6, Diploma work grade 6/6.
**Major:** Mobile Communications, **Eurecom Institute, France.**
Recipient of EPFL scholarship 1997-2000.

2000-2001 **Doctoral School, University of Sophia-Antipolis, France.**
DEA (specialized diploma) in Communication and Signal
Supervised by Prof. Pierre Comon.
Ranked first.

---

## WORK EXPERIENCE

| | |
|---|---|
| 2002–present | **Research and teaching assistant** LCM, EPFL.<br>**Research area**: Wireless Ad Hoc Networks, Multi-hop Communication, Cooperation Strategies, Cross-layer design.<br>**Teaching experience**: Advanced Digital Communications Software Defined Radio Course, Wireless Communications. |
| 09/2005–03/2006 | **6 months internship at Qualcomm, Standards group, Corp. R&D, San Diego, CA, USA**.<br>**Subject**: Mesh Networks, Continuous packet connectivity (CPC) for UL WCDMA, MC-CDMA.<br>Designed a cooperative MAC based on 802.11 for wireless mesh networks.<br>Submitted two contributions to the **3GPP standard**.<br>CPC proposal accepted as the baseline. |
| 03/2001–09/2001 | **6 months internship at Iospan Wireless PHY Group, in San Jose, CA, USA**.<br>**Subject**: Design of the receiver architecture for a MIMO-OFDM fixed wireless system. |
| February 2001 | **Alcatel at the Third World GSM Congress, France**.<br>Studied and presented various 3G products of Alcatel. |

## RESEARCH INTERESTS

Wireless Networks
Wireless Communications
Information Theory and Coding
Signal Processing
Graph Theory

## OTHER PROFESSIONAL ACTIVITIES

Member of the Terminodes project on self-organizing, mobile communication and information services (National Center of Competence in Research on Mobile Information and Communication Systems NCCR-MICS).

Reviewing Activity (Journals): IEEE Transaction on Information Theory, IEEE Transaction on Communications, IEEE Journal on Selected Areas on Communications, Springer Journal on Wireless Networks.

Reviewing Activity (Conferences): 2002 IEEE International Conference on Communications (ICC), 2004 Vehicular Technology Conference (VTC) Spring, Milano Italy.

## COMPUTER SKILLS

UNIX, Linux, Windows NT/XP platforms
C, C++, Java, Opnet, Mathematica, Matlab, SQL, assembler, Labview.

## LANGUAGES

French: bilingual; Arabic: mother tongue; English: fluent.

## OTHER INTERESTS AND ACTIVITIES

| | |
|---|---|
| Activities | Member of Agepoly (students' union of the EPFL ). |
| | Organization of Aquapub 97 (student event). |
| | Managed a Youth Hostel in Lausanne, Switzerland, August 1997. |
| Sport | Karting, Soccer, Badminton, Table tennis. |

## PUBLICATIONS

[1] T. Tabet and R.Knopp,
*Throughput analysis for slotted peer-to-peer wireless networks*, in IEEE Winter School on Coding and Information Theory 2003, Monte Verita, Ascona, Switzerland.

[2] T.Tabet and R.Knopp,
*Throughput analysis for decentralized slotted regular wireless networks*, 37th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, Nov 2003.

[3] T.Tabet and R.Knopp,
*Spatial throughput of multi-hop wireless networks under different retransmission protocols*, in Proc. 42th Annual Allerton Conference on Communication, Control and Computing, Monticello, Illinois, October 2004.

[4] C. Fragouli, T. Tabet, and C. Derdiyok,
*On the throughput of Ad Hoc wireless networks using connectivity graphs*, in Proc. 42th Annual Allerton Conference on Communication, Control and Computing, Monticello, Illinois, October 2004.

[5] T.Tabet and R.Knopp,
*Bounds on the throughput capacity of wireless Ad Hoc networks with non-uniform traffic*, The First IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON 2004), Santa Clara, CA, USA, 4-7 Oct 04.

[6] T. Tabet and R.Knopp,
*Spatial throughput of multi-hop wireless networks*, in IEEE Winter School on Coding and Information Theory Feb. 2005, Bratislava, Slovakia.

[7] T.Tabet, S. Dusad, R. Knopp,
*ARQ Based Half-Duplex Cooperative Diversity Protocol*, Canadian Workshop on Information Theory, Montreal, June 2005.

[8] T.Tabet, S. Dusad, R. Knopp,
*Achievable Diversity-Multiplexing-Delay Tradeoff in Half-Duplex ARQ Relay Channels*, IEEE International Symposium on Information Theory, Adelaide, Australia, Sept 2005.

[9] T.Tabet, T. Luo, J. Montojo,
*On The Demodulation Impact of DPCCH Gating IN HSUPA*, WMPC, San Diego, Sept 2006.

[10] C. Fragouli, T.Tabet,
*Conditions for Constant Throughput in Wireless Networks*, accepted for publication, ACM Transactions on Sensor Networks, July 2006.

[11] T.Tabet, S. Dusad, R. Knopp,
*Diversity-Multiplexing-Delay Tradeoff in Half-Duplex ARQ Relay Channels*, submitted to IEEE Transactions on Information Theory, Aug 2006.

[12] T. Tabet and R.Knopp,,
*ARQ Based Spatial Throughput of Multi-Hop Wireless Networks: Conclusions on Cross-Layer Design*, in preparation.