UNIVERSITÀ DI PADOVA FACOLTÀ DI INGEGNERIA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

SCUOLA DI DOTTORATO IN INGEGNERIA DELL'INFORMAZIONE

INDIRIZZO IN SCIENZA E TECNOLOGIA DELL'INFORMAZIONE

XXIII Ciclo

Interference management in wireless networks: MAC Protocols for MIMO Ad Hoc Networks and Cooperative Routing in Cognitive Radio Networks

Dottorando

DAVIDE CHIAROTTO

Supervisore:

Chiar.mo Prof. Michele Zorzi

Direttore della Scuola:

Chiar.mo Prof. Matteo Bertocco

Anno Accademico 2010/2011

A mia madre, a mio padre e a mia sorella,
per tutto quello che mi avete insegnato,
con poche parole, con tanto amore.

Ad Alessandra,
perché è successo,
nel modo che prima non sapevo.

Contents

List of Figures

Abstract

MIMO ad hoc networks and cognitive radio networks have enjoyed great interest in recent years. The performance gains of these frameworks depend on the ability to control the transmissions in the wireless channel. In particular, MIMO ad hoc networks exploit multiuser communications to favor simultaneous transmissions and cognitive radio networks regulate the access of unlicensed users to the licensed network in order to improve the available bandwidth utilization. In both cases, the main limiting factor is represented by the interference that could cause a performance degradation if not properly managed. In this work, we adopt a broad approach for the study of interference and propose new analytical approaches that make it possible to quantify and solve the adverse effects of interference in both MIMO ad hoc and cognitive radio networks.

The first scenario that we consider deals with ad hoc networks with multiple antennas and multiuser communications. We take into account a cross-layer networking protocol which integrates medium access control and physical layer with the objective to obtain a tradeoff between throughput and interference rejection. At the receiver side, the presence of several simultaneous signals is managed by using a Vertical-Bell Laboratories Layered Space-Time (V-BLAST) receiving scheme that decouples the superimposed signal to extract data through a successive interference cancellations process, which is potentially prone to interference, especially when the quality of the channel estimation is poor. In multiuser scenarios, these problems may lead to substantial loss of data. In this light, we propose an analytical technique that evaluates the statistics of the channel estimation errors and develop an analysis for both correlator-based and Minimum Mean-Square Error channel estimators, showing that there is a direct dependence of the channel estimation error on the instantaneous channel matrix, which also includes interfering transmitters. This is done to obtain precise expressions that can be used in analytical studies as well as realistic simulation ex-

periments. In this light, we include the effect of channel estimation errors in an ad hoc networking protocol simulator and thoroughly evaluate their impact.

A different point of view is required to handle cooperative routing techniques in cognitive radio networks, the second type of wireless network considered in this Thesis. Here, the problem of scarce radio spectrum availability and the inefficiency of traditional fixed spectrum management schemes call for new communications paradigms for spectrum sharing. Specifically, the presence of unlicensed (or secondary) users that transmit in the portion of spectrum licensed to primary users generates interference, avoidable only if secondary users adopt some sensing technique before they decide to transmit. A new paradigm, named spectrum leasing, allows licensed users to lease portions of the spectrum to unlicensed users, avoiding both interference and the need for spectrum sensing. Specifically, the secondary nodes may cooperate with the primary users serving as extra relays, but only in exchange for spectrum leasing. Namely, in return for their forwarding of primary packets, secondary nodes are awarded spectral resources by primary users for transmission of their own traffic. Thus, secondary nodes enforce minimal quality of service requirements in terms of rate and reliability when deciding whether to cooperate. Results demonstrate the advantages of the proposed spectrum leasing solution based on opportunistic routing and highlights the available trade-offs between primary throughput and energy consumption.

Sommario

Negli ultimi anni le tematiche riguardanti le reti MIMO di tipo “ad hoc” e le reti di tipo “cognitive” hanno riscosso grande interesse. Le prestazioni che i suddetti paradigmi comunicativi riescono ad ottenere dipendono, essenzialmente, dalla capacità dei medesimi di controllare le trasmissioni sul canale wireless. La peculiarità di ciascuno dei due sistemi consiste, nelle reti MIMO ad hoc, nel favorire le trasmissioni simultanee sfruttando le comunicazioni multi-utente; nelle reti cognitive, nel permettere l’accesso al canale wireless anche agli utenti privi di licenza, regolandone il comportamento in modo da migliorare l’utilizzo delle risorse radio. In entrambi i casi, tuttavia, l’interferenza può rappresentare, se non adeguatamente gestita, il principale fattore limitante delle prestazioni. Per tale ragione, nel presente elaborato, viene adottato un approccio di ampio respiro per lo studio dell’interferenza e vengono proposti nuovi metodi analitici che ci permettono di quantificare e risolvere gli effetti negativi generati dall’interferenza nelle reti MIMO ad hoc e nelle reti cognitive.

Il primo scenario considerato in questa Tesi sono le reti ad hoc con antenne multiple e comunicazioni multi-utente. Si prende, altresì, in considerazione l’utilizzo di un protocollo di tipo cross-layer in grado di integrare il controllo di accesso al mezzo e lo strato fisico, al fine di ottenere un buon compromesso tra velocità di trasmissione e controllo dell’interferenza. Questo viene portato a termine grazie all’utilizzo a lato ricevente di un particolare decodificatore (*Vertical-Bell Laboratories Layered Space-Time*, V-BLAST), il quale riesce a gestire le comunicazioni multi-utente disaccoppiando dal segnale in ricezione i dati dei vari utenti attraverso un processo di cancellazioni successive. Tale tecnica è molto promettente, ma è potenzialmente incline all’interferenza, soprattutto quando la qualità della stima di canale è scadente. Le sopra descritte problematiche, presenti anche nei scenari multi-utente, possono causare la perdita di molti dati. Ciò detto, nel presente lavoro viene proposta una nuova tec-

nica tesa a valutare analiticamente le statistiche degli errori di stima di canale per due tipi di stimatori (stimatore a correlazione e MMSE) ed a dimostrare la sussistenza di una dipendenza diretta dell'errore di stima di canale sulla matrice di canale istantanea, la quale, inoltre, include pure l'interferenza causata dalle trasmissioni simultanee. Si sono così ottenute delle espressioni matematiche che hanno permesso di effettuare sia precisi studi analitici sia esperimenti realistici, effettuati questi ultimi su uno simulatore di rete, inserendo l'effetto degli errori di stima di canale.

Per riuscire a gestire le tecniche di instradamento cooperative nelle reti radio cognitive, il secondo tipo di rete wireless considerato in questa Tesi, è necessario un approccio diverso rispetto al caso precedente. Qui, infatti, le principali problematiche, relative alla scarsa disponibilità di spettro radio e all'inefficienza dei classici sistemi statici di gestione dello spettro, sono state affrontate proponendo paradigmi innovativi di comunicazione per la condivisione dello spettro radio. In particolare, la presenza di utenti senza licenza (o utenti secondari), che trasmettono nella porzione dello spettro adibita esclusivamente a utenti primari, genera interferenza, evitabile solo se gli utenti secondari adottano particolari tecniche di rilevamento (da utilizzare prima di decidere se trasmettere o meno). Un nuovo paradigma, denominato "spectrum leasing", permette agli utenti autorizzati di allocare delle porzioni dello spettro di appartenenza agli utenti secondari, evitando così sia l'interferenza sia il bisogno di tecniche di rilevamento di altre trasmissioni. Dall'altro canto, i nodi secondari collaborano con gli utenti primari, proponendosi come ulteriori relay, solamente in cambio di una porzione dello spettro, ovvero in cambio della trasmissione di pacchetti primari, i nodi secondari richiedono agli utenti primari una porzione della risorsa spettrale per poter trasmettere il proprio traffico. Tale porzione viene calcolata in base ai propri requisiti di qualità di servizio, che devono essere rispettati. L'analisi del sopra descritto nuovo paradigma comunicativo, eseguita in questa tesi, ne evidenzia i vantaggi rispetto ad un approccio classico, mettendo, inoltre, luce sul compromesso tra velocità di trasmissione e consumo di energia primaria.

List of Acronyms

ACK	Acknowledgment
AWGN	Additive White Gaussian Noise Vector
BER	Bit Error Rate
CA	Collision Avoidance
CDMA	Code-Division Multiple Access
CRC	Cyclic Redundancy Check
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSMA	Carrier Sense Multiple Access
CTS	Clear-To-Send
FT	Follow Traffic
HARQ	Hybrid Automatic Repeat reQuest
iid	Independent and Identically Distributed
LSTC	Layered Space-Time Codes
MAC	Medium Access Control
MIMO	Multiple-Input-Multiple-Output
MMSE	Minimum Mean-Square Error

MPSK	M-ary Phase-Shift Keying
MQAM	M-ary Quadrature Amplitude Modulation
MSE	Mean-Square Error
PHY	Physical
PDU	Protocol Data Unit
QoS	Quality-of-Service
RC	Raised Cosine
RTS	Request-To-Send
SC	Superposition Coding
SER	Symbol Error Rate
SIC	Successive Interference Cancellation
SINR	Signal to Interference and Noise Ratio
SISO	Single-Input-Single-Output
SM	Spatial Multiplexing
SNR	Signal-to-Noise Ratio
STC	Space-Time Codes
TDM	Time Division Multiplexing
V-BLAST	Vertical-Bell Laboratories Layered Space-Time
ZF	Zero Forcing

Chapter 1

Introduction

Multiuser detection, multiple-input multiple-output (MIMO) systems, cooperative routing techniques and cognitive radios have recently gained increasing attention in the wireless networking community, due to network performance improvements that can be achieved in terms of throughput and full exploitation of the available radio spectrum. However, the presence of several simultaneous transmissions makes interference a key issue, that must be taken into account both for the network performance evaluation and especially in the design of protocols and routing policies. In this Thesis we adopt a broad approach for the study of multiuser interference and we propose possible solutions to this problem considering two types of wireless networks: ad hoc networks and cognitive radio networks.

The purpose of this introductory Chapter is to offer a broad perspective of the two wireless networks considered and to outline the problems that will be handled throughout the Thesis. In particular, Sections 1.1 and 1.2 present an introduction to ad hoc networks and cognitive radio networks, respectively. Finally, Section 1.3 summarizes the organization of this Thesis.

1.1 Ad Hoc Networks

A wireless ad hoc network is a decentralized wireless network where nodes, which can independently access (and leave) the network, communicate to each other without using a preexisting infrastructure (unlike in cellular networks where the base station coordinates the uplink and the downlink traffic in his cell). Such decentralized networks have enjoyed

significant interest in the research community in recent years, due to their capability of setting up a self-organizing wireless network in emergency or military scenarios, besides being suited for commercial applications and for quick communications setup in any environment where a cabled network is infeasible or not affordable.

Wireless communications are known to be strongly influenced by multipath fading and interference problems. In fact, the main effect of the multipath propagation of the signal is a time-varying behavior of the channel, especially in scenarios where line of sight is not available. This requires some additional complexity at the receiver, which is needed to successfully receive the incoming transmission irrespective of the particular channel realization. The second aspect is represented by the interference. The origin of this problem lies in the main characteristic of wireless transmissions: the wireless channel, shared by all users, is a broadcast medium. Each signal is received by all users that lie in the transmission range, which depends, e.g., on the transmit power and the thermal noise at the receiver. Therefore, each signal contains useful information for the intended receiver, but at the same time represents a source of interference for other users inside the transmission range. This can cause collisions between packets transmitted by different nodes, leading to a significant performance loss in terms of throughput and latency. So, suitable control mechanisms for the transmissions need to be designed. This problem has been handled by the ad hoc networking research community proposing smart protocols that regulate the access of users to the wireless channel.

One of the most common approaches considered is to use resource allocation strategies. This strategy is commonly used in cellular networks, in which the base station controls the channel access in a centralized method. In this case, each user can simultaneously transmit its signal only if the centralized controller gives it a portion of the shared resource. In the TDMA approach the shared resource is time, while in FDMA it is frequency, and in CDMA the knowledge of codes that permits to reduce the interference.

A second approach used to regulate the access of users to the wireless channel is based on the carrier sense multiple access with collision avoidance (CSMA/CA) protocol (described in the 802.11 standard [1]), which tries to regulate the channel access with a two step strategy. When a node is ready to transmit, it listens to the channel to ensure the absence of other communications with which it would probably interfere (first step) and, if the channel is sensed

free, it transmits the signal (second step). The channel is considered free (*idle* channel) when the perceived power by the node during the first step is below a certain threshold. Clearly, this scheme, that tries to avoid the occurrence of packet collisions, is conservative and does not exploit the spatial diversity offered by the wireless channel. Moreover, collisions are still possible in particular situations (hidden terminal and exposed terminal problem).

Finally, we consider a different solution that does not avoid multiple simultaneous transmissions in the same area as the previous strategies, but, instead, favors multiple transmissions, by improving the receiver capabilities with advanced Physical (PHY) layer architectures such as multiuser detection and MIMO systems. Efforts in this direction have led to the concept of interference cancellation: when interfering signals are present, it may be possible to decode them, and successively subtract their interference contribution from the received signal. In addition, with multiple antennas, nodes can handle interference even more easily, thus letting multiple links coexist within the same area.

The first part of the Thesis (Chapter 2) is focused on the latter strategy. The protocol considered for channel access management make it possible to have a balance between throughput and interference rejection by encouraging multiple access, while regulating simultaneous transmissions. At the receiver side, the presence of several simultaneous signals is managed by using a Vertical-Bell Laboratories Layered Space-Time (V-BLAST) receiving scheme that decouples the superimposed signal to extract data through successive cancellations. However, this scheme relies on channel knowledge at the receiver, and poor quality of channel estimates can critically impair signal reception. In multiuser scenarios, these problems may lead to interference and, hence, substantial loss of data. In this light, we propose an analytical technique that evaluates the statistics of the channel estimation errors in this scenario where the presence of several simultaneous, symbol-asynchronous signals makes the problem more complicated than in traditional channel estimation. We present this analysis of channel estimation errors using both correlator-based and Minimum Mean-Square Error (MMSE) channel estimators. We show that there is direct dependence of the channel estimation error on the instantaneous channel matrix, that also includes interfering transmitters. Then, with this approach, we directly evaluate the effect of multiuser interference in the expression of the channel estimation errors, taking into account all the unwanted transmissions that degrade the performance of the receiver. The model proposed makes it

possible to quickly evaluate the performance of channel estimation schemes as a function of the system parameters. In this light, we include the effect of channel estimation errors in an ad hoc networking protocol simulator and thoroughly evaluate their impact.

1.2 Cognitive Radio Networks

The growth of the number of wireless systems and services has caused a reduction of the wireless spectrum availability. This is confirmed by the frequency allocation chart provided by the National Telecommunications and Information Administration [2], which shows that almost all frequency bands have been assigned and there is only very little bandwidth available, that can be used only with emerging wireless products and services. The idea of cognitive radio meets this spectrum shortage. With this new paradigm, new wireless users (*secondary* users) can operate in the existing crowded spectrum without degrading the performance of entrenched users (*primary* users), by using devices with advanced radio and signal-processing technology along with novel spectrum-allocation policies. On the other hand, cognitive radio technology requires to collect and process information related to the presence of other users within the spectrum, which in turn requires advanced sensing and signal-processing capabilities. Moreover, another constraint is the requirement for significant changes in the way wireless spectrum is currently allocated to enable cognitive techniques.

In the previous section, we pointed out the main limitations of a wireless network, including interference. In fact, the presence of other users in the network has always been regarded as a problem so far. In a completely different approach, however, it can be seen as a way to improve the performance in wireless networks: the idle nodes that lie in the transmission range of another node can receive the transmitted signal and help the transmitter to forward it. With this form of cooperation, nodes that are not currently transmitting for their own purposes can offer their own resources to improve the performance of their transmitting neighbors. In doing so, they are consuming part of their energy to help other nodes, but an overall gain is likely to be achieved, since transmissions can be completed in a shorter time, and interference may be lowered as well.

In the second part of the Thesis (Chapter 3) we exploit the cognitive radio and cooperation principles to propose a new paradigm that allows the secondary users to transmit in the

wireless medium and introduces the concept of reward in the cooperation. Specifically, the secondary nodes may serve as extra relays, and hence as potential next hops for the primary users, but only in exchange for spectrum leasing. Namely, in return for their forwarding of primary packets, secondary nodes are awarded spectral resources for transmission of their own traffic. Secondary nodes enforce minimal Quality-of-Service (QoS) requirements in terms of rate and reliability when deciding whether to cooperate. With this approach, the presence of secondary users is exploited and regulated by the primary users, avoiding interference. Moreover, with respect to classical cognitive radio techniques, the spectrum leasing approach does not require significant changes in the wireless spectrum allocation (i.e., the resource is still managed by the primary users). This idea of "spectrum leasing via cooperation", combined with the principle of opportunistic routing, is the basis of our framework, which is studied in different types of wireless networks and with different PHY techniques, showing the advantages of the proposed solution.

1.3 Discussion and Organization of this Thesis

This Thesis addresses some interference-related issues in wireless networks, and presents two different studies regarding *(i)* the evaluation of the interference in ad hoc networks by considering the channel estimation errors and *(ii)* the proposal of a new cooperative paradigm in cognitive radio network for the coexistence between primary and secondary users.

In Chapter 2 we start the evaluation of the performance of a Medium Access Control (MAC) protocol for MIMO ad hoc networks under imperfect channel estimation, by presenting in Section 2.2 an analysis of channel estimation errors using both correlator-based and MMSE channel estimators. Unlike similar works, we specifically focus on a scenario where the presence of several simultaneous, symbol-asynchronous signals makes the problem more complicated than in traditional channel estimation. In particular, we show that there is direct dependence of the channel estimation error on the instantaneous channel matrix. In Section 2.2.2.1, for the correlator-based channel estimator only, we also extend the analysis by specifically accounting for transmit impulses of different shape and possibly infinite duration. Moreover, in Section 2.3 we integrate such formulas into a network simulator, in order to obtain a precise representation of PHY-level processes, as opposed to summarizing PHY layer performance into some compact expression. This aspect permits to

directly evaluate the interference in a complex simulator, without making any assumptions to make the model tractable. This is done obtaining precise expressions for the statistics of the estimated channel matrix to be used in analytical studies as well as realistic simulation experiments. Focusing on a cross layer MAC/PHY protocol we analyze its performance by means of simulation, where we employ the analysis outlined above to provide an accurate PHY model. We also highlight the tradeoffs that arise when tuning protocol behaviors and their interplay with PHY-level parameters, such as the length of the training sequences, the number of antennas and the type of transmit impulses. This constitutes a further step toward a more realistic model for PHY-level issues in distributed MIMO networks, so that higher-level protocols can be evaluated based on the resulting considerations.

In Chapter 3 we propose and evaluate an alternative approach to classic cognitive radio approach, based on a combination of the principles of opportunistic routing and of the spectrum leasing via cooperation framework. In order to evaluate the performance of this approach in cognitive radio networks, in Section 3.1 we first study the opportunistic routing technique. In particular, we study the throughput advantages of opportunistic routing over conventional multihop routing for linear multihop wireless networks with Type-I Hybrid Automatic Repeat reQuest (HARQ) and block-fading Rayleigh channel model. In detail, in Section 3.1.2.1 we derive the end-to-end throughput of opportunistic routing using Markov chain tools and accounting for fading statistics, considering both fixed-rate and optimal-rate transmissions. Furthermore, in Appendix A, we show analytical and non-recursive expressions of the end-to-end throughput for the opportunistic routing technique when the number of hops in the network is greater than two. Moreover, in Section 3.1.3 we present a deep throughput analysis using standard information-theoretic performance metrics for asymptotic signal-to-noise ratio (SNR) regimes, considering the multiplexing gain and energy efficiency (i.e., minimum energy per bit) of both opportunistic and multihop routing. Finally, in Section 3.1.4 we give some numerical results to corroborate the analysis.

In Section 3.2, we relate the cooperation principle, the spectrum leasing paradigm and the opportunistic routing technique. In particular, in Section 3.2.1, we propose a spectrum leasing mechanism for the coexistence between a primary and a secondary network that is based on cooperation and opportunistic routing. The primary network consists of a source and a destination communicating via a number of primary relay nodes. In each transmis-

sion slot, the next hop is selected in an on-line fashion based on the decoding outcomes in the previous transmissions according to the idea of opportunistic routing. The secondary nodes may serve as potential next hops for the primary network, but only in exchange for leasing of spectral resources so as to satisfy secondary QoS constraints. Four policies that exploit spectrum leasing via opportunistic routing to different degrees are proposed in Section 3.2.2. These policies are designed to span different operating points in the trade-off between gains in throughput and overall energy expenditure for the primary network. Analysis is carried out for networks with a linear geometry and quasi-static Rayleigh fading statistics by using Markov chain tools. In particular in Appendix B the transition probabilities of these policies are detailed. Furthermore, two physical layer techniques are considered for multiplexing of the primary and secondary traffic at the secondary nodes, namely time division (Section 3.2.1.4) and superposition coding (Section 3.2.1.5). The optimality in terms of both throughput and primary energy consumption of superposition coding over all possible multiplexing strategies, for the given routing techniques, is proved in Section 3.2.2.6. In Section 3.2.3 numerical results demonstrate the advantages of the proposed spectrum leasing solution based on opportunistic routing and illustrate the trade-offs between primary throughput and energy consumption. Finally, the analysis of this chapter is extended in Appendix C to a distributed network, showing the effectiveness of the proposed spectrum leasing via cooperative opportunistic routing technique also in a more general network.

Chapter 2

MIMO Ad Hoc Networks

As introduced in Section 1.2, one of the main goals of the ad hoc network research community has been to design effective and distributed protocols that yield a good throughput, taking into account practical limitations, such as decentralized channel access management, channel estimation errors and interference, that could decrease the network performance.

Network performance can be increased by considering the integration of multiple antennas in each terminal of the network. From a practical point of view, the adoption of multiple antennas can be feasible, due to higher communication frequencies used to communicate. This technique, which can be seen as a MIMO system between transmitter and receiver, has become very attractive after Foschini in [3] showed that the use of multiple antennas at both the receiver and the transmitter makes it possible to significantly increase the link capacity, effectively making use of multiple parallel radio channels in the same band, by separation in space. In fact, MIMO technology permits to achieve high spectral efficiency in Rayleigh fading environments and to enable the protection of communications in the space domain, by processing and transmitting signals through different antennas, according to predefined schemes such as Space-Time Codes (STC) (see, e.g., [4]). In the Layered STC (LSTC), a subset of STC, the independent coded streams are distributed throughout the transmission resource array in so-called layers. Then, the objective is to design the layering architecture and associated signal processing so that the receiver can efficiently separate the individual layers from one another and can decode each of the layers effectively. A special case of LSTC is V-BLAST [5], where the encoding component is absent and all resources are used for parallelizing transmissions (i.e., the vector encoding process is simply a demulti-

plex operation followed by independent bit-to-symbol mapping of each substream). This approach is also called Spatial Multiplexing (SM). It has been shown in [6] that there exists a tradeoff between diversity and SM gain in MIMO networks: V-BLAST achieves the greatest SM depth, whereas codes such as [7] are optimal in a diversity sense.

Ad hoc networks leverage MIMO techniques transmitting independent and separately encoded data signals (so-called *streams*) from each of the multiple transmit antennas. Specifically, if multiple streams are sent by different nodes, each using multiple antennas, all streams can be taken as a separate contribution by the intended receiver. If channel information is available at the receiver, the output of its antennas can be recombined and processed such that the data can finally be recovered. The primary consequence is the coexistence, without collisions, of multiple data packets in the network provided that some degree of coordination is maintained among transmitters. Moreover, by splitting a single packet transmission among multiple antennas (e.g., with V-BLAST), a node is allowed a higher bit rate, which is proportional to the number of antennas used [8, 9]. However, using a more powerful PHY layer in combination with existing MAC protocols for ad hoc networks (such as 802.11 [1]) may not necessarily be the best choice: a better design paradigm should jointly account for PHY and MAC features in a cross-layer fashion, by allowing some exchange of information between different layers. The need to regulate multiple access in MIMO ad hoc networks has led to the definition of protocols that balance between throughput and protection from interference, by encouraging multiple access while limiting simultaneous transmissions. The protocol we consider here is based on a framed channel access as detailed in [10], and will be summarized in the next Section 2.1. This protocol will be studied using the analysis presented in Section 2.2, that permits to include the effect of the interference due to multiuser communications in the expression of the channel estimation errors.[1]

2.1 Network Protocol Description

As introduced above, the reference scenario considered in this chapter is an ad hoc network formed of V-BLAST-capable nodes with multiple antennas. V-BLAST [5] is used here both to improve the communication performance (by increasing the bit rate over a link)

[1] The material presented in this chapter has been published in [11–14].

and to offer a greater level of protection against interference (through the cancellation of unwanted signals), both at the price of more complex signal processing. Since V-BLAST is based on SIC, it is potentially very prone to interference: if the detection of the first among a group of signals fails, its cancellation would increase the amount of interference affecting subsequent detections. Furthermore, V-BLAST operations rely on channel knowledge at the receiver, and poor quality of channel estimates can critically impair signal reception. In multiuser scenarios these problems may lead to substantial loss of data.

In order to regulate multiple access in MIMO ad hoc networks, we defined a protocol that balances between throughput and interference rejection, by encouraging multiple access while limiting simultaneous transmissions. This protocol is based on a framed channel access as detailed in [10], and will be summarized here for the reader's convenience. Each frame is divided into four phases, namely Request-To-Send (RTS), Clear-To-Send (CTS), Data and Acknowledgment (ACK). All packets are assumed to be split into fixed-length Protocol Data Units (PDUs), that are suitable to be sent in a frame using a single antenna. All control packets (RTS, CTS and ACK) are sent using one antenna. We assume an overall per-node power constraint, so that the transmit power is inversely proportional to the number of antennas used. The framed four-way handshake described in the following is then used to understand how many PDUs to send in parallel, and to which receivers.

The RTS-CTS handshake is used to reach an agreement on the transmissions to be performed, and is described as follows. Before composing RTSs, nodes must take into account the communication capabilities of their intended receivers. These capabilities are summarized in the concept of *class*. For what follows, suffice it to say that the class of a neighbor represents the maximum number of antennas that can be used when transmitting toward a group of nodes including that neighbor. This enforces a sufficiently low error rate [10]. The class is a restrictive constraint, and must be satisfied for each and every receiver. Each transmitter may send at most as many PDUs as the *minimum class* of all its current intended receivers. For instance, a node could send a total of 8 PDUs to three receivers, each of which has class 8; on the contrary, if two receivers have class 8 and one has class 4, the node is constrained to send at most 4 PDUs overall. This applies regardless of how PDUs are distributed among the receivers. At the end of this process, RTSs are generally made of a number of requests, each containing the receiver identifier and the number of PDUs to be

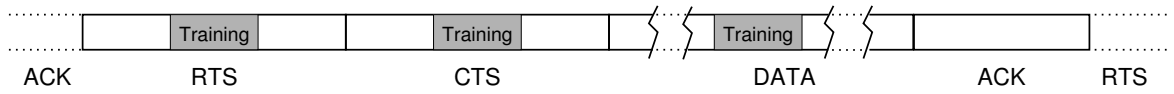

Figure 2.1. *Structure of the transmission frame.*

sent to it. These RTSs have the only purpose to carry transmission requests: hence, they do not block neighboring nodes, unlike in 802.11 [1] (see also the discussion in Section 1.1).

The receivers called upon by the RTSs have now sufficient information to decide whether to allow traffic or not. This choice is driven by a CTS policy, which strikes a balance between throughput and protection against interference, while being constrained by the maximum number of channels which can be tracked to operate the V-BLAST receiver: in other words, only a limited number of signals can be detected and canceled. In the following we have used the Follow Traffic (FT) policy [10]. With FT, each receiver grants at least one request directed to itself. Other requests are considered in order of decreasing received power: if they are directed towards the receiver, it prepares to estimate the corresponding channels and inserts a grant in the CTS; otherwise, the signals are potential interferers, and the receiver prepares to track their channels, in order to cancel them. This procedure goes on until all requests have been processed, or until no more channels can be tracked. CTSs are then transmitted, and PDUs are sent thereafter according to the grants in the CTSs. Finally, ACK messages are issued back to communicate correct receptions on a per PDU basis. Note that FT is a cross-layer policy, as it relies on the exchange of information between the PHY and MAC layers: PHY provides a measure of received power to the MAC, that in turn decides which PDUs to grant, and correspondingly arbitrates signal detection performed by PHY.

The MAC protocol is finally augmented with a standard exponential backoff, entered every time a CTS is not received in response to an RTS: namely, upon a missing CTS, a node refrains from transmissions for a number of frames uniformly drawn at random in the interval $[1, B_{max}]$, where $B_{max} = 2^{N_u - 1}$ and N_u is the number of consecutive unanswered RTSs. This helps limit access persistence and thus reduce congestion. The interested reader is referred to [15] for a thorough evaluation of backoff policies.

Unlike previous studies on the protocol described above [10, 15–17], in this chapter we explicitly account for channel training. Taking a practical standpoint, if channel estimation is performed in the middle of the packet, the coherence of the estimate throughout the re-

ception is improved, at the price of more buffering: therefore, we assume that a training sequence of length L is included within any transmitted packet (except ACKs, which may be transmitted through the same antenna previously used for the CTS). The resulting frame structure is shown in Figure 2.1.

2.2 Analysis of the Channel Estimation Errors

Recent research work and practical implementations (such as [18]) have shown the high potential of such MIMO techniques as spatial multiplexing [3] to achieve high bit rates in wireless communications. At the receiver side, superimposed signals must be decoupled to extract data, e.g., through successive cancellations as in V-BLAST [5]: however, this kind of signal processing requires accurate channel estimation [19, 20], [21, pp. 137–162] to work properly. Compared to point-to-point communications [22], little work has been done on MIMO channel estimation in networking scenarios [16, 23], where signals coming from different transmitters are non-orthogonal and asynchronous, and estimation errors due to multiuser interference are to be taken into account. Similarly, the effects of imperfect channel estimation on network protocols have also been often overlooked.

Our main contribution in this chapter is to provide a thorough analysis of the performance of a MAC protocol for MIMO ad hoc networks under imperfect channel estimation. To this end, we consider two different estimation methods, i.e., matched filtering on a known training sequence and MMSE channel estimation. In specific cases, we derive close-form expressions of their variance and integrate such formulas into a network simulator. This yields a precise representation of PHY-level processes, as opposed to summarizing PHY layer performance into some compact expression, e.g., the probability of symbol error as a function of the Signal to Interference and Noise Ratio (SINR); the latter figure, in particular, would require to make strong assumptions on interference models, potentially leading to an unsuitable representation of realistic network behaviors. As an example of application, an important part of this chapter consists of precise expressions for the statistics of the estimated channel matrix to be used in analytical studies as well as realistic simulation experiments. We focus on the protocol in [10] (described in Section 2.1) and thoroughly analyze its performance by means of simulation, where we employ the analysis outlined above to provide an accurate PHY model. We also highlight the tradeoffs that arise when tuning protocol

behaviors and their interplay with PHY-level parameters, such as the length of the training sequences, the shape of transmit impulses and the number of antennas. We consider the general case of a MIMO ad hoc network, where nodes employ V-BLAST [5], and are thus likely to receive multiple superimposed packets, typically sent through multiple antennas by different users. As these users are generally located at different distances, their packets bear different average power, and cannot be assumed to be symbol-synchronous. In more detail, only the signals coming from the antennas of the same node can be assumed to have the same average power and to be symbol-synchronous (path length differences between pairs of antennas are negligible compared to the symbol duration). These effects have an impact on channel estimation performance and must be taken into account. A direct consequence is that the channel estimation error can be shown to explicitly depend on the channel matrix as well, and not only on noise power and training sequence length.

A number of studies are available that assess the effects of channel estimation errors on the overall performance of single-antenna as well as MIMO systems. In the presence of a time-varying fading channel, a typical approach is to assume that the channel coefficients are perturbed by an additive random component, which accounts for channel estimation errors, such as an additive white Gaussian noise vector (AWGN) with known covariance. This approach is followed in [22], where the authors derive upper and lower bounds on the mutual information for Single-Input-Single-Output (SISO) channels. The same approach can be found in [24] for MIMO systems operating over an independent and identically distributed (iid) Rayleigh fading channel, and lower and upper bounds for mutual information under channel estimation errors are derived.

Comparatively less work can be found for the case of general networking scenarios, as it is difficult to assess the global performance of a whole network with respect to channel estimation errors. The typical model for errors affecting channel estimates is a noisy component that perturb estimated channel coefficients. In [25], a cross-layer distributed power control and scheduling protocol for delay-constrained applications over mobile Code-Division Multiple Access (CDMA)-based ad hoc wireless networks is proposed. The authors focus on a SISO channel, whose estimates are perturbed by adding a complex Gaussian component with known variance at the receiver. In [23], an analysis of multiuser scheduling in MIMO systems with imperfect channel estimation is presented using the same AWGN estimation

error model. Network performance is assessed as a function of the number of users and noise power: however, all signals are considered synchronous here, which is not realistic in distributed networks. In [26] the authors consider asynchronous transmissions in the context of cooperating base stations in a cellular system. Cooperation is implemented by linear precoding, assuming perfect channel knowledge at the transmitter. Base stations compensate for propagation delays during their cooperation phase, but can manage only a limited number of transmitters, and those signals whose delay cannot be compensated will cause interference.

In the specific case of V-BLAST, most papers focusing on channel estimation errors only treat point-to-point scenarios, and consider only noise power as an impairment to estimation. For instance, this is the case in [27], where Symbol Error Rate (SER) expressions are derived by taking into account error propagation and suboptimal substream ordering, before proposing an optimal ordering that reduces error propagation across substreams. In fact, the main problem affecting V-BLAST is the propagation of errors between subsequent detection steps, which limits the performance of the SIC procedure. In order to reduce this effect, an optimal substream ordering criterion to be used along with a Zero Forcing (ZF) or MMSE detector is presented. Similarly, in [28] channel estimation errors are modeled by adding an uncorrelated component with known covariance to the current MIMO channel realization. Unlike in [27], where the focus is on error propagation, in [28] the effects of optimal ordering are considered and accurate bit error rate (BER) expressions for M-ary Phase-Shift Keying (MPSK) and M-ary Quadrature Amplitude Modulation (MQAM) modulation at each stage of the cancellation procedure for a 2×2 MIMO link are derived. A drawback of this analysis is that it considers only two transmit and two receive antennas in a point-to-point scenario: therefore, the effects of error propagation can be ignored, because the number of substreams to be ordered is limited to two. In order to reduce the effects of error propagation, new methods for sorting and decoding the received streams are presented in [29]. In addition, a new ordering criterion is presented, whereby at each iteration the substream bearing the lowest mean square error under channel errors is chosen. Another study of a point-to-point V-BLAST system is presented in [30], where the estimation of MIMO channels using orthogonal training sequences is subject to errors due to noise power and channel variations in time. The analysis does not consider the effect of these errors on V-BLAST system, but they

only optimize the training length as functions of the Doppler shift intensity and the number of antennas, for a given BER.

Summarizing, in this chapter we consider a general networking scenario with symbol-asynchronous transmissions; in this case, we derive expressions for correlator-based (in Section 2.2.2) and MMSE (in Section 2.2.3) channel estimators, and we evaluate in Section 2.3 the performance of a MAC protocol specifically designed for a V-BLAST-capable network, described in Section 2.1. The V-BLAST architecture allows to achieve a significant portion of the theoretical capacity with reasonable implementation complexity [27], but is prone to errors if the channel state information (CSI) is not accurate: in this light, we also evaluate the tradeoffs between PHY-layer parameters (which affect estimation accuracy) and MAC protocol performance.

2.2.1 System Model

Assume that each node has N_A antennas. All antennas are used during reception, whereas in general a node i may use $n_i \leq N_A$ antennas for transmission. We focus on a single receiving node, which hears signals from all transmitting antennas at the various nodes $i = 1, \ldots, M$. The channel matrix (of size $N_A \times U$, where $U = \sum_{i=1}^{M} n_i$) contains the channel coefficients between all transmitting antennas at all nodes and the N_A antennas at the receiver. We assume flat fading, with iid realizations across the different antenna pairs, that is, the channel between antenna j of transmitting node i and antenna ℓ at the receiver is a complex scalar $h_{j\ell}^{(i)}$, with zero-mean complex Gaussian statistics and variance σ_i^2 (the same for all antennas j, ℓ since it only depends on the distance between the transmitting and receiving nodes). We remark that σ_i^2 includes both transmit power and path loss: thus, the case of equal receive power for all incoming signals usually considered in the literature can be included in the notation above by taking σ_i^2 equal for all is.

For channel estimation purposes, antenna j of user i sends a training sequence $s_{ij}(t)$ of L real binary symbols $b_{ij}[p] \in \{-1, +1\}$, $p = 0, \ldots, L-1$, namely

$$s_{ij}(t) = \sum_{p=0}^{L-1} b_{ij}[p] g(t - pT), \tag{2.1}$$

where $P[-1] = P[1] = 0.5$ and $g(t)$ is any impulse, including impulses of infinite duration. The random ± 1 training sequence does not minimize the variance of the estimation error in

the presence of interferers using the same transmit signal pulses as the sender of the training sequence: optimal sequences can be designed by solving a minimization problem over the Mean-Square Error (MSE) of the estimation [31–33]. Results show that if all signals (wanted and interfering) are symbol-synchronous, there is no gain from the use of optimal sequences; on the contrary, the gain is maximized if all *interferers* are symbol-synchronous and bear a delay of half a symbol period with respect to the wanted training sequence. In a more general scenario, the MSE is lower by $7\,\mathrm{dB}$ than achieved by using a standard Hadamard sequence. In any event, the optimal sequences can be obtained only by gathering feedback from the receiver; moreover, optimal sequences would turn only into a lower variance of channel estimates, making the network performance closer to that of the perfect CSI case, but would not alter the essence of the PHY-MAC tradeoffs discussed later. Therefore, in the following we still focus on a random ± 1 training sequence which, albeit sub-optimal, allows to keep the analysis simple.

Knowing that in the analysis we assume that the power of sequences $s_{ij}(t)$ is 1, as the actual received power is included in σ_i^2 (this incorporates the per antenna transmit power of node i, and is therefore inversely proportional to n_i), the signal received at antenna ℓ is given by

$$\begin{aligned} r_\ell(t) &= \sum_{i=1}^{M} \sum_{j=1}^{n_i} h_{j\ell}^{(i)} s_{ij}(t - \tau_i) + z_\ell(t) \\ &= \sum_{i=1}^{M} \sum_{j=1}^{n_i} h_{j\ell}^{(i)} \sum_{p=0}^{L-1} b_{ij}[p] g(t - pT - \tau_i) + z_\ell(t) \end{aligned} \tag{2.2}$$

where $z_\ell(t)$ is the thermal noise at the receiver, modeled as white circular complex Gaussian with zero mean and power spectral density $N_0/2$ (per dimension), and τ_i is the propagation delay of the signals transmitted by the antennas at user i (note that this propagation delay does not depend on j, ℓ but only on i, as the path length differences between pairs of antennas of a given transmitter/receiver pair are negligible compared to the symbol duration T). With no loss of generality, we focus on the estimation of the channels related to a specific user m, and assume $\tau_m = 0$. In the following we will consider the correlator- and MMSE-based channel estimators.

2.2.2 Correlator-based channel estimation

In this case, the channel coefficient between antenna k of user m and antenna $\ell = 1, \ldots, N_A$ of the receiver is estimated by matched filtering on the training sequence $s_{mk}(t)$ (see (2.1)), $k = 1, \ldots, n_m$:

$$\begin{aligned}
\hat{h}_{k\ell}^{(m)} &= \frac{1}{LT}\int_{-\infty}^{+\infty} r_\ell(t) s_{mk}(t)\,\mathrm{d}t \\
&= \frac{1}{LT}\int_{-\infty}^{+\infty} \sum_{j=1}^{n_m} h_{j\ell}^{(m)} s_{mj}(t) s_{mk}(t)\,\mathrm{d}t + \frac{1}{LT}\int_{-\infty}^{+\infty} \sum_{i\neq m}\sum_{j=1}^{n_i} h_{j\ell}^{(i)} s_{ij}(t-\tau_i) s_{mk}(t)\,\mathrm{d}t \\
&+ \frac{1}{LT}\int_{-\infty}^{+\infty} z_\ell(t) s_{mk}(t)\,\mathrm{d}t \qquad (2.3) \\
&= \frac{1}{LT}\int_{-\infty}^{+\infty} \sum_{j=1}^{n_m} h_{j\ell}^{(m)} \left\{\sum_{p=0}^{L-1} b_{mj}[p] g(t-pT) \sum_{q=0}^{L-1} b_{mk}[q] g(t-qT)\right\}\mathrm{d}t + \\
&+ \frac{1}{LT}\int_{-\infty}^{+\infty} \sum_{i\neq m}\sum_{j=1}^{n_i} h_{j\ell}^{(i)} \left\{\sum_{p=0}^{L-1} b_{ij}[p] g(t-pT-\tau_i) \sum_{q=0}^{L-1} b_{mk}[q] g(t-qT)\right\}\mathrm{d}t + \\
&+ \frac{1}{LT}\int_{-\infty}^{+\infty} z_\ell(t) \sum_{q=0}^{L-1} b_{mk}[q] g(t-qT)\,\mathrm{d}t. \qquad (2.4)
\end{aligned}$$

By calling $\phi_g(\tau)$ the auto-correlation function of $g(t)$, we have

$$\begin{aligned}
\hat{h}_{k\ell}^{(m)} &= \frac{1}{LT}\sum_{j=1}^{n_m} h_{j\ell}^{(m)} \sum_{p=0}^{L-1} b_{mj}[p] \sum_{q=0}^{L-1} b_{mk}[q] \int_{-\infty}^{+\infty} g(t-pT) g(t-qT)\,\mathrm{d}t + \\
&+ \frac{1}{LT}\sum_{i\neq m}\sum_{j=1}^{n_i} h_{j\ell}^{(i)} \sum_{p=0}^{L-1} b_{ij}[p] \sum_{q=0}^{L-1} b_{mk}[q] \int_{-\infty}^{+\infty} g(t-pT-\tau_i) g(t-qT)\,\mathrm{d}t + Z_\ell \qquad (2.5) \\
&= h_{k\ell}^{(m)} \frac{\phi_g(0)}{T} + \frac{1}{LT}\sum_{i\neq m}\sum_{j=1}^{n_i} h_{j\ell}^{(i)} \sum_{p=0}^{L-1} b_{ij}[p] \sum_{q=0}^{L-1} b_{mk}[q] \phi_g\left((p-q)T+\tau_i\right) + Z_\ell \qquad (2.6)
\end{aligned}$$

where we note that the first rhs term is zero whenever $p \neq q$ or $j \neq k$, and

$$Z_\ell = \frac{1}{LT}\int_{-\infty}^{+\infty} z_\ell(t) s_{mk}(t)\,\mathrm{d}t \qquad (2.7)$$

is the filtered noise term. For this term we have $\mathbb{E}[Z_\ell] = 0$ and

$$\begin{aligned}
\sigma_Z^2 =& \mathbb{E}[|Z_\ell|^2] = \frac{1}{(LT)^2}\int_{-\infty}^{+\infty} N_0 s_{mk}^2(t)\,\mathrm{d}t \\
=& \frac{N_0}{LT^2}\phi_g(0), \qquad (2.8)
\end{aligned}$$

so that Z_ℓ has circular complex Gaussian distribution with zero mean and variance $\frac{N_0}{LT^2}\phi_g(0)$.

In (2.3) through (2.6), the system model accounts only for the signals being sent at the same time as the training sequence used for estimation. Now, recall that we are considering frame-based channel access: the model in (2.3)–(2.6) applies therefore to one frame at a time, and would thus require to consider co-frame interference. However, as per the discussion at the end of Section 2.1, we note that it is convenient to place training sequences *within* transmitted packets, so that channel estimates are more likely to be coherent throughout the packet duration. This separates any training sequence from other frames by a significant number of symbols for any practical data and signaling packet length. Therefore, co-frame interference can be neglected, even in the case of long transmit impulses. By way of contrast, every other non-channel estimation-related symbol may create interference: however, if training sequences conform to pseudo-randomness criteria, interference from data is statistically equivalent to interference from other training sequences. This allows to approximate interference from data symbols by letting p go from $-\infty$ to $+\infty$ in the second rhs term in (2.5), as we will do from now on.

From (2.6) we define $J_{jk}^{(i,m)}$ as the matched filter output corresponding to antenna j of user i:

$$\begin{aligned} J_{jk}^{(i,m)} &= \frac{1}{LT}\int_{-\infty}^{+\infty} s_{ij}(t-\tau_i)s_{mk}(t)\,\mathrm{d}t \\ &= \frac{1}{LT}\sum_{p=-\infty}^{+\infty} b_{ij}[p]\sum_{q=0}^{L-1} b_{mk}[q]\phi_g((p-q)T+\tau_i). \end{aligned} \tag{2.9}$$

This term has zero mean, and its variance given τ_i can be found as

$$\begin{aligned} \mathrm{Var}[J_{jk}^{(i,m)}|\tau_i] = \frac{1}{(LT)^2}\mathbb{E}\left[\left(\sum_{q_1=0}^{L-1} b_{mk}[q_1]\sum_{p_1=-\infty}^{+\infty} b_{ij}[p_1]\phi_g((p_1-q_1)T+\tau_i)\right)\times\right. \\ \left.\times\left(\sum_{q_2=0}^{L-1} b_{mk}[q_2]\sum_{p_2=-\infty}^{+\infty} b_{ij}[p_2]\phi_g((p_2-q_2)T+\tau_i)\right)\right]. \end{aligned} \tag{2.10}$$

We note that the terms in the sums are non-zero only if $q_1 = q_2 = q$ and $p_1 = p_2 = p$; since $\mathbb{E}[b_{mk}^2] = \mathbb{E}[b_{ij}^2] = 1$, we have

$$\begin{aligned} \mathrm{Var}[J_{jk}^{(i,m)}|\tau_i] &= \frac{1}{(LT)^2}\sum_{q=0}^{L-1}\sum_{p=-\infty}^{+\infty}\phi_g^2((p-q)T+\tau_i) \\ &= \frac{1}{LT^2}\sum_{a=-\infty}^{+\infty}\phi_g^2(aT+\tau_i), \end{aligned} \tag{2.11}$$

Hence,

$$\mathrm{Var}[J_{jk}^{(i,m)}] = \int \mathrm{Var}[J_{jk}^{(i,m)}|\tau_i]\,\mathrm{d}F(\tau_i), \tag{2.12}$$

where $F(\tau_i)$ is the cumulative distribution function of τ_i.

Consider now the joint statistics of $J_{j_1,k_1}^{(i_1,m)}$ and $J_{j_2,k_2}^{(i_2,m)}$. In particular, we have that

$$\begin{aligned}
&\mathbb{E}\left[J_{j_1,k_1}^{(i_1,m)} J_{j_2,k_2}^{(i_2,m)}\right] = \\
&\quad = \mathbb{E}\left[\frac{1}{(LT)^2}\int_{-\infty}^{+\infty}\mathrm{d}t_1\int_{-\infty}^{+\infty}\mathrm{d}t_2 s_{i_1j_1}(t_1-\tau_{i_1})s_{mk_1}(t_1)s_{i_2j_2}(t_2-\tau_{i_2})s_{mk_2}(t_2)\right] \\
&\quad = \frac{1}{(LT)^2}\int_{-\infty}^{+\infty}\mathrm{d}t_1\int_{-\infty}^{+\infty}\mathrm{d}t_2\mathbb{E}\left[s_{i_1j_1}(t_1-\tau_{i_1})s_{mk_1}(t_1)s_{i_2j_2}(t_2-\tau_{i_2})s_{mk_2}(t_2)\right]
\end{aligned} \tag{2.13}$$

Consider the case $i_1, i_2 \neq m$. (Note that $i \neq m$ is the only case of interest in (2.6), and in any event $J_{jk}^{(i,m)}$ for $i = m$ is deterministic so that the computation is straightforward.) In this case we have

$$\begin{aligned}
&\mathbb{E}\left[s_{i_1j_1}(t_1-\tau_{i_1})s_{mk_1}(t_1)s_{i_2j_2}(t_2-\tau_{i_2})s_{mk_2}(t_2)\right] = \\
&\qquad = \mathbb{E}\left[s_{i_1j_1}(t_1-\tau_{i_1})s_{i_2j_2}(t_2-\tau_{i_2})\right]\mathbb{E}\left[s_{mk_1}(t_1)s_{mk_2}(t_2)\right],
\end{aligned} \tag{2.14}$$

which is clearly zero if $i_1 \neq i_2$, or $j_1 \neq j_2$, or $k_1 \neq k_2$ (modeling the used sequences as independent random binary sequences). Similarly, it is possible to show that $\mathbb{E}\left[J_{j_1,k_1}^{(i_1,m_1)} J_{j_2,k_2}^{(i_2,m_2)}\right] = 0, \quad \forall i_1, i_2, j_1, j_2, k_1, k_2$, if $m_1 \neq m_2$, because in this case $\mathbb{E}\left[s_{m_1k_1}(t_1)s_{m_2k_2}(t_2)\right] = 0$. Therefore, we can conclude that the variables $J_{jk}^{(i,m)}$ can be modeled as a set of iid random variables, each with approximately Gaussian statistics (by virtue of their being sums of several binary independent random variables, see (2.9)) of zero mean and variance $\mathrm{Var}[J_{jk}^{(i,m)}]$.

While the analysis above applies to any impulse $g(t)$, in the following we will consider the special case of rectangular impulses, that is

$$g(t) = \begin{cases} 1, & \text{for } -\frac{T}{2} \leq t < \frac{T}{2}, \\ 0, & \text{otherwise;} \end{cases} \tag{2.15}$$

A study on the effect of different transmit pulses (i.e., different shape and possibly infinite duration) is derived in Section 2.2.2.1. For the rectangular pulse, the auto-correlation function of $g(t)$ is given by

$$\phi_g(\tau) = \begin{cases} T - |\tau|, & \text{for } -T \leq \tau \leq T, \\ 0, & \text{otherwise;} \end{cases} \tag{2.16}$$

thus, (2.6) can be rewritten as

$$\hat{h}_{k\ell}^{(m)} = h_{k\ell}^{(m)} + \sum_{i \neq m} \sum_{j=1}^{n_i} h_{j\ell}^{(i)} J_{jk}^{(i,m)} + Z_\ell. \tag{2.17}$$

Note that, in the synchronous case, all received training sequences can be assumed to be orthogonal, and therefore the matched filter output in (2.9), $J_{jk}^{(i,m)}$, is equal to 1 only if $j = k$, regardless of the transmit node, and 0 otherwise. Under this assumption, the analysis yields the well-known close-form expression (e.g., see [34]):

$$\hat{h}_{k\ell}^{(m)} = h_{k\ell}^{(m)} + Z_\ell. \tag{2.18}$$

We are however interested in the asynchronous case, where sequences coming from different users have random time displacement, and therefore cannot be assumed orthogonal, whereas the signals from the antennas of the same user are synchronously received, and thus can be assumed to have zero cross-correlation. In other words, $J_{jk}^{(m,m)} = 1$ for $j = k$, and 0 otherwise. Because $\phi_g(\tau)$ is non-zero only in $(-T, T)$ if $g(t)$ is rectangular, the only terms of interest in (2.11) are those for $p - q = 0$ or -1, which yields

$$\mathrm{Var}[J_{jk}^{(i,m)} | \tau_i] = \frac{1}{(LT)^2} \sum_{p=0}^{L-1} [\tau_i^2 + (T - \tau_i)^2] \tag{2.19}$$

and

$$\begin{aligned} \sigma_J^2 = \mathrm{Var}[J_{jk}^{(i,m)}] &= \frac{1}{(LT)^2} 2L\mathbb{E}[\tau_i^2] \\ &= \frac{2}{3L}, \end{aligned} \tag{2.20}$$

where the distribution of τ_i is assumed to be uniform in $[0, T]$ $\forall i$ (or τ_i modulo T is uniform in $[0, T]$ $\forall i$), consistently with the approximation that there are infinite symbols per each interfering sequence.

Finally, the output of the matched filter in the asynchronous case is

$$\hat{h}_{k\ell}^{(m)} = h_{k\ell}^{(m)} + \Delta h_{k\ell}^{(m)} \tag{2.21}$$

where

$$\Delta h_{k\ell}^{(m)} = \sum_{i \neq m} \sum_{j=1}^{n_i} h_{j\ell}^{(i)} J_{jk}^{(i,m)} + Z_\ell \tag{2.22}$$

is the channel estimation error, and Z_ℓ is complex Gaussian distributed with zero mean and variance equal to

$$\sigma_Z^2 = \frac{N_0}{LT^2}\phi_g(0) = \frac{N_0}{LT} \tag{2.23}$$

for rectangular pulses.

It is useful to summarize the above calculations in matrix form. Let $\mathbf{H}^{(i)}$ be the channel matrix between user $i = 1, \ldots, M$ and the receiver, defined as

$$\mathbf{H}^{(i)} = \begin{bmatrix} h_{11}^{(i)} & \ldots & h_{n_i 1}^{(i)} \\ h_{12}^{(i)} & \ldots & h_{n_i 2}^{(i)} \\ \vdots & & \vdots \\ h_{1\ell}^{(i)} & \ldots & h_{n_i \ell}^{(i)} \\ \vdots & & \vdots \\ h_{1N_A}^{(i)} & \ldots & h_{n_i N_A}^{(i)} \end{bmatrix} = \begin{bmatrix} \underline{h}_1^{(i)} \\ \underline{h}_2^{(i)} \\ \vdots \\ \underline{h}_\ell^{(i)} \\ \vdots \\ \underline{h}_{N_A}^{(i)} \end{bmatrix}, \tag{2.24}$$

where we highlight row vectors through the underline notation $\underline{x}$, in order to differentiate from them column vectors, which are indicated using the boldface notation $\mathbf{x}$. Let the overall channel matrix from all transmitting antennas to the antennas of the receiver be

$$\begin{aligned} \mathbf{H} &= \left[\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \ldots, \mathbf{H}^{(M)}\right] \\ &= \begin{bmatrix} \underline{h}_1^{(1)} & \underline{h}_1^{(2)} & \ldots & \underline{h}_1^{(M)} \\ \vdots & & & \vdots \\ \underline{h}_{N_A}^{(1)} & \underline{h}_{N_A}^{(2)} & \ldots & \underline{h}_{N_A}^{(M)} \end{bmatrix} = \begin{bmatrix} \underline{h}_1 \\ \vdots \\ \underline{h}_{N_A} \end{bmatrix}. \end{aligned} \tag{2.25}$$

Define now

$$\widetilde{J}_{jk}^{(i,m)} = \begin{cases} J_{jk}^{(i,m)}, & i \neq m, \\ 0, & \text{otherwise.} \end{cases} \tag{2.26}$$

Let $\widetilde{\mathbf{J}}_k^{(m)}$ be the column vector collecting all variables $\widetilde{J}_{jk}^{(i,m)}$ for all antennas of the transmitting users:

$$\widetilde{\mathbf{J}}_k^{(m)} = \left[\widetilde{J}_{1k}^{(1,m)} \ldots \widetilde{J}_{n_1 k}^{(1,m)}\ \widetilde{J}_{1k}^{(2,m)} \ldots \widetilde{J}_{n_2 k}^{(2,m)} \quad \ldots \quad \widetilde{J}_{1k}^{(M,m)} \ldots \widetilde{J}_{n_M k}^{(M,m)}\right]^T, \tag{2.27}$$

and define the matrix

$$\widetilde{\mathbf{J}}^{(m)} = \left[\widetilde{\mathbf{J}}_1^{(m)}, \ldots, \widetilde{\mathbf{J}}_{n_m}^{(m)}\right]. \tag{2.28}$$

With this notation, and defining the column vectors

$$\hat{\mathbf{h}}_k^{(m)} = \left[\hat{h}_{k1}^{(m)}, \ldots, \hat{h}_{kN_A}^{(m)}\right]^T, \quad \mathbf{h}_k^{(m)} = \left[h_{k1}^{(m)}, \ldots, h_{kN_A}^{(m)}\right]^T \tag{2.29}$$

and

$$\mathbf{Z} = [Z_1, \ldots, Z_{N_A}]^T \tag{2.30}$$

for $k = 1, \ldots, n_m$ and $m = 1, \ldots, M$, we can rewrite (2.21) and (2.22) in vector form as

$$\hat{\mathbf{h}}_k^{(m)} = \mathbf{h}_k^{(m)} + \mathbf{H}\widetilde{\mathbf{J}}_k^{(m)} + \mathbf{Z}, \tag{2.31}$$

which shows that the channel estimation error vector $\mathbf{H}\widetilde{\mathbf{J}}_k^{(m)} + \mathbf{Z}$ can be generated by multiplying the channel matrix $\mathbf{H}$ by the training sequence correlation vector $\widetilde{\mathbf{J}}_k^{(m)}$ (whose non-zero elements can be drawn randomly and independently $\sim \mathcal{N}(0, \frac{2}{3L})$), and adding the noise vector. By grouping in a matrix all the above quantities with subscript $k = 1, \ldots, n_m$, we obtain

$$\hat{\mathbf{H}}^{(m)} = \mathbf{H}^{(m)} + \mathbf{H}\widetilde{\mathbf{J}}^{(m)} + \mathbf{Z}' \tag{2.32}$$

and, by further grouping all these results for all nodes $m = 1, \ldots, M$, we have

$$\hat{\mathbf{H}}_{\mathbf{CORR}} = \mathbf{H} + \mathbf{H}\widetilde{\mathbf{J}} + \mathbf{Z}'' \tag{2.33}$$

where

$$\widetilde{\mathbf{J}} = \left[\widetilde{\mathbf{J}}^{(1)}, \widetilde{\mathbf{J}}^{(2)}, \ldots, \widetilde{\mathbf{J}}^{(M)}\right], \tag{2.34}$$

and $\mathbf{Z}', \mathbf{Z}''$ are matrices of iid elements $\sim \mathcal{N}(0, \frac{N_0}{LT})$.

2.2.2.1 Different Transmit Waveforms

In this section we extend the analysis presented above by specifically accounting for transmit impulses of different shape and possibly infinite duration. This constitutes a further step toward a more realistic model for PHY-level issues in distributed MIMO networks, so that higher-level protocols (e.g., see Section 2.1) can be evaluated based on the resulting considerations. These results will be presented in Section 2.3.3.

We consider several choices for the transmit pulse $g(t)$ (see (2.1)):

- *rectangular* pulse:

$$g(t) = \begin{cases} 1, & \text{for } -\frac{T}{2} \leq t < \frac{T}{2}, \\ 0, & \text{otherwise;} \end{cases} \tag{2.35}$$

- *Sinc* pulse:

$$g(t) = \text{sinc}\left(\frac{t}{T}\right) = \frac{\sin\left(\frac{\pi t}{T}\right)}{\left(\frac{\pi t}{T}\right)}; \tag{2.36}$$

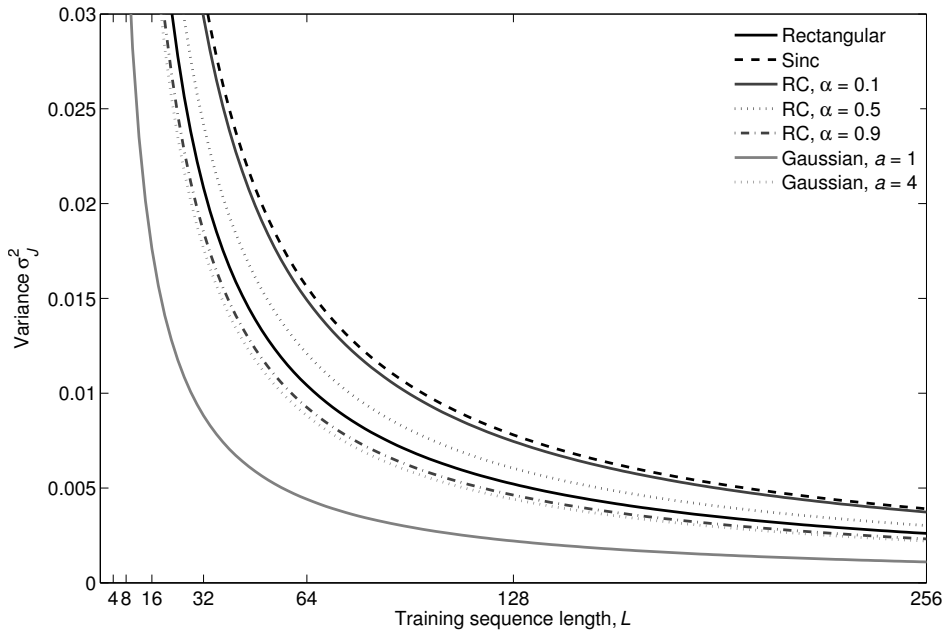

Figure 2.2. *Variance of the non-zero elements in $\widetilde{\mathbf{J}}$, for varying length of the training sequence, L, and different transmit pulses.*

- *raised cosine* (RC) pulse [35] of roll-off parameter α [2]

$$g(t) = \operatorname{sinc}\left(\frac{t}{T}\right) \frac{\cos\left(\frac{\pi\alpha t}{T}\right)}{1 - \frac{4\alpha^2 t^2}{T^2}}; \tag{2.38}$$

- *Gaussian* pulse of variance factor a

$$g(t) = \sqrt{\frac{a}{\pi}} e^{-a\left(\frac{t}{T}\right)^2}. \tag{2.39}$$

For the rectangular pulse the close-form variance formula has been obtained in (2.20): in fact the only terms of interest in (2.11) are those for $a = 0$, where $\phi_g(\tau_i) = T - \tau_i$ and $a = -1$, where $\phi_g(-T + \tau_i) = \tau_i$. Therefore

$$\mathrm{Var}[J_{jk}^{(i,m)} | \tau_i] = \frac{(T - \tau_i)^2 + \tau_i^2}{LT^2}, \tag{2.40}$$

and the variance of $J_{jk}^{(i,m)}$ is $\sigma_J^2 = 2/(3L)$ (see (2.20)).

[2] The raised cosine frequency characteristic is given as

$$G(f) = \begin{cases} T, & 0 \le |f| < \frac{1-\alpha}{2T} \\ \frac{T}{2}\left\{1 + \cos\left[\frac{\pi T}{\alpha}\left(|f| - \frac{1-\alpha}{2T}\right)\right]\right\}, & \frac{1-\alpha}{2T} \le |f| \le \frac{1+\alpha}{2T} \\ 0, & |f| \ge \frac{1+\alpha}{2T}; \end{cases} \tag{2.37}$$

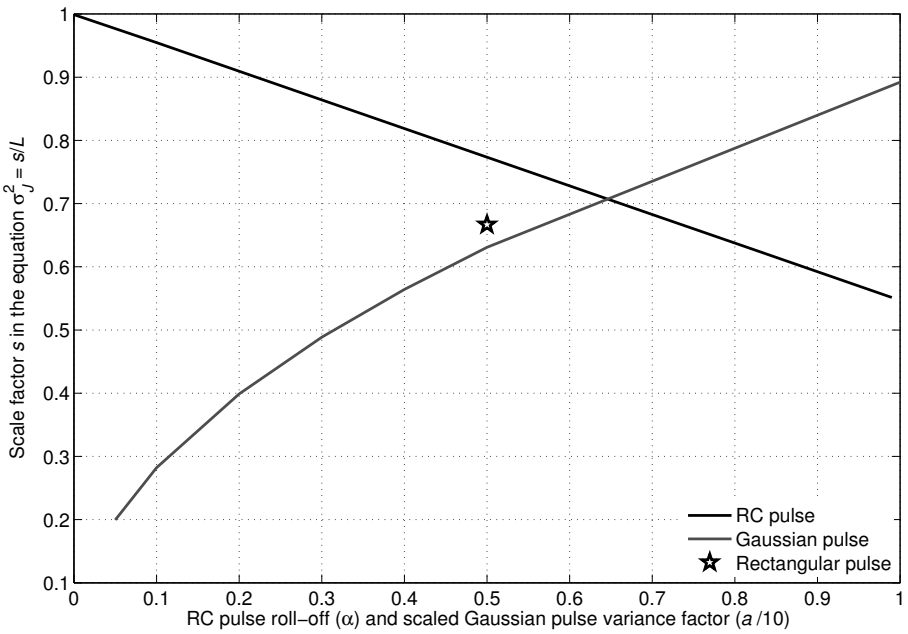

Figure 2.3. *Variation of the scale factor s in the equation $\sigma_J^2 = s/L$ with pulse parameters. For visualization purposes, the abscissa represents both the RC pulse roll-off factor α and a scaled version of the Gaussian pulse variance factor a, i.e., $a/10$. The value of σ_J^2 obtained using rectangular pulses is also shown for reference.*

The variance for the other type of pulses listed before is numerically derived and is shown in Figure 2.2. The figure shows that Gaussian pulses offer the best performance, by virtue of their autocorrelation, which is also Gaussian, and is either low-valued and wide (low a) or high-valued and narrow (high a). RC pulses offer different variance behavior depending on the value chosen for the roll-off factor α. In particular, a low α yields higher variance, as in the case of a Sinc pulse, but when α increases up to 0.9 the RC pulse decays to zero more rapidly, and hence has a narrower autocorrelation function, reaching the same performance as that of the Gaussian pulse with $a = 4$. In other words, a high roll-off factor reduces the variance of $\widetilde{\mathbf{J}}$, due to the fewer dominating terms in (2.10).

Plotting Figure 2.2 in a log-log scale confirms that the relation between σ_J^2 and the length of the training sequence L is linear of constant slope equal to -1 in log scale, and therefore of the type $\sigma_J^2 = s/L$ in linear scale. Figure 2.3 shows the variation of the scale factor s as a function of the parameters of RC and Gaussian pulses in comparison to a rectangular pulse, and confirms preceding comments on Figure 2.2: namely, RC pulses with a high roll-off

factor α or a Gaussian pulse with a low variance factor a yield better performance, mainly due to their narrower autocorrelation function. As a final observation, the choice of a transmit pulse is generally driven by other considerations than its impact on the performance of channel estimation in MIMO ad hoc networks; however, this impact is not negligible, and must be carefully taken into account in simulations: the methodology to calculate σ_J^2 presented so far, as well as the scale factors reported in Figure 2.3 for some specific pulse choices, are of help in this task. In Section 2.3.3 we study this aspect in the MAC protocol for MIMO ad hoc network introduced in Section 2.1.

2.2.2.2 Case Study

The analysis introduced in Section 2.2.2 allows to analytically compute various statistical metrics related to the channel estimation errors. Also, it can be readily used to simulate the effect of channel estimation errors in an ad hoc network, and to study their impact on network performance. As an example, consider a network with three nodes, i.e., two transmitters and one receiver, each equipped with two antennas. The transmitting nodes use both their antennas. The receiver wants to estimate the eight flat fading channel coefficients between each of the transmit antennas and its own receive antennas. The channel matrix can be written as $\mathbf{H} = (h_{kl}^{(m)})$, with $k, l, m = 1, 2$,

$$\mathbf{H} = \begin{pmatrix} h_{11}^{(1)} & h_{21}^{(1)} & h_{11}^{(2)} & h_{21}^{(2)} \\ h_{12}^{(1)} & h_{22}^{(1)} & h_{12}^{(2)} & h_{22}^{(2)} \end{pmatrix} \tag{2.41}$$

where $h_{k\ell}^{(m)}$ is the channel coefficient between the k-th antenna of the m-th transmitter and the ℓ-th antenna of the receiver. According to Eq. (2.33), the estimation error depends on the instantaneous responses of all channels. For a given realization of the channel conditions, we are interested in computing the mean and variance of the estimation error, $\Delta h_{k\ell}^{(m)}$, $m, k, \ell = 1, 2$, which is given by Equation (2.22). We have

$$\mathbb{E}\left[\Delta h_{k\ell}^{(m)}\right] = 0 \qquad \text{and} \qquad \mathbb{E}\left[|\Delta h_{k\ell}^{(m)}|^2\right] = \sum_{i \neq m} \sum_{j=1}^{n_i} |h_{j\ell}^{(i)}|^2 \frac{2}{3L} + \frac{N_0}{LT} \tag{2.42}$$

With a slight abuse of notation, we define the matrices $|\mathbf{H}|^2 = (|h_{kl}^{(m)}|^2)$ and $\mathbb{E}[|\Delta\mathbf{H}|^2] = (\mathbb{E}[|\Delta h_{kl}^{(m)}|^2])$ with $k, l, m = 1, 2$:

$$|\mathbf{H}|^2 = \begin{pmatrix} |h_{11}^{(1)}|^2 & |h_{21}^{(1)}|^2 & |h_{11}^{(2)}|^2 & |h_{21}^{(2)}|^2 \\ |h_{12}^{(1)}|^2 & |h_{22}^{(1)}|^2 & |h_{12}^{(2)}|^2 & |h_{22}^{(2)}|^2 \end{pmatrix} \tag{2.43}$$

$$\mathbb{E}[|\Delta\mathbf{H}|^2] = \begin{pmatrix} \mathbb{E}[|\Delta h_{11}^{(1)}|^2] & \mathbb{E}[|\Delta h_{21}^{(1)}|^2] & \mathbb{E}[|\Delta h_{11}^{(2)}|^2] & \mathbb{E}[|\Delta h_{21}^{(2)}|^2] \\ \mathbb{E}[|\Delta h_{12}^{(1)}|^2] & \mathbb{E}[|\Delta h_{22}^{(1)}|^2] & \mathbb{E}[|\Delta h_{12}^{(2)}|^2] & \mathbb{E}[|\Delta h_{22}^{(2)}|^2] \end{pmatrix} \tag{2.44}$$

The following equations give an example of results obtained for two independent and randomly generated channel realizations. The transmitters are at the same distance from the receiver, noise is neglected, and $L = 16$ (these results can be easily scaled to other values of L, as the variance of the channel estimation is inversely proportional to L in this case). First realization:

$$|\mathbf{H}|^2 = \begin{pmatrix} 6.63\text{e-}2 & 4.35\text{e-}3 & 0.554 & 1.32\text{e-}3 \\ 0.164 & 1.56\text{e-}3 & 2.12 & 0.645 \end{pmatrix}, \tag{2.45}$$

$$\mathbb{E}\left[|\Delta\mathbf{H}|^2\right] = \begin{pmatrix} 2.3\text{e-}2 & 2.3\text{e-}2 & 2.94\text{e-}3 & 2.94\text{e-}3 \\ 0.1152 & 0.1152 & 6.89\text{e-}3 & 6.89\text{e-}3 \end{pmatrix} \tag{2.46}$$

Second realization:

$$|\mathbf{H}|^2 = \begin{pmatrix} 3.04\text{e-}3 & 0.038 & 5.33\text{e-}5 & 1.58\text{e-}3 \\ 0.467 & 0.0308 & 5.33 & 16.0 \end{pmatrix}, \tag{2.47}$$

$$\mathbb{E}\left[|\Delta\mathbf{H}|^2\right] = \begin{pmatrix} 6.8\text{e-}5 & 6.8\text{e-}5 & 1.71\text{e-}3 & 1.71\text{e-}3 \\ 0.889 & 0.889 & 2.07\text{e-}2 & 2.07\text{e-}2 \end{pmatrix} \tag{2.48}$$

These results confirm that $\mathbb{E}[|\Delta h_{k\ell}^{(m)}|^2]$ does not depend on k, as expected. Also, they show that the ratio of the error variance to the channel strength greatly depends on the channel matrix, which poses significant challenges for accurate channel estimation in a multi-user scenario.

2.2.3 MMSE channel estimation

With reference to (2.2),

$$\begin{aligned} r_\ell(t) &= \sum_{i=1}^{M}\sum_{j=1}^{n_i} h_{j\ell}^{(i)} s_{ij}(t-\tau_i) + z_\ell(t) \\ &= \sum_{i=1}^{M}\sum_{j=1}^{n_i} h_{j\ell}^{(i)} \sum_{p=0}^{L-1} b_{ij}[p] g(t-pT-\tau_i) + z_\ell(t), \end{aligned}$$

let us now define $\underline{r}_\ell$ as the row vector of the samples of the received symbols at antenna ℓ:

$$\underline{r}_\ell = [r_\ell(T), \ldots, r_\ell(LT)], \quad \ell = 1, \ldots, N_A. \tag{2.49}$$

The scalar $r_\ell(qT)$ is suitable to be expressed using matrix notation. Let $\mathbf{s}_i(qT-\tau_i)$ be the training symbols sent by all antennas of node i as seen by the receiver at time qT, $q = 1,\ldots,L$:

$$\mathbf{s}_i(qT-\tau_i) = [s_{i1}(qT-\tau_i),\ldots,s_{in_i}(qT-\tau_i)]^T/\sqrt{LT}, \quad q=1,\ldots,L, \tag{2.50}$$

where we recall that $s_{ij}(qT-\tau_i) \in \{-1,+1\}$, $j=1,\ldots,n_i$, $i=1,\ldots,M$, and that τ_i is the propagation delay experienced by all signals of user i. Let us consider the qth symbol of all training sequences; by grouping the corresponding received versions into a vector, we get

$$\underline{s}(qT) = \left[\mathbf{s}_1(qT-\tau_1)^T,\ldots,\mathbf{s}_M(qT-\tau_M)^T\right], \qquad q=1,\ldots,L, \tag{2.51}$$

and by recalling that $U = \sum_{i=1}^{M} n_i$, we can write the $U \times L$ matrix of all training symbols received from all transmit antennas as

$$\begin{aligned}\mathbf{S} &= \left[\underline{s}(T)^T,\ldots,\underline{s}(LT)^T\right] = \\ &= \begin{bmatrix} \mathbf{s}_1(T-\tau_1) & \ldots & \mathbf{s}_1(LT-\tau_1) \\ \vdots & & \vdots \\ \mathbf{s}_M(T-\tau_M) & \ldots & \mathbf{s}_M(LT-\tau_M) \end{bmatrix}.\end{aligned} \tag{2.52}$$

Using (2.24), (2.25), (2.50) and (2.51) the signal received at antenna ℓ in (2.2) becomes:

$$\begin{aligned} r_\ell(qT) &= \sum_{i=1}^{M}\sum_{j=1}^{n_i} h_{j\ell}^{(i)} s_{ij}(qT) + z_\ell(qT) \\ &= \sum_{i=1}^{M}\left(h_{1\ell}^{(i)} s_{i1}(qT) + h_{2\ell}^{(i)} s_{i2}(qT) + \ldots + h_{n_i\ell}^{(i)} s_{in_i}(qT)\right) + z_\ell(qT) \\ &= \underline{h}_\ell^{(1)}\mathbf{s}_1(qT-\tau_1)^T + \ldots + \underline{h}_\ell^{(M)}\mathbf{s}_M(qT-\tau_M)^T + z_\ell(qT) \\ &= \underline{h}_\ell \underline{s}(qT)^T + z_\ell(qT), \quad \ell = 1,\ldots,N_A,\ q=1,\ldots,L. \end{aligned} \tag{2.53}$$

Focus for the moment on a perfectly synchronous case, whereby $\tau_i = 0$, $i=1,\ldots,M$. By calling

$$\underline{z}_\ell = [z_\ell(T),\ldots,z_\ell(LT)], \tag{2.54}$$

a compact form of (2.49) is then

$$\begin{aligned} \underline{r}_\ell &= [r_\ell(T),\ldots,r_\ell(LT)] = \\ &= \underline{h}_\ell\left[\underline{s}(T)^T,\ldots,\underline{s}(LT)^T\right] + \underline{z}_\ell \\ &= \underline{h}_\ell\mathbf{S} + \underline{z}_\ell, \quad \ell = 1,\ldots,N_A. \end{aligned} \tag{2.55}$$

By further grouping $\underline{r}_\ell$, $\ell = 1, \ldots, N_A$, we obtain

$$\begin{aligned}\mathbf{R_S} &= \begin{bmatrix} \underline{r}_1 \\ \vdots \\ \underline{r}_{N_A} \end{bmatrix} = \begin{bmatrix} \underline{h}_1 \\ \vdots \\ \underline{h}_{N_A} \end{bmatrix} \mathbf{S} + \begin{bmatrix} \underline{z}_1 \\ \vdots \\ \underline{z}_{N_A} \end{bmatrix} \\ &= \mathbf{HS} + \mathbf{Z_S}, \end{aligned} \tag{2.56}$$

where both $\mathbf{R_S}$ and $\mathbf{Z_S}$ have size $N_A \times L$. Let us now rewrite (2.56) in vectorized form [34]

$$\begin{aligned}\mathbf{r_S} &= \mathrm{vec}\left(\mathbf{R_S}\right) = \mathrm{vec}\left(\mathbf{I}_{N_A}\mathbf{HS}\right) + \mathrm{vec}\left(\mathbf{Z_S}\right) \\ &= \left(\mathbf{S}^T \otimes \mathbf{I}_{N_A}\right) \mathrm{vec}\left(\mathbf{H}\right) + \mathbf{z_S} \\ &= \overline{\mathbf{S}}\mathbf{h} + \mathbf{z_S}, \end{aligned} \tag{2.57}$$

where $\otimes$ denotes the Kronecker product. The linear MMSE estimator of the MIMO channel $\mathbf{H}$ is [36]

$$\begin{aligned}\widehat{\mathbf{h}}_{\mathbf{MMSE}} &= \mathrm{vec}\left(\widehat{\mathbf{H}}_{\mathbf{MMSE}}\right) \\ &= \Phi_{\mathbf{hr_S}} \Phi_{\mathbf{r_S}}^{-1} \mathbf{r_S}, \end{aligned} \tag{2.58}$$

where

$$\Phi_{\mathbf{hr_S}} = \mathbb{E}\left[\mathbf{h}\mathbf{r}_{\mathbf{S}}^H\right] \quad \text{and} \quad \Phi_{\mathbf{r_S}} = \mathbb{E}\left[\mathbf{r_S}\mathbf{r}_{\mathbf{S}}^H\right]. \tag{2.59}$$

Under the assumption that all signals are time synchronous, we get the well-known expressions

$$\Phi_{\mathbf{hr_S}} = \Phi_{\mathbf{h}}\overline{\mathbf{S}}^H \quad \text{and} \quad \Phi_{\mathbf{r_S}} = \overline{\mathbf{S}}\Phi_{\mathbf{h}}\overline{\mathbf{S}}^H + \Phi_{\mathbf{z_S}}, \tag{2.60}$$

where

$$\Phi_{\mathbf{h}} = \mathbb{E}\left[\mathbf{h}\mathbf{h}^H\right] \quad \text{and} \quad \Phi_{\mathbf{z_S}} = \mathbb{E}\left[\mathbf{z_S}\mathbf{z}_{\mathbf{S}}^H\right] = N_0\mathbf{I}_{LN_A}. \tag{2.61}$$

Therefore the MMSE estimator in the synchronous case becomes (see also [34])

$$\widehat{\mathbf{h}}_{\mathbf{MMSE}} = \Phi_{\mathbf{h}}\overline{\mathbf{S}}^H \left(\overline{\mathbf{S}}\Phi_{\mathbf{h}}\overline{\mathbf{S}}^H + \Phi_{\mathbf{z_S}}\right)^{-1} \mathbf{r_S}. \tag{2.62}$$

We now use the following variant of the Woodbury identity [37]: let A and B be positive definite matrices, then

$$AC^H\left(CAC^H + B\right)^{-1} = \left(A^{-1} + C^H B^{-1} C\right)^{-1} C^H B^{-1}. \tag{2.63}$$

In our tractation, we let $A = \Phi_{\mathbf{h}}$, $B = \Phi_{\mathbf{z_S}}$ and $C = \overline{\mathbf{S}}$. It is straightforward to show that A and B are positive definite.

Thus, we can rewrite (2.62) as follows

$$\widehat{\mathbf{h}}_{\mathbf{MMSE}} = \left(\Phi_{\mathbf{h}}^{-1} + \overline{\mathbf{S}}^H \Phi_{\mathbf{z_S}}^{-1} \overline{\mathbf{S}}\right)^{-1} \overline{\mathbf{S}}^H \Phi_{\mathbf{z_S}}^{-1} \mathbf{r}_{\mathbf{S}}, \tag{2.64}$$

where $\Phi_{\mathbf{h}}$ is a block-diagonal matrix with elements $\sigma_i^2 \mathbf{I}_{n_i N_A}$, $i = 1, \ldots, M$, that is,

$$\Phi_{\mathbf{h}} = \mathbb{E}\left[\mathbf{h}\mathbf{h}^H\right] = \begin{pmatrix} \sigma_1^2 \mathbf{I}_{n_1 N_A} & & \\ & \ddots & \\ & & \sigma_M^2 \mathbf{I}_{n_M N_A} \end{pmatrix} \tag{2.65}$$

By replacing (2.57) in (2.64), through straightforward algebra we obtain

$$\widehat{\mathbf{h}}_{\mathbf{MMSE}} = \left(\Phi_{\mathbf{h}}^{-1} + N_0^{-1} \overline{\mathbf{S}}^H \overline{\mathbf{S}}\right)^{-1} N_0^{-1} \left(\overline{\mathbf{S}}^H \overline{\mathbf{S}} \mathbf{h} + \overline{\mathbf{S}}^H \mathbf{z_S}\right). \tag{2.66}$$

Due to the properties of the Kronecker product, we have

$$\overline{\mathbf{S}}^H \overline{\mathbf{S}} = \Phi_{\mathbf{S}} \otimes \mathbf{I}_{N_A}, \tag{2.67}$$

where $\Phi_{\mathbf{S}} = \mathbf{S}\mathbf{S}^H$ is the cross-correlation matrix of the training sequences in the synchronous case.

Let us now consider the case of asynchronous received signals: in this scenario, the τ_is, $i = 1, \ldots, M$, are not zero in general. This condition reflects on $\Phi_{\mathbf{S}}$, which must be computed according to the model in (2.5)–(2.6). To make the task simpler, let us subdivide $\Phi_{\mathbf{S}}$ as follows:

$$\Phi_{\mathbf{S}} = \begin{bmatrix} \mathbf{\Sigma}^{(1,1)} & \ldots & \mathbf{\Sigma}^{(1,M)} \\ \vdots & & \vdots \\ \mathbf{\Sigma}^{(M,1)} & \ldots & \mathbf{\Sigma}^{(M,M)} \end{bmatrix}, \tag{2.68}$$

where the term in position (j, k) of $\mathbf{\Sigma}^{(i,m)}$, according to (2.6) and considering a rectangular $g(t)$, is

$$\begin{aligned} \mathbf{\Sigma}_{jk}^{(i,m)} &= \frac{T - (\tau_i - \tau_m)}{LT} \sum_{p=0}^{L-1} b_{mk}[p] b_{ij}[p] + \frac{\tau_i - \tau_m}{LT} \sum_{p=1}^{L-1} b_{mk}[p] b_{ij}[p-1] \\ &= \frac{1}{LT} \int_{-\infty}^{+\infty} s_{ij}(t - (\tau_i - \tau_m)) s_{mk}(t) \, \mathrm{d}t = J_{jk}^{(i,m)}, \end{aligned} \tag{2.69}$$

where $j = 1, \ldots, n_i$, $k = 1, \ldots, n_m$. Note that $\mathbf{\Sigma}^{(i,m)}$, $i, m = 1, \ldots, M$ represents the cross-correlation matrix of the sequences sent by all transmit antennas of users i and m. Collecting the elements of $\Phi_{\mathbf{S}}$ in (2.68) as in equation (2.27) yields

$$\begin{aligned}\overline{\mathbf{S}}^H\overline{\mathbf{S}} &= \left[\mathbf{J}_1^{(1)}, \ldots, \mathbf{J}_{n_1}^{(1)}, \ldots, \mathbf{J}_1^{(M)}, \ldots, \mathbf{J}_{n_M}^{(M)}\right] \otimes \mathbf{I}_{N_A} = \mathbf{J} \otimes \mathbf{I}_{N_A} \\ &= \widetilde{\mathbf{J}} \otimes \mathbf{I}_{N_A} + \mathbf{I}_{UN_A}, \end{aligned} \tag{2.70}$$

where matrix $\widetilde{\mathbf{J}}$ can be modeled as a set of iid random variables $\sim \mathcal{N}\left(0, \frac{2}{3L}\right)$ (see Section 2.2.2 for details). We note that if the received signals are perfectly synchronous the orthogonality among sequences is preserved (i.e. $i = m$ case, when the training sequences are sent from the same transmitter), hence

$$\mathbf{\Sigma}^{(i,m)} = \begin{cases} \mathbf{I}_{N_A}, & i = m, \\ \mathbf{0}_{N_A}, & \text{otherwise}, \end{cases} \tag{2.71}$$

and $\overline{\mathbf{S}}^H\overline{\mathbf{S}}$ reduces to

$$\overline{\mathbf{S}}^H\overline{\mathbf{S}} = \mathbf{I}_{UN_A}. \tag{2.72}$$

As in (2.8), it is straightforward to conclude that the noise vector in (2.66), $\overline{\mathbf{S}}^H\mathbf{z}_{\mathbf{S}} = \mathbf{w}$, has iid elements $\sim \mathcal{CN}\left(0, \frac{N_0}{LT}\right)$ when $g(t)$ is the rectangular pulse. To conclude, the linear MMSE estimator for asynchronous signals takes the following form

$$\boxed{\widehat{\mathbf{h}}_{\mathbf{MMSE}} = \left[\Phi_{\mathbf{h}}^{-1} + N_0^{-1}\left(\widetilde{\mathbf{J}} \otimes \mathbf{I}_{N_A} + \mathbf{I}_{UN_A}\right)\right]^{-1} N_0^{-1}\left(\mathbf{h} + \left(\widetilde{\mathbf{J}} \otimes \mathbf{I}_{N_A}\right)\mathbf{h} + \mathbf{w}\right)}. \tag{2.73}$$

The last expression of linear MMSE estimation, (2.73), is directly comparable with equation (2.33) derived for the correlator, which can be expressed in vectorized form as:

$$\boxed{\widehat{\mathbf{h}}_{\mathbf{CORR}} = \mathbf{h} + \left(\widetilde{\mathbf{J}} \otimes \mathbf{I}_{N_A}\right)\mathbf{h} + \mathbf{w}}. \tag{2.74}$$

Equations (2.73) and (2.74) clearly show that the error matrices related to a correlator-based or MMSE channel estimator have two components, one proportional to the channel matrix, $\left(\widetilde{\mathbf{J}} \otimes \mathbf{I}_{N_A}\right)\mathbf{h}$, which include interfering transmitters, and one independent of it, $\mathbf{w}$.

In the following section, we elaborate on the effects of imperfect channel estimation at the receiver by applying the analysis carried out so far to the evaluation of the performance of a MIMO ad hoc network.

2.3 Impact of Channel Estimation Errors in a PHY-MAC Protocol

The results in this section are derived adopting the MAC protocol described in Section 2.1 in a network with 25 nodes deployed in a square grid within a $100\,\mathrm{m} \times 100\,\mathrm{m}$ area, so that the distance between nearest neighbors is $25\,\mathrm{m}$. All nodes are static and frame-synchronous for the whole duration of the simulations. As in [10], we choose this specific topology because all nodes potentially receive a significant amount of interference from one another, a demanding scenario for the MAC protocols under test. Each node generates packets according to a Poisson process of rate λ packets per second per node. Each packet is formed of k PDUs, with k uniformly chosen in $\{1, 2, 3, 4\}$. Packets waiting to be sent are stored in a queue which can hold up to 120 PDUs. If the storage time exceeds a custom number of 2500 frames, the packet times out and is discarded. Each PDU is $1000\,\mathrm{bits}$ long, whereas signaling packets are $200\,\mathrm{bits}$ long. These values do not account for the length of the training sequence, L. For simulations with imperfect channel estimation, we assume that each packet contains such a sequence (see also the frame structure in Figure 2.1). In these cases, we use the models derived in Section 2.2.2, for correlator-based estimator, and in Section 2.2.3, for MMSE estimator, to consider the imperfect channel estimation. The results hereafter are obtained for $N_A = 4$ and $N_A = 8$ antennas at each node (we assume the antennas are sufficiently spaced, so as to yield a capacity very close to that an iid Rayleigh MIMO channel, see [38]). During the reception phase, a node uses all antennas to run the V-BLAST algorithm with the maximum number of degrees of freedom. On the contrary, the number of antennas to use during transmissions is chosen according to the directives of the received CTS packets (in particular, node i uses exactly n_i antennas, see also Section 2.1).

We employ the analytical approach devised in [17] in order to accurately model PHY-level details, without an excessive burden for the simulator. This technique approximates all contributions to the receive SINR as Gaussian, and separately accounts for the impairments caused by noise, imperfect cancellations, and signals that are still to be canceled in the V-BLAST stack.

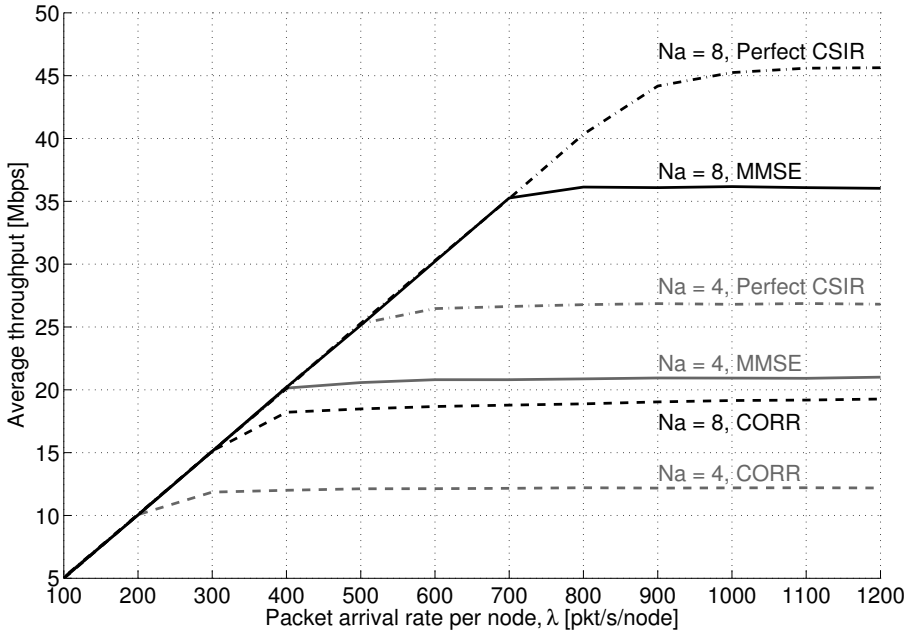

Figure 2.4. *Average throughput as a function of λ for the MMSE and CORR detectors, $L = 128$, compared to perfect channel estimation. The considered number of antennas is $N_A = 4$ (grey curves) and $N_A = 8$ (black curves).*

2.3.1 Study of MAC performance under perfect and imperfect channel estimation

Figure 2.4 shows the total average network throughput (in Mbps) as a function of traffic in case of both perfect and imperfect CSI at the Receiver (CSIR). Both the MMSE and correlator-based (CORR) estimators are shown; in addition, the study is performed for both $N_A = 4$ and $N_A = 8$. In this evaluation the length of the training sequences is fixed to $L = 128$. For a fair comparison, the perfect estimation case also assumes that a training sequence of the same length $L = 128$ is added to the transmission of RTSs, CTSs and PDUs. Figure 2.4 highlights the performance degradation incurred with imperfect CSIR, which translates into a reduced maximum throughput, about 25% lower when using MMSE, and up to 55% lower when using CORR with either number of antennas. In fact, imperfect CSIR makes the V-BLAST detection algorithm much more prone to interference, causing a greater probability of error and thus lower throughput. We stress that two negative effects add up to impact on the detection performance of V-BLAST. First, a wrong channel estimation leads to canceling an incorrect contribution from the received signal, even in the case of correct

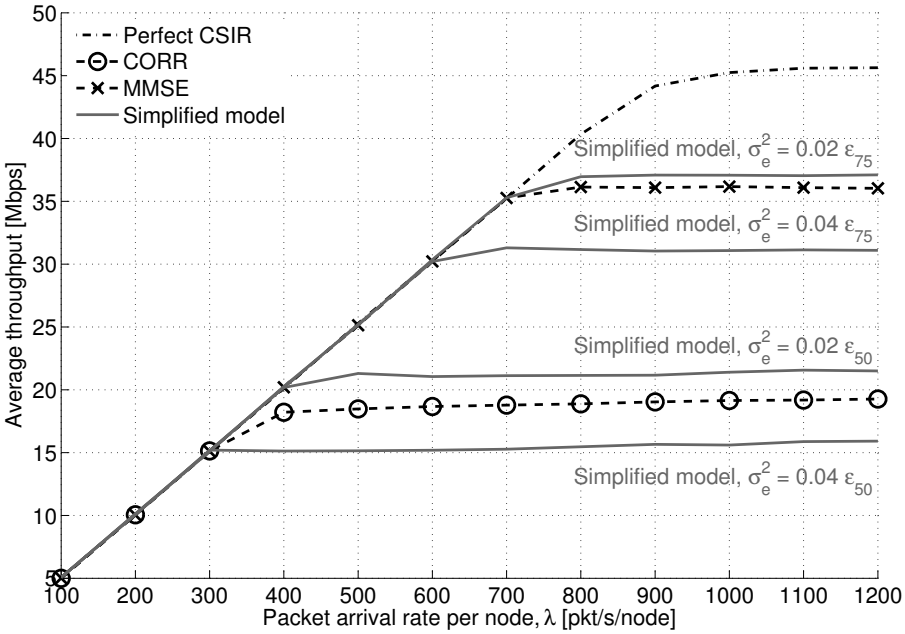

Figure 2.5. *Comparison of throughput as a function of λ for the estimation error model employed in this work against an additive Gaussian error model of given variance σ_e^2, normalized to the average power received from a given distance, chosen as $50\,\mathrm{m}$ (ε_{50}) and $75\,\mathrm{m}$ (ε_{75}), $N_A = 8$.*

detection. This leaves a trace of the canceled signal, which will decrease the SINR of all following detections. Second, even with perfect channel estimation, a wrong detection and cancellation (e.g., due to low SINR) may double the interference of the current signal on subsequent detection steps. Therefore the last signals in the V-BLAST detection order (that would experience the best SINR thanks to previous cancellations in the perfect CSIR case) are in fact the most affected. This limits the maximum achievable transmission parallelism. We also note that the throughput does not decrease to zero, but is instead maintained at a constant level for increasing λ because of MAC-level backoff, which limits channel access attempts and thus interference.

As expected, using $N_A = 4$ antennas results in lower network performance, while maintaining roughly the same proportion among the perfect CSIR, MMSE and CORR curves. For this reason, and because $N_A = 8$ allows more degrees of freedom to the MAC protocol, we will concentrate on the $N_A = 8$ case, with the understanding that for lower N_A the results are conceptually similar, though scaled down.

Our estimation error model is able to capture the effects of concurrent transmissions

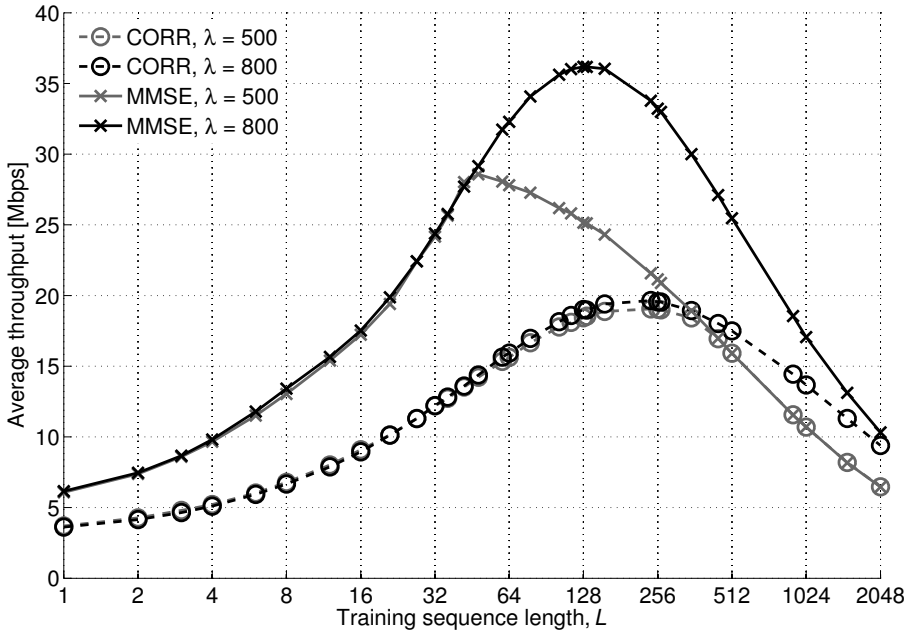

Figure 2.6. *Average throughput as a function of L for both MMSE and CORR,* $\lambda = 500, 800\,\mathrm{pkt/s/node}$, $N_A = 8$.

(that create interference to each other), whose interfering power varies on a per frame basis, on the accuracy of channel estimation. Simpler models of the kind $\hat{\mathbf{H}} = \mathbf{H} + \mathbf{\Delta H}$, where the error is modeled as an additive Gaussian term $\mathbf{\Delta H}$ of given variance σ_e^2 (independent of $\mathbf{H}$), are perhaps more straightforward, but do not consider the impact of multiple access interference. In addition, it is not clear how σ_e^2 is to be computed as a function of the system parameters. A comparison between our model and the Gaussian displacement model is shown in Figure 2.5, where the variance σ_e^2 is set to either 0.02 or 0.04, normalized to the average power received from a given distance, set to $50\,\mathrm{m}$ and $75\,\mathrm{m}$, and indicated in the figure with ε_{50} and ε_{75} respectively. We observe that only by empirically fine-tuning the normalized variance value can the grey curves (representing the simpler model) approach the black curves, where the estimation error model accounts for time-varying interference, and is thus tightly related to the system parameters. This shows that in a multiple-access context our model is significantly better and easier to use than the traditional Gaussian approach.

The length of the training sequence L is a key parameter, affecting both channel estimation and network performance in a non-trivial way. A longer training sequence improves the estimation accuracy, see (2.20), but also causes greater overhead, thus decreasing the effi-

ciency of all transmissions (recall that a training sequence is added to the transmission of all packets except ACKs, see Section 2.1). To quantify this concept, consider Figure 2.6, where throughput is plotted against L for $\lambda = 500$ and $\lambda = 800$ $\mathrm{pkt/s/node}$. The figure highlights that throughput is low in two opposite conditions, i.e., when L is too small and when it is too large. In the first case, the overhead imposed by channel estimation is negligible, but channel estimates are unreliable: therefore, nodes experience low transmission success ratio and throughput. Conversely, when L is large, channel estimates are very reliable, but a large fraction of the MAC frame is occupied by the training sequence: this leaves little room for data and results in low throughput as well. Referring to Figure 2.6, the optimal setting (from a throughput point of view) strikes a balance between low overhead and good channel estimation, and depends on the type of estimator, as well as on λ: for example, MMSE requires $L = 128$ if MAC-level throughput has reached saturation ($\lambda = 800$, see Figure 2.4), but only L between 32 and 64 for $\lambda = 500$, which causes lower traffic, hence interference: therefore, the required channel estimation accuracy is also lower. A similar argument holds for CORR as well: however both $\lambda = 500$ and 800 lie in the MAC throughput saturation region, unlike in the MMSE case: this explains the smaller differences between the two CORR curves.

In order to better assess the impact of the overhead due to channel estimation, we introduce two different efficiency metrics. The first, named *transmission efficiency*, is defined as the ratio of correctly received information bits over all sent bits, averaged over all nodes and frames. This definition allows to balance between the greater estimation accuracy (thus better probability of success) and the greater overhead incurred by increasing the length of the training sequence, L. However, it does not explicitly depend on the behavior of the protocol (e.g., on the way receivers are chosen and links are set up), but only on PHY-level parameters (i.e., L). For this reason we consider the *class efficiency*, which is more closely related to our MAC protocol. In each frame, we compute this metric as follows: we take the sum of the number of PDUs ACKed by all receivers (provided that the ACK is correctly received); then we consider the set of receivers of each transmitter, take the minimum class of the receivers within each set, and sum all minimum classes (the minimum class within a set of intended receivers in a frame has been introduced in Section 2.1). Finally, we take the ratio of the two, and average it over all frames. Note that the minimum class depends on which receivers a transmitter must send packets to (thus on the traffic pattern), and is determined during the

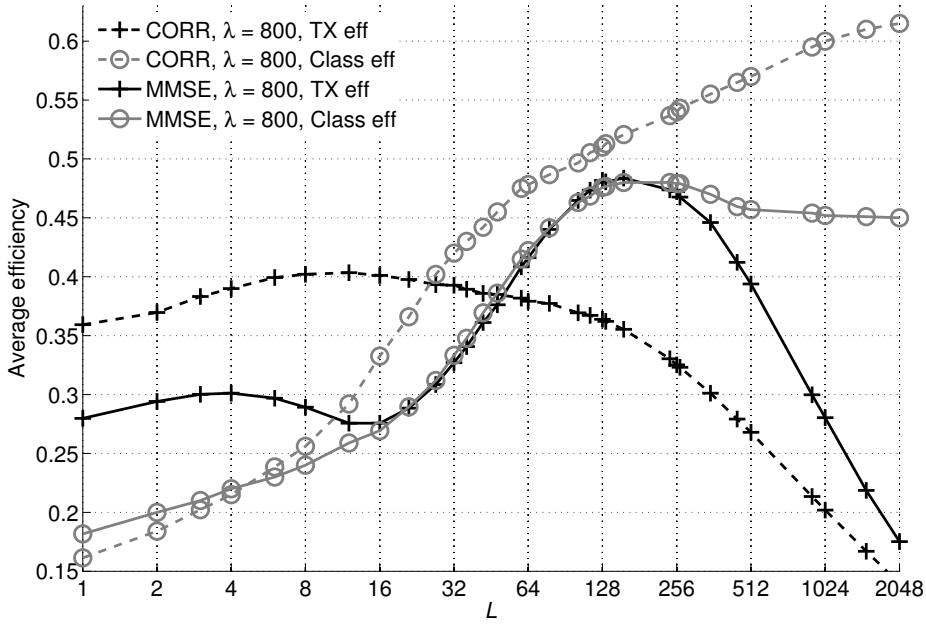

Figure 2.7. *Transmission efficiency and class efficiency as a function of L for both MMSE and CORR,* $\lambda = 800\,\mathrm{pkt/s/node}$, $N_A = 8$.

RTS phase; conversely the number of correctly received PDUs (which is calculated based on the information reported in ACKs) is a result of all protocol decisions within the frame. Therefore the class efficiency represents a very protocol-specific outcome.

We highlight that, given a minimum class c, the best a transmitter can do is to have c PDUs get through correctly to their receivers. In other words, the class efficiency metric is equal to 1 if all transmitters can satisfy the latter condition in all frames. The transmission efficiency and class efficiency metrics are plotted against L in Figure 2.7. Coherently with throughput curves, transmission efficiency shows a maximum around $L = 128$ for MMSE and $L = 32$ for CORR, but does not decrease to zero for low L. The latter represents a particular network situation where most of the RTS packets are not decoded due to bad channel estimation: this reduces traffic artificially, because nodes spend most of their time in backoff. Therefore, very few data transmissions actually occur, and due to the very low interference they almost always succeed. This makes the transmission efficiency significant, despite the very low throughput. For very high L, instead, the overhead required by channel estimation is very large, hence the transmission efficiency drops. The class efficiency exhibits a more regular behavior: as L increases, the better accuracy of channel estimation allows

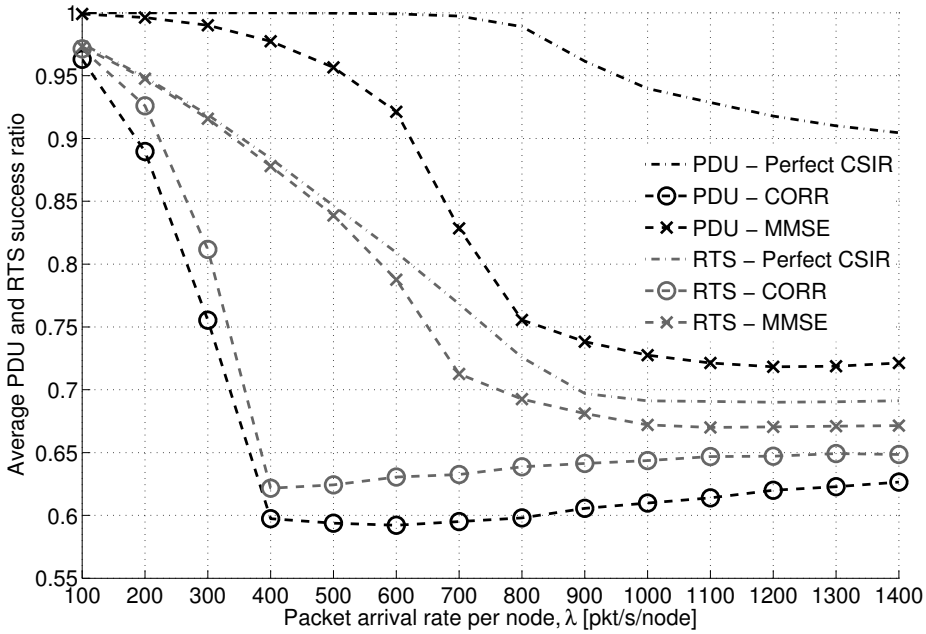

Figure 2.8. *Average PDU and RTS success ratio as a function of λ for both MMSE and CORR, for $L = 128$ and $N_A = 8$.*

both more signaling packets and more PDUs to be received correctly, and therefore a larger number of PDUs to be ACKed. Hence, class efficiency grows with L, is greatest at $L = 128$ for MMSE and stabilizes around a near-maximum value at very high L. For CORR, class efficiency is monotonically increasing for increasing L: this is explained in terms of the resulting larger traffic and lower probability of error, which makes it possible to correctly convey more traffic through activated links.

To sum up, a good choice for L is found by jointly considering Figures 2.6 and 2.7: for MMSE, the best value of L is around 128, which jointly optimizes throughput and efficiency; for the poorer CORR estimator, $L = 64$ is good for throughput and transmission efficiency, whereas class efficiency is suboptimal. However, a larger value would make little sense, given the short length of signaling packets ($200\,\mathrm{bits}$), and choosing a larger value for L would imply a waste of resources.

Up to this point, we have mostly considered the impact of imperfect CSIR on PDU reception. However, a second interesting effect is observed during the RTS/CTS exchange phases. To explain this, let us consider Figure 2.8, which shows the average transmission success ratios for RTSs and PDUs as a function of traffic for fixed $L = 128$, thus fixed channel esti-

mation accuracy. For comparison, the curves for the perfect CSIR case are also shown. Even in case of perfect CSIR, increasing traffic has a detrimental impact on RTS transmissions, which is due to more frequent channel access by all nodes. The PDU success ratio, instead, is more stable thanks to the way transmissions are granted (the CTS policy described in Section 2.1, which tends to control interference). With imperfect CSIR, the success ratio of both signaling packets and PDUs decreases even further, because of poor detections and wrong cancellations. (Choosing e.g., $L = 2048$ would make the two sets of curves more similar.) This shows that the throughput loss due to imperfect channel estimation is not simply due to more frequent PDU losses, but also to corrupted signaling packets, which do not allow links to be set up correctly.

2.3.2 The effect of different CSIR accuracy for signaling and data packets

Unlike before, let us allow the signaling packets and PDUs to use training sequences of different lengths (L_{sig} and L_{data}, respectively): by differentiating between estimation accuracy for signaling and data, we can get some more insight about how channel estimation affects MAC performance. In Figure 2.9, we consider the MMSE estimator, and plot contour curves of throughput and transmission efficiency as L_{sig} and L_{data} are independently varied from 1 to 2048. A greater L_{sig} yields better protection to RTSs: as more RTSs are correctly received, more links are established. Conversely, a greater L_{data} protects PDUs, but only those that have actually been sent. Therefore, a greater L_{data} can increase throughput significantly only if enough links can be established, i.e., only for high L_{sig} as well. For example, take $L_{sig} = 32$: increasing L_{data} causes throughput to increase up to roughly $40\,\mathrm{Mbps}$ (for L_{data} between 128 and 256), after which no further improvements are possible, because the number of active links is limited by L_{sig}. Conversely, take $L_{data} = 32$: increasing L_{sig} initially improves throughput, but only up to roughly $L_{sig} = 32$. In fact, for greater L_{sig} more RTSs are received correctly, hence more links are set up, and $L_{data} = 32$ cannot enforce a sufficiently accurate channel estimation to enable correct detections. Thanks to the superposition of throughput and efficiency, Figure 2.9 allows to tune the working point on the efficiency-throughput tradeoff by varying L_{sig} and L_{data}. For example, to ensure an efficiency of at least 0.5 and a throughput of at least $30\,\mathrm{Mbps}$ we can choose any pair of L_{sig} and L_{data} within the intersection of the corresponding contour curves. For example, the choice

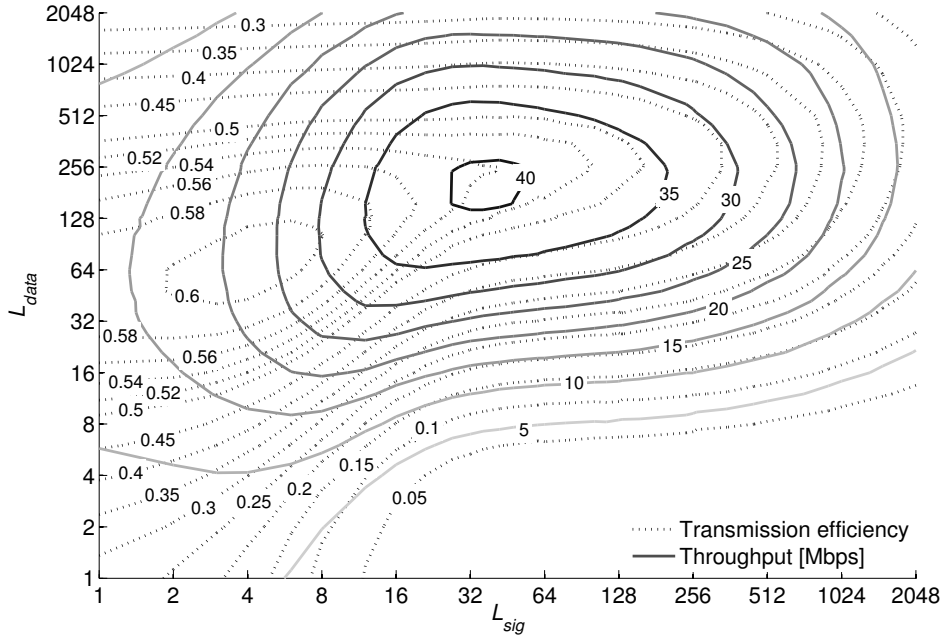

Figure 2.9. *Contour curves of average transmission efficiency and throughput when using the MMSE estimator. Curves are plotted as a function of the length of the training sequences used for signaling messages (L_{sig}) and PDUs (L_{data}), $\lambda = 800\,\mathrm{pkt/s/node}$, $N_A = 8$. Each curve in the Figure is obtained as the intersection of the efficiency or throughput surfaces as a function of L_{data} and L_{sig} with a horizontal plane corresponding to the efficiency or throughput value indicated by the label on each curve.*

$L_{sig} = 32$, $L_{data} = 128$ satisfies both constraints. Conversely, if the constraints are too strict (e.g., an efficiency of 0.6 and a throughput of $40\,\mathrm{Mbps}$) the requirements are incompatible and the curves do not intersect. For comparison, note that CORR (Figure 2.10) achieves worse performance, both in terms of maximum throughput and efficiency, and in terms of tradeoff between the values of the two metrics.

In order to get a better grasp on the tradeoff between different performance figures as a function of channel estimation accuracy, let us now choose two metrics between which we evaluate a tradeoff for fixed packet generation rate λ. In Figure 2.11 we consider the throughput and transmission efficiency values achieved by all (L_{sig}, L_{data}) pairs, and plot the convex hull of the resulting scatterplot: this allows to pick only the pairs that offer the best (i.e., outermost) metric values. Power-of-2 pairs are shown using circles and crosses and highlighted through labels. Note that, in order to get a denser scatterplot with a smoother convex hull, we have also considered pairs where (L_{sig} and L_{data}) are not powers of 2. Figure 2.11 shows the described plots for both MMSE and CORR, at $\lambda = 500$ (grey curves) and

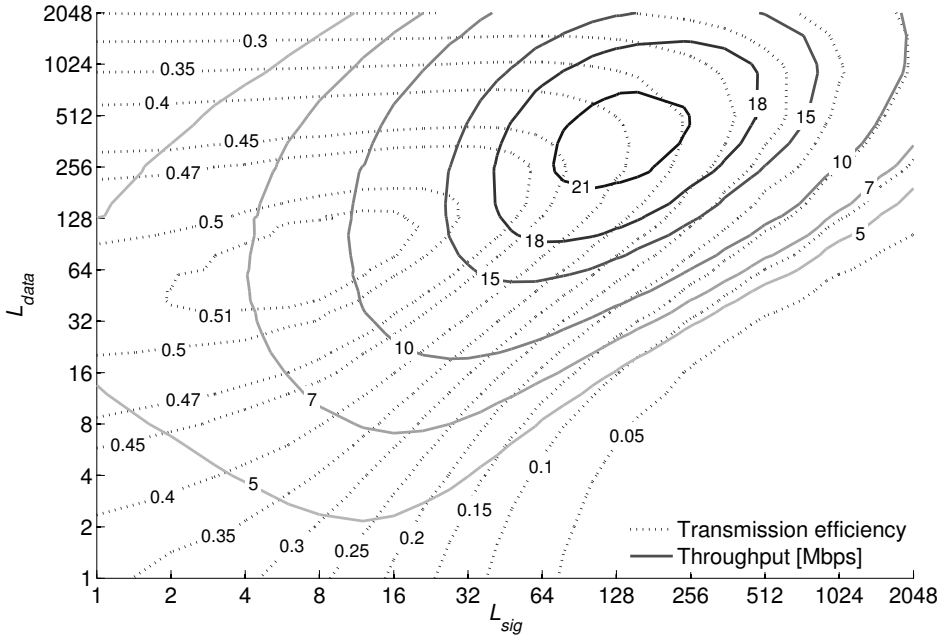

Figure 2.10. *Contour curves of average transmission efficiency and throughput when using the CORR estimator. Curves are plotted as in Figure 2.9.*

$\lambda = 800$ packets per second per node (black curves). The important section of the curves is the top-right corner, where the best combination of throughput and efficiency is achieved. A good choice of (L_{sig}, L_{data}) which optimizes throughput in spite of a suboptimal transmission efficiency, is $(32, 256)$ for MMSE and $(128, 256)$ for CORR, at $\lambda = 800$. Note that the $\lambda = 500$ curve for MMSE exhibits a flat region at the top-right corner, because this value of λ lies in the linear increase region of MAC-level throughput, see Figure 2.4: as a consequence, the maximum throughput achievable corresponds to delivering all generated PDUs; in turn, the optimum values of $(L_{\mathrm{sig}}, L_{\mathrm{data}})$ are just sufficient to ensure accurate channel estimation for all packets. We highlight that transmission efficiency and class efficiency (the capability of transmitters to saturate the minimum class among their set of receivers) are competing constraints in terms of training length. Figure 2.12 plots the convex hull of the scatterplot of class efficiency vs. throughput performance points, and clearly shows this fact. While the values reaching optimal throughput (uppermost part of the curves) are the same as in Figure 2.11, best class efficiency (rightmost part of the curves) requires quite higher values of L_{sig} and L_{data}.

The PDU transmission success ratio and average delivery delay are also important met-

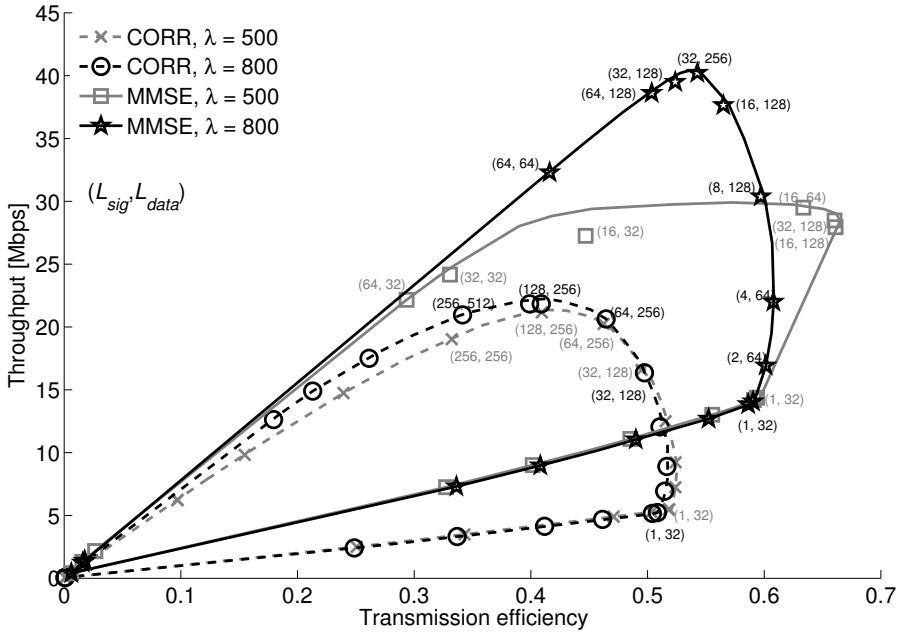

Figure 2.11. *Convex hull of the scatterplot of transmission efficiency vs. throughput performance points for varying pairs of L_{sig} and L_{data}, $N_A = 8$. Some relevant points corresponding to power-of-two training sequence lengths are highlighted for reference.*

rics in a wireless network. The contour plot of these metrics is shown in Figure 2.13 for MMSE. We observe that the optimal values for success ratio are expectedly found for high L_{data}; however, long training sequences excessively prolong the duration of the frame, leading to longer delivery delays: therefore, the best point for delay is around $(128, 512)$. As observed before, the CORR estimator requires longer training sequences to achieve sufficient estimation accuracy; therefore, lower success ratios and higher delays are experienced, the best point for delay being around $(256, 1024)$, see Figure 2.14. We highlight that the latter point has only a purely theoretical meaning, suggesting that minimum delay is achieved when links are set up as fast as possible ($L_{\mathrm{sig}} = 256$ gives very good estimates for RTSs) and data is protected from estimation errors ($L_{\mathrm{data}} = 1024$). The choice among the working points discussed so far should depend on the primary metric to be optimized. Considering the MMSE case, the best throughput is obtained at $(32, 256)$, the best efficiency at $(4, 64)$, the best delay at around $(128, 512)$. These could be jointly optimized by, e.g., defining objective functions which encompass different performance indications into a synthetic figure of merit.

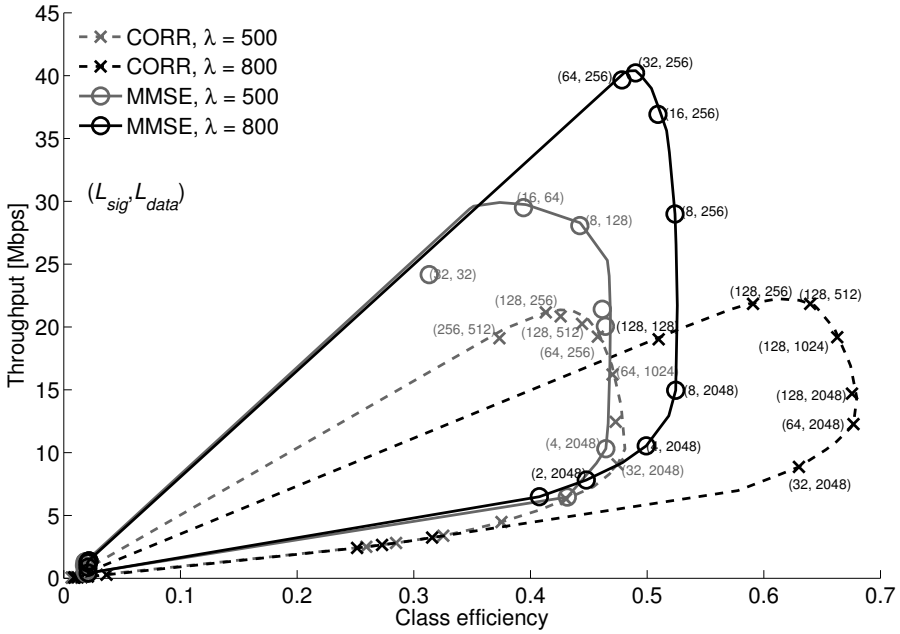

Figure 2.12. *Convex hull of the scatterplot of class efficiency vs. throughput performance points for varying pairs of L_{sig} and L_{data}. Some relevant points corresponding to power-of-two training sequence lengths are highlighted for reference.*

A standard way to do so is to normalize metrics so that they take values in the interval $[0, 1]$, and consider a linear combination of those metrics with weights ω_i such that $\sum_i \omega_i = 1$. This corresponds to establishing a linear relationship between the value of a metric and the "satisfaction" of the network designer with that metric [39, Chapter 3]. In this joint evaluation, we consider throughput, transmission efficiency, class efficiency and delivery delay, which we index from 1 to 4 and average using a weighing vector $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_4]$. In Figure 2.15 we consider three different vectors, namely $[0, 0.5, 0.5, 0]$ (equal importance of transmission efficiency and class efficiency), $[0.4, 0.3, 0.3, 0]$ (giving importance to throughput and then equally to the efficiency metrics) and $[0.25, 0.25, 0.25, 0.25]$ (weighing all metrics uniformly). We observe that a different choice of weights moves the optimum performance point from a low L_{sig} to a higher L_{sig} whenever the focus is on throughput rather than on pure efficiency. Similarly, if delay is brought into the average, the optimum working point is shifted toward even higher L_{sig} and L_{data}, so that more links are established, packets are more readily transmitted, and also correctly detected (by virtue of a high L_{data}).

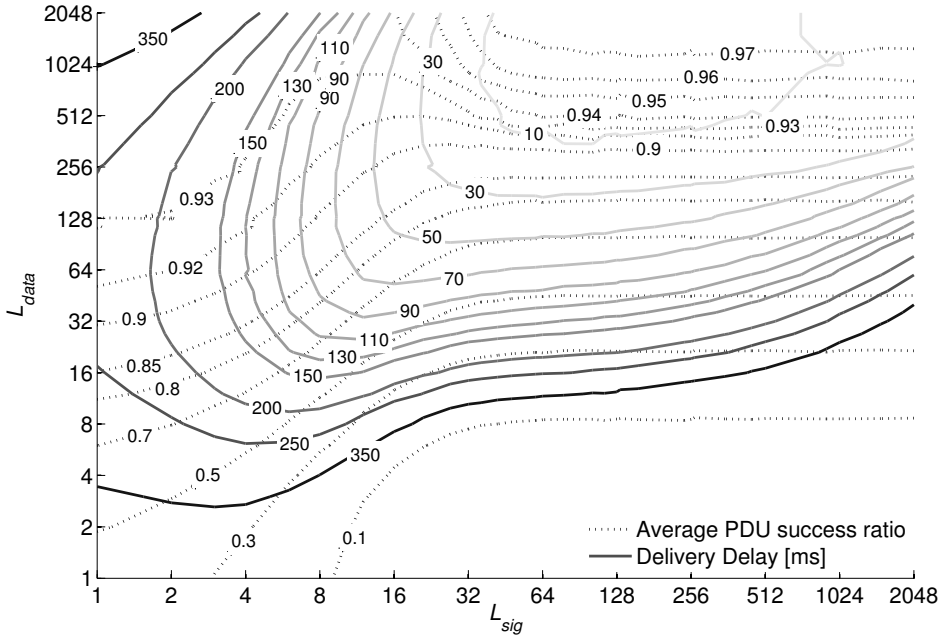

Figure 2.13. *Contour curves of average PDU transmission success ratio and delivery delay when using the MMSE estimator. Curves are plotted as a function of the length of the training sequences used for signaling messages (L_{sig}) and PDUs (L_{data}), $\lambda = 800\,\mathrm{pkt/s/node}$, $N_A = 8$.*

2.3.3 Effect of Different Transmit Waveforms

In the previous section we have analyzed a multiuser MIMO networks employing either correlator-based or MMSE channel estimators and we have then observed how PHY-level parameters, such as the length of the training sequences, affect higher-level metrics, such as throughput, efficiency, delay and success ratio. In the following, we consider the same model used before with correlator-based channel estimator only and we evaluate the impact of different pulse shapes introduced in Section 2.2.2.1 on the global performance of the network.

Figure 2.16 shows throughput (in Mbps) as a function of traffic in case of both perfect and imperfect CSIR (i.e., correlator-based channel estimator). All impulse types of Section 2.2.2.1 are considered. For this first comparison, the length of the training sequences is fixed to $L = 128$. Figure 2.16 highlights the performance degradation incurred with imperfect CSIR regardless of the particular choice of the transmit impulse $g(t)$, as compared to the case where perfect CSIR is available. In line with the observations in Figures 2.2 and 2.3, pulses yielding a lower value of σ_J^2 correspond to better network performance. Taking the

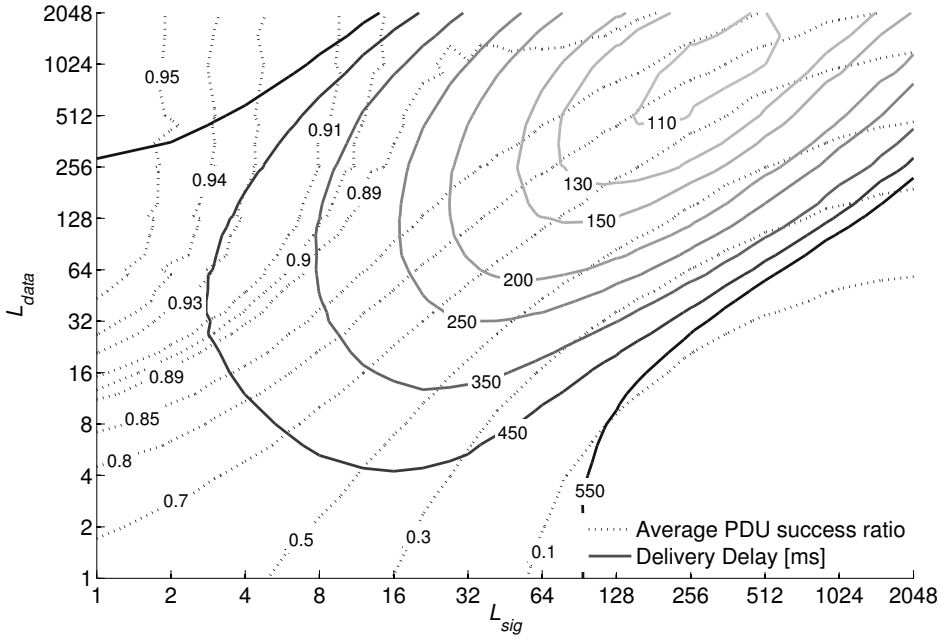

Figure 2.14. *Contour curves of average PDU transmission success ratio and delivery delay when using the CORR estimator. Curves are plotted as a function of the length of the training sequences used for signaling messages (L_{sig}) and PDUs (L_{data}), $\lambda = 800\,\mathrm{pkt/s/node}$.*

two extreme cases, for $L = 128$ the value of σ_J^2 with the Sinc pulse is 4 times higher than with the Gaussian pulse, $a = 1$: this translates into a difference of $8\,\mathrm{Mbps}$ in network performance at the MAC level.

Similarly to the previous section, we now perform a different analysis by letting training sequences for signaling and data packets have different length (i.e., L_{sig} and L_{data}, respectively). Figures 2.17 and 2.18 provide a representation of the tradeoff between throughput and efficiency as a function of L_{sig} and L_{data} by means of contour curves. In particular, Figure 2.17 refers to the use of a Gaussian pulse with $a = 1$, whereas Figure 2.18 corresponds to a Sinc pulse. Both figures give first of all some general information about the main impairments to a correct network behavior, namely the setup of a low number of links (due to a low L_{sig} which leads to high probability of error for RTSs, see also comments of Figure 2.9), and the insufficient accuracy of channel estimates for PDUs, due to a low L_{data}. If either is too low, both throughput and efficiency are limited. Note that throughput is low even if L_{data} is too small compared to L_{sig}, as the accuracy of channel estimates for PDUs becomes insufficient to support traffic over all links set up by correct RTSs. Because network metrics

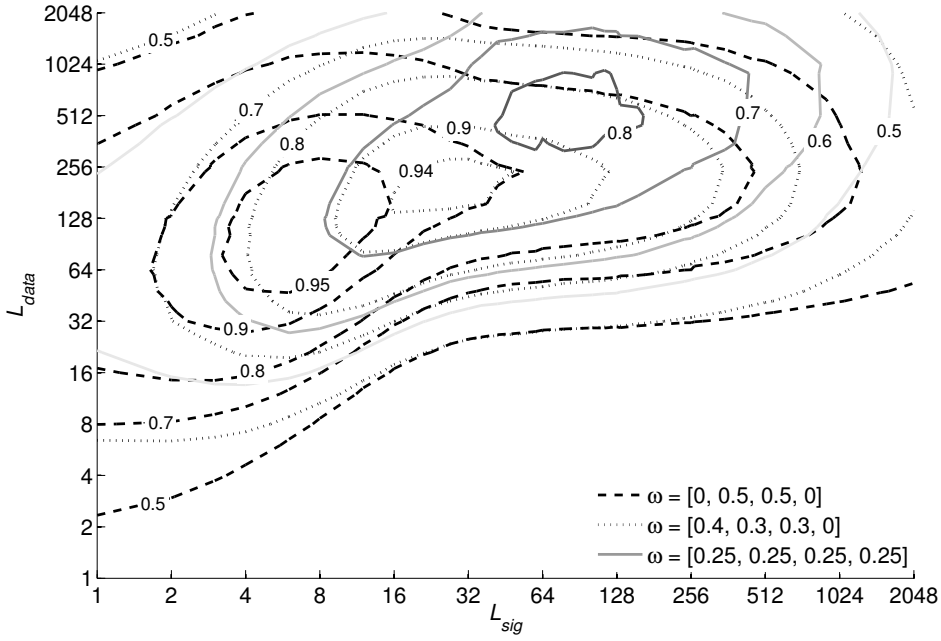

Figure 2.15. *Contour curves of mixed metrics using different weight vectors* $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_4]$*, where the four weights respectively refer to throughput, transmission efficiency, class efficiency and delivery delay. Curves are plotted as a function of the length of the training sequences used for signaling messages* (L_{sig}) *and PDUs* (L_{data})*,* $\lambda = 800\,\mathrm{pkt/s/node}$ *and* $N_A = 8$*.*

are deeply affected by the accuracy of channel estimation, and in particular by the value of σ_J^2, transmit pulses leading to high σ_J^2 explain the worse performance. For example, Figure 2.17 shows that by choosing $L_{\mathrm{sig}} = 64$ and $L_{\mathrm{data}} = 256$, a throughput of $25\,\mathrm{Mbps}$ and an efficiency of 0.47 can be achieved. Conversely, not only is this working point not available to the Sinc case (Figure 2.18), but the best tradeoff is also shifted toward $L_{\mathrm{sig}} = 128$, $L_{\mathrm{data}} = 256$, which indeed achieves both lower throughput and worse efficiency, in spite of longer training.

In order to make a better point about the tradeoffs involving all pulses discussed in Section 2.2.2.1, we plot in Figure 2.19 the convex hull of the scatterplot of all points having coordinates equal to the transmission efficiency–throughput pairs obtained for all pairs of L_{sig} and L_{data}. Figure 2.19 presents four such curves, related to a Gaussian pulse with $a = 1$, a Sinc pulse, a rectangular pulse, and an RC pulse with $\alpha = 0.5$. First, this figure allows to establish a ranking among pulses, in that it is easy to figure out whether the Pareto-optimal part of a curve (i.e., region of the curves closest to the upper-right corner of the graph)

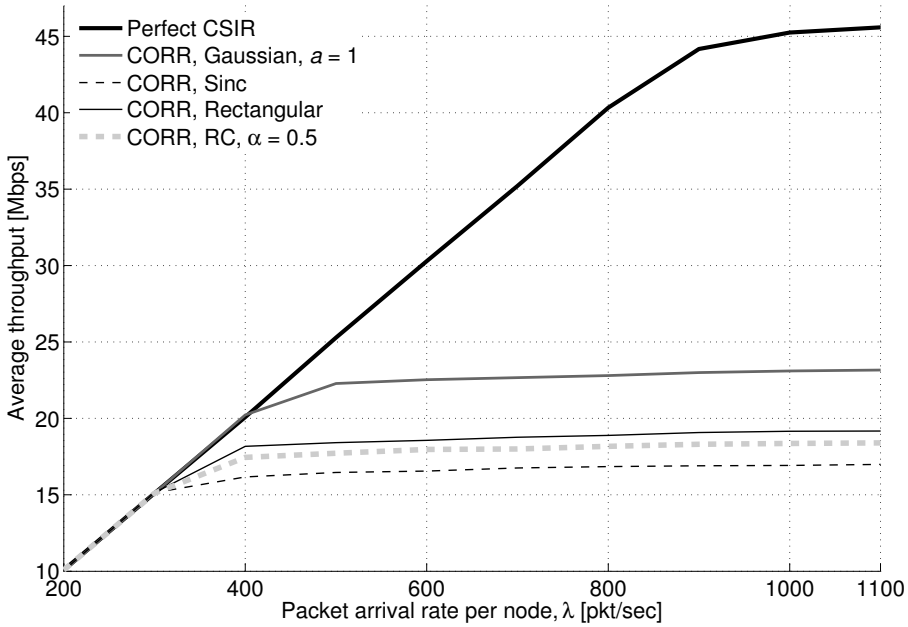

Figure 2.16. *Average throughput as a function of λ for different types of transmit impulses and with $L = 128$, compared to perfect channel estimation.*

is dominated by other curves or not: for example, the Gaussian pulse achieves again the best performance, whereas the Sinc pulse offers throughput-efficiency pairs with the lowest values. As a second outcome, Figure 2.19 also allows to pick the "best" configurations, i.e., those that allow to travel along the Pareto-optimal convex hull section and thus make a preference on throughput by sacrificing some efficiency, or vice-versa. In line with previous results, we observe that those points characterized by a high throughput are reached only through a higher L_{sig} when using pulses leading to higher σ_J^2, e.g., the Sinc and RC pulses, compared to the Rectangular and Gaussian pulses.

2.4 Conclusions

In this chapter, we have thoroughly assessed the impact of channel estimation accuracy on the performance of a MAC protocol for MIMO ad hoc networks and pointed out the main tradeoffs that arise. Channel estimation is a very relevant problem in wireless networks using advanced PHY techniques such as spatial multiplexing and layered multiuser detection; in addition, the presence of several simultaneous and asynchronous signals makes the

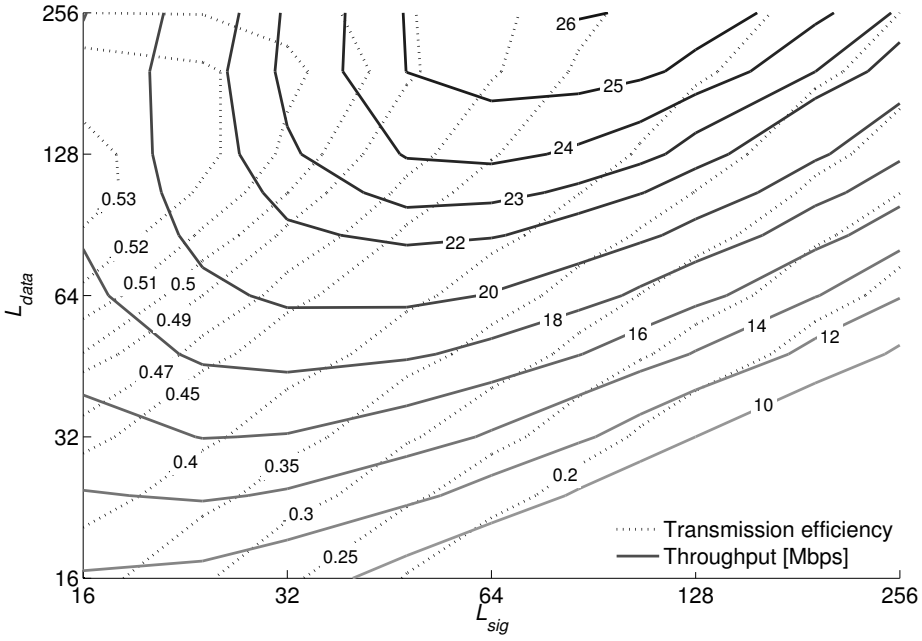

Figure 2.17. *Contour curves of throughput and transmission efficiency as a function of the length of the training sequences for signaling packets (L_{sig}) and for PDUs (L_{data}), for the Gaussian pulse, $a = 1$.*

problem more complicated than in traditional channel estimation. We have first described an analytical model for the computation of the statistics of the channel estimation error. Our analysis considers multiuser MIMO networks employing either correlator-based or MMSE channel estimators, and highlights the direct dependence of the channel estimation error on the instantaneous channel matrix, permitting an analytical evaluation of the interference directly in the channel estimation errors expression (see Equations (2.73) for MMSE and (2.74) for correlator-based channel estimator). Moreover, we showed the interplay between the transmit pulse shape and the variance of the channel estimates gathered by a correlator-based channel estimator. The bias of channel estimates due to cross-talk among different training sequences as well as among training sequences and data symbols has been quantified; the parameters of the inverse proportionality relationship tying the variance of this bias and the length of the training sequence have been quantified and related to pulse parameters such as the roll-off factor of RC pulses or the variance factor of Gaussian pulses.

The analyzed formulas for the statistics of the estimation errors have then been inserted into a simulator for MIMO ad hoc networks with a detailed MAC implementation and used to observe how PHY-level parameters, such as the length of the training sequences and the

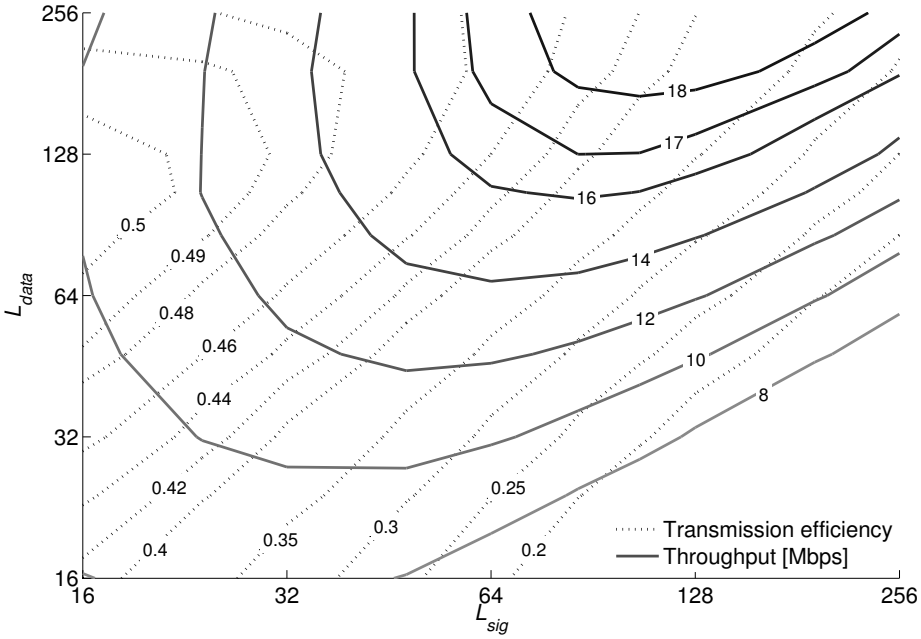

Figure 2.18. *Contour curves of throughput and transmission efficiency as a function of the length of the training sequences for signaling packets (L_{sig}) and for PDUs (L_{data}), for the Sinc pulse.*

number of antennas, affect higher-level metrics, such as throughput, efficiency, delay and success ratio. Our approach makes it possible to evaluate the robustness of networking protocols against channel estimation inaccuracies and interference, and to understand how to control the impact of channel errors in order to achieve prescribed tradeoff points among MAC-level metrics. Finally, we evaluate the impact of different pulse shapes on the performance of the network.

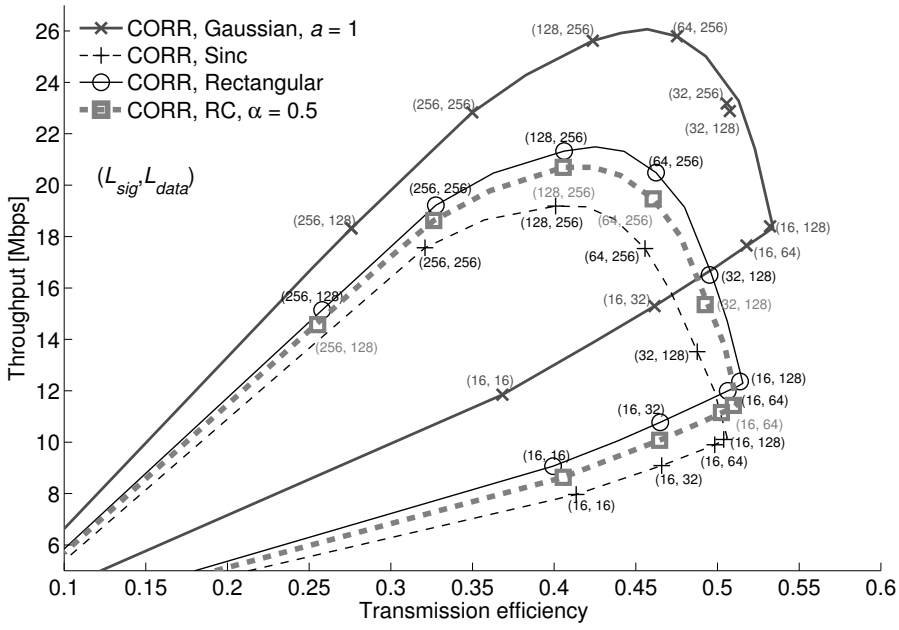

Figure 2.19. *Convex hull of the scatterplot of throughput-transmission efficiency points for different types of transmit impulses. The most important tradeoff choices on the hull are labeled for reference.*

Chapter 3

Cooperative Routing Techniques in Cognitive Radio Networks

The wireless medium opens up the possibility for the coexistence of different networks through appropriate interference management mechanisms. In particular, a scenario wherein a hierarchy exists between a "primary" network, whose performance should be guaranteed, and a "secondary" network, whose nodes must respect strict requirements so as not to interfere with the primary network, is attracting increasing attention under the label of "cognitive radio". One approach to cognitive radio prescribes the primary network to operate as if the secondary nodes were not present and the secondary nodes to limit their interference to the primary receivers below an acceptable level [40]. In this work, instead, we consider an alternative approach based on a combination of the principles of opportunistic routing and of the "spectrum leasing via cooperation" framework of [41,42]. In fact, the problem of scarce radio spectrum availability and the inefficiency of traditional fixed spectrum management schemes call for new communications paradigms for spectrum sharing [43]. Spectrum leasing is one such paradigm in which licensed users are allowed to lease portions of the spectrum to unlicensed users. In a standard implementation, spectrum leasing would be effected at a system level with "spectrum servers" allocating resources to secondary users [44, 45]. Moreover, secondary users would be charged for their use of the spectral resources. Instead, references [41,42] propose a novel approach in which spectrum leasing is performed locally and dynamically by primary devices and remuneration from secondary to primary takes place in the form of cooperation. In the approach of [41, 42], secondary nodes accept to

cooperate only if granted enough spectrum with respect to their desired QoS requirements.

This work proposes to implement spectrum leasing via cooperation in a multihop scenario by means of *opportunistic routing*.

Opportunistic routing is a well-known technique that aims at increasing the throughput of multihop networks over fading channels by exploiting the channel diversity offered by the availability of multiple possible next hops. In particular, selection of the next hop is made in an opportunistic fashion based on the channel conditions, and thus decoding outcomes, of previous transmissions of the given packet, thanks to appropriate feedback from the decoders. The next Section 3.1 studies the throughput advantages of opportunistic routing over conventional multihop routing for linear multihop wireless networks with Type-I HARQ and quasi-static Rayleigh fading channels. We recall that in Type-I HARQ error correction coding is used, but previous undecodable transmissions are discarded and detection is done only based on the current transmission [46]. The end-to-end throughput of opportunistic routing is derived using Markov chain tools and accounting for fading statistics. Both fixed-rate and optimal-rate transmissions are considered. Moreover, an investigation of the throughput using standard information-theoretic performance metrics for asymptotic SNR regimes is provided. Specifically, the multiplexing gain and energy efficiency (i.e., minimum energy per bit) of both opportunistic and multihop routing are analyzed. Numerical results are given to corroborate the analysis.

In Section 3.2, different policies that exploit spectrum leasing via opportunistic routing to different degrees are proposed. These policies are designed to span different operating points in the trade-off between gains in throughput and overall energy expenditure for the primary network. Moreover, two physical layer techniques are considered for multiplexing of the primary and secondary traffic at the secondary nodes, namely time division and Superposition Coding (SC). The optimality of the proposed routing techniques is proved, in terms of both throughput and primary energy consumption of SC over all possible multiplexing strategies. Finally, numerical results demonstrate the advantages of the proposed spectrum leasing solution based on opportunistic routing and the available trade-offs between primary throughput and energy consumption.[1]

[1]The material presented in this chapter has been published in [47–49].

3.1 Opportunistic Routing Analysis

Multihop routing is a conventional strategy used to forward a packet from source to destination through a number of hops in wireless ad hoc networks. Analysis of this class of protocols from a communication and information-theoretic standpoint has been pioneered by [50] under the assumption of links affected by additive white Gaussian noise only (i.e., no fading). This work shows that multihop transmission, with or without spatial reuse, performs very well in the *power-limited regime* (i.e., for *low*, SNRs), but becomes inefficient in the *bandwidth-limited regime* (i.e., for *high* SNRs). Analysis in the power-limited regime is performed using the standard measure of minimum energy per bit (over noise spectral density) required for a reliable transmission, $E_b/N_0|_{\min}$. Reference [51] extends these results to non-ergodic fading channels by studying the end-to-end outage probability, while ergodic fading is considered in [52]. Finally, [46] studies the end-to-end throughput for the same class of networks of [51] by assuming HARQ protocols to combat channel outages. However, no asymptotic SNR analysis is provided in [46].

As introduced above, a new routing paradigm has been introduced that potentially improves on standard multihop routing by exploiting the availability of multiple possible next hops in an adaptive manner: the next relay is selected based on the current channel conditions (and thus reception outcomes), as well as the distance to the destination [53–59]. These works focus on proposing different protocols to select the next hop based on alternative metrics.

In this section, we consider a linear multihop network over quasi-static fading channels as in [46, 51]. Our contributions are as follows:

(*i*) We derive the end-to-end throughput of opportunistic routing with Type-I HARQ;

(*ii*) We address the asymptotic regimes of high SNR (i.e., bandwidth-limited) and low SNR (i.e., power-limited) for both multihop and opportunistic routing, by studying the multiplexing gain and minimum energy per bit $E_b/N_0|_{\min}$ of the two schemes [60].

Throughout, we consider both the cases where the transmission rate is fixed and where the transmission rate can be optimized based on channel statistics.

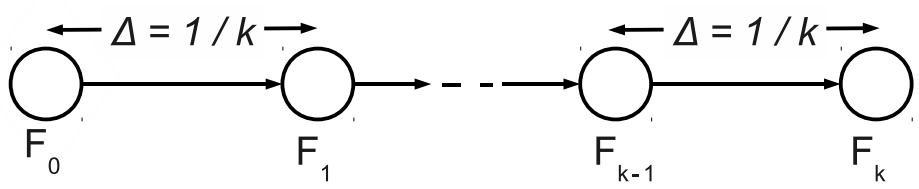

Figure 3.1. *A linear multihop network with k hops over quasi-static fading channels.*

3.1.1 System model

We consider a linear multihop network, where the source F_0 wants to communicate with the destination F_k at a normalized distance of one, possibly taking advantage of a set of $k-1$ relays, $F_1, ..., F_{k-1}$, equally spaced with inter-node distance $\Delta = 1/k$, as depicted in Figure 3.1. All nodes work in half-duplex mode (i.e., they cannot receive and transmit at the same time). Transmission is organized in blocks of n (complex) channel uses each. Recall that, from standard theory, n (complex) channel uses over a bandwidth of W Hz amount to n/W seconds in the absence of bandwidth expansion (see, e.g., [61]).

In each block, only one node (source or relay) is active, i.e., no spatial reuse is allowed. In the first block, the source F_0 transmits a packet of nR bits, where R is the rate of the first transmission in $bits/c.u.$ (channel use), or equivalently (in the absence of bandwidth expansion) in $bits/s/Hz$ (throughout the section, we will use $bits/c.u.$ or $bits/s/Hz$ interchangeably). In the following blocks, the source may retransmit the packet or else the relays, upon decoding previous transmissions, may forward or retransmit the packet, until the final destination F_k correctly receives it. We will discuss different transmission policies in the next sections. In general, it is assumed that transmission of each block is followed by some signaling, such as ACK/ Not ACK messages. When the current packet is successfully received, a new packet is transmitted by the source F_0, and the procedure repeats. Notice that this amounts to assuming the source is always backlogged.

Let $y_j(b,t)$ be the discrete-time (complex) baseband sample received by node j, $j \in \{F_1, \ldots, F_k\}$ during the b-th block, at channel use t, $t = 1, \ldots, n$:

$$y_j(b,t) = \left(\frac{k}{|j-i|}\right)^{\eta/2} h_{ij}(b)x_i(b,t) + z_j(b,t), \tag{3.1}$$

where $z_j(b,t)$ is the complex white Gaussian noise term with zero mean and power $\mathbb{E}[|z_j(b,t)|^2] = N_0$ and $x_i(b,t)$ is the symbol transmitted by the currently active node i, $i \in$

$\{F_0, \ldots, F_{k-1}\}$. We enforce the per-block power constraint

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left[|x_i(b,t)|^2\right] \leq P. \tag{3.2}$$

The channel coefficient between the i-th transmitter and the j-th receiver in (3.1), $h_{ij}(b)$, models quasi-static Rayleigh fading, i.e., it is a complex Gaussian random variable with zero mean and unit power, which is assumed to be constant within each block. Moreover, it is assumed to vary independently from block to block. Channels are known to the receivers but not to the transmitters. Finally, the term $(k/\left|j-i\right|)^{\eta/2}$ in (3.1) represents the path loss over the distance $d = \left|j-i\right|/k$ (i.e., $|j-i|$ hops) between transmitter i and receiver j with path-loss exponent η.

From (3.1), we define the SNR γ as the ratio between the maximum average power received by F_k directly from source F_0 and the noise power N_0, that is

$$\gamma = \frac{P}{N_0}. \tag{3.3}$$

With this definition, we have that γ/d^{η} is the SNR for a transmission that covers distance d.

Let $P_{out}(d)$ denote the probability that a certain packet transmitted by node i is not decoded correctly by node j with $d = \left|j-i\right|/k$. It is well known that this probability is given by the outage probability (see, e.g., [61]):

$$\begin{aligned} P_{out}(d) &= \Pr\left\{\log_2\left(1+|h_{ij}|^2\gamma d^{-\eta}\right) \leq R\right\} \\ &= 1-\exp\left(-\frac{2^R-1}{\gamma d^{-\eta}}\right), \end{aligned} \tag{3.4}$$

where we have used the fact that fading is Rayleigh.

We are interested in comparing the throughput, measured in $bits/s/Hz$, of multihop and opportunistic routing, both coupled with Type-I HARQ. The next section defines the throughput and evaluates it for these two strategies.

3.1.2 Throughput Analysis

The goal of this section is to determine the end-to-end throughput for both multihop and opportunistic routing. We define the throughput $T(k,R)$ as the average number of *successfully* delivered bits per second per Hz, given the total number of hops k and the transmission rate R. Using renewal theory, it is possible to show that (see, e.g., [62]):

$$T(k,R) = \frac{nR}{n\,\mathbb{E}[N]} = \frac{R}{\mathbb{E}[N]} \quad [bits/s/Hz], \tag{3.5}$$

where N is the number of transmission blocks necessary to transmit a given packet correctly, starting from the original transmission by the source F_0 until correct decoding at the destination F_k.

While definition (3.5) applies to the case where the transmission rate R is selected by the application and fixed, in many scenarios devices can tune their transmission rate. Therefore, we also consider an alternative definition of throughput $T^*(k)$, in which the transmission rate R is optimized:

$$T^*(k) = \sup_{R \geq 0} T(k, R). \tag{3.6}$$

Notice that there is a clear and well-known trade-off in the optimization of R: increasing R allows to send more bits to the destination (increasing the numerator in (3.5)), but also leads to an increased error probability and therefore to more transmissions (thus increasing the denominator in (3.5)). We also emphasize that optimization of rate R in (3.6) only requires knowledge of the channel statistics (and not of the instantaneous values) at the source.

The rest of the section derives the end-to-end throughput (3.5) and (3.6) for both multihop and opportunistic routing with Type-I HARQ.

3.1.2.1 Multihop Routing

With multihop routing, each packet goes through all the k hops: the source F_0 retransmits the packet until relay F_1 decodes it successfully; then, the link between F_1 and F_2 is operated in the same way, and so on, until the destination F_k decodes correctly. Assuming Type-I HARQ on each hop, previous retransmissions are discarded at each receiver and the probability of outage for any transmission is given by (3.4) with $d = \Delta$. Thus, we easily evaluate the throughput (3.5) with fixed rate R as:

$$\begin{aligned} T_{mh}(k, R) &= \frac{R}{\mathbb{E}[N]} = \frac{R}{k}\left(1 - P_{out}(\Delta)\right) \\ &= \frac{R}{k} \exp\left(-\frac{2^R - 1}{\gamma k^\eta}\right), \end{aligned} \tag{3.7}$$

where the number of retransmissions N is a geometric random variable with success probability given by $(1 - P_{out}(\Delta))$. Moreover, note that the result in Eq. (3.7) can also be interpreted by observing that each slot contains a successful transmission with probability $1 - P_{out}(\Delta)$, and since all transmitters use the same fraction of slots on average, the probability that a successful slot corresponds to a packet reaching the final destination is $1/k$.

For the throughput with optimized rate, we easily extend the result in [63] for a single-hop network ($k = 1$) to obtain the following Lemma.

Lemma 3.1.1. *The end-to-end throughput for a multihop scheme with k hops and optimized rate is given by*

$$T^*_{mh}(k) = \frac{R^*_{mh}}{k} \exp\left(-\frac{2^{R^*_{mh}} - 1}{\gamma k^\eta}\right), \tag{3.8}$$

with

$$R^*_{mh} = \frac{\mathcal{W}_0\left(k^\eta \gamma\right)}{\log_{\mathrm{e}} 2}, \tag{3.9}$$

where $\mathcal{W}_0(z)$, known as the Lambert $\mathcal{W}$ function, is the unique solution of the equation

$$\mathcal{W}(z)\, \mathrm{e}^{\mathcal{W}(z)} = z, \quad \textit{for} \quad z > 0. \tag{3.10}$$

3.1.2.2 Opportunistic Routing

With opportunistic routing, after each (re)transmission of the current packet, all decoding nodes issue an ACK message. If the destination is among the decoding nodes, transmission of the current packet is terminated and the next packet is transmitted by the source F_0. Otherwise, the transmitter for the next hop is selected opportunistically as the decoding relay that is the closest to the destination. The exact mechanism as to where and how the decision is made is not of concern here, and has been studied in [55–58]. With opportunistic routing, the average number $\mathbb{E}[N]$ of hops per packet can potentially be greatly reduced with respect to standard multihop routing, thus boosting the throughput (3.5) and (3.6).

To derive the throughput of opportunistic routing, we use the theory of Markov chains. Specifically, there are $k + 1$ states in the chain, one for each node in the linear network, with state S_0 referring to scenarios where the current packet is at the source F_0, states S_i, $i = 1, \ldots, k-1$, similarly defined, and S_k representing the state where the destination has successfully decoded. Recalling that we assume Type-I HARQ, the current transmitter retransmits the packet until at least one of the downstream nodes has successfully decoded. Based on this, the transition matrix can be found as

$$\mathbf{P} = \begin{bmatrix} P_{S,S}(0,0) & \ldots & P_{S,S}(0,k) \\ 0 & \ddots & \vdots \\ 0 & 0 & P_{S,S}(k,k) \end{bmatrix}, \tag{3.11}$$

where $P_{S,S}(i,j)$ is the probability that, given that the current state is i (i.e., the transmitter is node $i = 0, ..., k-1$), the next state is j (i.e., the next relay is j for $j = i, i+1, ..., k-1$ or the destination has decoded for $j = k$). The first k states are transient and the last state, corresponding to the packet having been received at the destination, is absorbing. The transition probabilities are given by:

$$P_{S,S}(k,k) = 1; \tag{3.12}$$

$$P_{S,S}(i,j) = (1 - P_{out}((j-i)\Delta)) \prod_{\ell=j+1}^{k} P_{out}((\ell-i)\Delta),$$

$$i = 0, \ldots, k-1, \quad j = i, \ldots, k; \tag{3.13}$$

$$P_{S,S}(i,j) = 0, \quad \text{otherwise.} \tag{3.14}$$

Proposition 3.1.2 (End-to-end Throughput of Opportunistic Routing)**.** *The throughput (3.5) for fixed transmission rate R of opportunistic routing is given by*

$$T_{opp}(k,R) = \frac{R}{v_0}, \tag{3.15}$$

where v_0 is the first entry of vector $\mathbf{v} = [v_0, \ldots, v_{k-1}]$, which is evaluated as

$$\mathbf{v} = (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}, \tag{3.16}$$

where $\mathbf{1}$ is a $(k-1) \times 1$ vector with all entries equal to 1, matrix $\mathbf{Q}$ is obtained from $\mathbf{P}$ by removing the last row and the last column and $\mathbf{I}$ is the $(k-1) \times (k-1)$ identity matrix.

Proof. The proposition follows from the theory of absorbing Markov chains (see, e.g., [64, Section. 4.5]). Let v_i be the expected number of steps before the chain is absorbed given that the chain starts in state S_i, $i = 0, \ldots, k-1$. Then, from standard first-step analysis, we have the set of equations (which is recursive, due to the triangular form of matrix (3.11)):

$$v_i = 1 + \sum_{j \neq k} P_{S,S}(i,j)\, v_j \;\text{ with }\; i \neq k, \tag{3.17}$$

Equation (3.16) is readily obtained from the matrix formulation for such set of equations (see, e.g., [64, Section. 4.5]). □

Remark 3.1.1. When k is large, closed-form (i.e., non-recursive) expressions for (3.15) are very involved. Here, we report the throughput for $k = 2$ (the expressions for $k = 3$ and $k = 4$ are

reported in Appendix A):

$$T_{opp}(2,R)=\frac{R\left[1-\left(1-\mathrm{e}^{-\frac{2^R-1}{2^\eta\gamma}}\right)\left(1-\mathrm{e}^{-\frac{2^R-1}{\gamma}}\right)\right]}{2-\mathrm{e}^{-\frac{2^R-1}{\gamma}}}. \tag{3.18}$$

Unfortunately, a closed-form expression for $T^*_{opp}(k)$ appears to be hard to find. In the next section, we shed some light on the performance analysis by focusing on asymptotic SNR regimes for both multihop and opportunistic routing.

3.1.3 Asymptotic Analysis

In this section, we focus on the asymptotic regimes of high and low SNR.

3.1.3.1 High SNR (Bandwidth-Limited Regime)

Consider first the case where the SNR is large (i.e., the bandwidth-limited regime). For a fixed transmission rate R, it is meaningful to consider the value of the throughput (3.5) as $\gamma\to\infty$, since the throughput remains necessarily finite, being bounded by R. However, when optimizing the transmission rate R, the throughput (3.6) scales with SNR, so that it is more meaningful to study the multiplexing gain, defined as (see, e.g., [61]):

$$\lim_{\gamma\to\infty}\frac{T^*(k)}{\log_2\gamma}. \tag{3.19}$$

Proposition 3.1.3 (High-SNR Characterizations)**.** *The high-*SNR *throughput (3.5) with fixed transmission rate R is given by:*

$$\lim_{\gamma\to\infty}T_{mh}(k,R)=\frac{R}{k}, \tag{3.20}$$

for multihop routing, whereas for opportunistic routing we have

$$\lim_{\gamma\to\infty}T_{opp}(k,R)=R. \tag{3.21}$$

When rate R is optimized, the multiplexing gain of throughput (3.6) is

$$\lim_{\gamma\to\infty}\frac{T^*_{mh}(k)}{\log_2\gamma}=\frac{1}{k}, \tag{3.22}$$

for multihop routing, whereas for opportunistic routing we have the bounds

$$\lim_{\gamma\to\infty}\frac{T^*_{opp}(k)}{\log_2\gamma}=1. \tag{3.23}$$

Proof. The results (3.20) and (3.21) follow easily from the fact that $P_{out}(d) \to 0$ for a fixed rate R and any d. Specifically, for opportunistic routing, this implies that direct transmission from source to destination is successful with high probability.

The multiplexing gain in (3.22) is derived from eq. (9) in [63], conveniently adapted.

Finally, to show (3.23), we prove that:

$$1 - \epsilon \leq \lim_{\gamma \to \infty} \frac{T^*_{opp}(k)}{\log_2 \gamma} \leq 1, \tag{3.24}$$

where $\epsilon > 0$ is arbitrarily small. The upper bound follows from cut-set arguments: the throughput of opportunistic routing cannot be larger than the throughput of a system where relays and destination fully cooperate for decoding. Moreover, the throughput of the latter system is upper bounded by its ergodic capacity, i.e., by

$$\mathbb{E}[\log_2(1 + \sum_{j=1}^{k} |h_{0j}|^2 \gamma (k/j)^\eta)], \tag{3.25}$$

whose multiplexing gain is one [61]. This follows immediately from the definition of throughput, noting that the latter can be written as

$$\mathbb{E}[R \cdot 1(\log_2(1 + \sum_{j=1}^{k} |h_{0j}|^2 \gamma (k/j)^\eta)) < R)] \leq \mathbb{E}[\log_2(1 + \sum_{j=1}^{k} |h_{0j}|^2 \gamma (k/j)^\eta))], \tag{3.26}$$

where $1(\cdot)$ is the indicator function. To obtain the lower bound in (3.24) it is enough to consider the following suboptimal transmission scheme: set the rate at

$$R(\gamma) = \log_2 \gamma^{(1-\epsilon)} \tag{3.27}$$

and consider only the link between source and destination. This scheme clearly sets a lower bound on the achievable throughput, namely

$$T^*_{opp}(k) \geq R(\gamma) \exp\left(-\frac{2^{R(\gamma)} - 1}{\gamma}\right). \tag{3.28}$$

Letting ϵ be arbitrarily small in (3.24) we conclude the proof. □

Discussion of the results of Proposition 2 is postponed to Section 3.1.4.

3.1.3.2 Low SNR (Power-Limited Regime)

We now focus on the energy efficiency of multihop and opportunistic routing. Specifically, we evaluate the minimum energy per bit required for reliable transmission, which is

defined as [60]

$$\left.\frac{E_b(k,R)}{N_0}\right|_{\min} = \inf_{\gamma} \frac{\gamma}{T(k,R)} \tag{3.29}$$

for transmission with fixed rate R and

$$\left.\frac{E_b(k)^*}{N_0}\right|_{\min} = \inf_{\gamma} \frac{\gamma}{T^*(k)} \tag{3.30}$$

for transmission with optimized rate. As already explained in Section 3.1, this is a standard measure on the performance of transmission schemes [60] and has been considered in related routing scenarios in [50,51].

Proposition 3.1.4 (Energy Efficiency of Multihop Routing)**.** *The minimum energy per bit for multihop routing with fixed transmission rate is given by:*

$$\begin{aligned}\left.\frac{E_b(k,R)}{N_0}\right|_{\min,mh} &= \inf_{\gamma} \frac{\gamma}{T_{mh}(k,R)} \\ &= \mathrm{e}\, k^{1-\eta} \frac{2^R-1}{R},\end{aligned} \tag{3.31}$$

whereas for optimized rate, we have:

$$\begin{aligned}\left.\frac{E_b(k)^*}{N_0}\right|_{\min,mh} &= \inf_{\gamma} \frac{\gamma}{T^*_{mh}(k)} \\ &= \mathrm{e}\, k^{1-\eta} \log_e 2.\end{aligned} \tag{3.32}$$

Proof. The first equality, (3.31), is obtained by noticing that the convexity of the exponential function implies that $\gamma/T_{mh}(k,R)$ is also convex, and therefore the minimum is found where its derivative is zero. The optimal value of the SNR, which maximizes energy efficiency, is

$$\gamma = (2^R-1)k^{-\eta}. \tag{3.33}$$

For (3.32) we first note that the quantity

$$\frac{\gamma}{T^*_{mh}(k)} = \frac{k\gamma \log_{\mathrm{e}} 2}{\mathcal{W}_0(k^\eta\gamma)} \exp\left(\frac{1}{\mathcal{W}_0(k^\eta\gamma)} - \frac{k^{-\eta}}{\gamma}\right), \tag{3.34}$$

where we have exploited the equation $\exp(\mathcal{W}(z)) = z/\mathcal{W}(z)$, is an increasing function of the SNR and therefore:

$$\inf_{\gamma} \frac{\gamma}{T^*_{mh}(k)} = \lim_{\gamma\to 0} \frac{\gamma}{T^*_{mh}(k)}. \tag{3.35}$$

We then expand (3.34) using the first two terms of the Taylor expansion of

$$\mathcal{W}_0(z) = \sum_{n=1}^{\infty} z^n(-n)^{n-1}/n! \tag{3.36}$$

(see, e.g. [63]) to obtain the approximation for small SNR

$$\frac{\gamma}{T^*_{mh}(k)} \cong \frac{k \log_{\mathrm{e}} 2}{k^\eta - k^{2\eta}\gamma} \exp\left(\frac{1}{1-k^\eta\gamma}\right).$$

The proof is concluded by evaluating the limit (3.35). □

Remark 3.1.2. We emphasize that, from the proof given above, the value of SNR that maximizes the energy efficiency (3.29) for multihop routing with fixed transmission rate is $\gamma = (2^R - 1)k^{-\eta}$, whereas if one allows optimal rate selection (3.30) the optimal $\gamma \to 0$. This is due to the well-known fact that energy efficiency is maximized at vanishing spectral efficiencies, that is $R \to 0$ and $\gamma \to 0$, if one can optimize the transmission rate [60].

Remark 3.1.3 (Wideband slope for multihop routing). Beside the minimum energy per bit $E_b/N_0|_{\min}$, reference [60] defines also the slope $\mathcal{S}_0$ of the spectral efficiency at $E_b/N_0|_{\min}$ in order to provide a more complete description of the rate behavior in the power-limited regime. This can be easily calculated for multihop routing with optimized rate and is given by [60]:

$$\begin{aligned}\mathcal{S}_{0,mh} &= -\frac{2\left[\dot{T}^*_{mh}(k)\Big|_{\gamma=0}\right]^2}{\ddot{T}^*_{mh}(k)\Big|_{\gamma=0}} \\ &= \frac{2}{\mathrm{e}\,k} \quad [bits/s/Hz/(3dB)],\end{aligned} \tag{3.37}$$

where $\dot{T}^*_{mh}(k)|_{\gamma=0}$ and $\ddot{T}^*_{mh}(k)|_{\gamma=0}$ denote the first and second derivative of the end-to-end throughput curve evaluated in $nats/s/Hz$ with $\gamma = 0$. Note that the slope of the throughput of the multihop (and opportunistic) routing for fixed rate turns out not to be well-defined due to the fact that (see Remark 3.1.2) the energy $E_b/N_0|_{\min}$ is not attained for vanishing throughput (see also discussion in [60]).

General expressions for energy efficiency in the case of opportunistic routing are difficult to obtain. In the rest of section, we consider some approximation for $k = 2$.

3.1.3.2.1 Energy-Efficiency Approximations for Opportunistic routing Here we consider analytical approximations for the minimum energy per bit with fixed rate and $k = 2$, $E_b(2,R)/N_0|_{\min,opp}$, and optimized rate, $E_b(2)^*/N_0|_{\min,opp}$ (recall (3.29) and (3.30), respectively). The approximations are based on suboptimal choices for the optimal SNR in (3.33)

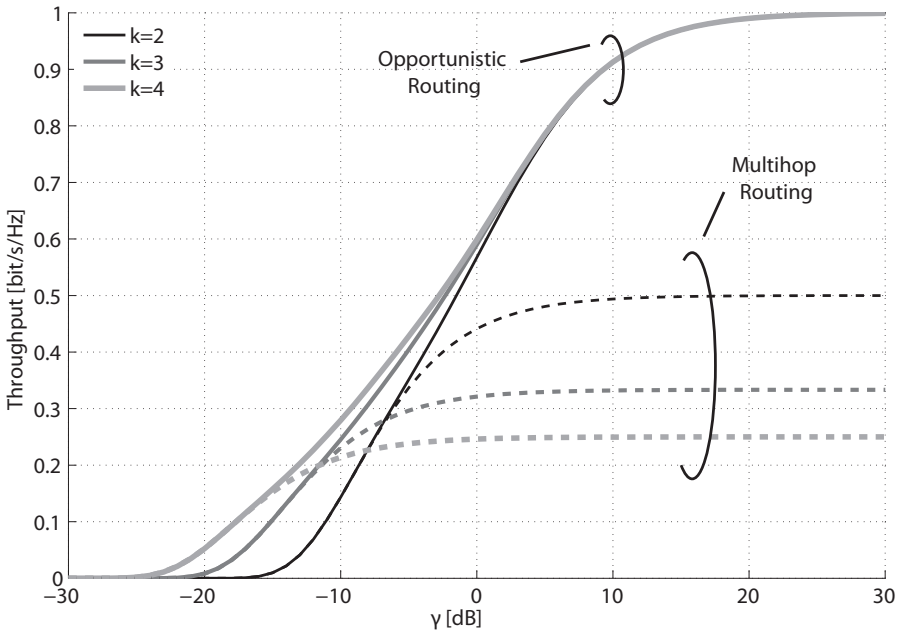

Figure 3.2. *End-to-end throughput of opportunistic routing, $T_{opp}(k,R)$ (solid lines), and multihop routing, $T_{mh}(k,R)$ (dashed lines), versus* SNR*, γ, for transmission rate $R=1$ $[bits/s/Hz]$ and $\eta=3$.*

and optimal rate in (3.9), that are expected to be close-to-optimal. The key observation is that, for low SNR, opportunistic routing will be often forced to use all the hops like multihop routing. Therefore, setting the optimal SNR and rate to the corresponding optimal SNR and rate for multihop routing (namely $\gamma=(2^R-1)k^{-\eta}$ and $R^*_{mh}=\mathcal{W}_0\left(k^\eta\gamma\right)/\log_{\mathrm{e}}2$) leads to potentially good approximations. Using these choices, we obtain

$$\begin{aligned}\left.\frac{E_b(2,R)}{N_0}\right|_{\min,opp} &= \inf_{\gamma}\frac{\gamma}{T_{opp}(2,R)}\\ &\simeq \frac{\mathrm{e}\left(2\,\mathrm{e}^{2\eta}-1\right)\left(2^R-1\right)}{2^\eta\left(\mathrm{e}+\mathrm{e}^{2\eta}-1\right)R}\end{aligned} \tag{3.38}$$

for fixed rate, and

$$\begin{aligned}\left.\frac{E_b(2)^*}{N_0}\right|_{\min,opp} &= \inf_{\gamma}\frac{\gamma}{T^*_{opp}(2)}\\ &\simeq \frac{\mathrm{e}\left(2\,\mathrm{e}^{2\eta}-1\right)\log_{\mathrm{e}}2}{2^\eta\left(\mathrm{e}+\mathrm{e}^{2\eta}-1\right)},\end{aligned} \tag{3.39}$$

for optimized rate. Our numerical results show that indeed these are good approximations.

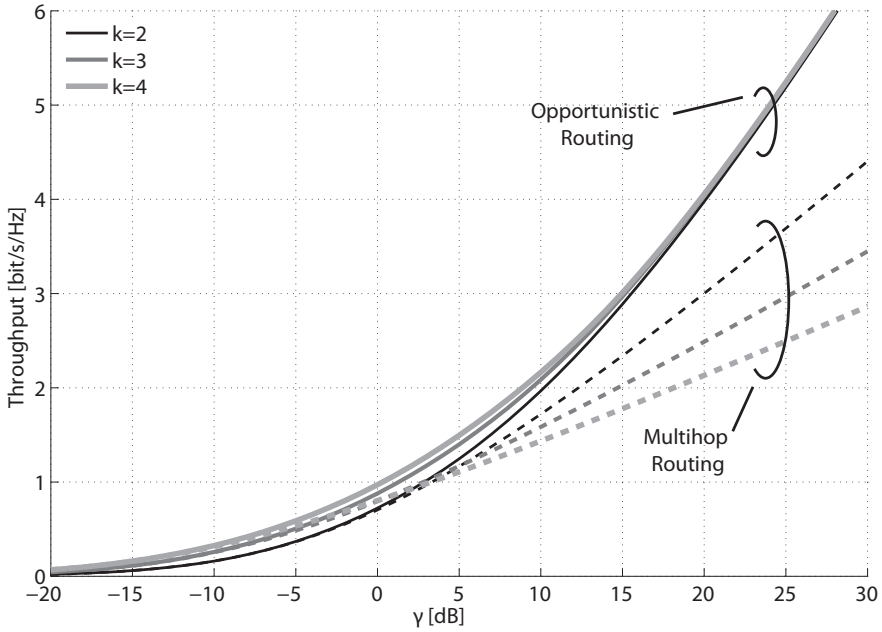

Figure 3.3. *End-to-end throughput of opportunistic routing, $T^*_{opp}(k,R)$ (solid lines), and multihop routing, $T^*_{mh}(k,R)$ (dashed lines), versus* SNR, γ, *for optimized rate and $\eta = 3$.*

3.1.4 Numerical Results

Here we provide some numerical results to corroborate the analysis above. We first study the end-to-end throughput versus SNR for both fixed rate (Figure 3.2) and optimized rate (Figure 3.3). Starting with fixed rate, Figure 3.2 shows the end-to-end throughput versus SNR for different numbers of hops $k = \{2, 3, 4\}$, path loss $\eta = 3$ and transmission rate $R = 1$ $[bits/s/Hz]$. It is seen that for sufficiently large SNR (i.e., $\gamma \gtrsim 10\ dB$), the asymptotic results of Proposition 3.1.3 apply. In particular, as per (3.20) and (3.21), while opportunistic routing is able to attain the maximum throughput of $T_{opp}(k,1) = 1$ $[bits/s/Hz]$ for every k, multihop routing shows the well-known performance degradation in the bandwidth-limited regime (recall beginning of Section 3.1), which amounts here to a factor of k, i.e., $T_{mh}(k,1) = 1/k$ $[bits/s/Hz]$. Turning to the performance with optimized rates, Figure 3.3 confirms the advantages of opportunistic routing and validates the results in Proposition 3.1.3: when the SNR is large enough (here, $\gamma \gtrsim 20\ dB$), the throughput of opportunistic routing increases with a slope independent of k and larger by a factor of k with respect to multihop routing.

We now explicitly consider the throughput ratio $\rho = T^*_{mh}(k)/T^*_{opp}(k)$ between multihop and opportunistic routing in Figure 3.4 for different path loss exponents $\eta = \{2.5, 3, 4\}$ and

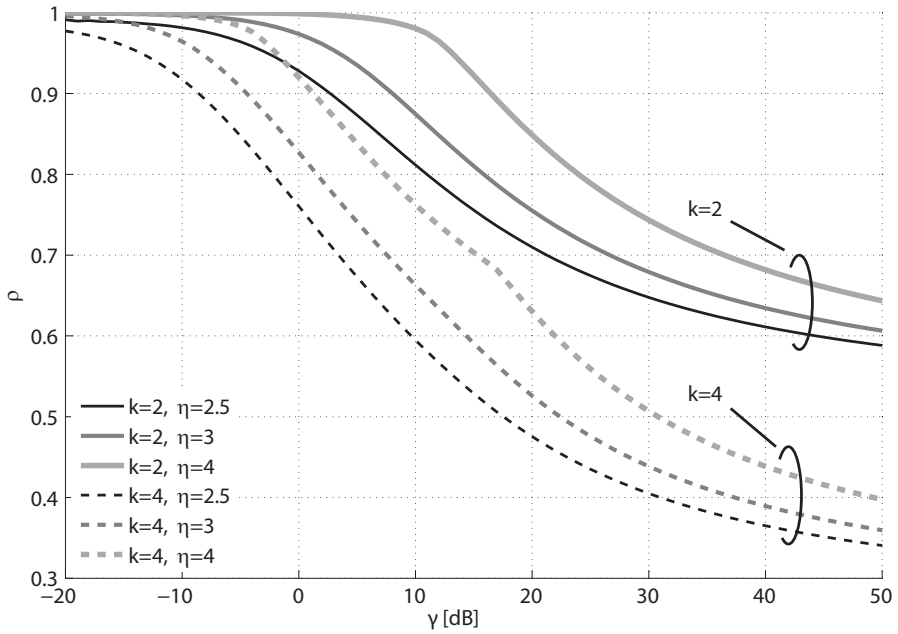

Figure 3.4. *Ratio $\rho = T^*_{mh}(k)/T^*_{opp}(k)$ between the throughput of multihop and opportunistic routing versus* SNR*, γ, for $k = \{2, 4\}$ and $\eta = \{2.5, 3, 4\}$ and optimized rate.*

$k = \{2, 4\}$ versus SNR for optimized rate. The figure points out that opportunistic routing has better gains for smaller path loss exponents ($\eta = 2.5$), due to the larger number of relays that are potentially reachable at each transmission and may thus serve as next hop.

Finally, Figure 3.5 focuses on the energy efficiency of the considered schemes by showing the throughput versus the energy per bit over noise power spectral density, i.e., $E_b/N_0 = \gamma/T^*(k)$ (recall (3.30)), for $k = \{2, 3, 4\}$, path loss $\eta = 3$ and optimized rate. First of all, we note that the simulation results confirm the values analytically derived in (3.32) for the minimum energy per bit ($E_b(k)^*/N_0|_{\min,mh} = \{-3.27, -6.79, -9.29\}\ dB$ for $k = \{2, 3, 4\}$) and in (3.37) for the slope $\mathcal{S}_{0,mh}$ of the multihop case (see Figure 3.5). We also note that the approximation given by (3.39) is close to the value found in the simulation, which uses a brute-force approach to find the optimum rate R^*_{opp} ($E_b(2)^*/N_0|_{\min,opp} \simeq -3.27\ dB$)[2]. It is also concluded that opportunistic routing fails to outperform multihop routing in the power-limited regime, being unable to exploit the path diversity, unlike in the bandwidth-limited regime.

[2]Also for the fixed rate case the simulation results are consistent with (3.38).

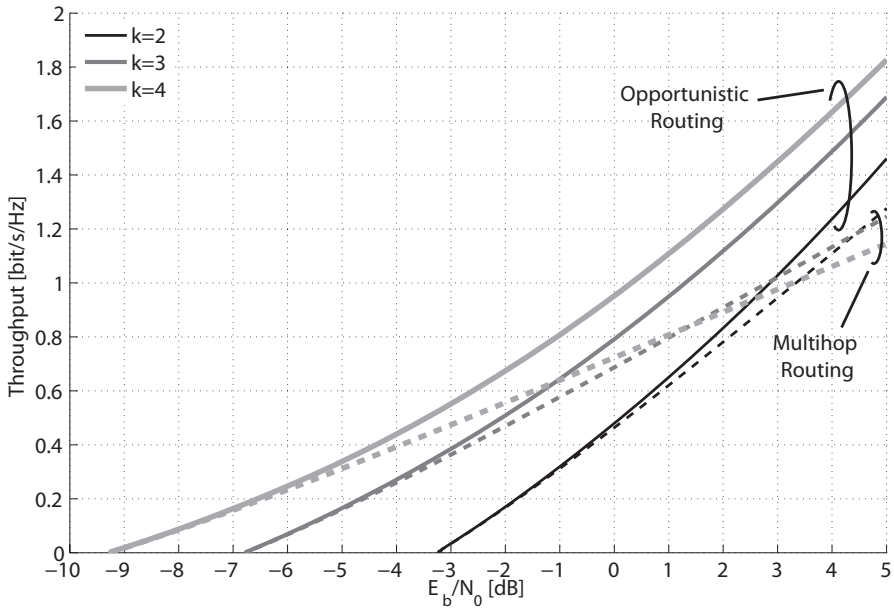

Figure 3.5. *End-to-end throughput of opportunistic routing, $T^*_{opp}(k)$ (solid lines), and multihop routing, $T^*_{mh}(k)$ (dashed lines), versus E_b/N_0 for $\eta = 3$ and optimized rate.*

3.2 Spectrum Leasing via Cooperative Opportunistic Routing

As we introduced at the beginning of this chapter, the main idea of this remaining section is that secondary nodes may serve as potential hops for a primary network that routes packets based on opportunistic routing, under the condition that the secondary nodes are leased enough spectral resources to satisfy their QoS requirements as well. In other words, secondary nodes may accept to serve as next hops, but, in return, if selected, they require to be leased part of the spectrum for their own transmissions. The primary network, thanks to spectrum leasing via cooperative opportunistic routing, may gain on two fronts:

(*i*) *Throughput*, due to the improved multiuser diversity in the selection of the next hop that is afforded by the availability of secondary nodes;

(*ii*) *Primary energy consumption*, due to the fact that transmissions can be delegated to the secondary network.

This work studies the trade-off between these two metrics by proposing protocols that work at different operating points of this trade-off. The analysis accounts for different possible multiplexing techniques at the secondary nodes, namely Time Division Multiplexing (TDM) and SC (see, e.g., in [61,65]). It is proved that SC is optimal in terms of both throughput and primary energy consumption over all possible multiplexing strategies.

3.2.1 System Model and Multiplexing Techniques

In Figure 3.6 we show a primary and a secondary network that coexist via spectrum leasing. The aim of the *primary source* P_0 is to communicate with the *primary destination* P_k, at a normalized distance of one, possibly taking advantage of multihop routing through two sets of additional nodes placed along two parallel linear geometries with vertical distance Δ_{V}. Both sets are composed of $k-1$ nodes: the first one is formed by *primary nodes*, denoted by $P_1, \ldots, P_{k-1}$ whose only role is that of forwarding information from P_0 to P_k; the second set of nodes, instead, consists of *secondary (unlicensed) nodes* $S_1, \ldots, S_{k-1}$ that can access the channel only if spectrum is leased by the primary network, as will be discussed below. Primary nodes have $\Delta_{\mathrm{H}} = 1/k$ inter-node distance. The secondary nodes are aligned with the primary nodes, and thus have the same inter-node distance. More generally, we will consider a *partial secondary deployment* in which only one every α secondary nodes in Figure 3.6

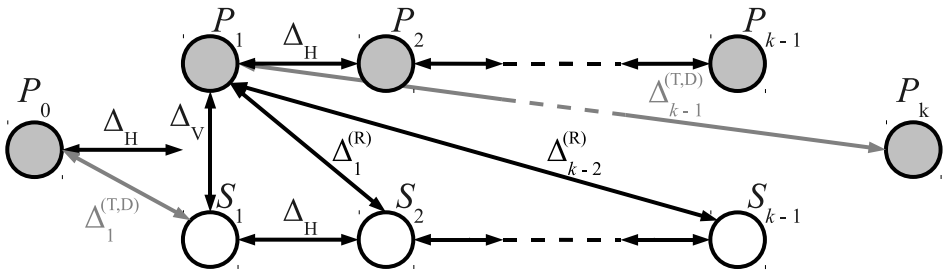

Figure 3.6. *A primary linear multihop network (grey circles) with k hops and a secondary network (white circles) aligned with respect to primary relay nodes.*

is active so that the number of secondary nodes is $k/\alpha - 1$ (assumed to be an integer) with inter-node distance $\alpha\Delta_{\mathrm{H}}$. For simplicity, where not stated otherwise, we will assume $\alpha = 1$ in the following. This work relies on geometrical simplifying assumptions with the objective of having both a solvable theoretical model and an insightful analysis of spectrum leasing via cooperative routing techniques. More general network topologies will be considered in Appendix C.

As in Section 3.1.1, all devices considered cannot receive and transmit at the same time, i.e., half-duplex mode is considered. Transmission is organized in blocks of n (complex) channel uses each, where only one node is active (i.e., no spatial reuse is allowed). In the first block, the source P_0 transmits a packet of $n\,\mathrm{R_P}$ bits, where $\mathrm{R_P}$ is the transmission rate of the original (primary) transmission in $bits/s/Hz$. In the following blocks, retransmissions take place, if necessary, according to a Type-I HARQ process (i.e., error correction coding is used, but previous undecodable transmissions are discarded and detection is only based on the current transmission [46]). Retransmissions in each block may be done by the source P_0, or by the primary relays $P_1, \ldots, P_{k-1}$ or secondary nodes $S_1, \ldots, S_{k-1}$, as long as the latter have correctly decoded in the previous block. After the packet is correctly delivered to the destination, the primary source P_0 transmits a new packet and the process repeats.

3.2.1.1 Opportunistic Routing and Spectrum Leasing

The next hop decisions in the network are made by primary nodes based on the feedback received at the end of the previous block from all nodes that have successfully received the packet. Thanks to this information (i.e., which nodes, both primary and secondary, were able to decode the packet transmitted in the previous block), the primary network

can schedule transmissions in an *opportunistic* fashion based on the channel conditions, and thus the decoding outcomes, in previous blocks. The exact mechanism as to where and how the decision is made is not of concern here, and has been studied in [56]. Channel resources for feedback allocation will not be included in the analysis (as usually considered in the literature on opportunistic routing), assuming that the feedback information will be available to the primary nodes as required by the different protocols to be introduced below.

As discussed above, secondary nodes may serve as relays for the current primary packet. However, secondary nodes follow the *spectrum leasing via cooperation* (relaying) principle stated in [41]: in fact, they do not cooperate for free, but they accept to serve as relays only if they are granted sufficient resource for their own traffic as well.

In this section we evaluate TDM and SC as two techniques to multiplex primary and secondary traffic. Specifically, in the TDM approach primary and secondary data are multiplexed by the secondary node S_i in the time domain: a portion β of the time slot is used to forward the primary packet and the remaining part to transmit own data in the secondary network. SC, instead, is a physical layer technique in which a transmitter can simultaneously send independent messages to multiple receivers. Specifically, the secondary node transmits a signal obtained by superimposing the packet carrying its own data and the primary packet intended for the primary network. In particular, the secondary transmitter encodes and modulates both packets at the selected rates and scales the *power* of each modulated symbol to match the chosen power split, assigning a portion $0 \leq \psi \leq 1$ to the primary packet and $1 - \psi$ to its own packet. Finally, the complex baseband symbols (or waveforms with pass band representation) are added to obtain the transmitted signal. So, the shared resource in this approach is the *transmitted power* [65]. Both the time fraction β for TDM and the power fraction ψ for SC are selected by the primary and secondary nodes so as to satisfy their own QoS requirements in terms of rate and reliability, which are expressed in terms of outage probabilities as explained below after introducing the signal model.

3.2.1.2 Signal Model and Secondary QoS Requirements

Considering a transmission from node $N_i \in \{P_0, \ldots, P_{k-1}, S_1, \ldots, S_{k-1}\}$, let $y_{N_i N_j}(b, t)$ denote the discrete-time (complex) baseband sample received by node N_j, with

$N_j \in \{P_1, \ldots, P_k, S_1, \ldots, S_{k-1}\}$, during the b-th block, at channel use t, $t = 1, \ldots, n$:

$$y_{N_i N_j}(b,t) = d_{N_i N_j}^{-\eta/2} h_{N_i N_j}(b) x_{N_i}(b,t) + z_{N_j}(b,t), \tag{3.40}$$

where $d_{N_i N_j}^{-\eta/2}$ is the path loss between the N_i-th transmitter and the N_j-th receiver with power path-loss exponent η. The distance $d_{N_i N_j}$ can assume three forms (see also Figure 3.6):

(*a*) if both nodes N_i and N_j lie on the same line, i.e., for transmissions between primary relays or secondary nodes, $d_{N_i N_j}$ is given by $|j - i|\,\Delta_{\mathrm{H}}$;

(*b*) if N_i is the source P_0 and the destination is a primary or secondary relay or N_i is a relay and N_j is the destination P_k, $d_{N_i N_j}$ is equal to $\Delta_{|j-i|}^{(\mathrm{T,D})}$, where $\Delta_a^{(\mathrm{T,D})} = \sqrt{(a\Delta_{\mathrm{H}})^2 + (\Delta_{\mathrm{V}}/2)^2}$;

(*c*) finally, if the transmission is between two relays, one in the primary network and one in the secondary, $d_{N_i N_j} = \Delta_{|j-i|}^{(\mathrm{R})}$ with $\Delta_a^{(\mathrm{R})} = \sqrt{(a\Delta_{\mathrm{H}})^2 + \Delta_{\mathrm{V}}^2}$.

The channel coefficient between transmitter N_i and receiver N_j is represented by $h_{N_i N_j}(b)$, and assumed to be quasi-static Rayleigh fading, i.e., it is a complex Gaussian random variable with zero mean and unit power, assumed to be constant within each block, but to vary independently from block to block. The channel state information $h_{N_i N_j}(b)$ is not known to the transmitter N_i, but only to the receiver N_j. Network geometry, and thus distances $d_{N_i N_j}$, are known to all nodes. The term $x_{N_i}(b,t)$ represents the discrete-time (complex) baseband sample transmitted by the scheduled node $N_i \in \{P_0, \ldots, P_{k-1}, S_1, \ldots, S_{k-1}\}$, with the following per-symbol power constraint:

$$\mathbb{E}\left[\left|x_{N_i}(b,t)\right|^2\right] \leq E_{\mathrm{N}}, \quad t = 1, \ldots, n \tag{3.41}$$

where E_{N} is equal to E_{P} or E_{S} when the transmitter is a primary or a secondary node, respectively (i.e., $N_i \in \{P_0, \ldots, P_{k-1}\}$ or $N_i \in \{S_1, \ldots, S_{k-1}\}$). Finally, we let $z_{N_j}(b,t)$ be the complex white Gaussian noise term with zero mean and power $\mathbb{E}[\left|z_{N_j}(b,t)\right|^2] = N_0$. We assume randomly generated Gaussian codebooks throughout.

We define γ_{P} as the SNR for primary users, which is given by the ratio between the maximum average energy directly received by P_k from the source P_0 and the noise power N_0:

$$\gamma_{\mathrm{P}} = \frac{E_{\mathrm{P}}}{N_0}. \tag{3.42}$$

Hence, for a transmission from a primary node that covers a distance d the average SNR is $\gamma_{\mathrm{P}} d^{-\eta}$. For consistency, the transmission from a secondary node that covers a distance d is given by $\gamma_{\mathrm{S}} d^{-\eta}$, with

$$\gamma_{\mathrm{S}} = \frac{E_{\mathrm{S}}}{N_0}. \tag{3.43}$$

To derive the time fraction β for TDM and the power fraction ψ for SC that meet the QoS requirements of the secondary users, we assume that each secondary node wants to transmit at rate $\mathrm{R_S}$ to a node at distance d_{S} with outage probability ϵ_{S}. In Sections 3.2.1.4 and 3.2.1.5 derivations of β and ψ are derived, respectively.

3.2.1.3 Outage Probabilities for Primary Transmission

Consider transmission from a primary node P_i. Assuming that the coding block is long enough, the probability that a packet transmitted by the primary node P_i is not decoded correctly by a node (primary, P_j, or secondary, S_j) at distance d[3] is given by [61]:

$$P_{\text{out,P}}(d) = \Pr\left\{\log_2\left(1+|h|^2\gamma_{\mathrm{P}} d^{-\eta}\right) \le \mathrm{R_P}\right\} = 1 - \exp\left(-\frac{2^{\mathrm{R_P}}-1}{\gamma_{\mathrm{P}} d^{-\eta}}\right). \tag{3.44}$$

3.2.1.4 Outage Probabilities for Secondary Transmission: TDM

Secondary nodes transmit both primary and secondary data according to the spectrum leasing via cooperation principle. So, we need to evaluate the two corresponding outage probabilities. Recalling that the secondary node transmits the primary packet for a fraction β of the time, we start with TDM.

Let $P^{(\text{TDM})}_{\text{out,SP}}(d)$ define the outage probability of a primary packet transmitted by a secondary node, at a distance d. Similarly to (3.44), this is given by:

$$P^{(\text{TDM})}_{\text{out,SP}}(d) = \Pr\left\{\beta\log_2\left(1+|h|^2\gamma_{\mathrm{S}} d^{-\eta}\right) \le \mathrm{R_P}\right\} = 1 - \exp\left(-\frac{2^{\mathrm{R_P}/\beta}-1}{\gamma_{\mathrm{S}} d^{-\eta}}\right). \tag{3.45}$$

Notice that the rate for secondary transmissions of the primary packet needs to be increased to $\mathrm{R_P}/\beta$ to compensate for the fact that only a fraction of time β is used for primary data.

[3] In the rest of this work, for simplicity, we do not write explicitly the expressions of distance, $d_{N_iN_j}$, and channel coefficient, $h_{N_iN_j}$, between transmitter N_i and receiver N_j, but only d and h, with the understanding that the subscript N_iN_j is implied.

Similarly, the outage probability of a secondary packet transmitted by a secondary node is given by

$$P_{\text{out,SS}}^{(\text{TDM})}(d) = 1 - \exp\left(-\frac{2^{\text{R}_\text{S}/(1-\beta)} - 1}{\gamma_\text{S} d^{-\eta}}\right). \tag{3.46}$$

The choice of β depends, as discussed above, on the QoS requirements of the secondary nodes. Recalling that a fraction $1-\beta$ of the time is used for the secondary's own traffic and imposing the condition on the outage probability in (3.46) as $P_{\text{out,SS}}^{(\text{TDM})}(d_\text{S}) = \epsilon_\text{S}$, we obtain:

$$\beta = 1 - \frac{\text{R}_\text{S}}{\log_2\left[1 - \log_e\left(1-\epsilon_\text{S}\right)\gamma_\text{S} d_\text{S}^{-\eta}\right]}. \tag{3.47}$$

3.2.1.5 Outage Probabilities for Secondary Transmission: SC

With SC, the secondary node sends the sum of two (complex) codewords, one for the primary with power ψE_S and one for the secondary with power $(1-\psi)E_\text{S}$.

We consider a receiver that employs two decoders in parallel. The first decoder attempts to decode the desired packet (primary for a primary node and secondary for a secondary) by treating the undesired packet, which is superimposed, as additive Gaussian noise. The second decoder, instead, first attempts to estimate the undesired packet, cancels it from the received signal and then decodes the desired packet from the interference-free signal. The overall decoder successfully obtains the desired message if either one of the two decoders discussed above decodes correctly (this can be checked via cyclic redundancy check (CRC), see for instance [62]). It is noted that this decoder is capacity-achieving for the Gaussian broadcast channel [65]. Further discussion on this issue can be found in Section 3.2.2.6.

For the SC approach, the outage probability for a primary packet transmitted by a secondary node S_i to a node N_j (primary or secondary) at distance d can be found to be given by

$$\begin{aligned} P_{\text{out,SP}}^{(\text{SC})}(d) = &\Pr\left\{\log_2\left(1 + \frac{|h|^2\psi E_\text{S} d^{-\eta}}{N_0 + |h|^2(1-\psi)E_\text{S} d^{-\eta}}\right) \le \text{R}_\text{P}\right. \\ &\left.\cap\left[\left(\log_2\left(1 + \frac{|h|^2(1-\psi)E_\text{S} d^{-\eta}}{N_0 + |h|^2\psi E_\text{S} d^{-\eta}}\right) \le \text{R}_\text{S}\right) \cup \left(\log_2\left(1 + \frac{|h|^2\psi E_\text{S} d^{-\eta}}{N_0}\right) \le \text{R}_\text{P}\right)\right]\right\} \end{aligned} \tag{3.48}$$

$$= \Pr\left\{|h|^2 \le \min\left(\mathcal{H}_\text{P}^{(1)}, \mathcal{H}_\text{P}^{(2)}\right)\right\} = 1 - \exp\left[-\min\left(\mathcal{H}_\text{P}^{(1)}, \mathcal{H}_\text{P}^{(2)}\right)\right], \tag{3.49}$$

where the threshold $\mathcal{H}_\text{P}^{(1)}$ and $\mathcal{H}_\text{P}^{(2)}$ are defined below. The first term in (3.48) represents the outage probability of the first decoder, in which the interference (i.e., secondary packet) is

treated as noise. The remaining term is the outage probability of the successive decoding scheme, where the receiver first decodes the secondary packet and then the primary one. The outage probability of the latter decoder measures the event that decoding of either the secondary packet in the first stage or the primary in the second stage is incorrect. The overall outage probability (3.48) is the probability that both the first and the second decoder fail. As such, in equation (3.49), $\mathcal{H}_{\mathrm{P}}^{(1)}$ and $\mathcal{H}_{\mathrm{P}}^{(2)}$ represent the minimum values that the channel coefficient $|h|^2$ can assume without causing an outage for the two decoders. These thresholds can be evaluated for the first decoder (which treats interference as noise) and for the successive decoding scheme as

$$\mathcal{H}_{\mathrm{P}}^{(1)} = \begin{cases} \infty, & 0 \leq \psi \leq 1 - 2^{-\mathrm{R_P}} \\ \frac{2^{\mathrm{R_P}}-1}{(1-(1-\psi)2^{\mathrm{R_P}})\gamma_{\mathrm{S}} d^{-\eta}}, & 1 - 2^{-\mathrm{R_P}} < \psi \leq 1 \end{cases} \tag{3.50}$$

$$\mathcal{H}_{\mathrm{P}}^{(2)} = \begin{cases} \max\left\{\frac{2^{\mathrm{R_S}}-1}{(1-\psi 2^{\mathrm{R_S}})\gamma_{\mathrm{S}} d^{-\eta}}, \frac{2^{\mathrm{R_P}}-1}{\psi \gamma_{\mathrm{S}} d^{-\eta}}\right\}, & 0 < \psi < 2^{-\mathrm{R_S}} \\ \infty, & \psi = 0 \text{ and } 2^{-\mathrm{R_S}} \leq \psi \leq 1 \end{cases} \tag{3.51}$$

Notice from (3.50), (3.51) that if the allocated power is too small, the channel gain threshold values for which there is no outage become infinite (i.e., outage occurs with probability one for all finite channel gains).

Similarly to (3.48), the outage probability that a secondary packet (superimposed with a primary message) transmitted by a secondary node S_i is not decoded correctly by a secondary node S_j placed at distance d is given by

$$P_{\text{out,SS}}^{(\mathrm{SC})}(d) = 1 - \exp\left[-\min\left(\mathcal{H}_{\mathrm{S}}^{(1)}, \mathcal{H}_{\mathrm{S}}^{(2)}\right)\right], \tag{3.52}$$

where $\mathcal{H}_{\mathrm{S}}^{(1)}$ and $\mathcal{H}_{\mathrm{S}}^{(2)}$ are

$$\mathcal{H}_{\mathrm{S}}^{(1)} = \begin{cases} \frac{2^{\mathrm{R_S}}-1}{(1-\psi 2^{\mathrm{R_S}})\gamma_{\mathrm{S}} d^{-\eta}}, & 0 \leq \psi < 2^{-\mathrm{R_S}} \\ \infty, & 2^{-\mathrm{R_S}} \leq \psi \leq 1 \end{cases} \tag{3.53}$$

$$\mathcal{H}_{\mathrm{S}}^{(2)} = \begin{cases} \infty, & 0 \leq \psi \leq 1 - 2^{-\mathrm{R_P}} \text{ and } \psi = 1 \\ \max\left\{\frac{2^{\mathrm{R_P}}-1}{(1-(1-\psi)2^{\mathrm{R_P}})\gamma_{\mathrm{S}} d^{-\eta}}, \frac{2^{\mathrm{R_S}}-1}{(1-\psi)\gamma_{\mathrm{S}} d^{-\eta}}\right\}, & 1 - 2^{-\mathrm{R_P}} < \psi < 1 \end{cases} \tag{3.54}$$

As with TDM in Section 3.2.1.4, given the secondary QoS requirements (d_{S}, $\mathrm{R_S}$, ϵ_{S}), one can obtain the resource allocation parameter ψ. While for TDM this could be easily done in closed-form (3.47), for SC we had to resort to a numerical solution of the equation $P_{\text{out,SS}}^{(\mathrm{SC})}(d_{\mathrm{S}}) = \epsilon_{\mathrm{S}}$, for a given rate pair $(\mathrm{R_P}, \mathrm{R_S})$.

The next section defines the performance criteria of interest in this work, i.e., primary throughput (recalling the definition given in Section 3.1.2, Equation (3.5)) and average primary energy consumption, and proves the superiority of the SC scheme over TDM.

3.2.2 Throughput and Primary Energy Analysis

The goal of this section is to first define the performance metrics of interest, then introduce four routing policies that exploit spectrum leasing via opportunistic routing to different degrees, and finally show the optimality of SC. We fix the primary rate $\mathrm{R_P}$ and the secondary QoS requirements (d_{S}, $\mathrm{R_S}$, ϵ_{S}) or equivalently the parameters β and ψ. Notice that these parameters can be calculated by the secondary network and made known to the primary, that can then decide which form of spectrum leasing (if any) is convenient.

Similarly to (3.5) in Section 3.1.2, let $T(k, \mathrm{R_P}, \mathcal{Q})$ be the *primary end-to-end throughput*, defined as the average number of *successfully* transmitted bits per second per Hz, given the total number of hops k, the primary transmission rate $\mathrm{R_P}$ and the parameter $\mathcal{Q}$, which represents the secondary QoS constraints (d_{S}, $\mathrm{R_S}$, ϵ_{S}) or equivalently the parameter β for TDM or ψ for SC. Using renewal theory, the throughput can be calculated as

$$T(k, \mathrm{R_P}, \mathcal{Q}) = \frac{\mathrm{R_P}}{\mathbb{E}[N]}, \tag{3.55}$$

where N is the total number of blocks, including *both primary and secondary transmissions*, necessary to transmit a given packet correctly from the source P_0 to the destination P_k. We also define the primary energy $E(k, \mathrm{R_P}, \mathcal{Q})$ as the average overall energy used by the primary network to deliver a packet successfully. We measure this quantity via the number of blocks involving *primary* transmissions that are necessary to correctly deliver a packet from the source P_0 to the destination P_k,

$$E(k, \mathrm{R_P}, \mathcal{Q}) = \mathbb{E}[N_{\mathrm{P}}], \tag{3.56}$$

where N_{P} represents the number of *primary* transmissions.

We now detail the four proposed transmission policies for the primary packets. We remark that all four policies are based on a Type-I HARQ so that decoding is performed in each block by discarding previous transmissions. Extension to more complex forms of HARQ is possible but would lead to a different analysis and is left as future work. All policies are implemented using both receiver techniques introduced above (i.e., TDM and

SC). Note that the policy descriptions below apply identically to both cases, as they only differ at the PHY level.

3.2.2.1 Policy 1: only Primary (only-P)

The *only-P* policy does not exploit spectrum leasing and is introduced here for reference. Only the primary nodes are involved in transmissions. A basic opportunistic routing strategy is assumed. Accordingly, in each block the transmitter is selected opportunistically as the primary node that has decoded the previous transmission and is the closest to the destination. Since we assume Type-I HARQ, the current transmitter retransmits the packet until at least one of the downstream nodes has successfully decoded. A new packet is transmitted by P_0 as soon as the destination P_k decodes successfully.

3.2.2.2 Policy 2: only Secondary (only-S)

The *only-S* policy aims minimizing the primary transmissions, and thus the primary energy consumption $E(k, \mathrm{R_P}, \mathcal{Q})$ (3.56), thanks to spectrum leasing, while possibly suffering some throughput loss. This is accomplished by forcing the source to send the information only through secondary nodes $\{S_1, \ldots, S_{k-1}\}$, i.e., without exploiting any primary relay $\{P_1, \ldots, P_{k-1}\}$ and thus the multiuser diversity arising from their presence as well. The only primary (re)transmissions allowed are thus from the source P_0. So, *only-S* has the same topology of *only-P*, but a different exploitation of the relays. In fact, when a secondary relay is considered as transmitter, only a portion β of the time-slot is used to transmit the primary packet and also the transmission power may be different (i.e., $E_{\mathrm{P}} \neq E_{\mathrm{S}}$ in general). An opportunistic routing scheme is used on the secondary network, where transmission is granted to the secondary node that has decoded the previous transmission and that is the closest to the destination.

3.2.2.3 Policy 3: Primary to Secondary (P-to-S)

The *only-S* policy introduced above minimizes the primary transmissions thanks to spectrum leasing, but may suffer from a poor throughput as, once the packet has entered the secondary network, the multiuser diversity arising from the presence of primary nodes, and their transmission power, is not leveraged. So the main limitation of *only-S* is the scarce

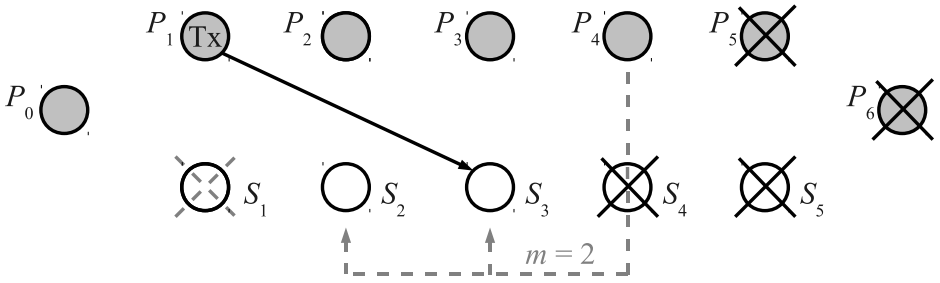

Figure 3.7. *An illustration of the* P-to-S *policy. Spectrum leasing is performed from node P_1 to S_3, with* backward window *with parameter $m = 2$.*

exploitation of the primary relays as the next hop. The *P-to-S* policy, proposed here, and the *P-and-S* policy, to be discussed in the next subsection, attempt, to different extents, to offer a better trade-off between primary throughput and energy that does not disregard throughput and multiuser diversity and that can be controlled via a parameter m.

In the *only-S* policy, the source P_0 retransmits until at least one secondary node, or the destination P_k, decodes. If a primary relay decodes, the policy does not use it. With *P-to-S*, the idea instead is to use primary relays unless a secondary node in a "sufficiently good" position, as dictated by m, has decoded. From that point on, the packet is handled by the secondary network as in the *only-S* strategy.

Specifically, at each block in which a primary node is transmitting, the primary network first determines the type of relay closest to the destination that has successfully decoded. If the latter node is a secondary, it is selected for the next hop. If it is a primary, in order to save primary energy, the node is selected only if the best secondary node is at least m hops behind. In other words, the next transmitter is selected as either the primary node at hand or the closest secondary node as long as the latter is within a window of m hops from the position of the primary node toward the source. This window will be referred to as *backward window* and we generally have $0 \leq m \leq k-2$.

The idea is illustrated in Figure 3.7, where a possible scenario in which P_1 is the current transmitter, P_4 is the most advanced decoding node (a black cross indicates a node which has not successfully decoded) and $m = 2$. Before selecting node P_4 as the next hop, the primary network checks whether any secondary relay within the *backward window* $\{S_4, S_3, S_2\}$ has decoded the packet. If this is the case, the transmitter picks the most advanced such node as the next hop. In this particular example, relay S_3 is selected, because node S_4 has not

successfully decoded. Note that S_1 is outside the *backward window*, and cannot be used even though it has successfully decoded the packet (a grey dotted cross is used to highlight that this node cannot be selected).

3.2.2.4 Policy 4: Primary and Secondary (P-and-S)

In the *P-to-S* policy, when a packet enters the secondary network, it cannot return to the primary one, except for the final destination P_k. This is again done in an attempt to save the primary energy, but it limits the multiuser diversity and the resource available to the secondary transmitter, causing a throughput reduction. The proposed *P-and-S* policy removes this constraint to favor throughput maximization. The policy is again described by a parameter m.

Let us start with $m = 0$. The idea here is simply to select in each block the node that is the closest to the destination among those that have decoded, irrespective of whether such node is primary or secondary. This strategy clearly privileges primary throughput, since it exploits all the transmission opportunities afforded by the network. In order to obtain a more controllable trade-off between throughput and energy, we generalize this policy by letting $m > 0$ and operating as follows.

Let $m > 0$. The policy extends *P-to-S* allowing transmissions from secondary back to primary relays, but with a constraint on the minimum progress given by the so called *forward window*. In particular, if the transmitter is a primary node, the strategy works as for the *P-to-S* policy. However, if the transmitter is a secondary, in order to enhance throughput, we enable the selection also of primary nodes, as long as the primary node to be selected is at least m hops ahead of the most advanced secondary decoding node. Thus, a primary relay can receive the packet from a secondary node only if it is outside the *forward window*, which is of size m hops[4] and starts from the most advanced decoding secondary node towards the destination.

We remark that the use of the *backward window* and the *forward window* is quite different: the *backward window* includes the secondary relays that can be selected as next hops when a primary node is the current transmitter; the *forward window* includes the primary relays that cannot be selected as the next hop when a secondary node transmits.

[4]In principle, one could choose two different sizes for forward and backward windows, but this is not further

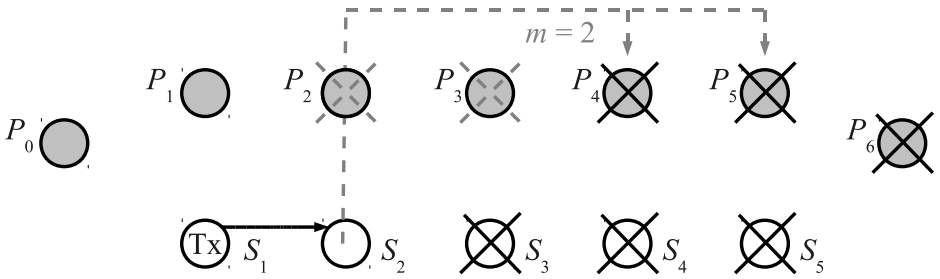

Figure 3.8. *An illustration of the* P-and-S *policy. The transmission granted is handed between secondary relays S_1 and S_2.*

In Figure 3.8 we illustrate a possible scenario, where relay S_1 is the current transmitter and $m = 2$. Node S_2 is the most advanced secondary decoding relay (nodes that have not decoded are marked with a black cross), whereas the best decoding primary relay is P_3. However, nodes P_2 and P_3 cannot be selected for the next hop because they are inside the *forward window* (a gray dotted cross indicates this fact). Therefore, the *P-and-S* policy selects node S_2 in this case.

3.2.2.5 Evaluating Primary Throughput and Energy

In order to evaluate the performance metrics throughput (3.55) and average primary energy (3.56) for the protocols discussed above, we use the theory of Markov chains, as similarly done in Section 3.1.2.2. In this case, we model the network with a chain of $2k$ states, one for each primary and secondary node. State P_0 refers to a situation where the current packet is at the source P_0, the primary states P_i and secondary states S_i, $i = 1, \ldots, k-1$, are similarly defined, and P_k represents the state where the destination has successfully decoded. Recalling that we assume Type-I HARQ, the current transmitter retransmits the packet until at least one of the nodes admitted by the specific policy has successfully decoded. So, the transition matrix is organized in four blocks as

$$\mathbf{\Phi} = \left[\begin{array}{c|c} \mathbf{\Phi}_{\mathrm{P,P}} & \mathbf{\Phi}_{\mathrm{P,S}} \\ \hline \mathbf{\Phi}_{\mathrm{S,P}} & \mathbf{\Phi}_{\mathrm{S,S}} \end{array}\right], \tag{3.57}$$

where the states are ordered as $P_0, P_1, \ldots, P_k, S_1, \ldots, S_{k-1}$, and $\mathbf{\Phi}_{\mathrm{A,B}}$, $A, B \in \{P, S\}$ are the submatrices that collect all the transition probabilities from nodes of type A (Primary, P, or

investigated here.

Secondary, S) to B. In general, in matrix $\mathbf{\Phi}_{\mathrm{A,B}}$ the term $\Phi_{\mathrm{A,B}}(i,j)$ represents the probability that, given the current state A_i (i.e., the transmitter is node A_i, with $i = 0, \ldots, k-1$ if $A = P$ and $i = 1, \ldots, k-1$ if $A = S$), the next state is B_j, with $j = i, i+1, \ldots, k$ when $B = P$ and $j = 1, \ldots, k-1$ if $B = S$ (this is because the policies allow backward transmissions if the receiver is a secondary node).

In matrix (3.57), the first k states and the last $k-1$ states are transient, whereas the $k+1$-th state, corresponding to the packet being received at the destination, is absorbing. Depending on the routing policy adopted, the transition probabilities will assume different expressions and will be detailed in the Appendix B. The average primary energy and throughput are derived as detailed in the Lemma below, which follows from standard Markov chain theory [66, Cap. 3] (see also Proposition 3.1.2).

Lemma 3.2.1. *The end-to-end throughput (3.55) and the primary energy (3.56) for fixed primary transmission rate* R_P *are given by*

$$T(k, \mathrm{R}_P, \mathcal{Q}) = \frac{\mathrm{R}_P}{v_{P_0}} \quad \textit{and} \quad E(k, \mathrm{R}_P, \mathcal{Q}) = w_{P_0}, \tag{3.58}$$

where v_{P_0} *and* w_{P_0} *are the first elements of vectors*

$$\mathbf{v} = [\mathbf{v}_P, \mathbf{v}_S] = [v_{P_0}, \ldots, v_{P_{k-1}}, v_{S_1}, \ldots, v_{S_{k-1}}] \tag{3.59}$$

and

$$\mathbf{w} = [\mathbf{w}_P, \mathbf{w}_S] = [w_{P_0}, \ldots, w_{P_{k-1}}, w_{S_1}, \ldots, w_{S_{k-1}}], \tag{3.60}$$

which are evaluated as

$$\mathbf{v} = (\mathbf{I} - \mathbf{Q})^{-1}\,\mathbf{1} \quad \textit{and} \quad \mathbf{w} = (\mathbf{I} - \mathbf{Q})^{-1}\,\mathbf{r}, \tag{3.61}$$

where $\mathbf{1}$ *is a* $(2k-1) \times 1$ *vector with all entries equal to* 1 *and* $\mathbf{r}$ *is the reward vector*

$$\mathbf{r} = [\mathbf{r}_P, \mathbf{r}_S] = [r_{P_0}, \ldots, r_{P_{k-1}}, r_{S_1}, \ldots, r_{S_{k-1}}] \tag{3.62}$$

where $\mathbf{r}_P$ *is a* $k \times 1$ *vector with all ones,* $\mathbf{r}_S$ *is a* $(k-1) \times 1$ *vector with all zero elements and* $\mathbf{I}$ *is the* $(2k-1) \times (2k-1)$ *identity matrix. Finally, matrix* $\mathbf{Q}$ *is obtained from* $\mathbf{\Phi}$ *by removing the* $(k+1)$*-th row and the* $(k+1)$*-th column.*

3.2.2.6 Optimality of SC approach

In [67] the authors consider the outage capacity of a two-user quasi-static fading broadcast channel when the transmitter has no information about the instantaneous state of the channel. It is proved that SC with Gaussian codewords is optimum. The following Proposition uses this result to prove that, in our system, the considered SC scheme is optimal for both throughput and primary energy consumption, that is, it is the best among all possible multiplexing schemes to be employed by the secondary nodes. We recall that, when transmitting, the secondary nodes multiplex primary and secondary messages.

Proposition 3.2.2 (Optimality of Superposition Coding)**.** *Fix primary rate* R_P*, secondary QoS requirements (*d_S*,* γ_{S}*,* R_S*,* ϵ_S*) and any of the proposed routing strategies. The following holds:*

(i) *Any throughput* $T(k, \mathrm{R}_P, \mathcal{Q})$ *that can be attained by any multiplexing scheme of primary and secondary codewords (not necessarily randomly generated according to a Gaussian codebook) at the secondary nodes can also be achieved by the SC scheme with Gaussian codewords studied in Section 3.2.1.5;*

(ii) *The primary energy* $E(k, \mathrm{R}_P, \mathcal{Q})$ *used by the SC scheme is no larger than the amount of energy expended by any other multiplexing scheme.*

Proof. Consider any multiplexing scheme at the secondary node S_i. In our model, let

$$\mathcal{M} \subset \{P_1, \ldots, P_k, S_1, \ldots, S_{k-1}\} \tag{3.63}$$

denote the set with cardinality $|\mathcal{M}| = M$ of primary and secondary receivers that, depending on the specific routing strategy, try to decode the primary message and let $N^{(2)}$ be the intended receiver of the secondary message. Fix the outage probability of such scheme at all nodes. We want to show that for given outage probabilities, SC can achieve rates as large as any other scheme. This would prove that SC is also throughput and primary energy optimal. It can be easily argued that, since the probability of outage at each receiver only depends on the corresponding fading channel, there are M thresholds $k_{N^{(1)}}$ that represent the minimum channel values $|h_{S_i N^{(1)}}|^2$, $\forall N^{(1)} \in \mathcal{M}$, that allow receivers to successfully decode the primary packet. Similarly, the secondary node is in outage if and only if $|h_{S_i N^{(2)}}|^2 < k_S$, where k_S is the threshold on its channel coefficient [67]. Moreover, it can be argued that if decoding

is successful at the threshold channel values, it must also be successful for larger channel values. So, fix such channel thresholds. Rates $(\mathrm{R_P}, \mathrm{R_S})$ will be correctly decoded if and only if they are inside the capacity region of the Gaussian (non-fading) broadcast channel with the given channel coefficients. In particular, one can focus on the two-user broadcast channel formed by the secondary user and the primary channel with the worst overall channel (since decoding of the primary packet at this node also implies decoding at the better nodes). It is known that Gaussian SC can achieve any point in the capacity region of the Gaussian broadcast channel with an outage probability no larger than any other scheme [65]. Therefore, SC is optimal for both throughput and primary energy: the minimal outage probability leads to both an optimal throughput, due to the fact that fewer retransmissions are necessary, and a decreased primary energy, due to the larger number of secondary relays that are potentially reachable at each transmission by the proposed routing schemes. □

3.2.3 Numerical Results

In this last part of the chapter we first provide some numerical evidence about the superiority of SC over TDM, which was proved in Proposition 3.2.2. Then, focusing on the SC scheme, we elaborate on the advantages of spectrum leasing and on the design of the proposed schemes. Throughout this section we fix the following parameters: number of hops $k = 12$, path loss $\eta = 3$, geometry of the network $\Delta_{\mathrm{V}} = \Delta_{\mathrm{H}} = 1/k$ and transmitting power of secondary users $E_{\mathrm{S}} = E_{\mathrm{P}}$ (then also the SNRs are equal, $\gamma_{\mathrm{P}} = \gamma_{\mathrm{S}} = \gamma$, see Section 3.2.1.2). We consider two secondary deployments: (*i*) Full ($\alpha = 1$) and (*ii*) Partial ($\alpha > 1$). As for the secondary QoS requirements $\mathcal{Q}$, we assume that each secondary node wants to transmit its own traffic at rate $\mathrm{R_S}$ to a node with SNR equal to γ at distance $d_{\mathrm{S}} = 1/10$ with outage probability $\epsilon_{\mathrm{S}} = 0.1$.

A comment on the calculation of the secondary transmission parameters based on the QoS requirements $\mathcal{Q}$ is in order. For TDM, from (3.47), the time fraction β is equal to 0.83, for $\mathrm{R_S} = 1\ bits/s/Hz$ and regardless of $\mathrm{R_P}$. As for SC, the value of ψ that satisfies the secondary QoS constraint $P^{(\mathrm{SC})}_{\mathrm{out,SS}}(d_{\mathrm{S}}) = \epsilon_{\mathrm{S}}$ in (3.52) is not necessarily unique, but depends on the particular decoder employed, namely treating interference as noise ($\mathcal{H}^{(1)}_{\mathrm{S}}$ in (3.53)) or adopting a successive decoding scheme ($\mathcal{H}^{(2)}_{\mathrm{S}}$ in (3.54)). We select the solution that maximizes the primary rate, i.e., the highest feasible ψ.

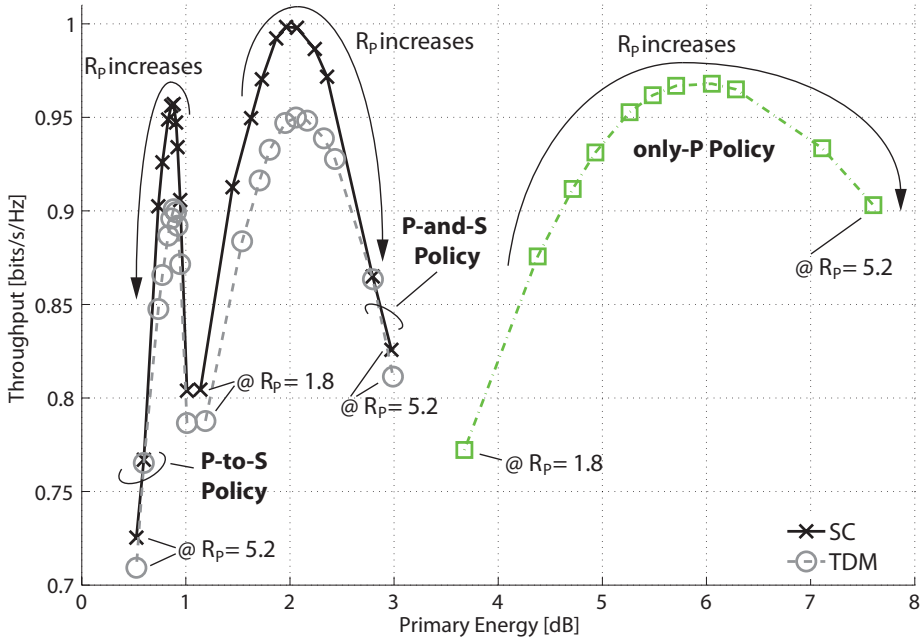

Figure 3.9. *End-to-end throughput and overall primary energy as a function of the primary transmission rate* R_P *for a network with the same number of primary and secondary relays* $(k-1)$ *(full secondary deployment) for* $k = 12$ *hops,* SNR $\gamma = -3dB$, $m = 1$ *and* $\mathrm{R}_S = 1\ bits/s/Hz$. *Each line is obtained by varying the transmission rate* R_P *as* $\{1.8, 2.4, 2.7, 2.9, 3.2, 3.4, 3.6, 3.9, 4.1, 4.8, 5.2\}\ bits/s/Hz$.

3.2.3.1 Comparison of SC and TDM

Figure 3.9 evaluates the gain of SC scheme over TDM for *P-to-S* and *P-and-S* policies, by varying the primary transmission rate $\mathrm{R_P}$ in a full secondary deployment for $\gamma = -3dB$, $m = 1$ and $\mathrm{R_S} = 1\ bits/s/Hz$. Each curve is obtained by evaluating the pair end-to-end throughput and primary energy of a given scheme (i.e., SC or TDM) and policy for different $\mathrm{R_P}$ and keeping all the other parameters fixed. For the rest of this section, the primary energy $E(k, \mathrm{R_P}, \mathcal{Q})$ is expressed in dB, i.e. $10\log_{10} E(k, \mathrm{R_P}, \mathcal{Q})$. Figure 3.9 confirms the optimality of the SC scheme for both *P-to-S* and *P-and-S*, regardless of $\mathrm{R_P}$. In fact, as already proved in Proposition 3.2.2, the better outage provided by SC results in a better throughput and in a primary energy saving. It is also seen that, for each policy, there exists a rate that maximizes the throughput. Such rate is different for distinct policies. For example, the throughput-optimal rates in the SC scheme for *only-P* (i.e., no spectrum leasing) is $\mathrm{R_P} = 3.9\ bits/s/Hz$, for *P-to-S* is $\mathrm{R_P} = 3.2\ bits/s/Hz$ and for *P-and-S* is $\mathrm{R_P} = 3.6\ bits/s/Hz$. In order to reduce the primary energy consumption, at the cost of a reduced throughput, one can decrease the

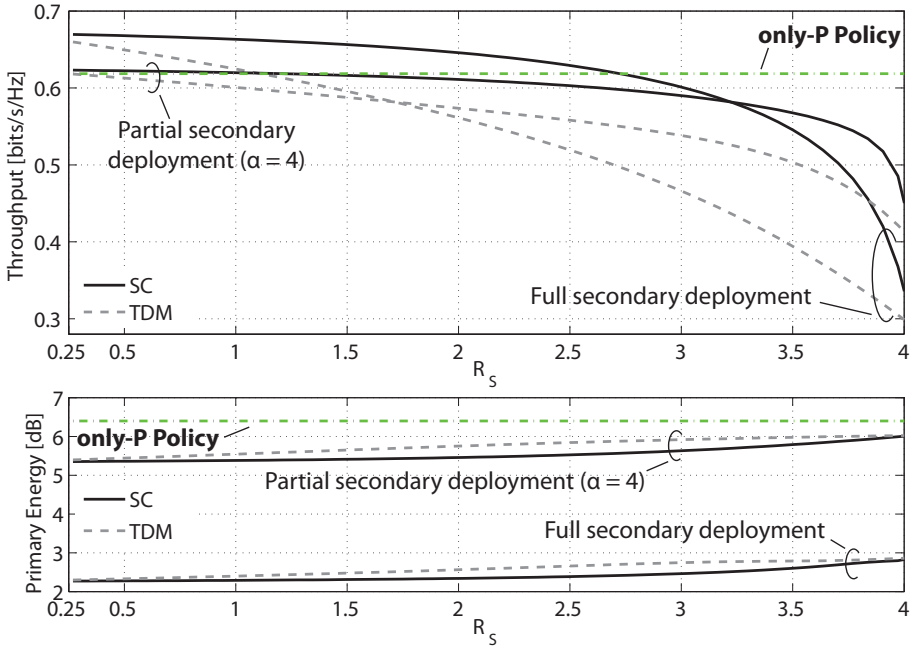

Figure 3.10. *End-to-end throughput and overall primary energy of* P-and-S *policy plotted varying* R_S *for a network with full secondary deployment and with less secondary relays* $(k/\alpha - 1)$ *than primary nodes* $(k-1)$ *(partial secondary deployment) for* $\alpha = 4$*,* $k = 12$ *hops, transmission rate* $\mathrm{R}_P = 2.7\ bits/s/Hz$*,* $\gamma = -8dB$ *and* $m = 1$*.*

transmission rate $\mathrm{R_P}$ for all policies except *P-to-S*. In fact, in general, a lower transmission rate causes a wider coverage range of the primary transmission. So, when the transmission rate is low the progress towards the destination increases, and this makes it less likely that a packet enters the secondary network. Thus, when the transmission rate decreases in the *P-to-S* policy, it is more likely to remain in the primary network, which leads to a higher primary energy consumption. It is finally noted that the energy gain of spectrum leasing over the *only-P* (no spectrum leasing) policy are substantial irrespective of the choice of the transmission rate. Moreover, spectrum leasing combined with the SC scheme outperforms *only-P* (no spectrum leasing) in terms of throughput for $\mathrm{R_P} < 4.8\ bits/s/Hz$.

The impact of secondary QoS requirements, quantified by parameter $\mathrm{R_S}$, for full and partial secondary deployment with $\alpha = 4$ is considered in Figure 3.10 for only *P-and-S* policy and for $\gamma = -8dB$, $m = 1$ and $\mathrm{R_P} = 2.7\ bits/s/Hz$. The throughput and primary energy superiority of SC scheme over TDM is confirmed also in this case, regardless of $\mathrm{R_S}$. Particularly, SC scheme is less affected by secondary rate variations, due to the better out-

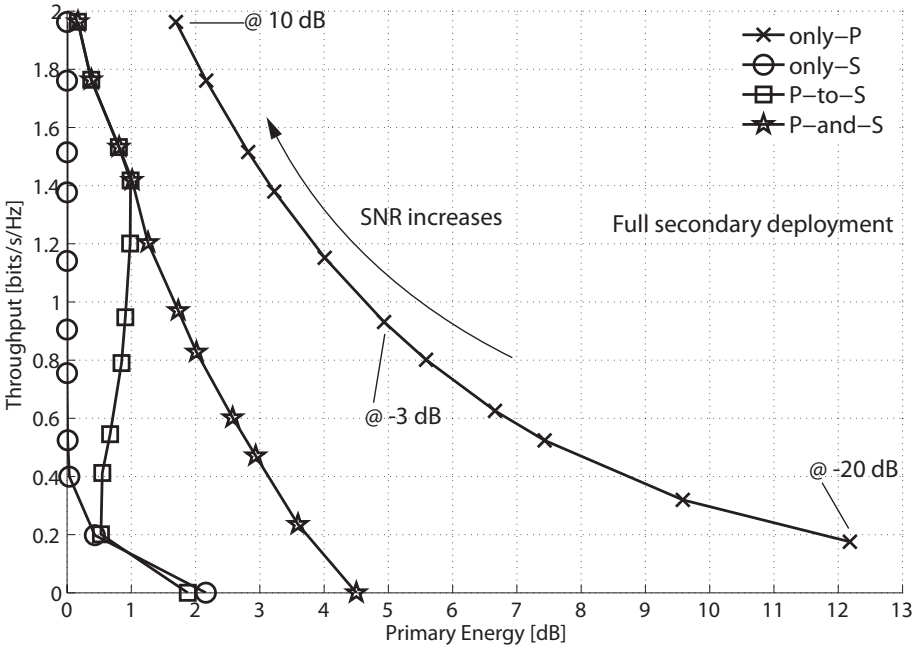

Figure 3.11. *End-to-end throughput and overall primary energy plotted varying the* SNR, γ, *for a network with full secondary deployment, for* $k = 12$ *hops, primary transmission rate* $\mathrm{R}_P = 2.9\ bits/s/Hz$, $m = 1$ *and* $\mathrm{R}_S = 1\ bits/s/Hz$. *Each line is obtained by varying* γ *as* $\{-20, -15, \pm 10, \pm 8, \pm 5, \pm 3, 0\} dB$.

age probability reached by this scheme respect to TDM, with equal constraints. Moreover, the throughput performances of *only-P* (no spectrum leasing) are outperformed by *P-and-S* policy with SC scheme for $\mathrm{R_S} < 2.6\ bits/s/Hz$. Finally, we note that partial deployment presents a better throughput when $\mathrm{R_S} > 3.2\ bits/s/Hz$ respect to full deployment. In fact, when $\mathrm{R_S}$ is high, avoid using secondary nodes (and the partial deployment does) leads to a gain in terms of outage probability. This gain is reduced when $\mathrm{R_S}$ decreases, (i.e. when secondary transmissions are preferable respect to primary transmissions) due to the secondary nodes which require a lower portion of the transmitting power to forward its own data.

3.2.3.2 Design and Advantages of Spectrum Leasing

In the rest of this section we numerically evaluate the impact of secondary relays on the primary network only. Given the discussion above for SC and TDM, we focus only on the optimal scheme SC to numerically study the spectrum leasing features. We first study the trade-off between end-to-end throughput and the overall primary energy consumption as a function of the SNR γ for two different secondary nodes deployments (full in Figure 3.11

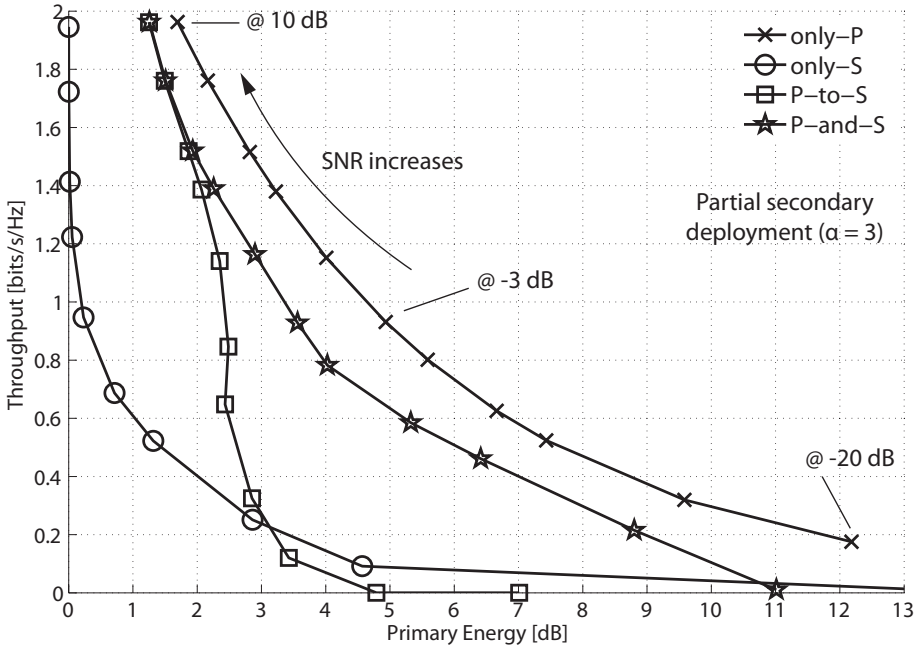

Figure 3.12. *End-to-end throughput and overall primary energy plotted varying the* SNR, γ, *for a network with partial secondary deployment for* $\alpha = 3$, $k = 12$ *hops, primary transmission rate* $\mathrm{R}_P = 2.9\ bits/s/Hz$, $m = 1$, $\mathrm{R}_S = 1\ bits/s/Hz$ *and* γ *that assumes the following values* $\{-20, -15, \pm 10, \pm 8, \pm 5, \pm 3, 0\} dB$.

and partial, with $\alpha = 3$, in Figure 3.12), fixing the remaining protocol parameters to $m = 1$, $\mathrm{R_P} = 2.9\ bits/s/Hz$ and $\mathrm{R_S} = 1\ bits/s/Hz$. Each curve is obtained by evaluating the pair end-to-end throughput and primary energy of a given policy for different γ (ranging from $-20\ dB$ to $10\ dB$), thus different ψ, and keeping all the other parameters fixed. From Figure 3.11, it is seen that, with the given parameters, spectrum leasing policies with full secondary deployment are more energy efficient with respect to the *only-P* policy (no spectrum leasing), especially as the SNR decreases, due to the larger benefits afforded by opportunistic routing.

Such gains, while still substantial, decrease with a partial secondary deployment as shown in Figure 3.12 for $\alpha = 3$. In both cases, however, when $\gamma \simeq -20dB$, the secondary QoS requirements are satisfied only with ψ close to 0, where the throughput of the spectrum leasing policies is almost 0, due to the low power assigned to the primary packet. Moreover, when the SNR decreases, the throughput of *only-S* and *P-to-S* is affected by the partial deployment due to the longer secondary hops. In this case, the *P-and-S* policy is to be preferred as it is able to keep the same level of throughput of Figure 3.11 (but with larger primary en-

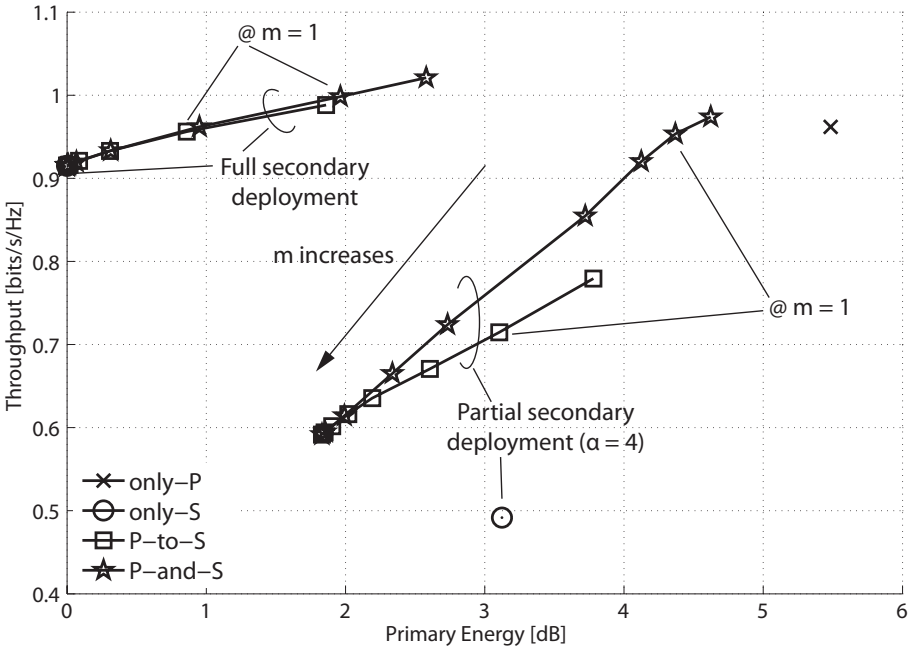

Figure 3.13. *End-to-end throughput and overall primary energy shown varying m for a network with full and partial secondary relays deployment for $\alpha = 4$, $k = 12$ hops, transmission rate $\mathrm{R_P} = 3.4\ bits/s/Hz$, $\gamma = -3dB$ and $\mathrm{R_S} = 1\ bits/s/Hz$. The lines are obtained by varying m from 0 to 10.*

ergy). Better performance can be obtained by optimizing on the *forward* and *backward window* size m, as discussed next.

Figure 3.13 shows end-to-end throughput and primary energy by varying m for full and partial secondary deployment with $\alpha = 4$ and for parameters $\gamma = -3\ dB$, $\mathrm{R_P} = 3.4\ bits/s/Hz$ and $\mathrm{R_S} = 1\ bits/s/Hz$. Similar to the discussion above, the *P-and-S* policy outperforms *only-S* and *P-to-S* from a throughput point of view, especially in the partial secondary deployment scenario. The *only-S* policy in the full secondary relay scenario is very close to the performance of *P-to-S* for high m, and is not visible in the graph. However, this behavior confirms the strict relationship between these two policies, especially for high m. Moreover, it is clear that the parameter m allows to trade-off energy and throughput. For the policies *P-to-S* and *P-and-S*, increasing m ($m \geq 4$) trades throughput for a decreased primary energy consumption, due to the larger number of secondary transmissions admitted. When m is sufficiently low ($m \leq 4$), the throughput increases differently in *P-to-S* and *P-and-S*. For the *P-to-S* policy, which employs only the *backward window* and blocks secondary transmissions to primary relays, the throughput and primary energy are larger due to the

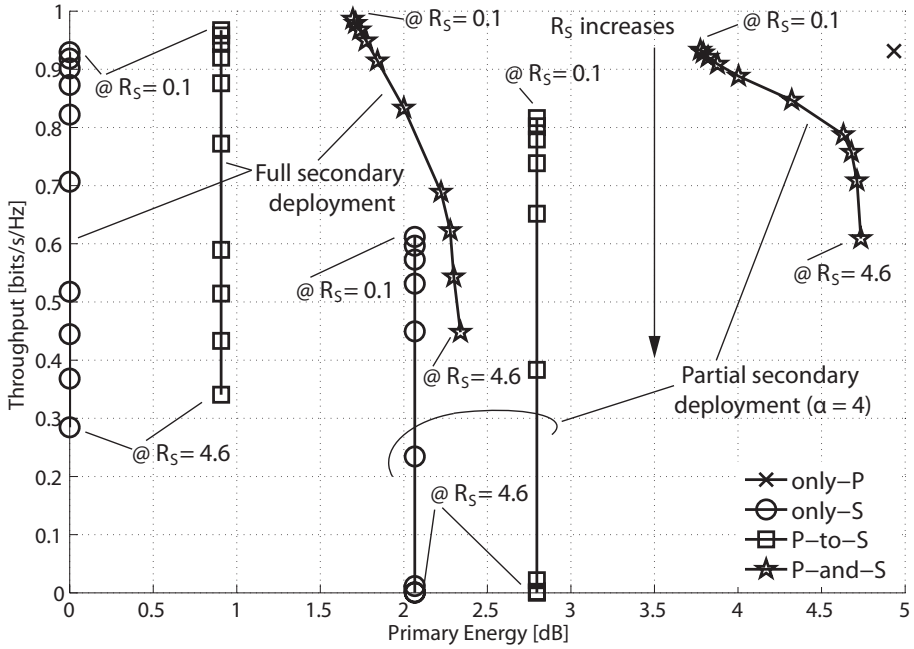

Figure 3.14. *End-to-end throughput and overall primary energy plotted varying* $\mathrm{R_S}$ *for a network with full and partial secondary deployment for* $\alpha = 4$, $k = 12$ *hops, transmission rate* $\mathrm{R_P} = 2.9\ bits/s/Hz$, $\gamma = -3dB$ *and* $m = 1$. *The lines are obtained by varying* $\mathrm{R_S}$ *from* $0.1\ bits/s/Hz$ *to* $4.6\ bits/s/Hz$ *with step* $0.5\ bits/s/Hz$.

lower number of secondary nodes available to lease the spectrum. In *P-and-S* this limit is overcome by removing the block from the secondary transmissions and by introducing the *forward window*. Thus, due to the capability of exploiting more path diversity, the *P-and-S* policy is able to obtain larger throughputs (for larger energy consumptions) than *P-to-S*.

Finally, in Figure 3.14, we consider the impact of $\mathrm{R_S}$ on the four policies for the SC scheme, for full and partial secondary deployment with $\alpha = 4$, $\gamma = -3\ dB$, $\mathrm{R_P} = 2.9\ bits/s/Hz$ and $m = 1$. We note that increasing the secondary QoS requirements (i.e., increasing $\mathrm{R_S}$) leads to a decreased throughput without affecting the primary energy for all policies, except *P-and-S*. Indeed, in all policies except *P-and-S* modifying $\mathrm{R_S}$ does not change the number of primary transmissions, but only the amount of power leased to the primary network. Instead, for *P-and-S* an higher $\mathrm{R_S}$ leads to both a decreased throughput and an increased primary energy, due to the larger number of secondary transmissions towards primary nodes. In fact, when $\mathrm{R_S}$ is high the number of secondary transmissions increases, but also the possibility to return on the primary network increases as well. If this happens,

the next hop will be covered by a primary transmission thus consuming primary energy. Moreover, with partial secondary deployment for $\alpha = 4$, *P-and-S* confirms to be able to best adapt to the lack of secondary nodes, by increasing primary transmissions. This causes a larger primary energy, but to a smaller extent than *only-S* and *P-to-S*.

3.3 Conclusions

The first goal of this chapter was to study the end-to-end throughput of opportunistic routing with Type-I HARQ over a linear multihop wireless network. Special emphasis was given on the impact of the number of hops and the path loss on the performance advantages with respect to multihop routing, by studying the asymptotic performance in the bandwidth-limited and power-limited regimes. Our study quantifies the gains achievable in the high-SNR (bandwidth-limited) regime for both fixed and optimized transmission rates, while showing negligible advantages in the low SNR (power-limited regime). Given the effectiveness of multihop routing in the power-limited regime, these results make opportunistic routing an excellent candidate to solve the performance limitations identified in [50,51] for multihop routing in the bandwidth-limited regime (see Section 3.1). We finally remark that this work relies on several simplifying assumptions, and does not account for spatial reuse, and thus interference, [50,52]. Moreover, it should be noted that opportunistic routing, when considering other criteria such as delay and congestion, may not always be desirable (see, e.g., [59]). These aspects will be considered in future work.

In the second part of the chapter, a novel approach to regulate the coexistence of primary and secondary nodes in multihop networks based on spectrum leasing and opportunistic routing is proposed. In particular, it is proposed that primary nodes may, in a local and dynamic fashion, select secondary nodes as next hops for primary traffic by allowing the latter to exploit the spectral resource for secondary data with QoS guarantees. This approach is an implementation of the previously proposed idea of spectrum leasing via cooperation. We have designed different routing strategies based on this principle that provide different trade-offs between gains in terms of primary throughput and energy. Moreover, we have shown that secondary nodes can optimally multiplex primary and secondary traffic using a physical layer strategy known as superposition coding. Our results demonstrate the effectiveness of the proposed paradigm.

Chapter 4

Conclusions

In this Thesis we discussed several issues concerning interference, multiuser communications and radio spectrum availability in wireless networks. We adopted a broad approach that permits to both highlight the main limitations in complex wireless networks, such as MIMO ad hoc networks with multiple simultaneous access, and develop a new cooperative paradigm in cognitive radio networks.

In the first part of the work we evaluated the performance of a MAC protocol for MIMO ad hoc networks in the presence of channel estimation errors, by analytically modeling the interference contribution due to multiuser communications in the expression of the channel estimation error. Indeed, the presence of several simultaneous and asynchronous signals in these types of networks, where advanced PHY techniques such as spatial multiplexing and layered multiuser detection are considered, makes the problem of the channel estimation significantly more complicated than in traditional channel estimation. We considered the analytical model for the computation of the statistics of the channel estimation error, and we developed two analysis both correlator-based and MMSE channel estimators. Moreover, we also highlighted the direct dependence of the channel estimation error on the instantaneous channel matrix. To make the analysis even more general, we also analyzed the interplay between the transmit pulse shape and the variance of the channel estimates gathered by a correlator-based channel estimator. Moreover, we quantified the bias of channel estimates due to cross-talk among different training sequences as well as among training sequences and data symbols and the parameters of the inverse proportionality relationship tying the variance of this bias and the length of the training sequence. These parameters are related to

pulse parameters such as the roll-off factor of RC pulses or the variance factor of Gaussian pulses. The analysis for the estimation errors has then been inserted into a simulator for MIMO ad hoc networks with a detailed MAC implementation and used to observe how PHY-level parameters (e.g., length of the training sequences and number of antennas) affect higher-level metrics (e.g., throughput, efficiency, delay and success ratio). The robustness of networking protocols against channel estimation inaccuracies and interference has been evaluated with our approach, which makes it also possible to understand how to control the impact of channel errors in order to achieve prescribed tradeoff points among MAC-level metrics. Finally, we analyzed the impact of different pulse shapes on the performance of the network.

In the second part of the Thesis we linked the cognitive radio and the cooperation principles, designing a new paradigm that both allows the secondary users to transmit in the wireless medium and exploits the concepts of opportunistic routing and reward in cooperation. At first, we studied the end-to-end throughput of opportunistic routing with Type-I HARQ over a linear multihop wireless network, focusing, in particular, on the impact of the number of hops and the path loss on the performance advantages with respect to multihop routing. This allows to prove and quantify the gains achieved by opportunistic routing in the high-SNR (bandwidth-limited) regime for both fixed and optimized transmission rates, while showing negligible advantages in the low SNR (power-limited regime). Given the effectiveness of multihop routing in the power-limited regime, these results make opportunistic routing both an excellent candidate to solve the performance limitations for multihop routing in the bandwidth-limited regime and a good starting point for the development of a novel approach that regulates the coexistence of primary and secondary nodes in multihop networks. More specifically, we proposed a framework where primary nodes may, following the spectrum leasing paradigm, select secondary nodes as next hops for primary traffic by allowing the latter to exploit the spectral resource for secondary data with QoS guarantees. We designed different routing strategies based on this principle, which allow to provide different trade-offs between gains in terms of primary throughput and energy consumption. Moreover, we showed the optimality of superposition coding, proving that it is the best among all possible multiplexing schemes to be employed by the secondary nodes. Finally, with numerical results we demonstrated the effectiveness of the proposed

paradigm in a linear network. The performance obtained by this new approach justified the extension to a more realistic scenario, the distributed network (see Appendix C). Here, two heuristic policies with low complexity have been proposed following the guidelines discovered in the linear network case. Also in this case, numerical results demonstrated the throughput and energy gains attainable by the proposed spectrum leasing approach by the primary network, while allowing also the secondary nodes to transmit. Moreover, we showed that the heuristic policies provide flexible solutions which permit to almost reach the performance of the optimal policy. In particular, we observed that the performance of the proposed spectrum leasing via cooperative opportunistic routing technique is better in a distributed network than in a network with linear topology. This gain is due to the improved multiuser diversity offered by the distributed network, where all nodes are placed at different distance with respect to each other.

4.1 Future directions

Results in the first part of the Thesis show that one of the main limiting factors on the network performance is the presence of unwanted signals at the receiving side. The mechanisms to control transmissions in order to limit interference in scenarios where channel estimation errors are considered remain as a future area of research. To do this, one would have to evaluate the interference level during the signaling phase to maximize the success reception probability during the data transmission, thus implementing some sort of interference prediction. With this approach, receivers would be able to better control their own transmissions in a distributed way. In order to further increase the ability of the receivers to confirm during the signaling phase only transmissions that will be successfully decoded during the data phase (thereby increasing success ratio and throughput), one would need to derive the probability that a node fails the reception of a signal, given that the previous signals are correctly received (here we are considering a V-BLAST SIC scheme for multiuser detection).

In the second part of the Thesis we proposed a new paradigm for cognitive radio networks. We note that our work relies on several assumptions, that permit, however, to obtain insightful results on the framework introduced. At first, we studied the performance of opportunistic routing in a linear network, adopting as a metric to choose the next relay its

distance to the destination. Extensions of this approach to more general cases remain as a future area of work. For example, other relay selection criteria could be considered to balance the tradeoff between routing the packets along the shortest path and distributing the traffic across the networks. Moreover, one could take into account the effects of spatial reuse, and thus interference, in networks where opportunistic routing is adopted. Given these future directions for opportunistic routing, we plan to insert these improvements in cognitive radio networks, where spectrum leasing is adopted. Moreover, we plan to relax the secondary QoS requirements from the actual parameter ψ (see Section 3.2.1.1), which depend on a single-hop transmission, to a more general constraint that depends on the end-to-end transmission of the secondary packet in a multihop scenario.

From the discussions and the results presented in this thesis it is clear that interference plays a fundamental role in the design of wireless communication systems. In fact, the study of its effects in complex networks and the proposal of counter-measures still remain as a very promising future area of research. In this Thesis some work in this direction has been done, mainly on the quantification of the interference effects in ad hoc networks and the interference management in cognitive radio networks.

Appendix A

End-to-End Throughput Expression of Opportunistic Routing

The non-recursive closed-form expression of the end-to-end throughput for fixed transmission rate R of opportunistic routing given in (3.15) in Section 3.1.2.2 is very difficult to obtain in general (see Remark 3.1.1). In the following we give the throughput expression for $k = 3$ and $k = 4$.

A.1 Throughput for $k = 3$

The closed-form expression of the throughput for a linear network with opportunistic routing (see (3.15)) for $k = 3$ is given by

$$T_{opp}(3,R) = \frac{T_{opp,NUM}(3,R)}{T_{opp,DEN}(3,R)} \tag{A.1}$$

where

$$\begin{aligned} T_{opp,NUM}(3,R) = {} & \mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}} \left(1-\left(1-\mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}}\right)\left(1-\mathrm{e}^{\frac{2^R-1}{\left(\frac{2}{3}\right)^{-\eta}\gamma}}\right)\right) \times \\ & \times \left(1-\left(1-\mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}}\right)\left(1-\mathrm{e}^{\frac{2^R-1}{\gamma}}\right)\left(1-\mathrm{e}^{\frac{2^R-1}{\left(\frac{2}{3}\right)^{-\eta}\gamma}}\right)\right) R \end{aligned} \tag{A.2}$$

and

$$T_{opp,DEN}(3,R) = \mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}}\left(1-\mathrm{e}^{\frac{2^R-1}{\gamma}}\right)\left(1-\mathrm{e}^{\frac{2^R-1}{\left(\frac{2}{3}\right)^{-\eta}\gamma}}\right)\left(-\mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}}-\mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}}\left(1-\mathrm{e}^{\frac{2^R-1}{\left(\frac{2}{3}\right)^{-\eta}\gamma}}\right)\right)+$$
$$+\left(\mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}}+\mathrm{e}^{\frac{2^R-1}{\left(\frac{2}{3}\right)^{-\eta}\gamma}}\left(1-\mathrm{e}^{\frac{2^R-1}{\gamma}}\right)\right)\left(-1+\left(1-\mathrm{e}^{-\frac{2^R-1}{3^\eta\gamma}}\right)\left(1-\mathrm{e}^{\frac{2^R-1}{\left(\frac{2}{3}\right)^{-\eta}\gamma}}\right)\right) \tag{A.3}$$

A.2 Throughput for $k = 4$

The throughput for a linear network with opportunistic routing and $k = 4$ is given by

$$T_{opp}(4,R) = \frac{T_{opp,NUM}(4,R)}{T_{opp,DEN}(4,R)} \tag{A.4}$$

where

$$T_{opp,NUM}(4,R) = e^{-\frac{2^R-1}{4^\eta\gamma}}\left(1-\left(1-e^{-\frac{2^R-1}{4^\eta\gamma}}\right)\left(1-e^{\frac{2^R-1}{2^\eta\gamma}}\right)\right)\times$$
$$\times\left(1-\left(1-e^{-\frac{2^R-1}{4^\eta\gamma}}\right)\left(1-e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\right)\left(1-e^{\frac{2^R-1}{2^\eta\gamma}}\right)\right)\times$$
$$\times\left(1-\left(1-e^{-\frac{2^R-1}{4^\eta\gamma}}\right)\left(1-e^{\frac{2^R-1}{\gamma}}\right)\left(1-e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\right)\left(1-e^{\frac{2^R-1}{2^\eta\gamma}}\right)\right)R \tag{A.5}$$

and

$$T_{opp,DEN}(4,R) = \left(1-\left(1-e^{-\frac{2^R-1}{4^\eta\gamma}}\right)\left(1-e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\right)\left(1-e^{\frac{2^R-1}{2^\eta\gamma}}\right)\right)\times$$
$$\times\left(e^{\frac{2^R-1}{2^\eta\gamma}}\left(1-e^{\frac{2^R-1}{\gamma}}\right)\left(1-e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\right)\left(e^{-\frac{2^R-1}{4^\eta\gamma}}+e^{-\frac{2^R-1}{4^\eta\gamma}}\left(1-e^{\frac{2^R-1}{2^\eta\gamma}}\right)\right)+\right.$$
$$\left.+\left(e^{-\frac{2^R-1}{4^\eta\gamma}}+e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\left(1-e^{\frac{2^R-1}{\gamma}}\right)\right)\left(1-\left(1-e^{-\frac{2^R-1}{4^\eta\gamma}}\right)\left(1-e^{\frac{2^R-1}{2^\eta\gamma}}\right)\right)\right)+$$

$$
\begin{aligned}
&+ e^{-\frac{2^R-1}{4\eta\gamma}} \left(1 - e^{\frac{2^R-1}{\gamma}}\right) \left(1 - e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\right) \left(1 - e^{\frac{2^R-1}{2\eta\gamma}}\right) \times \\
&\times \left(e^{-\frac{2^R-1}{4\eta\gamma}} \left(1 - e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\right) \left(1 - e^{\frac{2^R-1}{2\eta\gamma}}\right) \left(e^{-\frac{2^R-1}{4\eta\gamma}} + e^{-\frac{2^R-1}{4\eta\gamma}} \left(1 - e^{\frac{2^R-1}{2\eta\gamma}}\right)\right) + \right. \\
&\left. + \left(e^{-\frac{2^R-1}{4\eta\gamma}} + e^{\frac{2^R-1}{2\eta\gamma}} \left(1 - e^{\frac{2^R-1}{\left(\frac{3}{4}\right)^{-\eta}\gamma}}\right)\right) \left(1 - \left(1 - e^{-\frac{2^R-1}{4\eta\gamma}}\right) \left(1 - e^{\frac{2^R-1}{2\eta\gamma}}\right)\right)\right)
\end{aligned} \tag{A.6}
$$

Appendix B

Transition Probabilities for Spectrum Leasing Policies

In this Appendix, we define the transition probabilities of the policies described in Section 3.2.2 needed to calculate the matrix (3.57).[1] In order to keep the expressions simple, we will use the following notations:

1. $P_{\text{out,TP}}(a) = P_{\text{out,TS}}(a) = P_{\text{out,P}}(\Delta_a^{(\text{T,D})})$ for transmissions from source (T) to primary (P) or secondary (S) relays;

2. $P_{\text{out,TD}}(k) = P_{\text{out,P}}(k\Delta_{\text{H}})$ for transmissions from source (T) to destination (D);

3. $P_{\text{out,PP}}(a) = P_{\text{out,P}}(a\Delta_{\text{H}})$, $P_{\text{out,SS}}(a) = P_{\text{out,S}}(a\Delta_{\text{H}})$ for transmissions between primary relays (PP) and between secondary relays (SS);

4. $P_{\text{out,PS}}(a) = P_{\text{out,P}}(\Delta_a^{(\text{R})})$, $P_{\text{out,SP}}(a) = P_{\text{out,S}}(\Delta_a^{(\text{R})})$ for transmissions between the two sets of relays, primary to secondary (PS) and vice versa (SP);

5. $P_{\text{out,PD}}(a) = P_{\text{out,P}}(\Delta_a^{(\text{T,D})})$, $P_{\text{out,SD}}(a) = P_{\text{out,S}}(\Delta_a^{(\text{T,D})})$ for primary (P) or secondary (S) transmissions to destination (D).

[1] The material presented in this chapter has been published in [49].

B.1 only-P

The only non-zero submatrix in *only-P* (no spectrum leasing) is $\mathbf{\Phi}_{\text{P,P}}$, that describes the transition probabilities between primary nodes. We have:

$$\mathbf{\Phi}_{\text{P,P}} = \begin{bmatrix} \Phi_{\text{P,P}}(0,0) & \dots & \Phi_{\text{P,P}}(0,k) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{\text{P,P}}(k,k) \end{bmatrix}; \tag{B.1}$$

$$\Phi_{\text{P,P}}(0,0) = \prod_{\ell=1}^{k-1} P_{\text{out,TP}}(\ell) P_{\text{out,TD}}(k); \tag{B.2}$$

$$\Phi_{\text{P,P}}(0,k) = 1 - P_{\text{out,TD}}(k); \tag{B.3}$$

$$\Phi_{\text{P,P}}(0,j) = (1 - P_{\text{out,TP}}(j)) \prod_{\ell=j+1}^{k-1} P_{\text{out,TP}}(\ell) P_{\text{out,TD}}(k), \quad j = 1, \dots, k-1; \tag{B.4}$$

$$\Phi_{\text{P,P}}(i,j) = (1 - P_{\text{out,PP}}(j-i)) \prod_{\ell=j+1}^{k-1} P_{\text{out,PP}}(\ell - i) P_{\text{out,PD}}(k-i),$$
$$i = 1, \dots, k-1, \quad j = i, \dots, k-1; \tag{B.5}$$

$$\Phi_{\text{P,P}}(i,k) = 1 - P_{\text{out,PD}}(k-i), \quad i = 1, \dots, k-1; \tag{B.6}$$

$$\Phi_{\text{P,P}}(k,k) = 1; \tag{B.7}$$

$$\Phi_{\text{P,P}}(i,j) = 0, \text{otherwise.} \tag{B.8}$$

The other submatrices are zero, that is

$$\mathbf{\Phi}_{\text{P,S}} = \mathbf{0}_{[k+1,k-1]}, \quad \mathbf{\Phi}_{\text{S,P}} = \mathbf{0}_{[k-1,k+1]} \quad \text{and} \quad \mathbf{\Phi}_{\text{S,S}} = \mathbf{0}_{[k-1,k-1]}, \tag{B.9}$$

where $\mathbf{0}_{[c,d]}$ is a zero matrix with c rows and d columns.

B.2 only-S

In *only-S* the only primary transmissions allowed are from the source, which leads to submatrices $\mathbf{\Phi}_{\text{P,P}}$ and $\mathbf{\Phi}_{\text{P,S}}$:

$$\mathbf{\Phi}_{\text{P,P}} = \left[\begin{array}{ccccc} \Phi_{\text{P,P}}(0,0) & 0 & \dots & 0 & \Phi_{\text{P,P}}(0,k) \\ \hline & & \mathbf{0}_{[k-1,k+1]} & & \\ \hline 0 & & \dots & 0 & \Phi_{\text{P,P}}(k,k) \end{array}\right]; \tag{B.10}$$

$$\Phi_{\text{P,P}}(0,0) = P_{\text{out,TD}}(k) \prod_{q=1}^{k-1} P_{\text{out,TS}}(q); \tag{B.11}$$

$$\Phi_{\text{P,P}}(0,k) = 1 - P_{\text{out,TD}}(k); \tag{B.12}$$

$$\Phi_{\text{P,P}}(k,k) = 1. \tag{B.13}$$

$$\mathbf{\Phi}_{\text{P,S}} = \left[\begin{array}{ccc} \Phi_{\text{P,S}}(0,1) & \dots & \Phi_{\text{P,S}}(0,k-1) \\ \hline \multicolumn{3}{c}{\mathbf{0}_{[k,k-1]}} \end{array}\right]; \tag{B.14}$$

$$\Phi_{\text{P,S}}(0,j) = P_{\text{out,TD}}(k)\,(1 - P_{\text{out,TS}}(j)) \prod_{q=j+1}^{k-1} P_{\text{out,TS}}(q), \quad j = 1,\dots,k-1. \tag{B.15}$$

Finally, submatrices $\mathbf{\Phi}_{\text{S,P}}$ and $\mathbf{\Phi}_{\text{S,S}}$ reflect the fact that secondary transmissions can reach only other secondary relays or the destination:

$$\mathbf{\Phi}_{\text{S,P}} = \left[\begin{array}{c|c} & \Phi_{\text{S,P}}(1,k) \\ \mathbf{0}_{[k-1,k]} & \vdots \\ & \Phi_{\text{S,P}}(k-1,k) \end{array}\right]; \tag{B.16}$$

$$\Phi_{\text{S,P}}(i,k) = 1 - P_{\text{out,SD}}(k-i), \quad i = 1,\dots,k-1. \tag{B.17}$$

$$\mathbf{\Phi}_{\text{S,S}} = \begin{bmatrix} \Phi_{\text{S,S}}(1,1) & \dots & \Phi_{\text{S,S}}(1,k-1) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{\text{S,S}}(k-1,k-1) \end{bmatrix}; \tag{B.18}$$

$$\Phi_{\text{S,S}}(i,j) = (1 - P_{\text{out,SS}}(j-i))\, P_{\text{out,SD}}(k-i) \prod_{q=j+1}^{k-1} P_{\text{out,SS}}(q-i),$$
$$i = 1,\dots,k-1, \quad j = i,\dots,k-1. \tag{B.19}$$

$\Phi_{\text{A,B}}(i,j) = 0$ with $A, B \in \{P, S\}$, in all other cases.

B.3 P-to-S

Submatrix $\mathbf{\Phi}_{\text{P,P}}$ assumes the same structure of the *only-P* (no spectrum leasing) policy, but the transition probabilities have to consider the presence of the unlicensed network. We

have:

$$\mathbf{\Phi}_{\text{P,P}} = \begin{bmatrix} \Phi_{\text{P,P}}(0,0) & \dots & \Phi_{\text{P,P}}(0,k) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{\text{P,P}}(k,k) \end{bmatrix}; \tag{B.20}$$

$$\Phi_{\text{P,P}}(0,0) = \prod_{\ell=1}^{k-1} P_{\text{out,TP}}(\ell) P_{\text{out,TD}}(k) \prod_{q=1}^{k-1} P_{\text{out,TS}}(q); \tag{B.21}$$

$$\Phi_{\text{P,P}}(0,k) = 1 - P_{\text{out,TD}}(k); \tag{B.22}$$

$$\Phi_{\text{P,P}}(0,j) = (1 - P_{\text{out,TP}}(j)) \prod_{\ell=j+1}^{k-1} P_{\text{out,TP}}(\ell) P_{\text{out,TD}}(k) \prod_{q=j-m}^{k-1} \left[1 + \mathbf{1}_{\{q>0\}} \left(P_{\text{out,TS}}(q) - 1\right)\right], \quad j = 1, \dots, k-1; \tag{B.23}$$

$$\Phi_{\text{P,P}}(i,j) = (1 - P_{\text{out,PP}}(j-i)) \prod_{\ell=j+1}^{k-1} P_{\text{out,PP}}(\ell - i) P_{\text{out,PD}}(k-i) \cdot \prod_{q=j-m}^{k-1} \left[1 + \mathbf{1}_{\{q>0\}} \left(P_{\text{out,PS}}(|q-i|) - 1\right)\right], \quad i = 1, \dots, k-1, \quad j = i, \dots, k-1; \tag{B.24}$$

$$\Phi_{\text{P,P}}(i,k) = 1 - P_{\text{out,PD}}(k-i), \quad i = 1, \dots, k-1; \tag{B.25}$$

$$\Phi_{\text{P,P}}(k,k) = 1. \tag{B.26}$$

In the following submatrix the effect of spectrum leasing and of the *backward window* with parameter m are taken into account to express the transition between primary and secondary relays:

$$\mathbf{\Phi}_{\text{P,S}} = \begin{bmatrix} \Phi_{\text{P,S}}(0,1) & \dots & \Phi_{\text{P,S}}(0,k-1) \\ \vdots & & \vdots \\ \Phi_{\text{P,S}}(k-1,1) & \dots & \Phi_{\text{P,S}}(k-1,k-1) \\ 0 & \dots & 0 \end{bmatrix}; \tag{B.27}$$

$$\Phi_{\text{P,S}}(0,j) = (1 - P_{\text{out,TS}}(j)) \prod_{\ell=(j+m)+1}^{k-1} P_{\text{out,TP}}(\ell) P_{\text{out,TD}}(k) \prod_{q=j+1}^{k-1} P_{\text{out,TS}}(q),$$

$$j = 1, \ldots, k-1; \tag{B.28}$$

$$\Phi_{\text{P,S}}(i,j) = (1 - P_{\text{out,PS}}(j-i)) \prod_{\ell=(j+m)+1}^{k-1} P_{\text{out,PP}}(\ell - i) P_{\text{out,PD}}(k-i) \prod_{q=j+1}^{k-1} P_{\text{out,PS}}(q-i),$$

$$i = 1, \ldots, k-1, \quad j = i, \ldots, k-1; \tag{B.29}$$

$$\Phi_{\text{P,S}}(i,j) = \mathbf{1}_{\{(i-j)\leq m\}} \left(1 - P_{\text{out,PS}}(i-j)\right) P_{\text{out,PD}}(k-i) \prod_{\ell=(j+m)+1}^{k-1} P_{\text{out,PP}}(\ell - i) \cdot$$

$$\prod_{q=j+1}^{k-1} P_{\text{out,PS}}(|i-q|), \quad i = 2, \ldots, k-1, \quad j = 1, \ldots, i-1. \tag{B.30}$$

and $\Phi_{\text{A,B}}(i,j) = 0$ with $A, B \in \{P, S\}$ in all other cases. $\mathbf{\Phi}_{\text{S,P}}$ and $\mathbf{\Phi}_{\text{S,S}}$ are equal to the submatrices of the *only-S* policy in (B.16) and (B.18), respectively.

B.4 P-and-S

Submatrices $\mathbf{\Phi}_{\text{P,P}}$ and $\mathbf{\Phi}_{\text{P,S}}$ are equal to those of the *P-to-S* policy, so they are not reported. However, in this policy the behavior of the secondary network is different, so $\mathbf{\Phi}_{\text{S,P}}$ and $\mathbf{\Phi}_{\text{S,S}}$ are derived considering the presence of primary relays, as limited by the *forward window*. We have:

$$\mathbf{\Phi}_{\text{S,P}} = \begin{bmatrix} 0 & 0 & \Phi_{\text{S,P}}(1,2) & \ldots & \Phi_{\text{S,P}}(1,k) \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \Phi_{\text{S,P}}(k-1,k) \end{bmatrix}; \tag{B.31}$$

$$\Phi_{\text{S,P}}(i,j) = \mathbf{1}_{\{m\leq(j-i)\}} \left(1 - P_{\text{out,SP}}(j-i)\right) \prod_{\ell=j+1}^{k-1} P_{\text{out,SP}}(\ell - i) P_{\text{out,SD}}(k-i) \prod_{q=j}^{k-1} P_{\text{out,SS}}(q-i) \cdot$$

$$\left[1 + \mathbf{1}_{\{m\geq 2\}} \left(-1 + \prod_{t=1}^{m-1} P_{\text{out,SS}}(j-t-i)\right)\right], \quad i = 1, \ldots, k-2, \quad j = i+1, \ldots, k-1; \tag{B.32}$$

$$\Phi_{\text{S,P}}(i,k) = 1 - P_{\text{out,SD}}(k-i), \quad i = 1, \ldots, k-1. \tag{B.33}$$

$$\mathbf{\Phi}_{\text{S,S}} = \begin{bmatrix} \Phi_{\text{S,S}}(1,1) & \ldots & \Phi_{\text{S,S}}(1,k-1) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{\text{S,S}}(k-1,k-1) \end{bmatrix}; \tag{B.34}$$

$$\Phi_{\text{S,S}}(i,j) = (1 - P_{\text{out,SS}}(j-i)) \prod_{\ell=m+j}^{k-1} P_{\text{out,SP}}(\ell - i) P_{\text{out,SD}}(k-i) \prod_{q=j+1}^{k-1} P_{\text{out,SS}}(q-i),$$
$$i = 1, \ldots, k-1, \quad j = i, \ldots, k-1. \tag{B.35}$$

$\Phi_{\text{S,B}}(i,j) = 0$ with $B \in \{P, S\}$, in all other cases.

B.5 Partial secondary deployment

The previous expressions can be derived also for a secondary deployment with $\alpha > 1$. In this case matrix $\mathbf{\Phi}$ (3.57) and vectors $\mathbf{v}$, $\mathbf{1}$, $\mathbf{w}$ and $\mathbf{r}$ in Lemma 3.2.1 (Section 3.2.2.5) have to be reduced in accordance to the number of active secondary relays. This change has effects on the calculation of the transition probabilities derived above. In particular, it is sufficient to modify $P_{\text{out,TS}}(a)$ and $P_{\text{out,SS}}(a)$ by accounting for α and the number of hops covered by packet transmission and $P_{\text{out,PS}}(a)$ as follows

$$P_{\text{out,TS}}(a) = \begin{cases} P_{\text{out,TS}}(a), & \text{rem}(a,\alpha) = 0 \\ 1, & \text{otherwise} \end{cases},$$
$$P_{\text{out,SS}}(a) = \begin{cases} P_{\text{out,SS}}(a), & \text{rem}(a,\alpha) = 0 \\ 1, & \text{otherwise} \end{cases}$$

and, supposing transmission between P_i and S_j

$$P_{\text{out,PS}}(a) = \begin{cases} P_{\text{out,PS}}(a), & \text{rem}(j,\alpha) = 0 \\ 1, & \text{otherwise} \end{cases},$$

and $\text{rem}(a,\alpha)$ is the remainder of the integer division a/α. The other probabilities remain unchanged.

Appendix C

Spectrum Leasing via Cooperative Opportunistic Routing in a General Network Topology

In Section 3.2 we described a novel approach to regulate the coexistence of primary and secondary nodes in multihop networks based on spectrum leasing via cooperative opportunistic routing. With the objective of having both a solvable theoretical model and an insightful analysis of this technique, we considered a simple and tractable linear network topology.[1]

In collaboration with Cristiano Tapparello, a PhD Student of the University of Padova, we considered the issues related to the coexistence of primary and secondary users in a distributed network (see [C6] of List of Publications). The objective of this work has been to both find optimal spectrum leasing policies and design heuristic policies that route a primary packet through primary and secondary transmitters, in an arbitrary topology, in order to maximize the desired trade-off between end-to-end throughput and primary energy consumption for the given secondary QoS requirements.

In this chapter our interest is focused only on the performance of spectrum leasing via cooperative opportunistic routing in a distributed network, with respect to the case where only the primary network is used. So, in the following sections we will not discuss about

[1] The material presented in this chapter has been published in [68].

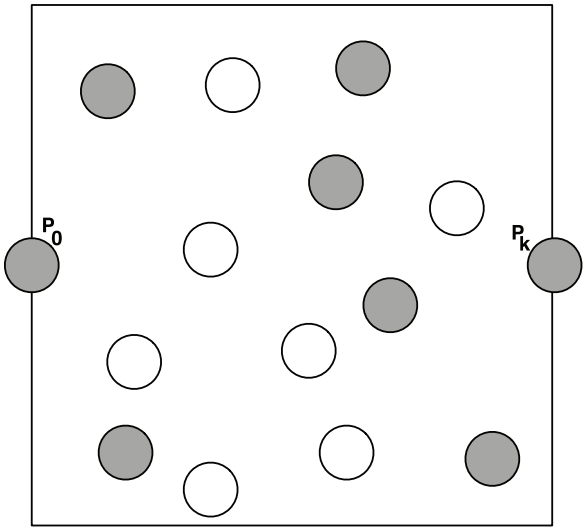

Figure C.1. *A possible representation of a primary distributed network (grey circles) with $N_P = 6$ relay nodes and a secondary network (white circles) with $N_S = 7$ relay nodes.*

the optimal policies (see [C6] and [69] for a complete analysis), but we will focus only on both the description of the heuristic policies and the effectiveness of the proposed paradigm, with respect to the case where spectrum leasing is not considered.

C.1 System Model

Similarly to Section 3.2.1, we consider the communication between the primary source P_0 and the primary destination P_k through two sets of relays. The two sets are composed by N_P primary nodes and N_S secondary nodes, respectively. With respect to Section 3.2.1, here we consider only SC to multiplex primary and secondary messages when a secondary node is the transmitter, since this strategy was proved to be optimal (see Section 3.2.2.6). Moreover, the node arrangement is different from Figure 3.6. In fact, we consider here two sets of relays arbitrarily placed in a square area with normalized side equal to one, where source P_0 and destination P_k are positioned in the middle of two opposite sides (see Figure C.1). The position of each node is static and known by all nodes in the network.

In the next section we describe two heuristic policies that exploit the spectrum leasing via opportunistic routing paradigm. Before doing that, we briefly recall the outage probability for a packet transmitted by the primary node, $P_{\text{out,P}}(d)$ (see (3.44)), or by the secondary node,

$P^{(\mathrm{SC})}_{\mathrm{out,SP}}(d)$ (see (3.49)):

$$P_{\mathrm{out,P}}(d) = 1 - \exp\left(-\frac{2^{\mathrm{R_P}}-1}{\gamma_{\mathrm{P}} d^{-\eta}}\right) \tag{C.1}$$

$$P^{(\mathrm{SC})}_{\mathrm{out,SP}}(d) = 1 - \exp\left[-\min\left(\mathcal{H}^{(1)}_{\mathrm{P}}, \mathcal{H}^{(2)}_{\mathrm{P}}\right)\right], \tag{C.2}$$

where

$$\mathcal{H}^{(1)}_{\mathrm{P}} = \begin{cases} \infty, & 0 \le \psi \le 1 - 2^{-\mathrm{R_P}} \\ \frac{2^{\mathrm{R_P}}-1}{(1-(1-\psi)2^{\mathrm{R_P}})\gamma_{\mathrm{S}} d^{-\eta}}, & 1 - 2^{-\mathrm{R_P}} < \psi \le 1 \end{cases} \tag{C.3}$$

$$\mathcal{H}^{(2)}_{\mathrm{P}} = \begin{cases} \max\left\{\frac{2^{\mathrm{R_S}}-1}{(1-\psi 2^{\mathrm{R_S}})\gamma_{\mathrm{S}} d^{-\eta}}, \frac{2^{\mathrm{R_P}}-1}{\psi\gamma_{\mathrm{S}} d^{-\eta}}\right\}, & 0 < \psi < 2^{-\mathrm{R_S}} \\ \infty, & \psi = 0 \text{ and } 2^{-\mathrm{R_S}} \le \psi \le 1 \end{cases} \tag{C.4}$$

In addition, we define the outage probability $p_{\mathrm{out}}(a,b)$ for any transmission between a transmitter a and a receiver b:

$$p_{\mathrm{out}}(a,b) = \begin{cases} P_{\mathrm{out,P}}(d_{a,b}) & \text{when } a \text{ is a } \textit{primary} \text{ node} \\ P^{(\mathrm{SC})}_{\mathrm{out,SP}}(d_{a,b}) & \text{when } a \text{ is a } \textit{secondary} \text{ node,} \end{cases} \tag{C.5}$$

where $d_{a,b}$ is their distance.

C.2 Heuristic Routing Policies

In this section we detail two heuristic policies that adopt the spectrum leasing via opportunistic routing technique and are suitable to be implemented in real distributed scenarios as they have a reduced complexity with respect to the optimal policies described in [C6].

In order to control the trade-off between the primary energy consumption and the end-to-end throughput we define the parameter K: it represents the maximum number of primary relays that can be used to route a primary packet from the source to the destination, i.e., and not the number of retransmissions performed by the primary nodes.

The K information is contained in the packet header and is decremented by one unit each time a new primary relay is selected. Every time the source transmits a new primary packet (time slot $k = 0$), we have $K_{\mathrm{res}} = K$ and $K_{\mathrm{used}} = 0$ ($K = K_{\mathrm{used}} + K_{\mathrm{res}}$), where K_{res} is the number of primary relays that can still be used and K_{used} represents the number of primary relays already used in the current routing path. In the subsequent time slot $k = 1, 2, 3, \ldots$, if K_{res} is greater than 0 the next relay can be either a primary or a secondary

node, otherwise (i.e., $K_{\rm res} = 0$), the current primary transmitter is the last primary node that can be used along the routing path and subsequent relays must all be secondary nodes. The effect of K is to limit the available multiuser diversity, due to the low number of possible receivers that can act as the next hop. Moreover, secondary users only allocate a portion ψ of the total power for their primary transmissions, so that they can cover a shorter distance with respect to primary transmissions for the same outage probability (assuming they use the same transmission power). Thus, a reduction of the primary energy consumption, which we enforce through K, comes at the expense of a further decreased performance in terms of geographical advancement toward P_k and thus of a throughput low. This will be further discussed in Section C.3.

We now detail the two heuristic routing policies for primary packets.

C.2.1 K-Closer

The objective of the K-Closer policy is to minimize the overall number of transmissions (N in (3.55) in Section 3.2.2) in the network while limiting the maximum energy consumption of primary users through the budget parameter K. Let us consider a generic transmitter node at time slot k, which sends a copy of the packet. All nodes that correctly receive this packet are ranked by the transmitter according to their distance from the destination P_k so that closer nodes have a higher rank.[2] The transmitter chooses as next relay the receiver with the highest rank only if $K_{\rm res} > 0$ (and $K_{\rm res}$ is decremented by one if this receiver is a primary node), otherwise (i.e., $K_{\rm res} = 0$) it grants the secondary node with the highest rank. This process is iterated until the primary packet is correctly received by the destination P_k.

C.2.2 K-One Step Look Ahead (K-OSLA)

The potential drawback of K-Closer is to choose a relay with a small number of neighbors in its proximity, due to the limited amount of information that it uses. Notably, this leads to an increase in the average number of retransmissions that are necessary to reach the next relay. In what follows, we extend the K-Closer heuristic to avoid this situation.

We assume that each relay node a can estimate its proximity δ_a to the destination P_k, that

[2]This implies a feedback mechanism from the receivers to the transmitter, whose design is out of the scope of this work.

is $\delta_a = d_{a,P_k}$. Moreover, relay node a can both collect this metric from all other nodes (both primary and secondary) that are closer to the destination with respect to itself, and build an ordered set $\mathcal{B}(a) = \left\{b_1, b_2, \ldots, b_{|\mathcal{B}(a)|}\right\}$ with them which satisfy the following order: $\delta_{b_i} \leq \delta_{b_{i+1}} \leq \delta_a$ with $i = 1, \ldots, |\mathcal{B}(a)| - 1$. At the same time, node a also determines the ordered subset $\mathcal{B}^S(a) \subseteq \mathcal{B}(a)$, with $\mathcal{B}^S(a) = \left\{c_1, c_2, \ldots, c_{|\mathcal{B}^S(a)|}\right\}$, which only contains the secondary nodes in $\mathcal{B}(a)$. This procedure is carried out for each relay node a of the distributed network, except the destination P_k. Given these informations, node a can determine the geographical advancement toward P_k provided by a relay node b (if it is chosen as next transmitter) as

$$g_{a,b} = \delta_a - \delta_b. \tag{C.6}$$

Moreover, node a can also compute the expected geographical advancement toward the destination P_k by considering the following two metrics:

$$\mathbb{E}[g_a] = \sum_{i=1}^{|\mathcal{B}(a)|} g_{a,b_i}[1 - p_{\text{out}}(a, b_i)] \prod_{j=i+1}^{|\mathcal{B}(a)|} p_{\text{out}}(a, b_j)\,, \tag{C.7}$$

$$\mathbb{E}[g_a^S] = \sum_{i=1}^{|\mathcal{B}^S(a)|} g_{a,c_i}[1 - p_{\text{out}}(a, c_i)] \prod_{j=i+1}^{|\mathcal{B}^S(a)|} p_{\text{out}}(a, c_j)\,. \tag{C.8}$$

In the first case, node a takes into account both primary and secondary nodes, while in the latter only secondary are used.

Finally, we introduce $G_{a,b}$, that represents, with respect to a, the overall expected advancement provided in the next two transmission hops, if node b is selected

$$G_{a,b} = g_{a,b} + \mathbb{E}[g_b]. \tag{C.9}$$

Similarly defined is $G_{a,b}^S$, which considers only secondary nodes as possible relay nodes:

$$G_{a,b}^S = g_{a,b} + \mathbb{E}[g_b^S]. \tag{C.10}$$

The K-OSLA policy is described in the following. Node a has to transmit the primary packet, so it sends a copy of it. Let $\{r_1, \ldots, r_M\}$ be the M receiving nodes that successfully decoded the copy of the packet. If $K_{\text{res}} > 1$, the transmitter a rearranges this set according to the metrics $\{G_{a,r_1}, \ldots, G_{a,r_M}\}$ (see (C.9)) and selects as relay the receiver node $r^* \in \{r_1, \ldots, r_M\}$ with the highest metric G_{a,r^*} (i.e., $G_{a,r^*} \geq G_{a,r_i}\ \forall\ i = 1, \ldots, M$). If r^* is a primary user, K_{res} is decremented by one. When $K_{\text{res}} = 1$, the transmitter a orders the

set $\{r_1, \ldots, r_M\}$ using the metrics defined both in (C.9) and in (C.10). In particular, if r_1 is a primary node the G_{a,r_1} metric is calculated, otherwise (i.e., r_1 is a secondary node) G^S_{a,r_1} is used. This process is done for all receivers r_i, with $i = 1, \ldots, M$. Afterwards, the transmitter a selects as relay the receiver node (primary or secondary) with the highest metric, and if it is a primary user, K_{res} is decremented by one. Finally, if $K_{\text{res}} = 0$, only secondary nodes of the set $\{r_1, \ldots, r_M\}$ are ranked according to the metric defined in (C.10) and the secondary node having the highest metric is selected by the transmitter as the next relay. This procedure is iterated until the packet is correctly received by P_k.

C.3 Results

In this section we present the performance of the heuristic routing policies described above, emphasizing the advantages of spectrum leasing in a distributed network.

For the results in this section we consider a random network with one source P_0, one destination P_k, $N_P = 8$ primary nodes, an equal transmission power for primary and secondary users, i.e., $E_{\text{P}} = E_{\text{S}}$, which yields $\gamma_{\text{P}} = \gamma_{\text{S}} = \gamma$, where we set $\gamma = -5$ dB. Relay nodes are uniformly placed at random in a square area with normalized side equal to one, where source P_0 and destination P_k are positioned in the middle of two opposite sides. For the QoS of the secondary network we consider the following (see also Section 3.2.3). The fraction of power allocated to primary transmissions ψ is computed by imposing an outage probability of $\epsilon_{\text{S}} = 0.1$ for the transmission for a secondary packet, at rate R_{S}, between any two secondary nodes placed at a distance $d_{\text{S}} = 0.1$ (the same of Section 3.2.3). When there are multiple values of ψ that satisfy the QoS requirements (as in 3.2.3), we always select the value of ψ that maximizes the primary throughput, i.e., the highest ψ. We plot the performance of the considered routing schemes in terms of primary end-to-end throughput (3.55) *vs* primary energy consumption (expressed in dB, i.e., $10 \log_{10} E(k, \text{R}_{\text{P}}, \mathcal{Q})$, see (3.56)).

In Figure C.2 we set $\text{R}_{\text{P}} = 3\ bits/s/Hz$, $\text{R}_{\text{S}} = 1\ bits/s/Hz$ and $N_S = 8$. The points in this figure have been obtained by varying ξ in $[0, 1]$ for the optimal policy[3] (Optimal) and

[3]With the term optimal we refer here to policies that minimize, across all the possible evolutions of the system, the expected throughput (*throughput optimal*), the expected total transmission energy expended by primary users (*energy optimal*) or a combination of throughput and primary energy through a weighting factor $\xi \in [0, 1]$ (see [C6] for details).

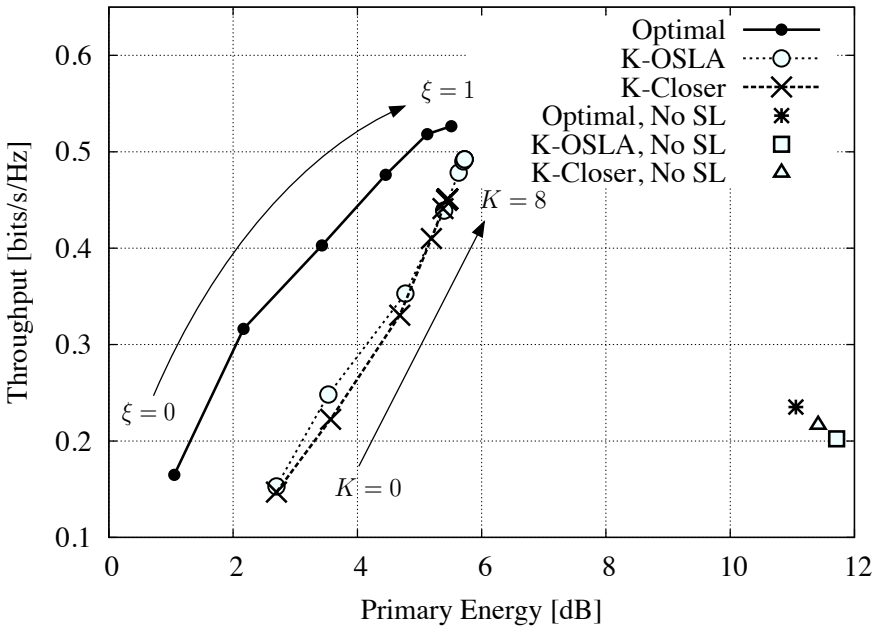

Figure C.2. *End-to-end throughput vs overall primary energy plotted varying $\xi \in [0, 1]$ for the optimal policy (solid line) and $K \in \{0, \ldots, N_S\}$ for the heuristic policies (dotted lines). The results are obtained for $N_P = N_S = 8$, $\gamma = -5$ dB, $\mathrm{R}_P = 3$ $bits/s/Hz$ and $\mathrm{R}_S = 1$ $bits/s/Hz$.*

K in $\{0, \ldots, N_S\}$ for the heuristic policies (K-Closer and K-OSLA). The performance of optimal and heuristic policies when spectrum leasing is not used (indicated in the figure as "No SL") is also shown for comparison. We observe that cooperation via spectrum leasing allows for improved performance in terms of throughput and energy. Both K-Closer and K-OSLA for increasing K provide better throughput performance at the cost of a slightly increased primary energy consumption. This is due to the fact that larger values of K enable the selection of a large number of primary relay nodes. As expected, K-OSLA improves over K-Closer in terms of throughput performance, especially for high values of K ($K \geq 3$ in the figure). In fact, for increasing K the multiuser diversity is higher as more primary nodes can be selected along the path from P_0 to P_k. Note that, as we discuss below, see Figure C.4, the throughput increase of K-OSLA can even be much larger than the one shown in Figure C.2 if we increase $\mathrm{R_S}$ (i.e., the secondary QoS requirements). For the primary energy consumption, as expected, for $K = 0$ (i.e., the relays are all secondary nodes) the energy expenditure of the two schemes is the same. Instead, for $K \geq 1$, K-OSLA has a slightly higher energy consumption with respect to K-Closer and this is due to the fact that the expected

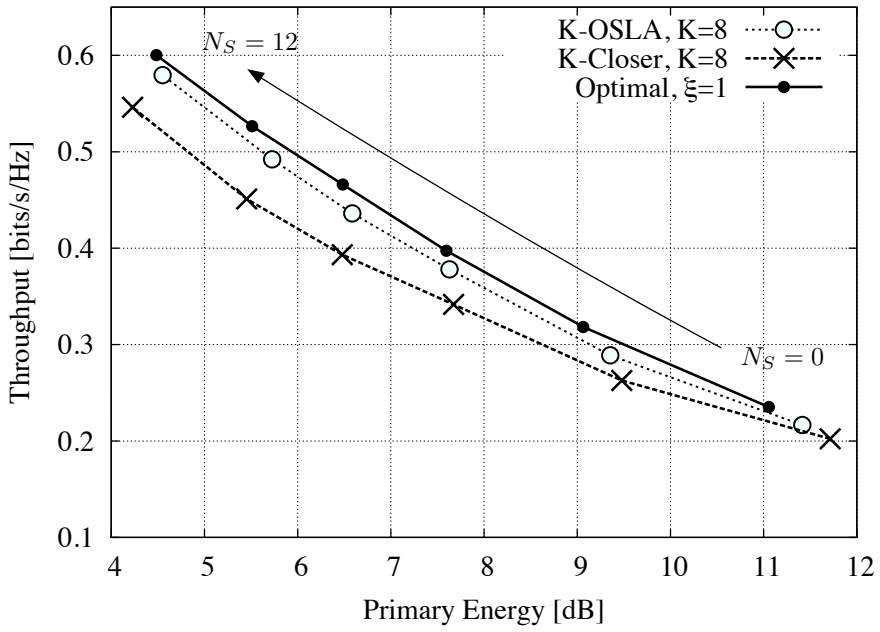

Figure C.3. *End-to-end throughput vs overall primary energy: comparison of optimal throughput policy ($\xi = 1$) and the two heuristic policies with $K = 8$. Each point in the graph represents the pair end-to-end throughput and overall primary energy plotted varying the number of secondary nodes deployed $N_S \in \{0, 2, 4, 6, 8, 12\}$, with $N_P = 8$, $\gamma = -5$ dB, $\mathrm{R}_P = 3$ $bits/s/Hz$ and $\mathrm{R}_S = 1$ $bits/s/Hz$. $N_S = 0$ represents the case where spectrum leasing is not used.*

advancement metric slightly favors primary nodes. In fact, these nodes provide higher expected advancements due to the higher transmission power they use for the transmission of primary packets.

With Figure C.3 we investigate how close heuristic policies can get to the optimal throughput performance ($\xi = 1$). The curves in this figure have been obtained setting $K = 8$ and varying the number of secondary nodes $N_S \in \{0, 2, 4, 6, 8, 12\}$. Here, we note that the usage of spectrum leasing allows for a substantial increase in the throughput (twofold increase) and primary energy performance (gains as high as 6 dB) with respect to the case where only primary transmissions are allowed (i.e., $N_S = 0$). Moreover, K-OSLA approaches the optimal throughput performance for nearly all values of N_S.

In the last figure, we look at the impact of the QoS requirements of secondary users, expressed through $\mathrm{R_S}$, considering $\xi = 1$ for the optimal routing policy and $K = 8$ for the heuristic policies (so as to maximize their throughput performance). Figure C.4 is obtained

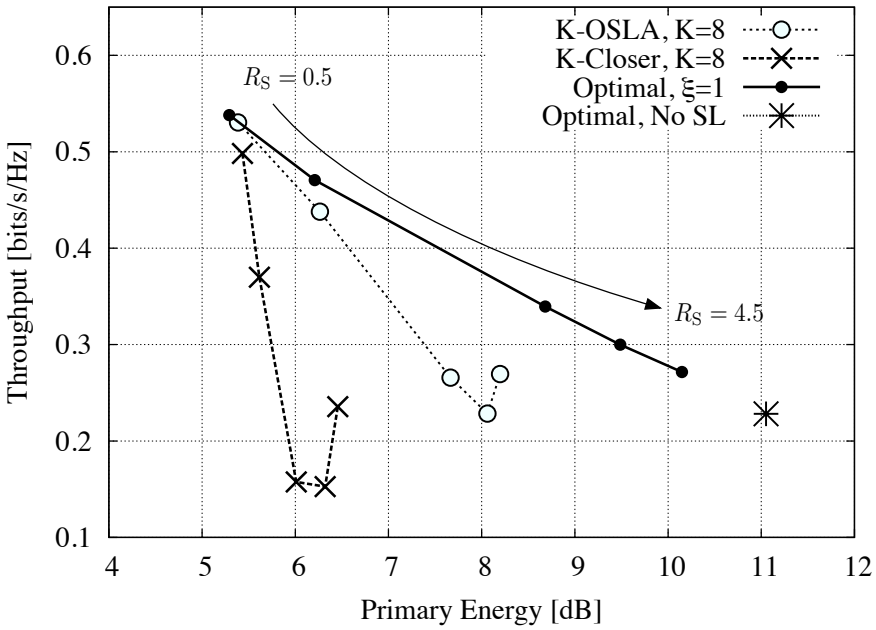

Figure C.4. *End-to-end throughput vs overall primary energy plotted varying* $\mathrm{R_S}$ *for the throughput optimal policy (*$\xi = 1$*) and for the heuristic policies with* $K = 8$, $N_P = 8$, $\gamma = -5$ *dB and* $\mathrm{R}_P = 3$ $bits/s/Hz$*. The curves are obtained by varying* $\mathrm{R_S} \in \{0.5, 1.5, 2.5, 3.5, 4.5\}$ $bits/s/Hz$.

for a fixed number of secondary nodes $N_S = 8$, $\mathrm{R_P} = 3$ $bits/s/Hz$ and varying the secondary transmission rate $\mathrm{R_S} \in \{0.5, 1.5, 2.5, 3.5, 4.5\}$ $bits/s/Hz$. First of all, for increasing $\mathrm{R_S}$ the primary throughput decreases, up to a certain point, for all schemes. This is because an increasing $\mathrm{R_S}$ leads to a smaller coverage for the secondary transmissions, which means that more secondary transmissions (and secondary relays) are needed to reach P_k. However, when $\mathrm{R_S}$ becomes too high (i.e., $\mathrm{R_S} \geq 2.5$ $bits/s/Hz$ in the figure) the secondary nodes are no longer used as relays because the fraction of power ψ that they allocate to primary transmissions is too small to allow the correct reception, at any node, of the primary packets they send. Hence, more and more primary nodes are used and the end-to-end throughput increases. For this same reason, the primary energy always increases for increasing $\mathrm{R_S}$. The latter aspects are emphasized by K-OSLA: this scheme obtains a much higher throughput at the cost of a higher primary energy consumption, especially when $\mathrm{R_S}$ is high. In fact, with respect to K-Closer, K-OSLA uses the expected advancement metric which selects with higher probability a primary node as the next relay (the expected advancement of secondary nodes is smaller than that of primary users). Finally, we note that, for the consid-

ered value of R_S, K-OSLA outperforms the optimal routing policy with no spectrum leasing (here referred to as "No SL"), providing better throughput as well as energy performance.

C.4 Conclusions

In this chapter we extended the approach presented in Section 3.2 for a simpler linear network topology to a more generic distributed network. Two heuristic policies with low complexity were also proposed. Our numerical results lend evidence to the throughput and energy gains that can be attained by the proposed spectrum leasing approach by the primary network, while allowing also the secondary nodes to transmit. Moreover, the heuristic policies are shown to provide flexible solutions that perform close to the optimal policy. Finally, we note that the spectrum leasing via cooperative opportunistic routing technique in distributed scenarios is even more effective than the same approach used in linear network topologies. This is due to the increased multiuser diversity, which comes from the additional degrees of freedom of a two-dimensional node deployment.

List of Publications

The work presented in this thesis has appeared in the articles reported below.

Journal papers

[J1] **D. Chiarotto**, P. Casari, M. Zorzi, "On the impact of channel estimation errors on MAC protocols for MIMO ad hoc networks", *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3290-3300, Oct. 2010.

[J2] **D. Chiarotto**, O. Simeone, M. Zorzi, "Spectrum Leasing via Cooperative Opportunistic Routing Techniques", preliminarily accepted for publication in *IEEE Trans. Wireless Commun.*, Jan. 2011.

Conference papers

[C1] **D. Chiarotto**, P. Casari, M. Zorzi, "On the statistics and MAC implications of channel estimation errors in MIMO ad hoc networks", in *Proc. of IEEE GlobeCom*, New Orleans, LA, 30 Nov. – 4 Dec. 2008, pp. 1–6.

[C2] P. Casari, **D. Chiarotto**, M. Zorzi, "On the tradeoff between MAC-level performance metrics in MIMO ad hoc networks with imperfect channel estimation", in *Proc. of IWCLD*, Palma de Mallorca, Spain, 11–12 Jun. 2009, pp. 1–6.

[C3] P. Casari, **D. Chiarotto**, M. Zorzi, "On the impact of transmit waveforms on channel estimation inaccuracies in distributed MIMO ad hoc networks", in *Proc. of CISS*, Princeton, NJ, 17–19 Mar. 2010, pp. 1–6.

[C4] **D. Chiarotto**, O. Simeone, M. Zorzi, "Throughput and Energy Efficiency of Oppor-

tunistic Routing with type-I HARQ in Linear Multihop Networks", *IEEE GlobeCom*, Miami, FL, 6–10 Dec. 2010.

[C5] **D. Chiarotto**, O. Simeone, M. Zorzi, "Spectrum Leasing via Cooperative Opportunistic Routing", *Asilomar*, Pacific Grove, CA, 7–10 Nov. 2010.

[C6] C. Tapparello, **D. Chiarotto**, O. Simeone, M. Rossi, M. Zorzi, "Optimal Spectrum Leasing via Cooperative Opportunistic Routing in Distributed Ad Hoc Networks", submitted to *IEEE SECON*, Salt Lake City, UT, 27–30 Jun. 2011.

Bibliography

[1] IEEE Standards Department, *ANSI / IEEE Standard 802.11*. IEEE Press, 1999.

[2] U.S. Department of Commerce, "United States Frequency Allocations, The Radio Spectrum." [Online]. Available: www.ntia.doc.gov/osmhome/allochrt.pdf

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[4] H. Jafarkhani, *Space–Time Coding: Theory and Practice*. Cambridge University Press, Sep. 2005.

[5] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V–BLAST: an architecture for realizing very high data rates over the rich–scattering wireless channel," in *Proc. of IEEE ISSSE*, Pisa, Italy, Sep. 1998, pp. 295–300.

[6] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[7] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Trans. Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[8] I. E. Telatar, "Capacity of multi–antenna Gaussian channels," *Eur. Trans. Tel.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[9] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications: a key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[10] P. Casari, M. Levorato, and M. Zorzi, "MAC/PHY cross–layer design of MIMO ad hoc networks with layered multiuser detection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4596–4607, Nov. 2008.

[11] D. Chiarotto, P. Casari, and M. Zorzi, "On the statistics and MAC implications of channel estimation errors in MIMO ad hoc networks," in *Proc. of IEEE GlobeCom*, New Orleans, LA, Nov. 2008, pp. 1–6.

[12] P. Casari, D. Chiarotto, and M. Zorzi, "On the tradeoff between MAC-level performance metrics in MIMO ad hoc networks with imperfect channel estimation," in *Proc. of IWCLD*, Palma de Mallorca, Spain, Jun. 2009, pp. 1–6.

[13] ——, "On the impact of transmit waveforms on channel estimation inaccuracies in distributed MIMO ad hoc networks," in *Proc. of CISS*, Princeton, NJ, Mar. 2010, pp. 1–6.

[14] D. Chiarotto, P. Casari, and M. Zorzi, "On the impact of channel estimation errors on MAC protocols for MIMO ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3290–3300, Oct. 2010.

[15] M. Levorato, P. Casari, and M. Zorzi, "On the performance of access strategies for MIMO ad hoc networks," in *Proc. of IEEE GlobeCom*, Nov. 2006, pp. 1–5.

[16] M. Zorzi, J. Zeidler, A. Anderson, B. Rao, J. Proakis, A. L. Swindlehurst, M. Jensen, and S. Krishnamurthy, "Cross–layer issues in MAC protocol design for MIMO ad hoc networks," *IEEE Wireless Commun. Mag.*, vol. 13, no. 4, pp. 62–76, Aug. 2006.

[17] M. Levorato, S. Tomasin, P. Casari, and M. Zorzi, "Physical layer approximations for cross–layer performance analysis in MIMO–BLAST ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 11, pp. 4390–4400, Dec. 2007.

[18] NTT DoCoMo, "Main technologies of 5Gbps packet transmission experiment." [Online]. Available: www.nttdocomo.com/pr/files/20070209_attachment01.pdf

[19] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advances in Wireless and Mobile Communications - Volume I, Trends in Channel Estimation and Equalization*. Prentice-Hall, September 2000.

[20] S. Verdú, *Multiuser Detection*. Cambridge University Press, 1998.

[21] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[22] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[23] X. Zhang and W. Wang, "Multiuser scheduling analysis of MIMO STBC with imperfect channel estimation," in *Proc. of IEEE MAPE*, Hangzhou, China, Aug. 2007, pp. 217–220.

[24] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 2203–14, May 2006.

[25] Q. Qu, L. B. Milstein, and D. R. Vaman, "Distributed power and scheduling management for mobile ad hoc networks with delay constraints," in *Proc. of IEEE MILCOM*, Washington, DC, Oct. 2006, pp. 1–7.

[26] H. Zhang, N. Mehta, A. Molisch, J. Zhang, and H. Dai, "Asynchronous interference mitigation in cooperative base station systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 155–165, Jan. 2008.

[27] R. Narasimhan, "Error propagation analysis of V-BLAST with channel-estimation errors," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 27–31, Jan. 2005.

[28] M. Guo and S. Yuehong, "Performance analysis of V-BLAST system with two transmit antennas in the presence of channel estimation error," in *Proc. of WiCOM*, Oct. 2008, pp. 1–4.

[29] K. Lee and J. Chun, "Symbol detection in V-BLAST architectures under channel estimation errors," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 593–597, Feb. 2007.

[30] Q. Sun, D. C. Cox, H. C. Huang, and A. Lozano, "Estimation of continuous flat fading MIMO channels," in *Proc. of IEEE WCNC*, Orlando, FL, Mar. 2002, pp. 189–193.

[31] Y. Liu, T. Wong, and W. Hager, "Training signal design for estimation of correlated MIMO channels with colored interference," *IEEE Trans. Signal Processing*, vol. 55, no. 4, pp. 1486–1497, Apr. 2007.

[32] B. Park and T. Wong, "Optimal training sequence in MIMO systems with multiple interference sources," in *Proc. of IEEE GLOBECOM*, Dallas, TX, Dec. 2004.

[33] T. Wong and B. Park, "Training sequence optimization in MIMO systems with colored interference," *IEEE Trans. Wireless Commun.*, vol. 52, no. 11, pp. 1939–1947, Nov. 2004.

[34] N. Czink, G. Matz, D. Seethaler, and F. Hlawatsch, "Improved MMSE estimation of correlated MIMO channels using a structured correlation estimator," in *Proc. of IEEE SPAWC*, New York, NY, Jun. 2005, pp. 595–599.

[35] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1999.

[36] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[37] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.

[38] J. Lv, Y. Lu, Y. Wang, H. Zhao, and C. Y. Han, "Antenna spacing effect on indoor MIMO channel capacity," in *Proc. of IEEE APMC*, Suzhou, China, Dec. 2005.

[39] R. L. Keeney and H. Raffa, *Decisions with Multiple Objectives*. Cambridge, UK: Cambridge University Press, 1993.

[40] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking Spectrum Gridlock With Cognitive Radios: An Information Theoretic Perspective," *Proc. of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[41] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U.Spagnolini, and R. Pickholtz, "Spectrum Leasing to Cooperating Secondary Ad Hoc Networks," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, Jan. 2008.

[42] I. Stanojev, O. Simeone, U. Spagnolini, Y. Bar-Ness, and R. Pickholtz, "Cooperative ARQ via auction-based spectrum leasing," *IEEE Trans. Commun.*, vol. 58, no. 6, Jun. 2010.

[43] Spectrum Efficiency Working Group, "Report of the Spectrum Efficiency Working Group," FCC, Tech. Rep., Nov. 2002.

[44] J. M. Peha, "Approaches to spectrum sharing," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 10–12, Feb. 2005.

[45] J. Acharya and R. Yates, "A Framework for Dynamic Spectrum Sharing Between Cognitive Radios," in *Proc. of IEEE ICC*, Glasgow, Scotland, UK, Jun. 2007, pp. 5166–5171.

[46] I. Stanojev, O. Simeone, Y. Bar-Ness, and C. Myeon-gyun, "On the Optimal Number of Hops in Linear Wireless Ad Hoc Networks with Hybrid ARQ," in *Proc. of WiOPT*, Berlin, Germany, Apr. 2008, pp. 369–374.

[47] D. Chiarotto, O. Simeone, and M. Zorzi, "Throughput and Energy Efficiency of Opportunistic Routing with type-I HARQ in Linear Multihop Networks," in *Proc. of IEEE GlobeCom*, Miami, FL, Dec. 2010.

[48] ——, "Spectrum Leasing via Cooperative Opportunistic Routing," in *Proc. of Asilomar*, Pacific Grove, CA, Nov. 2010.

[49] ——, "Spectrum Leasing via Cooperative Opportunistic Routing Techniques," preliminarily accepted for publication in *IEEE Trans. Wireless Commun.*, Jan. 2011.

[50] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, and T. E. Fuja, "Bandwidth- and Power-Efficient Routing in Linear Wireless Networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2624–2633, Jun. 2006.

[51] O. Oyman and S. Sandhu, "A Shannon-Theoretic Perspective on Fading Multihop Networks," in *Proc. of CISS*, Princeton, NJ, Mar. 2006, pp. 525–530.

[52] A. Bader and E. Ekici, "Performance Optimization of Interference-Limited Multihop Networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 5, pp. 1147–1160, Oct. 2008.

[53] P. Larsson, "Selection Diversity Forwarding in a Multihop Packet Radio Network with Fading Channel and Capture," *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 5, no. 4, pp. 47–54, Oct. 2001.

[54] C. Lott and D. Teneketzis, "Stochastic routing in ad hoc wireless networks," in *Proc. of IEEE Conference on Decision and Control*, Sydney, NSW Australia, Dec. 2000.

[55] M. Zorzi and R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance," *IEEE Trans. Mobile Comput.*, vol. 2, no. 4, pp. 337–348, Oct. 2003.

[56] S. Biswas and R. Morris, "Opportunistic Routing in Multi-Hop Wireless Networks," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 1, pp. 69–74, Jan. 2004.

[57] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading Structure for Randomness in Wireless Opportunistic Routing," in *Proc. of ACM SIGCOMM*, Kyoto, Japan, Aug. 2007, pp. 169–180.

[58] K. Zeng, W. Lou, and H. Zhai, "On End-to-end Throughput of Opportunistic Routing in Multirate and Multihop Wireless Networks," in *Proc. of IEEE INFOCOM*, Phoenix, AZ, Apr. 2008, pp. 816–824.

[59] M. Naghshvar, H. Zhuang, and T. Javidi, "A General Class of Throughput Optimal Routing Policies in Multi-hop Wireless Networks," preprint, http://arxiv.org/pdf/0908.1273.

[60] S. Verdú, "Spectral Efficiency in the Wideband Regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[61] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[62] G. Caire and D. Tuninetti, "The Throughput of Hybrid-ARQ Protocols for the Gaussian Collision Channel," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1971–1988, Jul. 2001.

[63] I. Bettesh and S. Shamai, "Optimal Power and Rate Control for Minimal Average Delay: The Single-User Case," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4115–4141, Sep. 2006.

[64] R. Gallager, *Discrete Stochastic Processes*. Kluwer Academic Publishers, 1996.

[65] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., New York, 2006.

[66] H. Taylor and S. Karlin, *An introduction to stochastic modeling*, 3rd ed. Academic Press, 1998.

[67] M. Uppal, A. Høst-Madsen, and Z. Xiong, "Outage Capacity of the Broadcast Channel in the Low Power Regime," in *Proc. of IEEE ISIT*, Austin, TX, Jun. 2010, pp. 609–613.

[68] C. Tapparello, D. Chiarotto, O. Simeone, M. Rossi, and M. Zorzi, "Optimal Spectrum Leasing via Cooperative Opportunistic Routing in Distributed Ad Hoc Networks," submitted to *IEEE SECON*, Salt Lake City, UT, 27–30 Jun. 2011.

[69] M. Rossi, C. Tapparello, and S. Tomasin, "On Optimal Cooperator Selection Policies for Multi-Hop Ad Hoc Networks," *IEEE Transactions on Wireless Communication, Accepted for Publication*, 2011.

Questa Tesi non ci sarebbe senza Paolo Casari, Osvaldo Simeone, Prof. Michele Zorzi.

Grazie

Desidero ringraziare Riccardo, Giorgio, Marco Rossi, Marco Rotoloni. Ognuno di voi sa perché.