ABSTRACT

Title of dissertation:	MEDIUM ACCESS CONTROL AND NETWORK CODING FOR WIRELESS INFORMATION FLOWS
	Yalin Evren Sagduyu, Doctor of Philosophy, 2007
Dissertation directed by:	Professor Anthony Ephremides Department of Electrical and Computer Engineering

This dissertation addresses the intertwined problems of medium access control (MAC) and network coding in ad hoc wireless networks. The emerging wireless network applications introduce new challenges that go beyond the classical understanding of wireline networks based on layered architecture and cooperation. Wireless networks involve strong interactions between MAC and network layers that need to be jointly specified in a cross-layer design framework with cooperative and non-cooperative users.

For multi-hop wireless networks, we first rediscover the value of scheduled access at MAC layer through a detailed foray into the questions of throughput and energy consumption. We propose a distributed time-division mechanism to activate dynamic transmitter-receiver assignments and eliminate interference at non-intended receivers for throughput and energy-efficient resource allocation based on stable operation with arbitrary single-receiver MAC protocols.

In addition to full cooperation, we consider competitive operation of selfish

users with individual performance objectives of throughput, energy and delay. We follow a game-theoretic approach to evaluate the non-cooperative equilibrium strategies at MAC layer and discuss the coupling with physical layer through power and rate control. As a cross-layer extension to multi-hop operation, we analyze the non-cooperative operation of joint MAC and routing, and introduce cooperation stimulation mechanisms for packet forwarding. We also study the impact of malicious transmitters through a game formulation of denial of service attacks in random access and power-controlled MAC.

As a new networking paradigm, network coding extends routing by allowing intermediate transmitters to code over the received packets. We introduce the adaptation of network coding to wireless environment in conjunction with MAC. We address new research problems that arise when network coding is cast in a cross-layer optimization framework with stable operation. We specify the maximum throughput and stability regions, and show the necessity of joint design of MAC and network coding for throughput and energy-efficient operation of cooperative or competitive users. Finally, we discuss the benefits of network coding for throughput stability in single-hop multicast communication over erasure channels. Deterministic and random coding schemes are introduced to optimize the stable throughput properties. The results extend our understanding of fundamental communication limits and trade-offs in wireless networks.

MEDIUM ACCESS CONTROL AND NETWORK CODING FOR WIRELESS INFORMATION FLOWS

by

Yalin Evren Sagduyu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:

Professor Anthony Ephremides, Chair/Advisor
Professor Prakash Narayan
Professor Dianne O'Leary
Professor Sennur Ulukus
Professor Min Wu

DEDICATION

To my family.

ACKNOWLEDGMENTS

First, I would like to thank my advisor Prof. Ephremides for his guidance and encouragement. I am very fortunate to have him as my mentor. Also, I would like thank the members of my dissertation committee, Prof. Narayan, Prof. O'Leary, Prof. Ulukus and Prof. Wu, for reading my thesis and taking part in my doctoral defense.

I have the endless gratitude for my family. I thank my father Prof. Halil Sagduyu, my mother Nihal Sagduyu, my brother Emre and my love Ariunaa. Nothing would be possible without their support and love.

<div align="center">Table of Contents</div>

List of Tables

List of Figures

Chapter 1

Introduction

It is a fundamental, and still open, problem to characterize and compute the ultimate and achievable communication limits in wireless networks. The general operation of wireless networks as well as the resulting cross-layer interactions have not been fully developed or clearly understood yet. The wireless communication properties introduce new cross-layer design problems [1] that do not exist in wireline networks based on the layered (OSI) network reference model (shown in Figure 1.1) [2]. The classical network operation relies in an essential way on cooperation of nodes for common network tasks. However, wireless networks consist of entities that may cooperate or possibly compete with each other for limited network resources such as bandwidth and energy [3]. Hence, it is necessary to analyze selfish and malicious user behavior in wireless network operation.

This dissertation research is focused on the joint design considerations at the medium access control (MAC) and network layers (together with several implications on the physical layer) in ad hoc wireless networks with cooperative and non-cooperative users. The MAC layer of wireless networks coordinates simultaneous transmissions for reliable communication, whereas network layer operations select routes and forward information units of data packets between source-destination pairs. Power and rate control have strong connections to MAC and network layer

Network Layers	Network Tasks	Network Layers
Application	high level application	Application
Presentation	encryption, compression	Presentation
Session	session coordination	Session
Transport	end–to–end data transfer	Transport
Network	switching and routing	Network
Data link control	error detection	Data link control
MAC	multiple access	MAC
Physical	physical connection	Physical
User 1	Interactions	User 2

Figure 1.1: Open Systems Interconnection (OSI) reference model for wireline networks.

operations, and therefore they need to be jointly designed.

As the ad hoc wireless network model, we assume omnidirectional transmissions that are synchronized into unit time slots. Each node is equipped with a single transceiver and hence cannot simultaneously transmit and receive packets in the same time slot. Therefore, it is necessary to partition the nodes into the disjoint transmitter and receiver sets at any time instant. We do not allow multiple packet transmissions by any node in a single time slot. We assume that nodes receive immediate and correct feedback on the channel outcomes from their receivers through separate conflict-free channels.

We consider two different models to represent the interference effects among simultaneous transmissions. (1) Collision Channel Model: We assume circular transmission (reception) ranges with sharp boundaries. No successful transmission or interference is possible beyond those ranges. We model each link as a classical collision channel with three possible channel outcomes: idle, success, and collision that occur,

respectively, if none, one or more than one packets simultaneously reach the same receiver in the same time slot. This model is also extended to allow probabilistic packet captures of simultaneous transmissions. (2) Physical Channel Model: Any transmission is successfully received, if the Signal-to-Interference-plus-Noise-Ratio (SINR) achieved at the receiver exceeds a fixed threshold. This model incorporates transmission powers, channel gains and additive (Gaussian) noise power.

To have a good understanding of the cross-layer design issues and complex trade-offs imposed by wireless communication properties, we first consider the problem of MAC in multi-destination networks from the different perspectives of stability, distributed operation, throughput and energy efficiency.

We propose a two-step time-division mechanism as a polynomial-time solution to throughput and energy-efficient link scheduling and resource allocation in wireless networks. As the first step, a topology-based greedy heuristic determines the receiver sets to be activated over separate time intervals that depend on the residual battery energies at the individual transmitters. After activating each receiver group separately, the Group TDMA scheme of the second step allocates interfering transmitter groups (with packets addressed to different destinations) within disjoint time fractions to decouple the feedback from different destinations for stable operation. A linear program formulation chooses the time allocation of different transmitter groups to optimize the stable throughput rates with reliable feedback. For any activated transmitter group, we rely on an arbitrary single-receiver (either contention-based or conflict-free) MAC protocol to coordinate transmissions at intended receivers. Distributed implementation is possible through the use of

graph-coloring arguments.

As the network size grows with arbitrary traffic demands and capabilities, it becomes more difficult to manage the cooperation of nodes for common MAC and network layer tasks (for which the individual objectives of nodes may strongly conflict with each other). The cooperation can be realized externally by a central authority, or we can let nodes make individual decisions for distributed operation to optimize their performance objectives that involve throughput rewards, transmission energy costs and packet delay costs. In this context, distributed cooperation reinforcement mechanisms are necessary to improve the network performance.

We formulate a stochastic game framework to extend the wireless network analysis to non-cooperative operation with nodes competing for limited network resources of bandwidth and energy. At the MAC layer, we formulate a non-cooperative random access game of selecting individual probabilities of transmitting packets to a common receiver. Specifically, we derive the transmission strategies in Nash equilibrium depending on the throughput rewards, energy and delay costs. Adaptive best-response update mechanisms are introduced for distributed implementation and the results are compared with the social equilibrium strategies based on full cooperation of transmitters. The analysis of non-cooperative operation is extended to a repeated game model for backlogged packet transmissions. We also incorporate power and rate control in MAC games by further exploiting the capture effects in the communication channel.

Non-cooperative nodes may not be only selfish but may also pursue malicious objectives of blocking random access of the other selfish nodes. The next objective

is to evaluate the impact of malicious transmitters that have the dual objectives of jamming the packet transmissions of the other selfish nodes as well as optimizing their individual performance measures. The analysis provides insights for the optimal strategies to block random access of the other nodes as well as the optimal defense mechanisms against the possible denial of service attacks of malicious nodes. We also extend the game model to power-controlled MAC with a more general channel model based on SINR criterion. The goal is to formulate a non-cooperative game of selecting the individual transmission powers and to evaluate the interactions between selfish and malicious transmitters with throughput and energy objectives.

Next, we extend the non-cooperative MAC operation to a simple form of multi-hop communication over a relay channel. The MAC and plain routing operations are formulated as a joint stochastic game among selfish source and relay nodes competing over collision channels to deliver packets to a common destination node using alternative paths. We rely on a reward mechanism to stimulate cooperation for packet forwarding. In this context, we evaluate the conflicting MAC and routing strategies of direct communication and relaying packets over an intermediate node.

The next question is how we can effectively change the fundamental operation of wireless networks (e.g. how to code, decode and route packets, etc.) to extend the communication limits beyond the classical understanding of replicating and forwarding packets along the path from sources to destinations. We answer this fundamental question using the emerging idea of network coding that extends plain routing operation by allowing intermediate relay nodes to code and decode packets rather than simply forwarding them over predetermined routes. We formulate wireless network

coding in conjunction with conflict-free scheduled access. In this context, we intro-duce the graph-theoretic notion of wireless broadcast cuts and define wireless flows on network graphs by taking into account the omnidirectional nature of wireless transmissions. We derive the optimal linear network codes for wireless networks with a single source node and discuss their properties in conjunction with the un-derlying MAC protocols. Then, we use the subtree graph decomposition methods to present distributed implementation of joint MAC and wireless network coding. The performance is compared to the classical plain routing approach through a de-tailed foray into the questions of coding complexity, throughput, delay, and energy efficiency. We also combine network coding with contention-based random access through the application of Group TDMA method.

The basic assumption of uninterrupted availability of source packets in periodic network coding operation guarantees saturated queues at source and relay nodes. This assumption is reasonable to understand the fundamental capacity limits but cannot result in stable network operation with finite packet delay. Therefore, we need to relax the assumption of saturated queues and explore the network stability problem with packet underflow through a cross-layer optimization framework of MAC and network coding. A single scalar criterion is not sufficient to reflect all communication demands of multiple source-destination pairs and it is necessary to construct a region of attainable transmission rates at which reliable communication can occur. In this context, we address the problem of specifying the maximum throughput and stability regions for multiple source nodes. We also evaluate the trade-offs between throughput and (transmission and processing) energy measures.

For non-cooperative operation, we extend the game formulation of plain routing to network coding operation (in conjunction with MAC) and present reward-based cooperation stimulation mechanisms.

The benefits of wireless network coding are not limited to multi-hop operation. We apply network coding to specify the stable throughput region for single-hop multicast erasure channels with probabilistic reception. We show that the plain policy of retransmitting uncoded packets to multiple receivers is suboptimal. Instead, we introduce dynamic network coding policies based on the instantaneous queue content to optimize the stable operation to the achievable bounds of the maximum throughput rates. We also evaluate the performance of several low-complexity solutions based on random or deterministic network coding.

The dissertation is organized as follows. Chapter 2 addresses the problem of stable MAC in multi-destination wireless networks and proposes a time-division solution to throughput and energy-efficient resource allocation. In addition to cooperative strategies, a game-theoretic approach is followed in Chapter 3 to look at the problem of non-cooperative MAC for single-receiver random access of selfish and malicious transmitters. The analysis is also extended to combine random access with power and rate control for non-cooperative transmissions of backlogged packets. Chapter 4 incorporates a more general SINR-based channel model in the game formulation. Then, Chapter 5 extends the analysis to multi-hop communication and develops a stochastic game of joint MAC and routing among selfish source and relay nodes in a simple relay channel.

As an extension of plain routing, we introduce wireless network coding in

Chapter 6 and jointly develop network coding and MAC in a cross-layer framework of optimizing throughput rates, energy costs and packet delay. Then, we replace the model of continuous packet traffic with stable operation of possibly emptying packet queues and derive the maximum throughput and stability regions for multiple source nodes in Chapter 7. We also evaluate the fundamental trade-offs among throughput and (transmission and processing) energy costs, and study the non-cooperative operation of MAC and wireless network coding. The benefits of network coding are further discussed in Chapter 8 for single-hop multicast erasure channels. Specifically, we show the equivalence of the maximum throughput and stability regions through the application of coded packet retransmissions. Finally, we summarize the dissertation work and discuss future research directions in Chapter 9.

Chapter 2

Medium Access Control (MAC) in Multi-Hop Wireless Networks: A New Look at Multiple Access and Time Division

We study the problem of MAC based on scheduled access through a detailed foray into the questions of throughput and energy consumption in multi-hop wireless networks. We consider destructive interference effects represented by collision channels and rule out simultaneous transmission and reception by any node. This requires partitioning of nodes into disjoint sets of transmitters and receivers at any time instant. Under the assumption of circular transmission (reception) ranges with sharp boundaries, a greedy receiver activation heuristic is developed relying on the network connectivity map to determine distinct receiver groups to be activated within disjoint time intervals. For energy efficiency, we choose the time allocation to each receiver group depending on the residual battery energy available at the respective transmitters. Upon activating each receiver group separately, the additional mechanism of Group TDMA (Time Division Multiple Access) schedules transmissions interfering at non-intended destinations over separate time intervals. This method preserves the reliable feedback information for stable operation. The two-step time-division structure of receiver activation and Group TDMA offers a distributed link scheduling solution to throughput and energy-efficient resource allocation in multi-hop wireless networks.

2.1 Introduction

In wireless access, whether in cellular or general ad hoc networks, throughput and energy efficiency have paramount importance and involve trade-offs that have not been clearly developed or understood yet. For efficient resource allocation, we strip out all complexities of the multi-hop routing operation at the network layer and consider throughput and energy-efficient scheduling of one-hop transmissions. Throughput-optimal channel access scheduling in wireless networks is known to be an NP-complete problem without polynomial-time solutions [4, 5]. For practical use, it is necessary to develop suboptimal heuristics including energy consumptions and node lifetimes as the additional performance measures.

In this chapter, we outline a throughput and energy-efficient MAC approach that allows distributed implementation and supports multi-hop communication as required by autonomous and large scale wireless ad-hoc or sensor networks with high throughput needs and energy constraints.

The extent of studies on multiple access has been traditionally limited to simple networks with multiple transmitters and a single destination. This model is clearly not sufficient to represent the self-organizing wireless networks with multiple number of dynamically changing transmitter-receiver pairs. As an extension of MAC operation to multi-destination networks, the interference effects at non-intended destinations need to be eliminated for feedback reliability in stable operation (as outlined by [6] in conjunction with energy-efficient wireless access). For the case of two fixed receivers, the problem of contention-based access has been studied in [7, 8]

for wireless networks and conflict-resolution algorithms have been used to explore the bounds on the maximum stable throughput. The resulting Group TDMA algorithm separates in time the interfering groups of nodes with packets addressed to different destinations [7]. Each group is assigned separate fractions of time depending on the packet traffic needs. The Group TDMA method was analyzed in terms of throughput properties in [8] for a two-destination network and the optimal time allocation was determined as function of the offered loads independently of the underlying MAC protocol employed within each receiver's area.

We need to extend the model of fixed transmitter-receiver assignments to dynamic wireless networks, where all nodes are both capable and obligated to transmit and receive packets either as parts of source-destination pairs or for relaying purposes (as required by the multi-hop operation in large scale wireless networks). If we further assume that only a single transceiver per node is available, we need to activate nodes either as a transmitter or receiver for disjoint time intervals.

We introduce a greedy receiver activation method based on partial knowledge of the network connectivity map to partition nodes into disjoint transmitter-receiver sets. Rather than ensuring conflict-free schedules as in standard link scheduling, we allow multiple transmission assignments to each receiver and rely on an arbitrary single-receiver (either contention-based or conflict-free) MAC protocol to resolve the unavoidable packet conflicts while satisfying performance measures, such as throughput, energy efficiency or complexity. To obtain reliable feedback information from each receiver, Group TDMA eliminates the secondary conflicts in terms of packet collisions due to transmissions at the non-intended receivers.

11

The predetermined receiver groups are separately activated in a time-division mechanism. We can use the battery energies and node lifetimes as decision criteria in time allocation for distinct receiver groups in order to make best use of finite energy resources. The intuitive idea is to extend node lifetimes by allocating more time to transmissions by those nodes that have the higher residual energy.

In summary, we outline a resource allocation and link scheduling scheme based on a two-step time-division operation. The first step allocates disjoint fractions of time (depending on residual energy at transmitters) to activate distinct sets of receivers (predetermined on the basis of network topology). For each receiver group, the second step creates time orthogonality (based on throughput properties) between transmitter groups interfering at non-intended destinations.

The chapter is organized as follows. We introduce the Group TDMA algorithm for a two-destination network in section 2.2. We extend the analysis to more realistic interference models and arbitrary multi-destination networks in sections 2.3 and 2.4, respectively. We introduce the receiver activation heuristic in section 2.5 and discuss the throughput and energy-efficient time allocation methods in section 2.6. The distributed implementation issues are discussed in section 2.7. We present numerical results in section 2.8 and draw conclusions in section 2.9.

2.2   Group TDMA in Two-Destination Networks

Consider a fixed assignment of disjoint sets of transmitter and receiver nodes. Although scheduled access has distinct advantages (especially at heavy traffic con-

ditions), some form of random access is unavoidable in wireless systems. At a minimum, it is necessary on the reservation sub-channel in dynamic allocation protocols. Therefore, we consider a random-access-based (single-receiver) MAC protocol, although almost any MAC protocol can be assumed. In random access, the splitting-based collision resolution algorithms provide higher stable throughput than the stabilized slotted Aloha algorithm [9]. Splitting a group of collided packets can be implemented based on various criteria, such as coin toss, node or packet ID, time of arrival or residual energy [6]. In this chapter, we use the FCFS (First-Come-First-Served) algorithm as the single-receiver MAC protocol to resolve the primary packet collisions. A new collision resolution period is initiated, whenever a packet collision occurs. All packets that arrive within a specified time allocation interval are transmitted in the first period of time allocation. If there is another collision, the time allocation window is further shortened and the same procedure is repeated, until all packets involved in the original collision are successfully received.

We extend the problem of multiple access to the simple network in Figure 2.1 with two receivers and multiple number of transmitters that are within the reception range of at least one of the receivers. Transmission ranges are circular with sharp boundaries and beyond that range no transmission or interference can be observed. Transmitters generate packets at a common rate and have immediate access to the ternary channel feedback only from their intended destinations. We assume infinite number of unbuffered transmitters in each region. Nodes follow an arbitrary MAC protocol with the maximum stable throughput $S_{\max}$ achievable in the single-destination case.
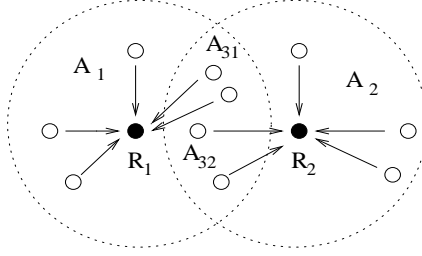
Figure 2.1: A simple two-destination network example for Group TDMA algorithm.

We denote the receivers in Figure 2.1 as R_1 and R_2. For each time instant (or slot), we identify four distinct groups of transmitters. We define A_1 and A_2 as the disjoint transmitter groups in the reception range of only R_1 and R_2, respectively. Nodes from groups A_1 and A_2 randomly generate packets that are transmitted only to R_1 and R_2, respectively. We define A_3 as the group of transmitters that have both receivers in their transmission ranges. Nodes of group A_3 transmit either to R_1 or to R_2 with equal probability. At each time we divide node group A_3 into groups A_{31} and A_{32}, where nodes in A_{3i}, $i = 1, 2$, are transmitting to the receiver R_i.

2.2.1 Group TDMA Algorithm

The problems of deriving optimal channel access schedules for multi-hop networks and network partitioning into activation sets are both NP-complete, and require heuristic suboptimal solutions for practical use. If we use FCFS with the so called first improvement [9] as the single-destination MAC protocol, there is the potential instability problem created by the misinterpretation of the channel feedback, since collisions at any receiver R_i can be caused by packets that are destined for either receiver. This has the equivalent effect of introducing errors in the feedback

signals. If a collision is followed by an idle slot, the current allocation interval is split into smaller subintervals. However, if an idle is misinterpreted as a collision at any receiver R_i, the allocation interval is split indefinitely, resulting in a deadlock situation [9, 7, 8]. For instance, the feedback error at receiver R_1 occurs, if nodes from A_1 and A_{31} are idle, whereas multiple nodes from A_{32} transmit to R_2 and also interfere at receiver R_1.

Group TDMA algorithm solves the feedback reliability problem in two-destination networks, by scheduling transmissions of node groups $\{A_1, A_2\}$, A_{31} and A_{32} over three non-overlapping time fractions of x_1, x_2 and x_3, respectively. Thus, cross-collisions among transmitters with different destinations are ruled out resulting in multiple access operations as in two separate single-destination systems. We define f_1, f_2, f_{31} and f_{32} as fractions of the traffic load generated by groups A_1, A_2, A_{31} and A_{32}, respectively. Then, the total packet arrival rate λ must satisfy

$$\lambda f_i \leq x_1 S_{\max}, \ i = 1, 2, \quad \lambda f_{3i} \leq x_{i+1} S_{\max}, \ i = 1, 2, \tag{2.1}$$

where $0 \leq f_1, f_2, f_{31}, f_{32} \leq 1$, $\sum_{i=1}^{2} f_i + f_{3i} = 1$, $0 \leq x_1, x_2, x_3 \leq 1$, and $\sum_{i=1}^{3} x_i = 1$. The optimal temporal allocation and the maximum stable throughput are given by

$$x_1^* = \frac{\max(f_1, f_2)}{\max(f_1, f_2) + f_{31} + f_{32}}, \tag{2.2}$$

$$x_{i+1}^* = \frac{f_{3i}}{\max(f_1, f_2) + f_{31} + f_{32}}, \ i = 1, 2, \tag{2.3}$$

$$\lambda^* = \frac{S_{\max}}{\max(f_1, f_2) + f_{31} + f_{32}}. \tag{2.4}$$

Clearly, λ^* can improve, if we allow node group A_1 (or A_2) to transmit also during A_{31}'s time x_2 (or A_{32}'s time x_3). Hence, λ^* is actually a lower bound on the general maximum throughput.

2.2.2 Throughput Efficiency of Group TDMA Algorithm

For a network of two receivers and unlimited transmitter population, we consider alternative multiple access schemes of (a) activating all transmitter-receiver pairs simultaneously, (b) activating receivers one at a time. We define S'_{max} and S_{max} as the maximum stable throughput achievable by the unlimited node population (for a single destination) with the neighboring node interference and without interference from the adjoining groups.

Theorem 2.2.1 *The maximum stable throughput rates λ_a^* and λ_b^* under MAC schemes (a) and (b) are given by*

$$\lambda_a^* \leq \frac{S'_{\text{max}}}{\max(f_1, f_2) + f_{31} + f_{32}}, \tag{2.5}$$

$$\lambda_b^* = S_{\text{max}}. \tag{2.6}$$

Proof: MAC scheme (a): Because of the coupling between MAC operations we have to use a single-receiver MAC protocol with stable throughput $S'_{\text{max}} \leq S_{\text{max}}$ (e.g. FCFS algorithm without the first improvement. We can find an upper bound on the stable throughput, if we combine the transmissions of A_{31} and A_{32} with the other (intended or non-intended) transmissions to R_2 and R_1, respectively (i.e. we assume that A_{31} and A_{32} also contribute to the throughput for receiver R_2 and R_1, respectively). The resulting stable throughput λ_c satisfies

$$\lambda_c(f_i + f_{31} + f_{32}) \leq S'_{\text{max}} \leq S_{\text{max}} , \; i = 1, 2. \tag{2.7}$$

The optimal solution for $\lambda_a^* \leq \lambda_c^* \leq \lambda^*$ is given by Eq. (2.5).

16

MAC scheme (b): If we separately activate the receivers R_1 and R_2 for y_1 and y_2 fractions of time ($0 \leq y_1, y_2 \leq 1$, $y_1 + y_2 = 1$), then the resulting stable throughput λ_b satisfies

$$\lambda_b(f_i + f_{3i}) \leq y_i \, S_{\max} \,, \ i = 1, 2, \tag{2.8}$$

and the optimal solution for $\lambda_b^* \leq \lambda^*$ is given by Eq. (2.6). $\qquad\square$

Theorem 2.2.2 *The Group TDMA algorithm can achieve a higher maximum stable throughput than the MAC schemes (a) and (b).*

Proof: Since $\max(f_1, f_2) \leq f_1 + f_2$ and $S'_{\max} \leq S_{\max}$, λ^* in Eq. (2.4) is greater than the throughput rates λ_a^* and λ_b^* given by Eqs. (2.5) and (2.6), respectively. $\qquad\square$

The time allocation could be also based on the energy efficiency objectives. Let $\mathcal{E}_i(\lambda, x_i)$ and $\mathcal{E}_{3i}(\lambda, x_3)$ denote the rates of the cumulative energy consumptions of transmitter groups A_i and A_{3i}, $i = 1, 2$, respectively. For any fixed throughput rate $\lambda < \lambda^*$, this would lead to the alternative problem of choosing the time fractions $\underline{x}$ to minimize the total energy consumption rate given by

$$\mathcal{E}(\lambda, \underline{x}) = \sum_{i=1}^{2} x_1 \left(\mathcal{E}_i(\lambda, x_1) + x_{i+1} \, \mathcal{E}_{3i}(\lambda, x_{i+1}) \right) \tag{2.9}$$

subject to the stability conditions (2.1). For numerical results, we let the FCFS algorithm operate as the single-receiver algorithm under Group TDMA. We define $S = \frac{\lambda}{2}$ as the stable throughput per destination and assume unit energy consumption per transmission. We illustrate in Figure 2.2 the maximum stable throughput per destination and energy consumption rate per time slot (to achieve the maximum stable throughput) as function of f_3 (provided that $f_1 = f_2$ and $f_{31} = f_{32}$).
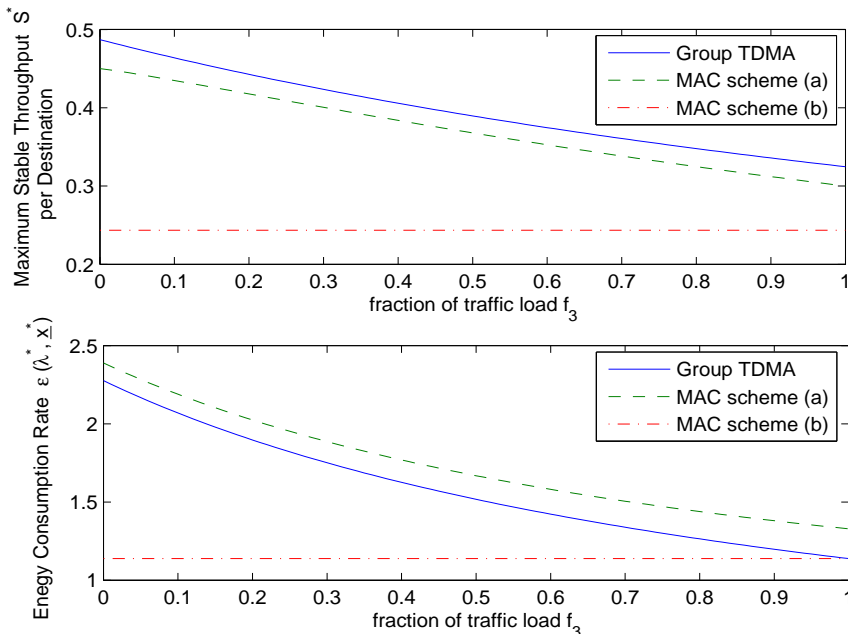
Figure 2.2: The maximum stable throughput and energy consumption rate.

2.3   Group TDMA under Protocol Model

We can also adapt a more realistic criterion for successful packet reception, namely the Protocol model [10] that extends the interference effects beyond transmission ranges such that all nodes have the common range r for transmissions and any node X_i successfully transmits to the intended receiver $X_{R(i)}$, if and only if

$$|X_i - X_{R(i)}| \leq r \ \text{ and } \ |X_k - X_{R(i)}| \geq (1 + \Delta)\, r \qquad (2.10)$$

for every other concurrent transmitter X_k, $k \neq i$, where the quantity $\Delta \geq 0$ accounts for a guard zone that prevents a neighboring node from transmitting over the same single channel at the same time. Figure 2.3 shows the two-destination network to illustrate the Group TDMA operation under Protocol model. We par-
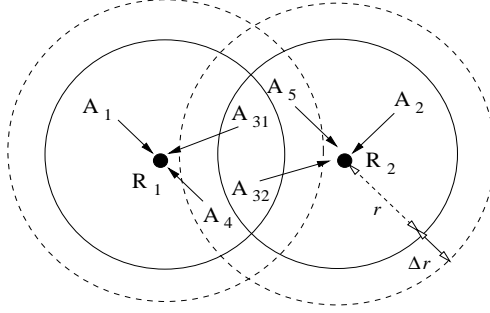
Figure 2.3: Two-destination network operating under Protocol model.

tition the network into five subregions with distinct transmission and interference properties. Regions 1 and 2 contain nodes that have only receivers R_1 and R_2 as their destinations, respectively, and cannot cause interference at the other receiver. Nodes in the reception ranges of both receivers are included in region 3. On the other hand, region 4 consists of nodes that are in the reception range of R_1 but can also interfere at R_2. Similarly, region 5 consists of nodes that are in the reception range of R_2 but can also interfere at R_1.

We denote by f'_i the fraction of the traffic load generated by transmitters in region $i = 1, ..., 5$. The set of nodes in region i is denoted by A_i. We partition A_3 into two subgroups $A_{3,1}$ and $A_{3,2}$ such that $A_{3,i}$, $i = 1, 2$, with traffic load $f'_{3,i}$ contains nodes with packets destined to receiver R_i. Group TDMA allocates the transmissions from $\{A_1, A_2\}$, $\{A_{3,1}, A_4\}$ and $\{A_{3,2}, A_5\}$ within disjoint time fractions x'_1, x'_2 and x'_3, respectively. Thus, the interference effects at non-intended receivers are eliminated and we obtain reliable feedback in the presence of additional interference effects of the Protocol model. The total stable throughput λ (achievable at

two destinations by the Group TDMA algorithm) must satisfy

$$\lambda f_i' \leq x_1' S_{\max}, i = 1, 2, \tag{2.11}$$

$$\lambda(f_{3i}' + f_{i+3}') \leq x_{i+1}' S_{\max}, \ i = 1, 2. \tag{2.12}$$

The optimal temporal allocation and the maximum stable throughput are

$$x_1'^* = \frac{\max(f_1', f_2')}{\max(f_1', f_2') + f_3' + f_4' + f_5'}, \tag{2.13}$$

$$x_{i+1}'^* = \frac{f_{3i}' + f_{i+3}'}{\max(f_1', f_2') + f_3' + f_4' + f_5'}, i = 1, 2, \tag{2.14}$$

$$\lambda^* = \frac{S_{\max}}{\max(f_1', f_2') + f_3' + f_4' + f_5'}. \tag{2.15}$$

Since $f_i = f_i' + f_{i+3}'$ and $f_{3i} = f_{3i'}$, $i = 1, 2$, we have $\max(f_1', f_2') + f_4' + f_5' \geq \max(f_1, f_2) = \max(f_1' + f_4', f_2' + f_5')$, and the maximum stable throughput is a decreasing function of $\Delta \geq 0$. Thus, the increase in interference effects reduces the maximum throughput of the Group TDMA algorithm. For $f_1' = f_2'$, $f_4' = f_5'$ and $f_{3,1}' = f_{3,2}'$, Figure 2.4 depicts the maximum stable throughput per destination S^*, namely $\frac{\lambda^*}{2}$, as function of the load fraction f_3' for different values of Δ. For the rest of the chapter, we continue with the initial assumption of $\Delta = 0$, although similar results can be derived for the general case with $\Delta \geq 0$.

2.4 Group TDMA in General Multi-Destination Systems

We illustrate the Group TDMA operation in a multi-destination network using the simple network with three activated receivers shown in Figure 2.5. If nodes $\{1, 6, 10\}$ are the activated receivers, we groups nodes $\{2, 5\}$ in A_1, node 7 in A_2, and nodes $\{9, 11, 12\}$ in A_3. Nodes 3 and 4 belong to $A_{4,1}$ or $A_{4,2}$ depending on
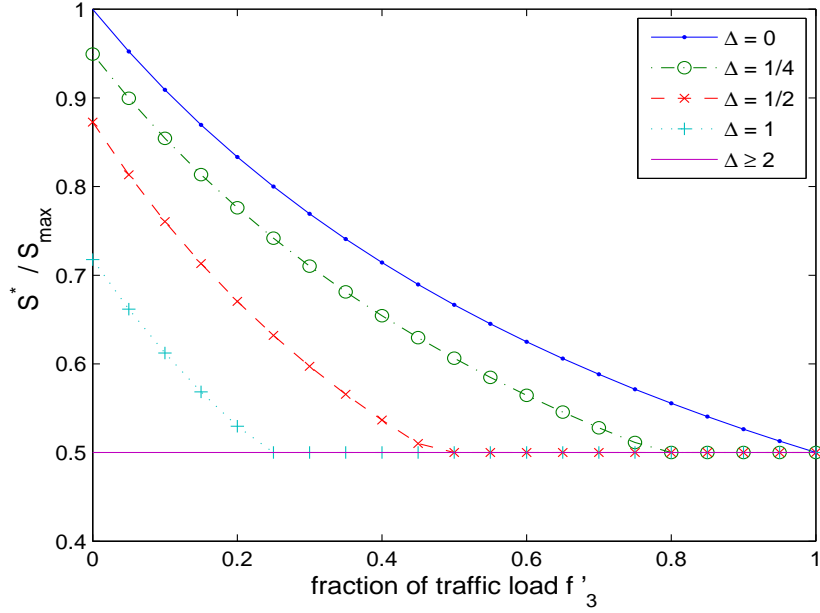
Figure 2.4: The maximum value of the normalized stable throughput $\frac{S^*}{S_{\max}}$.

whether their packets are addressed to node 1 or 6, respectively. Node 8 belongs to $A_{5,2}$ or $A_{5,3}$, if its packets are addressed to node 6 or 10, respectively. Then, nodes from three distinct transmitter groups $\{A_1, A_2, A_3\}$, $\{A_{4,1}, A_{5,3}\}$ and $\{A_{4,2}, A_{5,2}\}$ are activated over non-overlapping time fractions of x_1, x_2 and x_3, respectively, where $0 \le x_i \le 1$, $i = 1, 2, 3$, and $\sum_{i=1}^{3} x_i = 1$.
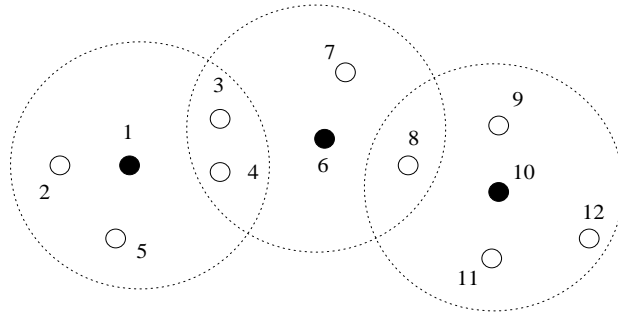
Figure 2.5: Simple multi-destination network model.

2.4.1 Throughput-Optimal Time Allocation

Consider a given set of receivers G_i and a fixed distribution of traffic loads. Nodes transmitting to any receiver in group G_i are divided into disjoint groups $G_{i,k}$, $k = 1, ..., c_i$, such that nodes in the same group can transmit without causing interference to each other at non-intended receivers. The transmitter group $G_{i,k}$ is activated within $x_{i,k}$ fraction of time such that $\sum_{k=1}^{c_i} x_{i,k} = 1$. We define $R_{i,k}^{(j)}$ as the jth element of the receiver set $R_{i,k}$ that can be reached by $G_{i,k}$. We define $A_{i,k}^{(j)}$ as the subset of nodes that belong to the kth transmitter group $G_{i,k}$ and have packets destined to the receiver $R_{i,k}^{(j)}$. The fraction of the traffic load generated by $A_{i,k}^{(j)}$ is $f_{i,k}^{(j)}$. If λ_i is the total rate of packet arrival to nodes transmitting to the receiver group G_i, nodes in $G_{i,k}$, $k = 1, ..., c_i$, jointly satisfy

$$\lambda_i f_{i,k}^{(j)} \leq x_{i,k} S_{\max}, \quad \forall j : R_{i,k}^{(j)} \in R_{i,k}. \tag{2.16}$$

The optimal value of $x_{i,k}$ and λ_i are given by

$$x_{i,k}^* = \frac{\max_{j:R_{i,k}^{(j)} \in R_{i,k}} f_{i,k}^{(j)}}{\sum_{k=1}^{c_i} \max_{j:R_{i,k}^{(j)} \in R_{i,k}} f_{i,k}^{(j)}}, \quad k = 1, ..., c_i, \tag{2.17}$$

$$\lambda_i^* = \frac{S_{\max}}{\sum_{k=1}^{c_i} \max_{j:R_{i,k}^{(j)} \in R_{i,k}} f_{i,k}^{(j)}}. \tag{2.18}$$

As an example, consider a tandem network topology with the fixed set G_i ordered from left to right. Transmitters that are only in the reception range of one receiver belong to the transmitter group $G_{i,1}$. The rest of transmitters are divided into subgroups $G_{i,2}$ and $G_{i,3}$ that consist of nodes transmitting only to the odd and even-numbered receivers, respectively. In general, we have $c_i \leq 3$ and $c_i \leq 13$ for tandem and planar networks, respectively. We consider infinite number of nodes

uniformly distributed on a tandem network of length L or on a two-dimensional planar network of area A. The common transmission radius is r. We define λ_i^* as the maximum stable throughput achievable (under the Group TDMA algorithm) for receiver group G_i, and we define $S_i^* = \frac{\lambda_i^*}{N_i}$ as the maximum stable throughput per destination in receiver group G_i, where N_i is the size of the given receiver set G_i. We assume that the distance between activated receivers is at least r in accordance with the receiver activation that we will introduce in section 2.5.

Theorem 2.4.1 *Consider a tandem network with nodes distributed on a length of $L \gg r$. The quantities λ_i^* and S_i^* satisfy*

$$\frac{S_{\max}L}{4r} \le \lambda_i^* \le \frac{S_{\max}L}{2r}, \tag{2.19}$$

$$\frac{S_{\max}}{4} \le S_i^* \le S_{\max}. \tag{2.20}$$

Proof: (I) Consider the case of tandem networks with the minimum overlapping between reception ranges (i.e. with the minimum number of receivers) such that there exist several non-interfering single-destination systems, as depicted in Figure 2.6-(a). We have $N_i = \frac{L}{2r}$, $f_{i,2}^{(j)} = 0$, $f_{i,3}^{(j)} = 0$, for all j such that $R_{i,j} \in G_i$, and $\sum_{j:R_{i,j} \in G_i} f_{i,1}^{(j)} = 1$. The maximum achievable stable throughput can be computed from Eq. (2.18) as $\lambda_i^* = \frac{S_{\max}}{\frac{1}{N_i}} = \frac{S_{\max}L}{2r}$. The maximum value of the stable throughput per destination, namely $S_i^* = \lambda_i^*/N_i$, is given by $S_{\max}$. This particular case imposes an upper bound on S_i^* for general multi-destination systems.

(II) Next, consider the case with the maximum possible overlapping between reception ranges for all receivers, as shown in Figure 2.6-(b). We have $N_i = \frac{L}{r} + 1$, $f_{i,1}^{(j)} = 0$ for all j such that $R_{i,j} \in G_i$, and $\sum_{j:R_{i,j} \in G_i} (f_{i,2}^{(j)} + f_{i,3}^{(j)}) = 1$. The maximum
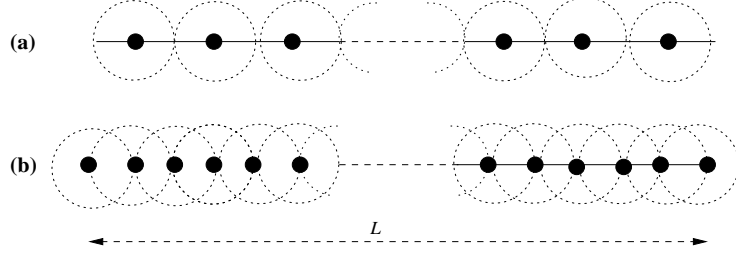
Figure 2.6: Tandem networks with (a) minimum, (b) maximum number of receivers.

value of the total stable throughput can be expressed for the particular tandem network configuration as $\lambda_i^* = \frac{S_{\max}}{\frac{2r}{L}\frac{1}{2}+\frac{2r}{L}\frac{1}{2}} = \frac{S_{\max}L}{2r}$. The maximum stable throughput per destination is computed as $S_i^* = S_{\max}\frac{L}{2(L+r)} \approx \frac{S_{\max}}{2}$, for $L \gg r$.

(III) Finally, consider the hybrid case that includes both non-overlapping and maximally overlapping reception ranges, i.e. $f_{i,k}^{(j_k)} > 0$, $k = 1, 2, 3$, for at least one combination of $\{j_1, j_2, j_3\}$ such that $R_{i,j_k} \in G_i$, $k = 1, 2, 3$. This network configuration maximizes $\sum_{k=1}^{3} \max_{j:R_{i,j}\in G_i} f_{i,k}^{(j)}$ and minimizes λ_i^* according to Eq. (2.18). Figure 2.7 depicts hybrid cases of (a) minimum and (b) maximum number of destinations. The lower bound is $\lambda_i^* = \frac{S_{\max}}{\frac{2r}{L}+\frac{2r}{2L}+\frac{2r}{2L}} = \frac{S_{\max}L}{4r}$. Since $\frac{L-r}{r} \geq N_i \geq \frac{L+3r}{2r}$, the upper and lower bounds on S_i^* are $\frac{2}{3}S_{\max}\frac{L}{L+3r} \approx \frac{2}{3}S_{\max}$ and $S_{\max}\frac{L}{4(L-r)} \approx \frac{S_{\max}}{4}$ for $L \gg r$. The comparison of cases (I)-(III) leads to conditions (2.19)-(2.20). □

Theorem 2.4.2 *Consider a planar network with nodes distributed on an area of* $A \gg \pi r^2$. *The quantities* λ_i^* *and* S_i^* *satisfy*

$$\frac{S_{\max}A}{(3\sqrt{3}+\pi)r^2} \leq \lambda_i^* \leq \frac{S_{\max}A}{\pi r^2}, \tag{2.21}$$

$$\frac{S_{\max}\pi}{3(3\sqrt{3}+\pi)} \leq S_i^* \leq S_{\max}. \tag{2.22}$$

Proof: (I) Consider the case of planar networks with the minimum overlapping be-
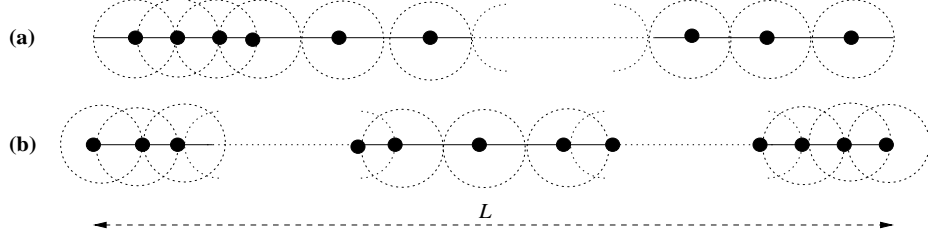
Figure 2.7: The hybrid configurations for tandem networks with (a) minimum, (b) maximum number of receivers.

tween reception ranges (i.e. with the minimum number of the receivers) as depicted in Figure 2.8-(a). If we assume $A \gg \pi r^2$ to eliminate the boundary effects on the distribution of nodes, we obtain $N_i \approx \frac{A}{\pi r^2}$, $f_{i,1}^{(j)} = \frac{\pi r^2}{A}$, $f_{i,k}^{(j)} = 0$, $k = 2, ..., 13$, for all $j : R_{i,k}^{(j)} \in R_{i,k}$. From Eq. (2.18), we obtain $\lambda_i^* = \frac{S_{\max} A}{\pi r^2}$ and $S_i^* = \frac{\lambda_i^*}{N_i}$ is $S_{\max}$.
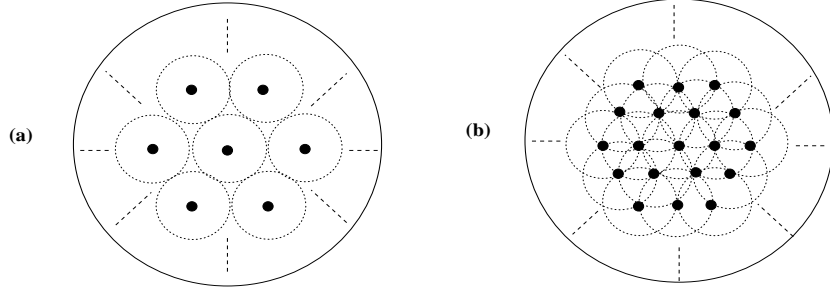
Figure 2.8: Planar networks with (a) minimum, (b) maximum number of receivers.

(II) Next, consider the case with the maximum overlapping between reception ranges of receivers, as shown in Figure 2.8-(b). The non-negative load fractions are $f_{i,k}^{(j)} = \frac{1}{4}\frac{(4\pi - 6\sqrt{3})r^2}{A} + \frac{1}{3}\frac{(6\sqrt{3} - 3\pi)r^2}{A} = \frac{\sqrt{3}r^2}{2A}$ for $2 \leq k \leq 7$ and j such that $R_{i,k}^{(j)} \in R_{i,k}$.

From Eq. (2.18), we have $\lambda_i^* = \frac{S_{\max} A}{3\sqrt{3}r^2}$. If we assume that the distance between receivers is at least r, then successful transmissions to any receiver node should disable transmissions to at most 6 other neighbor receivers. The area of $(mr)^2\pi$

includes at most $1 + \sum_{x=1}^{m} 6x = 1 + 3m + 3m^2$ receivers, where m is a positive integer. For $A \gg \pi r^2$, we have $N_i \approx \frac{3A}{\pi r^2}$, if we ignore the boundary effects. Then, the maximum stable throughput per destination is given by $S_i^* = \frac{S_{\max}\pi}{9\sqrt{3}}$.

(III) Finally, consider the hybrid case that includes both non-overlapping and maximally overlapping reception ranges, i.e. $f_{i,k}^{(j_k)} > 0$, $k = 1, ..., 7$, for at least one combination of $\{j_1, ..., j_7\}$ such that $R_{i,k}^{j_k} \in G_i$ for $k = 1, ..., 7$. We obtain the lower bound $\lambda_i^* = \frac{S_{\max}A}{(3\sqrt{3}+\pi)r^2}$. For $A \gg \pi r^2$, we have at most $\frac{3A}{\pi r^2}$ receivers and the lower bound on S_i^* is $\frac{S_{\max}\pi}{3(3\sqrt{3}+\pi)}$. The comparison of cases (I)-(III) leads to (2.21)-(2.22). $\square$

We cannot directly apply the throughput-optimal solutions of Group TDMA to finite node population, since $S_{\max}$ strongly depends on the number of transmitters.

2.4.2 Group TDMA for Finite Node Population

Instead of revisiting the MAC stability problem with finite number of transmitters [11, 12, 13], we assume that the maximum stable throughput of MAC protocol $S_{\max}(T_i)$ is known as function of the number of transmitters T_i. We define $|A_{i,j}^{(j)}| = T_i f_{i,j}^{(k)}$ as the fixed cardinality of transmitter group $A_{i,k}^{(j)}$. If λ_i denotes the rate of packet generation by nodes transmitting to G_i, then nodes in transmitter group $G_{i,k}$, $k = 1, ..., c_i$, with packets destined to receiver $R_{i,k}^{(j)}$ must jointly satisfy

$$\lambda_i f_{i,k}^{(j)} \le x_{i,k}\, S_{\max}(|A_{i,k}^{(j)}|), \quad \forall j : R_{i,k}^{(j)} \in R_{i,k}, \tag{2.23}$$

where $x_{i,k}$ is the fraction of time allocated to group $G_{i,k}$. If we define $S_{i,k}$ as $\max_{j:R_{i,k}^{(j)}\in R_{i,k}} \{f_{i,k}^{(j)}\,(S_{\max}(|A_{i,k}^{(j)}|))^{-1}\}$.

The optimal value of $x_{i,k}$ is

$$x_{i,k}^* = \frac{S_{i,k}}{\sum_{k=1}^{c_i} S_{i,k}} \tag{2.24}$$

and the maximum value of λ_i is

$$\lambda_i^* = \frac{1}{\sum_{k=1}^{c_i} S_{i,k}}. \tag{2.25}$$

As an example, we consider the activated receiver group G_i with two receivers. Each receiver has two transmitters in their reception ranges, whereas the intersection of reception ranges includes one transmitter. Transmitters are divided into four groups each of size one. Nodes from groups $\{A_{i,1}^{(1)}, A_{i,1}^{(2)}\}$, $A_{i,2}^{(1)}$ and $A_{i,3}^{(2)}$ are activated for $x_{i,1}$, $x_{i,2}$ and $x_{i,3}$ fractions of time, respectively. If traffic load is homogeneously distributed among transmitters, we have $f_{i,1}^{(1)} = f_{i,1}^{(2)} = 1/3$, $f_{i,2}^{(1)} = f_{i,3}^{(2)} = 1/6$. From Eqs. (2.24)-(2.25), the optimal time allocation is $x_{i,1}^* = 1/2, x_{i,2}^* = x_{i,3}^* = 1/4$ and the maximum achievable throughput is $\lambda_i^* = \frac{3}{2} S_{\max}(1)$ where $S_{\max}(1) = 1$ is feasible. On the average, $\frac{3}{4}$ packets per slot can be transmitted to each receiver.

We define $\lambda_{i,a}$ and $\lambda_{i,b}$ as the stable throughput of multiple access schemes (a) and (b) from section 2.2. We find an upper bound on $\lambda_{i,a}$, if we combine transmissions from $A_{i,1}^{(1)}$ with $A_{i,2}^{(1)}$ and transmissions from $A_{i,1}^{(2)}$ with $A_{i,3}^{(2)}$. The stable throughput $\lambda_{i,c}$ of the improved system must satisfy $\lambda_{i,c}(f_{i,1}^{(1)} + f_{i,1}^{(2)} + f_{i,j+2}^{(j)}) \leq S_{\max}(2)$, for $j = 1, 2$, where the optimal value is $\lambda_{i,c}^* = \frac{3}{2} S_{\max}(2) \geq \lambda_{i,a}^*$, which is the maximum throughput of MAC scheme (a). The transmissions to receiver R_j, $j = 1, 2$, are separately activated by MAC scheme (b) over fractions of time $y_{i,j}$ so that we have $\lambda_{i,b}\frac{1}{2} \leq y_{i,j} S_{\max}(2)$, for $j = 1, 2$. The maximum stable throughput is $\lambda_{i,b}^* = S_{\max}(2)$ and $\lambda_i^* \geq \max(\lambda_{i,a}^*, \lambda_{i,b}^*)$, which indicates the throughput efficiency of Group TDMA.

2.5 Receiver Activation in Wireless Networks

We need to partition nodes into disjoint transmitter and receiver sets at any time instant. We assume that each node lies within the transmission (reception) range of at least one other node. An arbitrary node is chosen to initiate the first receiver group. The decision is either random or follows any priority-based rule. Then, the activated receiver designates any node within a fixed receiving range as transmitter. We exclude nodes already selected as transmitter or receiver from the list of receiver candidates and continue with sequential assignments of transmitters and receivers, until all nodes are chosen either as receiver or transmitter at least once. We determine other distinct receiver activation groups, until every node is included in at least one receiver group.

We use the network in Figure 2.5 to illustrate the receiver activation heuristic. We pick node 1 as the first activated receiver. Nodes $\{2, 3, 4, 5\}$ that are in the receiving range of node 1 are selected as transmitters. If node 6 is the second activated receiver, nodes $\{7, 8\}$ become the corresponding transmitters. Similarly, if node 10 is the third activated receiver, nodes $\{9, 11, 12\}$ are selected as transmitters. We exclude nodes $\{1, 6, 10\}$ as receiver candidates and repeat the same procedure, until all nodes are activated as receiver at least once. If network topology does not allow a node to become a transmitter, it is activated as a receiver. For instance, the node sets $\{1, 6, 10\}$, $\{2, 4, 7, 12\}$, $\{5, 3, 8, 12\}$, $\{9, 11, 6, 1\}$ form receiver groups to be activated over disjoint time intervals. Consider a tandem network of length L and two-dimensional planar network of area $A \gg \pi r^2$ with n nodes and transmission

(reception) radius r. We define N_i as the size of receiver group G_i, $1 \le i \le N$, and N as the number distinct receiver groups.

Theorem 2.5.1 *(a) For finite number of n nodes,*

$$1 \le N_i \le \min\left(g(n), \left\lfloor \frac{L}{r} \right\rfloor + 1\right) \quad \text{in tandem networks,} \qquad (2.26)$$

where $g(1) = 1, g(n) = \left\lceil \dfrac{2(n-1)}{3} \right\rceil, n > 1$.

$$1 \le N_i \le \min\left(h(n), \left\lfloor \frac{3A}{\pi r^2} \right\rfloor\right) \quad \text{in planar networks,} \qquad (2.27)$$

where $h(1) = 1, h(n) = \left\lceil \dfrac{12(n-1)}{13} \right\rceil, n > 1$.

For uniform distribution of unlimited node population,

$$\left\lceil \frac{L}{2r} \right\rceil \le N_i \le \left\lfloor \frac{L}{r} \right\rfloor + 1 \text{ in tandem networks,} \qquad (2.28)$$

$$\left\lceil \frac{A}{\pi r^2} \right\rceil \le N_i \le \left\lfloor \frac{3A}{\pi r^2} \right\rfloor \text{ in planar networks.} \qquad (2.29)$$

(b) For finite number of n nodes,

$$\left\lceil \frac{n}{\min(g(n), \lfloor \frac{L}{r} \rfloor + 1)} \right\rceil \le N \le n \text{ in tandem networks,} \qquad (2.30)$$

$$\left\lceil \frac{n}{\min(h(n), \lfloor \frac{3A}{\pi r^2} \rfloor)} \right\rceil \le N \le n \text{ in planar networks.} \qquad (2.31)$$

For uniform distribution of unlimited node population,

$$\frac{1}{\lfloor \frac{L}{r} \rfloor + 1} \le \lim_{n \to \infty} \frac{N}{n} \le 1 \text{ in tandem networks,} \qquad (2.32)$$

$$\frac{1}{\lfloor \frac{3A}{\pi r^2} \rfloor} \le \lim_{n \to \infty} \frac{N}{n} \le 1 \text{ in planar networks.} \qquad (2.33)$$

Proof: (a) We can locate at most $\left\lfloor \frac{L}{r} \right\rfloor + 1$ activated receivers on a tandem network of length L, since the distance between activated receivers is at least r. There are

at most $\left\lfloor \frac{3A}{\pi r^2} \right\rfloor$ activated receivers on a planar network of area A. We define d_i as the maximum number of neighboring receivers that any receiver from G_i can have (such that their reception ranges intersect). We can show by plain geometry that we have at least one node and at most $\lceil d_i(n-1)/(d_i+1) \rceil$ nodes for $n > 1$ (or just one node for $n = 1$) in any receiver group. Since the distance between activated receivers is at least r, we have $d_i = 2$ and $d_i = 12$ for tandem and planar networks. Thus, we get conditions (2.26) and (2.27).

For the case of infinite node population with uniform distribution, all portions of the network must be covered by reception ranges of any activated receiver group to ensure that every node is designated either as receiver or transmitter at any time instant. For a tandem network of length L, we must have at least $\left\lceil \frac{L}{2r} \right\rceil$ activated receivers in each receiver group. This is only possible, if the reception ranges of activated receivers cover the entire network with minimum overlapping. We can locate at most $\left\lfloor \frac{L}{r} \right\rfloor + 1$ activated receivers on a tandem network of length L such that the distance between activated receivers is equal to r. Hence, we obtain condition (2.28). Similarly, we must have at least $\left\lceil \frac{A}{\pi r^2} \right\rceil$ activated receivers in each receiver group for a planar network of area A. Ignoring the boundary effects, we can locate at most $\left\lfloor \frac{3A}{\pi r^2} \right\rfloor$ activated receivers on a planar network of area A with the maximum overlapping between reception ranges and the minimum distance r between receivers. Thus, we obtain condition (2.29).

(b) Receiver activation produces minimum number of distinct receiver groups, if we include maximum number of receivers (with minimum overlapping between reception ranges) in each receiver group and every node appears only in one re-

ceiver group. The lower bound on N is $\left\lceil \frac{n}{\max_i N_i} \right\rceil$. We have $\min(g(n), \lfloor \frac{L}{r} \rfloor + 1)$ and $\min(h(n), \lfloor \frac{3A}{\pi r^2} \rfloor)$ as the maximum values of N_i for tandem and planar networks. We obtain the maximum number of distinct receiver groups, if each group includes only one receiver such that we have n as the upper bound on N. This leads to conditions (2.30) and (2.31). For the unlimited node population, we have $\max_i N_i = \lfloor \frac{L}{r} \rfloor + 1$ and $\max_i N_i = \lfloor \frac{3A}{\pi r^2} \rfloor$ in tandem and planar networks, respectively. The upper bound on N is still n. If we evaluate $\frac{N}{n}$ as n goes to infinity, we obtain conditions (2.32) and (2.33). $\qquad \square$

For tandem networks with uniform node distribution, we illustrate N/n and $\frac{1}{N}(\sum_{i=1}^{N} N_i)$ (averaged over different network topologies) as functions of the number of nodes n in Figure 2.9.

2.6 Time Allocation for Receiver Activation

2.6.1 Throughput-Efficient Time Allocation

We assume infinite-energy systems and apply the Group TDMA algorithm separately for each activated receiver group. We define $F_{i,k}^{(j)}$ as the fraction of traffic load generated by $A_{i,k}^{(j)}$ for the entire network operation such that we obtain $\sum_{i=1}^{N} \sum_{k=1}^{c_i} \sum_{j:R_{i,k}^{(j)} \in R_{i,k}} F_{i,k}^{(j)} = 1$ and $f_{i,k}^{(j)} = F_{i,k}^{(j)} \ (\sum_{k=1}^{c_i} \sum_{j:R_{i,k}^{(j)} \in R_{i,k}} F_{i,k}^{(j)})^{-1}$ for all values of i, j and k. Transmitter group $G_{i,k} = \{A_{i,k}^{(j)}, j : R_{i,k}^{(j)} \in R_{i,k}\}$ is activated for $X_{i,k}$ fraction of time such that $\sum_{i=1}^{N} \sum_{k=1}^{c_i} X_{i,k} = 1$, $X_{i,k} = \tau_i x_{i,k}$ and $\tau_i = \sum_{k=1}^{c_i} X_{i,k}$. If λ is the overall packet arrival rate, then all transmitter nodes in $G_{i,k}$, $k = 1, ..., c_i$,
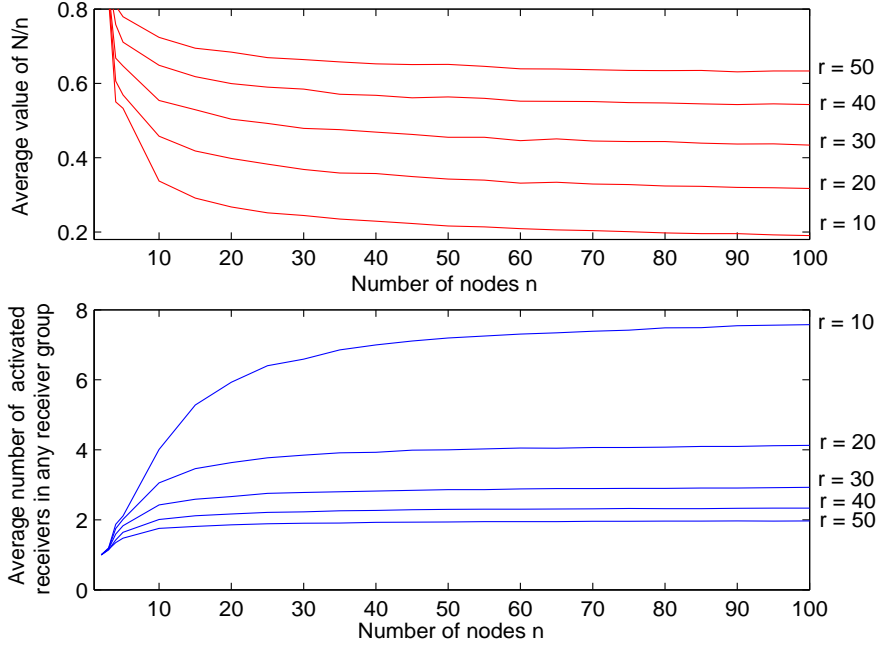
Figure 2.9: Average value of N/n and $\frac{1}{N}(\sum_{i=1}^{N} N_i)$ in tandem networks of length $L = 100$ as functions of the number of nodes n and transmission radius r.

must jointly satisfy

$$\lambda F_{i,k}^{(j)} \leq X_{i,k} S_{\max}, \ \ \forall j : R_{i,k}^{(j)} \in R_{i,k}. \tag{2.34}$$

The optimal temporal allocation can be expressed as

$$X_{i,k}^* = \frac{\max_{j:R_{i,k}^{(j)} \in R_{i,k}} F_{i,k}^{(j)}}{\sum_{i=1}^{N} \sum_{k=1}^{c_i} \max_{j:R_{i,k}^{(j)} \in R_{i,k}} F_{i,k}^{(j)}} \tag{2.35}$$

for any $i = 1, ..., N$ and $k = 1, ..., c_i$. The total maximum throughput achievable at all receivers is

$$\lambda^* = \frac{S_{\max}}{\sum_{i=1}^{N} \sum_{k=1}^{c_i} \max_{j:R_{i,k}^{(j)} \in R_{i,k}} F_{i,k}^{(j)}}. \tag{2.36}$$

The throughput-optimal time allocation $\tau_i^* = \sum_{k=1}^{c_i} X_{i,k}^*$ to receiver group G_i, $i = 1, ..., N$, is

$$\tau_i^* = \frac{\sum_{k=1}^{c_i} \max_{j:R_{i,k}^{(j)} \in R_{i,k}} F_{i,k}^{(j)}}{\sum_{i=1}^{N} \sum_{k=1}^{c_i} \max_{j:R_{i,k}^{(j)} \in R_{i,k}} F_{i,k}^{(j)}}. \tag{2.37}$$

2.6.2 Energy-Efficient Time Allocation

We propose to use residual energies and node lifetimes as measures of time allocation to distinct receiver groups in energy-limited systems. We define RG_m as the receiver group activated in the mth activation period with allocated time of length t_m. The energy of each node is equally dedicated for transmissions to each receiver in its transmission range. We denote by $E_m(G_i)$ the total energy available for transmissions to receiver group G_i before the mth activation period. The objective of the energy-efficient temporal allocation is to maximize the residual system lifetime $LT = \min_{1 \leq i \leq N} LT_i$, where LT_i is the lifetime of energy supplies for transmissions to receiver group G_i. The intuitive solution is $RG_m = \arg \max_i E_m(G_i)$, i.e. we activate at any time instant only the receiver group for which the transmitters have the highest residual cumulative energy.

The optimal solution is $RG_m = \arg \max_i E_m(G_i)$ for all m with $\lim t_m \to 0$. The quantity LT is optimized by load balancing that equalizes the cumulative residual energies $E_m(G_i)$, $i = 1, ..., N$, over successive periods m so that no node group (transmitting to any receiver group) runs out of energy earlier than other groups. Receiver activation period $m + 1$ is initiated, if $E_{m+1}(RG_k) < E_{m+1}(G_i)$ for all i such that $G_i \neq RG_m$. The fraction of time allocated to G_i is given by $\tau_i = \frac{1}{LT} \sum_{m=0}^{LT} t_m 1(RG_m = G_i)$, where 1 is an indicator function. To resolve the

ambiguity in activation order after the period $m^* = \arg\min_{m \in Z_0}\{E_m(G_i) = E_{m^*}$,
$i = 1, ..., N\}$, for a positive constant E_{m^*} and the set of nonnegative integers Z_0,
the optimal value of t_m, $m \geq m^*$, is chosen inversely proportional to $\mathcal{E}_m(G_i)$ for
$RG_m = G_i$, where $\mathcal{E}_m(G_i)$ denotes the rate of change in $E_m(G_i)$ per unit time during
the mth period. Instead of the optimal solution with the infinitesimal activation
durations t_m for $m \geq m^*$, a sub-optimal but practical solution is to activate first
the receiver group with the highest total energy of corresponding transmitters, i.e.
$RG_m = \arg\max_i E_m(G_i)$, and then to replace RG_m with another receiver group for
receiver activation period $m+1$, if $E_{m+1}(RG_m)$ falls below $\min_{G_i \neq RG_m} E_{m+1}(G_i) - \kappa$.
The constant κ prevents rapid changes in the activation process. The length of the
mth activation period t_m is $\frac{\kappa}{\mathcal{E}_m(G_i)}$ for $RG_m = G_i$. A sample solution is shown in
Figure 2.10 for four receiver groups. If we let $\kappa = 0$, the value of LT_i for the optimal
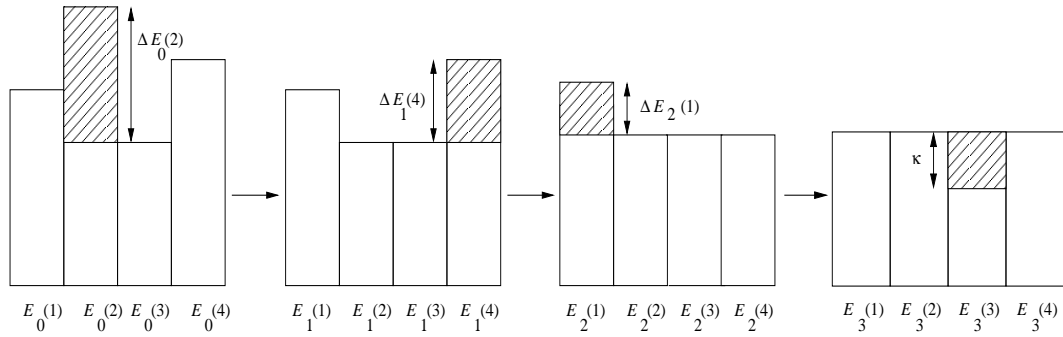
Figure 2.10: Example of time allocation among four receiver groups.

time allocation is $\arg\min_{\{p \in Z_0\}}\{E_0(G_i) - \sum_{m=0}^{p} \mathcal{E}_m(G_i)t_m 1(RG_m = G_i) \leq 0\}$ and
the maximum value of LT is found as

$$LT^* = \max_{\{(RG_m, t_m), m \in Z_0\}} \min_{1 \leq i \leq N} LT_i \, , \tag{2.38}$$

34

which has the particular solution $LT_i = LT^*$, $i = 1, ..., N$.

We assume $\mathcal{E}_m(G_i) = \mathcal{E}(G_i)$, $m \in Z_0$, for any receiver group G_i with the maximum (initial) energy $E_0(G_i)$. This is valid for systems with arbitrarily large (or renewable) energy supplies or unlimited node population. The quantity $\tau_i \cdot LT = \sum_{m=0}^{LT} t_m \, 1(RG_m = G_i)$ denotes total time allocated to the operation of receiver group G_i over the time interval $[0, LT]$. The system lifetime LT is maximized by choosing $\tau_i \cdot LT = \frac{E_0(G_i)}{\mathcal{E}(G_i)}$. The optimal value of τ_i is

$$\tau_i^* = \frac{E_0(G_i)(\mathcal{E}(G_i))^{-1}}{\sum_{i=1}^{N} E_0(G_i)(\mathcal{E}(G_i))^{-1}} \,, \quad i = 1, ..., N. \tag{2.39}$$

Transmitter group $G_{i,k}$ achieves stable throughput rate of $\lambda_i f_{i,k}^{(j)}$ over $x_{i,k}$ fraction of time such that the effective stable rate is $\frac{\lambda_i f_{i,k}^{(j)}}{x_{i,k}}$. The values of $\mathcal{E}(G_i)$, $i = 1, ..., N$, depend on the underlying MAC protocol, load fractions and temporal allocation for Group TDMA algorithm, and are given by

$$\mathcal{E}(G_i) = \sum_{k=1}^{c_i} x_{i,k} \sum_{j:R_{i,k}^{(j)} \in R_{i,k}} \epsilon \left(\lambda_i f_{i,k}^{(j)} / x_{i,k} \right), \tag{2.40}$$

where $\epsilon(y)$ is the minimum rate of energy consumption by any single-receiver MAC protocol operating under Group TDMA algorithm with the effective stable rate of y packets per slot.

2.7 Distributed Operation of Receiver Activation and Group TDMA

The standard link scheduling [4, 5, 14] assigns channels (i.e. time slots, frequencies or codes) to connecting links between nodes so that all links assigned to the same channel are conflict-free. The network topology is described by a directed

graph where directional links between nodes are only possible if nodes are within each other's transmission-reception ranges. For conflict-free packet transmission, (I) nodes cannot simultaneously transmit and receive packets, (II) nodes cannot transmit packets to multiple destinations in the same time slot, (III) primary packet conflicts, i.e. multiple number of simultaneous transmissions to the same receiver, are not allowed, (IV) secondary packet conflicts, i.e. interference effects at the non-intended receivers, are not allowed. Conflict-free scheduling can be formulated as a link coloring problem. The problems of determining the edge chromatic number of graph (i.e. the fewest number of colors necessary to color each graph edge so that no two graph edges incident on any graph vertex have the same color) [15] and optimal link scheduling [4, 5] are equivalent and NP-complete. Instead of solving the standard scheduling problem, we rely on a receiver activation heuristic to determine disjoint subsets of transmitter and receivers at each time instant (so that condition (I) is satisfied and possible violations of other conditions are reduced, but not eliminated, for all links) and on the Group TDMA method to create time orthogonality between links violating conditions (II) and (IV).

We can set up the transmitter group classification as a link coloring problem. We assume that transmitters can discover receivers up to a two-hop distance. Two receivers are called neighbors, if there is at least one transmitter in the intersection of their reception ranges. Interfering transmitter groups are assigned to distinct fractions of time, i.e. different colors are assigned to links from different transmitter groups. Transmitters with only one receiver in their transmission ranges acquire membership in group A_1 and all links from A_1 are given color C_1. Next, an arbitrary

36

receiver R_1 is selected such that any transmitter that has multiple receivers in its transmission range including R_1 as the intended destination initiates a transmitter group A_2, i.e. links from A_2 to R_1 are given a new color C_2. Next, we consider all neighbors of R_1. If R_2 is a neighboring receiver of R_1, we assign different colors to all links from transmitters in the intersection of R_1 and R_2 to the particular receiver R_2. We continue with coloring links to receivers one by one. Transmitters that originate links with new colors initiate new transmitter groups.

For receiver activation group G_i, we denote by a_i the maximum number of intersections of reception ranges (i.e. the maximum number of neighbors) for each receiver and we denote by e_i the modified edge chromatic number, which is the minimum number of colors necessary to color graph edges so that no two graph edges violating condition (IV) have the same color, i.e. links to neighbor receivers are assigned different colors. Note that $a_i + 1 \leq e_i$ which follows from plain geometry. Since the separation between receiver nodes is greater than the reception radius r (as a consequence of receiver activation), there exists a fixed upper-bound on the number of intersections of reception ranges for each receiver. At most 13 different colors are needed for planar networks (with $a_i = 12$) and at most 3 different colors are needed for tandem networks (with $a_i = 2$). As a result, transmitters with any intended destination R choose one of the finite number of available group memberships different than those previously acquired by other transmitter groups with intended destinations that are neighbors to R. Packets are addressed randomly to any of the receivers in the transmission range so that condition II is also satisfied.

If the receiver activation has already partitioned nodes to subsets of trans-

mitters and receivers (so that condition (I) is satisfied for all links), the remaining problem of creating time orthogonality among transmitter groups (so that conditions (II) and (IV) are satisfied for all possible links) can be solved in polynomial time by the distributed Group TDMA method.

2.8 Performance Evaluation

We consider static tandem and planar networks of 1000 unbuffered nodes. We consider systems with first unlimited energy supply and then with hard finite energy-constraints. For the latter case, we assume that each node has an amount of initial battery energy $E_{\max} = 10^6$ (unit energy). Each packet transmission consumes πr^2 units of battery energy. We assume that nodes generate packet transmissions with the same rate according to a common Poisson process and employ the FCFS collision resolution algorithm to resolve the primary packet conflicts.

The value of the common transmission (reception) radius characterizes the distribution of the activated transmitter-receiver pairs on the network and specifies the overlapping between the reception regions, on which the operation of receiver activation and Group TDMA strongly depends. We introduce the quantities $\frac{2r}{L}$ and $\frac{\pi r^2}{A}$, which denote the ratios of the transmission range to the network length and to the network area in tandem and planar networks, respectively. We first apply the topology-based receiver activation heuristic (without energy-efficient solutions) to unlimited energy systems and compare the Group TDMA algorithm with the simultaneous operation of the activated receivers. For both cases, equal fractions of

time are allocated to each receiver group. The maximum (single-hop) throughput per destination is shown in Figure 2.11 as functions of $\frac{2r}{L}$ and $\frac{\pi r^2}{A}$. Group TDMA achieves higher throughput than the simultaneous operation of transmitter-receiver pairs. The achievable stable throughput is larger in tandem networks, since the reception ranges overlap less. In general, the network approaches a single one-destination system for large values of r, whereas the number of one-destination systems increases with smaller values of r.
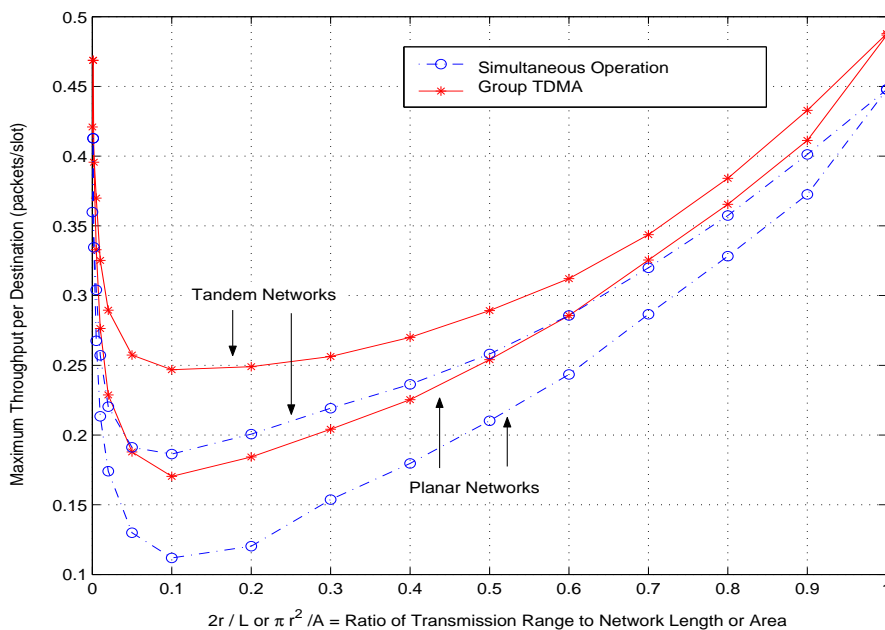
Figure 2.11: The maximum stable throughput for Group TDMA and simultaneous operation of transmitter-receiver pairs.

We also impose hard energy constraints such that the initial battery energy of each node is 10^6 energy units. Each packet transmission consumes πr^2 energy units. We compare the energy-efficient receiver activation to the topology-based receiver activation that allocates equal time fractions to each receiver group. We let the
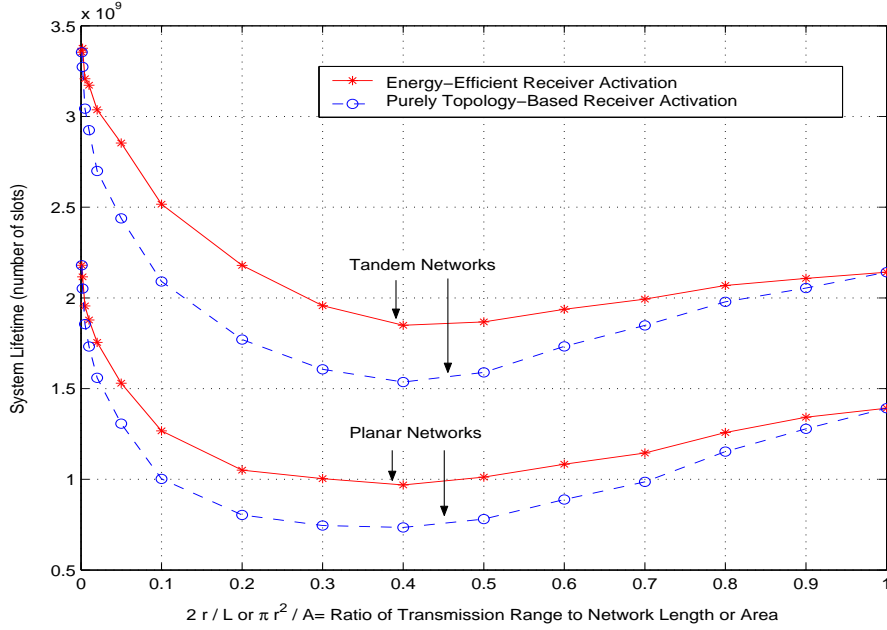
Figure 2.12: The system lifetime for energy-efficient and topology-based receiver activation.

Group TDMA algorithm operate with the maximum stable throughput under both receiver activation methods. The system lifetime is defined as the duration of time interval from the start of the network operation until the first time when the energy supplies for transmissions to any activated receiver group are depleted. Figure 2.12 depicts the system lifetimes for both receiver activation heuristics in tandem and planar networks. Simulation results verify that the energy-efficient receiver activation outperforms the solutions based on equal time allocations in terms of energy properties. The gap between the two heuristics increases for intermediate values of r, where there are several potentially interfering multi-destination systems. The performance of both methods becomes identical as r increases, so in the end we have a single activated receiver in each receiver activation group.

2.9   Summary and Conclusions

We rediscovered the value of scheduled access in wireless networks from the perspectives of stable throughput and energy efficiency. We proposed a two-step time-division mechanism based on receiver activation and Group TDMA as distributed link scheduling with suboptimal but polynomial-time solutions. We developed a topology-based greedy heuristic to determine distinct receiver groups to be activated within disjoint fractions of time, and determined temporal allocations based on cumulative battery energies left at transmitter groups to extend the node lifetimes. We used the Group TDMA method to formulate a linear programming solution to the problem of throughput-optimal temporal allocation for transmissions to each activated receiver group, and derived bounds on the maximum stable throughput for tandem and planar networks. We also evaluated the performance improvement by energy-efficient receiver activation and throughput-efficient Group TDMA. We discussed distributed implementation based on the full cooperation of nodes. The associated synchronization and overhead issues should be further considered. The extension to the non-cooperative operation at the MAC layer will be discussed in Chapters 3-5. The analysis will be extended to the joint design of MAC and network layers in Chapters 5-7.

Chapter 3

A Game-Theoretic Look at MAC in Wireless Networks:

Non-Cooperative Random Access for Selfish and Malicious Users

In wireless access, transmitter nodes need to make individual decisions for distributed operation and do not necessarily cooperate with each other. The non-cooperative transmitter behavior at the MAC layer may be purely selfish or may also reflect malicious objectives of generating interference to prevent the successful transmissions of the other nodes. In this chapter, we address the single-receiver random access problem for non-cooperative transmitters with individual performance objectives of optimizing throughput rewards, transmission energy costs and delay costs. We evaluate the interactions between selfish and malicious transmitters by formulating a non-cooperative game of selecting the individual probabilities of transmitting packets to a common receiver. Specifically, we derive the non-cooperative transmission strategies in Nash equilibrium. The analysis provides insights for the optimal strategies to block random access of selfish nodes as well as the optimal defense mechanisms against the possible denial of service attacks of malicious nodes in wireless networks. The results are also compared with the cooperative equilibrium strategies that optimize the total system utility (separately under random access and scheduled access). In this context, a pricing scheme is presented to improve the non-cooperative operation. For distributed implementation, we formulate an adap-

tive mechanism of the best-response strategy updates. We also introduce adaptive heuristics based on the channel feedback only provided that the system parameters are not explicitly known at the individual transmitters. Alternatively, we consider a repeated game model and let nodes randomly choose to transmit or to wait depending on the number of backlogged nodes in the random access system. The capture effects in the communication channel are further exploited to develop a joint power and rate control game in conjunction with random access.

3.1   Introduction

Game theory is a powerful tool to analyze the interactions among decision-makers with conflicting interests and finds a rich extent of applications in communication systems including network routing, load balancing, resource allocation, flow and power control. There is an emerging interest in game-theoretic studies for random access of selfish users contending for collision channels [16, 17, 3, 18, 19].

Random access has been extensively studied as a cooperative throughput optimization problem [11, 12, 13, 20]. The analysis of the achievable throughput and stability properties has been extended from the classical collision channel model without packet captures to the multi-packet reception channels with a more realistic model for packet captures and interference effects [21, 22].

As the number of transmitters increases, it becomes more difficult to coordinate reliable transmissions between transmitter-receiver pairs. Instead, we can consider the non-cooperative operation, in which selfish nodes select the transmis-

sion probabilities to optimize their individual performance objectives. A system in which nodes behave selfishly is inherently distributed and therefore more scalable compared to the centralized form of MAC. Selfish behavior of nodes can also increase the fairness of the system, since no node can achieve performance gains by breaking the rules of the underlying channel access scheme.

For finite number of transmitters without queue buffers, the problem of non-cooperative random access has been studied in [3, 17, 19, 23] for collision channels under the assumption that each packet arrives at a new node, and whenever a packet is successfully transmitted, it leaves the system. The effects of packet captures in non-cooperative random access have been evaluated in [24, 25]. The stability properties have been discussed in [26] and [18] for the ALOHA system model with and without queue buffers.

A distributed game has been proposed in [16] to achieve any given requirements of fixed throughput rates. For arbitrary throughput objectives, non-cooperative random access has been formulated in [27] jointly with the network layer decisions in a network utility optimization framework. Additional costs for unsuccessful packet transmissions have been considered in [28] for non-cooperative random access systems, and the optimal pricing strategies have been derived in [29] for selfish transmitters to improve the non-cooperative equilibrium.

In this chapter, our first objective is to extend the throughput and stability studies of random access with finite number of transmitter nodes to the non-cooperative operation. We strip off all complexities introduced by multi-hop operation and analyze the fundamental interactions of selfish nodes randomly transmitting

packets to a single receiver at the MAC layer. As the starting point, we consider the case of saturated queues at selfish transmitter nodes with the objectives of maximizing their individual utilities that reflect throughput rewards, transmission energy and delay costs. We evaluate the non-cooperative transmission probabilities in non-cooperative Nash equilibrium such that no node in equilibrium can unilaterally improve the individual performance.

The results are also compared with the cooperative equilibrium strategies that optimize the total utility. In this context, we introduce a pricing scheme to stimulate the cooperation of transmitters and improve the non-cooperative operation to the cooperative equilibrium. The random access results are also compared with the scheduled access solutions and the resulting performance loss is evaluated.

Next, we allow packet queues to empty and determine the non-cooperative equilibrium strategies to minimize the energy costs while keeping the packet queues stable for fixed packet arrival rates. Finally, we extend the model to multiple receiver nodes and evaluate the transmission strategies for broadcast communication with each packet addressed to multiple receiver nodes.

Non-cooperative nodes may also pursue malicious objectives such as blocking the packet transmissions of the other selfish nodes. In this context, the channel jamming effects of malicious transmitters have been evaluated in [30] in terms of the best worst case (minimax) performance for multi-hop operation in ALOHA multiple access systems. From a network security point of view, a non-cooperative game framework has been developed in [31] for intrusion detection systems without MAC interactions. The possible misbehavior of transmitters has been formulated

45

in [32] for IEEE 802.11 MAC protocol and different methods have been introduced in [33] to detect and prevent the operation of colluding selfish transmitters.

The second objective of this chapter is to develop a framework for the denial of service attacks as a stochastic game among non-cooperative selfish nodes (that randomly transmit packets to a common receiver) and malicious nodes (that attack the other selfish nodes by jamming their packet transmissions). The results lead to the optimal denial of service attack and defense strategies for random access systems. In particular, we evaluate the possible performance loss, if selfish nodes defect by becoming malicious in their transmission strategies.

For distributed implementation, we present a repeated play of the random access game and develop a best-response update mechanism, in which nodes independently adjust their transmission strategies in response to the channel outcome. For the case of imperfect knowledge of system parameters, we also introduce a stochastic update mechanism based on the channel feedback only and point at the deviation of the results from the non-cooperative Nash equilibrium.

The third objective of this chapter is to formulate a repeated game model of selecting transmission probabilities depending on the contention for the channel, namely the number of potential transmitters. For that purpose, we consider a fixed number of backlogged packets waiting for retransmissions without additional packet arrivals until all packet conflicts are resolved. In addition, we propose to include additional power and rate control strategies in random access games as a more efficient way to exploit the capture phenomena in the communication channel. Following a single packet capture model, we evaluate by numerical results the pos-

sible performance improvement of joint power and rate control strategies over the limited actions of choosing retransmission probabilities.

The chapter is organized as follows. In section 3.2, we introduce the random access system model and describe the non-cooperative and cooperative games together with the underlying reward and cost structures. In section 3.3, we consider saturated queues and derive the non-cooperative equilibrium strategies for the case of two selfish nodes and compare the results with the full cooperation as well as with the cooperation stimulation based on pricing. We extend the analysis in section 3.4 to stable operation of two selfish transmitter nodes with possibly emptying packet queues. For the case of saturated queues, we introduce malicious node behavior in section 3.5 and evaluate non-cooperative equilibrium with one selfish and one malicious transmitter node with or without throughput and delay objectives. In section 3.6, we describe distributed adaptive mechanisms for different cases with and without explicit information on system parameters. For the case of saturated queues, we also extend the random access problem to arbitrary number of selfish and malicious transmitters and compare the random access strategies with (cooperative) scheduled access solutions in section 3.7. We present extensions to broadcast communication with multiple receivers in section 3.8. Alternatively, we formulate in section 3.9 a repeated random access game depending on the number of backlogged transmitters. As an extension, we introduce power and rate control games in conjunction with random access in section 3.10. We draw conclusions and discuss future work in section 3.11.

3.2 System Model, Rewards and Costs

We assume that multiple nodes randomly transmit packets with fixed proba-bilities to a common receiver, as shown in Figure 3.1. Each transmitter has a packet queue of infinite buffer capacity. Let p_i denote the transmission probability of node i from the set N of n transmitters. We assume multi-packet reception channels with possible packet captures. A packet of node i is successfully received with probability $q_{i|J}$, if nodes in set J (including node i) transmit in the same time slot. Throughout this chapter, we will often specialize the results to the particular case of classical collision channels with $q_{i|J} = 1$ for $J = \{i\}$ and $q_{i|J} = 0$ for $J \neq \{i\}$.
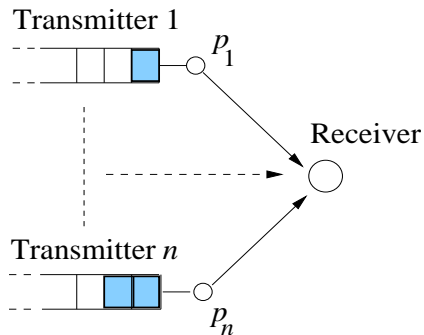
Figure 3.1: Random access system with multiple transmitters and single receiver.

Any selfish node i has the objective of choosing the transmission probability p_i to maximize the individual utility function u_i that reflects the difference between the throughput rewards and costs of of transmission energy and delay per time slot. We define λ_i as the throughput rate of node i, r_i as the reward for any successful packet transmission of node i, $\mathcal{E}_i$ as the average transmission energy cost of node i per time slot, and $\mathcal{D}_i$ as the average delay-type cost of node i per time slot. Let $M_i \subseteq N - \{i\}$ denote the set of nodes attacked by malicious node i that has the

objective of blocking the packet transmissions of any node $j \in M_i$. Node i receives the reward $c_{i,j}^{(1)}$, if the transmission of node $j \in M_i$ fails. On the other hand, node i incurs cost $c_{i,j}^{(2)}$, if the transmission of node $j \in M_i$ is successful. We define $\mathcal{C}_i$ as the difference between the reward of node i for blocking random access of the other selfish nodes in M_i and the cost of node i for missing the opportunity to prevent successful packet transmissions of the other selfish nodes in M_i in any time slot. The utility of any (selfish or malicious node i) as function of transmission probabilities $\underline{p}$ is defined as

$$u_i(\underline{p}) = r_i \lambda_i - \mathcal{E}_i - \mathcal{D}_i + \mathcal{C}_i . \tag{3.1}$$

The non-cooperative optimization problem for any node i is defined as

$$\max_{p_i} \ u_i(\underline{p}) . \tag{3.2}$$

For saturated queues, λ_i denotes the average number of successful packet transmission per time slot. For stable operation with random packet arrivals, λ_i denotes the packet arrival rate at node i such that the packet queue at node i is stable, i.e. it is asymptotically finite with finite packet delay. For node i, we let E_i denote the energy cost of any packet transmission and D_i denote the delay-type cost for each slot of waiting or unsuccessful transmission under the assumption of saturated queues. The throughput reward r_i normalizes the values of E_i and D_i in the utility function u_i of node i with respect to the throughput rate λ_i.

We can also consider the cooperative transmission strategies that maximize the weighted sum of utilities over all transmission probabilities. For the non-negative

weight constants β_i, $i \in N$, the cooperative optimization problem is defined as

$$\max_{\underline{p}} \sum_{i=1}^{n} \beta_i \, u_i(\underline{p}) \, . \tag{3.3}$$

3.3 Two Selfish Transmitters with Saturated Packet Queues

Consider $N = \{1, 2\}$ with two selfish transmitters (i.e. $M_i = \{\emptyset\}$, $i = 1, 2$).

3.3.1 Non-Cooperative Equilibrium

Node i transmits a packet with probability p_i and it is successfully received with probability $q_{i|i}$, if the other node decides not to transmit, or captured with probability $q_{i|\{1,2\}}$, if the other node also transmits in the same time slot. We assume $0 \le q_{i|\{1,2\}} \le q_{i|i} \le 1$ for $i = 1, 2$. The region of the achievable throughput rates is given by

$$\lambda_1 \le p_1 \left(p_2 q_{1|\{1,2\}} + (1 - p_2) q_{1|1} \right), \tag{3.4}$$

$$\lambda_2 \le p_2 \left(p_1 q_{2|\{1,2\}} + (1 - p_1) q_{2|2} \right). \tag{3.5}$$

We consider two transmitter nodes 1 and 2 with the selfish objectives of maximizing the individual utilities u_1 and u_2, respectively. Nodes 1 and 2 receive the respective throughput rewards $r_1 \lambda_1$ and $r_2 \lambda_2$ per time slot on the average. The throughput rates λ_1 and λ_2 need to satisfy the conditions (3.4)-(3.5) with equality to maximize the individual utility $u_i(p_1, p_2) = r_i \lambda_i - \mathcal{E}_i - \mathcal{D}_i$. In any time slot, any selfish node incurs a delay cost, if it decides to wait or its packet transmission is unsuccessful. The delay costs are $\mathcal{D}_1 = D_1(1 - (1 - p_2)q_{1|1} - p_2 q_{1|1,2})$ and

$\mathcal{D}_2 = D_2(1 - (1-p_1)q_{2|2} - p_1 q_{2|1,2})$ for unsuccessfully transmitting nodes 1 and 2, respectively, or $\mathcal{D}_1 = D_1$ and $\mathcal{D}_2 = D_2$ for waiting nodes 1 and 2, respectively.

The average energy cost $\mathcal{E}_i$ per time slot is $p_i E_i$ for any node $i = 1, 2$, since node i transmits one of the available packets with probability p_i in any time slot. As a result, the expected utilities of nodes 1 and 2 are given by

$$u_1(p_1, p_2) = p_1 \left((r_1 + D_1) \left((1-p_2)q_{1|1} + p_2 q_{1|1,2} \right) - D_1 - E_1 \right) + (1-p_1)(-D_1), \quad (3.6)$$

$$u_2(p_1, p_2) = p_2 \left((r_2 + D_2) \left((1-p_1)q_{2|2} + p_1 q_{2|1,2} \right) - D_2 - E_2 \right) + (1-p_2)(-D_2), \quad (3.7)$$

respectively. We define $A_i = \{W, T\}$ as the set of actions of node i, where W and T are actions of waiting and transmitting. Let p_{-i} denote the strategy of the node other than node i. The non-cooperative Nash equilibrium strategies p_1^* and p_2^* satisfy

$$u_i(p_i^*, p_{-i}^*) \geq u_i(p_i, p_{-i}^*) \quad , \quad i = 1, 2, \quad (3.8)$$

for any node i and strategy p_i such that no node i can unilaterally improve the individual utility u_i beyond the non-cooperative Nash equilibrium.

Given p_{-i}^*, the expected utility u_i of every action $a_i \in A_i$, which is in support of p_i^*, should be the same for any non-cooperative node i in Nash equilibrium. The resulting mixed strategies $0 \leq p_i^* \leq 1$, satisfy $u_i(1, p_{-i}^*) = u_i(0, p_{-i}^*)$ for any i. Otherwise, node i can shift the probability distribution to the action with higher conditional expected utility and end up with a pure strategy of $p_i^* = 0$ or $p_i^* = 1$.

Theorem 3.3.1 *There exist possibly multiple non-cooperative strategies in Nash equilibrium.*

(a) The mixed equilibrium strategies are given by

$$p_1^* = \left(q_{2|2} - q_{2|1,2}\right)^{-1} \left(q_{2|2} - \frac{E_2}{r_2 + D_2}\right), \tag{3.9}$$

$$p_2^* = \left(q_{1|1} - q_{1|1,2}\right)^{-1} \left(q_{1|1} - \frac{E_1}{r_1 + D_1}\right), \tag{3.10}$$

if $q_{i|1,2} \leq \frac{E_i}{r_i + D_i} \leq q_{i|i}$, $i = 1, 2$.

(b) The pure equilibrium strategies are given by $p_i^* = 1$, $i = 1, 2$, if $\frac{E_i}{r_i + D_i} \leq q_{i|1,2}$, or

$p_i^* = 0$, $i = 1, 2$, if $\frac{E_i}{r_i + D_i} \geq q_{i|i}$, or $p_1^* = 1$, $p_2^* = 0$, if $\frac{E_1}{r_1 + D_1} \leq q_{1|1}$ and $\frac{E_2}{r_2 + D_2} \geq q_{2|1,2}$,

or $p_1^* = 0$, $p_2^* = 1$, if $\frac{E_2}{r_2 + D_2} \leq q_{2|2}$ and $\frac{E_1}{r_1 + D_1} \geq q_{1|1,2}$.

Proof: The conditional expected utilities of node 1 are given by

$$u_1(1, p_2) = (r_1 + D_1)((1 - p_2)q_{1|1} + p_2 q_{1|1,2}) - D_1 - E_1, \tag{3.11}$$

$$u_1(0, p_2) = -D_1, \tag{3.12}$$

and the conditional expected utilities of node 2 are given by

$$u_2(p_1, 1) = (r_2 + D_2)((1 - p_1)q_{2|2} + p_1 q_{2|2,2}) - D_2 - E_2, \tag{3.13}$$

$$u_2(p_1, 0) = -D_2. \tag{3.14}$$

Given the strategy p_{-i}^* of the node other than node i, the expected utility u_i of every action $a_i \in A_i$, which is in support of p_i^*, should be the same for each node i. The resulting indifference condition of $u_i(0, p_{-i}^*) = u_i(1, p_{-i}^*)$, $i = 1, 2$, for non-cooperative Nash equilibrium leads to the mixed equilibrium strategies given by Eqs. (3.9)-(3.10) provided that $q_{i|1,2} \leq \frac{E_i}{r_i + D_i} \leq q_{i|i}$, $i = 1, 2$. Otherwise, node i can shift the probability distribution to the action with higher conditional expected utility and end up with a pure strategy given by $p_i = 0$ or $p_i = 1$.

If the indifference condition cannot be satisfied for any feasible set of the transmission probabilities such that $0 \leq p_i \leq 1$, $i = 1, 2$, each node i performs with probability 1 the action of higher conditional expected utility. The resulting pure equilibrium strategy of node $i = 1, 2$ is $p_i^* = 1$, if $u_i(1, p_{-i}^*) > u_i(0, p_{-i}^*)$, or $p_i^* = 0$, if $u_i(1, p_{-i}^*) < u_i(0, p_{-i}^*)$. □

The Nash equilibrium strategies may not be unique depending on systems parameters. For instance, if $q_{i|i} = 1$, $q_{i|1,2} = 0$, $r_i = 1$, $D_i = 0$ and $0 < E_i < 1$, $i = 1, 2$, the Nash equilibrium strategies (p_1^*, p_2^*) are $(1 - E_2, 1 - E_1)$, $(1, 0)$ or $(0, 1)$.

For the classical collision channels, the non-cooperative equilibrium strategies of Eqs. (3.9)-(3.10) are given by $p_1^* = 1 - \frac{E_2}{r_2 + D_2}$ and $p_2^* = 1 - \frac{E_1}{r_1 + D_1}$, and the through-put rate region is described as $\lambda_1 \leq \frac{E_1}{r_1 + D_1}\left(1 - \frac{E_2}{r_2 + D_2}\right)$ and $\lambda_2 \leq \frac{E_2}{r_2 + D_2}\left(1 - \frac{E_1}{r_1 + D_1}\right)$ for $0 \leq \frac{E_i}{r_i + D_i} \leq 1$, $i = 1, 2$. The envelope of the throughput rate pairs λ_1 and λ_2 over $0 \leq E_i \leq r_i + D_i$, $i = 1, 2$, coincides with the boundary of the achievable throughput region $\sqrt{\lambda_1} + \sqrt{\lambda_2} = 1$ [12], if $\sum_{i=1}^{2} \frac{E_i}{r_i + D_i} = 1$. If $E_1 = 0$ and $E_2 = 0$, nodes choose aggressive transmission strategies of $p_i^* = 1$, $i = 1, 2$, such that $\lambda_i = 0$, $i = 1, 2$. The non-zero energy costs prevent zero throughput rates, whereas the delay costs punish the idle slots and motivate nodes to transmit with higher probability, i.e. the energy and delay costs mutually balance the non-cooperative equilibrium strategies.

3.3.2 Comparison with Cooperative Equilibrium

Transmitter nodes in full cooperation have the common objectives of maximizing the weighted utility sum u_Σ that is defined to be $\sum_{i=1}^{2} \beta_i u_i$ for non-negative

weight constants β_i, $i = 1, 2$.

Theorem 3.3.2 *The cooperative equilibrium strategies that maximize u_Σ are given*

by

$$p_1^* = \frac{\beta_2(q_{2|2}(r_2 + D_2) - E_2)}{\sum_{i=1}^{2} \beta_i(q_{i|i} - q_{i|1,2})(r_i + D_i)}, \tag{3.15}$$

$$p_2^* = \frac{\beta_1(q_{1|1}(r_1 + D_1) - E_1)}{\sum_{i=1}^{2} \beta_i(q_{i|i} - q_{i|1,2})(r_i + D_i)}. \tag{3.16}$$

Proof: The weighted utility sum u_Σ of Eqs. (3.6) and (3.7) is concave in p_i, $i = 1, 2$.

The optimal solutions of Eqs. (3.15)-(3.16) follow from $\frac{\partial u_\Sigma}{\partial p_i} = 0$ for $i = 1, 2$. $\qquad \square$

The total utility is optimized by the strategies from Eqs. (3.15) and (3.16) to

$$u_\Sigma^c = \frac{\prod_{i=1}^{2} \beta_i(q_{i|i}(r_i + D_i) - E_i)}{\sum_{i=1}^{2} \beta_i(q_{i|i} - q_{i|1,2})(r_i + D_i)} - \sum_{i=1}^{2} \beta_i D_i, \tag{3.17}$$

whereas the non-cooperative equilibrium utility u_i of node i is $-D_i$ such that the

total utility u_Σ^{nc} is $- \sum_{i=1}^{n} \beta_i D_i$, if $q_{i|\{1,2\}} \leq \frac{E_i}{r_i + D_i} \leq q_{i|i}$, $i = 1, 2$. Whether the

cooperative equilibrium strategies that maximize the total utility can also maximize

the throughput rates depends on the packet capture probabilities as well as on the

energy and delay costs. As an example, we consider the classical collision channels

with $q_{i|i} = 1$ and $q_{i|1,2} = 0$, $i = 1, 2$, and assume symmetric rewards and costs,

i.e. $r_i = r$, $D_i = D$ and $E_i = E$, $i = 1, 2$. The equilibrium transmission proba-

bility of each selfish transmitter is $1 - \frac{E}{r+D}$ and throughput of each transmitter is

$\left(1 - \frac{E}{r+D}\right) \frac{E}{r+D}$ packets per time slot. The transmission probability in cooperative

equilibrium with $\beta_i = 1$, $i = 1, 2$, is $\frac{1}{2}\left(1 - \frac{E}{r+D}\right)$ and the resulting throughput rate

of each transmitter is $\frac{1}{4}\left(1 - \frac{E}{r+D}\right)\left(1 + \frac{E}{r+D}\right)$ packets per time slot. As a result, the

non-cooperative equilibrium can achieve higher throughput rates compared to the

utility-optimal cooperative operation, if $3E > r + D$, i.e. higher energy costs can make the non-cooperative operation more throughput-efficient.

3.3.3 Improvement of Non-cooperative Equilibrium through Pricing

Non-cooperative nodes are more aggressive in their transmission strategies than cooperative nodes and cannot necessarily achieve the optimal utilities in cooperative equilibrium. In this context, we can introduce a linear pricing mechanism to improve the performance of the non-cooperative operation. We assume that the receiver charges a price e_i in any time slot for any successful packet transmission by node i. This is reflected in the utility function of any node $i = 1, 2$ by replacing the throughput reward r_i with a discounted amount $r_i - e_i$ such that the utility u_i of any node i is reduced to $u_i - \lambda_i e_i$. The non-cooperative equilibrium strategies are

$$p_1^* = \frac{q_{2|2} - \frac{E_2}{r_2 - e_2 + D_2}}{q_{2|2} - q_{2|1,2}}, \tag{3.18}$$

$$p_2^* = \frac{q_{1|1} - \frac{E_1}{r_1 - e_1 + D_1}}{q_{1|1} - q_{1|1,2}}. \tag{3.19}$$

Next, we answer the question of how the receiver should choose the prices e_1 and e_2 for the efficient network operation. First, note that there exist linear prices for which the non-cooperative equilibrium strategies can be made equal to the cooperative equilibrium strategies that can be achieved in the absence of pricing. These utility-optimal prices are given by

$$e_1^* = \frac{\beta_2(q_{2|2} - q_{2|1,2})(r_2 + D_2)(q_{1|1}(r_1 + D_1) - E_1)}{\beta_1 E_1(q_{1|1} - q_{1|1,2}) + \beta_2 q_{1|1}(q_{2|2} - q_{2|1,2})(r_2 + D_2)}, \tag{3.20}$$

$$e_2^* = \frac{\beta_1(q_{1|1} - q_{1|1,2})(r_1 + D_1)(q_{2|2}(r_2 + D_2) - E_2)}{\beta_2 E_2(q_{2|2} - q_{2|1,2}) + \beta_1 q_{2|2}(q_{1|1} - q_{1|1,2})(r_1 + D_1)}. \tag{3.21}$$

From the receiver point of view, the desirable measures to be optimized are the total achievable throughput rate and the total revenue (namely the total amount of prices collected at the receiver). First, we consider the problem of optimizing the weighted sum of throughput rates $\sum_{i=1}^{2} \alpha_i \lambda_i$ for the non-negative weight constants α_i, $i = 1, 2$. The throughput-optimal transmission probabilities in full cooperation are given by

$$p_1^* = \frac{\alpha_2 q_{2|2}}{\alpha_1(q_{1|1} - q_{1|1,2}) + \alpha_2(q_{2|2} - q_{2|1,2})}, \tag{3.22}$$

$$p_2^* = \frac{\alpha_1 q_{1|1}}{\alpha_1(q_{1|1} - q_{1|1,2}) + \alpha_2(q_{2|2} - q_{2|1,2})}. \tag{3.23}$$

There exist linear prices such that the total throughput rate $\sum_{i=1}^{n} \alpha_i \lambda_i$ can be maximized by the non-cooperative equilibrium strategies given by Eqs. (3.18)-(3.19) with pricing to the throughput-optimal cooperative equilibrium value. The corresponding prices are chosen as

$$e_1^* = r_1 + D_1 - \frac{E_1 \left(\alpha_1(q_{1|1} - q_{1|1,2}) + \alpha_2(q_{2|2} - q_{2|1,2}) \right)}{\alpha_2(q_{2|2} - q_{2|1,2})q_{1|1}}, \tag{3.24}$$

$$e_2^* = r_2 + D_2 - \frac{E_2 \left(\alpha_1(q_{1|1} - q_{1|1,2}) + \alpha_2(q_{2|2} - q_{2|1,2}) \right)}{\alpha_1(q_{1|1} - q_{1|1,2})q_{2|2}}. \tag{3.25}$$

If we consider the classical collision channels and assume symmetric rewards and costs such as $r_i = r$, $D_i = D$, $E_i = E$ and $\alpha_i = \alpha$, $i = 1, 2$, the throughput-optimal symmetric prices are given by $r + D - 2E$ for both transmitters. The resulting pricing mechanism can maximize the individual throughput rate to $\frac{1}{2}$ packets per time slot. On the other hand, the non-cooperative equilibrium strategies with prices e_1^* and e_2^* that satisfy

$$E_i \sum_{k=1}^{2} e_k^*(q_{k|k} - q_{k|1,2}) = q_{i|i}(r_i - e_i^* + D_i)e_j^*(q_{j|j} - q_{j|1,2}) \tag{3.26}$$

for $i = 1, 2$, $j = 1, 2$ and $i \neq j$ can optimize the total revenue $\sum_{i=1}^{2} e_i \lambda_i$ achievable by the receiver per time slot.

3.4 Extensions to Stable Operation with Two Selfish Transmitters

Assume that packets arrive at transmitters 1 and 2 with fixed rates λ_1 and λ_2, respectively, according to independent and stationary random processes. Consider a dominant system S_i, $i = 1, 2$, in which node i transmits dummy packets with probability p_i, if the queue at node i is empty [12]. For S_1, provided that node 2 transmits with probability p_2, the transmission is successful with probability $q_{2|2}$, if node 1 decides not to transmit with probability $1 - p_1$, or with probability $q_{2|\{1,2\}}$, if node 1 decides to transmit with probability p_1. As a result, we obtain the condition $\lambda_2 \leq p_2(q_{2|2} - p_1(q_{2|2} - q_{2|1,2}))$ for the throughput rate of node 2. On the other hand, node 1 sees a probability of success rate that is $p_1 q_{1|1}$, when node 2 does not have packets (which occurs with probability $1 - \frac{\lambda_2}{p_2(q_{2|2} - p_1(q_{2|2} - q_{2|1,2}))}$ according to Little's result [9]), and that is $p_1((1 - p_2)q_{1|1} + p_2 q_{1|\{1,2\}})$, when node 2 has packets (which occurs with probability $\frac{\lambda_2}{p_2(q_{2|2} - p_1(q_{2|2} - q_{2|1,2}))}$). Hence, we obtain the condition $\lambda_1 < p_1 q_{1|1}(1 - \frac{\lambda_2}{p_2(q_{2|2} - p_1(q_{2|2} - q_{2|1,2}))}) + p_1(q_{1|1} - p_2(q_{1|1} - q_{1|1,2}))\frac{\lambda_2}{p_2(q_{2|2} - p_1(q_{2|2} - q_{2|1,2}))}$ for the throughput rate of node 1. As given in [21], the stable throughput rates for dominating system S_1 satisfy

$$\lambda_1 < p_1 q_{1|1} - \frac{p_1(q_{1|1} - q_{1|1,2})\lambda_2}{q_{2|2} - p_1(q_{2|2} - q_{2|1,2})} \quad \text{for} \quad \lambda_2 \leq p_2 q_{2|2} - p_1 p_2(q_{2|2} - q_{2|1,2}). \quad (3.27)$$

Similarly, the stable throughput rates for dominating system S_2 satisfy

$$\lambda_2 < p_2 q_{2|2} - \frac{p_2(q_{2|2} - q_{2|1,2})\lambda_1}{q_{1|1} - p_2(q_{1|1} - q_{1|1,2})} \quad \text{for} \quad \lambda_1 \leq p_1 q_{1|1} - p_1 p_2(q_{1|1} - q_{1|1,2}). \quad (3.28)$$

Since the queue lengths in systems S_1 and S_2 are at least as long as the queue lengths in original system, the union of rates given by (3.27) and (3.28) defines the stability region $\mathcal{S}(p_1, p_2)$.

Assume that the packet arrival rates λ_1 and λ_2 are fixed and known. We assume the classical collision channels for simplicity, although the results can be easily extended to the general multi-packet reception channels. We consider the stability of packet queues as a constraint for both transmitters. We do not consider finite packet delay costs in stable operation. A detailed discussion of the delay properties in random access systems has been presented in [21]. The non-cooperative optimization problem of any transmitter node $i = 1, 2$ is given by

$$\max_{p_i} (r_i \lambda_i - p_i E_i) \quad \text{s.t. } (\lambda_1, \lambda_2) \in \mathcal{S}(p_1, p_2). \tag{3.29}$$

Theorem 3.4.1 *For fixed packet arrival rates λ_1 and λ_2 and arbitrarily small positive constants ϵ_1 and ϵ_2, the non-cooperative transmission probabilities p_1^* and p_2^* in Nash equilibrium satisfy*

$$p_1^*(q_{1|1} - p_2^*(q_{1|1} - q_{1|1,2})) = \lambda_1 - \epsilon_1, \tag{3.30}$$

$$p_2^*(q_{2|2} - p_1^*(q_{2|2} - q_{2|1,2})) = \lambda_2 - \epsilon_2, \tag{3.31}$$

where $p_1 q_{1|1}(q_{2|2} - q_{2|1,2}) + p_2 q_{2|2}(q_{1|1} - q_{1|1,2}) < q_{1|1} q_{2|2}$.

Proof: According to Loynes theorem [34], a queue is stable, if the arrival process and service process of the queue are all stationary, and the average arrival rate is less than the average service rate. We do not analyze the stability at the boundary points of equality between the average arrival rates and service rates. Therefore,

we need to convert the problem to a standard constrained optimization problem by replacing the stability conditions (3.27) and (3.28) by

$$\lambda_1 \leq p_1 q_{1|1} - \frac{p_1(q_{1|1} - q_{1|1,2})\lambda_2}{q_{2|2} - p_1(q_{2|2} - q_{2|1,2})} - \epsilon_1, \tag{3.32}$$

$$\lambda_2 \leq p_2 q_{2|2} - \frac{p_2(q_{2|2} - q_{2|1,2})\lambda_1}{q_{1|1} - p_2(q_{1|1} - q_{1|1,2})} - \epsilon_2, \tag{3.33}$$

where ϵ_1 and ϵ_2 are arbitrarily small positive constants. Since r_i, λ_i and E_i are all non-negative and fixed, the optimization problem (3.29) is equivalent to

$$\min_{p_i} \ p_i \quad \text{s.t. } (\lambda_1, \lambda_2) \in \mathcal{S}(p_1, p_2). \tag{3.34}$$

We use the Lagrange multiplier method to solve (3.34). Let $c_j = \eta_j$ and $g_j(p_1, p_2) = \lambda_j - p_j \left(q_{j|j} - \frac{(q_{j|j} - q_{j|1,2})\lambda_k}{q_{k|k} - p_j(q_{k|k} - q_{k|1,2})} \right)$, $j = 1, 2$, and $k = 1, 2$, $k \neq j$. Then, the Lagrange Multiplier function is defined as

$$L_i(p_1, p_2) = p_i - \sum_{j=1}^{m} \eta_j (g_j(\underline{x}) - c_j). \tag{3.35}$$

The Karush-Kuhn-Tucker (KKT) conditions for the optimal solutions of (3.34) are given by

$$\frac{\partial L(p_1, p_2)}{\partial p_i} = 0, i = 1, 2, \ \eta_j \geq 0, g_j(x) \leq c_j, \ \eta_j(g_j(x) - c_j) = 0, j = 1, ..., m. \tag{3.36}$$

If $\eta_j = 0$, $j = 1, 2$, we have $\frac{\partial L(p_1, p_2)}{\partial p_i} = 1$, $i = 1, 2$ and we cannot satisfy KKT conditions (3.36). Therefore, we need $\eta_j > 0$, $j = 1, 2$, and conditions (3.32)-(3.33) need to be satisfied with equality, i.e. the given arrival rates λ_1 and λ_2 must be located arbitrarily close to the boundary of the stability region $\mathcal{S}$. The constraints are convex in p_1 and p_2 (provided that $p_1 q_{1|1}(q_{2|2} - q_{2|1,2}) + p_2 q_{2|2}(q_{1|1} - q_{1|1,2}) < q_{1|1}q_{2|2}$. Then, this condition is both necessary and sufficient. The utility of node i can be

increased only if a decrease in p_i does not violate the stability conditions (3.32)-(3.33). Assume that the arrival rates λ_1 and λ_2 are arbitrarily close to the boundary of $\mathcal{S}$ for given p_1 and p_2.

If $p_1 q_{1|1}(q_{2|2} - q_{2|1,2}) + p_2 q_{2|2}(q_{1|1} - q_{1|1,2}) < q_{1|1} q_{2|2}$, node i can reduce the transmission probability p_i without violating the stability conditions until the arrival rates λ_1 and λ_2 are located at the intersection of the two stability conditions for both transmitters. Any further reduction in p_i violates the stability conditions, and the packet arrival rates λ_1 and λ_2 fall out of the stability region $\mathcal{S}$. As a result, the values of p_1 and p_2 chosen arbitrarily close to stability region $\mathcal{S}$ constitute a Nash equilibrium. This statement would not hold for the condition $p_1 q_{1|1}(q_{2|2} - q_{2|1,2}) + p_2 q_{2|2}(q_{1|1} - q_{1|1,2}) > q_{1|1} q_{2|2}$, which however cannot exist in Nash equilibrium, since nodes can reduce their transmission probabilities unilaterally without violating the stability conditions until this condition is violated. The arrival rates λ_1 and λ_2 for any transmission probabilities p_1 and p_2 such that $p_1 q_{1|1}(q_{2|2} - q_{2|1,2}) + p_2 q_{2|2}(q_{1|1} - q_{1|1,2}) > q_{1|1} q_{2|2}$ are located at the intersection of the stability conditions in another stability region for different transmission probabilities p_1' and p_2' such that $p_1 q_{1|1}(q_{2|2} - q_{2|1,2}) + p_2 q_{2|2}(q_{1|1} - q_{1|1,2}) < q_{1|1} q_{2|2}$, where $p_1' = \frac{q_{2|2}(q_{1|1} - p_2(q_{1|1} - q_{1|1,2}))}{q_{1|1}(q_{2|2} - q_{2|1,2})}$ and $p_2' = \frac{q_{1|1}(q_{2|2} - p_1(q_{2|2} - q_{2|1,2}))}{q_{2|2}(q_{1|1} - q_{1|1,2})}$. This equivalent stability region is achievable by smaller transmission probabilities and energy costs.

As a result, the non-cooperative transmission probabilities need to be chosen such that the given arrival rates λ_1 and λ_2 are located arbitrarily close to the busy service rates $p_1(q_{1|1} - p_2(q_{1|1} - q_{1|1,2}))$ and $p_2(q_{2|2} - p_1(q_{2|2} - q_{2|1,2}))$, where $p_1 q_{1|1}(q_{2|2} - q_{2|1,2}) + p_2 q_{2|2}(q_{1|1} - q_{1|1,2}) < q_{1|1} q_{2|2}$. $\qquad\square$

In addition to selfish or cooperative operation, a node may also pursue malicious objectives of blocking random access of the other selfish nodes, as we discuss next for saturated queues.

3.5   Effects of Malicious Operation on Non-Cooperative Equilibrium

Assume $N = \{1, 2\}$ and consider one selfish transmitter 1 and one malicious transmitter 2 (i.e. $M_1 = \{\emptyset\}$ and $M_2 = \{1\}$). First, we assume that malicious node 2 does not have throughput or delay objectives, i.e. $r_2 = 0$ and $D_2 = 0$.

Theorem 3.5.1 *There exist possibly multiple non-cooperative strategies in Nash equilibrium.*

(a) The mixed equilibrium strategies are given by

$$p_1^* = E_2 \left(\left(q_{1|1} - q_{1,1,2} \right) \left(c_{2,1}^{(1)} + c_{2,1}^{(2)} \right) \right)^{-1}, \qquad (3.37)$$

$$p_2^* = \left(q_{1|1} - q_{1|1,2} \right)^{-1} \left(q_{1|1} - \frac{E_1}{r_1 + D_1} \right), \qquad (3.38)$$

if $q_{1|1,2} \le \frac{E_1}{r_1 + D_1} \le q_{1|1}$ and $0 \le E_2 \le (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} + c_{2,1}^{(2)})$.

(b) The pure equilibrium strategies are given by $p_i^* = 1$, $i = 1, 2$, if $\frac{E_1}{r_1 + D_1} \le q_{1|1,2}$ and $E_2 \le (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} + c_{2,1}^{(2)})$, or $p_i^* = 0$, $i = 1, 2$, if $\frac{E_1}{r_1 + D_1} \ge q_{1|1}$ and $E_2 \ge 0$, or $p_1^* = 1$, $p_2^* = 0$, if $\frac{E_1}{r_1 + D_1} \le q_{1|1}$ and $E_2 \ge (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} + c_{2,1}^{(2)})$, or $p_1^* = 0$, $p_2^* = 1$, if $\frac{E_1}{r_1 + D_1} \ge q_{1|1,2}$ and $E_2 \le 0$.

Proof: The conditional expected utilities of node 1 are the same as in the case of two selfish nodes. The conditional expected utilities of node 2 are changed to

$$u_2(p_1, 1) = p_1(1 - q_{1|1,2})c_{2,1}^{(1)} - p_1 q_{1|1,2} c_{2,1}^{(2)} - E_2, \qquad (3.39)$$

61

$$u_2(p_1, 0) = p_1(1 - q_{1|1})c_{2,1}^{(1)} - p_1 q_{1|1} c_{2,1}^{(2)}. \tag{3.40}$$

The indifference condition of $u_i(0, p_{-i}^*) = u_i(1, p_{-i}^*)$, $i = 1, 2$, for non-cooperative Nash equilibrium leads to the mixed equilibrium strategies given by Eqs. (3.37)-(3.38), if $q_{1|1,2} \leq \frac{E_1}{r_1 + D_1} \leq q_{1|1}$ and $0 \leq E_2 \leq (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} + c_{2,1}^{(2)})$. If the indifference condition cannot be satisfied for any feasible set of transmission probabilities such that $0 \leq p_i \leq 1$, $i = 1, 2$, each node i performs with probability 1 the action of higher conditional expected utility. The pure equilibrium strategy of node $i = 1, 2$ is $p_i^* = 1$, if $u_i(1, p_{-i}^*) >_i (0, p_{-i}^*)$, or $p_i^* = 0$, if $u_i(1, p_{-i}^*) < u_i(0, p_{-i}^*)$. $\qquad \square$

The equilibrium strategies are independent of any reward that is offered to the malicious node for the purpose of forcing the selfish node to stay idle. If we assign a reward $c_{2,1}^{(3)}$ to node 2 for each idle slot of node 1, the additional reward term $c_{2,1}^{(3)}(1 - p_1)$ cancels in the indifference condition given by $u_2(1, p_1) = u_2(0, p_1)$ in non-cooperative Nash equilibrium. Therefore, there is no need for this extra reward to describe the malicious transmitter behavior.

The mixed non-cooperative strategies p_2^* of node 2 in Nash equilibrium (given by Eqs. (3.10) and (3.38)) are the same regardless of whether node 2 is selfish or malicious. In other words, malicious operation of node 2 only changes the equilibrium strategy of the selfish node 1 to be attacked. If node 2 is selfish, we have

$$\lambda_1 = \frac{q_{2|2} - \frac{E_2}{r_2 + D_2}}{q_{2|2} - q_{2|1,2}} \left(1 - \left(q_{1|1} - q_{1|1,2} \right)^{-1} \left(q_{1|1} - \frac{E_1}{r_1 + D_1} \right) \right). \tag{3.41}$$

If node 2 is malicious, the throughput rate λ_1 of node 1 is changed to

$$\lambda_1 = E_2 \frac{1 - \left(q_{1|1} - q_{1|1,2} \right)^{-1} \left(q_{1|1} - \frac{E_1}{r_1 + D_1} \right)}{\left(q_{1|1} - q_{1|1,2} \right) \left(c_{2,1}^{(1)} + c_{2,1}^{(2)} \right)}. \tag{3.42}$$

Node 2 can reduce the throughput rate λ_1 of node 1 by switching to malicious operation, i.e. the denial of service attack of node 2 is successful, if and only if

$$\left(q_{1|1} - q_{1|1,2}\right)\left(c_{2,1}^{(1)} + c_{2,1}^{(2)}\right) > E_2 \left(q_{2|2} - q_{2|1,2}\right)\left(q_{2|2} - \frac{E_2}{r_2 + D_2}\right)^{-1}, \qquad (3.43)$$

i.e. the reward for jamming transmissions of node 1 should be high enough for node 2 to compensate the increase in energy cost and the decrease in individual throughput rate.

If both nodes are in selfish equilibrium and node 2 switches to malicious operation, the utility u_2 becomes an increasing function of p_2 under condition (3.43). Node 2 would increase p_2 to improve u_2 and node 1 would react by decreasing p_1. For large values of p_2, the energy cost of node 2 dominates the rewards for blocking transmissions of node 1. Therefore, node 2 chooses the same transmission probability p_2 as in the selfish equilibrium case, whereas p_1 is reduced below the selfish equilibrium value, if condition (3.43) holds. For the classical collision channels with $q_{i|i} = 1$ and $q_{i|1,2} = 0$, $i = 1, 2$, Figure 3.2 depicts the minimum value of $c_{1,2}^{(1)} + c_{1,2}^{(2)}$ (as function of the energy cost E_2) such that node 2 can reduce the throughput rate λ_1 of node 1. For symmetric costs $E_i = E$, $r_i + D_i = 1$, $i = 1, 2$, and $c_{1,2}^{(1)} = c_{1,2}^{(2)}$, Figures 3.3 depicts the throughput rates in non-cooperative equilibrium as function of $c_{1,2}^{(1)} + c_{1,2}^{(2)}$ for different values of energy cost E.

Consider the mixed equilibrium strategies for $E \leq r + D$ and $E \leq c_{2,1}^{(1)} + c_{2,1}^{(2)}$. The transmission probability of any node in cooperative equilibrium is reduced to half of the value in selfish operation. If transmitter 2 becomes malicious, the only change occurs in the transmission probability of selfish transmitter 1. The
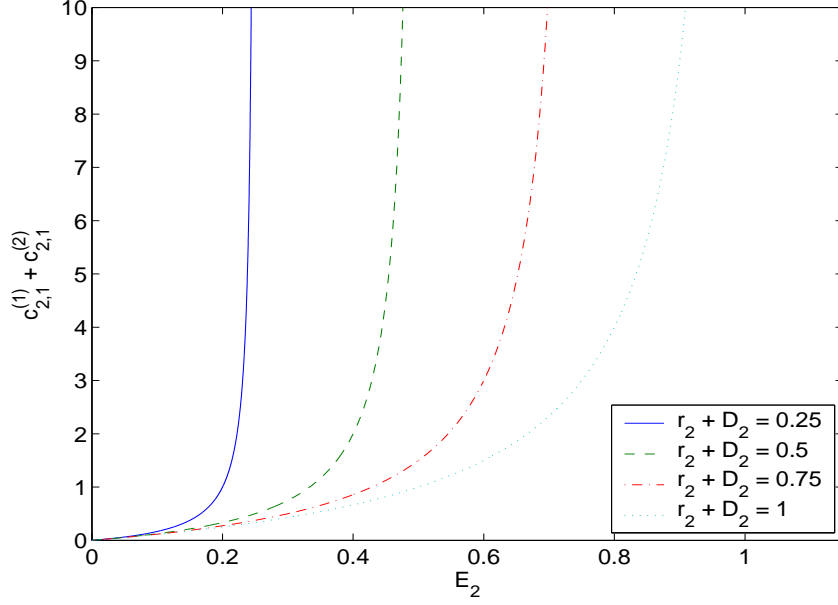
Figure 3.2: The minimum value of attack parameters $c_{1,2}^{(1)} + c_{1,2}^{(2)}$ as function of energy cost E_2 such that malicious node 2 can reduce the throughput of selfish node 1.

utility in cooperative equilibrium is strictly greater than the case of selfish operation. However, the throughput rates under selfish operation can be higher than those under cooperative operation (particularly for large values of energy costs provided that $3E > R+D$). On the other hand, when transmitter 2 becomes malicious, it may increase its own utility depending on the cost and reward parameters. Furthermore, the malicious operation of transmitter 2 may be better than selfish operation in terms of reducing the throughput rate of transmitter 1, which occurs only if the malicious attack parameters are relatively greater than the energy cost provided that $(c_{2,1}^{(1)} + c_{2,1}^{(2)}) \left(1 - \frac{E}{r+D}\right) > E$.

Next, we assume that malicious node 2 has also packets to transmit and pursues the additional objectives of optimizing the throughput rewards and delay costs
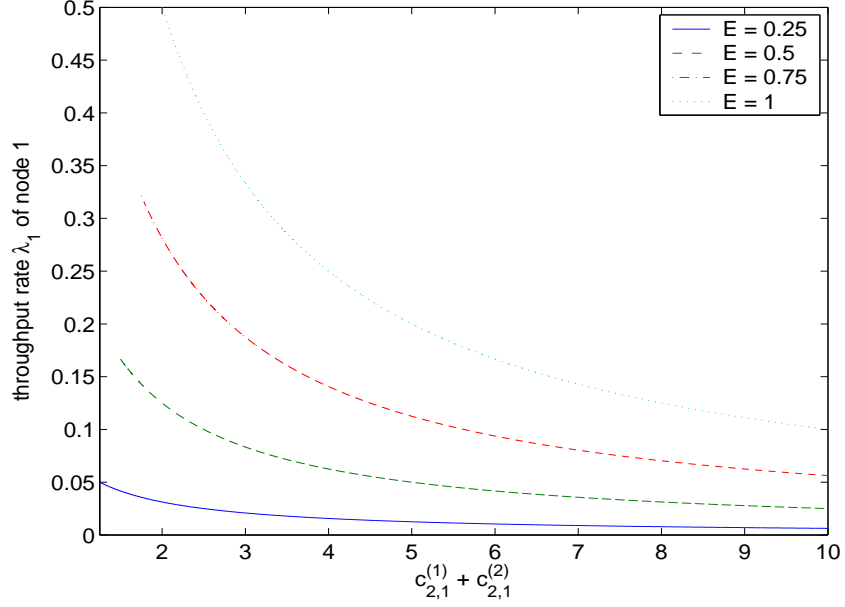
Figure 3.3: The throughput λ_1 of selfish node 1 as function of attack parameters $c_{1,2}^{(1)} + c_{1,2}^{(2)}$ in non-cooperative equilibrium.

(i.e. $r_2 > 0$ and $D_2 > 0$) in addition to the energy costs and malicious objectives of blocking random access of selfish node 1.

Theorem 3.5.2 *There exist possibly multiple non-cooperative strategies in Nash equilibrium.*

(a) The mixed equilibrium strategies are given by

$$p_1^* = \frac{E_2 - (r_2 + D_2)q_{2|2}}{(q_{1|1} - q_{1|1,2})\left(c_{2,1}^{(1)} + c_{2,1}^{(2)}\right) - (q_{2|2} - q_{2|1,2})(r_2 + D_2)}, \qquad (3.44)$$

$$p_2^* = \left(q_{1|1} - q_{1|1,2}\right)^{-1}\left(q_{1|1} - \frac{E_1}{r_1 + D_1}\right), \qquad (3.45)$$

if $q_{1|1,2} \leq \frac{E_1}{r_1 + D_1} \leq q_{1|1}$ and $0 \leq E_2 - (r_2 + D_2)q_{2|2} \leq (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} + c_{2,1}^{(2)}) - (q_{2|2} - q_{2|1,2})(r_2 + D_2)$.

(b) The pure equilibrium strategies are given by $p_i^* = 1$, $i = 1, 2$, if $\frac{E_1}{r_1 + D_1} \leq q_{1|1,2}$ and

$E_2 \leq (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} + c_{2,1}^{(2)}) + q_{2|1,2}(r_2 + D_2)$, or $p_i^* = 0$, $i = 1, 2$, if $\frac{E_1}{r_1 + D_1} \geq q_{1|1}$ and

$E_2 \geq (r_2 + D_2)q_{2|2}$, or $p_1^* = 1$, $p_2^* = 0$, if $\frac{E_1}{r_1 + D_1} \leq q_{1|1}$ and $E_2 \geq (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} +$

$c_{2,1}^{(2)}) + q_{2|1,2}(r_2 + D_2)$, or $p_1^* = 0$, $p_2^* = 1$, if $\frac{E_1}{r_1 + D_1} \geq q_{1|1,2}$ and $E_2 \leq (r_2 + D_2)q_{2|2}$.

Proof: The conditional expected utilities of node 1 are the same as the case of two selfish nodes or the case of one selfish node and one malicious node without throughput or delay objectives. The conditional expected utilities of node 2 are changed to

$$u_2(p_1, 1) = (r_2 + D_2)((1 - p_1)q_{2|2} + p_1 q_{2|1,2}) \tag{3.46}$$

$$+ p_1(1 - q_{1|1,2})c_{2,1}^{(1)} - p_1 q_{1|1,2} c_{2,1}^{(2)} - E_2 - D_2,$$

$$u_2(p_1, 0) = p_1(1 - q_{1|1})c_{2,1}^{(1)} - p_1 q_{1|1} c_{2,1}^{(2)} - D_2. \tag{3.47}$$

The indifference condition of $u_i(0, p_{-i}^*) = u_i(1, p_{-i}^*)$, $i = 1, 2$, for non-cooperative Nash equilibrium leads to the mixed equilibrium strategies given by Eqs. (3.44)-(3.45), if $q_{1|1,2} \leq \frac{E_1}{r_1 + D_1} \leq q_{1|1}$ and $0 \leq E_2 - (r_2 + D_2)q_{2|2} \leq (q_{1|1} - q_{1|1,2})(c_{2,1}^{(1)} + c_{2,1}^{(2)}) - (q_{2|2} - q_{2|1,2})(r_2 + D_2)$.

If the indifference condition cannot be satisfied for any feasible set of transmission probabilities such that $0 \leq p_i \leq 1$, $i = 1, 2$, each node i performs with probability 1 the action of higher conditional expected utility. The pure equilibrium strategy of node $i = 1, 2$ is $p_i^* = 1$, if $u_i(1, p_{-i}^*) > u_i(0, p_{-i}^*)$, or $p_i^* = 0$, if $u_i(1, p_{-i}^*) < u_i(0, p_{-i}^*)$. □

Malicious node 2 cannot reduce the throughput rate of node 1 compared to selfish operation for any given value of $c_{2,1}^{(1)} > 0$ and $c_{2,1}^{(2)} > 0$. The denial of service attack of node 2 is unsuccessful, unless node 2 receives a smaller throughput reward

66

r_2' (compared to the selfish operation with throughput reward r_2). For the classical collision channels with $q_{i|i} = 1$ and $q_{i|1,2} = 0$, $i = 1, 2$, the value of r_2' needs to satisfy

$$r_2' < r_2 \left(1 - \left(c_{2,1}^{(1)} + c_{2,1}^{(2)}\right) E_2^{-1}\right) - \left(c_{2,1}^{(1)} + c_{2,1}^{(2)}\right) \left(D_2 E_2^{-1} - 1\right). \qquad (3.48)$$

Hence, node 2 needs to compromise between selfish throughput and malicious attack objectives.

3.6 Adaptive Update Mechanisms for Distributed Operation

We consider adaptive update mechanisms as distributed random access algorithms. Each node independently chooses the transmission probability based on the channel feedback from the receiver. We distinguish two separate cases with and without availability of information on system parameters (such as rewards, costs and capture probabilities) at individual transmitter nodes.

Nodes can observe the channel outcome, namely whether it is an idle slot, a successful packet transmission or a packet collision. Therefore, nodes are aware of the actual action of each other in each time slot but do not know the transmission probabilities of each other. As a distributed algorithm, we consider a repeated play such that nodes adjust their transmission strategies in response to the observations of the channel outcome by playing the optimized best response to the empirical frequencies of each other's actions. The equilibrium strategies coincide with the Nash equilibrium strategies, if the empirical frequencies converge [35]. Let $B_i(p_{-i}) = \arg\max_{p_i \in [0,1]} u_i(p_i, p_{-i})$ denote the best-response function of node i, where p_{-i} is the transmission probability of the node other than node i. Let $A_i(t)$ be 1 or 0, if

node i transmits a packet at time slot t or not. The empirical frequency $\bar{p}_i(t+1)$ of node i at time $t+1$ is defined as the running average of actions of node $i = 1, 2$:

$$\bar{p}_i(t+1) = \min\left(\max\left(\bar{p}_i(t) + \frac{1}{t+1}(A_i(t) - \bar{p}_i(t)), 0\right), 1\right). \tag{3.49}$$

The strategy of node $i = 1, 2$ at time $t+1$ is the optimal response to the running average of the opponent's actions:

$$p_i(t+1) = B_i(\bar{p}_{-i}(t)) \quad, \quad i = 1, 2. \tag{3.50}$$

Consider the equilibrium transmission probabilities p_i^*, $i = 1, 2$, given in Eqs. (3.9)-(3.10) or (3.37)-(3.38) or (3.44)-(3.45) depending on whether the system consists of two selfish nodes, or one selfish node and one malicious node in the absence or presence of throughput and delay objectives, respectively. Provided that $p_i^* \in [0, 1]$, $i = 1, 2$, we have $A_1(t) = 1$, if $\bar{p}_2(t) \leq p_2^*$, or $A_1(t) = 0$, otherwise, and we have $A_2(t) = 1$, if $\bar{p}_1(t) \geq p_1^*$, or $A_2(t) = 0$, otherwise.

The choice of the initial transmission probabilities affects which of the possibly multiple Nash equilibrium strategies the algorithm will track and converge to. The analysis of the convergence properties of the best-response update mechanism is outside the scope of this chapter. For the classical collision channels with $q_{i|i} = 1$ and $q_{i|1,2} = 0$, $i = 1, 2$, and system parameters $E_i = 0.25$, $D_i = 0$, $i = 1, 2$, and $r_1 = 1$, we illustrate in Figure 3.4 how the transmission probabilities evolve from the equilibrium of two selfish nodes with $p_1^* = 0.75$ and $p_2^* = 0.75$ (for $r_2 = 1$ and $c_{2,1}^{(1)} = c_{2,1}^{(2)} = 0$) to the equilibrium of one selfish and node one malicious node with $p_1^* = 0.5$ and $p_2^* = 0.75$ (for $r_2 = 0$ and $c_{2,1}^{(1)} = c_{2,1}^{(2)} = 0.25$).
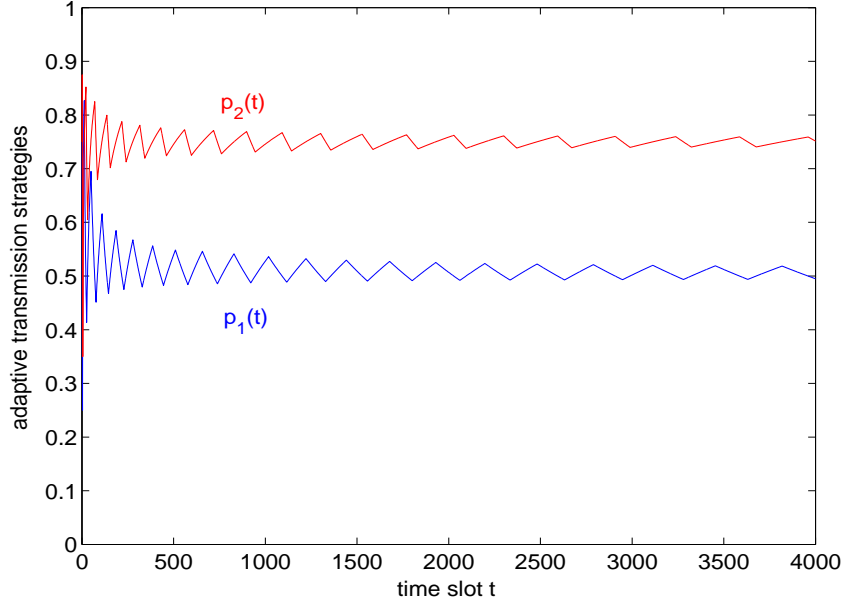
Figure 3.4: Best-response updates.

The proposed update mechanism given by Eqs. (3.49)-(3.50) depends on the availability of the reward and cost information at all transmitter nodes. Next, we assume that the reward and cost parameters are unknown in a distributed operation and nodes vary their transmission strategies as function of time depending on the channel feedback only. We define the step-size sequence $\epsilon(t)$ that satisfies $\lim_{t\to\infty} \epsilon(t) = 0$, $\sum_{t=1}^{\infty} \epsilon(t) = \infty$ and $\sum_{t=1}^{\infty} \epsilon(t)^2 < \infty$ as suggested by the stochastic approximation theory, and we define the update sequence $\zeta_i(t) \in \{-1, 0, 1\}$, $i = 1, 2$. The strategy of node $i = 1, 2$ is updated at time $t + 1$ as

$$p_i(t + 1) = \min\left(\max\left(p_i(t) + \epsilon(t)\,\zeta_i(t)\,,\,0\right)\,,\,1\right). \tag{3.51}$$

We consider the approach that nodes do not change strategies, as long as their decisions in the previous time slot have been successful given the opponent's decision. If any node i is successful given the opponent's decision at time t, $\zeta_i(t) = 0$.

69
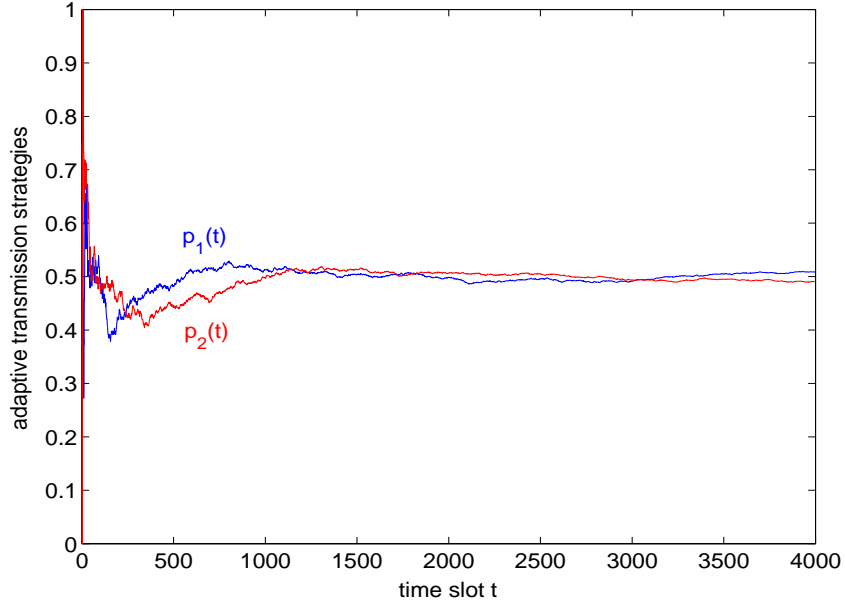
Figure 3.5: Heuristic updates based on channel feedback only.

If any node i transmits at time slot t and it is an unsuccessful decision given the opponent's decision, node i reduces the transmission probability, i.e. $\zeta_i(t) = -1$. Similarly, if any node i waits at time slot t and it is an unsuccessful decision given the opponent's decision, node i increases the transmission probability, i.e. $\zeta_i(t) = 1$.

Any selfish node prefers successful transmission over waiting and prefers waiting over packet collision. Similarly, any malicious node prefers blocking random access of another node over waiting and prefers unnecessary transmission without blocking over missing a denial of service attack opportunity. Consider the case of one selfish node and one malicious node without throughput or delay objectives. In any time slot, node 1 is successful in its decision, if node 1 transmits, while node 2 is waiting, or if node 1 waits, while node 2 is transmitting. Similarly, we assume that node 2 is successful in its decision, if node 2 transmits, while node 1 is transmitting,

70

or if node 2 waits, while node 1 is waiting. The update sequences are defined as $\zeta_1(t) = -1, \zeta_2(t) = 0$, if $A_i(t) = 1$, $i = 1, 2$, or $\zeta_1(t) = 0, \zeta_2(t) = 1$, if $A_1(t) = 1$, $A_2(t) = 0$, or $\zeta_1(t) = 0, \zeta_2(t) = -1$, if $A_1(t) = 0$, $A_2(t) = 1$, or $\zeta_1(t) = 1, \zeta_2(t) = 0$, $A_i(t) = 0$, $i = 1, 2$. Due to the constraints on the availability of parameter information, the resulting transmission strategies can significantly deviate from the Nash equilibrium, as shown in Figure 3.5 for the step-size sequence $\epsilon(t) = \frac{1}{t}$, $t \geq 1$

3.7 Arbitrary Number of Transmitters

In this section, we extend the results to arbitrary sets of selfish and malicious transmitter nodes in random access for the case of saturated queues. Following the conjecture in [22], future work should investigate the equivalence of the non-cooperative strategies for saturated queues and stable operation. Let $\underline{p}_{-i}$ denote the transmission probabilities of nodes other than node i. For any p_i, $i \in N$, the non-cooperative Nash equilibrium strategies p_i^*, $i \in N$, satisfy

$$u_i(p_i^*, \underline{p}_{-i}^*) \geq u_i(p_i, \underline{p}_{-i}^*) \quad , \quad i \in N. \tag{3.52}$$

3.7.1 Arbitrary Number of Selfish Transmitters

Theorem 3.7.1 *Consider transmitter set N of n selfish nodes. If $q_{i|N} \leq \frac{E_i}{r_i + D_i} \leq q_{i|i}$, $i \in N$, the non-cooperative mixed strategies p_i^*, $i \in N$, in Nash equilibrium satisfy*

$$\sum_{J \subseteq N : i \in J} q_{i|J} \prod_{j \in J - \{i\}} p_j^* \prod_{k \notin J} (1 - p_k^*) = \frac{E_i}{r_i + D_i}, \quad i \in N, \tag{3.53}$$

71

and the non-cooperative equilibrium throughput of node i is given by

$$\lambda_i^* = p_i^* \frac{E_i}{r_i + D_i}, \quad i \in N. \tag{3.54}$$

Proof: The expected utility $u_i(0, \underline{p}_{-i})$ of waiting node i is $-D_i$ and the expected utility $u_i(1, \underline{p}_{-i})$ of transmitting node i is $\lambda_i' - (1 - \lambda_i')D_i - E_i$, where $\lambda_i = p_i \lambda_i'$ satisfies

$$\lambda_i \leq \sum_{J \subseteq N: i \in J} q_{i|J} \prod_{j \in J - \{i\}} p_j^* \prod_{k \notin J} (1 - p_k^*), \quad i \in N. \tag{3.55}$$

The indifference condition given by $u_i(0, \underline{p}_{-i}) = u_i(1, \underline{p}_{-i})$ for non-cooperative Nash equilibrium leads to the equalities of $\lambda_i' = \sum_{J \subseteq N: i \in J} q_{i|J} \prod_{j \in J - \{i\}} p_j^* \prod_{k \notin J} (1 - p_k^*) = \frac{E_i}{r_i + D_i}$, $i \in N$, and the equilibrium throughput of node i follows from $\lambda_i^* = p_i^* \lambda_i'$ and it is given by Eq. (3.54). $\square$

Next, we consider classical collision channels with $q_{i|J} = 1$, $J = \{i\}$, and $q_{i|J} = 0$, $J \neq \{i\}$.

Theorem 3.7.2 *For classical collision channels, the mixed strategies in Nash equilibrium are*

$$p_i^* = 1 - \left(\frac{r_i + D_i}{E_i} \right) \left(\prod_{j=1}^n \frac{E_j}{r_j + D_j} \right)^{\frac{1}{n-1}}, \, i \in N. \tag{3.56}$$

Proof: Let $x_i = 1 - p_i$ and $y_i = \frac{E_i}{1 + D_i}$ for $i = 1, ..., n$. Define $\lambda_i = p_i \lambda_i'$. From Eq. (3.54), we have $\lambda_i' = \prod_{j \neq i} x_j = y_i$ for any i such that $x_k y_k = x_l y_l$ for any k and l. If condition $\prod_{j \neq k} x_j = y_k$ is rewritten as $\prod_{j \neq k} \frac{x_i y_i}{y_j} = y_k$, i.e. $(x_i y_i)^{n-1} = \prod_{j=1}^n y_j$, Eq. (3.56) follows by setting $y_i = \frac{E_i}{r_i + D_i}$ and $x_i = \frac{1}{y_i} (\prod_{j=1}^n y_j)^{\frac{1}{n-1}}$ in $p_i = 1 - x_i$ for $i = 1, ..., n$. $\square$

From Eqs. (3.54) and (3.56), the non-cooperative throughput of node i in Nash Equilibrium is

$$\lambda_i^* = \left(\frac{E_i}{r_i + D_i}\right) - \left(\prod_{j=1}^{n} \frac{E_j}{r_j + D_j}\right)^{\frac{1}{n-1}} \tag{3.57}$$

for classical collision channels. The utility u_i of node i in Nash equilibrium is equal to $-D_i$ and the non-cooperative total utility u_Σ^{nc} is $-\sum_{i=1}^{n} D_i$, whereas the cooperative total utility is

$$u_\Sigma = \sum_{i=1}^{n} p_i((r_i + D_i) \prod_{j \neq i}(1 - p_j) - E_i - D_i \tag{3.58}$$

for $\beta_i = 1$, $i = 1, ..., n$. Eq. (3.58) is concave in p_i and can be maximized by the transmission probabilities p_i, $i = 1, ..., n$, that satisfy $\frac{\partial u_\Sigma}{\partial p_i} = 0$, $i = 1, ..., n$. The total utility improvement $u_\Sigma^c - u_\Sigma^{nc}$ is $\sum_{i=1}^{n} p_i((r_i + D_i) \prod_{j \neq i}(1 - p_j) - E_i)$. For the symmetric energy and delay costs such that $\frac{E_i}{r_i + D_i} = \frac{E_S}{r_S + D_S}$ for $i = 1, ..., n$, the improvement $u_\Sigma^c - u_\Sigma^{nc}$ is maximized by the common transmission probability p^* that satisfies $(1 - p^*)^{n-2}(1 - np^*) = \frac{E_S}{r_S + D_S}$. From Eq. (3.54), the common non-cooperative throughput is given by $\frac{E_S}{r_S + D_S} - \left(\frac{E_S}{r_S + D_S}\right)^{\frac{n}{n-1}}$ for the symmetric energy and delay costs and is maximized to $\frac{1}{n}\left(1 - \frac{1}{n}\right)^{n-1}$ provided that $E_S = (r_S + D_S)\left(1 - \frac{1}{n}\right)^{n-1}$.

The total throughput $\sum_{i=1}^{n} \lambda_i$ is illustrated in Figure 3.6. If the energy costs are high and the delay costs are low, the non-cooperative operation can improve the total throughput in the system compared to utility-optimal cooperation. For $r_S = 1$ and $D_S = 0$, Figures 3.7, 3.8 and 3.9 depict the transmission probability, individual throughput rate and utility, respectively, as function of the number of transmitters n in non-cooperative and cooperative equilibrium for different values of energy cost E_S. Selfish nodes are more aggressive in their transmission strategies compared to

73

cooperative nodes. The equilibrium transmission probability, throughput rate and cooperative utility decrease as the number of selfish transmitters competing for the channel increases.
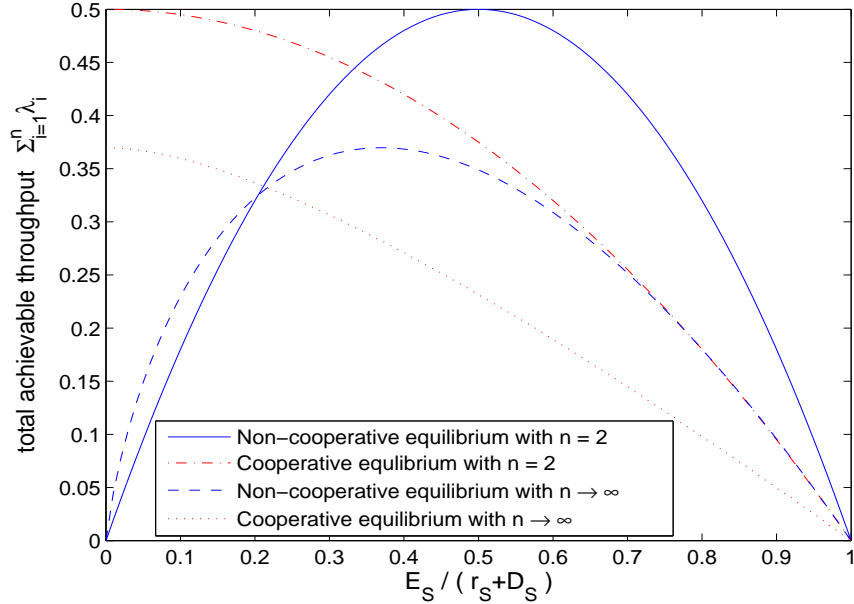
Figure 3.6: The total throughput $\sum_{i=1}^{n} \lambda_i$ as function of rewards and costs.

3.7.2 Comparison with Scheduled Access

Consider deterministic scheduled access such that nodes are centrally scheduled in any time slot to transmit a packet or to wait.

Theorem 3.7.3 *The utility-optimal scheduled access dedicates the channel to node group $J \subseteq N$ with the largest positive value of $\sum_{i \in J \subseteq N} \beta_i((r_i + D_i)q_{i|J} - E_i)$ and can improve the total utility $u_\Sigma = \sum_{i=1}^{n} \beta_i u_i$ of non-cooperative random access by the amount of*

74
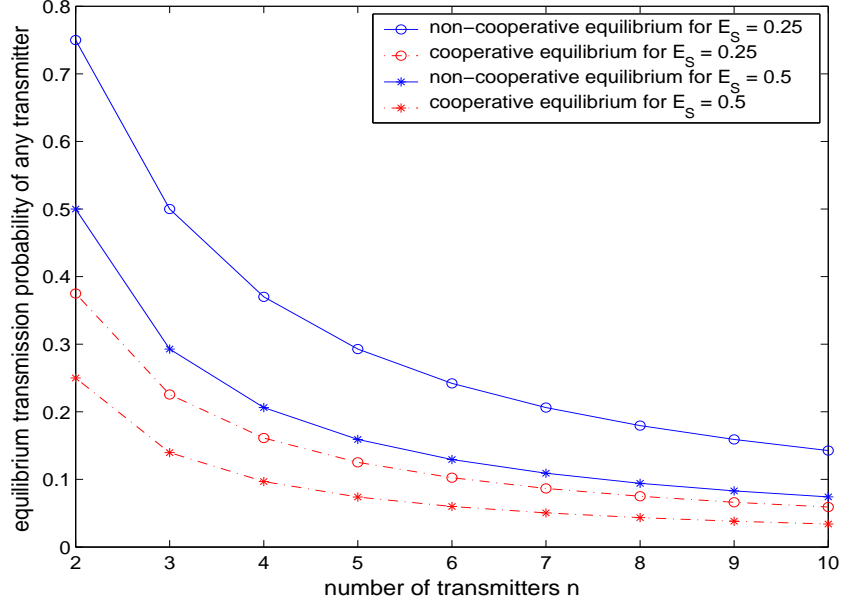
Figure 3.7: The equilibrium transmission probability of selfish transmitters.

$$\max_{J \subseteq N} \left[\sum_{i \in N} \beta_i((r_i + D_i)q_{i|i} - E_i) \right]^+, \qquad (3.59)$$

where $[x]^+ = x$, if $x > 0$, and $[x]^+ = 0$, if $x \leq 0$.

Proof: Let τ_J denote the disjoint time fraction allocated to transmissions of node group $J \subseteq N$ such that $0 \leq \tau_J \leq 1$ and $\sum_{J \subseteq N} \tau_J \leq 1$. The utility of node i is

$$u_i = \sum_{J \subseteq N: i \in J} \tau_J \left(r_i q_{i,J} - E_i - (1 - q_{i|J})D_i \right) - (1 - \sum_{J \subseteq N: i \in J} \tau_J)D_i \qquad (3.60)$$

and the resulting value of total utility is

$$u_\Sigma = \sum_{i \in N} \left(\sum_{J \subseteq N: i \in J} \tau_J \, \beta_i \left((r_i + D_i)q_{i,J} - E_i \right) \right) - \sum_{i \in N} \beta_i D_i \qquad (3.61)$$

$$= \sum_{J \subseteq N} \tau_J \left(\sum_{i \in J} \beta_i \left((r_i + D_i)q_{i,J} - E_i \right) \right) - \sum_{i \in N} \beta_i D_i. \qquad (3.62)$$

We need to dedicate the channel to node group J with the largest positive value of $\sum_{i \in J \subseteq N} \beta_i((r_i + D_i)q_{i|J} - E_i)$ to maximize u_Σ. If a non-zero time fraction is
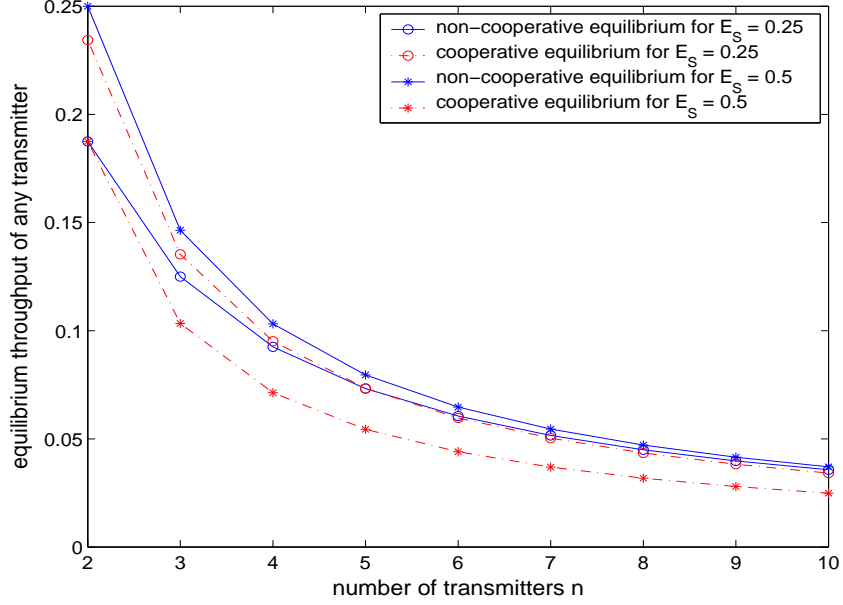
Figure 3.8: The equilibrium throughput of selfish transmitters.

allocated to another node group J', we can improve u_Σ by shifting the time allocation of node group J' to the particular node group J. As a result, u_Σ is maximized to

$$u_\Sigma = \max_{J \subseteq N} \left(\sum_{i \in N} \beta_i \left[((r_i + D_i)q_{i|J} - E_i) \right]^+ \right) - \sum_{i \in N} \beta_i D_i. \tag{3.63}$$

If $\sum_{i \in J \subseteq N} \beta_i((r_i + D_i)q_{i|J} - E_i) < 0$ for all $J \subseteq N$, then no node transmits. The improvement $u_\Sigma^c - u_\Sigma^{nc}$ is given by Eq. (3.63), where $u_\Sigma^{nc} = -\sum_{i \in N} D_i$. $\qquad \square$

For the classical collision channels and symmetric costs of $\frac{E_i}{r_i + D_i} = \frac{E_S}{r_S + D_S}$, $i \in N$, Figure 3.10 compares the non-cooperative equilibrium utility u_Σ^{nc} of random access with the cooperative equilibrium utility u_Σ^c of random or scheduled access.

3.7.3 Arbitrary Sets of Selfish and Malicious Transmitters

Any malicious node $i \in N$ has the objective of blocking random access of the other selfish nodes in node set $M_i \subseteq N - \{i\}$. The conditional expected utilities of
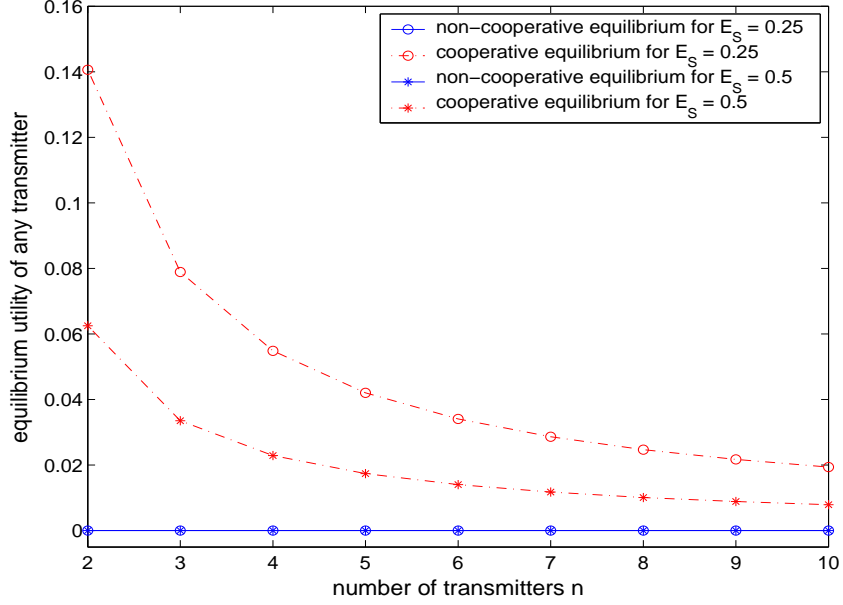
Figure 3.9: The equilibrium utility of selfish transmitters.

any node $i \in N$ are

$$u_i(1, \underline{p}_{-i}) = (r_i + D_i) \sum_{J \subseteq N : i \in J} q_{i|J} \prod_{j \in J - \{i\}} p_j \prod_{k \notin J} (1 - p_k) - E_i - D_i \qquad (3.64)$$

$$+ \sum_{j \in M_i} c_{i,j}^{(1)} p_j \left(1 - \sum_{J \subseteq N : (i,j) \subseteq J} q_{j|J} \prod_{k \in J - \{i,j\}} p_k \prod_{l \notin J} (1 - p_l) \right)$$

$$- \sum_{j \in M_i} c_{i,j}^{(2)} p_j \sum_{J \subseteq N : i \in J, j \in J} q_{j|J} \prod_{k \in J - \{i,j\}} p_k \prod_{l \notin J} (1 - p_l),$$

$$u_i(0, \underline{p}_{-i}) = \sum_{j \in M_i} c_{i,j}^{(1)} p_j \left(1 - \sum_{J \subseteq N : i \notin J, j \in J} q_{j|J} \prod_{k \in J - \{j\}} p_k \prod_{l \notin J, l \neq i} (1 - p_l) \right) \quad (3.65)$$

$$- D_i - \sum_{j \in M_i} c_{i,j}^{(2)} p_j \sum_{J \subseteq N : i \notin J, j \in J} q_{j|J} \prod_{k \in J - \{j\}} p_k \prod_{l \notin J, l \neq i} (1 - p_l)$$

for transmitting and waiting, respectively. The mixed non-cooperative strategies p_i^*, $i \in N$, in Nash equilibrium satisfy the indifference condition of $u_i(1, \underline{p}_{-i}^*) = u_i(0, \underline{p}_{-i}^*)$, whereas the pure non-cooperative strategy of node i in Nash equilibrium is given by $p_i^* = 0$, if $u_i(1, \underline{p}_{-i}^*) < u_i(0, \underline{p}_{-i}^*)$, and $p_i^* = 1$, if $u_i(1, \underline{p}_{-i}^*) > u_i(0, \underline{p}_{-i}^*)$. Consider
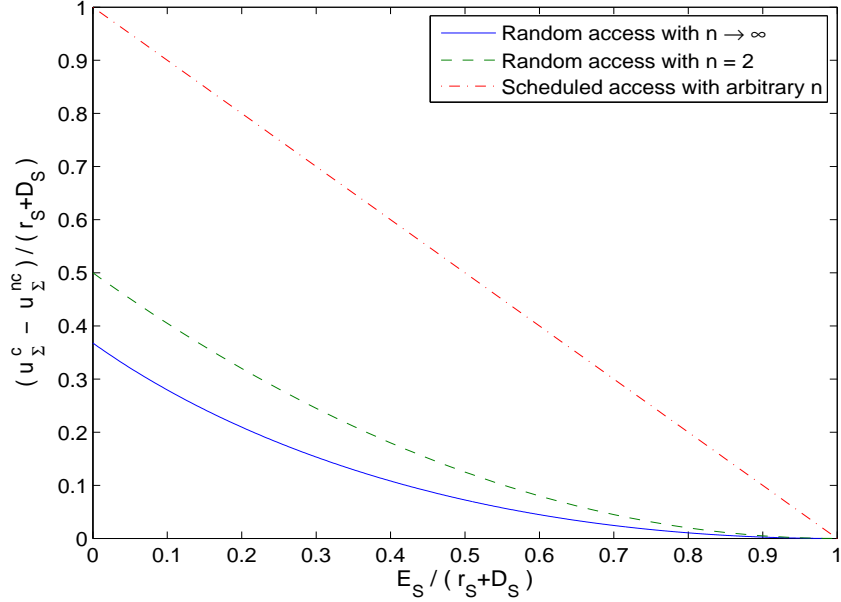
Figure 3.10: The improvement of cooperation over non-cooperative equilibrium.

the classical collision channels with $q_{i|J} = 1$ for $J = \{i\}$ and $q_{i|J} = 0$ for $J \neq \{i\}$. Partition the node set N into the disjoint sets S and M of selfish and malicious transmitters. Assume symmetric costs $E_i = E_S$, $D_i = D_S$, $r_i = r_S$, $c_{i,j}^{(1)} = 0$, $c_{i,j}^{(2)} = 0$ for $i \in S$ and $E_i = E_M$, $D_i = D_M$, $r_i = r_M$, $M_i = S$, $c_{i,j}^{(1)} = c^{(1)}$, $c_{i,j}^{(2)} = c^{(2)}$ for $i \in M$ and $j \in S$. Let m denote the number of malicious nodes. The mixed non-cooperative strategies in Nash equilibrium have the symmetric form of $p_i^* = p_S$, $i \in S$, and $p_i^* = p_M$, $i \in M$, and satisfy

$$(1 - p_S)^{n-m-1}(1 - p_M)^m = \frac{E_S}{r_S + D_S}, \tag{3.66}$$

$$(r_M + D_M)(1 - p_S)^{n-m}(1 - p_M)^{m-1}$$

$$+(n - m)p_S c^{(1)}(1 - p_S)^{n-m-1}(1 - p_M)^{m-1} \tag{3.67}$$

$$+(n - m)p_S c^{(2)}(1 - p_S)^{n-m-1}(1 - p_M)^{m-1} = E_M.$$

From Eq. (3.68), the non-cooperative throughput of any selfish node $i \in S$ is

$$\lambda_S = \frac{E_S}{r_S + D_S}\, p_S \,.\tag{3.68}$$

For the classical collision channels, only one transmission is sufficient to block the random access of all other selfish nodes, i.e. independent transmissions of multiple malicious nodes increases the energy costs without extra benefits in terms of reducing the throughput rate of selfish nodes. The average utility of any malicious node is maximized for $m = 1$. Therefore, if malicious nodes are allowed to cooperate with each other, they should form a coalition such that only one node transmits at a time to minimize the energy costs and maximize the utilities. If we consider multi-packet reception channels such that the packet capture probability decreases as the number of simultaneous transmissions increases, it may be desirable that multiple malicious nodes transmit simultaneously to block the transmissions of the other selfish nodes.

Next, we present numerical results to investigate the interactions between selfish and malicious nodes. We assume $E_M = 0.1$, $c^{(1)} = 0.25$, $c^{(2)} = 0.25$, $D_S = 0$, $D_M = 0$, $r_S = 1$, $r_M = 0$ and $m = n - 1$, i.e. there is only one selfish transmitter. Figure 3.11 shows the symmetric transmission probability in non-cooperative equilibrium as function of the number of transmitters n for different values of energy cost E_S. Selfish node needs to increase the transmission probability to maintain the throughput rate for increasing number of malicious nodes. On the other hand, malicious nodes need to transmit less frequently, as more nodes switch to the malicious operation. Both equilibrium transmission probabilities of selfish and malicious
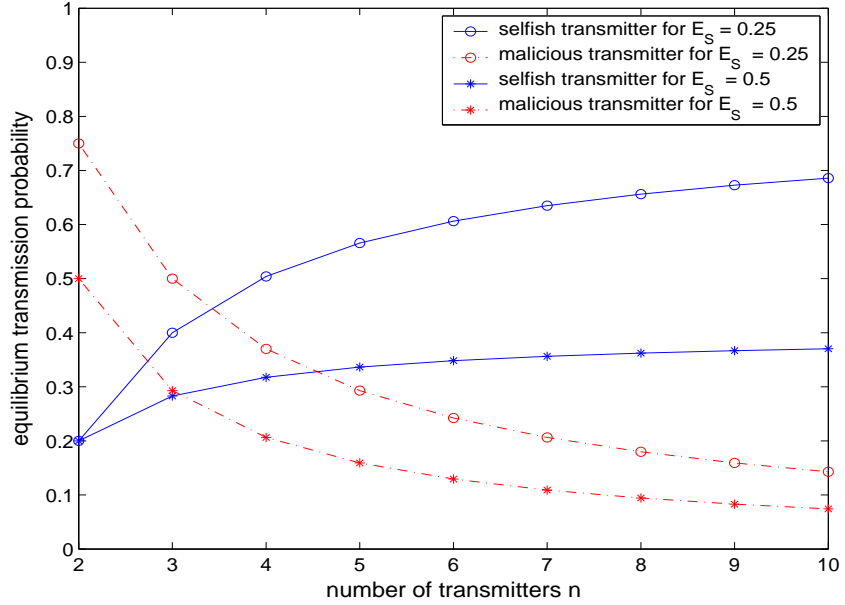
Figure 3.11: The non-cooperative equilibrium transmission probabilities.

nodes decrease with increasing energy costs, i.e. nodes become less aggressive in their transmission strategies, when the energy costs are high.

Figure 3.12 shows the throughput rate of the selfish node in non-cooperative equilibrium. The throughput rate of selfish node increases with the number of transmitters n, since the selfish node starts transmitting with higher probability, while malicious nodes transmit with smaller probabilities. We also observe that both selfish and malicious nodes transmit less frequently, when the energy costs increase, and the overall effect is an increase in the throughput rate.

Figure 3.13 shows the individual utility in non-cooperative equilibrium. The utility of the selfish node is equal to the zero delay cost and the utility of any malicious node increases with n. Malicious nodes with smaller transmission probabilities are more successful in denial of service attacks. The simultaneous operation of ma-
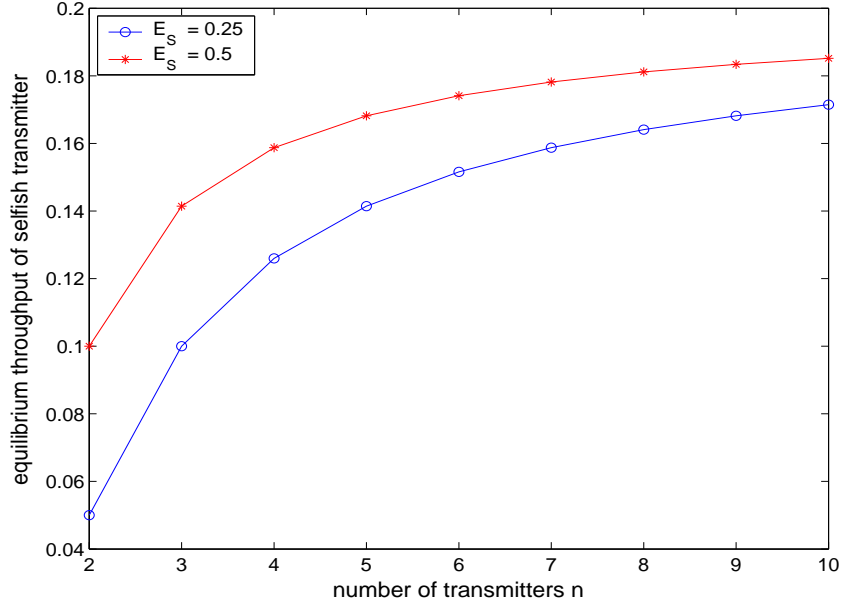
Figure 3.12: The non-cooperative equilibrium throughput of the selfish transmitter.

licious nodes reduces the individual energy cost and increases the individual utility. A single malicious node is more effective in reducing the throughput of selfish node at the expense of increasing the individual energy cost compared to the simultaneous operation of malicious nodes. Malicious nodes can reduce the throughput of the selfish node further for smaller energy costs and can increase their utilities.

3.8 Extensions to Multiple Receivers in Broadcast Communication

Let $q_{i|J}^{(j)}$ denotes the probability that the transmission of node i is successfully received by receiver node j, if nodes in set J transmit in the same time slot. We consider next the broadcasting operation of two selfish nodes randomly transmitting packets addressed to two receivers. A packet transmission contributes to the throughput only if it is received by both receivers. Broadcast communication in
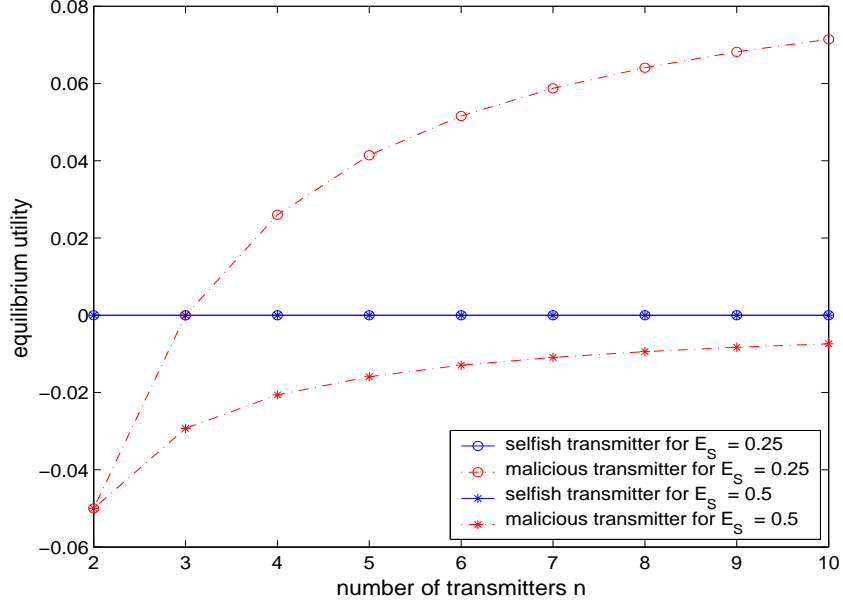
Figure 3.13: The non-cooperative equilibrium utilities.

random access will be addressed in Chapters 6-8 through the application of network coding. Each node is required to transmit linear combinations of packets based on the channel feedback from receivers [36] such that the achievable throughput of node $i = 1, 2$ can be optimized to the Max-flow Min-cut bound of

$$\lambda_1 \leq p_1 (\min_{j=1,2} (p_2 q_{1|\{1,2\}}^{(j)} + (1 - p_2) q_{1|1}^{(j)})), \tag{3.69}$$

$$\lambda_2 \leq p_2 (\min_{j=1,2} (p_1 q_{2|\{1,2\}}^{(j)} + (1 - p_1) q_{2|2}^{(j)})). \tag{3.70}$$

The conditional expected utilities are given by $u_i(1, p_{-i}) = (\min_{j=1,2} (p_k q_{i|\{1,2\}}^{(j)} + (1 - p_k) q_{i|i}^{(j)}))(r_i + D_i) - E_i - D_i$ for $k \neq i$ and $u_i(0, p_{-i}) = -D_i$. The indifference condition of $u_i(1, p_{-i}) = u_i(0, p_{-i})$ for non-cooperative Nash equilibrium leads to the equality of $\min_{j=1,2} g_{i,j}(p_k) = \frac{E_i}{1+D_i}$, where $g_{i,j}(p_k) = p_k q_{i|\{1,2\}}^{(j)} + (1 - p_k) q_{i|i}^{(j)}$. So, the mixed strategies in Nash equilibrium are given by

82

$$p_1^* = \frac{q_{2|2}^{(j^*)} - \frac{E_2}{r_2+D_2}}{q_{2|2}^{(j^*)} - q_{2|1,2}^{(j^*)}}, \quad \text{for} \ \ j^* = \arg\min_{j=1,2} g_{2,j}(p_1^*), \tag{3.71}$$

$$p_2^* = \frac{q_{1|1}^{(j^*)} - \frac{E_1}{r_1+D_1}}{q_{1|1}^{(j^*)} - q_{1|1,2}^{(j^*)}}, \quad \text{for} \ \ j^* = \arg\min_{j=1,2} g_{1,j}(p_2^*). \tag{3.72}$$

The resulting throughput rate λ_i is $\frac{E_i}{r_i+D_i}p_i^*$ and the utility u_i is $-D_i$, $i = 1, 2$.

3.9 Repeated Random Access Games for Backlogged Transmissions

We consider a conflict resolution situation with a fixed number of n transmitters with backlogged packets waiting for retransmission (without additional packet arrivals until all backlogged packets are successfully transmitted). In this case, each packet corresponds to a new transmitter node and whenever the packet is successfully received, it leaves the system. We allow nodes to select their retransmission probabilities depending on the contention for the channel, namely the number of potential transmitters. For any node i, the cost of transmission is E_i and the reward for successful transmission is r_i. As a delay cost, we decrease the expected future payoff of any node i by a discount factor of δ_i over each time slot of unsuccessful transmission or waiting. In the presence of n backlogged nodes, we define $p_{i,n}$ as the transmission probability and $u_{i,n}$ as the expected utility of node i. We define T and W as the actions of transmitting and waiting with the corresponding expected utilities of $u_{i,n}(T)$ and $u_{i,n}(W)$ for node i, respectively. The transmission probabilities vary depending on the number of backlogged packets n. This approach is based on the perfect knowledge of n at each transmitter in any time slot and can be realized

by the immediate and correct receiver feedback over separate conflict-free channels.

When players face uncertainty (which occurs at every decision point of the game), they need to evaluate their uncertain future payoffs reflected in their utility functions and players attempt to maximize the expected values of their utilities. The expected utility is determined by randomizing the decisions between transmitting and waiting, and it is given by

$$u_{i,n} = p_{i,n} u_{i,n}(T) + (1 - p_{i,n}) u_{i,n}(W) \tag{3.73}$$

for node i. In random access game $G(n)$, node i selects in any time slot the probability of transmitting, $p_{i,n}$, to maximize the expected utility $u_{i,n}$ in the presence of n backlogged nodes. We consider a symmetric capture model and define q_k as the probability that any node i successfully transmits, when there are total of $k + 1$ nodes (including node i) simultaneously transmitting in the given slot. The expected utilities of node i are given by

$$u_{i,n}(T) = -E_i + \sum_{k=0}^{n-1} P(K_{i,n} = k)(r_i q_k + \delta_i(1 - q_k) \sum_{j=0}^{k} u_{i,n-j} \binom{k}{j} q_k{}^j (1 - q_k)^{k-j}), \tag{3.74}$$

$$u_{i,n}(W) = \delta_i \sum_{k=0}^{n-1} P(K_{i,n} = k) \sum_{j=0}^{k} u_{i,n-j} \binom{k}{j} q_{k-1}{}^j (1 - q_{k-1})^{k-j} \tag{3.75}$$

for transmitting and waiting, respectively, where the random variable $K_{i,n}$ denotes the number of transmitting nodes except node i. For tractable analysis, we consider symmetric games with identical users and common reward and cost parameters of $r_i = r$, $E_i = E$ and $\delta_i = \delta$. Nodes have the same possible actions, utilities, feedback and capture statistics such that they have the same retransmission strategies given the same state of the game. We have $p_{i,n} = p_n$, $u_{i,n} = u_n$ and $P(K_{i,n} = k) = P(K_n =$

$k) = \binom{n-1}{k} p_n^k (1 - p_n)^{n-1-k}$. The symmetric expected utilities $u_{i,n}(T) = u_n(T)$ and $u_{i,n}(W) = u_n(W)$ depend only on the values of p_n, u_k, q_k, $1 \leq k \leq n-1$.

3.9.1 Non-Cooperative Equilibrium Strategies

We consider the non-cooperative strategies in Nash equilibrium such that no node can improve the utility through individual effort, if the strategies of the other nodes remain the same. Random access can be described by multi-stage games of extensive form, which specify the complete order of moves in time in addition to the complete list of payoffs and the available information at each decision point in time. At each stage of the game, nodes follow a mixed strategy with the possible actions of transmitting and waiting. The equilibrium of extensive form games is characterized by a subgame perfect equilibrium, where nodes play the Nash equilibrium strategies at each possible subgame that can be reached over successive stages.

The decision at each stage of the random access game can be exactly described by a state (i.e. the number of backlogged nodes), which is a general property of the class of Markov games [17]. For Markov games, the subgame perfect equilibrium can be replaced by the Markov perfect equilibrium, in which each node's decision solely depends on the current state of the game. Given the state information, the Markov perfect equilibrium at each stage of the game can be computed as the Nash equilibrium of the mixed strategy game of normal form.

The symmetric Markov perfect equilibrium for the non-cooperative random access game $G(n)$ (in the presence of n backlogged nodes) can be computed as

the transmission probability p_n^* that makes the payoffs of transmission and waiting

equal to each other, i.e. $u_n(T) = u_n(W)$, provided that Markov perfect equilibrium

strategies are played for all possible future games $G(k)$, $1 \leq k < n$. Since the

best strategy in the absence of other backlogged nodes is to transmit, the initial

conditions are $u_1^* = rq_0 - E$ and $p_1^* = 1$ (provided that $rq_0 > E$). Starting with

$n = 2$, we set $u_n(T) = u_n(W) = u_n$, and recursively solve the equality of symmetric

Eqs. (3.74) and (3.75) to obtain the Markov perfect equilibrium strategy p_n^* and the

utility u_n^* for each game state $n \geq 2$. We consider a physical layer capture model

based on SINR criterion. A packet of node i is successfully received, if

$$\frac{P_i}{\frac{1}{L} \sum_{k \in S \setminus \{i\}} P_k + \sigma^2} \geq \gamma^*, \tag{3.76}$$

where P_i is the power level of node i, γ^* is the predetermined fixed SINR threshold,

σ^2 is the variance of Gaussian background noise, S denotes the set of transmitter

nodes, and L is the processing gain. We have $L = 1$ for narrowband systems,

and $L > 1$ for spread-spectrum CDMA systems. We assume a symmetric model

with $P_i = P$. In section 3.10, we will modify the game model by allowing nodes

to choose the probabilities of transmitting at different powers (and rates) from a

fixed set. The resulting successful transmission probability depends on the number

transmitters such that $q_k = 1$, for $0 \leq k \leq \left\lfloor \frac{L(P - \gamma^* \sigma^2)}{\gamma^* P} \right\rfloor$ and $q_k = 0$ otherwise.

3.9.2 Comparison of Selfish and Cooperative Strategies

The Markov perfect equilibrium strategies p_n^* for the non-cooperative random

access game $G(n)$ is computed by solving the non-linear equations of form $u_n(T)$

$= u_n(W)$ with the constrained solution space $p_n^* \in [0,1]$. Numerical solutions in the least squares sense can be found by minimizing the penalty function $|u_n(T) - u_n(W)|^2$. For numerical solutions, we can use the logarithmic barrier method to eliminate the constraints through Lagrangians and then apply the Newton's method to perform the necessary unconstrained optimization.

On the other hand, the cooperative random access game can be cast as a team problem of maximizing the total system utility $u_{\Sigma,n} = \sum_{i=1}^{n} u_{i,n}$ for each game state n. From (3.73)-(3.75), the recursive equation for $u_{\Sigma,n}$ is given by

$$u_{\Sigma,n} = -Enp_n + \sum_{k=0}^{n} \binom{n}{k} p_n^k (1-p_n)^{n-k} \sum_{j=0}^{k} \binom{k}{j} q_{k-1}^j (1-q_{k-1})^{k-j} (rj + \delta u_{\Sigma,n-j}) \quad (3.77)$$

with the initial conditions of $u_{\Sigma,1} = rq_0 - E$ and $p_1 = 1$.
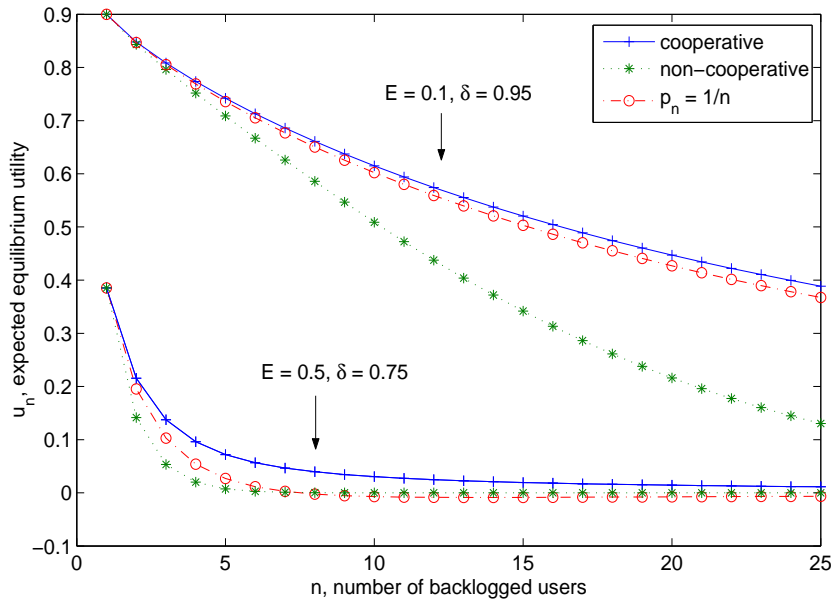
Figure 3.14: Performance comparison of the cooperative and non-cooperative random access strategies with the retransmission probability assignment of $p_n = 1/n$.

For numerical results, we assume $L = 5$, $\gamma^* = 0.5$ and $\sigma^2 = 0.1P$. Figure 3.14 compares the performance of the cooperative and non-cooperative random access games with the probability assignments of $p_n = \frac{1}{n}$, which is known to be the optimal strategy for classical collision channel without packet captures, if the throughput is the only performance measure. We first assume cost parameters of $E = 0.1$ and $\delta = 0.95$. The retransmission strategy $p_n = \frac{1}{n}$ achieves performance results close to the cooperative operation of nodes, and clearly outperforms the non-cooperative equilibrium retransmission strategies. If we impose strict cost parameters $\delta = 0.75$ and $E = 0.5$, the performance difference between the cooperative and non-cooperative strategies decreases, as illustrated in Figure 3.14. For large number of backlogged nodes, selfish operation achieves higher expected utilities than dictating nodes to use $\frac{1}{n}$ as the probability of retransmitting, whereas the cooperation among nodes leads to strictly higher utilities than the Nash equilibrium strategies.

3.10   Randomized Power and Rate Control Game

We allow nodes to employ distributed power and rate control schemes to adjust their capture probabilities in a stochastic game for random access. A cooperative scheme for selecting random power levels has been introduced in [37, 38], where nodes randomly transmit at discrete power levels to maximize the channel utilization in each time slot without taking the transmission and delay costs into account. The power and rate control problems have been analyzed before for non-cooperative cellular access [39, 50]. However, these studies are based on one-stage pure strategy

games and cannot capture the essence of the repeated random access operation. Instead, we allow selfish nodes to randomize their power and rate control strategies (over a discrete and finite strategy space) over successive time slots.

We consider the same feedback information and game state as before. In the joint power and rate control game $PRG(n)$, each node randomly chooses transmission power from the set $\mathcal{P}$ and transmission rate from the set $\mathcal{R}$. The choice of zero transmission power or rate is equivalent to the action of waiting. We define $p_{i,n}(S_j)$ as the probability that node i chooses power and rate pair $S_j \in J$, where $J = \mathcal{P}\mathrm{x}\mathcal{R}$. As a result, the overall expected utility is given by

$$u_{i,n} = \sum_{S_j \in J} u_{i,n}(S_j) p_{i,n}(S_j), \tag{3.78}$$

where $u_{i,n}(S_j)$ is the utility of node i that chooses action S_j for the game state n. Each node i selects the probability $p_{i,n}(S_j)$ for all $S_j = (P_j, R_j)$, where $P_j \in \mathcal{P}$ and $R_j \in \mathcal{R}$, to maximize the expected utility $u_{i,n}$ in the presence of n backlogged nodes. For the game $PRG(n)$, we define $K_{j,n}$ as the number of nodes that select the transmission power and rate pair $S_j \in J$ given game state n, and we define the capture probability $q_n(S_j | \{K_{j,n} = k_j, S_j \in J\})$ as the probability that any node with action S_j is successful, if $K_{j,n} = k_j$ nodes employ each action $S_j \in J$.

A node selects any transmission rate $R_j \in \mathcal{R}$ (bits per time slot) by employing a different adaptive modulation scheme. Specifically, we assume that each node employs m-ary Quadrature Amplitude Modulation (QAM) (with constellation size of $m = 2^{R_j}$) or chooses to wait ($m = 1$). The probability of successful packet reception of a node depends on the SINR value achieved at the receiver and on

the modulation scheme (i.e. transmission rate). We consider a symmetric system with identical nodes and denote by $f_m(\gamma)$ the probability that a node with action $S_j = (P_j, R_j)$, where $R_j = \log_2 m$, satisfies the SINR condition of $\gamma \geq \gamma^*$. We use the same utility functions of (3.74) and (3.75) as before. For any node with action S_j, the energy cost is changed from E to EP_j, if $R_j > 0$, and the throughput reward is changed from r to rR_j, whereas the capture probability is redefined as

$$q_n(S_j|\{K_{j,n} = k_j, S_j \in J\}) \tag{3.79}$$

$$= f_m \left(\frac{P_j \, 1(R_j > 0)}{\frac{1}{L}((k_j - 1)P_j 1(R_j > 0) + \sum_{i \neq j}^{|J|} k_i E_i \, 1(R_i > 0) + \sigma^2} \right),$$

where $1(\cdot)$ denotes the indicator function, which assigns the value 1 to a true statement and the value 0 to a false statement. The Markov perfect equilibrium in $PRG(n)$ game corresponds to the power and rate control strategies $P(x, y)$, for all $(x, y) \in J$, that yield equal expected utilities. To find the symmetric equilibrium strategies, we need to equate the expected utilities of all joint power and rate strategies. For numerical results, we use MATLAB to solve the resulting equalities.

The power levels (measured in unit power) are chosen from the set $\mathcal{P} = \{0, 1, 2, 3, 4, 5\}$ and the ambient noise power is 0.1 unit power. The transmission rates (measured in number of bits per time slot) are selected from the set $\mathcal{R} = \{0, 1, 2, 3, 4\}$. The cost parameters are $E = 0.1$ and $\delta = 0.95$. For comparison, we also consider the non-cooperative game of selecting retransmission probabilities for fixed and common transmission powers and rates, and modify the throughput rewards and costs accordingly. We then evaluate the performance of the original random access game $G(n)$ by averaging or optimizing the expected utilities over the
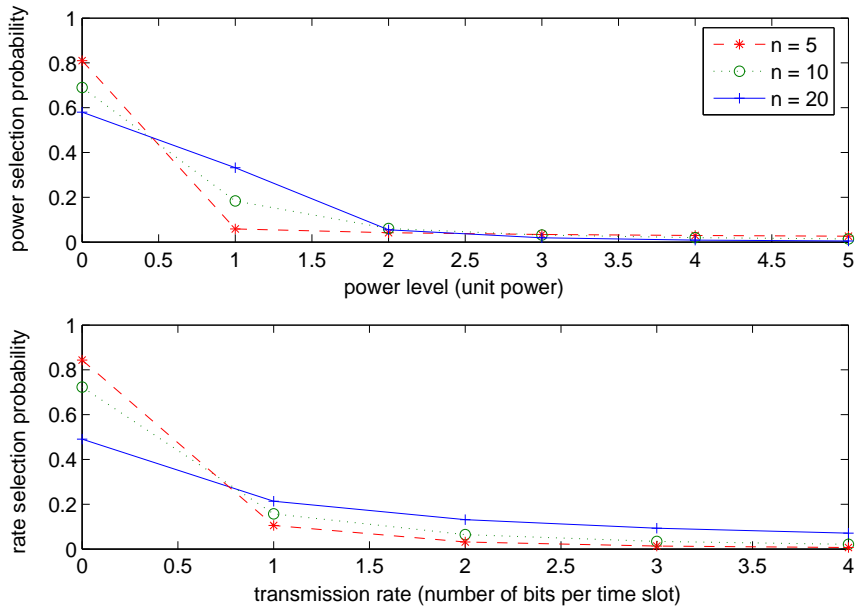
Figure 3.15: Non-cooperative equilibrium strategies for power and rate control game.

possible transmission power and rate levels common for each node.

The strategies and the corresponding expected utilities of the joint power and rate control game $PRG(n)$ are shown in Figures 3.15 and 3.16, respectively. We also include in Figure 3.16 the expected utilities obtained in game $G(n)$ that are averaged or optimized over the common assignments of power and rate levels. Numerical results indicate that the power and rate control game $PRG(n)$ improves the non-cooperative performance over the plain random access game $G(n)$.

3.11 Summary and Conclusions

We addressed the problem of non-cooperative random access with selfish and malicious transmitters. First, we modeled selfish nodes as individual decision mak-
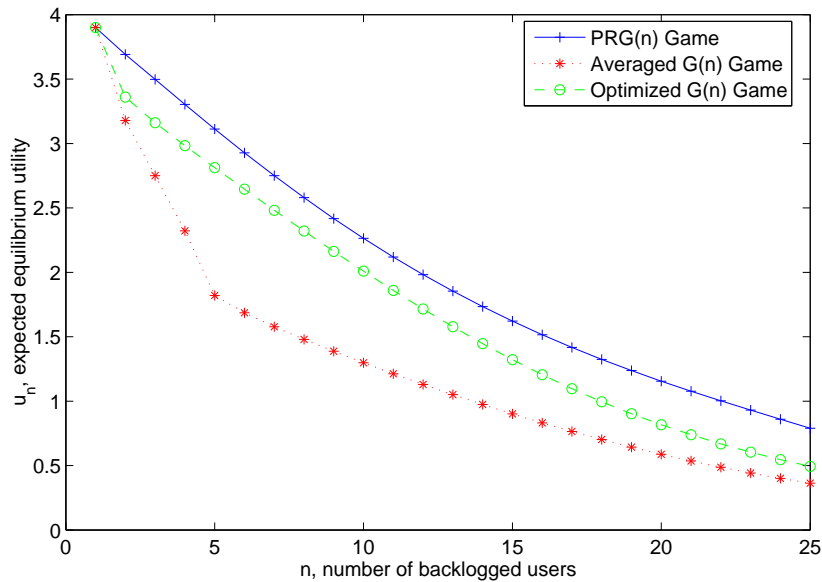
Figure 3.16: Non-cooperative equilibrium utilities for the random access game $G(n)$ and joint power and rate control game $PRG(n)$.

ers with throughput, energy and delay objectives, and derived the non-cooperative transmission probabilities in Nash Equilibrium. Then, we evaluated the effects of malicious nodes that have the additional objectives of blocking the transmissions of the other nodes. A distributed implementation followed from the adaptive best-response update algorithms. We also illustrated the performance loss compared to the cooperative strategies that maximize the weighted utility sum. In addition, a linear pricing scheme was introduced to improve the non-cooperative equilibrium and distributed implementation was outlined. We also formulated a repeated random access game of selecting retransmission probabilities depending on the contention for the channel. Finally, we extended the model to a power and rate control game.

The game model is based on the perfect knowledge of the number of packets

contending for the channel at each time slot. This is realized by immediate and correct feedback from the receiver. Future work needs to look at the effects of partial or erroneous feedback on the non-cooperative MAC operation. We will extend the model to power-controlled MAC in Chapter 4. The ultimate goal is to envision a passage to the multi-hop operation that would extend the strategy space and require more complicated utility functions [40]. We will formulate a cross-layer framework to represent the non-cooperative network operation through the joint consideration of MAC with plain routing in Chapter 5 and with network coding in Chapter 7.

Chapter 4

Power-Controlled MAC Games for Non-Cooperative Wireless Access

We extend the non-cooperative operation to power-controlled MAC with SINR-based channel model. We formulate a non-cooperative game of selecting the individual transmission powers and evaluate the interactions between selfish and malicious transmitters at the MAC layer. We consider the performance objectives of optimizing the throughput rewards and energy costs reflected in the node utilities subject to power constraints. We show the existence and uniqueness of the non-cooperative Nash equilibrium strategies for two different classes of utility functions that represent either the linear combination or ratio of successful transmission rate and power consumption. We also discuss the social equilibrium strategies that maximize the total system utility. For distributed operation, we introduce adaptive best-response update mechanisms that converge to the non-cooperative equilibrium. Finally, we consider an alternative successful packet reception criterion based on fixed SINR threshold and show that the existence and uniqueness of Nash equilibrium strongly depend on the system parameters.

4.1 Introduction

The problem of power-controlled MAC has been extensively studied before [41, 42, 43] for multiple nodes simultaneously transmitting to a single receiver. The

main objective of distributed power control is to minimize the power consumption while achieving the target SINR at the receiver for each transmitter. Power control can be formulated as a non-cooperative game between selfish transmitters [44, 45, 46] and can be combined with random access or rate control in a cross-layer design, as outlined in Chapter 3.

Any transmitter may also have the malicious objective of reducing the transmission rate or increasing the power consumption of the other selfish transmitter nodes while minimizing the individual power consumption. The interactions of selfish and malicious transmitters have been evaluated in Chapter 3 for random access [25]. For Gaussian multi-access channel, the problem of fair rate allocation has been studied in [47] as a cooperative game of selfish nodes that can form coalition to threaten each other with jamming the channel. In this chapter, we assume a SINR-based successful packet reception criterion and formulate a non-cooperative power control game between selfish and malicious transmitters with performance objectives of throughput and power efficiency. We consider the problem of maximizing two different utility functions that represent the ratio or linear combination of the successful packet transmission rate and power consumption subject to power constraints. We evaluate the unique non-cooperative strategies of transmission powers in Nash equilibrium such that no node can unilaterally change the strategy to improve the individual performance. We also consider another successful packet reception criterion such that a packet is successfully received if the transmitter's signal power exceeds a fixed SINR threshold. In that case, we show that the Nash equilibrium solutions may not be unique or may not exist depending on the throughput

reward and power cost parameters.

For distributed operation, we introduce adaptive update mechanisms such that each node repeatedly plays the best-response strategy. We also discuss the social equilibrium strategies that maximize the weighted sum of the individual node utilities in the system. Our objective is to provide insights for the optimal strategies of jamming the channels of the selfish transmitter nodes as well as for the optimal defense mechanisms against the denial of service attacks of malicious transmitter nodes in power-controlled MAC operation.

The chapter is organized as follows. In section 4.2, we formulate the non-cooperative power control game and introduce throughput rewards and energy costs. We discuss two classes of utility functions in sections 4.3 and 4.4, and derive the on-cooperative strategies in Nash equilibrium for two transmitters. We introduce the best response updates in section 4.5 for distributed implementation. Then, we discuss in section 4.6 the social equilibrium strategies that optimize the weighted sum utility. The results are extended to the arbitrary number of selfish and malicious transmitters in section 4.7. We discuss in section 4.8 an alternative measure for the transmission rates based on a fixed SINR threshold, and evaluate the existence and uniqueness of Nash equilibrium strategies. We collect final remarks in section 4.9.

4.2   System Model, Rewards and Costs

Consider set N of (selfish and malicious) transmitters and one common receiver. Each node i transmits one packet in each time slot with power P_i. We

assume saturated queues and continuous packet transmissions. The channel gain from transmitter i to the receiver is h_i and the additive noise power is σ^2. We consider a SINR-based channel model such that the successful transmission probability of any selfish node i is expressed as the throughput reward function $f_i(\gamma_i)$, where $\gamma_i = \frac{h_i P_i}{\sum_{k \neq i} h_k P_k + \sigma^2}$ is the SINR value achieved at the receiver. We focus on a simple receiver model based on matched filter. However, the results can be extended to any linear receiver. The throughput reward function f_i depends on modulation, coding and packet size. We assume that f_i is a continuous, differentiable and increasing function of γ_i, $f_i(0) = 1$, $f_i(\infty) = 0$ and f_i has a sigmoidal shape such that there exists γ_i^* below which f_i is convex and above which f_i is concave. The desirable properties of functions to represent the transmission rates have been discussed in [45]. As an example, we have $f_i(\gamma_i) = (1 - e^{-\gamma_i})^K$ for DPSK modulation with K bits in each packet. We will relax some of these assumptions in section 4.8, and consider an alternative function f_i to show that the Nash equilibrium strategies do not need to exist or to be unique.

Let M_i denote the set of selfish transmitters attacked by malicious transmitter i. Any selfish or malicious transmitter i has the objective of maximizing $f_i(\gamma_i)$ or $f_i(\gamma_j, j \in M_i) = \sum_{j \in M_i} (1 - f_j(\gamma_j))$, respectively. Also, any node i needs to minimize the power consumption P_i. We consider $\mathcal{P}_i = [P_i^{\min}, P_i^{\max}]$ as the constraint set for the transmission power P_i of any node i. The objectives of non-cooperative transmitters can be extended to include the delay-type costs [48]. Let D_i denote the average number of time slots necessary to deliver the packet of node i under transmission to the receiver. For saturated queues, D_i has an

exponential distribution such that $\Pr(D_i = d) = f_i(\gamma_i)(1 - f_i(\gamma_i))^{d-1}$. The delay constraint $\Pr(D_i \leq D_{\min}) \geq \beta_{\min}$ for node i can be written as a lower bound $\gamma_i > \gamma_i^{\min}$, where $\gamma_i^{\min} = f_i^{-1}\left(1 - (1 - \beta_{\min})^{\frac{1}{D_{\min}}}\right)$. Any malicious node may have the objective of imposing $\Pr(D_i \leq D_{\max}) \leq \beta_{\max}$ for selfish transmitter i such that $\gamma_i^{\max} = f_i^{-1}\left(1 - (1 - \beta_{\max})^{\frac{1}{D_{\max}}}\right)$. We will continue with the power constraints only, although the results can be easily adapted to include SINR constraints [49]. Let $\underline{P}$ denote the set of transmission powers. The non-cooperative optimization problem for transmitter i is given by

$$\max_{P_i \in \mathcal{P}_i, i \in N} u_i(\underline{P}) . \tag{4.1}$$

Our objective is to find the non-cooperative Nash equilibrium strategies $P_i^* \in \mathcal{P}_i$, $i \in N$, such that no node can unilaterally improve individual performance, i.e.

$$u_i(\underline{P}^*) \geq u_i(P_i, \underline{P}_{-i}^*) \quad \text{for any } P_i \in \mathcal{P}_i , \tag{4.2}$$

where $\underline{P}_{-i}$ is the set of transmission powers of nodes except i.

4.3   Utility Function Type 1

Each node i has the objective of choosing $P_i \in \mathcal{P}_i$ to maximize the utility function u_i. Consider the utility function u_i as the ratio of $f_i(\gamma_i)$ and P_i such that

$$u_i(\underline{P}) = \frac{f_i(\gamma_i)}{P_i} . \tag{4.3}$$

4.3.1   Two Selfish Transmitters

Assume that each transmitter $i = 1, 2$ is selfish with the objective of optimizing utility function $u_i(P_1, P_2) = \frac{f_i(\gamma_i)}{P_i}$. The problem of maximizing $u_i(P_1, P_2)$ over $P_i \in$

$\mathcal{P}_i$ has the unique solution given by $\max(\min(P_i^*, P_i^{\max}), P_i^{\min})$ such that P_i^* satisfies

$\frac{\partial u_i(P_1, P_2)}{\partial P_i} = 0$, since $\frac{f_i(\gamma_i}{P_i}$ is a quasi-concave function of P_i (i.e. there exists a value

of P_i below which $\frac{f_i(\gamma_i)}{P_i}$ is non-decreasing and above which $\frac{f_i(\gamma_i)}{P_i}$ is non-increasing).

We assume that the receiver uses matched filter such that $\gamma_i = \frac{h_i p_i}{h_j p_j + \sigma^2}$ for $j \neq i$.

Since $\frac{\partial \gamma_i}{\partial P_i} = \frac{\gamma_i}{P_i}$ for any linear receiver, the condition $\frac{\partial u_i(P_1, P_2)}{\partial P_i} = 0$ is equivalent to

$f_i'(\gamma_i)\gamma_i = f_i(\gamma_i)$, where f_i' is the partial derivative of f_i with respect to γ_i. Since

$u_i(P_1, P_2)$ is a quasi-concave in P_i for any given transmission power P_{-i}, there exists

a unique equivalent SINR solution $\gamma_i^* > 0$ that satisfies $f_i'(\gamma_i^*)\gamma_i^* = f_i(\gamma_i^*)$. For

$i = 1, 2$, we can solve $\gamma_i^* = \frac{h_i P_i^*}{h_j P_j^* + \sigma^2}$, $j \neq i$, for P_i^*, $i = 1, 2$. As a result, there exist

unique non-cooperative Nash equilibrium strategies given by

$$P_i^* = \max\left(\min\left(\frac{\sigma^2}{h_i}\gamma_i^* \frac{1 + \gamma_j^*}{1 - \gamma_1^*\gamma_2^*}, P_i^{\max}\right), P_i^{\min}\right), \quad (4.4)$$

where $f_i'(\gamma_i^*)\gamma_i^* = f_i(\gamma_i^*)$, $i = 1, 2$. For different values of γ_i^*, $i = 1, 2$, we evaluate in

Table 4.1 the non-cooperative Nash equilibrium strategies P_i^*, $i = 1, 2$.

Table 4.1: The Nash Equilibrium Strategies for Two Selfish Transmitters.

$\gamma_1^* = 0$	$P_1^* = P_1^{\min}$, $P_2^* = \max\left(\min\left(\frac{\sigma^2}{h_2}\gamma_2^*, P_2^{\max}\right), P_2^{\min}\right)$
$\gamma_2^* = 0$	$P_1^* = \max\left(\min\left(\frac{\sigma^2}{h_1}\gamma_1^*, P_1^{\max}\right), P_1^{\min}\right)$, $P_2^* = P_2^{\min}$
$\gamma_1^*\gamma_2^* = 0$	$P_1^* = P_1^{\max}$, $P_2^* = P_2^{\max}$
$\gamma_1^*\gamma_2^* > 1$	$P_1^* = P_1^{\min}$, $P_2^* = P_2^{\min}$

The equilibrium SINR values γ_i^*, $i = 1, 2$, follow from the equations $f_i'(\gamma_i)\gamma_i = f_i(\gamma_i)$, $i = 1, 2$, and are independent of the equilibrium transmissions powers P_i^*,

$i = 1, 2$, which, however, strongly depend on the equilibrium SINR values, as Table

4.1 illustrates. If the equilibrium SINR $\gamma_i = 0$ of any selfish transmitter i diminishes to 0, then it transmits with the minimum power $P_i^{\min}$ to minimize the power cost, while the other selfish node needs to combat the background noise only to achieve the individual equilibrium SINR γ_k^*, $k \neq i$, and transmits with the minimum power $\frac{\sigma^k}{h_k}\gamma_k^*$, provided that the equilibrium transmission power satisfies the power constraints.

As the equilibrium SINR values increase, nodes start increasing their transmission powers to combat the noise and additional interference from the other selfish node. The equilibrium transmission powers increase up to the maximum levels, as $\gamma_1^*\gamma_2^*$ approaches the value 1.

If the equilibrium SINR values are high enough such that $\gamma_1^*\gamma_2^* > 1$, then both nodes achieve sufficiently high throughput rewards such that their only objective is to minimize the transmission power cost such that both nodes transmit with the minimum values of their transmission powers.

4.3.2   One Selfish and One Malicious Transmitter

Assume that transmitter 1 is selfish with utility function $u_1(P_1, P_2) = \frac{f_1(\gamma_1)}{P_1}$ and transmitter 2 is malicious with utility function $u_2(P_1, P_2) = \frac{1 - f_1(\gamma_1)}{P_2}$. The problem of maximizing $u_1(P_1, P_2)$ over $P_1 \in \mathcal{P}_1$ has the unique solution given by $\max(\min(P_1^*, P_1^{\max}), P_1^{\min})$ such that P_1^* satisfies $\frac{\partial u_1(P_1, P_2)}{\partial P_1} = 0$ for any P_2, since $\frac{f_1(\gamma_1)}{P_1}$ is a quasi-concave function of P_1. The problem of maximizing $u_2(P_1, P_2)$ over $P_2 \in \mathcal{P}_2$ has the unique solution given by $P_2^* = \max(\min(\overline{P}_2, P_2^{\max}), P_2^{\min})$ such that $\overline{P}_2$ satisfies $\frac{\partial u_2(P_1, P_2)}{\partial P_2} = 0$ for any P_1. The condition $\frac{\partial u_1(P_1, P_2)}{\partial P_2} = 0$ is equivalent to

$f_1'(\gamma_1)\gamma_1 \frac{h_2 P_2}{h_2 P_2 + \sigma^2} = 1 - f_1(\gamma_1)$, since $\frac{\partial \gamma_1}{\partial P_2} = -\frac{\gamma_1 h_2}{h_2 P_2 + \sigma^2}$ for the matched filter receiver.

If $u_1(P_1, P_2)$ is quasi-concave in P_1, $u_2(P_1, P_2)$ is quasi-concave in P_2, i.e. there exist unique non-cooperative Nash equilibrium solutions $P_i^* = \max(\min(\overline{P}_i, P_i^{\max}), P_i^{\min})$, $i = 1, 2$, such that $\overline{P}_1 = \frac{\sigma^2}{h_1}\gamma_1^* \frac{f_1(\gamma_1^*)}{2f_1(\gamma_1^*)-1}$ and $\overline{P}_2 = \frac{\sigma^2}{h_2} \frac{1-f_1(\gamma_1^*)}{2f_1(\gamma_1^*)-1}$, where the resulting SINR value is the unique solution to $f_1'(\gamma_1^*)\gamma_1^* = f_1(\gamma_1^*)$. In summary, the unique non-cooperative Nash equilibrium solutions are given by

$$P_1^* = \max\left(\min\left(\frac{\sigma^2}{h_1}\gamma_1^* \frac{f_1(\gamma_1^*)}{(2f_1(\gamma_1^*)-1)}, P_1^{\max}\right), P_1^{\min}\right), \qquad (4.5)$$

$$P_2^* = \max\left(\min\left(\frac{\sigma^2}{h_2} \frac{(1-f_1(\gamma_1^*))}{(2f_1(\gamma_1^*)-1)}, P_2^{\max}\right), P_2^{\min}\right), \qquad (4.6)$$

where $f_1'(\gamma_1^*)\gamma_1^* = f_i(\gamma_1^*)$. The equilibrium transmission powers depend on the throughput reward function evaluated at the equilibrium SINR γ_1^* of selfish transmitter 1, as Table 4.1 indicates.

For $f_1(\gamma_1^*) < 0.5$, the throughput reward of selfish node 1 is small enough such that node 1 transmits only with the minimum power to minimize the power cost and consequently individual utility. In the mean time, any further reduction of the throughput reward of selfish node 1 cannot increase the utility of malicious node 2 and therefore node 2 also needs to transmit with the minimum power only. For $f_1(\gamma_1^*) > 0.5$, selfish node 1 needs to increase transmission power to increase the equilibrium SINR γ_1^* and consequently throughput reward $f(\gamma_1^*)$. In the mean time, malicious node 2 also increases transmission power to decrease the throughput reward of node 1. If the equilibrium SINR maximizes $f_1(\gamma_1^*)$ to 1, then malicious node 2 needs to transmit with the minimum power level to minimize the power cost without any effect on the utility of selfish node 1 and node 1 only needs to combat

the background noise to achieve the equilibrium SINR γ_1^*.

We compare three different cases of (a) single transmitter 1 (stand-alone), (b) selfish transmitter 1 and selfish transmitter 2, and (c) selfish transmitter 1 and malicious transmitter 2. The value of P_1^* in case (b) is greater than or equal to the value of P_1^* in case (c), as the comparison of Eqs. (4.4) and (4.6) reveals. For different values of $f_i(\gamma_i^*)$, $i = 1, 2$, we evaluate in Table 4.2 the non-cooperative Nash equilibrium strategies P_i^*, $i = 1, 2$.

Table 4.2: The Nash Equilibrium Strategies for One Selfish and One Malicious Transmitter.

$f_1(\gamma_1^*) = 0$	$P_1^* = P_1^{\min}, \ P_2^* = P_2^{\min}$
$f_1(\gamma_1^*) = 0.5$	$P_1^* = P_1^{\max}, \ P_2^* = P_2^{\max}$
$f_1(\gamma_1^*) = 1$	$P_1^* = \max\left(\min\left(\frac{\sigma^2}{h_1}\gamma_1^*, P_1^{\max}\right), P_1^{\min}\right) \ , \ P_2^* = P_2^{\min}$

Introduction of selfish transmitter 2 does not affect the SINR values of transmitter 1. Transmitter 2 (selfish or malicious) in cases (b) and (c) increases power cost P_1 and reduces utility u_1 compared to case (a). Malicious operation of transmitter 2 in case (c) increases the power cost P_1 compared to selfish operation in case (b), if and only if

$$f_1(\gamma_1^*) < \frac{1 + \gamma_2^*}{1 + \gamma_2^*(2 + \gamma_1^*)}. \tag{4.7}$$

For numerical results, we assume $h_i = 1$, $P_i^{\min} = 0$, $P_i^{\max} = \infty$, $i = 1, 2$, $\sigma^2 = 0.1$ and $f_1(\gamma_1) = (1 - e^{-\gamma_1})^K$, which denotes the success rate, if nodes use DPSK (Differentiable Phase Shift Keying) modulation, where K represents the number of

bits in each packet. We compare two different cases before and after malicious transmitter 2 attacks selfish transmitter 1. Figure 4.1 depicts the transmission powers P_1 and P_2 as function of K. The SINR value achievable by transmitter 1 is shown in Figure 4.2 as function of K.
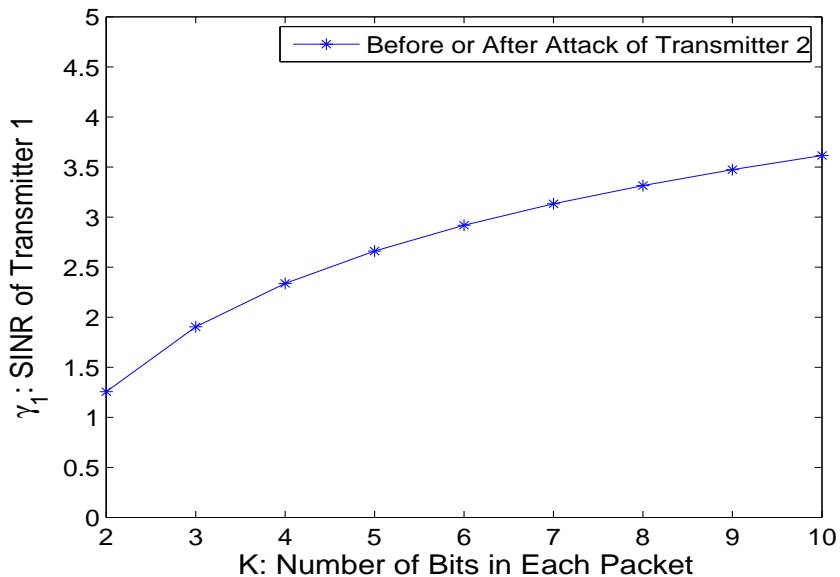
Figure 4.1: SINR value γ_1 of selfish transmitter 1 for utility function type 1.

The individual utilities u_1 and u_2 are depicted in Figure 4.3 as function of K. Malicious operation of transmitter 2 increases power cost P_1 and reduces utility u_1 of transmitter 1 (but cannot change γ_1).

4.4 Utility Function Type 2

As considered in [50] and [51] for selfish transmitters, we redefine the utility function u_i of node i as the linear combination of $f_i(\gamma_i)$ and P_i with non-negative
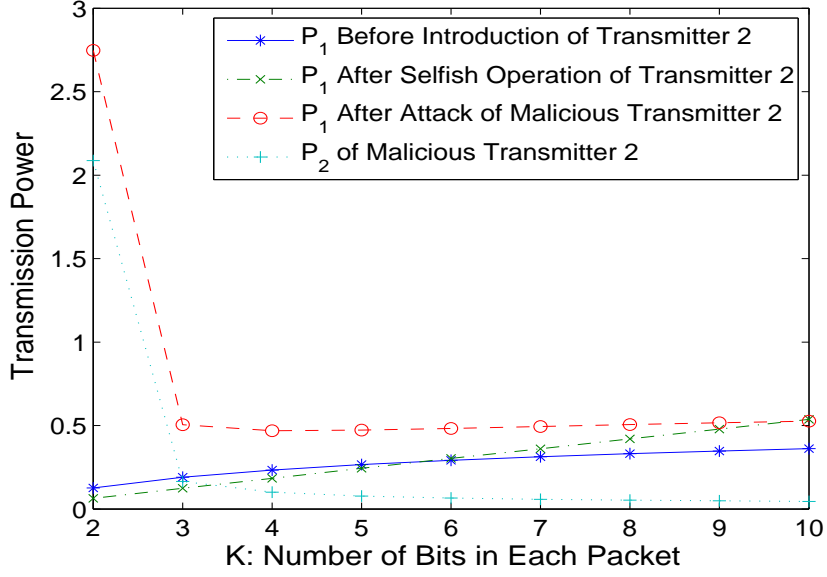
Figure 4.2: Transmission powers P_1 and P_2 for utility function type 1.

constant α_i such that

$$u_i(\underline{P}) = f_i(\gamma_i) - \alpha_i p_i . \tag{4.8}$$

4.4.1 Two Selfish Transmitters

Assume that each transmitter $i = 1, 2$ is selfish with utility function $u_i(P_1, P_2) = f_i(\gamma_i) - \alpha_i P_i$. For transmitter $i = 1, 2$, the problem of maximizing $u_i(P_1, P_2)$ over $P_i \in \mathcal{P}_i$ has the unique solution given by $P_i^* = \max(\min(\overline{P}_i, P_i^{\max}), P_i^{\min})$ such that $\overline{P}_i$ satisfies $\frac{\partial u_i(P_1, P_2)}{\partial P_i} = 0$ for any given transmission power level P_{-i} of the node other than i. The condition $\frac{\partial u_i(P_1, P_2)}{\partial P_i} = 0$ is equivalent to $f_i'(\gamma_i)\gamma_i = \alpha_i P_i$, $i = 1, 2$, such that for $\gamma_i^* = \frac{h_i P_i^*}{h_j P_j^* + \sigma^2}$ for $j \neq i$ there exist unique non-cooperative Nash equilibrium solutions given by

$$P_i^* = \max \left(\min \left(f_i'\left(\gamma_i^* \right) \frac{\gamma_i^*}{\alpha_i}, P_i^{\max} \right), P_i^{\min} \right) . \tag{4.9}$$

104
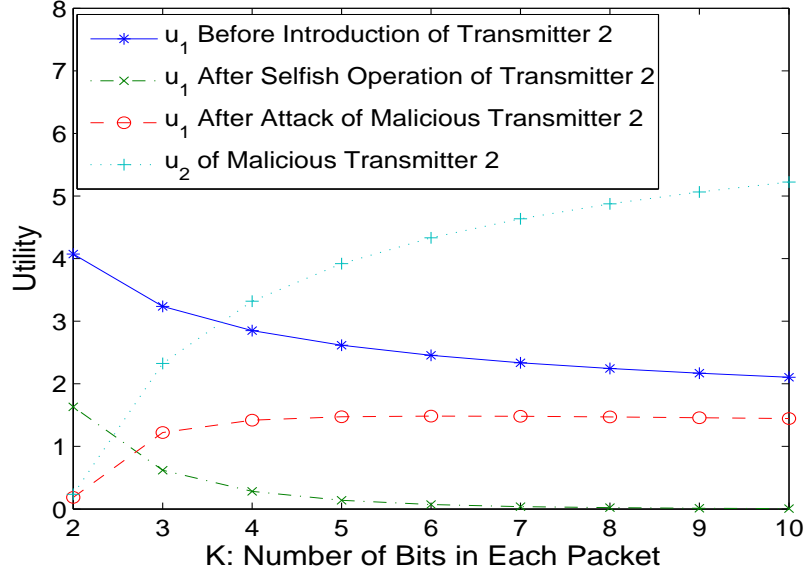
Figure 4.3: Utilities u_1 and u_2 for utility function type 1.

4.4.2 One Selfish and One Malicious Transmitter

Assume that transmitter 1 is selfish with utility function $u_1(P_1, P_2) = f_1(\gamma_1) - \alpha_1 P_1$ and transmitter 2 is malicious with utility function $u_2(P_1, P_2) = 1 - f_1(\gamma_1) - \alpha_2 P_2$. The condition $\frac{\partial u_1(P_1,P_2)}{\partial P_1} = 0$ is equivalent to $f_1'(\gamma_1)\gamma_1 = \alpha_1 P_1$. By using $f_2(\gamma_1) = 1 - f_1(\gamma_1)$, the condition $\frac{\partial u_2(P_1,P_2)}{\partial P_2} = 0$ is equivalent to $f_1'(\gamma_1)\gamma_1 = \frac{\alpha_2}{h_2}(h_2 P_2 + \sigma^2)$. If we combine the conditions $\frac{\partial u_i(P_1,P_2)}{\partial P_i} = 0$, $i = 1, 2$, we obtain the unique SINR value $\gamma_1^* = \max\left(\min\left(\frac{h_1 \alpha_2}{h_2 \alpha_1}, \gamma_1^{\max}\right), \gamma_1^{\min}\right)$ of transmitter 1. Therefore, there exist unique Nash equilibrium solutions given by

$$P_1^* = \max\left(\min\left(f_1'(\gamma_1^*)\frac{\gamma_1^*}{\alpha_1}, P_1^{\max}\right), P_1^{\min}\right), \qquad (4.10)$$

$$P_2^* = \max\left(\min\left(f_1'(\gamma_1^*)\frac{\gamma_1^*}{\alpha_2} - \frac{\sigma^2}{h_2}, P_2^{\max}\right), P_2^{\min}\right), \qquad (4.11)$$

where $\gamma_1^* = \frac{h_1 \alpha_2}{h_2 \alpha_1}$. For numerical results, we assume $h_i = 1$, $\alpha_i = 1$, $P_i^{\min} = 0$, $P_i^{\max} = \infty$, $i = 1, 2$, $\sigma^2 = 0.1$ and $f_1(\gamma_1) = (1 - e^{-\gamma_1})^K$. We compare two different

cases before and after malicious transmitter 2 attacks selfish transmitter 1. Figure 4.4 depicts the transmission powers P_1 and P_2 as function of K. The SINR value achievable by transmitter 1 is shown in Figure 4.5 as function of K. The individual utilities u_1 and u_2 are depicted in Figure 4.6. Malicious operation of transmitter 2 reduces SINR value γ_1 and utility u_1 of transmitter 1 (but may or may not increase power cost P_1 depending on function f_1).
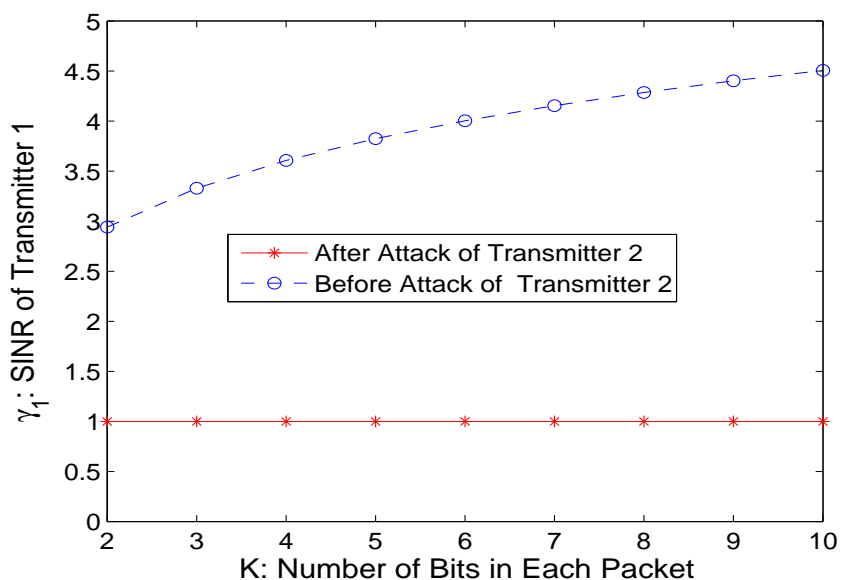
Figure 4.4: SINR value γ_1 of selfish transmitter 1 for utility function type 2.

4.5 Best Response Update Mechanisms for Distributed Operation

We consider one selfish and one malicious transmitter, and assume that each node explicitly knows SINR value γ_1. We consider the adaptive best-response update mechanisms such that at any iteration time $k+1$ node $i = 1, 2$ chooses

$$P_i(k+1) = \arg \max_{P_i \in \mathcal{P}_i, i=1,2} u_i(P_i, P_j(k)), \ \ j \neq i. \tag{4.12}$$
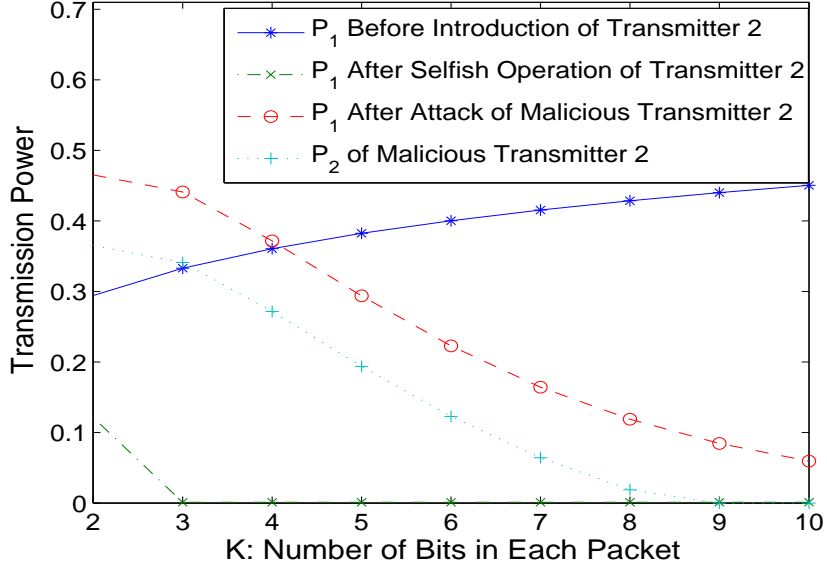
Figure 4.5: Transmission powers P_1 and P_2 for utility function type 2.

An iterative power control algorithm $\underline{P}(k + 1) = A(\underline{P}(k))$, where $\underline{P}(k) = [P_1(k), P_2(k)]$, is said to be standard [43], if matrix A satisfies $A(\underline{P}) > 0$, $A(\underline{P}') \geq A(\underline{P})$, if $\underline{P}' \geq \underline{P}$, and $\mu A(\underline{P}) > A(\mu \underline{P})$ for $\mu > 1$. The best-response updates given by Eq. (4.12) result in standard power control algorithm and converge to the non-cooperative Nash equilibrium solutions.

4.6   Comments on Social Equilibrium

We consider the problem of jointly maximizing the weighted utility sum, i.e.

$$\max_{P_i \in \mathcal{P}_i, \, i=1,2} \sum_{i=1,2} \beta_i \, u_i(P_1, P_2), \tag{4.13}$$

where β_i is the weight assigned to the utility u_i of node i. We assume one selfish and one malicious transmitter, and consider utility function type 2 (although similar results hold for utility function type 1). We use the Lagrangian multiplier method

107
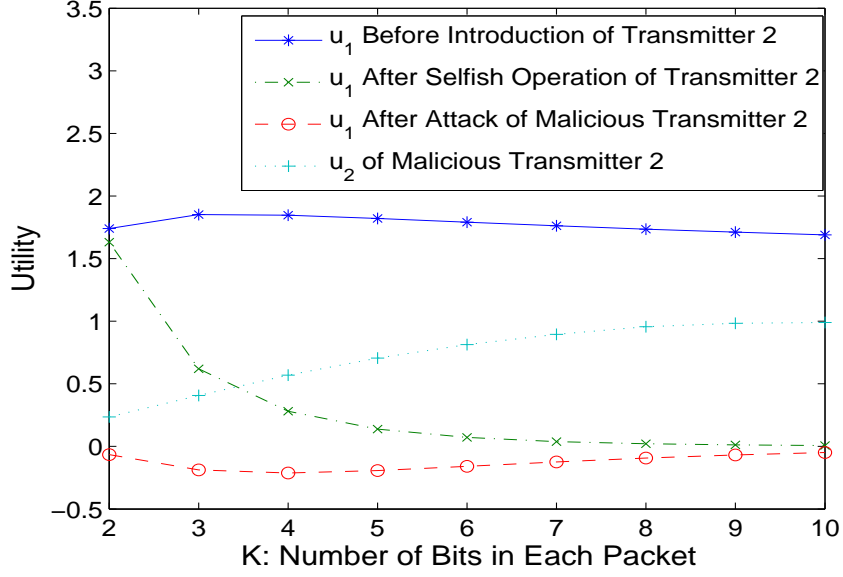
Figure 4.6: Utilities u_1 and u_2 for utility function type 2.

such that the objective is to maximize

$$L(P_1, P_2) = \sum_{i=1}^{2} (\beta_i u_i(P_1, P_2) - \lambda_{1,i}(P_i - P_i^{\max}) - \lambda_{2,i}(-P_i + P_i^{\min})) \quad (4.14)$$

over $P_i \in \mathcal{P}_i = [P_i^{\min}, P_i^{\max}]$, $i = 1, 2$, for non-negative constants $\lambda_{i,j}$, $i = 1, 2$ and

$j = 1, 2$. The KKT conditions for the optimal solutions are given by

$$\frac{\partial L(P_1, P_2)}{\partial P_i} = 0, \quad (4.15)$$

$$P_i \leq P_i^{\max}, \lambda_{1,i} > 0, \lambda_{1,i}(P_i - P_i^{\max}), \quad (4.16)$$

$$P_i \geq P_i^{\min}, \lambda_{2,i} > 0, \lambda_{2,i}(-P_i + P_i^{\min}) \quad (4.17)$$

for $i = 1, 2$. Condition (4.15) can be expressed as

$$(\beta_1 - \beta_2)f_1'(\gamma_1)\frac{\gamma_1}{P_1} - \beta_1\alpha_1 - \lambda_{1,1} + \lambda_{2,1} = 0, \quad (4.18)$$

$$-(\beta_1 - \beta_2)f_1'(\gamma_1)\frac{\gamma_1 h_2}{h_2 P_2 + \sigma^2} - \beta_2\alpha_2 - \lambda_{1,2} + \lambda_{2,2} = 0. \quad (4.19)$$

108

Conditions (4.18)-(4.19) are satisfied, if and only if

$$\beta_1 \alpha_i + \lambda_{1,i} - \lambda_{2,i} = 0, \ i = 1, 2. \tag{4.20}$$

We need to satisfy $\lambda_{2,1} > 0$ and $\lambda_{2,2} > 0$, whereas $\lambda_{1,1} = 0$ and $\lambda_{1,2} = 0$. Therefore, the KKT conditions (4.17) require that $P_i = P_i^{\min}$, $i = 1, 2$. The same solutions also hold for the case of two selfish transmitters.

4.7 Arbitrary Number of Selfish and Malicious Transmitters

Consider disjoint sets S and M (of arbitrary cardinality) of selfish and malicious transmitters. Any transmitter $i \in M$ attacks all transmitters in set S. Define $f_{i,j}(\gamma_j)$ as the reward or cost received by node i, when node j achieves the SINR value of γ_j.

4.7.1 Utility Function Type 1

The utility of transmitter $i \in S$ is $u_i(\underline{P}) = \frac{f_i(\gamma_i)}{P_i}$ and the utility of node $i \in M$ is $u_i(\underline{P}) = \frac{\sum_{j \in S} f_{i,j}(\gamma_j)}{P_i}$. For simplicity, we assume $f_{i,j}(\gamma_j) = 1 - f_j(\gamma_j)$ for any malicious node i, where $\gamma_j = \frac{h_j P_j}{\sum_{k \neq j} h_k P_k + \sigma^2}$. For any selfish transmitter $i \in S$, the condition $\frac{\partial u_i(\underline{P})}{\partial P_i} = 0$ is equivalent to $f_i'(\gamma_i)\gamma_i = f_i(\gamma_i)$. By definition, the utility function $\frac{f_i(\gamma_i)}{P_i}$ is quasi-concave in P_i. Therefore, there exists a unique equilibrium SINR value of γ_i^* that satisfies $f_i'(\gamma_i^*)\gamma_i^* = f_i(\gamma_i^*)$. For any malicious transmitter $i \in M$, the condition $\frac{\partial u_i(\underline{P})}{\partial P_i} = 0$ is equivalent to

$$\sum_{j \in S} f_j'(\gamma_j)\gamma_j \left(\frac{h_i}{\sum_{k \neq j} h_k P_k + \sigma^2} \right) = \sum_{j \in S}(1 - f_j(\gamma_j)). \tag{4.21}$$

109

The non-cooperative Nash equilibrium strategies for $f_i'(\gamma_i^*)\gamma_i^* = f_i(\gamma_i^*)$, $i \in S$,

are given by $P_i^* = \max(\min(\overline{P}_i, P_i^{\max}), P_i^{\min})$, where

$$\overline{P}_i = \frac{1}{h_i} \frac{\sum_{j \in S}(1 - f_j(\gamma_j^*))}{\sum_{j \in S} \frac{f_j(\gamma_j^*)\gamma_j^*}{h_j P_j^*}} \, , \ i \in M, \tag{4.22}$$

$$\overline{P}_i = \frac{1}{h_i}\gamma_i^*\left(\sum_{k \neq i} h_k P_k + \sigma^2\right), \ i \in S. \tag{4.23}$$

4.7.2   Utility Function Type 2

The utility of node $i \in S$ is $u_i(\underline{P}) = f_i(\gamma_i) - \alpha_i P_i$ and the utility of node

$i \in M$ is $u_i(\underline{P}) = \sum_{j \in S}(1 - f_j(\gamma_j)) - \alpha_i P_i$. From $\frac{\partial u_i(\underline{P})}{\partial P_i} = 0$, $i \in N$, the non-

cooperative Nash equilibrium strategies such that $\sum_{j \in S} \frac{\alpha_j}{h_j}\gamma_j^* = \frac{\alpha_i}{h_i}$, $i \in M$, are given

by $P_i^* = \max(\min(\overline{P}_i, P_i^{\max}), P_i^{\min})$, where

$$\overline{P}_i = \frac{1}{\alpha_i}f_i'(\gamma_i^*)\gamma_i^* \, , \ i \in S, \tag{4.24}$$

$$\overline{P}_i = \frac{h_j P_j^*}{h_i \gamma_j^*}\left(\sum_{k \neq i, k \neq j} h_k P_k^* + \sigma^2\right), \ i \in M, \ \text{for any } j \in S. \tag{4.25}$$

4.7.3   Numerical Results

Assume $\alpha_i = 1$, $h_i = 1$, $P_i^{\min} = 0$, $P_i^{\max} = \infty$, $i \in N$, $\sigma^2 = 0.1$ and

$f_i(\gamma_i) = (1 - e^{-\gamma_i})^2$, $i \in S$. The symmetric non-cooperative Nash equilibrium strate-

gies are depicted as function of the number of malicious transmitters in Figure 4.7 for

one selfish transmitter and utility function type 1. The symmetric non-cooperative

Nash equilibrium strategies are depicted as function of the number of selfish trans-

mitters in Figure 4.8 for one malicious transmitter and utility function type 2. As

the number of malicious nodes increases, a smaller power level is sufficient to jam

transmissions of selfish nodes, whereas the power level of selfish node is not affected
for utility function type 1 and decreases for utility function type 2.
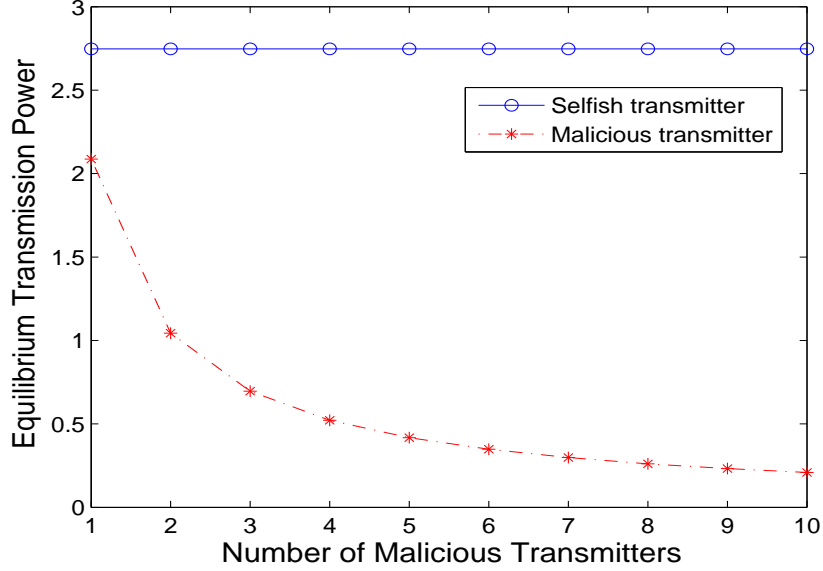
Figure 4.7: Transmission powers for utility function type 1.

4.8 Alternative Measures for Transmission Rates

So far, we modeled the transmission rate of transmitter i by a continuous and
differentiable function f_i. In this section, we consider a successful packet reception
criterion based on SINR threshold that restricts function f_i to

$$f_i(\gamma_i) = \begin{cases} 1 & \text{, if } \gamma_i \geq \Gamma_i \\ 0 & \text{, otherwise} \end{cases} \qquad (4.26)$$

such that a packet is successfully received, if transmitter i exceeds the predetermined
SINR threshold. We consider one selfish and one malicious transmitter with utility
function type 1. If $\mathcal{P}_i = [0, \infty)$, $i = 1, 2$, there exists no finite power level in non-

111
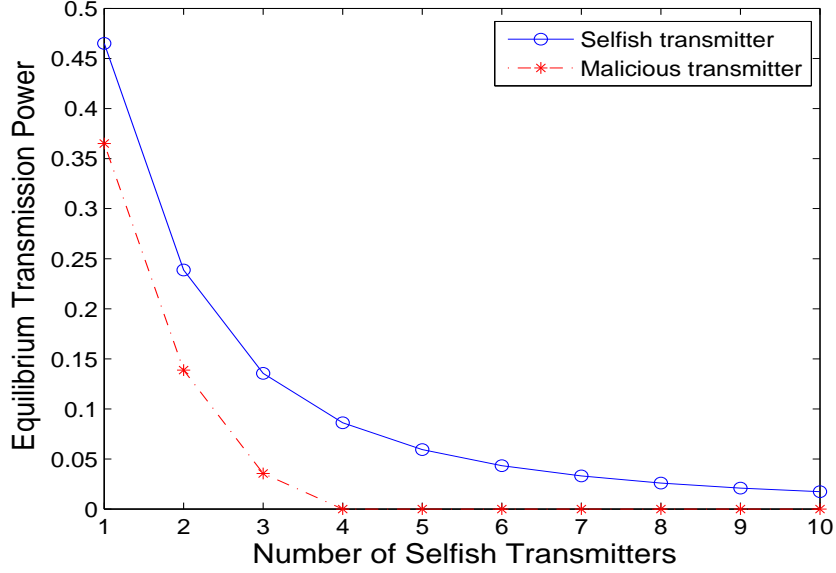
Figure 4.8: Transmission powers for utility function type 2.

cooperative Nash equilibrium, since given any P_i, $i = 1, 2$, node $j \neq i$ needs to increase transmission power P_j to improve performance. If $\mathcal{P}_i = [0, P_i^{\max}]$, $i = 1, 2$, there may exist multiple Nash equilibrium strategies

$$P_1^* = \Gamma_1 \frac{h_2 P_2^{\max} + \sigma^2}{h_1}, P_2^* \in [0, P_2^{\max}], \text{if } \Gamma_1 \frac{h_2 P_2^{\max} + \sigma^2}{h_1} < P_1^{\max} < \infty, \quad (4.27)$$

$$P_1^* \in [0, P_1^{\max}], P_2^* = \frac{\frac{h_1 P_1^{\max}}{\Gamma_1} - \sigma^2}{h_2}, \text{if } \frac{\frac{h_1 P_1^{\max}}{\Gamma_1} - \sigma^2}{h_2} < P_2^{\max} < \infty, \quad (4.28)$$

$$P_i^* = 0, i = 1, 2, \text{if } \frac{\Gamma_1 \sigma^2}{h_1} > P_1^{\max} \quad (4.29)$$

depending on system parameters. On the other hand, if we consider utility function type 2, there exist unique Nash equilibrium strategies

$$P_1^* = \Gamma_1 \frac{h_2 P_2^{\max} + \sigma^2}{h_1}, P_2^* = 0, \text{ if } \Gamma_1 \frac{h_2 P_2^{\max} + \sigma^2}{h_1} < \min\left(P_1^{\max}, \frac{1}{\alpha_1}\right), \quad (4.30)$$

$$P_1^* = 0, P_2^* = \frac{\frac{h_1 P_1^{\max}}{\Gamma_1} - \sigma^2}{h_2}, \text{ if } \frac{\frac{h_1 P_1^{\max}}{\Gamma_1} - \sigma^2}{h_2} < \min\left(P_2^{\max}, \frac{1}{\alpha_2}\right). \quad (4.31)$$

112

4.9 Summary and Conclusions

We addressed the problem of power control for selfish and malicious transmitters, and formulated a non-cooperative game of selecting transmission powers to optimize the individual objectives of throughput and energy efficiency. We used two different utility functions to derive the non-cooperative Nash equilibrium strategies. For distributed implementation, we presented the best-response update mechanisms that converge to Nash equilibrium. We also discussed the social equilibrium strategies that optimize the total system utility. The extension to multi-hop communication requires more complicated utility functions, as we will consider in Chapter 5 for the case of random access with collision channels. Further implications on energy and delay efficiency should be also explored. Our analysis was based on the assumption of uninterrupted availability of packets to be transmitted by selfish transmitters. Future work should consider the stable operation with random packet arrivals and possibly emptying packet queues, as we discussed in Chapter 3 for random access.

Chapter 5

A Game-Theoretic Analysis of Joint MAC and Routing in Non-Cooperative Wireless Networks

We extend the non-cooperative MAC operation to multi-hop communication and address the cross-layer problem of joint MAC and routing over a single relay channel. A stochastic game is formulated for transmitter and relay nodes competing over collision channels to deliver packets to a common destination node using alternative paths. We rely on a reward mechanism to stimulate cooperation for packet forwarding and evaluate the conflicting multiple access and routing strategies of direct communication and relaying through a detailed foray into the questions of cooperation incentives, throughput, delay and energy efficiency. We assume random packet arrivals such that the transmission strategies depend on the availability of packets at source and relay queues. We study the interactions among the equilibrium strategies under the separate models of selfish and cooperative network operation. We also developed a distributed adaptive algorithm for the case of unknown system parameters.

5.1 Introduction

We have studied in Chapters 3 and 4 the non-cooperative wireless network operation at the MAC layer. The extension to multi-hop operation introduces new

challenges and requires game models with extended strategy spaces and more complicated utility functions. If the transmission signal is modeled as decaying with distance, transmitter nodes might prefer relaying packets through intermediate nodes and over alternative paths (depending on node locations) in order to preserve transmission energy and reduce the interference effects throughout the network. However, the additional distance traveled by packets in multi-hop operation may increase the packet delay and reduce the throughput, especially if we do not allow simultaneous packet transmission and reception by any node. In addition, packet forwarding increases the energy consumption and decreases the throughput of relay nodes. The throughput objectives of (transmitter and relay) nodes conflict further, if the packet transmissions share a single collision channel based on random access.

To understand the tradeoffs that involve energy, interference and distance limitations, we strip out all complexities introduced by multiple source-destination pairs and focus on a simple relay channel topology, namely an ad hoc wireless network of a transmitter and relay node both with individual packet traffic addressed to a common destination. The relaying operation is only possible over the relay node that is chosen as the one closest to the destination. The relay channel has been extensively analyzed in terms of throughput and (information-theoretic) capacity properties [52] for the case of full cooperation of transmitter and relay nodes. Instead, we formulate the communication incentives over the relay channel as a stochastic game between transmitter and relay nodes with conflicting interests of throughput, energy efficiency and delay.

For single-hop transmissions, the game theory has been applied to study the

information-theoretic multi-access channels [47], practical collision channels based on random access (in Chapter 3) and power-controlled MAC (in Chapter 4). However, some form of multi-hop routing operation is unavoidable in full-fledged networks with topology and energy limitations. The classical routing problem inherently assumes full cooperation of nodes for relaying purposes in wireline networks, whereas game formulations are needed to study the conflicting routing objectives [53]. The extension to ad hoc wireless networks should incorporate the joint operations of MAC and routing with distributed implementation. There is a need to develop cooperation stimulation mechanisms for packet forwarding [54, 55, 56].

In this chapter, we evaluate the fundamental tradeoffs between the two extremes of cooperation and selfishness in terms of two basic (joint MAC and routing) strategies of direct communication and relaying. Our ultimate objective is to apply the basic game-theoretic tools to understand the efficient operational modes of multi-hop wireless networks starting from the simplest building block of relay channels. In our game formulation, the transmitter node selects one of the three possible actions of transmitting packets directly to the final destination, delivering them first to the closer relay node, or simply waiting. On the other hand, the relay node chooses between direct transmission and waiting, if it has its own packet to transmit. Otherwise, it decides on whether to accept or reject a packet of the transmitter node. Instead of using external enforcement, we introduce a reward mechanism to stimulate the cooperation of relay node to forward packets. The proposed non-cooperative game formulation allows selfish nodes to select the probability distributions over their available actions with the objective of optimizing their

individual performance measures of throughput, delay and energy consumption. On the other hand, the full cooperation among nodes is formulated as a team problem of optimizing the total system performance.

The chapter is organized as follows. Section 5.2 describes the relay channel model, possible strategies, rewards and costs of transmitter and relay nodes. After introducing the general framework for two-person stochastic games in section 5.3, we devote section 5.4 to formulate non-cooperative communication over the relay channel as a stochastic game. Section 5.5 provides numerical results for cooperative and non-cooperative equilibrium strategies. Then, section 5.6 describes a distributed adaptive algorithm under the assumption of limited information on the system parameters. In section 5.7, we improve the relaying mechanism by offering higher transmission priority to the new and forwarded packets. We collect the final remarks in section 5.8.

5.2   Network Model for the Simple Relay Channel

We consider a static relay channel topology of a transmitter node (node 1), a relay node (node 2) and a common destination node (node 3), as shown in Figure 5.1. Nodes 1 and 2 independently generate packets destined to a common destination at each time slot with probability λ_1 and λ_2, respectively, provided that their individual queues were empty at the end of the previous slot. To simplify the analysis, we assume that the maximum buffer capacity is one.

As the routing strategy, node 1 chooses between sending packets directly to
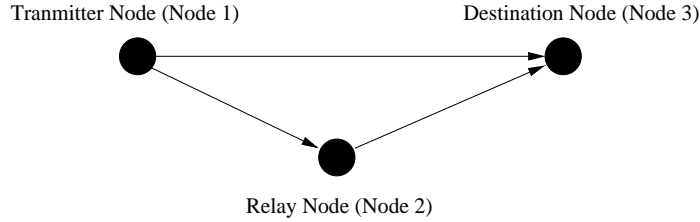
Figure 5.1: The simple relay channel model.

node 3 or relying on node 2 to forward them to the final destination, whereas node 2 has the only option of transmitting packets directly to node 3. We consider the classical collision channel model for both transmissions to the relay node and common destination node. Transmitter and relay nodes also have to choose between transmitting and waiting as means of random access to collision channels.

At each receiver node (namely node 2 or 3), transmissions are subject to one of the three possible channel outputs, namely idle, success or collision, whenever 0, 1 or more than one packet reach the particular receiver in the given time slot. We assume that the simultaneous transmissions to different receivers destructively interfere with each other. Also, we do not allow simultaneous transmission and reception by any node at any time slot as an extension of the single-receiver MAC model in Chapters 3 and 4. Another form of unsuccessful transmission (i.e. packet collision) occurs, if node 1 transmits to node 2, which has already a packet in its queue or does not accept the packet from node 1 to forward to the final destination. In any case of packet collision, the backlogged nodes attempt to retransmit their packets in the subsequent slots for reliable communication.

5.2.1 Cooperation Stimulation Mechanism for Relaying

Node 2 cannot accept the packet of node 1, whenever it has a packet to transmit in its queue. Otherwise, node 2 decides on whether to accept or reject a packet (possibly) arriving from node 1. If node 2 accepts a packet of node 1, node 2 does not generate any new packet during that particular time slot and undertakes all future rewards and costs of the accepted packet by paying in return an immediate reward c to node 1. The proposed strategy of charging the relay node an immediate payoff for the forwarded packet and offering instead a future throughput reward of value 1 (common for all delivered packets) will stimulate the cooperation of node 2 to deliver the forwarded packet to the final destination. We assume that node 1 is immediately informed of whether the transmitted packet is accepted by node 2 or not.

For each successful transmission to the common destination, the transmitting node (1 or 2) receives a throughput reward of value 1. Node 1 receives a reward c from node 2 for delivering a packet to node 2, which can obtain the full throughput credit of value 1 only after successfully transmitting that particular packet to the final destination in a subsequent time slot. Each packet transmission attempt from node i to node j incurs an energy cost $E_{i,j} \geq 0$. The transmission power is chosen as the smallest value that would result in successful packet reception in the absence of interfering transmissions. If the signal power decays with distance r as $\frac{1}{r^\alpha}$, where $\alpha \geq 2$ is the path loss exponent, the energy cost $E_{i,j}$ can be expressed as $E_0\, r_{i,j}^\alpha$ for some positive constant E_0, where $r_{i,j}$ denotes the distance between nodes i and

j. Each slot of packet delay results in an additive cost D for the node that has the particular packet in its queue (regardless of the source of that packet). The immediate utility of any node is defined as the difference between the reward and cost obtained during the preceding time slot.

5.3 A Framework for Two-User Stochastic Games

A two-user game is described by the tuple $\{\mathcal{K}, \mathcal{A}_1, \mathcal{A}_2, T, u_1, u_2\}$, where $\mathcal{K}$ is the state space, $\mathcal{A}_i$ is the action space of user $i = 1, 2$, $T : \mathcal{K} \times \mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{K} \rightarrow [0, 1]$ is a transition probability function, and $u_i : \mathcal{K} \times \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \Re$ is an immediate utility function of user i. We consider a slotted system with all users synchronously selecting their actions at the beginning of each time slot. Suppose that the game is in state $k^t \in \mathcal{K}$ in time slot t and the users play actions $A_1^t \in \mathcal{A}_1$ and $A_2^t \in \mathcal{A}_2$. Then, each user i receives the immediate utility $u_i(k^t, A_1^t, A_2^t)$ and the game moves in the next time slot $t + 1$ to state $k^{t+1} \in \mathcal{K}$ with the probability given by the transition function $T(k^{t+1}|k^t, A_1^t, A_2^t)$. We define a history h^t at time t to be the sequence of the current state, previous states and previous actions such that $h^t = (k^1, A_1^1, A_2^1, ..., k^{t-1}, A_1^{t-1}, A_2^{t-1}, k^t)$. Let H^t be the set of all possible histories until time t. A mixed strategy $\underline{s}_i$ for user i is a sequence of $s^t : H^t \rightarrow P(\mathcal{A}_i)$, which is a function that assigns to H^t a probability measure over the action set of user i. Let $\underline{s} = (\underline{s}_1, \underline{s}_2)$ denote the joint mixed strategies of the two users. Any initial state distribution β and strategy $\underline{s}$ jointly define the probability measure $P_{\underline{s}, \beta}$, which determines the distributions of the stochastic process $\{k^t, A_1^t, A_2^t\}$ of states

and actions for all $t \geq 1$. The (undiscounted) time-average utility (per time slot) of user i is defined as

$$u_i^{\beta}(\underline{s}) = \lim_{t_f \to \infty} \frac{1}{t_f} E_{\underline{s},\beta} \left[\sum_{t=1}^{t_f} u_i(k^t, a_1^t, a_2^t) \right], \tag{5.1}$$

where the expectation $E_{\underline{s},\beta}$ is taken over $P_{\underline{s},\beta}$. In the non-cooperative game, user i independently selects $\underline{s}_i$ to maximize $u_i^{\beta}(\underline{s})$. In the cooperative team problem, users jointly select $\underline{s}$ to maximize the total utility $\sum_{i=1}^{2} u_i^{\beta}(\underline{s})$.

The given formulation of the stochastic games is based on the assumption that all states of the game are perfectly observable by all users (i.e. transmitter and relay nodes). However, the particular problem of communication over the relay channel topology does not fit in this framework, since nodes can only have partial information on the state of the game, which should in some form represent the evolution of the packet queues. As a simple but effective solution, we define the content of (transmitter and relay) queues (e.g. whether each queue has a new, a backlogged, a forwarded or no packet) as the state of the game (on which the node strategies will be based) and restrict the information of each node to its own packet queue (ignoring the partial information that nodes can have on the content of each other's queue). For tractable analysis, we need to modify the previous game model and redefine the strategies and expected utilities as follows.

We consider only the stationary (mixed) strategies of the form $\underline{s}_{i,k}(\mathcal{A}_i) = \{s_{i,k}(A), A \in \mathcal{A}_i\}$ such that $s_{i,k}(A)$ assigns the (stationary) probability distribution to action $A \in \mathcal{A}_i$ of each user i at state $k \in \mathcal{K}$. We follow the Markov game assumption such that the decisions of users are only based on the current state $k^t \in \mathcal{K}$

(independent of time t) instead of the entire history H^t. For user i, the random strategy space is defined as the shortest probability vector $\underline{s}_i$ that can uniquely describe the non-deterministic strategies $\{s_{i,k}(A), A \in \mathcal{A}_i, k \in \mathcal{K}\}$, which are not (necessarily) limited to the probability 0 or 1. We denote by $\underline{s} = (\underline{s}_1, \underline{s}_2)$ the joint space of the random stationary strategies of both users. For any stationary strategy $\underline{s}$, the state $k_j \in \mathcal{K}$ has the stationary distribution $\pi_j(\underline{s})$ such that the vector $\underline{\pi}(\underline{s}) = \{\pi_j(\underline{s}), k_j \in \mathcal{K}\}$ satisfies the global balance equations $\underline{\pi}(\underline{s}) = \underline{\pi}(\underline{s})\mathrm{T}(\underline{s})$, where $\mathrm{T}(\underline{s})$ denotes the state transition matrix. The (k_1, k_2)th entry of $\mathrm{T}(\underline{s})$ gives the probability of transition from state k_1 to k_2 under the stationary strategy $\underline{s}$ (i.e. the probability is averaged over the possible actions of nodes and the channel outcomes) and can be explicitly derived as $\sum_{A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2} T(k_2|k_1, A_1, A_2) s_{1,k_1}(A_1) s_{2,k_2}(A_2)$, where $T(k_2|k_1, A_1, A_2)$ is the stationary state transition probability from state $k_1 \in \mathcal{K}$ to state $k_2 \in \mathcal{K}$, if the actions of users are $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$. For the case of partially observable states, we use the expectation over the state distributions (rather than the time-average form of Eq. (5.1)) to define $u_i(\underline{s})$ as the stationary utility per time slot expected by user i:

$$u_i(\underline{s}) = \sum_{k_j \in \mathcal{K}} \pi_j(\underline{s}) \, E[u_i(k_j, \underline{s})], \tag{5.2}$$

where $E[u_i(k_j, \underline{s})]$ denotes the immediate utility (over a single time slot) expected by user i, if the joint strategy $\underline{s}$ is played at state k_j. The objective of the non-cooperative user i is to select the strategy $\underline{s}_i$ independently to maximize the expected utility $u_i(\underline{s})$. In the cooperative case, both users jointly select $\underline{s}$ to maximize the total utility $\sum_{i=1}^{2} u_i(\underline{s})$.

5.4 Communication over Relay Channel as a Stochastic Game

5.4.1 State Definition

For communication over the simple relay channel, the state of the joint multiple access and routing game between transmitter and relay nodes is defined as (Q_1, Q_2), where Q_1 and Q_2 denote the queue contents of node 1 and 2, respectively. For $i = 1, 2$, the quantity Q_i takes the value 0 or 1, if no or one packet (independent of whether it is a new, forwarded or backlogged packet) is present at the queue of node i. We assume that new and forwarded packets are immediately backlogged, before they are transmitted for the first time, i.e. nodes do not distinguish between the transmission strategies of the packets in their queues.

The four states of the resulting game are ordered in the state space as given by $\mathcal{K} = \{(0,0), (0,1), (1,0), (1,1)\}$. Since nodes have reliable information only on their own queues, we assume that the strategy of each node i is based on Q_i rather than on the partial information of the complete state (Q_1, Q_2). In section 5.7, we will extend the state space to include forwarded and backlogged packets as additional types of queue contents to exploit more efficiently the side information of nodes on each other's packet queues.

5.4.2 Action Space and Mixed Stationary Strategies

The actions available to node 1 are A_1^1 (transmitting to node 3), A_2^1 (transmitting to node 2) and A_3^1 (waiting), i.e. we have $\mathcal{A}_1 = \{A_1^1, A_2^1, A_3^1\}$. The mixed stationary strategies are $\underline{s}_{1,k}(A_1^1, A_2^1, A_3^1) = (p_1^1, p_2^1, 1 - p_1^1 - p_2^1)$ for states $k \in \{(1,0), (1,1)\}$,

123

where p_1^1, p_2^1 and $1 - p_1^1 - p_2^1$ denote the probabilities of selecting actions A_1^1, A_2^1 and A_3^1, respectively. Node 1 has the stationary strategy $\underline{s}_{1,k}(A_1^1, A_2^1, A_3^1) = (0, 0, 1)$ for states $k \in \{(0, 0), (0, 1)\}$. The random strategy space of node 1 is uniquely described by $\underline{s}_1 = (p_1^1, p_2^1)$.

Node 2 has the available actions A_1^2 (transmitting to node 3) and A_2^2 (waiting) for states $(0, 1), (1, 1)$, and has the available actions A_3^2 and A_4^2 (accepting and rejecting a packet transmitted from node 1) for states $(0, 0), (1, 0)$. Thus, we have $\mathcal{A}_2 = \{A_1^2, A_2^2, A_3^2, A_4^2\}$. The mixed stationary strategies of node 2 are $\underline{s}_{2,k}(A_1^2, A_2^2, A_3^2, A_4^2) = (p_1^2, 1 - p_1^2, 0, 1)$ for states $k \in \{(0, 1), (1, 1)\}$, where p_1^2 and $1 - p_1^2$ denote the probabilities of selecting actions A_1^2, A_2^2 at states $(0, 1)$ and $(1, 1)$, and $\underline{s}_{2,k} = (0, 1, p_3^2, 1 - p_3^2)$ for states $k \in \{(0, 0), (1, 0)\}$, where p_3^2 and $1 - p_3^2$ denote the probabilities of selecting actions A_3^2, A_4^2 at states $(0, 0)$ and $(1, 0)$. The random strategy space of node 2 is uniquely described by the stationary distribution $\underline{s}_2 = (p_1^2, p_3^2)$ and the joint set of random stationary strategies is given by $\underline{s} = (p_1^1, p_2^1, p_1^2, p_3^2)$.

5.4.3 State Transition Matrix and Utility Functions

For any stationary strategy $\underline{s}$, the evolution of (Q_1, Q_2) follows a two-dimensional ergodic Markov chain, which is irreducible and aperiodic with finite number of states such that there exists a unique stationary distribution $\underline{\pi}(\underline{s})$. If the strategy $\underline{s} = (p_1^1, p_2^1, p_1^2, p_3^2)$ is played, the ith row $\mathrm{T}_i(\underline{s})$ of the state transition matrix $\mathrm{T}(\underline{s})$ is given by

$$T_1(\underline{s}) = [(1 - \lambda_1)(1 - \lambda_2), (1 - \lambda_1)\lambda_2, \lambda_1(1 - \lambda_2), \lambda_1\lambda_2],$$

$$T_2(\underline{s}) = [(1 - \lambda_1)p_1^2(1 - \lambda_2), (1 - \lambda_1)(1 - p_1^2 + p_1^2\lambda_2), \lambda_1 p_1^2(1 - \lambda_2),$$

$$\lambda_1(1 - p_1^2 + p_1^2\lambda_2)],$$

$$T_3(\underline{s}) = [p_1^1(1 - \lambda_1)(1 - \lambda_2), (1 - \lambda_1)(p_1^1\lambda_2 + p_2^1 p_3^2),$$

$$(1 - \lambda_2)(1 - p_1^1 - p_2^1 + p_1^1\lambda_1 + p_2^1(1 - p_3^2)),$$

$$(1 - p_1^1 - p_2^1)\lambda_2 + p_2^1 p_3^2\lambda_1 + p_2^1(1 - p_3^2)\lambda_2 + p_1^1\lambda_1\lambda_2],$$

$$T_4(\underline{s}) = [0, p_1^1(1 - \lambda_1)(1 - p_1^2), (1 - p_1^1 - p_2^1)p_1^2(1 - \lambda_2),$$

$$(1 - p_1^1 - p_2^1)(1 - p_1^2 + p_1^2\lambda_2 + p_1^1 p_1^2 + p_2^1 + p_1^1\lambda_1(1 - p_1^2)].$$

If nodes follow the strategy $\underline{s} = (p_1^1, p_2^1, p_1^2, p_3^2)$, the expected utilities $u_1(\underline{s})$ and $u_2(\underline{s})$ are given by

$$u_1(\underline{s}) = \pi_3(\underline{s})[\, p_1^1(1 - E_{1,3}) + p_2^1(p_3^2 c - (1 - p_3^2)D - E_{1,2}) - (1 - p_1^1 - p_2^1)D] \quad (5.3)$$

$$+\pi_4(\underline{s})[p_1^1(-E_{1,3} + 1 - p_1^2 - p_1^2 D) + p_2^1(-E_{1,2} - D) - (1 - p_1^1 - p_2^1)D],$$

$$u_2(\underline{s}) = \pi_2(\underline{s})[p_1^2(1 - E_{2,3}) - (1 - p_1^2)D] + \pi_3(\underline{s})[-p_2^1 p_3^2 c] \quad (5.4)$$

$$+\pi_4(\underline{s})[p_1^2(-E_{2,3} + 1 - p_1^1 - p_2^1 - (p_1^1 + p_2^1)D) - (1 - p_1^2)D].$$

The term $u_2(\underline{s})$ seems to decrease monotonically with increasing p_3^2 such that accepting packets is malicious to node 2. However, if p_3^2 increases, then $\pi_2(\underline{s})$ and $\pi_4(\underline{s})$ can increase, while $\pi_3(\underline{s})$ is decreasing (depending on λ_1 and λ_2). This will compensate the decrease in $u_2(\underline{s})$ because of the term $p_2^1 p_3^2 c$. Although node 1 seems to benefit from an increase in p_3^2, the increase in $T_{32}(\underline{s})$ can shift distribution from

state $(1, 0)$ to $(0, 1)$ and consequently decrease $u_1(\underline{s})$. Thus, we cannot necessarily expect the pure strategy $p_3^2 \in \{0, 1\}$ for all values of c.

The collision channel with two uplink nodes is equivalent to the case of direct communication ($p_2^1 = 0$ or $p_3^2 = 0$) in the proposed game model (as was studied in Chapter 3 for saturated queues with infinite buffer capacities). On the other hand, the pure strategy of two-hop relaying (i.e. $p_1^1 = 0$) can be used to model the special case of very large values of $E_{1,3}$ compared to $E_{1,2}$.

For selfish users, we evaluate the strategies in Nash Equilibrium such that no user can improve its own utility, if the strategies of the other users remain the same. Thus, we are interested in the problem of finding the equilibrium strategies $\underline{s}^* = (\underline{s}_1^*, \underline{s}_2^*)$ such that $u_1(\underline{s}^*) \geq u_1(\underline{s}_1, \underline{s}_2^*)$ and $u_2(\underline{s}^*) \geq u_2(\underline{s}_1^*, \underline{s}_2)$ for any strategy $\underline{s} = (\underline{s}_1, \underline{s}_2)$. The best response correspondence of node 1 (if node 2 plays the strategy $\underline{s}_2$) is defined as $B_1(\underline{s}_2) = \arg\max_{\{\underline{s}_1\}} u_1(\underline{s}_1, \underline{s}_2)$, and the best response correspondence of node 2 (if node 1 plays the strategy $\underline{s}_1$) is defined as $B_2(\underline{s}_1) = \arg\max_{\{\underline{s}_2\}} u_2(\underline{s}_1, \underline{s}_2)$. As a result, the non-cooperative Nash equilibrium $\underline{s}^* = (\underline{s}_1^*, \underline{s}_2^*)$ is simply given by $\underline{s}_1^* \in B_1(\underline{s}_2^*)$ and $\underline{s}_2^* \in B_2(\underline{s}_1^*)$. The cooperation between transmitter and relay nodes can be set up as a team problem of maximizing $u_\Sigma(\underline{s}) = \sum_{i=1}^{2} u_i(\underline{s})$ over $\underline{s}$ such that $\underline{\pi}(\underline{s}) = \underline{\pi}(\underline{s}) \mathrm{T}(\underline{s})$, $0 \leq \pi_j(\underline{s}) \leq 1$ for any $k_j \in \mathcal{K}$, and $\sum_{k_j \in \mathcal{K}} \pi_j(\underline{s}) = 1$.

5.5 Performance Evaluation of Relay Channel Communication

For numerical results, we consider the common system parameters: $E_{1,3} = 0.2$, $E_{1,2} = 0.1$, $E_{2,3} = 0.1$, $D = 0.1$, $c = 0.5$, $\lambda_1 = 0.25$ and $\lambda_2 = 0.25$. For the non-

cooperative case, we illustrate the equilibrium strategies (namely the probabilities p_1^1, p_2^1, p_1^2 and p_3^2) as functions of $E_{1,3}, c, \lambda_1$ and λ_2 in Figures 5.2, 5.3, 5.4 and 5.5, respectively.
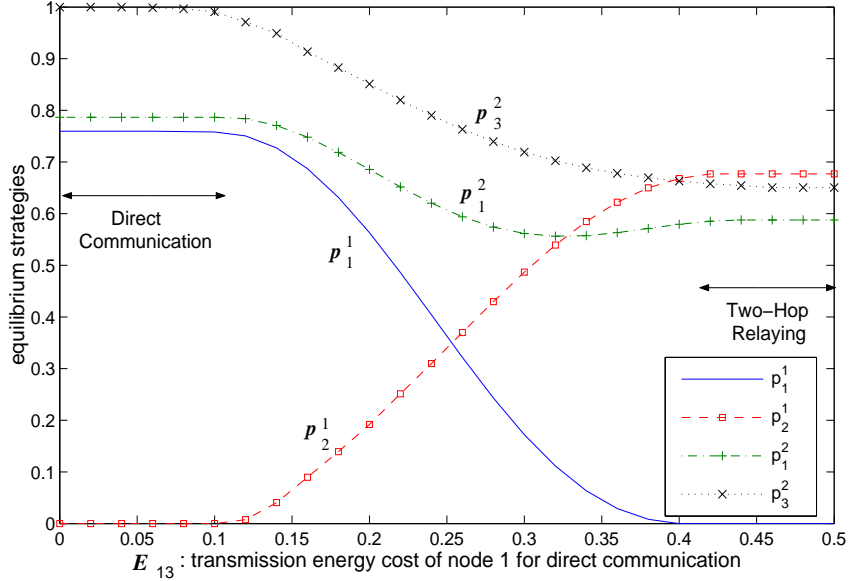
Figure 5.2: Effects of energy cost $E_{1,3}$ on non-cooperative equilibrium strategies.

We observe from Figure 5.2 that low values of $E_{1,3}$ support the direct communication (with $p_2^1 = 0$), whereas node 1 transfers the distributions from p_1^1 to p_2^1 with increasing $E_{1,3}$ and starts operating only in two-hop relaying mode (with $p_1^1 = 0$) for high values of $E_{1,3}$. As shown in Figure 5.3, node 1 selects $p_2^1 = 0$ for $c < 0.37$ and node 2 selects $p_3^2 = 0$ for $c > 0.71$, whereas node 1 counteracts by switching to direct communication strategy with $p_2^1 = 0$. The mixed strategies of direct communication and relaying are only possible for intermediate values of c between 0.37 and 0.71.

Figures 5.4 and 5.5 show that both strategies p_3^2 and p_2^1 are inversely proportional to the packet generation probabilities λ_1 and λ_2. The probabilities $p_1^1 + p_2^1$
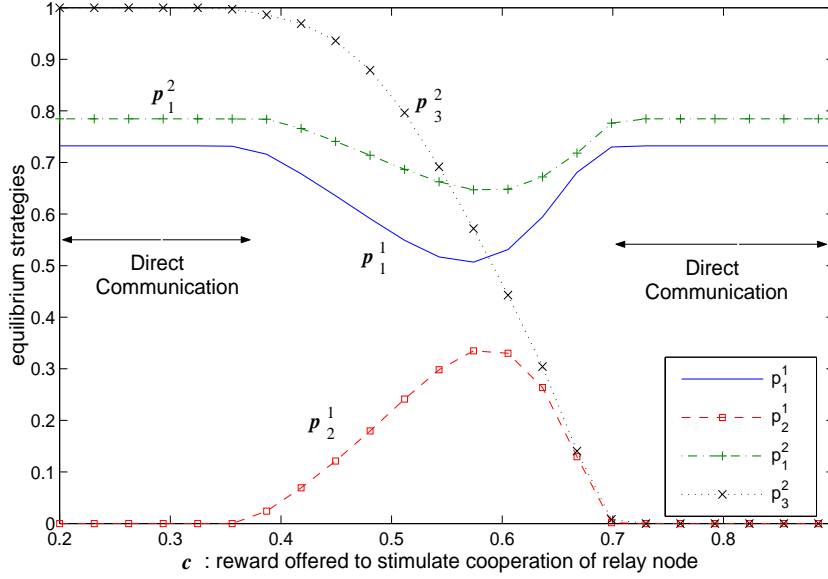
Figure 5.3: Effects of reward c for cooperation stimulation on non-cooperative equilibrium strategies.

and p_1^2 of multiple access strategies decrease first with increasing λ_1 and λ_2, until both nodes respond to each other's strategy by becoming more aggressive in their transmission decisions (i.e. increasing p_1^1 and p_1^2) to keep pace with high arrival rates. Further increases in λ_1 or λ_2 cause intensive packet accumulation at both nodes that increases $\pi_4(\underline{s})$ close to 1 and restricts node 1 to the strategy $p_2^1 = 0$ (i.e. direct communication) and node 2 to the strategy p_3^2 (i.e. no packet forwarding). The strategies p_1^1 and p_1^2 continue to increase with λ_1, since node 1 (exposed to high level of packet accumulation) switches to the aggressive transmission policy by increasing p_1^1 and node 2 can easily respond to that by increasing p_1^2 because of the relatively low energy cost $E_{2,3}$. However, if we further increase λ_2, node 2 needs to increase p_1^2 as before, whereas node 1 cannot react similarly because of the high

128

value of the energy cost $E_{1,3}$ and instead starts decreasing p_1^1.
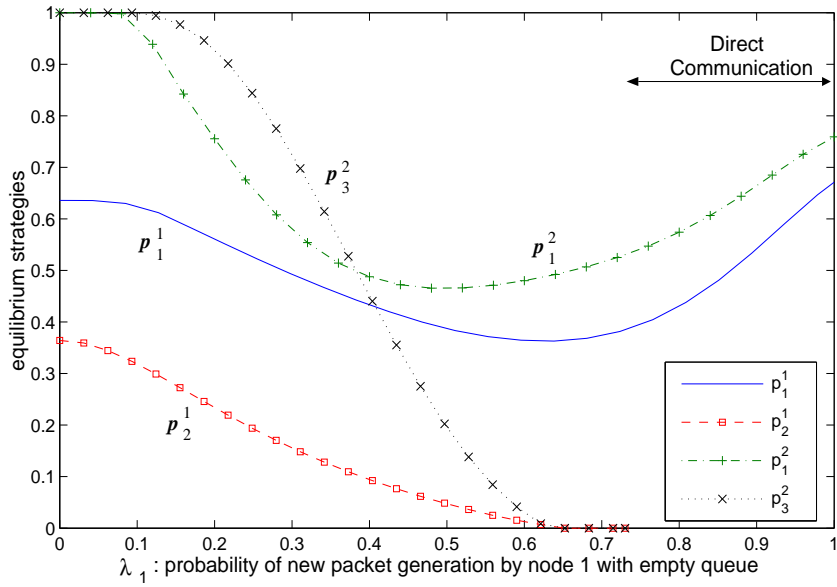
Figure 5.4: Effects of new packet generation probability λ_1 at node 1 on non-cooperative equilibrium strategies.

We evaluate the non-cooperative and social equilibrium utilities as functions of $E_{1,3}, c, \lambda_1$ and λ_2, and summarize the results in Table 5.1. We define $u_{i,nc}^*$ and $u_{i,c}^*$ as the expected utilities of node i in non-cooperative and social equilibrium, respectively. For both cases, the total expected utilities decrease with increasing $E_{1,3}$ (since more energy is wasted to keep the same throughput level) and increase with increasing λ_2 (since the contribution from the additional throughput by node 2 not only compensates the emerging packet collisions for node 1 but also extends the total system utility).

The gap between the values of $\sum_{i=1}^2 u_{i,c}^*$ and $\sum_{i=1}^2 u_{i,nc}^*$ increases with increasing λ_2, as the selfish nodes prefer lower transmission probabilities, although relatively

129
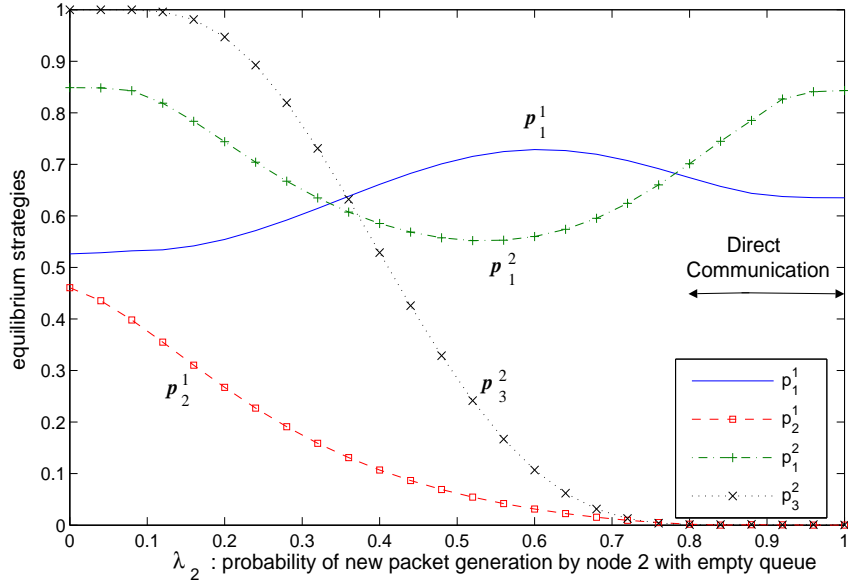
Figure 5.5: Effects of new packet generation probability λ_2 at node 2 on non-cooperative equilibrium strategies.

higher transmission probabilities can be chosen to increase the total utility under the assumption of full cooperation. The quantities $u_{1,nc}^*, u_{2,nc}^*$ and $\sum_{i=1}^{2} u_{i,nc}^*$ strongly depend on c, although the total expected utility and the equilibrium strategies are independent of c in the cooperative equilibrium case, as we can easily observe by adding Eqs. (5.3) and (5.4). Compared to the case of $c = 0.5$, the selfish and cooperating nodes achieve larger utilities for $c = 0.6$ with a larger value of $p_2^1 p_3^2$ (namely the probability of successful packet forwarding) such that the alternative route over the relay node becomes more feasible and the routing possibilities are extended for the transmitter node.

The comparison of the Nash equilibrium solutions with the cooperative team solutions in Table 5.1 indicates that the Nash equilibrium of the given game formula-

Table 5.1: Comparison of Non-cooperative and Cooperative Expected Utilities.

$c = 0.5$	$\lambda_2 = 0$ $E_{1,3} = 0.1$	$\lambda_2 = 0.25$ $E_{1,3} = 0.1$	$\lambda_2 = 0$ $E_{1,3} = 0.2$	$\lambda_2 = 0.25$ $E_{1,3} = 0.2$
$u_{1,nc}^*$	0.1500	0.0834	0.1048	0.0738
$u_{2,nc}^*$	0	0.1135	0.0227	0.1182
$u_{1,nc}^* + u_{2,nc}^*$	0.1500	0.1969	0.1276	0.1920
$u_{1,c}^*$	0.1500	0.0812	0.1119	0.0703
$u_{2,c}^*$	0	0.1580	0.0316	0.1606
$u_{1,c}^* + u_{2,c}^*$	0.1500	0.2392	0.1435	0.2309
$c = 0.6$	$\lambda_2 = 0$ $E_{1,3} = 0.1$	$\lambda_2 = 0.25$ $E_{1,3} = 0.1$	$\lambda_2 = 0$ $E_{1,3} = 0.2$	$\lambda_2 = 0.25$ $E_{1,3} = 0.2$
$u_{1,nc}^*$	0.1500	0.0867	0.1066	0.0755
$u_{2,nc}^*$	0	0.1170	0.0285	0.1201
$u_{1,nc}^* + u_{2,nc}^*$	0.1500	0.2036	0.1351	0.1956
$u_{1,c}^*$	0.1500	0.0859	0.1141	0.0714
$u_{2,c}^*$	0	0.1533	0.0294	0.1595
$u_{1,c}^* + u_{2,c}^*$	0.1500	0.2392	0.1435	0.2309

tion is not Pareto-optimal, i.e. nodes 1 and 2 can jointly improve their performance, although any improvement by unilateral changes in strategies is not possible at the stable operating points of Nash Equilibrium.

Ws assume that the strategies of nodes are based only on the present state rather than the entire history of the game. It is possible to improve the overall performance by allowing a larger state space with additional memory of a fixed number of time slots, as discussed in detail in [40]. However, the number of states will

increase substantially with increasing number of memory elements (in the history).

5.6   Distributed Adaptive Algorithm with Limited Information

In previous sections, we assumed that (both) nodes have the perfect information on the system parameters so that they can use the feedback information to compute (in advance) the expected values of the immediate utilities and the state transition probabilities for playing any action. However, it is also of interest to study the algorithms that do not depend on the system parameters, which cannot be explicitly known in some distributed applications.

Stochastic games with imperfect information have been extensively studied in the literature and traditional Q-Learning algorithms for one-person Markov Decision Processes has been extended to the zero and general-sum stochastic games [57, 58]. The general assumption is that the model of rewards and state transition probabilities are not known but must be observed through experience. There are several algorithms (such as Nash-Q for pure strategies [58] and Policy Hill Climbing for mixed strategies [59]) that converge the system to non-cooperative equilibrium points. However, the previous analysis as well the proposed algorithms are based on the assumption that the state of the game is fully observable, which is not the case in the relay game model studied in this chapter.

In this section, we follow a different approach to take the effects of imperfect state monitoring into account. We extend the stochastic adaptive algorithm of [19] (originally developed for the slotted Aloha systems) to the relay channel model and

allow nodes to vary their strategies as function of time t depending on the feedback information, i.e. we have the strategy $\underline{s}_1 = (p_1^1(t), p_2^1(t))$ and $\underline{s}_2 = (p_1^2(t), p_3^2(t))$ at time t. We assume that any successful transmission yields higher immediate utility than the channel outcomes of waiting or collision. We also assume that the system parameters are unknown but fixed, and define the sequence $\epsilon(t)$ that satisfies $\lim_{t \to \infty} \epsilon(t) = 0$ and $\sum_{t=1}^{\infty} \epsilon(t) = \infty$, as suggested by the stochastic approximation theory. If node 1 has a packet to transmit at time t, then the strategies $p_1^1(t)$ and $p_2^1(t)$ are updated at time t depending on the feedback information as follows:

$$p_1^1(t+1) = \min \left(\max \left(p_1^1(t) + \epsilon(t)\, \xi_1(t)\, ,\, 0 \right),\, 1 \right), \qquad (5.5)$$

$$p_2^1(t+1) = \min \left(\max \left(p_2^1(t) + \epsilon(t)\, \xi_2(t)\, ,\, 0 \right),\, 1 \right). \qquad (5.6)$$

The control parameters ξ_1 and ξ_2 take values $\xi_1(t) = 1, \xi_2(t) = -\frac{1}{2}$ or $\xi_1(t) = -1, \xi_2(t) = \frac{1}{2}$, if a transmission of node 1 to node 3 at time t is successful or fails. Similarly, $\xi_1(t) = -\frac{1}{2}, \xi_2(t) = 1$ or $\xi_1(t) = \frac{1}{2}, \xi_2(t) = -1$, if a transmission of node 1 to node 2 at time t is successful or fails. For each idle slot of waiting, we set $\xi_1(t) = \xi_2(t) = \frac{1}{2}$. For the rare situation, in which Eqs. (5.5) and (5.6) yield invalid distributions such that $\sum_{i=1}^{2} p_i^1(t+1) > 1$, we need a further mapping such that $\xi_1(t)$ and $\xi_2(t)$ are both reduced with the same scale to the largest values that satisfy $\sum_{i=1}^{2} p_i^1(t+1) = 1$. If node 2 has a packet to transmit at time t, then the strategy $p_1^2(t)$ is updated at time $t+1$ depending on the feedback information as follows:

$$p_1^2(t+1) = \min \left(\max \left(p_1^2(t) + \epsilon(t)\, \xi_3(t)\, ,\, 0 \right),\, 1 \right), \qquad (5.7)$$

where $\xi_3(t) = -1$, if a transmission of node 2 fails at time t, and $\xi_3(t) = 1$, if a transmission of node 2 is successful or node 2 waits. If the queue of node 2 is empty

at time t, then the strategy $p_3^2(t)$ is updated at time $t+1$ depending on the feedback information as follows:

$$p_3^2(t+1) = \min \left(\max \left(p_3^2(t) + \epsilon(t)\, \xi_4(t)\, ,\, 0 \right),\, 1 \right), \tag{5.8}$$

where $\xi_4(t) = -1$ or 1, if there is a forwarding request at time t or not. Update (5.8) stimulates node 1 to send more (or less) packets to node 2 in the case of low (or high) rates of forwarding requests.
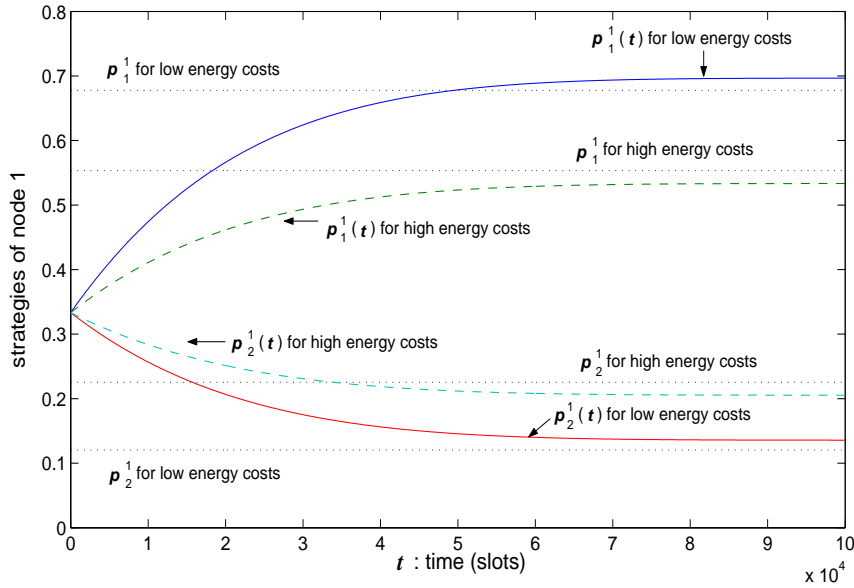
Figure 5.6: Temporal evolution of $p_1^1(t)$ and $p_2^1(t)$ for distributed adaptive algorithm.

The proposed algorithm represents the throughput and delay properties but cannot distinguish between the different energy costs. Therefore, the algorithm can easily deviate from the equilibrium strategies, since the expected utilities strongly depend on the unknown values of the energy costs. If the node decisions are dominated by the energy costs, the performance of the algorithm is expected to be poor, unless the control parameters are properly selected.

134
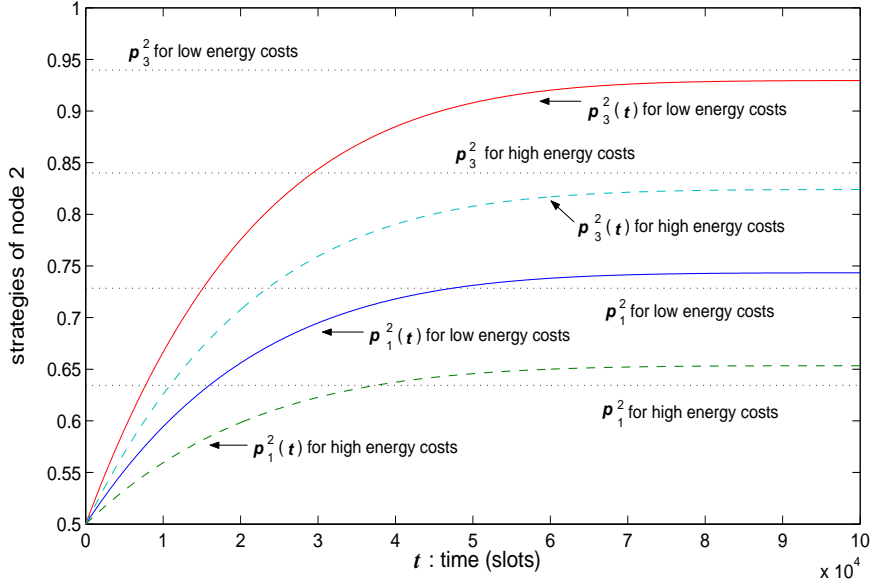
Figure 5.7: Temporal evolution of $p_1^2(t)$ and $p_3^2(t)$ for distributed adaptive algorithm.

We run the proposed algorithm for 10^5 time slots. We assume $p_1^1(1) = p_2^1(1) = \frac{1}{3}$, $p_1^2(1) = p_3^2(1) = \frac{1}{2}$ as the initial conditions for the algorithm and let $\epsilon(t) = \frac{1}{10t}$ for $t \geq 1$. The system parameters are $D = 0.1$, $c = 0.5$, $\lambda_1 = 0.25$ and $\lambda_2 = 0.25$. We consider two different sets of low and high energy costs: (I) $E_{1,3} = 0.02$, $E_{1,2} = E_{2,3} = 0.01$, and (II) $E_{1,3} = 0.2$, $E_{1,2} = E_{2,3} = 0.1$. Figures 5.6 and 5.7 illustrate the temporal evolution of the simulated values of $p_1^1(t)$, $p_2^1(t)$, $p_1^2(t)$ and $p_3^2(t)$ for low and high energy costs, and compared them to the exact numerical solutions obtained for the equilibrium strategies. For low energy costs, the algorithm converges close to the non-cooperative equilibrium with perfect information. However, the algorithm should deviate from the equilibrium points with increasing energy costs. For high energy costs, it is advisable to use control parameters lower than -1 for strategy updates in the case of packet collisions such that nodes are more

135

effectively stimulated to preserve energy. For instance, we can choose the control parameter -2 (instead of -1 as before), if a packet collision occurs, whereas control parameters for $p_1^1(t)$, $p_2^1(t)$ and $p_1^2(t)$ can be doubled for success and idle slots. Exhaustive search over the control parameters can make the proposed algorithm converge close to the equilibrium strategies even for high energy costs, as shown in Figure 5.6 and 5.7. However, the systematic way of finding the optimal control parameters (depending on the energy or other costs) remains as an open problem.

5.7 Improvement of the Game Model: A New Relaying Rule

If a packet of node 1 is accepted by node 2, node 1 is informed of this decision (i.e. the queue content of node 2) and has the opportunity to avoid transmitting in the next time slot to node 2 that has the increased probability of transmitting because of the recent packet arrival. In complex network scenarios, node 1 may receive only limited (or delayed) information on the acceptance decision of node 2 because of the possible interference effects of concurrent transmissions. However, the successful operation of wireless networks should be inherently based on the reliable exchange of control packets among the neighboring nodes for distributed control and self organization. Therefore, we assume that the decision of the relay node on whether to accept or reject a packet is immediately and correctly received by the corresponding transmitter node. This can be realized by dedicating a separate channel based on scheduled access to the feedback control packets that carry information on the channel outcome and packet forwarding decision of the relay node in the preceding

time slot. To eliminate the packet collisions that are more likely to happen after packet forwarding, we impose the additional rule that node 2 immediately forwards the packet received from node 1 to the destination, whereas node 1 waits for the next time slot independent of a new packet generation.

5.7.1 Modified State Definition and Mixed Stationary Strategies

We assume that nodes 1 and 2 agree on the new relaying rule as a form of co-operative packet scheduling (i.e. node 1 waits, while node 2 transmits the forwarded packet) to stimulate the cooperation of node 2 as well as to avoid the possible packet collisions. Then, the state of the game needs to specify further whether a packet at the queue of node 2 is a new generated one or has been forwarded before from node 2 in the previous time slot. We define the state of the game as (Q_1, Q_2), where $Q_1 \in \{0, 1\}$ and $Q_2 \in \{0, f, \overline{f}\}$. Specifically, we have $Q_2 = f$, if a packet of node 1 has been accepted to the queue of node 2 in the previous time slot, whereas $Q_2 = \overline{f}$ represents a new generated or already backlogged packet at queue of node 2. Thus, the state space has been extended to $\mathcal{K} = \{(0, 0), (0, f), (0, \overline{f}), (1, 0), (1, f), (1, \overline{f})\}$.

The action spaces $\mathcal{A}_1$ and $\mathcal{A}_2$ remain the same as defined in section 5.4. The stationary strategies of node 1 are $\underline{s}_{1,k}(A_1^1, A_2^1, A_3^1) = (p_1^1, p_2^1, 1 - p_1^1 - p_2^1)$ for $k \in \{(1, 0), (1, \overline{f})\}$, where p_1^1, p_2^1 and $1 - p_1^1 - p_2^1$ denote the probabilities of selecting actions A_1^1, A_2^1 and A_3^1, and $\underline{s}_{1,k}(A_1^1, A_2^1, A_3^1) = (0, 0, 1)$ for $k \in \{(0, 0), (0, \overline{f}), (0, f), (1, f)\}$. The random stationary strategy of node 1 is given by $\underline{s}_1 = (p_1^1, p_2^1)$.

Node 2 has the stationary strategies of $\underline{s}_{2,k}(A_1^2, A_2^2, A_3^2, A_4^2) = (p_1^2, 1 - p_1^2, 0, 1)$

for $k \in \{(0, \overline{f}), (1, \overline{f})\}$, where p_1^2, and $1 - p_1^2$ denote the probabilities of selecting actions A_1^2, A_2^2. We have $\underline{s}_{2,k}(A_1^2, A_2^2, A_3^2, A_4^2) = (1, 0, 0, 1)$ for $k \in \{(0, f), (1, f)\}$, and $\underline{s}_{2,k} = (0, 1, p_3^2, 1 - p_3^2)$ for $k \in \{(0, 0), (1, 0)\}$, where p_3^2, and $1 - p_3^2$ denote the probabilities of selecting actions A_3^2 and A_4^2. The random stationary strategy of node 2 is given by $\underline{s}_2 = (p_1^2, p_3^2)$.

5.7.2   Modified State Transition Matrix and Utility Functions

The joint set of the random stationary strategies of nodes 1 and 2 is given by $\underline{s} = (p_1^1, p_2^1, p_1^2, p_3^2)$. For $i = 1, ..., 6$, the ith row of the state transition matrix $\mathrm{T}(\underline{s})$, i.e. the row vector $\mathrm{T}_i(\underline{s})$, can be expressed as

$$\mathrm{T}_1(\underline{s}) = \mathrm{T}_2(\underline{s}) = [(1 - \lambda_1)(1 - \lambda_2), 0, (1 - \lambda_1)\lambda_2, \lambda_1(1 - \lambda_2), 0, \lambda_1\lambda_2],$$

$$\mathrm{T}_3(\underline{s}) = [(1 - \lambda_1)p_1^2(1 - \lambda_2), 0, (1 - \lambda_1)(1 - p_1^2 + p_1^2\lambda_2), \lambda_1 p_1^2(1 - \lambda_2), 0,$$

$$\lambda_1(1 - p_1^2 + p_1^2\lambda_2)],$$

$$\mathrm{T}_4(\underline{s}) = [p_1^1(1 - \lambda_1)(1 - \lambda_2), (1 - \lambda_1)p_2^1 p_3^2, p_1^1(1 - \lambda_1)\lambda_2,$$

$$(1 - \lambda_2)(1 - p_1^1 - p_2^1 + p_1^1\lambda_1 + p_2^1(1 - p_3^2)),$$

$$p_2^1 p_3^2\lambda_1, \lambda_2(1 - p_1^1 - p_2^1 + p_1^1\lambda_1 + p_2^1(1 - p_3^2))],$$

$$\mathrm{T}_5(\underline{s}) = [0, 0, 0, 1 - \lambda_2, 0, \lambda_2],$$

$$\mathrm{T}_6(\underline{s}) = [0, 0, p_1^1(1 - \lambda_1)(1 - p_1^2), (1 - p_1^1 - p_2^1)p_1^2(1 - \lambda_2), 0,$$

$$1 - p_1^1(1 - \lambda_1)(1 - p_1^2) - (1 - p_1^1 - p_2^1)p_1^2(1 - \lambda_2)].$$

If nodes 1 and 2 follow the strategy $\underline{s} = (p_1^1, p_2^1, p_1^2, p_3^2)$, the expected utilities

$u_1(\underline{s})$ and $u_2(\underline{s})$ are given by

$$u_1(\underline{s}) = \pi_4(\underline{s})[p_1^1(1 - E_{1,3}) + p_2^1(p_3^2 c - (1 - p_3^2)D - E_{1,2}) - (1 - p_1^1 - p_2^1)D] \qquad (5.9)$$

$$+\pi_5(\underline{s})[-D] + \pi_6(\underline{s})[p_1^1(-E_{1,3} + 1 - p_1^2 - p_1^2 D)$$

$$+p_2^1(-E_{1,2} - D) - (1 - p_1^1 - p_2^1)D],$$

$$u_2(\underline{s}) = \pi_2(\underline{s})[1 - E_{2,3}] + \pi_3(\underline{s})[p_1^2(1 - E_{2,3}) - (1 - p_1^2)D] + \pi_4(\underline{s})[-p_2^1 p_3^2 c] \qquad (5.10)$$

$$+\pi_5(\underline{s})[1 - E_{2,3}] + \pi_6(\underline{s})[p_1^2(-E_{2,3} + 1 - p_1^1 - p_2^1 - (p_1^1 + p_2^1)D) - (1 - p_1^2)D].$$

We can also consider an alternative cooperation stimulation mechanism, in which the transmitter node pays a reward of value c_1 to relay node, if relay node accepts the packet to forward to the destination node, and only the transmitter node receives a throughput reward of value 1, whenever the relay node successfully delivers the forwarded packet to the destination. Node 2 receives the reward c_1 immediately (i.e. node 1 incurs the cost c_1) and node 1 receives (throughput) reward 1 in the next slot such that the utilities obtained over two successive time slots are $1 - c_1$ and c_1 for transmitter and relay nodes, respectively. In the original reward mechanism, node 2 incurs the cost c after accepting a packet (i.e. node 1 receives the reward c) and node 2 receives the full throughout reward 1 alone in the next time slot such that the utilities obtained over two successive time slots are c and $1 - c$ for transmitter and relay nodes, respectively. Thus, the original and reverse mechanisms are equivalent with the same utility functions, if $c = 1 - c_1$. The only difference is that the reliable operation of the original and reverse mechanisms is based on the separate assumptions that node 1 or node 2 follows the improved relaying rule without any external stimulation.

5.7.3 Numerical Analysis of the Performance Improvement

In this section, we evaluate the effects of the modified game model on the expected utilities. We consider the system parameters $E_{1,2} = 0.1$, $E_{2,3} = 0.1$, $D = 0.1$ and $\lambda_1 = 0.25$. The results summarized in Table 5.2 verify the expected performance improvement. Since the packet arrival rate $\lambda_1 = 0.25$ is low, the particular case of packet collisions to be prevented by the proposed improvement is rare and the performance improvement is rather limited (especially for $\lambda_2 = 0$).

5.7.4 An Analytical Look at the Equilibrium Strategies

In this section, we follow an analytical approach to reveal (at least partially) the interactions among the equilibrium strategies and dependence on system parameters under the assumption of $\lambda_2 = 0$ for tractable solutions. This will also provide several insights on selecting the equilibrium strategies for the general case with arbitrary values of $0 \le \lambda_2 \le 1$.

From Eqs. (5.9) and (5.10), we can express the expected utilities for $\lambda_2 = 0$ as follows:

$$u_1(\underline{s}) = \pi_4(\underline{s})[p_1^1(1 - E_{1,3}) + p_2^1(p_3^2 c - (1 - p_3^2)D - E_{1,2} - p_3^2 \lambda_1 D) \quad (5.11)$$

$$-(1 - p_1^1 - p_2^1)D],$$

$$u_2(\underline{s}) = \pi_4(\underline{s})p_2^1 p_3^2(-c + 1 - E_{2,3}), \quad (5.12)$$

where $\pi_4 = \frac{\lambda_1}{p_1^1(1-\lambda_1)+p_2^1 p_3^2(1-\lambda_1+\lambda_1^2)+\lambda_1}$. The random strategy space does not include p_1^2, since there is neither backlogged nor new generated packet at the queue of node 2, i.e. $\pi_3(\underline{s}) = \pi_6(\underline{s}) = 0$. The utility $u_2(\underline{s})$ is monotonically increasing with

Table 5.2: Expected Equilibrium Utilities with the First Improvement.

$c = 0.5$	$\lambda_2 = 0$ $E_{1,3} = 0.1$	$\lambda_2 = 0.25$ $E_{1,3} = 0.1$	$\lambda_2 = 0$ $E_{1,3} = 0.2$	$\lambda_2 = 0.25$ $E_{1,3} = 0.2$
$u_{1,nc}^*$	0.1500	0.0946	0.1077	0.0851
$u_{2,nc}^*$	0	0.1260	0.0268	0.1310
$u_{1,nc}^* + u_{2,nc}^*$	0.1500	0.2206	0.1345	0.2161
$u_{1,c}^*$	0.1500	0.0924	0.1102	0.0819
$u_{2,c}^*$	0	0.1717	0.0351	0.1758
$u_{1,c}^* + u_{2,c}^*$	0.1500	0.2641	0.1453	0.2577
$c = 0.6$	$\lambda_2 = 0$ $E_{1,3} = 0.1$	$\lambda_2 = 0.25$ $E_{1,3} = 0.1$	$\lambda_2 = 0$ $E_{1,3} = 0.2$	$\lambda_2 = 0.25$ $E_{1,3} = 0.2$
$u_{1,nc}^*$	0.1500	0.1004	0.1088	0.0879
$u_{2,nc}^*$	0	0.1294	0.0289	0.1328
$u_{1,nc}^* + u_{2,nc}^*$	0.1500	0.2298	0.1377	0.2207
$u_{1,c}^*$	0.1500	0.0968	0.1113	0.0843
$u_{2,c}^*$	0	0.1673	0.0336	0.1734
$u_{1,c}^* + u_{2,c}^*$	0.1500	0.2641	0.1449	0.2577

p_3^2, if $1 - E_{2,3} > c$, and monotonically decreasing with p_3^2, if $1 - E_{2,3} < c$. To maximize $u_2(\underline{s})$, node 2 selects strategy $p_3^2 = 1$, if $1 - E_{2,3} > c$, or strategy $p_3^2 = 0$, if $1 - E_{2,3} < c$. An arbitrary value of p_3^2 is chosen, if $1 - E_{2,3} = c$, since we have $u_2(\underline{s}) = 0$ independent of p_3^2.

For $p_3^2 = 0$, i.e. if $1 - E_{2,3} < c$, the necessary condition for the pure strategy of direct communication, i.e. $p_1^1 = 1$ and $p_2^1 = 0$, is given by $\lambda_1(1 - E_{1,3}) > -D$, whereas the necessary condition for the strategy of waiting only, i.e. $p_1^1 = 0$ and

$p_2^1 = 0$, is $\lambda_1(1 - E_{1,3}) < -D$. Note that the pure strategy of the two-hop relaying, i.e. $p_1^1 = 0$ and $p_2^1 = 1$, is not feasible for any positive value of $E_{1,2}$. For $p_3^2 = 1$, i.e. if $1 - E_{2,3} > c$, the necessary condition for the pure strategy of direct communication is $\lambda_1(1 + \lambda_1^2)(1 - E_{1,3}) \geq \max(\lambda_1(c - E_{1,2} - \lambda_1 D), -D(1 + \lambda_1^2))$. The pure strategy of two-hop relaying is possible only if $\lambda_1(c - E_{1,2} - \lambda_1 D) \geq (1 + \lambda_1^2)\max(\lambda_1(1 - E_{1,3}), -D)$, whereas the necessary condition for the pure strategy solution of waiting is $D(1 + \lambda_1^2) \leq \lambda_1 \min(-c + E_{1,2} + \lambda_1 D, (1 + \lambda_1^2)(-1 + E_{1,3}))$.

For the full cooperative case, the strategy $p_3^2 < 1$ would only reduce the total system utility (by causing extra energy waste and packet delay that can be prevented by selecting $p_3^2 = 1$), since node 2 (with zero packet generation rate) cannot otherwise contribute to the total system utility. Thus, we must have $p_3^2 = 1$ under the assumption of full cooperation and the resulting total expected utility $u_\Sigma(\underline{s}) = u_1(\underline{s}) + u_2(\underline{s})$ is given by

$$u_\Sigma(\underline{s}) = \frac{\lambda_1(p_1^1(1 - E_{1,3} + D) + p_2^1(1 - E_{2,3} - E_{1,2} + D - \lambda_1 D) - D)}{p_1^1(1 - \lambda_1) + p_2^1(1 - \lambda_1 + \lambda_1^2) + \lambda_1}. \quad (5.13)$$

The condition for direct communication at the cooperative equilibrium is $\lambda_1(1 + \lambda_1^2)(1 - E_{1,3}) \geq \max(\lambda_1(1 - E_{2,3} - E_{1,2} - \lambda_1 D), -D(1 + \lambda_1^2))$. The pure strategy of two-hop relaying has the necessary condition of $\lambda_1(1 - E_{2,3} - E_{1,2} - \lambda_1 D) \geq (1 + \lambda_1^2)\max(\lambda_1(1 - E_{1,3}), -D)$, whereas we need $D(1 + \lambda_1^2) \leq \lambda_1 \min(-1 + E_{2,3} + E_{1,2} + \lambda_1 D, (1 + \lambda_1^2)(-1 + E_{1,3}))$ for the pure strategy solution of waiting.

The mixed strategy solutions might be possible depending on the complex interactions among the system parameters (e.g. there is no mixed strategy solution, if $p_3^2 = 0$). If the objective function to be maximized is expressed as $(k_1 p_1^1 + k_2 p_2^1 +$

$k_3)/(k_4 p_1^1 + k_5 p_2^1 + k_6)$, then a possible mixed strategy solution is $\underline{s} = \{p_1^1, p_2^1\} = \{(k_3 k_5 - k_2 k_6)/(k_2 k_4 - k_1 k_5), (k_1 k_6 - k_3 k_4)/(k_2 k_4 - k_1 k_5)\}$ for the values of p_1^1 and p_2^1 such that $p_1^1 \geq 0, p_2^1 \geq 0, p_1^1 + p_2^1 \leq 1$. For the general case of $\lambda_2 \geq 0$, we need lower values of c to stimulate the cooperation of node 2 and the results of this section can be referred as useful bounds to select the appropriate values of c.

5.7.5 Improvement by Immediate Transmissions of New Packets

A common rule in the distributed random MAC protocols (as implemented in [19] for slotted Aloha systems) is to transmit new packets immediately in the next possible time slot such that different strategies are needed for the transmissions of the new and backlogged packets. We can enforce the immediate transmissions of new generated packets to avoid the unnecessary packet delays. This might otherwise cause successive collisions in subsequent time slots, since nodes with new and backlogged packets are subject to channel conditions that are statistically different. Therefore, it is more efficient to allow them to select different distributions over their actions. This formulation requires an extended state space, as discussed in [40], and leads to performance gains, since delays in transmissions of new packets possibly lead to high levels of packet accumulation in both queues and consequently reduce the expected utilities. We consider the common parameters $E_{1,2} = 0.1$, $E_{2,3} = 0.1$, $D = 0.1$ and $\lambda_1 = 0.25$. The expected utilities given in Table 5.3 verify the improvement in both cases of the non-cooperative and social equilibrium.

Table 5.3: Expected Equilibrium Utilities with the First and Second Improvements

$c = 0.5$	$\lambda_2 = 0$ $E_{1,3} = 0.1$	$\lambda_2 = 0.25$ $E_{1,3} = 0.1$	$\lambda_2 = 0$ $E_{1,3} = 0.2$	$\lambda_2 = 0.25$ $E_{1,3} = 0.2$
$u_{1,nc}^*$	0.1500	0.0954	0.1101	0.0870
$u_{2,nc}^*$	0	0.1262	0.0275	0.1328
$u_{1,nc}^* + u_{2,nc}^*$	0.1500	0.2216	0.1376	0.2198
$u_{1,c}^*$	0.1500	0.0934	0.1102	0.0844
$u_{2,c}^*$	0	0.1723	0.0357	0.1794
$u_{1,c}^* + u_{2,c}^*$	0.1500	0.2647	0.1459	0.2638
$c = 0.6$	$\lambda_2 = 0$ $E_{1,3} = 0.1$	$\lambda_2 = 0.25$ $E_{1,3} = 0.1$	$\lambda_2 = 0$ $E_{1,3} = 0.2$	$\lambda_2 = 0.25$ $E_{1,3} = 0.2$
$u_{1,nc}^*$	0.1500	0.1017	0.1106	0.0890
$u_{2,nc}^*$	0	0.1378	0.0291	0.1376
$u_{1,nc}^* + u_{2,nc}^*$	0.1500	0.2395	0.1390	0.2266
$u_{1,c}^*$	0.1500	0.0973	0.1113	0.0863
$u_{2,c}^*$	0	0.1684	0.0346	0.1775
$u_{1,c}^* + u_{2,c}^*$	0.1500	0.2657	0.1459	0.2638

5.8   Summary and Conclusions

In this chapter, we extended the analysis of non-cooperative operation to multi-hop communication and looked at the problem of joint MAC and routing from the perspective of stochastic games in a simple ad hoc wireless network that consists of a transmitter and relay node with conflicting interests of transmitting to a common destination. As routing decisions, the transmitter node randomizes actions between

directly transmitting to the destination and relying on a relay node to forward packets, whereas the relay node decides on whether to accept packets from the transmitter node or to transmit its own packets to the destination. The decision whether to transmit or to wait represents the MAC strategies intertwined with the routing decisions of direct transmission or relaying over the intermediate node. We assumed the model of selfish nodes without any external enforcement for cooperation and stimulated packet forwarding by offering a future reward to the relay node for successfully delivering the forwarded packets in subsequent time slots. We presented a detailed comparison of the non-cooperative and social equilibrium strategies. For distributed operation with unknown system parameters, we developed an adaptive algorithm and pointed at the possible deviations from the equilibrium solutions for high energy costs. Finally, we improved the game model by possibly giving higher priority to the transmissions of the forwarded and new generated packets.

We can extend the model to incorporate arbitrary buffer capacities and designate the relay node also as a potential destination for packets of the transmitter node. This would extend the cooperation incentives beyond the reward mechanism for cooperation stimulation. The simple relay channel topology can be further generalized to a layered network model with nodes in different layers competing for the network resources to forward packets among different layers. The ultimate goal aims at the joint analysis of MAC and network layer operation as non-cooperative games under the assumptions of arbitrary network topology and packet traffic. This goal will be partially realized in Chapter 7 by the joint design of MAC and network coding (as extension of plain routing).

Chapter 6

Cross-Layer Design of MAC and Network Coding in Wireless Networks

We extend the plain routing operation at the network layer and consider wireless network coding in conjunction with MAC. It is known that coding over wired networks enables connections with rates that cannot be achieved by routing. However, the properties of wireless networks (such as omnidirectional transmissions, destructive interference, single transceiver per node) and the additional performance criteria (such as energy efficiency) modify the formulation of time-varying network coding in a way that reflects strong interactions with the underlying MAC protocols and deviates from the classical approach used in wired network coding. We propose to separately activate predetermined (conflict-free) network realizations in a time-division mechanism and derive the content of network flows through network coding so as to optimize performance measures such as achievable throughput and energy costs. We present a method to construct linear wireless network codes and discuss their properties under wireless assumptions and interactions with MAC schedules. We also outline how to operate network coding with arbitrary (e.g. contention-based) MAC protocols. Then, we obtain conflict-free transmission schedules jointly with network codes by decomposing the wireless network into subtrees and employing graph coloring on simplified subtree graphs. Finally, the performance of network

coding is compared with plain routing (in terms of throughput, energy and delay properties) in conjunction with underlying MAC solutions.

6.1   Introduction

As a network layer problem, classical routing involves simply replicating and forwarding the received packets by the intermediate relay nodes in multi-hop communication, whether in wired or wireless networks. Network coding extends routing by allowing relay nodes to combine the information received from multiple links in the subsequent transmissions. It is known that network coding in wired networks enables connections with rates that are higher than those achieved by plain routing only [60].

To extend network coding to a wireless network, we need to take into account the additional properties of omnidirectional transmissions, no simultaneous packet transmission and reception by any node, and possible destructive interference effects among concurrent transmissions (such that MAC, e.g. scheduling, is necessary to coordinate transmissions). These wireless communication properties introduce new cross-layer interactions that have not been addressed in the context of wired network coding. The interdependence of MAC and network coding (just like that of MAC and plain routing) requires their joint specification for efficient wireless network operation.

In this chapter, we extend network coding to operate in ad hoc wireless networks and study the cross-layer design possibilities of joint MAC and network coding

(or plain routing as a special case). There have been efforts [61, 62, 63, 64] to extend network coding to randomized environments with distributed implementation for wired networks. The problem of network coding in the presence of omnidirectional transmissions has been studied in [65] through the use of linear programs to optimize the network resources based on link costs. Decentralized solutions to achieve the minimum cost have been introduced in [66]. Omnidirectional transmissions have been modeled in [67] and [68] as hyperlinks (with additional constraint sets that can prevent nodes from transmitting and receiving packets simultaneously). Network coding for simple (tandem and grid) network topologies has been studied in [69] for the case of energy-efficient broadcast communication with omnidirectional transmissions.

The interference effects have been incorporated by [70] for joint optimization of MAC and network flows while preventing simultaneous transmission and reception by any node. The transmission-reception pairs are based on the SINR threshold criterion. This implies that a node would not be able to receive and transmit at the same time as the SINR at the receiver would very small, since the signal of the transmitter would dominate the interference. As a special case, network coding has been addressed in [71] for energy-efficient multicasting by considering only one isolated link at a time without interference effects. For energy efficiency purposes only, there is no need to consider interference effects and simultaneous transmissions should not be activated. However, if we consider the additional objectives such as throughput or delay efficiency, then network codes must be jointly designed with MAC.

We incorporate in the network coding analysis realistic wireless network properties such as omnidirectional transmissions, interference and delay effects, practical constraints of single transceiver per node and multiple performance criteria (such as throughput and energy). We eliminate explicitly from the sets of transmitter nodes those nodes that receive packets. Therefore, we can reduce the number of possible transmission sets considerably and thus, even though their number remains exponential, it reduces the complexity of any heuristics that allow their consideration in the pool of the possible transmission sets. In addition, we consider dynamic operation of multi-hop packet communication in contrast to previous models of connection-oriented traffic. This introduces an operational advantage for conflict-free transmission scheduling. Also, we avoid the use of link costs, and, rather, model the wireless performance measures such as throughput and energy consumption in terms of node (rather than link) costs due to omnidirectional transmissions. Our ultimate goal is to analyze and design wireless network codes in conjunction with conflict-free transmission schedules. Random network codes have been suggested before for wireless networks [68]. Instead, we construct deterministic wireless network codes jointly optimized with MAC schedules as solutions to a time-dependent flow optimization problem.

The first objective of the chapter is to evaluate the interplay between MAC schedules and network codes. Particularly, we discuss the time-varying properties of network codes under conflict-free scheduling and specify the interactions of MAC schedules (specifically time allocation) and wireless network codes. We formulate joint MAC and network coding as a flow optimization problem. For this purpose,

we consider a two-step solution: 1. Predetermine feasible (conflict-free) wireless network realizations, and assign minimum costs (e.g. transmission power) to each node for any network realization, 2. Assign time fractions to network realizations and choose the flows between transmitter-receiver pairs (addressed to different destinations) to optimize the performance objectives such as maximizing the multicast rate or minimizing the average cost (e.g. energy) achievable by network coding.

In this context, we introduce the notion of time-varying network codes (since nodes either encode and transmit packets or receive and decode packets or remain idle to avoid packet conflicts) and present a practical solution to derive linear network codes that are compatible with the wireless communication properties. We present an alternative formulation of wireless cuts to take into account the omni-directional transmissions and specify the conditions on network codes imposed by wireless network assumptions and conflict-free MAC schedules.

The second objective of the chapter is to use the results of network code assignments (in a possibly distributed analysis for scalable network operation) to jointly derive conflict-free transmission schedules. We follow the network coding framework of [72] based on the subtree decomposition that can support omnidirectional transmissions within each subtree network. Graph coloring is applied on a simplified subtree graph to assign the network codes while imposing conflict-free link scheduling within each subtree using three separate methods. Network coding is also extended to operate with arbitrary (e.g. contention-based) MAC protocols within each receiver's area through the application of the Group TDMA algorithm [73, 8]. Throughout this chapter, we consider the classical collision channel model.

The results can be also extended to the physical channel model.

The chapter is organized as follows. In section 6.2, we present a two-step basic method for joint network coding and conflict-free scheduling. We present in section 6.3 the method to construct wireless network codes over a given set of transmission schedules by taking the effects of omnidirectional transmissions into account. We also specify the properties of wireless network codes in conjunction with the underlying MAC protocol. Then, in section 6.4 we provide different methods to derive network codes and conflict-free link schedules in a joint analysis by applying graph coloring on simplified subtree representations of wireless networks. We present numerical results in section 6.5 that evaluate the throughput, energy efficiency and delay properties of network coding and plain routing methods in conjunction with the proposed MAC solutions. We draw conclusions in section 6.6.

6.2   Cross-Layer Design of Network Coding and MAC

6.2.1   An Example of Wireless Network Coding

We start with the wireless version of the classic network coding example that was used in [60], and illustrate the advantages of network coding over plain routing solutions under wireless multicasting for the network topology shown in Figure 6.1. We will extend the results to general wireless networks in subsequent sections. We compare two different strategies for multi-hop communication: (1) classical plain routing that limits nodes to act as forwarding switches, (2) network coding that allows nodes to code over the received information.
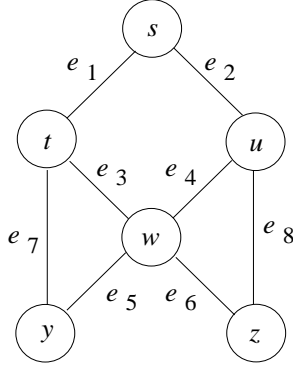
Figure 6.1: Wireless network example for network coding.

The objective of the wireless multicasting problem is to deliver packets originating at the source node s to both destination nodes y and z. We assume the classical collision channel model, although the results can be also extended to the SINR-based channel model. A packet transmission reaches all nodes connected through a single link (as shown in Figure 6.1) to the transmitter node. If multiple transmissions reach a node in the same time slot, a packet collision occurs. We schedule conflict-free transmissions to avoid the performance loss (resulting from the packet collisions) with respect to throughput, delay and energy efficiency. We assume that each packet contains one bit, the source s has always a packet to transmit, and each transmission consumes an amount of $\mathcal{E}_t$ energy units. The performance measures are: (1) λ: throughput per destination (average number of packets successfully delivered to each destination per unit time), (2) $\mathcal{E}_{avg}$: average energy consumed to deliver a packet to any destination, and (3) D_{avg}: average delay per packet. We denote by $c \vec{b} d$ the transmission of bit b from node c to node d. The throughput and energy-optimal plain routing solution schedules transmissions $s \vec{b} t$, $s \vec{b} u$ in odd time slots and $t \vec{b} y$, $u \vec{b} z$ in even time slots for any bit b. This plain routing solution

152

achieves $\lambda = \frac{1}{2}$ bits/slot, $D_{avg} = 2$ time slots and $\mathcal{E}_{avg} = \frac{3}{2}\mathcal{E}_t$ energy units per bit.

The conflict-free transmissions for optimal network coding are shown in Table 6.1 and have the period of three slots (after the initial time slot). A simple form of linear network coding consists of node w performing the bit additions $b_{2k-1} + b_{2k}$ at time slots $3k+1$, $k = 1, 2, ...$, and using the wireless multicast advantage to send the bit sum to nodes y and z in a single transmission. Since bits b_{2k-1} and b_{2k} have been delivered to nodes y and z in the previous transmissions, nodes y and z can combine b_{2k-1} with $b_{2k-1} + b_{2k}$ to decode b_{2k} and can combine b_{2k} with $b_{2k-1} + b_{2k}$ to decode b_{2k-1}. As time evolves, λ approaches $\frac{2}{3}$ bits/slot, D_{avg} approaches $\frac{13}{4}$ time slots and $\mathcal{E}_{avg}$ approaches $\frac{5}{4}\mathcal{E}_t$ energy units per bit. Since multiple input and output links (for

Table 6.1: Network Coding Solution for the Network in Figure 6.1.

Time Slot	1	2	3	4	5	6	7	8
Transmission	$s \overrightarrow{b_1} t$	$t \overrightarrow{b_1} w$	$u \overrightarrow{b_2} w$	$w \overrightarrow{b_1 + b_2} y$	$t \overrightarrow{b_3} w$	$u \overrightarrow{b_4} w$	$w \overrightarrow{b_3 + b_4} y$	$t \overrightarrow{b_5} w$
Schedule		$t \overrightarrow{b_1} y$	$u \overrightarrow{b_2} z$	$w \overrightarrow{b_1 + b_2} z$	$t \overrightarrow{b_3} y$	$u \overrightarrow{b_4} z$	$w \overrightarrow{b_3 + b_4} z$	$t \overrightarrow{b_5} y$
		$s \overrightarrow{b_2} u$	$s \overrightarrow{b_3} t$		$s \overrightarrow{b_4} u$	$s \overrightarrow{b_5} t$		$s \overrightarrow{b_6} u$

Table 6.2: Performance Comparison of Plain Routing and Network Coding for the Network in Figure 6.1.

Performance Measure	λ	$\mathcal{E}_{avg}$	D_{avg}
Best Plain Routing Solution	$\frac{1}{2}$ bit/slot	$\frac{3}{2}\mathcal{E}_t$/bit	2 slots
Network Coding Solution in Table 6.1	$\frac{2}{3}$ bits/slot	$\frac{5}{4}\mathcal{E}_t$/bit	$\frac{13}{4}$ slots

node w) are not simultaneously possible in wireless communication, packets need to

153

be stored (by node w) over successive time slots to perform network coding. This is the root cause of the increase in average packet delay. The throughput, energy cost and delay measures are compared in Table 6.2 for the best plain routing strategy and the network coding strategy presented in Table 6.1.

Network coding involves the extra processing cost for coding and decoding. To quantify this, we introduce $\mathcal{E}_c$ as the energy cost for each of the network coding and decoding operations, i.e. the energy cost of binary addition. The cost $\mathcal{E}_c$ is separately incurred at node w and destinations (y and z) at time slots $3k + 1$ for positive integer values of k. The extra coding/decoding energy cost of the network coding solution is $\mathcal{E}_c$ per time slot and $\frac{3\mathcal{E}_c}{4}$ per successfully decoded bit. If we scale this cost with respect to the transmission energy cost $\mathcal{E}_t$, we obtain $\frac{5\mathcal{E}_t + 3\mathcal{E}_c}{4}$ as the energy cost per bit of the network coding solution compared to the total transmission energy cost of $\frac{3\mathcal{E}_t}{2}$ of the plain routing solution. Hence, network coding is preferable, if $\mathcal{E}_c < \frac{\mathcal{E}_t}{3}$.

6.2.2 Joint Design of Network Coding and Conflict-free Scheduling

We assume that a single source s wishes to send packet traffic with common rate λ to each node in the destination set $\mathcal{D}$. We use the fact that network coding achieves maximum flows [60] and convert the network coding problem to a flow optimization problem, as done before for wired networks in [65] and extended to wireless operation in [67, 68, 70] according to a continuous flow model. The constraint of no simultaneous transmission and reception by any node and the application of

scheduling to eliminate packet conflicts impose time-dependent network flows on links. The total flow on link (i,j) at time slot t is defined as $z_{i,j}(t)$. The fraction of $z_{i,j}(t)$ destined to destination node $d \in \mathcal{D}$ is defined as $x_{i,j}(d,t)$.

We can formulate the problem of selecting flows $z_{i,j}(t)$ and $x_{i,j}(d,t)$ to maximize the achievable throughput per destination (averaged over all time slots) or minimize an average cost criterion (e.g. energy). The resulting formulation involves optimization at each time slot and is rater complex. Instead, we periodically activate distinct network realizations (i.e. conflict-free link sets) over non-overlapping time intervals. Then, we consider flow optimization problem individually for each network realization rather than for each time slot.

6.2.2.1 Network Flow Optimization

Let $P_{i,j}$ denote the transmission power required by node i to reach node j over a single hop (Note that $P_{i,j} = \infty$ means that there is no direct link possible from node i to node j.). We define $N^f = \{N_m^f\}_{m=1}^M$ as the set of M feasible network realizations, where the network realization $N_m^f = (V_m^f, E_m^f)$ with node set V_m^f and conflict-free link set E_m^f is allocated for τ_m fraction of the total time over time interval $\mathcal{T}_m$ such that the flows on link (i,j) are $z_{i,j}(t) = z_{i,j}^{(m)}$ and $x_{i,j}(d,t) = x_{i,j}^{(m)}(d)$ and the transmission power of node i is $P_i(t) = P_i^{(m)}$ for all $t \in \mathcal{T}_m$. The network realizations ensure conflict-free transmissions on all activated links and partition the network into the disjoint sets of transmitters and receivers at any time slot. The construction of network realizations also ensures omnidirectional transmissions (The

effects of omnidirectional transmissions on wireless flows will be taken into account in section 6.2.2.3.). In section 6.2.2.2, we will discuss how to construct an adequate set of network realizations.

For any $d \in \mathcal{D}$, $i \in V$, $j \in V$ and $m = 1, ..., M$, we have the following network flow conditions for wireless network realizations $\{N_m^f\}_{m=1}^M$:

$$x_{i,j}^{(m)}(d) \in \{0,1\}, z_{i,j}^{(m)} \in \{0,1\}, x_{i,j}^{(m)}(d) \leq z_{i,j}^{(m)}, z_{i,j}^{(m)} = \begin{cases} 1, & (i,j) \in E_m^f \\ 0, & \text{else} \end{cases}, \qquad (6.1)$$

$$(i,j) \in E_m^f \text{ if and only if } P_i^{(m)} \geq P_{i,j} \text{ and } P_k^{(m)} < P_{k,j}, k \neq i \text{ and } P_j^{(m)} = 0, \qquad (6.2)$$

$$\sum_{m=1}^M \tau_m \left(\sum_{j:(i,j) \in E_m^f} x_{i,j}^{(m)}(d) - \sum_{j:(j,i) \in E_m^f} x_{j,i}^{(m)}(d) \right) = \begin{cases} \lambda, & i = s \\ -\lambda, & i \in \mathcal{D} \\ 0, & \text{else} \end{cases}, i \in V, d \in \mathcal{D}. \, (6.3)$$

We propose a two-step solution to this time-dependent flow optimization problem of joint MAC and network coding:

Step 1: Predetermine the conflict-free network realizations $N^f = \{N_m^f\}_{m=1}^M$ with the minimum transmission power assignments $\{P_i^{(m)}\}_{i \in V}$ that ensure conflict-free transmissions for each network realization N_m^f. The flow $z_{i,j}^{(m)}$ for any node pair i and j is uniquely determined by the choice of N_m^f and $\{P_i^{(m)}\}_{i \in V}$ subject to condition given by (6.2).

Step 2: Assign the time fractions $\{\tau_m\}_{m=1}^M$ to network realizations $\{N_m^f\}_{m=1}^M$ and choose the flows $x_{i,j}^{(m)}(d)$ (subject to conditions given by (6.1) and (6.3)) through network coding (or plain routing as a special case) in order to either

(I) maximize the achievable throughput per destination λ, or

(II) minimize the average cost $a = \sum_{m=1}^M \tau_m a^{(m)}$ for fixed λ, where $a^{(m)}$ is

156

the total node cost incurred during network realization N_m^f (Particularly, we have $a^{(m)} = \sum_{i \in V} P_i^{(m)}$, if we consider transmission energy cost $P_i^{(m)}$ for node i in network realization N_m^f.), or

(III) minimize the average cost per successfully decoded packet, namely the quantity $\frac{a}{\lambda |\mathcal{D}|}$.

We will show in section 6.3 how to derive wireless network codes in order to achieve the maximum flows resulting from the solution to the presented flow optimization problem.

The optimal solution requires extensive search over all network realizations that result in conflict-free transmissions through partitioning of the nodes into the disjoint sets of transmitters and receivers. When we consider all possible sets of network partitioning and choose the best schedule (along with network codes), then the result will be globally optimal but will not scale well with increasing number of nodes in the network. Therefore, we use the following greedy heuristic to determine a distinct set of network realizations.

6.2.2.2 Step 1 (Construction of Conflict-Free Network Realizations)

We start by constructing the first network realization; we choose a node arbitrarily as receiver, and designate as the corresponding transmitter the node with the smallest power to reach the chosen receiver node. We then choose arbitrarily as the second receiver node one that has not been chosen so far as either transmitter or receiver. Similarly, we designate as its transmitter a node that has not been previously

chosen as receiver and has the smallest power to reach the second receiver. We admit this new transmitter-receiver pair provided that the activation of this link does not destructively interfere with the previously admitted transmitter-receiver pairs. For that purpose, we only need to check whether the transmission range of the chosen transmitter includes a non-intended receiver that has been previously activated. If so, then we choose another transmitter node and run the same admissibility check algorithm. We proceed in this fashion and determine the transmitter-receiver pairs until no link can be admitted without distorting the already admitted conflict-free link assignments.

Subsequently, we repeat the same procedure by choosing as receiver a node previously designated as transmitter and running the same algorithm to determine the complete set of network realizations $N^f = \{N^f_m\}^M_{m=1}$, until each node is designated as transmitter and receiver at least once (with the exception of the source node that should be only activated as transmitter). Each network realization N^f_m partitions the nodes into the disjoint transmitter and receiver sets $T^{(m)}$ and $R^{(m)}$, respectively. The link set E^f_m is conflict-free, if the condition (6.2) holds for all flows on links $(i, j) \in E^f_m$ such that $i \in T^{(m)}$ and $j \in R^{(m)}$. For the network realization N^f_m, the power cost of node i is given by $P^{(m)}_i = \max_{j:(i,j)\in E^f_m} P_{i,j}$.

6.2.2.3 Step 2 (Time Allocation for MAC Schedules)

The next problem is to find the time fraction τ_m allocated to each network realization τ_m. We define $C(s, y)$ as the set of cuts between nodes s and y and
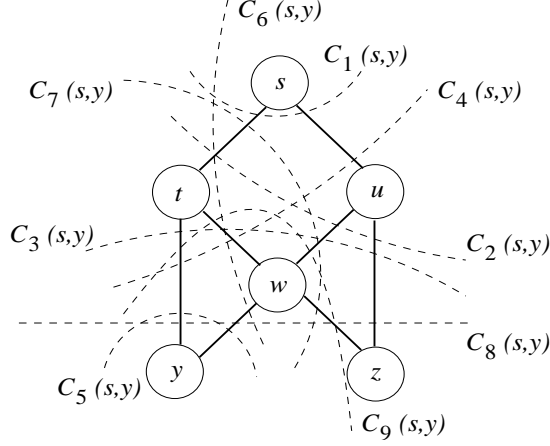
Figure 6.2: Example of cuts between nodes s and y for the network in Figure 6.1.

$C_i(s, y)$ as the ith cut in $C(s, y)$. Figure 6.2 depicts some of the possible cuts between the source s and the destination y of the network in Figure 6.1. For wired networks, the value of a cut is the sum of the capacities of the links that cross the given cut. For wireless networks, we introduce the average cut value $c_i^{N^f}(s, y)$, which is the maximum number of successful transmissions (time-averaged over all network realizations in N^f) across the cut $C_i(s, y)$ per unit time. To incorporate omnidirectional transmissions, the contribution of a node to any cut is limited to at most <u>one</u> per unit time, since a single node can transmit at most one distinct information packet over any cut at any time slot. Allowing higher values of cuts to reflect broadcast advantages creates complications or necessitates the introduction of artificial nodes as done by [70].

The maximum flow from the source s to destination nodes is given by the Max-flow Min-cut Theorem as

$$\lambda = \max_{\{\tau_m,\, m=1,\dots,M\}} \left(\min_{d \in \mathcal{D}} \left(\min_{k:C_k(s,d) \in C(s,d)} c_k^{N^f}(s, d) \right) \right). \tag{6.4}$$

159
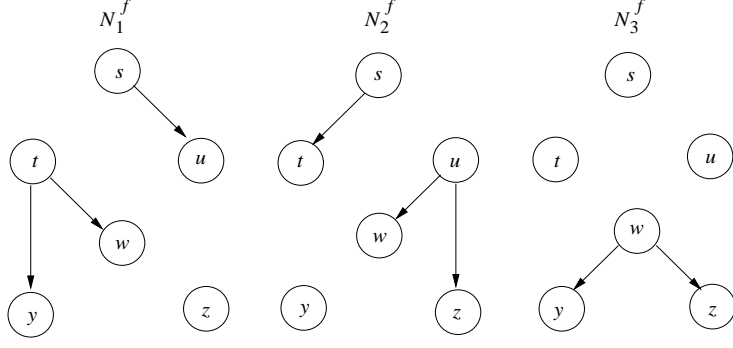
Figure 6.3: Wireless network realizations $N^f = \{N_1^f, N_2^f, N_3^f\}$ for the network in Figure 6.1.

We define $a^{(m)}$ as the total (e.g. transmission energy) cost for network realization N_m^f. The average cost over all network realizations is given by $a = \sum_{m=1}^{M} \tau_m a^{(m)}$.

We consider alternative optimization problems of selecting $\{\tau_m\}_{m=1}^{M}$ in order to either maximize the achievable throughput per destination λ, or minimize the average cost a for fixed throughput per destination λ, or minimize the average cost per successfully decoded packet $\frac{a}{\lambda |\mathcal{D}|}$.

6.2.2.4 Example for Cross-Layer Design

One adequate set of conflict-free network realizations for the network in Figure 6.1 is shown in Figure 6.3. Figure 6.2 depicts some of the cuts between nodes s and y with the following cut values: $c_1^{N^f}(s, y) = \tau_1 + \tau_2$, $c_2^{N^f}(s, y) = 2\tau_2$, $c_3^{N^f}(s, y) = \tau_1 + \tau_2$, $c_4^{N^f}(s, y) = 2\tau_1$, $c_5^{N^f}(s, y) = \tau_1 + \tau_3$, $c_6^{N^f}(s, y) = \tau_2 + \tau_3$, $c_7^{N^f}(s, y) = 2\tau_2$, $c_8^{N^f}(s, y) = \tau_1 + \tau_2 + \tau_3$ and $c_9^{N^f}(s, y) = \tau_1 + \tau_2$. The maximum flow from the source s to any destination y or z is given by $\min(2\tau_1, 1 - \tau_1)$ for the case of $\tau_1 = \tau_2$, and maximized by the time allocation $\tau_m = \frac{1}{3}$, $m = 1, 2, 3$. This is equivalent to the

network coding solution in Table 6.1. For plain routing, we can remove coding node w, as shown in section 6.2.1, such that the activated links in two network realizations are $E_1^f = \{(s,t), (s,u)\}$ and $E_2^f = \{(t,y), (u,z)\}$. Then, the two cuts that separate source s from destination y or z have the values τ_1 and τ_2, where $\sum_{i=1}^{2} \tau_i = 1$. The throughput-optimal plain routing solution activates two network realizations over disjoint time fractions with equal lengths $\tau_i = \frac{1}{2}$, $i = 1, 2$.

Assume unit energy cost for each transmission. The energy-optimal network coding solution employs the network realizations in Figure 6.3 with the transmission energy costs $a^{(1)} = 2$, $a^{(2)} = 2$ and $a^{(3)} = 1$. This network coding solution achieves the average transmission energy cost $a = 1 + 2\tau_1$ for the achievable throughput per destination $\lambda = \min(2\tau_1, 1 - \tau_1)$. The average energy cost per successfully decoded packet $\mathcal{E}_{avg} = \frac{a}{2\lambda}$ is minimized to $\frac{5}{4}$ energy units per packet and the throughput per destination r is maximized to $\frac{2}{3}$ packets per time slot, if we choose $\tau_m = \frac{1}{3}$, $m = 1, 2, 3$. On the other hand, the best plain routing solution is equivalent for the throughput and energy-optimal operations and can only achieve the minimum value of $\mathcal{E}_{avg} = \frac{3}{2}$ energy units per packet and the maximum value of $\lambda = \frac{1}{2}$ packets per time slot.

6.3   Construction of Wireless Network Codes

Next, we present a systematic method to derive network codes depending on the MAC schedules that result as solutions to the flow optimization problem considered in section 6.3.
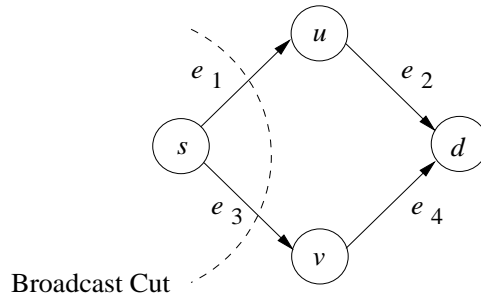
Figure 6.4: Example to illustrate the difference between the wired and wireless information flows.

6.3.1 Wireless Network Flows with Omnidirectional Transmissions

Consider the simple network topology of Figure 6.4. Source node s wishes to deliver packets to destination node d using the relay nodes u and v. If this was a wired network, the cut that separates s from the rest of the network (namely the broadcast cut) could carry the maximum flow of 2 packets per unit time slot (if each link has the capacity of 1 packet per unit time). According to the Max-flow Min-cut Theorem, the maximum flow is upper-bounded by the minimum cut value, which would be 2 in this case. If source transmits two disjoint packet streams error-free over two alternative paths $\{e_1, e_2\}$ and $\{e_3, e_4\}$, then the network can carry the maximum flow of 2 packets per unit time.

If we use the given topology in a network with broadcast transmissions as in a wireless network (still with error-free links), then the source node can transmit simultaneously to the two relay nodes u and v at the same time. Thus, the minimum cut is 2, although at most one distinct packet can be delivered to the destination resulting in the throughput value of 1 packet per unit time, since the original Max-

flow Min-cut Theorem inherently assumes that nodes can carry distinct information over the outgoing links from any node. The value of a cut was defined as the sum of the capacities of the links that cross the given cut. However, the wireless cut value needs to be formulated such that contribution of a node to any cut is at most one (that is equal to the value of the distinct information that can be transmitted by any node over each cut) due to the broadcast nature of transmissions (that have been also addressed in [74] for erasure networks). This is a standard application of the information-theoretic cut-set bounds [75] and introduces conceptual simplicity to joint design of network codes and conflict-free transmission schedules. In addition, wireless formulation of broadcast cuts enables node-based characterization of flows rather than a link-based one and node-based costs associated with each transmission (e.g. energy cost) instead of using link-based costs (as proposed in [65, 66]).

6.3.2 Basic Method for Constructing Wireless Network Codes

In wireless networks, nodes encode and transmit packets or receive and decode packets at different time instants. Hence, we need time-varying network coding that distinguishes when information is generated or received and when information is encoded or decoded at any node. We consider linear network codes [76, 77] and extend them to wireless network operation.

If we impose omnidirectional transmissions on the network flows such that all links out of a single node carry the same information at any time slot, we need to consider node-based (rather than link-based) network encoding. For a link $e =$
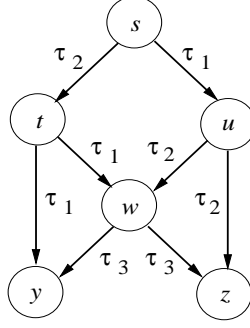
Figure 6.5: The hypothetical wired network graph N^g.

(i, j), we define $head(e)$ as j and $tail(e)$ as i. The flow out of node v in network realization N_m^f is denoted by $Y^{(m)}(v)$. For transmissions in network realization N_n^f, the coefficient $\beta_e^{(m,n)}$ weights the flow $Y^{(m)}(tail(e))$ that has arrived on link e in network realization N_m^f onto each of the links out of node $head(e)$ and the coefficient $\alpha_j^{(m,n)}(v)$ weights the jth flow $X_j^{(m)}(v)$, $j = 1, ..., \mu(v)$, that has been generated by node v in network realization N_m^f onto each of the links out of node v. The coefficient $\epsilon_k^{(m,n)}(e)$ weights the flow $Y^{(m)}(tail(e))$ that has arrived on link e in network realization N_m^f to reconstruct the kth flow $Z_k^{(n)}(head(e))$ in network realization N_n^f. The flow $Y^{(n)}(v)$ is encoded as

$$Y^{(n)}(v) = \sum_m \left(\sum_{u:(u,v) \in E_m^f} \beta^{(m,n)}((u,v)) Y^{(m)}(u) + \sum_{j=1}^{\mu(v)} \alpha_j^{(m,n)}(v) X_j^{(m)}(v) \right).$$

The kth output flow is decoded by node $v \in V$ in network realization N_n^f as

$$Z_k^{(n)}(v) = \sum_m \sum_{u:(u,v) \in E_m^f} \epsilon_k^{(m,n)}((u,v)) Y^{(m)}(u) . \tag{6.5}$$

We will discuss the general constraints imposed by wireless properties on coding coefficients in section 6.3.4. Assume that the wireless network realizations $N^f = \{N_m^f\}_{m=1}^M$ are predetermined (as in section 6.2.2.2) and the time alloca-

tions $\{\tau_m\}_{m=1}^M$ are specified (as in section 6.2.2.3). We define a hypothetical connected (wired) network $N^g = (V^g, E^g)$, as shown in Figure 6.5, with the node set $V^g = \bigcup_{m=1}^M V_m^f$ and link set $E^g = \bigcup_{m=1}^M E_m^f$. The capacity of any link out of node $i \in V^g$ is equal to $c_i^{(m)} = \tau_m$, if that link is activated during network realization N_m^f.

Figure 6.5 depicts N^g for the network realizations in Figure 6.3. The value of a cut in the network N^g is the sum of the capacities of the links that cross the given cut. For a wireless network, the value of a cut is the maximum number of successful transmissions across the given cut time-averaged over all (predetermined) network realizations. To model omnidirectional transmissions, the contribution of a node to any cut is limited to the value of at most <u>one</u> (per unit time) for both N^f and N^g.

Lemma 6.3.1 *The wireless network realizations N^f and the hypothetical wired network graph N^g have the same cut values and the same maximum flows.*

Proof: The proof follows directly from the construction of the graph N^g and the definition of the cut values. We denote by $C(s,d)$ the set of cuts between nodes s and d. We define $c_k^{N^g}(s,d)$ and $c_k^{N^f}(s,d)$ as the values of cut $C_k(s,d) \in C(s,d)$ (between nodes s and d) for the wired network graph N^g and for the wireless network realizations $N^f = \{N_m^f\}_{m=1}^M$ (with time allocations $\{\tau_m\}_{m=1}^M$), respectively. Since we assign the value of 1 to a cut that has multiple concurrent links out of a single node, we define the cut values as follows:

$$c_k^{N^g}(s,d) = \sum_{i \in V^g} \sum_{m=1}^M 1\left(i \rightarrow_{(m)} C_k(s,d)\right) c_i^{(m)}, \qquad (6.6)$$

$$c_k^{N^f}(s,d) = \sum_{m=1}^M \tau_m \sum_{i \in T^{(m)}} 1\left(i \rightarrow_{(m)} C_k(s,d)\right), \qquad (6.7)$$

165

where $i \rightarrow_{(m)} C_k(s, d)$ is the event that at least one link out of node i crosses the cut $C_k(s, d)$ during network realization N_m^f and $1(\cdot)$ is the indicator function that assigns the value 1 to a true statement and the value 0 to a false statement. Eq. (6.6) can be rewritten as Eq. (6.8) by using the definition of $c_i^{(m)} = \tau_m 1 (i \in T^{(m)})$, where $T^{(m)}$ is the set of transmitters activated during network realization N_m^f:

$$c_k^{N^g}(s, d) = \sum_{i \in V^g} \sum_{m=1}^{M} 1(i \rightarrow_{(m)} C_k(s, d)) \tau_m 1(i \in T^{(m)}) \tag{6.8}$$

$$= \sum_{m=1}^{M} \tau_m \sum_{i \in V^g} 1(i \rightarrow_{(m)} C_k(s, d)) 1(i \in T^{(m)}) \tag{6.9}$$

$$= \sum_{m=1}^{M} \tau_m \sum_{i \in T^{(m)}} 1(i \rightarrow_{(m)} C_k(s, d)). \tag{6.10}$$

Eq. (6.9) follows from exchanging the sum of cut capacities over transmitting nodes and network realizations. Eq. (6.10) follows from $T^{(m)} \subseteq V^g = \bigcup_{m=1}^{M} V_m^f$. Note that Eq. (6.10) is equivalent to $c_k^{N^f}(s, d)$ in Eq. (6.7). Thus, the wireless network realizations N^f and the corresponding hypothetical wired network graph N^g have the same cut values. According to the Max-flow Min-cut Theorem, the maximum flow between node s and destination set $\mathcal{D}$ is $\min_{d \in \mathcal{D}} \min_{k:C_k(s,d) \in C(s,d)} c_k^{N^f}(s, d)$ for wireless network realizations N^f and is $\min_{d \in \mathcal{D}} \min_{k:C_k(s,d) \in C(s,d)} c_k^{N^g}(s, d)$ for network graph N^g. Since networks N^f and N^g have the same cut values, they have the same properties of the maximum flows between any source-destination node pair s and d. □

Assume that we would like to solve the problem of omnidirectional transmissions while keeping the "link-based" notion of the connectivity graphs with the original definition of flows and cuts. A possible solution to incorporate omnidirec-
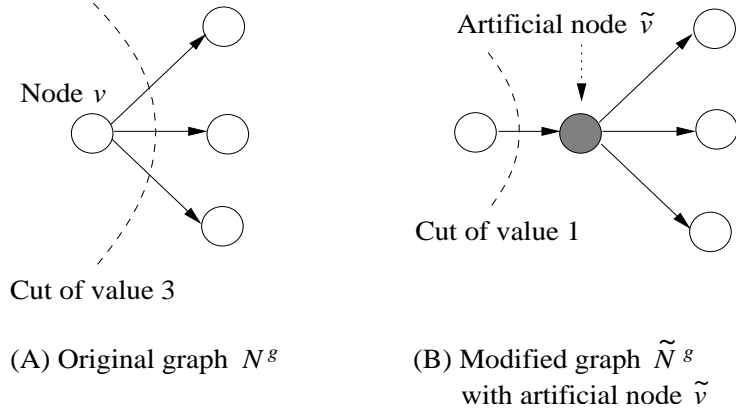
Figure 6.6: Illustrative example of introducing an artificial node $\widetilde{v}$ for each transmitting node v.

tional transmissions [70] is to introduce an artificial node for each transmitting node and connect the original node to that artificial node with an error-free link of capacity 1 (information flow per unit time) such that this artificial node has the same outgoing links as the original node, as shown in Figure 6.6. The idea is to separate the coder from transmitter so that we have the maximum flow of 1 over any cut crossed by the outgoing links of any transmitting node.

Suppose that we introduce the artificial node $\widetilde{v}$ for transmitter node v with outgoing links E_v on the original graph N^g. On the modified graph $\widetilde{N}^g$, the only link out of node v is $(v, \widetilde{v})$, whereas the links out of the artificial node $\widetilde{v}$ are equivalent to links E_v. For each cut on the graph N^g that is crossed by links E_v, we obtain a new cut on the modified graph $\widetilde{N}^g$ that is crossed by link $(v, \widetilde{v})$ but not by links E_v. This is illustrated by Figure 6.6. Thus, the contribution of node v to the new cut on $\widetilde{N}^g$ is 1. Since we are interested in achieving the maximum flow through the cut with the minimum value, the artificial nodes have the same effect on $\widetilde{N}^g$

as assigning the value of 1 to a cut on N^g that has multiple links out of a single node. As far as the flows are concerned, we can either use (a) the original graph N^g with the cut value definitions based on omnidirectional transmissions or use (b) the modified graph $\widetilde{N}^g$ with artificially introduced nodes. Both formulations lead to the same (correct) generalization of the minimum cut capacity notion for wireless networks. However, the formulation (b) cannot handle MAC scheduling, whereas the formulation (a) has the advantage of offering an operational difference in that it facilitates the definition of schedules and can also incorporate the node-based metrics (such as energy expenditures) in wireless networks.

Next, we construct the wireless network codes using the linear network codes that have been determined for the hypothetical wired network graph N^g.

Theorem 6.3.1 *There exist wireless linear network codes that achieve the maximum flows for a given set N^f of wireless network realizations.*

Proof: There are low complexity polynomial-time algorithms [78] available to provide linear network codes on the corresponding hypothetical wired network graph N^g. Suppose that we find linear network codes on network graph N^g that achieve the maximum flow $\min_{d \in \mathcal{D}} \min_{k:C_k(s,d) \in C(s,d)} c_k^{N^g}(s,d)$ between the source node s and destination nodes $d \in \mathcal{D}$. Note that the cut value achievable on network graph N^g (with respect to the definition of cuts based on omnidirectional transmissions) can be smaller than what would be achievable on the same topology with respect to the traditional definition of cuts. However, the minimum cut value on N^g is actually equal to the value of the maximum flow achievable over the wireless network

realizations N^f, as Lemma 6.3.1 indicates.

For the simplicity of illustration, we assume the finite field $\mathcal{F}_2$ for the network coding operations. The notation for wired network codes is the same as the notation for wireless network codes with the exceptions that coding and decoding coefficients of wired network codes do not involve time indices and wired network codes have link-based properties. The flow $Y(e')$ that arrives on link e' is weighted by $\beta(e', e)$ onto link e. For $j = 1, ..., \mu(tail(e))$, the input flow $X_j(tail(e))$ generated by node $tail(e)$ is weighted by $\alpha_j(e)$ onto link e. The incoming flow $Y(e)$ on link e is weighted by $\epsilon_k(e)$ to reconstruct the kth flow $Z_k(v)$ for destination node $v = head(e)$.

To minimize the packet delay, we can assume that a packet to be transmitted during the network realization N_n^f is generated during the previous network realization N_m^f, where $n = m + 1 \pmod{M}$, i.e. we have $\alpha_j^{(m,n)}(v) = 0$ for all $n \neq m + 1 \pmod{M}$. Nodes store the arriving packets and decode any flow only after receiving all the necessary information. Therefore, we can impose the constraint $\epsilon_k^{(m,n)}(e) = 0$ unless $n = \max(p \in \{1, ..., M\} : \epsilon_k(e) = 1, e \in E_p^f)$, i.e. any flow k should be decoded only after activating once each of the network realizations that can deliver all of the information packets necessary to decode the flow k. This ensures that nodes accumulate all the necessary information for decoding purposes at the expense of possibly increasing the packet delay compared to the classical plain routing approach. Since no node can transmit and receive packets in the same network realization, we also need to impose the constraint $\beta^{(m,n)}(e) = 0$ for any link e, if $m \neq n$. We construct the non-zero wireless network codes from the linear network

codes predetermined for the hypothetical wired network graph N^g as follows

$$\alpha_j^{(m,n)}(tail(e)) = 1, \text{ if } \alpha_j(e) = 1, e \in E_n^f \text{ and } n = m + 1 \,(\text{mod } M), \quad (6.11)$$

$$\beta^{(m,n)}(e') = 1, \text{ if } \beta(e',e) = 1, e' \in E_m^f, e \in E_n^f \text{ and } m \neq n, \quad (6.12)$$

$$\epsilon_k^{(m,n)}(e) = 1, \text{ if } \epsilon_k(e) = 1, e \in E_m^f \text{ and } n = \max\{p : \epsilon_k(e) = 1, e \in E_p^f\}. \quad (6.13)$$

Averaged over the entire set of network realizations N^f, the wireless network codes perform exactly the same encoding and decoding operations as wired network codes for the hypothetical wired network graph N^g, and achieve the maximum flows for wireless network realizations N^f. $\qquad\square$

The wireless network codes are constructed such that:

(a) Nodes accumulate generated or incoming packets over at most one schedule period,

(b) Nodes perform the same encoding and decoding operations as those for the hypothetical wired network (Particularly, each source or relay node sends the same linear combination to the same neighbor node as determined by wired network codes but only during the network realization for which the node is activated as transmitter and the neighbor node is activated as receiver.), and

(c) Nodes either encode and transmit or receive and decode packets during the network realization for which they are activated as transmitters or receivers.

6.3.3 Example for Constructing Wireless Network Codes

To illustrate how to construct wireless network codes, we consider the network example in Figure 6.1 with the network realizations shown in Figure 6.3. First, we

construct the wired network graph N^g and find the non-zero linear network codes on N^g. The non-zero coefficients of linear network coding solution for the finite field $\mathcal{F}_2$ are given by $\alpha_1(e_1) = 1$, $\alpha_2(e_2) = 1$ and $\beta(e_1, e_7) = 1$, $\beta(e_1, e_3) = 1$, $\beta(e_2, e_4) = 1$, $\beta(e_2, e_8) = 1$, $\beta(e_3, e_5) = 1$, $\beta(e_3, e_6) = 1$, $\beta(e_4, e_5) = 1$, $\beta(e_4, e_6) = 1$ and $\epsilon_1(e_6) = 1$, $\epsilon_1(e_7) = 1$, $\epsilon_1(e_8) = 1$, $\epsilon_2(e_5) = 1$, $\epsilon_2(e_7) = 1$, $\epsilon_2(e_8) = 1$. Using the method given in the proof of Theorem 6.3.1, we derive the wireless linear network codes for the specific network realizations $N^f = \{N_m^f\}_{m=1}^3$ in Figure 6.3 as follows: For N_1^f, we have $\beta^{(2,1)}(e_1) = 1$, $\alpha_2^{(3,1)}(s) = 1$ and $\epsilon_1^{(1,1)}(e_7) = 1$. For N_2^f, we have $\beta^{(1,2)}(e_2) = 1$, $\alpha_1^{(1,2)}(s) = 1$ and $\epsilon_2^{(2,2)}(e_8) = 1$. For N_3^f, we have $\beta^{(1,3)}(e_3) = 1$, $\beta^{(2,3)}(e_4) = 1$, $\epsilon_2^{(3,3)}(e_5) = 1$, $\epsilon_2^{(1,3)}(e_7) = 1$, $\epsilon_1^{(3,3)}(e_6) = 1$ and $\epsilon_1^{(2,3)}(e_8) = 1$.

6.3.4 Properties of Linear Wireless Network Codes and Interactions with MAC Schedules

We assume that the wireless network realizations $\{N_m^f\}_{m=1}^M$ are predetermined with conflict-free link sets $\{E_m^f\}_{m=1}^M$ for a network graph with node set V and link set E. We assume binary network coding operations in finite field $\mathcal{F}_2$.

First, we consider the effects of the constraint that nodes cannot simultaneously transmit and receive packets. Suppose that node $v \in V$ transmits a packet in network realization N_m^f, i.e. $\beta^{(n,m)}(e_1) = 1$ or $\alpha_j^{(n,m)}(v) = 1$ for any incoming link $e_1 \in E_n^f : head(e_1) = v$ and flow j generated in network realization N_n^f. Then, node v cannot receive any information flow (on any incoming link $e \in E$) to be encoded for possible transmissions in another network realization N_p^f, i.e. $\beta^{(m,p)}(e) = 0$ for

any link $e : head(e) = v$. Thus, the linear encoding operations must satisfy

$$\beta^{(m,p)}(e) = 0, \text{ if } \beta^{(n,m)}(e_1) = 1, head(e) = head(e_1), \qquad (6.14)$$

$$\beta^{(m,p)}(e) = 0, \text{ if } \alpha_j^{(n,m)}(v) = 1, head(e) = v \qquad (6.15)$$

for m, n, $p \in \{1, ..., M\}$. Also, node v does not receive any information that can be used to decode any flow in network realization N_f^r, i.e. $\epsilon_k^{(m,r)}(e) = 0$ for any link $e : head(e) = v$, network realization N_f^r and flow k. Thus, we obtain the conditions

$$\epsilon_k^{(m,r)}(e) = 0, \text{ if } \beta^{(n,m)}(e_1) = 1, head(e) = head(e_1), \qquad (6.16)$$

$$\epsilon_k^{(m,r)}(e) = 0, \text{ if } \alpha_j^{(n,m)}(v) = 1, head(e) = v \qquad (6.17)$$

for m, n, $r \in \{1, ..., M\}$. Next, we consider the effects of conflict-free scheduling to eliminate simultaneous transmission and reception by any node such that $T^{(m)}$ and $R^{(m)}$ denote the disjoint sets of transmitters and receivers activated in network realization N_m^f. Suppose that node v is activated as receiver in network realization N_m^f, i.e. $v \in R^{(m)}$. Node v should not encode and transmit any information that has arrived in any network realization N_n^f on incoming link $e : head(e) = v$, i.e. $\beta^{(n,m)}(e) = 0$. Otherwise, node v cannot successfully receive packets in N_m^f. Similarly, node v should not transmit in network realization N_m^f any information flow j generated in any network realization N_n^f, i.e. $\alpha_j^{(n,m)}(v) = 0$. Thus, we obtain the conditions

$$\beta^{(n,m)}(e) = 0, \text{ if } head(e) \in R^{(m)}, \qquad (6.18)$$

$$\alpha_j^{(n,m)}(v) = 0, \text{ if } v \in R^{(m)} \qquad (6.19)$$

for m, $n \in \{1, ..., M\}$. If node v is activated as transmitter in N_m^f (i.e. if $v \in T^{(m)}$),

172

then it must encode at least one information flow with at least one non-zero encoding coefficient, i.e. there exist $e \in E$ and $p \in \{1, ..., M\}$ such that $\beta^{(p,m)}(e) = 1$ for $head(e) = v$ or there exist $j \in \{1, ..., \mu(v)\}$ and $p \in \{1, ..., M\}$ such that $\alpha_j^{(p,m)}(v) = 1$, if $v \in T^{(m)}$ for $m \in \{1, ..., M\}$. Otherwise, node v would stay idle or act as receiver. Also, the transmitting node v receives in N_m^f no new information that can be encoded for subsequent transmissions in N_r^f, i.e. $\beta^{(m,r)}(e) = 0$ for any link $e : head(e) = v$. Therefore, we obtain the condition

$$\beta^{(m,r)}(e) = 0, \text{ if } head(e) \in T^{(m)} \tag{6.20}$$

for $m, r \in \{1, ..., M\}$. In addition, there is no new information that can be decoded to any flow k in N_r^f, i.e. $\epsilon_k^{(m,r)}(e) = 0$ for any link $e : head(e) = v$. Thus, we obtain the condition

$$\epsilon_k^{(m,r)}(e) = 0, \text{ if } head(e) \in T^{(m)} \tag{6.21}$$

for $m, r \in \{1, ..., M\}$. Also, there is no need to change the decoding coefficients of node $v \in T^{(m)}$.

Next, we consider the effects of interference on network codes. The flows on links e_1 and e_2 arriving at node v in the same network realization N_m^f cannot be successfully received and therefore cannot be encoded by node v for subsequent transmissions and cannot be decoded to any flow in subsequent network realizations N_n^f and N_p^f, respectively. Therefore, the encoding coefficients $\beta^{(m,n)}(e_1)$ and $\beta^{(m,p)}(e_2)$ as well as the decoding coefficients $\epsilon_k^{(m,n)}(e_1)$ and $\epsilon_k^{(m,p)}(e_2)$ cannot be both non-zero for $e_1 \neq e_2$ and $head(e_1) = head(e_2)$, i.e.

$$\beta^{(m,n)}(e_1)\beta^{(m,p)}(e_2) = 0, \epsilon_k^{(m,n)}(e_1)\epsilon_k^{(m,p)}(e_2) = 0, \text{if } e_1 \neq e_2, head(e_1) = head(e_2) \tag{6.22}$$

for $m, n, p \in \{1, ..., M\}$. Also, if $v = head(e_1)$ transmits a packet to node $head(e_2)$ in N_m^f (i.e. $(head(e_1), head(e_2)) \in E$ and $\beta^{(n,m)}(e_1) = 1$ or $\alpha_j^{(n,m)}(v) = 1$), then node $head(e_2)$ cannot successfully receive packets on any link e_2 for $head(e_1) \neq tail(e_2)$. Therefore, node $head(e_2)$ cannot use $Y^{(m)}(tail(e_2))$ for the encoding and decoding purposes in any subsequent network realization N_p^f, i.e. $\beta^{(m,p)}(e_2) = 0$ and $\epsilon_k^{(m,p)}(e_2) = 0$. For $m, n, p \in \{1, ..., M\}$, we obtain the conditions

$$\beta^{(m,p)}(e_2) = \epsilon_k^{(m,p)}(e_2) = 0, \text{if } \beta^{(n,m)}(e_1) = 1, \tag{6.23}$$

$$head(e_1) \neq tail(e_2), (head(e_1), head(e_2)) \in E,$$

$$\beta^{(m,p)}(e_2) = \epsilon_k^{(m,p)}(e_2) = 0, \text{if } \alpha_j^{(n,m)}(v) = 1, v \neq tail(e_2), (v, head(e_2)) \in E. \tag{6.24}$$

Consider conflict-free scheduling to eliminate interference. Assume that node $v = head(e_2)$ is scheduled to transmit in N_f^m such that $\beta^{(r,m)}(e_2) = 1$ or $\alpha_j^{(s,m)}(v) = 1$, i.e. $v = head(e_2) \in T^{(m)}$, and can reach the non-intended receiver node d, i.e. $(v, d) \in E$ and $(v, d) \notin E_m^f$. Then, node u should not transmit a packet to node d in network realization N_f^m, since this transmission cannot be successful, i.e. $(u, d) \notin E_m^f$. Node d cannot receive in network realization N_m^f any packet flow from node u that can be encoded for subsequent transmissions or decoded to output flows in subsequent network realizations. For $m, n \in \{1, ..., M\}$, we obtain the conditions

$$\beta^{(m,n)}((u,d)) = 0, \epsilon_k^{(m,n)}((u,d)) = 0, \text{ if } \exists\, v \in T^{(m)} \text{ s.t. } v \neq u \text{ and } (v,d) \in E. \tag{6.25}$$

6.3.5   Network Coding with Arbitrary MAC through Group TDMA

The next problem is to operate network coding with arbitrary MAC protocols within each receiver's area. The network is partitioned into separately activated

174

transmitter and receiver sets. Then, for each receiver set, Group TDMA scheme of Chapter 2 separately activates interfering node groups (with packets addressed to different receivers) to decouple feedback from multiple receivers and prevent the instability of MAC operations. We use an arbitrary single-receiver protocol (e.g. FCFS algorithm) for multiple access within each receiver's area. The time fractions allocated to different node groups are chosen to optimize the stable throughput or energy costs. We construct the network coding (or routing) solution over the one-hop transmission schedules of Group TDMA, as we did for the conflict-free network realizations in section 6.2.

6.4   Improved Joint MAC and Network Coding Methods

In this section, we introduce alternative solutions for joint MAC (scheduling) and network coding under the assumption of classical collision channel model. We assume single common network method based on subtree decomposition that has been proposed in [72] for wired networks and consider three separate methods for the MAC part. The objective is to use the results of network coding to construct conflict-free transmission schedules in a joint analysis (with possibly distributed implementation) while incorporating omnidirectional transmissions.

6.4.1   Subtree Decomposition Method

We outline three steps of the subtree decomposition method [72].

Step 1: We construct from the original network graph G a dual line graph (as shown
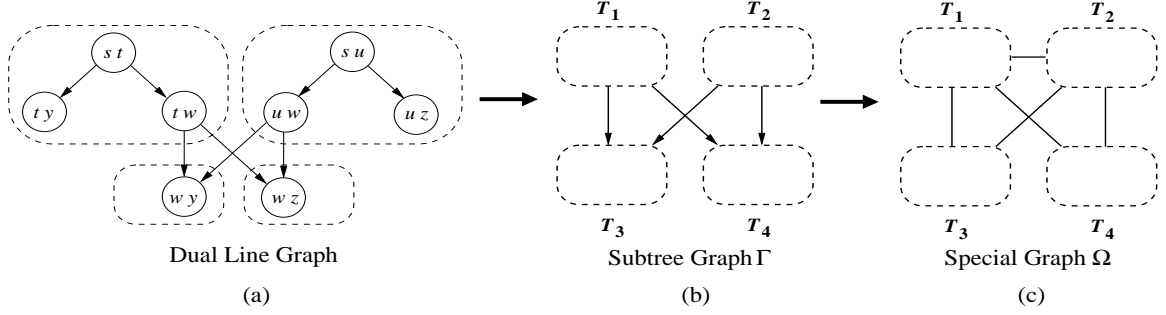
Figure 6.7: (a) Dual line graph for the original network graph in Figure 6.1 and subtree decomposition of the dual line graph, (b) Subtree graph Γ, (c) Special graph Ω.

in Figure 6.7 for the network topology in Figure 6.1), where links become vertices and these vertices are connected, if the corresponding links in the original graph G are adjacent (i.e. share a common vertex). We define three special types of nodes in dual line graph: Source nodes in dual line graph are links in original graph that are outgoing from a source. Destination nodes in dual line graph are links in original graph that are incoming to a destination. Coding nodes in dual line graph are links in original graph that are downstream and adjacent to more than one incoming link. For the dual line graph in Figure 6.7, the source nodes are st and su, and the coding and destination nodes are both wy and wz.

Step 2: We partition the dual line graph to a disjoint union of connected subtrees $\{T_i\}$ (as shown in Figure 6.7 for the network topology in Figure 6.1) such that
(a) each subtree contains exactly one source or exactly one coding node, and
(b) any other node in dual line graph belongs to the subtree that contains its first ancestral (i.e. closest upstream) source or coding node. Network coding operation

176

depends on how subtrees are connected and which destination nodes are included in each subtree.

Step 3: We construct the subtree graph Γ from the dual line graph (as shown in Figure 6.7-(b) for the network topology in Figure 6.1) by

(a) contacting each subtree in the dual line graph to a node, and

(b) retaining only the edges that connect the subtrees.

For minimal subtree decomposition, we check if we can further remove any link in the subtree graph and merge the subtrees without reducing the minimum cut capacity of the original graph.

6.4.2 Common Network Coding Method by Subtree Decomposition

We assume that there are h source nodes in dual line graph and there are h edge-disjoint paths from source to any destination in the original graph G. The source node in G generates h-dimensional information vector $\underline{x}$ that should be jointly decoded by each destination. Each component of vector $\underline{x}$ carries a symbol from a finite field F_q. Any node in the original graph G can transmit at most one symbol per time slot. Each transmission on link e in graph G (or node e in dual line graph) carries a single symbol $y(e)$. This symbol is constructed as $y(e) = \underline{f}(e) \, \underline{x}^T$, where $\underline{f}(e)$ is the h-dimensional code vector for node e in dual line graph.

We assign the same code vector to all nodes in each subtree T_i, i.e. $\underline{f}(e)$ is constant for $e \in T_i$. As a result, the same symbol is transmitted in each subtree, i.e. $y(e)$ is constant for all $e \in T_i$. This is also consistent with the assumption of

omnidirectional transmissions and equivalent to assigning $\underline{f}(T_i)$ as the code vector for each subtree T_i. The symbol $y(e) = \underline{f}(e)\,\underline{x}^T$ is transmitted on $e \in T_i$ only when e is activated by the underlying MAC method (that we will specify in sections 6.4.3, 6.4.4, and 6.4.5 in three different ways). We construct linear network codes such that each destination receives all necessary and sufficient information as independent vectors to recover the information vector $\underline{x}$. Successful decoding by destination nodes is assured if

(a) the code vector of any subtree is in the linear span of code vectors of parent subtrees, and

(b) the code vectors of subtrees are linearly independent provided that

    (i) the subtrees are connected to a common child subtree, or

    (ii) the subtrees contain nodes (links in original graph) such that these links lead to the same destination node of the original graph (i.e. they share a destination node).

The problem of assigning codes $f(T_i)$ to the subtree nodes $T_i \in \Gamma$ has been formulated in [79] as vertex coloring of a special graph Ω. We construct Ω from the subtree graph Γ (as shown in Figure 6.7-(c) for the network topology in Figure 6.1) by keeping the original links and introducing new links between (subtree) nodes, if (i) these nodes are connected to a common child (subtree) node, or (ii) the corresponding subtrees contain nodes (links in the original graph) such that these links lead to the same destination node in the original graph.

We color the special graph Ω (by one of the well-known graph coloring heuristics [80] such that no adjacent subtree has the same color) and assign linearly inde-

pendent network codes to the subtrees that have been colored with different colors. As an example, linear network codes for the network in Figure 6.1 with $h = 2$ are given by $f(T_1) = [0, 1]$, $f(T_2) = [1, 0]$, $f(T_3) = f(T_4) = [1, 1]$.

It is an NP-complete problem to find the smallest coding alphabet size, i.e. the smallest number of colors needed so that no two adjacent vertices share the same color. We need at most $|\Omega|$ colors and the alphabet size has the upper bound of $|\Omega| - 1$. As shown in [79], the upper bound on $|\Omega|$ is $n_d + 1$, where n_d is the number of destination nodes in the dual line graph. For $h = 2$, we can employ a greedy coloring algorithm that sequentially colors each vertex of graph Ω with a color that has not been used before for any of its neighbors. If $\Delta(\Omega)$ denotes the maximum degree of the vertices on graph Ω, we need at least $\Delta(\Omega) + 1$ colors and the alphabet size is at least $\Delta(\Omega)$. The distributed implementation issues have been discussed in [81] for (wired) network coding solutions based on subtree decomposition.

6.4.3 Scheduling by Subtree Decomposition: Method A

We use the results of subtree decomposition to construct conflict-free transmission schedules. We divide the time frame into subframes and match each subframe with exactly one subtree (as shown in Figure 6.8). Nodes in a given subtree are activated only in the matched subframe and remain idle for the rest of the time. We divide each subframe into three time slots and assign each node in the corresponding subtree into one of the three disjoint node groups according to its depth level, which is defined as the hop-distance from the source or coding node of the subtree.
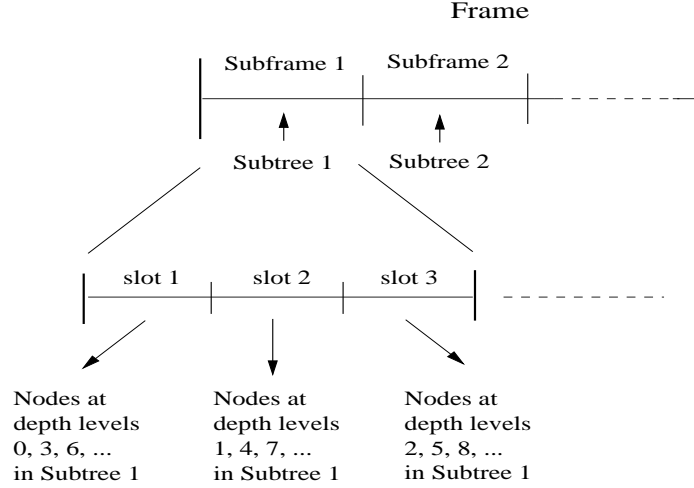
Figure 6.8: Subframe and slot assignment by scheduling Method A.

For each subtree $T_i \in \Gamma$ and $j = 0, 1, 2$, nodes at depth levels $3\,l + j$, $l = 0, 1, 2, ...$, form a different node group $N_{i,j}$ (as shown in Figure 6.9 for an isolated subtree i). Then, each node in group $N_{i,j}$ for $T_i \in \Gamma$ and $j = 0, 1, 2$ is activated in slot $j + 1$ of the subframe matched to the subtree T_i. This spatial reuse approach enables conflict-free transmission schedules between subtrees (since transmissions in different subtrees are separately activated) and also within each subtree (since nodes at every third depth level can be simultaneously activated without interfering with each other). Scheduling method A is summarized as follows:

(a) Divide time frame into subframes and combine each subtree with exactly one distinct subframe.

(b) Use additional spatial reuse to combine nodes at every third depth level of any subtree in the same time slot of the corresponding subframe.

The resulting time frame is activated in a time-division fashion and has the length of $3|\Omega|$ time slots. If a subtree contains less than three nodes, some time slots
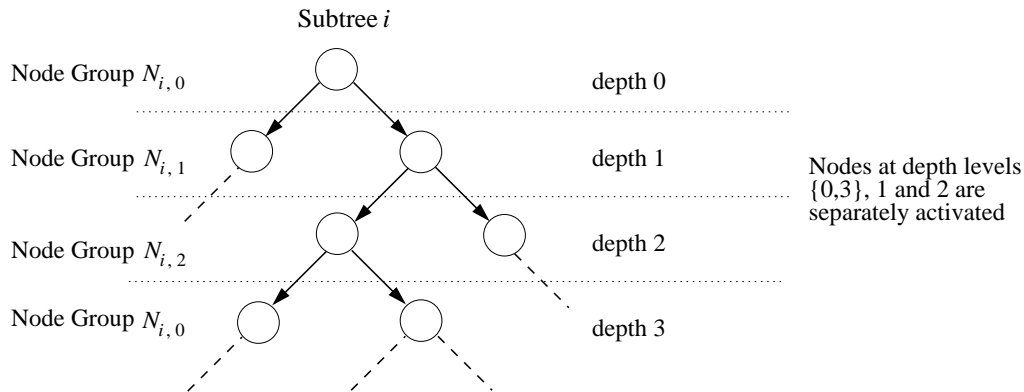
Figure 6.9: Conflict-free scheduling within a given subtree.

become idle and need to be shared by the nodes present in that subtree. In general, the fixed subframe length of three time slots is not optimal. Instead, the subframe lengths can be chosen to optimize the performance objectives such as maximizing throughput or minimizing cost (e.g. energy). For that purpose, we construct a hypothetical wired graph with the same node set as the original graph G and define the link capacity as the proportion of time during which the link is activated (as we did before in section 6.3). Then, we express the cut capacities as functions of the subframe lengths that are further chosen to optimize the performance objectives.

6.4.4 Scheduling by Subtree Graph Coloring: Method B

Method A is not optimal and needs to be refined, since different subtrees (i.e. subframes) may include nodes that can be simultaneously activated without any conflict, e.g. Method A separately activates links wy and wz of the network in Figure 6.1, although these links do not conflict with each other.

An improved scheduling method follows directly from the network coding so-

lution based on subtree decomposition and subtree graph coloring:

(a) Determine network codes by assigning colors to subtree nodes in Ω.

(b) Divide time frame into subframes and combine all subtrees of the same color (i.e. code) into the same subframe.

(c) Use additional spatial reuse to combine nodes at every third depth level of any subtree in the same time slot of the corresponding subframe.

Transmissions in subtrees of the same color (i.e. code) cannot interfere with each other and therefore can be simultaneously activated. We also apply spatial reuse by separating transmissions of nodes within hop-distance of two or less and simultaneously activating nodes at every third depth level within each subtree to construct conflict-free transmission schedules.

If a subtree with a single node connects a pair of parent and child subtrees that are not directly connected on graph Ω, then these parent and child subtrees are possibly assigned the same network code and the same time schedule that will result in packet collisions because of the secondary interference effects at non-intended receivers. We need to identify such single-node subtrees during the initial coloring process (via communication over the feedback channel) and reassign the time schedules to their child subtrees while keeping the network code assignments.

As the coloring heuristic [80], we order the colors and use the smallest color to color the subtree that has the largest number of colored adjacent subtrees. We need $\Delta(\Omega) + 1$ colors (i.e. subframes) and the length of the resulting time frame is at least $3(\Delta(\Omega) + 1)$ slots. In addition, the subframe lengths can be further specified for optimality, as described for Method A. Scheduling Method B is based on apply-

ing coloring on simple subtree decompositions, whereas the traditional scheduling problem is related to high-complexity coloring of the entire network graph [14].

6.4.5 Scheduling by Exhaustive Search: Method C

Scheduling Method B reduces the number of subframes and improves time reuse; however, it is not optimal and does not simultaneously activate all possible conflict-free transmissions, e.g. links st and uw of the network in Figure 6.1 are separately activated by Method B, although these links do not conflict with each other.

For better spatial reuse, scheduling Method C performs exhaustive search for nodes that can be also activated in other subframes without any conflict as follows:
(a) Perform scheduling Method B.
(b) Visit subframes sequentially.

(i) For each subframe, check each node if it is a destination or a coding node.

(ii) If not, assign the node also to slots in other subframes, unless the subframe contains a coding node connected to the particular node.

Thus, we keep the schedules assigned by Method B to destination nodes or coding nodes while separately activating coding nodes and their one-hop neighbor nodes. This cancels the interference effects between different subtrees by separately activating the incoming links of different destination nodes (in different subtrees) and separately activating the incoming and outgoing links of any source or coding node. Since the rest of the nodes in the dual line graph cannot conflict with nodes

in other subtrees, they are assigned additional time schedules in other subframes.

6.4.6   Properties of Scheduling Methods A, B and C

For the network in Figure 6.1, the scheduling solutions are given in Table 6.3. For the scheduling methods A and B, we define t_i as the fraction of time allocated to subframe i and $t_{i,k}$ as the fraction of time allocated to nodes at depth levels $k \pmod 3$ of the corresponding subtree. The only difference of scheduling Method B from scheduling Method A is that links wy and wz are allocated within the common time fraction $t_{3,0}$. Scheduling Method C provides the most efficient reuse of the time schedules by separately activating the three network realizations in Figure 6.2.

Table 6.3: Transmission Schedule Assignments by Methods A, B, C.

Links:	$s\,t$	$t\,y$	$t\,w$	$s\,u$	$u\,w$	$u\,z$	$w\,y$	$w\,z$
Scheduling Method A	$t_{1,0}$	$t_{1,1}$	$t_{1,1}$	$t_{2,0}$	$t_{2,1}$	$t_{2,1}$	$t_{3,0}$	$t_{4,0}$
Scheduling Method B	$t_{1,0}$	$t_{1,1}$	$t_{1,1}$	$t_{2,0}$	$t_{2,1}$	$t_{2,1}$	$t_{3,0}$	$t_{3,0}$
Scheduling Method C	t_1	t_2	t_2	t_2	t_1	t_1	t_3	t_3

The problem of finding the optimal subframe lengths is equivalent to finding the optimal time fractions allocated to subframes. For the subtree decomposition in Figure 6.7, we have from symmetry $t_1 = t_2$, $t_3 = t_4$, $t_{i,k} = \frac{t_i}{2}$ for $i = 1, 2$, $k = 0, 1$, and $t_{i,0} = t_i$ for $i = 3, 4$ under scheduling Method A or $t_{i,0} = t_3$ for $i = 3, 4$ under scheduling Method B.

The packets arrive at each destination y and z with rate $\frac{t_1}{2}$ in subtrees T_1 and T_2 and with rate t_3 in subtrees T_3 and T_4. The minimum cut carries packets

184

with rate $\frac{t_1}{2} + \min\left(\frac{t_1}{2}, t_3\right)$, where $2(t_1 + t_3) = 1$ under scheduling method A and

$2t_1 + t_3 = 1$ under scheduling method B. The maximum throughput per destination

achievable by the scheduling methods A and B is, respectively, $\frac{1}{3}$ and $\frac{2}{5}$ packets

per time slot (with $t_1 = t_2 = \frac{1}{3}$, $t_3 = t_4 = \frac{1}{6}$ under scheduling method A and

$t_1 = t_2 = \frac{2}{5}$, $t_3 = \frac{1}{5}$ under scheduling method B). The scheduling method C activates

three network realizations (equivalent to those given in Figure 6.3) in three distinct

subframes with the optimal time allocation of $t_i = \frac{1}{3}$, $i = 1, 2, 3$, and maximizes the

throughput rate λ to $\frac{2}{3}$ packets per time slot. The energy-optimal solutions can be

obtained similarly and omitted for brevity.

6.4.7 Distributed Implementation Issues

Wireless ad hoc network operation requires distributed methods that should

only use local exchange of information among neighbor nodes. Graph coloring and

conflict-free scheduling (among ordered nodes) are equivalent and both have well-

known distributed solutions [14, 5]. We need to find distributed methods to identify

the subtrees and construct the subtree graph. The existing methods are centralized

and either require global network information such as the number of sources and

destinations or exchange subtree assignments throughout the network [81]. If the

subtrees are already identified, the remaining problem is to color the subtrees for

the network code assignments. For that purpose, source and coding nodes of dif-

ferent subtrees must communicate with each other through arbitrary number relay

nodes (depending on the depths of the subtrees). However, this may increase the

communication burden and needs to be carefully organized.
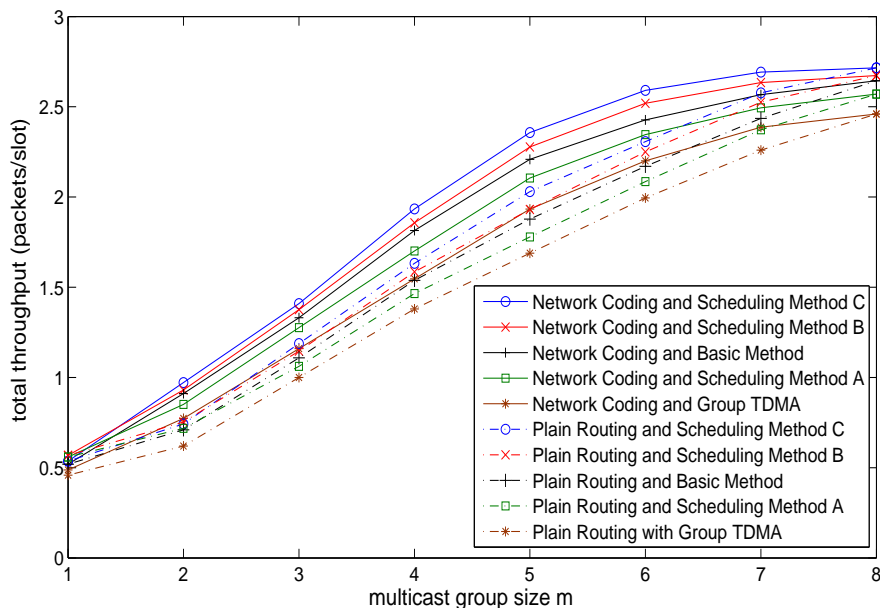
Figure 6.10: Throughput of network coding and plain routing.

6.5   Performance Comparison of Network Coding and Plain Routing

We compare the performance of network coding with plain routing under the proposed MAC solutions. We consider a network of 9 nodes uniformly distributed on a regular square grid with at least one and at most four-neighbor connectivity. We consider the classical collision channel model and assume that each packet contains one bit and source node has always a packet to be delivered to randomly chosen destination nodes in the multicast group of size m.

Figure 6.10 evaluates the total throughput achievable (for all destination nodes) by network coding and plain routing as function of m under the proposed MAC solutions (namely Methods A, B and C of section 6.4, Basic Method of section 6.3 and
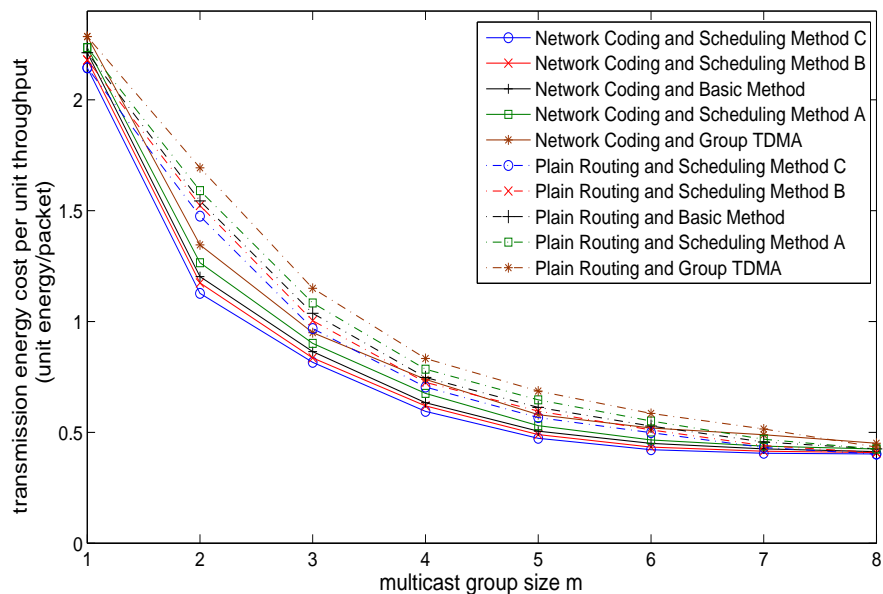
Figure 6.11: Transmission energy cost of network coding and plain routing.

Group TDMA algorithm that uses FCFS as single-receiver MAC protocol within each receiver's area).

Table 6.4: Maximum Improvement of Network Coding over Plain Routing.

Perfomance Measure	Total Throughput	Transmission Energy Cost
Basic Method	25 %	24 %
Group TDMA	18 %	15 %
Method C	31 %	28 %
Method B	28 %	26 %
Method A	21 %	19 %

Next, we assume unit energy cost for each packet transmission and evaluate the transmission energy efficiency of network coding and plain routing solutions that have the common objective to minimize the average transmission energy consump-

tion per successfully decoded packet. Figure 6.11 depicts the transmission energy cost per unit throughput for different values of m, respectively. The maximum improvement of network coding over plain routing is presented in Table 6.4 in terms of throughput and transmission energy cost.
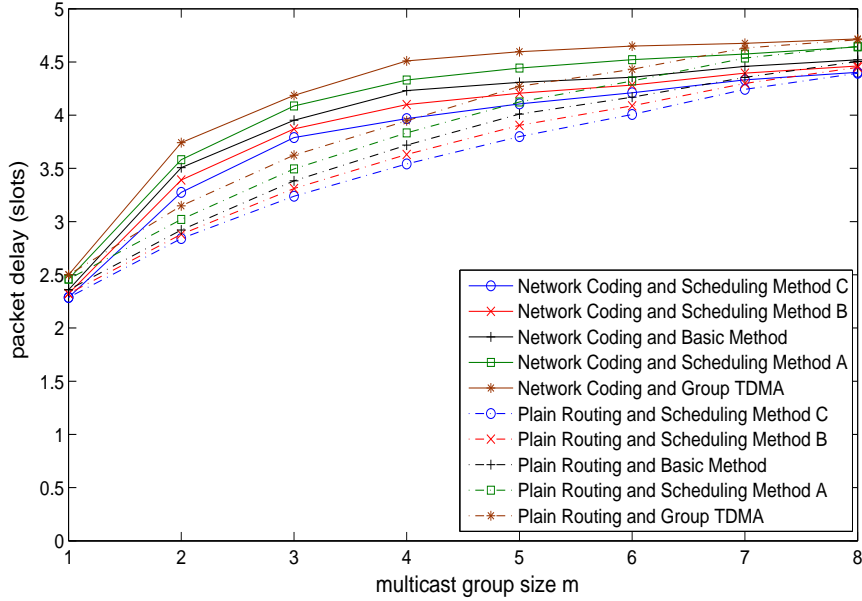
Figure 6.12: Packet delay of throughput-optimal network coding and plain routing.

Numerical results show that network coding outperforms plain routing and that the scheduling Method C described in section 6.4 performs better than both the solution of section 6.3 that is based on predetermined network realizations and also the Group TDMA algorithm. Network coding increases the average packet delay compared to plain routing, as shown in Figure 6.12 for the case of throughput-optimal solutions, since nodes need to accumulate packets (from different neighbor nodes) over successive time slots to perform network coding. Finally, we assume unit cost for each coding (or decoding) operation and evaluate in Figure 6.13 the
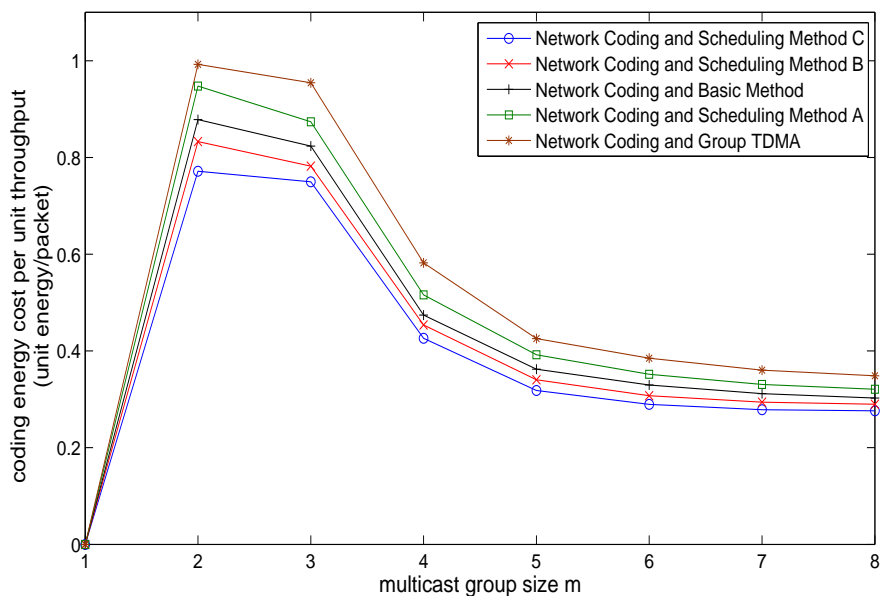
Figure 6.13: Coding energy cost of throughput-optimal network coding.

total network coding cost per unit throughput for the case of throughput-optimal solutions. We expect similar results for more general network topologies and packet traffic levels.

6.6   Summary and Conclusions

In this chapter, we considered joint design of network coding (or plain routing as a special case) and MAC in wireless ad hoc networks. We introduced time-varying linear network codes and evaluated their properties under wireless network assumptions. After that we presented a basic method to derive the conflict-free network realizations and separately activate them in order to maximize the achievable throughput or minimize the node costs (such as energy consumption) through network coding. We also outlined an extension of network coding to operate with

189

arbitrary MAC protocols through the application of the Group TDMA algorithm. Then, we introduced practical solutions that rely on graph coloring to derive network codes and conflict-free transmission schedules on simplified subtree network representations. Finally, we evaluated the performance of network coding over plain routing in conjunction with the proposed MAC solutions.

Our analysis is based on allowing nodes to accumulate relay packets in periodic network coding operation with saturated packet queues at source and relay nodes. However, this approach can increase packet delay compared to plain routing. The network coding results should be extended to wireless queueing networks with possibly emptying packet queues and resultant underflow. In this context, the notions of delay and buffer overflow should be revisited to design network codes based on the instantaneous queue contents. We will address the stability issues in Chapter 7.

Chapter 7

Joint Optimization of MAC and Network Coding in Wireless

Queueing Networks with Multiple Sources

Throughput optimization is an essential performance objective that cannot be
adequately characterized by a single criterion (such as the minimum transmitted or
sum-delivered throughput) and should be specified over all source-destination pairs
as a region of throughput rates. For a simple and yet fundamental model of tandem
networks, we formulate a cross-layer optimization framework across the medium
access control (MAC) and the network layers to derive the maximum throughput
region achievable by saturated multicast traffic. The results illuminate multidimen-
sional throughput gains of network coding over plain routing. If the network model
also incorporates bursty sources and allows packet queues to empty, the objective
is changed to specify the stability region as the set of maximum throughput rates
that can be sustained with finite packet delay for all source-destination pairs. Dy-
namic queue management strategies are used to distinguish source and relay packet
transmissions based on instantaneous queue contents and expand the stability re-
gion towards the maximum throughput region by jointly specifying network coding
(or plain routing) and MAC operations. The throughput optimization results im-
pose fundamental trade-offs with the energy efficiency objectives. We evaluate both
transmission and processing energy costs to show that the throughput-optimal op-

eration of network coding is not necessarily energy-efficient. Finally, the analysis is extended to non-cooperative network operation with selfish nodes competing for limited network resources. We point at the inefficiency of competitive MAC and network coding (or plain routing) decisions, and introduce cooperation stimulation mechanisms to improve the throughput and energy efficiency performance.

7.1 Introduction

A single scalar criterion is not sufficient to reflect all communication demands of multiple source-destination pairs and it is necessary to construct a region of attainable transmission rates at which reliable communication can occur. The performance limits of reliable data transmission can be formulated as an information-theoretic capacity problem that is based on the assumption of uninterrupted availability of information symbols to be coded (and transmitted) without regard to delay. This problem has not been fully extended yet from the point-to-point transmission paradigm to general network operation [82]. Alternatively, we can consider packets as information units to be transmitted and evaluate the maximum region of throughput rates achievable under the assumption of saturated queues with continuously generated packet traffic. We can also assume bursty packet arrivals and allow packet queues to empty. The resulting performance measure of interest is the stability region, which is defined as the largest collection of packet traffic rates (originating at multiple sources) for which the queue sizes remain finite. In general, the capacity, maximum throughput and stability regions can be different from each

other in wireless networks [22, 83].

For point-to-point or unicast communication, the back-pressure algorithms achieve the maximum stable throughput region at the expense of poor delay performance [84]. The capacity analysis of wireless networks has been limited to saturated queues with infinite delay [10]. The extension of the network capacity problem to general multicast communication is facilitated through network coding that we discussed in Chapter 6 for wireless domain [85]. The classical objective of network coding studies is to maximize the Max-flow Min-cut capacity for all source-destination pairs and therefore it is sufficient to assume saturated queues that guarantee continual availability of packets for transmission without risk of underflow or concern about delay build-up. For stable operation, the potential use of the back-pressure algorithms in conjunction with network coding has been presented in [86] for the case of separating the multicast traffic of different source nodes.

The network coding objective of maximizing a single common multicast rate cannot fully represent the aggregate throughput performance, since the throughput demands of different source-destination pairs may differ and conflict with each other. As an extension, the throughput rates achievable at different destinations have been separately considered in [87] to study the average throughput properties of wired networks.

In this chapter, we distinguish the throughput rates achievable by multiple source nodes in a wireless network. For general communication demands (that include multicast, broadcast and multiple unicast scenarios), the main objective is to specify the achievable and stable throughput regions (with common rate for each

193

destination in the same multicast group) for the separate cases of (a) saturated queues and (b) not overloaded systems with finite packet delay. We formulate a general cross-layer design framework for optimizing the achievable throughput or stability regions through the joint selection of MAC and network layer strategies. We evaluate the multidimensional performance gains of network coding over plain routing in terms of the maximum throughput and stability regions.

Because of the complexity introduced by the presence of multiple sources and the requirement of stable operation, we restrict our attention to a simple tandem network topology with at most two-node connectivity such that any packet transmission reaches only its left and right neighbors, respectively. The linear tandem network model with node set $N = \{1, ..., n\}$ is shown in Figure 7.1.
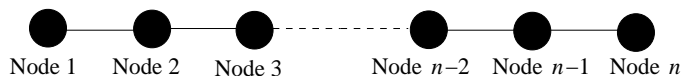
Figure 7.1: Tandem network model with node set $N = \{1, ..., n\}$.

For stable operation of the queues, the possible underflow of relay packets opens up new questions regarding the optimal queue management and the optimal use of network coding based on the instantaneous queue contents. Practical network coding heuristics have been presented in [88, 89]. Instead, we derive the achievable throughput and stability regions in a theoretical cross-layer optimization framework for joint MAC and network coding, and discuss how the question of MAC protocol affects and is affected by the use of network coding (or plain routing as a special case). For stable operation, we present network coding and plain routing strategies based on different priorities assigned to source and (either network-coded

or uncoded) relay packets. Specifically, we show that the stable operation in tandem networks can approach the maximum throughput rates achievable by saturated packet traffic.

We also consider the single aggregate throughput performance measure, in addition to the entire region of throughput rates. In this context, we formulate the cross-layer optimization problems of maximizing the sum-delivered throughput or the minimum transmitted throughput. We also evaluate the transmission and processing energy costs for network coding and plain routing operations to sustain a given set of (achievable or stable) throughput rates. We highlight the cross-layer optimization trade-offs between different measures of throughput and energy efficiency.

The previous (wired or wireless) network coding studies rely on cooperation of nodes to jointly optimize the network performance objectives. However, it is difficult to coordinate a large number of nodes under cooperation-based MAC and network coding (or plain routing). If selfish nodes with individual performance objectives compete for MAC and network layer tasks, they are subject to performance loss compared to centralized cooperation [53, 90, 54, 55, 56, 3, 19, 40]. The cooperation can be realized externally by a central authority, or we can impose distributed cooperation reinforcement mechanisms and let nodes operate selfishly in a distributed manner to optimize the individual performance measures involving the throughput rewards and (transmission and processing) energy costs.

In this chapter, we also extend the analysis to the non-cooperative operation with nodes competing for limited network resources of bandwidth and energy.

We formulate a joint MAC and network coding game with the node utilities that reflect the individual throughput and energy efficiency objectives. We introduce distributed cooperation stimulation mechanisms and evaluate the improvement in non-cooperative equilibrium strategies of MAC and network coding (or plain routing). We introduce a rewarding mechanism to charge credits for packet relaying. We show that it is possible to let nodes make selfish MAC and network coding (or plain routing) decisions in a distributed fashion while approaching the same performance results as in the centralized cooperation. We also specify the dependence of the non-cooperative equilibrium strategies of joint MAC and network coding on the throughput credits, energy costs and rewards for packet relaying.

The chapter is organized as follows. We present the wireless network model in section 7.2 and formulate the cross-layer optimization problem in section 7.3. For scheduled access, we specify the achievable throughput region in section 7.4 and the stability region in section 7.5. We discuss the throughput optimization trade-offs in section 7.6 and incorporate the energy efficiency measures of transmission and processing energy costs in section 7.7. Then, the results are extended to contention-based random access in section 7.8 and to the non-cooperative network operation in section 7.9. Finally, we draw conclusions in section 7.10.

7.2   MAC and Network Layer Model

We assume the classical collision channel model and separately consider (a) conflict-free transmission schedules and (b) contention-based random access at the

MAC layer.

(a) *Scheduled Access*: We order n nodes in a tandem network from left to right and divide them into three groups such that node i is included in group $m = (i-1)(\bmod 3) + 1$, where $m = 1, 2, 3$. Nodes in group m are activated for a time fraction t_m, where $0 \leq t_m \leq 1$ and $\sum_{m=1}^{3} t_m = 1$. The activation times of the three groups are disjoint.

(b) *Random Access*: Each node i transmits a packet in any time slot with a constant and fixed probability p_i. The collided packets remain backlogged until they are successfully received. Any transmission randomly chooses either a source packet or a relay packet that arrived from a neighboring node.

We also distinguish between the case of saturated queues (with uninterrupted availability of source and relay packets) and stable operation (with possibly emptying packet queues). Each relay packet coming from one of a node's neighbors must be transmitted only to the neighbor node on the opposite side, whereas a source packet may need to be delivered to one or both neighbor nodes, depending on whether its destinations are located on one or both sides of the node in question.

Each node i has three separate packet queues of infinite capacities: Queue Q_i^1 stores the source packets that node i generates, while queues Q_i^2 and Q_i^3 store the relay packets that are incoming from its right and left neighbors, respectively. At the network layer, we separately consider (a) plain routing and (b) network coding operation:

(a) *Plain Routing*: Each node i either transmits a packet from its source queue Q_i^1 or a packet from one of its relay queues Q_i^2 and Q_i^3; these two queues may be

combined into a single queue (in a first-come-first-served fashion).

(b) *Network Coding*: Each node i either transmits a source packet from queue Q_i^1, or a relay packet from one of the other two queues, or the coded combination of two relay packets one from each of the relay queues Q_i^2 and Q_i^3. We consider each packet as a vector of bits and assume $\mathcal{F}_2$ to be the field for linear network coding operations such that the bit-sum $x + y$ of two packets x and y is a modulo-2 vector addition of the corresponding vectors of each packet.

Lemma 7.2.1 *In a tandem network, nodes can separately transmit source and relay packets (instead of combining them by network coding) without loss of optimality in terms of throughput, energy efficiency or packet delay.*

Proof: Consider the case in which a node i has at least one source packet s_i to transmit and at least one relay packet $r_{i,i-1}$ in queue Q_i^2 and at least one relay packet $r_{i,i+1}$ in queue Q_i^3 to deliver to neighbor nodes $i - 1$ and $i + 1$, respectively. Node $i - 1$ has already packet $r_{i,i+1}$ but not yet packets s_i and $r_{i,i-1}$. Similarly, node $i + 1$ has already packet $r_{i,i-1}$ but not yet packets s_i and $r_{i,i+1}$.

First, we assume that node i does not separate but, rather, combines the transmissions of source and relay packets. If node i transmits coded packet $s_i + r_{i,i-1}$ (or packet $s_i + r_{i,i+1}$), node $i + 1$ can decode packet s_i by computing $(s_i + r_{i,i-1}) + r_{i,i-1}$ (or node $i - 1$ can decode packet s_i by computing $(s_i + r_{i,i+1}) + r_{i,i+1}$). In a subsequent time slot, node i needs to transmit coded packet $r_{i,i-1} + r_{i,i+1}$ so that node $i + 1$ can decode packet $r_{i,i+1}$ by computing $(r_{i,i-1} + r_{i,i+1}) + r_{i,i-1}$ and node $i - 1$ can decode packets $r_{i,i-1}$ and s_i by computing $(r_{i,i-1} + r_{i,i+1}) + r_{i,i+1}$ and

198

$(s_i + r_{i,i-1}) + r_{i,i-1}$, respectively (or so that node $i - 1$ can decode packet $r_{i,i-1}$ by computing $(r_{i,i-1} + r_{i,i+1}) + r_{i,i+1}$ and node $i+1$ can decode packets $r_{i,i+1}$ and s_i by computing $(r_{i,i-1} + r_{i,i-1}) + r_{i,i-1}$ and $(s_i + r_{i,i-1}) + r_{i,i+1}$, respectively).

This operation delivers packets s_i, $r_{i,i-1}$ and $r_{i,i+1}$ to nodes $i-1$ and $i+1$ using two packet transmissions in two time slots and performing two coding operations with a total of two bit additions at node i and two decoding operations with a total of two bit additions at each of the nodes $i - 1$ and $i + 1$. The packet delay is two time slots for both relay packets $r_{i,i-1}$ and $r_{i,i+1}$, whereas the delay of source packet s_i is one time slot for one neighbor node and two time slots for the other one, i.e. the average packet delay is $\frac{7}{4}$ time slots. Alternatively, node i can first transmit packet $s_i + r_{i,i-1}$ (or packet $s_i + r_{i,i+1}$), and then transmit packet $s_i + r_{i,i+1}$ (or packet $s_i + r_{i,i-1}$). This results in the same performance as the previous operation.

Instead, node i can separate the transmissions of source and relay packets. Node i transmits source packet s_i and coded packet $r_{i,i-1} + r_{i,i+1}$ separately such that both neighbor nodes can decode packets s_i, $r_{i,i-1}$ and $r_{i,i+1}$. Nodes $i - 1$ and $i + 1$ directly receive uncoded packet s_i and can decode packets $r_{i,i-1}$ and $r_{i,i+1}$ by computing $(r_{i,i-1} + r_{i,i+1}) + r_{i,i+1}$ and $(r_{i,i-1} + r_{i,i+1}) + r_{i,i-1}$, respectively. This operation requires two transmissions, two time slots, only one coding operation with one bit addition at node i and only one decoding operation with one bit addition at each of nodes $i-1$ and $i+1$. The packet delay is either one or two time slots for any packet depending on the order of transmissions of the source and relay packets s_i and $r_{i,i-1} + r_{i,i+1}$, i.e. the average packet delay is $\frac{3}{2}$ time slots. The same argument can be made, if node i has packets to relay only in one direction rather than in

both directions, as we considered so far. As a result, we can conclude that the separation of source and relay packet transmissions is as good as combining them for network coding purposes in terms of throughput and transmission energy cost and even better in terms of processing energy cost and packet delay. □

According to Lemma 7.2.1, nodes do not need to combine source and relay packets through a network coding operation. The separation of the source and relay packet transmissions offers the operational advantage of facilitating simple strategies for optimal queue management that we will introduce in sections 7.4 and 7.5.

7.3   Cross-Layer Throughput Optimization Problem

For saturated queues, we define $\lambda_{i,j}$ to be the achievable throughput rate from source node i to destination node j in the multicast group M_i. Under stable operation, each source node i independently generates packets to be delivered to destination j in multicast group M_i with rate $\lambda_{i,j}$ according to Bernoulli process.

For the separate operations of network coding and plain routing at the network layer, we need to specify the constraints on the achievable or stable throughput rates $\underline{\lambda} = \{\lambda_{i,j}, i \in N, j \in M_i\}$ depending on whether we consider saturated packet queues or stable operation. These constraints determine the achievable through-put region $\mathcal{A}$ and stability region $\mathcal{S}$ as functions of either the transmission schedules $\underline{t} = \{t_m\}_{m=1}^3$ under scheduled access, or of the transmission probabilities $\underline{p} = \{p_i\}_{i \in N}$ under random access. The dependence of the achievable throughput and stability regions, $\mathcal{A}$ and $\mathcal{S}$, respectively, on $\underline{t}$ or $\underline{p}$ results in the following cross-layer opti-

mization problem:

$$\text{Select } \underline{t} \text{ or } \underline{p} \text{ to find the largest set } \underline{\lambda} \in \mathcal{A} \text{ or } \underline{\lambda} \in \mathcal{S} \qquad (7.1)$$

for network coding or plain routing. Alternatively, we can assess the aggregate throughput performance through a single criterion such as the total sum-delivered throughput

$$\lambda_\Sigma = \sum_{i \in N} \sum_{j \in M_i} \lambda_{i,j} \qquad (7.2)$$

or the minimum transmitted throughput

$$\lambda_{\min} = \min_{i \in N, j \in M_i} \lambda_{i,j}. \qquad (7.3)$$

We consider multicast communication such that each source node i has a destination group M_i to each member of which it wants to transmit at the same rate; i.e. $\lambda_{i,j} = \lambda_i$, $j \in M_i$, for any node $i \in N$, such that $\lambda_\Sigma = \sum_{i \in N} \lambda_i |M_i|$, where $|M_i|$ denotes the size of the multicast group M_i, and $\lambda_{\min} = \min_{i \in N} \lambda_i$. The resulting optimization problem can be formulated as:

$$\text{Select } \underline{\lambda}, \text{ and } \underline{t} \text{ or } \underline{p} \text{ to maximize } \lambda_\Sigma \text{ or } \lambda_{\min} \text{ subject to } \underline{\lambda} \in \mathcal{A} \text{ or } \underline{\lambda} \in \mathcal{S} \qquad (7.4)$$

for network coding or plain routing. The regions $\mathcal{A}$ and $\mathcal{S}$ will serve as the optimization constraints for the aggregate throughput measures λ_Σ and $\lambda_{\min}$.

7.4 Achievable Throughput Region for Saturated Queues

First, we assume saturated packet queues and consider scheduled access. We define N_i^r to be the set of nodes whose packets arrive at node i from the right

direction and need to be forwarded to the left neighbor of node i, and we define N_i^l to be the set of nodes whose packets arrive at node i from the left direction and need to be forwarded to the right neighbor of node i. Let $\Lambda_i^r = \sum_{j \in N_i^r} \lambda_j$ and $\Lambda_i^l = \sum_{j \in N_i^l} \lambda_j$ denote the rate of relay traffic incoming from the right and left neighbor nodes of node i, respectively.

Theorem 7.4.1 *The achievable throughput region $\mathcal{A}$ with rates $\lambda_i \geq 0$, $i \in N$, is given by*

$$\sum_{m=1}^{3} \max_{i \in N: m(i) = m} \left(\lambda_i + \max(\Lambda_i^r, \Lambda_i^l) \right) \leq 1, \tag{7.5}$$

$$\sum_{m=1}^{3} \max_{i \in N: m(i) = m} \left(\lambda_i + \Lambda_i^r + \Lambda_i^l \right) \leq 1 \tag{7.6}$$

for network coding and plain routing, respectively, where $m(i) = (i-1)\,(mod\,3) + 1$, $i \in N$.

Proof: Each node i separately transmits packets it generates and (plain or coded) relay packets for τ_i and $1 - \tau_i$ fractions of time (whenever it is scheduled to transmit). The achievable throughput rates $\underline{\lambda} = \{\lambda_i, i \in N\}$ satisfy

$$0 \leq \lambda_i \leq t_{m(i)} \tau_i, \quad \Lambda_i^r \leq t_{m(i)}(1 - \tau_i), \quad \Lambda_i^l \leq t_{m(i)}(1 - \tau_i), \tag{7.7}$$

$$0 \leq \lambda_i \leq t_{m(i)} \tau_i, \quad \Lambda_i^r + \Lambda_i^l \leq t_{m(i)}(1 - \tau_i) \tag{7.8}$$

for network coding and plain routing, respectively, where $m(i)$ is the time schedule assigned to node i. For any time fraction τ_i, the achievable throughput rates $\lambda_i \geq 0, i \in N$, satisfy

$$\lambda_i + \Lambda_i^r \leq t_{m(i)}, \quad \lambda_i + \Lambda_i^l \leq t_{m(i)}, \tag{7.9}$$

$$\lambda_i + \Lambda_i^r + \Lambda_i^l \leq t_{m(i)} \tag{7.10}$$

for network coding and plain routing, respectively. We sum up the traffic loads over all disjoint time fractions t_1, t_2 and t_3 to obtain the achievable multicast throughput region $\mathcal{A}$ with $\lambda_i > 0$, $i \in N$, described by Eqs. (7.5) and (7.6) for network coding and plain routing, respectively. $\qquad \square$

The achievable throughput region $\mathcal{A}$ described by conditions (7.5) and (7.6) involves only linear constraints for the tandem network model independent of the transmission schedules $\underline{t}$. Thus, the problem (7.4) of maximizing performance measures λ_Σ and $\lambda_{\min}$ can be formulated as a linear optimization problem with linear constraints.

Example: Assume $n = 3$ and consider broadcast communication with multicast groups $M_i = N - \{i\}$, $i = 1, 2, 3$. For network coding, the conditions on the achievable throughput rates $\lambda_i \geq 0$, $i = 1, 2, 3$, are given by $\lambda_i \leq t_i$, $i = 1, 3$, and $\lambda_2 + \lambda_i \leq t_2$, $i = 1, 3$. If $\lambda_1 = 0$ or $\lambda_3 = 0$, we have $\lambda_2 + 2\lambda_3 \leq 1$ or $\lambda_2 + 2\lambda_1 \leq 1$ as the boundary conditions. If $\lambda_2 = 0$, we need to consider two cases: $2\lambda_1 + \lambda_3 \leq 1$ for $\lambda_1 > \lambda_3$ and $\lambda_1 + 2\lambda_3 \leq 1$ for $\lambda_3 > \lambda_1$. The achievable throughput region $\mathcal{A}$ is illustrated in Figure 7.2-(a). Consider the problem of optimizing $\lambda_\Sigma = 2(\lambda_1 + \lambda_2 + \lambda_3)$ subject to the condition $3\lambda_1 + 2\lambda_2 + 3\lambda_3 \leq 2$ on achievable throughput rates $\lambda_i \geq 0$, $i = 1, 2, 3$. The maximum value of λ_Σ is equal to 2 and is achieved by throughput rates $\lambda_1 = \lambda_3 = 0$ and $\lambda_2 = 1$ (i.e. by time allocation of $t_1 = t_3 = 0$ and $t_2 = 1$). The resulting value of $\lambda_{\min}$ is 0. On the other hand, the minimum transmitted throughput rate $\lambda_{\min}$ is maximized for values of λ_i, $i = 1, 2, 3$, such that $\lambda_{\min} \leq t_1$, $2\lambda_{\min} \leq t_2$ and $\lambda_{\min} \leq t_3$. The maximum value of $\lambda_{\min}$ is equal to $\frac{1}{4}$ and is achieved through time allocation given by $t_1 = \frac{1}{4}, t_2 = \frac{1}{2}$ and $t_3 = \frac{1}{4}$. The resulting value of

λ_Σ is $6\lambda_{\min}$, which is equal to $\frac{3}{2}$.

For plain routing, the conditions on the achievable throughput rates $\lambda_i \geq 0$, $i = 1, 2, 3$, are $\lambda_i \leq t_i$, $i = 1, 3$, and $\lambda_1 + \lambda_2 + \lambda_3 \leq t_2$. If $\lambda_1 = 0$, $\lambda_2 = 0$ or $\lambda_3 = 0$, we have $\lambda_2 + 2\lambda_3 \leq 1$, $\lambda_1 + \lambda_2 \leq \frac{1}{2}$ or $\lambda_2 + 2\lambda_1 \leq 1$ as the boundary conditions. The achievable throughput region $\mathcal{A}$ is illustrated in Figure 7.2-(b). Consider the problem of optimizing $\lambda_\Sigma = 2(\lambda_1 + \lambda_2 + \lambda_3)$ subject to the condition $2\lambda_1 + \lambda_2 + 2\lambda_3 \leq 1$ on achievable throughput rates $\lambda_i \geq 0$, $i = 1, 2, 3$. The value of λ_Σ is maximized again to 2, as in the case of network coding, by the achievable throughput rates $\lambda_1 = \lambda_3 = 0$ and $\lambda_2 = 1$ (i.e. by time allocation of $t_1 = t_3 = 0$ and $t_2 = 1$). The resulting value of $\lambda_{\min}$ is 0. On the other hand, the minimum transmitted throughput rate $\lambda_{\min}$ is optimized for values of λ_i, $i = 1, 2, 3$, such that $\lambda_{\min} \leq t_1, 3\lambda_{\min} \leq t_2$ and $\lambda_{\min} \leq t_3$. The value of $\lambda_{\min}$ is maximized to $\frac{1}{5}$ by time allocation of $t_1 = \frac{1}{5}, t_2 = \frac{3}{5}$ and $t_3 = \frac{1}{5}$. The resulting value of λ_Σ is $6\lambda_{\min}$, which is equal to $\frac{6}{5}$.

This example illustrates the trade-off involving the throughput measures λ_Σ and $\lambda_{\min}$ that cannot be simultaneously optimized. Thus, the throughput gains of network coding over plain routing vary. However, these gains can be completely represented in the 3-dimensional throughput region space, as shown in Figure 7.2.

7.5   Stability Region for Possibly Emptying (Non-Saturated) Queues

If we allow packet queues to empty, which is to say that we consider stable operation (i.e. finite queue size operation with external arrivals), it is possible to
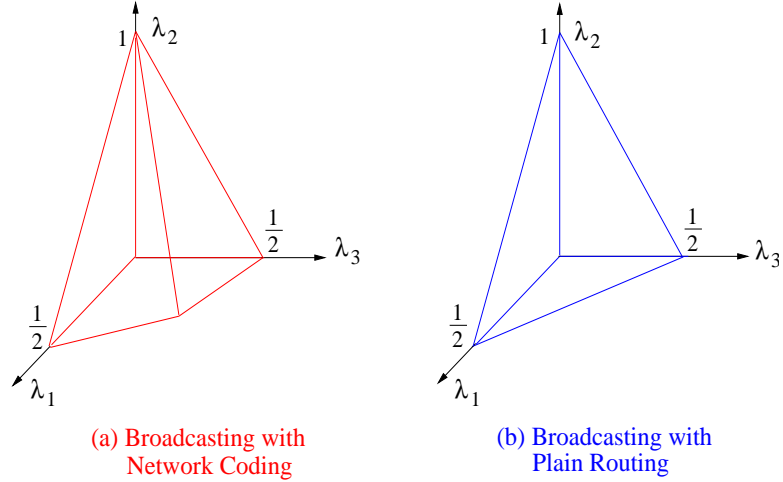
Figure 7.2: Achievable throughput regions for network coding and plain routing solutions under broadcast communication and scheduled access with $n = 3$.

have packet underflow. Thus, any relay node may have to wait for incoming packets over subsequent time slots (and therefore possibly increase the packet delay and reduce the achievable throughput), if it must perform network coding, or proceeds with plain routing of relay packets (and loses the throughput and energy efficiency of the possible network coding solutions). Therefore, conditions (7.5)-(7.6) provide the upper bounds on the stability region $\mathcal{S}$ under scheduled access, if we allow packet queues to empty. We consider only stationary network operation, in which the queue distributions reach steady state. By assuming stationary input processes, a queue is stable, if the arrival and service processes of the queue are all stationary, and the average arrival rate is less than the average service rate [34]. We introduce dynamic network coding (and plain routing) strategies based on the instantaneous queue sizes and allow different priorities (in terms of order of transmission) to relay and source packets (rather than assuming fixed decisions of time-divisioned network

coding as in the case of saturated queues):

Strategy 1: Any node first transmits relay packet(s) by simple forwarding as in plain routing operation, if only one relay queue contains packets, or by network coding, if both relay queues contain packets. Otherwise, that is if both left and right queues are empty, the node transmits a packet from the source queue.

Strategy 2: Any node first transmits a source packet. Only if the source queue is empty, the relay packets are transmitted by either simply forwarding as in plain routing operation, if only one relay queue contains packets, or by network coding, if both relay queues contain packets.

Both strategies 1 and 2 perform the network coding decisions over the first available packets in the relay queues (rather than over all previously received relay packets as done in [68]) to avoid network stability problems and related issues of coding complexity. The stability properties of strategies 1 and 2 are as follows:

Theorem 7.5.1 *(a) For network coding, the stability conditions of strategies 1 and 2 (on the stable throughput rates $\lambda_i \geq 0$, $i \in N$) are given by*

$$\lambda_i < t_{m(i)} \left(1 - \frac{\Lambda_i^r}{t_{m(i)}}\right) \left(1 - \frac{\Lambda_i^l}{t_{m(i)}}\right), \max(\Lambda_i^r, \Lambda_i^l) < t_{m(i)}, i \in N, \quad (7.11)$$

$$\lambda_i + \max(\Lambda_i^r, \Lambda_i^l) < t_{m(i)} \quad, \quad i \in N. \tag{7.12}$$

(b) For plain routing, the stability conditions of strategies 1 and 2 (on the stable throughput rates $\lambda_i \geq 0$, $i \in N$) are both given by

$$\lambda_i + \Lambda_i^r + \Lambda_i^l < t_{m(i)} \quad, \quad i \in N. \tag{7.13}$$

(c) The stability region $\mathcal{S}$ for strategy 1 is strictly suboptimal for network coding, whereas strategy 2 expands the stability region $\mathcal{S}$ to the boundary of the achiev-

able throughput region $\mathcal{A}$:

$$\sum_{m=1}^{3} \max_{i \in N : m(i)=m} \left(\lambda_i + \max(\Lambda_i^r, \Lambda_i^l)\right) < 1. \qquad (7.14)$$

(d) The stability region $\mathcal{S}$ for strategies 1 and 2 is the same under plain routing:

$$\sum_{m=1}^{3} \max_{i \in N : m(i)=m} \left(\lambda_i + \Lambda_i^r + \Lambda_i^l\right) < 1. \qquad (7.15)$$

Proof: Under strategy 1, relay queues Q_i^2 and Q_i^3 have arrival rates Λ_i^r and Λ_i^l, respectively. Under network coding, the service rate for both relay queues is $t_{m(i)}$, since node i can successfully transmit one packet from each relay queue (for $t_{m(i)}$ fraction of time). Therefore, relay queues Q_i^2 and Q_i^3 are empty with probabilities $1 - \frac{\Lambda_i^r}{t_{m(i)}}$ and $1 - \frac{\Lambda_i^l}{t_{m(i)}}$, respectively. Since a source packet is transmitted only if both relay queues are empty, the multicast stability conditions under network coding are given by Eq. (7.11).

Under strategy 2, the source queue is empty with probability $1 - \frac{\lambda_i}{t_{m(i)}}$. Hence, each of the relay queues Q_i^2 and Q_i^3 has the service rate $t_{m(i)} \left(1 - \frac{\lambda_i}{t_{m(i)}}\right)$. The resulting stability condition is the same as the achievable throughput conditions without equalities and given by Eq. (7.12). For plain routing, both queues can be merged such that the total arrival rate is $\Lambda_i^r + \Lambda_i^l$ and the service rate is $t_{m(i)}$ such that the stability condition of Eq. (7.13) holds for both strategies 1 and 2. Next, we evaluate the traffic loads over all disjoint time fractions t_i, $i = 1, 2, 3$, such that $\sum_{i=1}^{3} t_i = 1$. Then, we obtain Eqs. (7.14) and (7.15) by summing up both sides of Eqs. (7.12) and (7.13), respectively, over $\underline{t}$. Also, note that the region described by Eq. (7.11) is strictly smaller than the region described by Eq. (7.12) for any given $\underline{t}$. Therefore, strategy 1 is strictly suboptimal for network coding. $\qquad \square$

An alternative strategy would give higher priority in transmission order to network-coded packets instead of source packets. This may result in the same throughput performance of strategy 2 but cannot improve it further. On an intuitive basis the optimal rule should give the least priority to forwarding uncoded relay packets. One throughput-optimal way of queue management is realized through the use of strategy 2.

Example: Assume $n = 3$ and consider broadcast communication with multicast groups $M_i = N - \{i\}$, $i = 1, 2, 3$. For network coding, the stability conditions are given by $\lambda_1 < t_1$, $\lambda_2 < t_2 \left(1 - \frac{\lambda_1}{t_2}\right)\left(1 - \frac{\lambda_3}{t_3}\right)$ and $\lambda_3 < t_3$ under strategy 1 and $\lambda_1 < t_1$, $\lambda_2 + \max(\lambda_1, \lambda_3) < t_2$ and $\lambda_3 < t_3$ under strategy 2. Strategies 1 and 2 achieve the throughput rates with the common performance bound of $\lambda_\Sigma < 2$ and individual performance bounds of $\lambda_{\min} < \frac{3-\sqrt{5}}{2(4-\sqrt{5})}$ and $\lambda_{\min} < \frac{1}{4}$, respectively. For plain routing, the common stability conditions for strategies 1 and 2 are given by $\lambda_1 < t_1$, $\lambda_1 + \lambda_2 + \lambda_3 < t_2$ and $\lambda_3 < t_3$. We see that strategies 1 and 2 achieve throughput rates with the common performance bounds of $\lambda_\Sigma < 2$ and individual performance bounds of $\lambda_{\min} < \frac{1}{5}$. This example highlights the trade-off between strategies 1 and 2 in terms of throughput measures λ_Σ and $\lambda_{\min}$ under stable operation.

7.6 Throughput Optimization Trade-offs

The achievable and stable throughput regions $\mathcal{A}$ and $\mathcal{S}$ have been derived in sections 7.4 and 7.5, respectively, to specify the multidimensional throughput

properties of a wireless tandem network. From a practical point of view, it is still of interest to summarize the throughput properties in a single variable such as the sum-delivered throughput λ_Σ or the minimum transmitted throughput $\lambda_{\min}$. In this section, we consider the saturated packet queues case and derive the resulting linear optimization problem with linear constraints, which leads to a trade-off between the throughput measures λ_Σ and $\lambda_{\min}$. We consider separately three different traffic profiles, namely multicast, broadcast, and unicast.

7.6.1 Upper Bounds on Multicast Communication

For multicast communication, the most favorable traffic demand (in terms of λ_Σ) is one-hop (closest-neighbor) communication, i.e. multicast communication with multicast groups $M_i = \{i - 1, i + 1\}$, $i \in N - \{1, n\}$, $M_1 = 2$ and $M_n = n - 1$. For both the network coding and the plain routing cases, the value of λ_Σ is maximized to $\frac{2n}{3}$ (by time allocation of $t_1 = 0$, $t_2 = 1$ and $t_3 = 0$), if $n(\mod 3) = 0$, or maximized to $\frac{2n-2}{3}$ (by time allocation of $t_1 = 0$ and $t_2 + t_3 = 1$), if $n(\mod 3) = 1$, or maximized to $\frac{2n-1}{3}$ (by time allocation of $t_1 + t_2 = 1$ and $t_3 = 0$), if $n(\mod 3) = 2$.

The value of $\lambda_{\min}$ is maximized to $\frac{1}{3}$ (by time allocation of $t_m = \frac{1}{3}$, $m = 1, 2, 3$), if $n \geq 3$ (and maximized to $\frac{1}{2}$ if $n = 2$) such that $\lambda_\Sigma = \frac{2(n-1)}{3}$, if $n \geq 3$ (and $\lambda_\Sigma = 1$, if $n = 2$) . These results provide upper bounds on any multicast communication problem including the special cases of broadcast and unicast communication, which we separately consider next in detail.

7.6.2 Broadcast Communication

We consider broadcast communication with $\lambda_{i,j} = \lambda_i$, $j \in M_i = N - \{i\}$ and $i \in N$. For network coding, the value of λ_Σ is maximized to $\frac{2(n-1)}{3}$ by the achievable throughput rates $\lambda_1 = \lambda_n = \frac{1}{3}$, $\lambda_i = 0$, $i = N - \{1,n\}$ (and time allocation of $t_m = \frac{1}{3}, m = 1,2,3$) for $n > 4$. For plain routing, the value of λ_Σ is maximized to $\frac{(n-1)}{3}$ by the achievable throughput rates $\lambda_1 = \lambda_n = \frac{1}{6}$, $\lambda_i = 0$, $i = N - \{1,n\}$ (and time allocation of $t_m = \frac{1}{3}, m = 1,2,3$) for $n > 4$. The optimal value of λ_Σ is equal to $n - 1$ for $n \leq 4$, under both network coding and plain routing. Note that as n increases, the value of λ_Σ increases first for $2 \leq n \leq 4$. Then, the increasing interference effects decrease the value of λ_Σ for $n = 5$ and subsequently slow down the increase in λ_Σ for $n > 5$.

Note that $\lambda_{\min} = 0$, if we optimize λ_Σ. On the other hand, the value of $\lambda_{\min}$ is maximized to $\frac{1}{3n-5}$, if $n(\mathrm{mod}\ 3) = 0$ or 1, and to $\frac{1}{3n-4}$, if $n(\mathrm{mod}\ 3) = 2$. The resulting value of λ_Σ is then $\frac{n(n-1)}{3n-5}$, if $n(\mathrm{mod}\ 3) = 0$ or 1, and $\frac{n(n-1)}{3n-4}$, if $n(\mathrm{mod}\ 3) = 2$. Note that in this case λ_Σ can only approach 50% of its optimal value, as n increases. For plain routing, the value of $\lambda_{\min}$ is maximized to $\frac{1}{3n}$, if $n > 4$, $\frac{1}{2}$, if $n = 2$, $\frac{1}{5}$, if $n = 3$, or $\frac{1}{9}$, if $n = 4$. The optimal values of $\lambda_{\min}$ under network coding and plain routing regimes approach each other, as n increases. Network coding doubles the value of λ_Σ compared to plain routing without any improvement in $\lambda_{\min}$, as n increases. As a result, the objectives of maximizing $\lambda_{\min}$ and λ_Σ cannot be achieved simultaneously.

Figure 7.3 depicts the throughput rate per source-destination pair (namely the value of λ_i averaged over all source nodes $i = 1, ..., n$) that is obtained by separately

optimizing λ_Σ and $\lambda_{\min}$ under broadcast communication.

Figure 7.3: Achievable throughput rates per source-destination pair under broadcast communication and scheduled access.

7.6.3 Unicast Communication

We consider unicast communication with $|M_i| = 1$, $i \in N$. The least favorable unicast demand is that the destination of each packet is chosen as the node that has the largest distance (in number of "hops") from the source node. For network coding, we have $\lambda_\Sigma = n\lambda_{\min}$, where the value of $\lambda_{\min}$ is $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{7}$ for $n = 2, 3, 4$, $\frac{1}{2n}$ for $n = 5, 6, 7$, and $\frac{1}{3\lceil \frac{n}{2} \rceil + 1}$ for $n \geq 8$. For plain routing, the value of $\lambda_{\min}$ is $\frac{1}{2}$, $\frac{1}{5}$, $\frac{1}{9}$ for $n = 2, 3, 4$, and $\frac{1}{3n-1}$ for $n > 4$. The most favorable unicast demand consists of each destination being the one-hop neighbor of the source node. For both network coding and plain routing, the optimal value of $\lambda_{\min}$ is $\frac{1}{3}$, if $n \geq 3$ (and $\frac{1}{2}$, if $n = 2$),

and the optimal value of λ_Σ is $\lfloor \frac{n}{3} \rfloor$, for $n \geq 3$ (and 1, if $n = 2$).
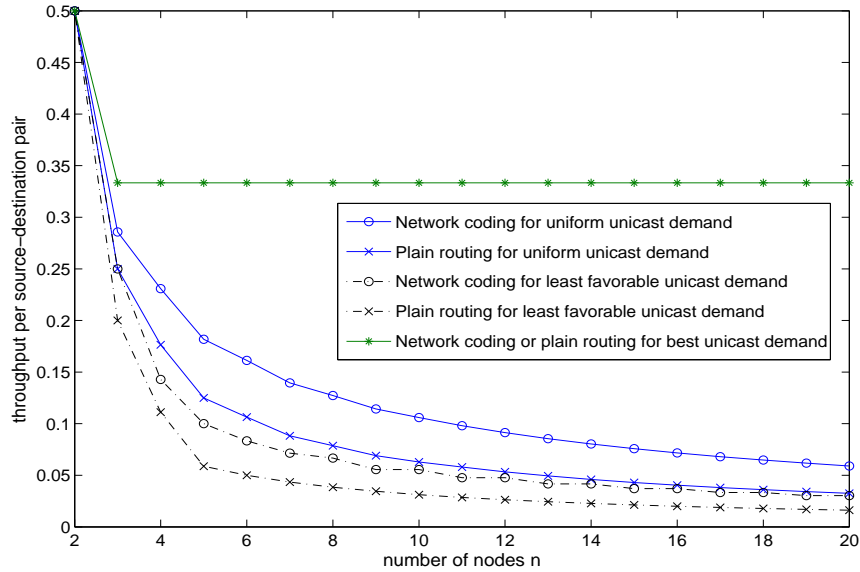
Figure 7.4: Achievable throughput rates per source-destination pair under unicast communication and scheduled access with the optimal throughput rate of λ_Σ.

We consider also the uniform demand case, in which the destination of each source node is randomly chosen from the rest of the nodes. We compare in Figure 7.4 the throughput rates per source-destination pair that are obtained by separately optimizing λ_Σ for different unicast traffic demands. We evaluate in Figure 7.5 the throughput rates per source-destination pair for the cases of optimal temporal allocation (in terms of λ_Σ) and the suboptimal uniform allocation with $t_m = \frac{1}{3}, m = 1, 2, 3$. Any deviation from optimal MAC operation (that needs to be separately chosen for network coding or plain routing) can result in significant throughput losses. These results highlight the interdependence of the MAC and network layer operations and the need for joint cross-layer design. The throughput gains of network coding over
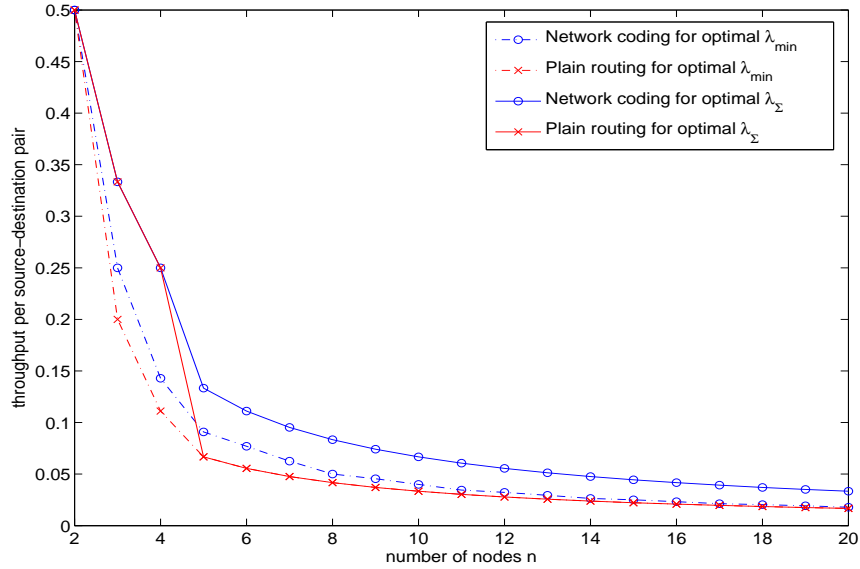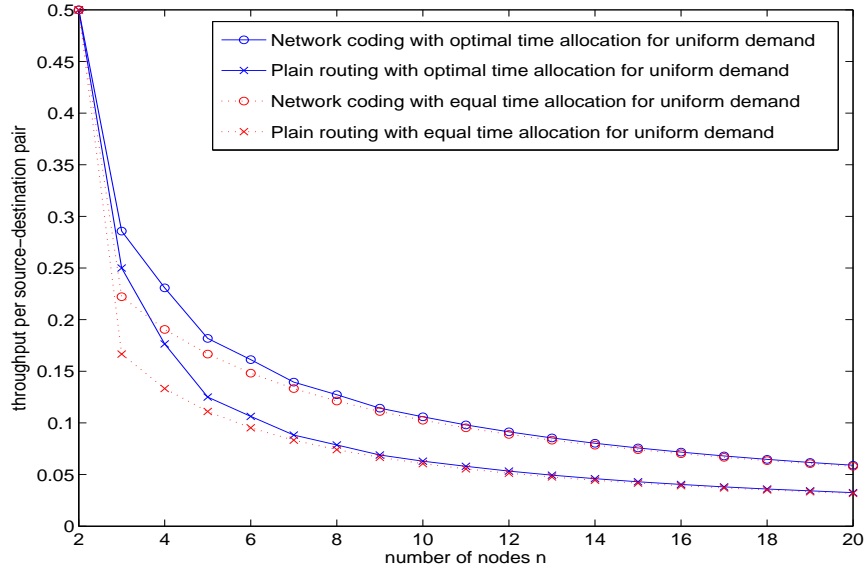
Figure 7.5: Achievable throughput rates per source-destination pair under unicast communication and scheduled access with uniform demand for the optimal temporal allocation (in terms of λ_Σ) and for $t_m = \frac{1}{3}, m = 1, 2, 3$.

plain routing achieve the largest values for intermediate values of the number of nodes in the network and may diminish, as the network size grows. Since broadcast communication generates more packet traffic to relay at each node and fills relay queues faster than unicast communication, there are more opportunities in the broadcast case to combine relay packets; therefore network coding is expected to yield higher throughput benefits in that case.

7.6.4 Joint Optimization of Throughput Rates λ_Σ and $\lambda_{\min}$

The objectives of maximizing λ_Σ and $\lambda_{\min}$ cannot be achieved simultaneously. As we saw, the value of $\lambda_{\min}$ can be equal to 0 under broadcast communication,

whenever λ_Σ is optimized. For broadcast communication, network coding doubles the value of λ_Σ under plain routing, as n goes to infinity, whereas the improvement in $\lambda_{\min}$ diminishes to zero. For unicast communication, network coding can double both λ_Σ and $\lambda_{\min}$, as n goes to infinity. These results illustrate that the throughput trade-offs depend on the traffic profile.

Another approach is to maximize the weighted sum of λ_Σ and $\lambda_{\min}$. This is equivalent to maximizing λ_Σ subject to $\lambda_{\min} \geq \alpha$ for some positive constant α, i.e. subject to $\lambda_{i,j} \geq \alpha$ for all $i \in N$ and $j \in M_i$. We can solve the resulting optimization problem by the Lagrange multipliers method. For $n > 4$, the optimal value of λ_Σ is $\frac{n-1}{3}(2 - \alpha(3n - 2c))$, where $c = 5$, if $n(\mathrm{mod}\ 3) = 0$ or 1, and $c = 4$, if $n(\mathrm{mod}\ 3) = 2$ for network coding, and $\lambda_\Sigma(\alpha) = \frac{n-1}{3}, 0 \leq \alpha \leq \frac{1}{3n}$, for plain routing. We depict in Figure 7.6 the optimal value of λ_Σ as a function of the parameter α.
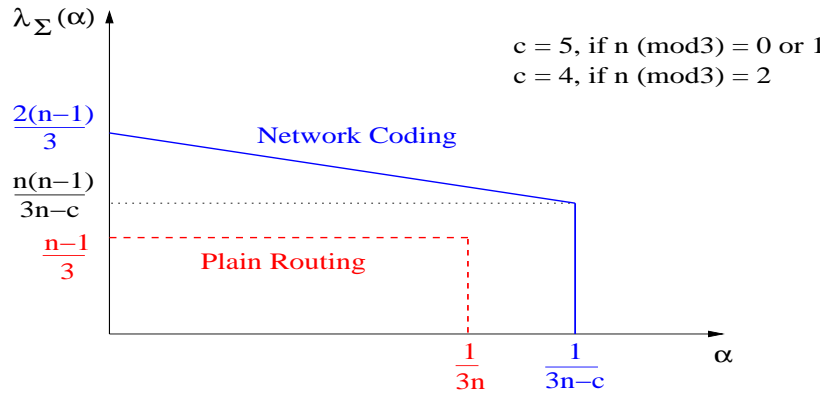
Figure 7.6: The optimal value of λ_Σ for $\lambda_{\min} \geq \alpha$ under broadcast communication and scheduled access for $n > 4$.

If we require stable operation, and hence, allow packet queues to empty, then, under network coding, strategy 2 approaches the same throughput rates as in the

case of saturated queues, whereas strategy 1 has a strictly worse throughput performance. On the other hand, if we consider only plain routing, both strategies 1 and 2 can approach the same throughput performance as in the case of saturated queues. We evaluate in Figure 7.7 the effects of saturated queues and stable operation on the throughput objectives of λ_Σ and $\lambda_{\min}$.
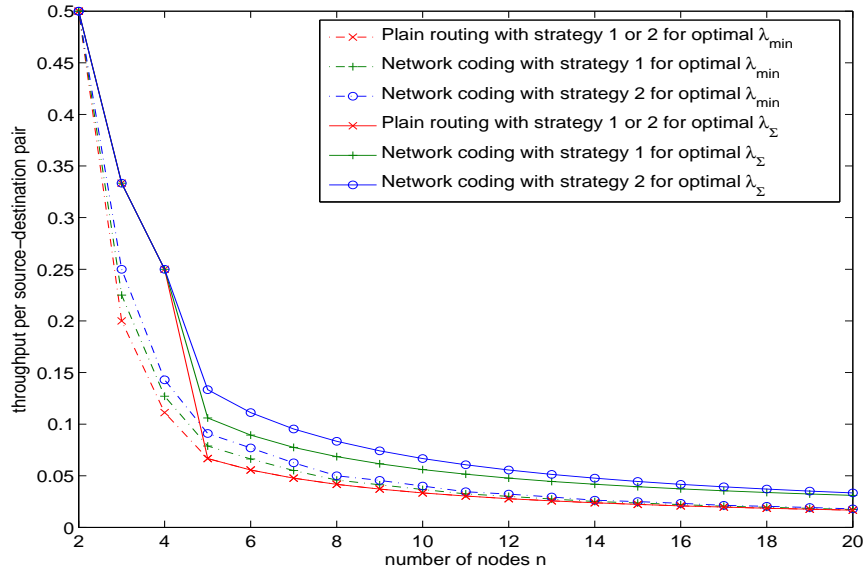
Figure 7.7: Stable throughput rates per source-destination pair under broadcast communication and scheduled access.

7.7 Energy Properties and Trade-offs with Throughput Objectives

Energy consumption in wireless systems is of course of great concern. Thus, it is useful to consider the energy efficiency of the alternative strategies and objectives that we have considered so far. Let $\mathcal{E}_t$ be the transmission energy cost of each packet transmission and let $\mathcal{E}_c$ be the processing energy cost of a coding or decoding

operation (namely the energy cost of binary vector addition) and let $\mathcal{E}_f$ be the processing energy cost of plain forwarding.

7.7.1 Transmission and Processing Energy Costs for Saturated Queues

First, we consider the case of saturated queues with uninterrupted availability of source and relay packets at each node.

Theorem 7.7.1 *(a) The total transmission energy cost $E_t(\underline{\lambda})$ per time slot to achieve throughput rates $\underline{\lambda} \in \mathcal{A}$ is given by*

$$E_t(\underline{\lambda}) = \sum_{i=1}^{n} \mathcal{E}_t \left(\lambda_i + \max(\Lambda_i^r, \Lambda_i^l) \right), \qquad (7.16)$$

$$E_t(\underline{\lambda}) = \sum_{i=1}^{n} \mathcal{E}_t \left(\lambda_i + \Lambda_i^r + \Lambda_i^l \right) \qquad (7.17)$$

for network coding and plain routing, respectively.

(b) The total processing energy cost $E_p(\underline{\lambda})$ per time slot to achieve throughput rates $\underline{\lambda} \in \mathcal{A}$ is given by

$$E_p(\underline{\lambda}) = \sum_{i=1}^{n} (3\mathcal{E}_c - \mathcal{E}_f) \min(\Lambda_i^r, \Lambda_i^l) + \mathcal{E}_f \max(\Lambda_i^r, \Lambda_i^l), \qquad (7.18)$$

$$E_p(\underline{\lambda}) = \sum_{i=1}^{n} \mathcal{E}_f(\Lambda_i^r + \Lambda_i^l) \qquad (7.19)$$

for network coding and plain routing, respectively.

Proof: First, we consider the total transmission energy cost $E_t(\underline{\lambda})$ per time slot to achieve throughput rates $\underline{\lambda}$. Node i transmits source packets with rate $\lambda_i \in \mathcal{A}$ incurring transmission energy cost $\mathcal{E}_t \lambda_i$ per time slot. Node i receives relay packets with rates Λ_i^r and Λ_i^l from the right and left neighbors. For network coding, node i

consumes $\mathcal{E}_t$ amount of energy to transmit one coded packet from both relay queues Q_i^2 and Q_i^3. Since the relay packets arrive at queues Q_i^2 and Q_i^3 with rates Λ_i^r and Λ_i^l, respectively, the total amount of energy per time slot consumed by node i to relay packets is $\mathcal{E}_t \max(\Lambda_i^r, \Lambda_i^l)$. For plain routing, each relay packet from any of the relay queues is separately transmitted with cost $\mathcal{E}_t$ such that the total amount of energy per time slot consumed by node i to relay packets is $\mathcal{E}_t(\Lambda_i^r + \Lambda_i^l)$. As a result, we obtain the energy cost given by Eq. (7.16) for network coding operation and by Eq. (7.17) for plain routing operation.

Next, we consider the processing energy cost $E_p(\underline{\lambda})$ per time slot to achieve the throughput rates $\underline{\lambda}$. For network coding, the total amount of energy consumed to achieve $\underline{\lambda}$ is given by $\sum_{i=1}^n 3\mathcal{E}_c \min(\Lambda_i^r, \Lambda_i^l) + \mathcal{E}_f(\max(\Lambda_i^r, \Lambda_i^l) - \min(\Lambda_i^r, \Lambda_i^l))$, since relay node i performs a coding operation with rate $\min(\Lambda_i^r, \Lambda_i^l)$ and each coding operation is accompanied by two decoding operations at neighboring nodes, whereas the other packets are simply forwarded with rate $\max(\Lambda_i^r, \Lambda_i^l) - \min(\Lambda_i^r, \Lambda_i^l)$ and processing cost $\mathcal{E}_f$. For plain routing, the total amount of energy consumed to achieve $\underline{\lambda}$ is $\sum_{i=1}^n \mathcal{E}_f(\Lambda_i^r + \Lambda_i^l)$, since any node i forwards relay packets with rate $\Lambda_i^r + \Lambda_i^l$, respectively. As a result, we obtain Eq. (7.18) for network coding and Eq. (7.19) for plain routing operation. □

As shown in section 7.4, network coding strictly improves the achievable throughput region $\mathcal{A}$ over plain routing. However, whether network coding is more energy-efficient than plain routing, strongly depends on the particular method of managing (source and relay) queues and hardware constraints that are reflected in the cost parameters $\mathcal{E}_t$, $\mathcal{E}_c$ and $\mathcal{E}_f$. Network coding reduces the total (transmission

and processing) energy cost per packet (namely $E_t(\underline{\lambda}) + E_p(\underline{\lambda})$ for any given $\underline{\lambda} \in \mathcal{A}$) compared to plain routing, if and only if

$$3\mathcal{E}_c < \mathcal{E}_t + 2\mathcal{E}_f \ . \tag{7.20}$$

As a result, the throughput and energy costs may not be simultaneously optimized depending on the energy cost parameters. For $\mathcal{E}_t = 1$, $\mathcal{E}_c = \frac{1}{3}$ and $\mathcal{E}_f = 0$, we evaluate in Figure 7.8 the transmission and processing energy costs per source-destination pair to achieve the optimal throughput rates (in terms of $\lambda_{\min}$) under unicast and broadcast communication. We may need higher energy costs in order for the throughput of network coding to exceed that of plain routing.
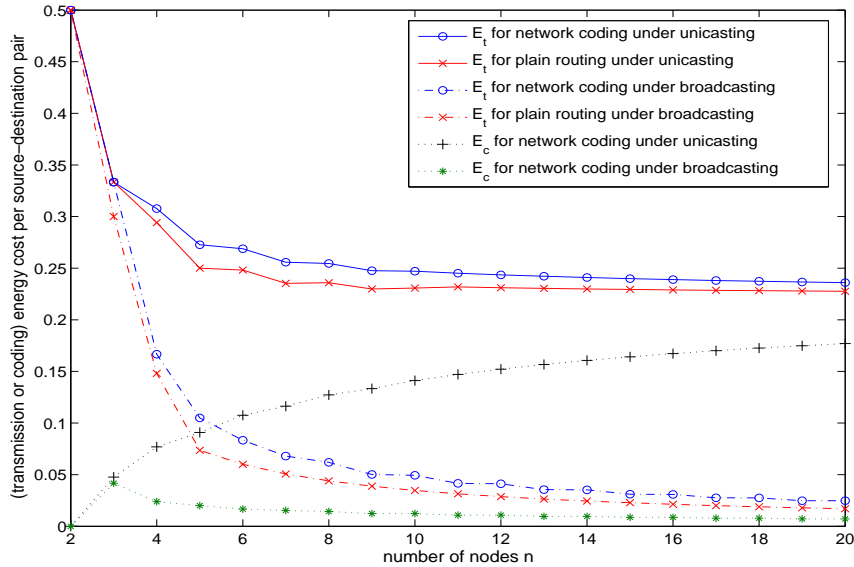
Figure 7.8: Transmission and processing energy costs per source-destination pair to achieve the optimal throughput rate of $\lambda_{\min}$ under unicast and broadcast communication with scheduled access.

7.7.2 Transmission and Processing Energy Costs in Stable Operation

The energy costs under stability are given by the next theorem.

Theorem 7.7.2 *(a) For network coding, the total transmission energy cost $E_t(\underline{\lambda})$ per time slot to achieve throughput rates $\underline{\lambda} \in \mathcal{S}$ is given by*

$$E_t(\underline{\lambda}) = \sum_{i=1}^{n} \mathcal{E}_t \left(\lambda_i + \Lambda_i^r + \Lambda_i^l - \frac{\Lambda_i^r \Lambda_i^l}{t_{m(i)}} \right), \tag{7.21}$$

$$E_t(\underline{\lambda}) = \sum_{i=1}^{n} \mathcal{E}_t \left(\lambda_i + \Lambda_i^r + \Lambda_i^l - \frac{\Lambda_i^r \Lambda_i^l}{t_{m(i)} - \lambda_i} \right) \tag{7.22}$$

under strategy 1 and 2, respectively.

(b) For network coding, the total processing energy cost $E_p(\underline{\lambda})$ per time slot to achieve throughput rates $\underline{\lambda} \in \mathcal{S}$ is given by

$$E_p(\underline{\lambda}) = \sum_{i=1}^{n} \frac{3\mathcal{E}_c \Lambda_i^r \Lambda_i^l}{t_{m(i)}} + \mathcal{E}_f \left(\left(1 - \frac{\Lambda_i^r}{t_{m(i)}} \right) \Lambda_i^l + \left(1 - \frac{\Lambda_i^l}{t_{m(i)}} \right) \Lambda_i^r \right), \tag{7.23}$$

$$E_p(\underline{\lambda}) = \sum_{i=1}^{n} \frac{3\mathcal{E}_c \Lambda_i^r \Lambda_i^l}{t_{m(i)} - \lambda_i} + \mathcal{E}_f \left(\left(1 - \frac{\Lambda_i^r}{t_{m(i)} - \lambda_i} \right) \Lambda_i^l + \left(1 - \frac{\Lambda_i^l}{t_{m(i)} - \lambda_i} \right) \Lambda_i^r \right) \tag{7.24}$$

under strategy 1 and 2, respectively.

(c) For plain routing, both strategies achieve the values given by Eqs. (7.22) and (7.24) for the transmission and processing energy costs $E_t(\underline{\lambda})$ and $E_p(\underline{\lambda})$.

Proof: Node i transmits a source packet with rate λ_i and incurs transmission energy cost $\mathcal{E}_t$. Under strategy 1, node i transmits a relay packet, if there exists at least one packet in relay queue Q_i^2 or Q_i^3. This occurs with probability $1 - \left(1 - \frac{\Lambda_i^l}{t_{m(i)}} \right) \left(1 - \frac{\Lambda_i^r}{t_{m(i)}} \right)$, since relay queues Q_i^2 and Q_i^3 have arrival rates Λ_i^r and Λ_i^l and common service rate $t_{m(i)}$ and they are empty with probabilities $1 - \frac{\Lambda_i^r}{t_{m(i)}}$ and $1 - \frac{\Lambda_i^l}{t_{m(i)}}$, respectively. Hence, the relay packet transmissions incur the energy cost

$\mathcal{E}_t \left(\frac{\Lambda_i^l + \Lambda_i^r}{t_{m(i)}} - \frac{\Lambda_i^l \Lambda_i^r}{t_{m(i)}^2} \right)$. Since node i is activated for $t_{m(i)}$ fraction of the time, the total

transmission energy cost per time slot under network coding is given by Eq. (7.21)

for strategy 1.

Under strategy 2, node i transmits a relay packet, if there exists no source

packet at queue Q_i^1 and there exists at least one packet in Q_i^2 or Q_i^3. The service

rate is $t_{m(i)}$ for queue Q_i^1 and $t_{m(i)} \left(1 - \frac{\lambda_i}{t_{m(i)}} \right)$ for both queues Q_i^1 and Q_i^3, since node

i is scheduled to transmit for $t_{m(i)}$ fraction of the time and queue Q_i^1 is idle with

probability $1 - \frac{\lambda_i}{t_{m(i)}}$, and the relay packets from Q_i^2 and Q_i^3 are transmitted only

with probability $1 - \frac{\lambda_i}{t_{m(i)}}$ during $t_{m(i)}$ fraction of time. Since queues Q_i^1, Q_i^2 and Q_i^3

are empty with respective probabilities $1 - \frac{\lambda_i}{t_{m(i)}}$, $1 - \frac{\frac{\Lambda_i^r}{t_{m(i)}}}{1 - \frac{\lambda_i}{t_{m(i)}}}$ and $1 - \frac{\frac{\Lambda_i^l}{t_{m(i)}}}{1 - \frac{\lambda_i}{t_{m(i)}}}$, the total

transmission energy cost per time slot under network coding is given by Eq. (7.22)

for strategy 2.

For plain routing, both strategies 1 and 2 approach the same throughput rates

and therefore the same transmission and processing energy costs as in the case of

saturated queues. Next, we consider the case of network coding. Under strategy

1, the coding operation is performed at node i for $t_{m(i)}$ fraction of the time, if

both relay queues have at least one packet, which occurs with probabilities $\frac{\Lambda_i^r}{t_{m(i)}}$

and $\frac{\Lambda_i^l}{t_{m(i)}}$, i.e. the processing cost for network coding is $\frac{3\mathcal{E}_c \Lambda_i^r \Lambda_i^l}{t_{m(i)}}$. Otherwise, the

relay packets packets are forwarded for $t_{m(i)}$ fraction of the time with probability

$\left(1 - \frac{\Lambda_i^r}{t_{m(i)}} \right) \Lambda_i^l + \left(1 - \frac{\Lambda_i^l}{t_{m(i)}} \right) \Lambda_i^r$ (namely with the probability that only one relay queue

has packet(s)). On the other hand, node i forwards the source packets with rate λ_i

and incurs the processing energy cost $\mathcal{E}_f$. As a result, the total processing energy

cost per time slot under network coding is given by Eq. (7.23) for strategy 1.

Under strategy 2, the coding operation is performed at node i for $t_{m(i)} \left(1 - \frac{\lambda_i}{t_{m(i)}} \right)$ fraction of the time (since source queue Q_i^1 is empty with probability $1 - \frac{\lambda_i}{t_{m(i)}}$ and node i is scheduled to transmit for $t_{m(i)}$ fraction of the time), if both relay queues have packets (which occurs with probabilities $\frac{\frac{\Lambda_i^r}{t_{m(i)}}}{1 - \frac{\lambda_i}{t_{m(i)}}}$ and $\frac{\frac{\Lambda_i^l}{t_{m(i)}}}{1 - \frac{\lambda_i}{t_{m(i)}}}$); that is, the processing cost for network coding operation is $\frac{3\mathcal{E}_c \frac{\Lambda_i^r \Lambda_i^l}{t_{m(i)}}}{1 - \frac{\lambda_i}{t_{m(i)}}}$. Otherwise, relay packets are forwarded for $t_{m(i)} \left(1 - \frac{\lambda_i}{t_{m(i)}} \right)$ fraction of the time with probability $\left(1 - \frac{\frac{\Lambda_i^r}{t_{m(i)}}}{1 - \frac{\lambda_i}{t_{m(i)}}} \right) \Lambda_i^l + \left(1 - \frac{\frac{\Lambda_i^l}{t_{m(i)}}}{1 - \frac{\lambda_i}{t_{m(i)}}} \right) \Lambda_i^r$ (namely with the probability that only one relay queue has packet(s)). Node i forwards source packets with rate λ_i and incurs the processing energy cost $\mathcal{E}_f$. As a result, the total processing energy cost per time slot under network coding is given by Eq. (7.24) for strategy 2. $\square$

As was shown in section 7.5, strategy 2 strictly improves the stable throughput region $\mathcal{S}$ compared to strategy 1 for the case of network coding. However, whether strategy 1 or 2 is more energy-efficient, strongly depends on the values of energy costs $\mathcal{E}_t$, $\mathcal{E}_c$ and $\mathcal{E}_f$. Under network coding, for any set of stable throughput rates $\underline{\lambda} \in \mathcal{S}$, the transmission energy cost $E_t(\underline{\lambda})$ is lower for strategy 2 when $\mathcal{E}_t > 0$, whereas the processing energy cost $E_p(\underline{\lambda})$ is lower for strategy 1, if and only if $\mathcal{E}_c > \frac{2\mathcal{E}_f}{3}$. The total (transmission and processing) energy cost per packet (namely $E_t(\underline{\lambda}) + E_p(\underline{\lambda})$) is lower for strategy 2, if and only if $3\mathcal{E}_c < \mathcal{E}_t + 2\mathcal{E}_f$.

The energy costs are non-linear functions of the time allocation (i.e. of the MAC rules) for the case of stable operation (compared to the case of saturated queues) and lead to a non-linear objective function for energy cost optimization.

However, the optimization constraints are non-linear for strategy 1, but linear for strategy 2. If nodes wait to accumulate relay packets to perform network coding (rather than proceeding with plain routing) in the case of packet underflow, the transmission energy costs can be reduced at the expense of additional packet delay and poor stable throughput properties compared to strategies 1 and 2.

7.8   Extension to Random Access

Although scheduled access has distinct advantages (especially at heavy traffic conditions), some form of random access is unavoidable in wireless systems. At a minimum, it is necessary on the reservation sub-channel in dynamic allocation protocols. Therefore, the cross-layer design framework for throughput region optimization should also incorporate contention-based random access. We will restrict our attention to the case of saturated queues, because, otherwise, the node queues interact (see [12, 11, 13]) and lead to formidable difficulties of analysis.

We assume that each node randomly chooses between transmitting a source or a relay packet. Specifically, any transmission of node i consists of a relay packet (coded or uncoded) with probability β_i and a source packet with probability $1 - \beta_i$. Each relay packet (in coded or plain form) needs to be delivered to the neighbor receiver at one side only (since the receiver at the other side already has that packet). However, any source packet may need to be delivered to the two neighbor receivers at both sides (as in broadcast communication). Clearly, under random access, a source packet that is successfully received by one neighbor node may fail at the

other neighbor node. Therefore, we need to carefully track the methods for the transmissions of source and relay packets under random access (whereas under the conflict-free scheduling that was considered so far, there was no possibility of such partially successful transmission).

7.8.1 Methods for Source Packet Transmissions in Random Access

We consider three separate methods for transmitting source packets.

Method A: Each node retransmits a source packet, until it is successfully received by all intended neighbor nodes in the same time slot (i.e. partially successful transmissions are ignored).

Method B: Nodes transmit a new source packet only if the previous source packet has been received by all intended neighbor nodes. This method is based on the repetition of transmissions until all intended receivers have successfully received the transmitted source packet over successive time slots (but not necessarily in the same time slot).

Method C: Each node computes a linear combination of the source packets that have not been decoded yet by the intended neighbor nodes, and transmits the corresponding coded packet. Based on receiver feedback, each source packet of node i is put to a virtual queue depending on which receiver has successfully received that packet. We assume three virtual queues within the source queue Q_i^1 for node i. All source packets arrive at queue $Q_i^{1(a)}$. If the packet has been successfully received by both receivers, the packet leaves the system. If the packet has not been received

by either one of the two receiver nodes, the packet remains in queue $Q_i^{1(a)}$. Packets that have been received by the left neighbor receiver but not by the right neighbor receiver enter queue $Q_i^{1(b)}$. Similarly, packets that have been received by the right neighbor receiver but not by the left neighbor receiver enter queue $Q_i^{1(c)}$. Then, a packet leaves queue $Q_i^{1(b)}$ and $Q_i^{1(c)}$, if it is received by the right or left receiver, respectively. We assume that all queues are saturated and packets from queues $Q_i^{1(b)}$ and $Q_i^{1(c)}$ are combined by linear coding before transmission.

The coding benefits can be illustrated for the case in which source packets x and y of node i must be delivered to both neighbor nodes $i - 1$ and $i + 1$. If the transmission of packet x by node i is only received by node $i + 1$, node i transmits packet $x + y$ (instead of packet x as in method B). If packet $x + y$ is successfully received, node i needs to deliver packet y only to node $i - 1$ (rather than to both nodes $i - 1$ and $i + 1$ as in method B). This is the generalization of the concept of network coding for rateless communication [91] to the case of packet overflow with infinite delay by taking into account the mutual interference effects.

7.8.2 Achievable Throughput Region in Random Access

Let s_i^r and s_i^l be the probability that a transmission of node i is successfully received by the right and left neighbor node, respectively, i.e. we have $s_i^r = (1 - p_{i+1})(1 - p_{i+2})$ for $1 \leq i \leq n - 2$, $s_{n-1}^r = 1 - p_n$, $s_n^r = 1$, $s_i^l = (1 - p_{i-2})(1 - p_{i-1})$ for $3 \leq i \leq n$, $s_1^l = 1$ and $s_2^l = 1 - p_1$.

Theorem 7.8.1 *Under random access, the achievable throughput region $\mathcal{A}$ with*

throughput rates $\lambda_i \geq 0$, $i \in N$, satisfy

$$s_i^l \lambda_i + \gamma_i(s_i^r, s_i^l)\Lambda_i^r \leq p_i s_i^l \gamma_i(s_i^r, s_i^l), \;\; s_i^r \lambda_i + \gamma_i(s_i^r, s_i^l)\Lambda_i^l \leq p_i s_i^r \gamma_i(s_i^r, s_i^l), (7.25)$$

$$s_i^r s_i^l \lambda_i + s_i^r \gamma_i(s_i^r, s_i^l)\Lambda_i^r + s_i^l \gamma_i(s_i^r, s_i^l)\Lambda_i^l \leq p_i s_i^l s_i^r \gamma_i(s_i^r, s_i^l) \tag{7.26}$$

for network coding and plain routing, respectively, where

$$\gamma_i(s_i^r, s_i^l) = \begin{cases} s_i^r s_i^l & , \text{ for method } A \\ \frac{s_i^{l2} s_i^{r2} + s_i^{l2} s_i^r (1-s_i^r) + s_i^{r2} s_i^l (1-s_i^l)}{s_i^l s_i^r + s_i^{l2}(1-s_i^r) + s_i^{r2}(1-s_i^l)} & , \text{ for method } B \\ \min(s_i^r, s_i^l) & , \text{ for method } C \end{cases} \tag{7.27}$$

Proof: Let α_i^l and α_i^r be the probability that the packet relayed by node i is intended for the left neighbor node $i - 1$ or right neighbor node $i + 1$, respectively. We have $\alpha_i^r = \alpha_i^l = 1$ and $\alpha_i^r + \alpha_i^l = 1$, $i \in N$, for network coding and plain routing, respectively.

For network coding, the relay queues Q_i^2 and Q_i^3 have arrival rates Λ_i^r and Λ_i^l according to the Max-flow Min-cut Theorem. The service rates of the queues Q_i^2 and Q_i^3 are given by $p_i \beta_i s_i^l$ and $p_i \beta_i s_i^r$. If the coded relay packet is not successfully received by both neighbor nodes, the collided packet needs to be retransmitted. However, if only one neighbor receives the coded relay packet successfully, then the packet intended for the unsuccessful receiver is combined with a new packet from the queue of the successfully transmitted packet for the next network coding operation. The achievable throughput rates satisfy $\sum_{j \in N_i^r} \lambda_j \leq p_i \beta_i s_i^l$ and $\sum_{j \in N_i^l} \lambda_j \leq p_i \beta_i s_i^r$ for the transmissions from relay queues Q_i^2 and Q_i^3, respectively.

Since node i transmits with probability p_i and any transmission of node i carries a source packet with probability $1 - \beta_i$, the source packet transmissions

impose the constraint $\lambda_i \leq p_i(1 - \beta_i)\gamma_i(s_i^r, s_i^l)$, where $\gamma_i(s_i^r, s_i^l)$ is the rate at which

a source packet transmission is successfully delivered to both neighbor nodes.

If the transmission of a source packet from node i is not successfully received by

all intended neighbor nodes, a retransmission is required, although that particular

packet may have been already received by one of the neighbor nodes. In method A,

the probability of successful transmission of a source packet is given by $s_i^r s_i^l$ provided

that node i transmits a source packet, which happens with probability $p_i(1 - \beta_i)$.

Thus, the achievable throughput satisfies the condition $\lambda_i \leq p_i(1 - \beta_i)\gamma_i(s_i^r, s_i^l)$,

where $\gamma_i(s_i^r, s_i^l) = s_i^r s_i^l$. In method B, the transmission of source packets of node

i follows a Markov Chain shown in Figure 7.9, where $\mu_i^r = p_i(1 - \beta_i)s_i^r$ and $\mu_i^l = p_i(1 - \beta_i)s_i^l$. In state $(0,1)$ or $(1,0)$, only node $i - 1$ or $i + 1$ has received the packet

of node i. In state $(0,0)$, the packet has been received by both nodes $i - 1$ and $i + 1$

or equivalently node i is ready to transmit a new source packet. Let Π_j denote the

stationary distribution of state $j \in \{(0,0), (0,1), (1,0)\}$.
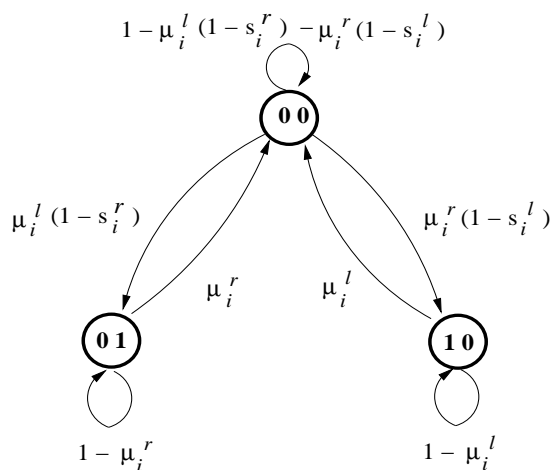
Figure 7.9: Markov chain model for the transmission of packets from source queue Q_i^1 at node i.

226

The achievable throughput rate for source packets of node i satisfies

$$\lambda_i \leq \Pi_{(0,0)}\mu_i^l s_i^r + \Pi_{(0,1)}\mu_i^r + \Pi_{(1,0)}\mu_i^l, \tag{7.28}$$

which is given by $\lambda_i \leq p_i(1-\beta_i)\gamma_i(s_i^r, s_i^l)$, where $\gamma_i(s_i^r, s_i^l) = \frac{s_i^{l\,2}s_i^{r\,2}+s_i^{l\,2}s_i^r(1-s_i^r)+s_i^{r\,2}s_i^l(1-s_i^l)}{s_i^l s_i^r + s_i^{l\,2}(1-s_i^r) + s_i^{r\,2}(1-s_i^l)}$.

The rate at which a source packet is successfully transmitted to both neighbor nodes has the upper bound $\min(s_i^r, s_i^l)$, which is equal to the maximum flow rate over the minimum broadcast cut. In method C, saturated queues $Q_i^{1(b)}$ and $Q_i^{1(c)}$ are served with rates s_i^r and s_i^l, such that $\lambda_i \leq s_i^r$ and $\lambda_i \leq s_i^l$. Method C can achieve the rate $\min(s_i^r, s_i^l)$ for source packet transmissions provided that node i transmits some source packet. The throughput rate λ_i achievable by Method C satisfies the condition $\lambda_i \leq p_i(1-\beta_i)\gamma_i(s_i^r, s_i^l)$, where $\gamma_i(s_i^r, s_i^l) = \min(s_i^r, s_i^l)$.

As a result, the achievable throughput rates satisfy

$$\Lambda_i^r \leq p_i\beta_i s_i^l \alpha_i^l, \quad \Lambda_i^l \leq p_i\beta_i s_i^r \alpha_i^r \quad \text{and} \quad \lambda_i \leq p_i(1-\beta_i)\gamma_i(s_i^r, s_i^l), \tag{7.29}$$

where $\gamma_i(s_i^r, s_i^l) = s_i^r s_i^l$ for method A, $\gamma_i(s_i^r, s_i^l) = \frac{s_i^{l\,2}s_i^{r\,2}+s_i^{l\,2}s_i^r(1-s_i^r)+s_i^{r\,2}s_i^l(1-s_i^l)}{s_i^l s_i^r + s_i^{l\,2}(1-s_i^r) + s_i^{r\,2}(1-s_i^l)}$ for method B and $\gamma_i(s_i^r, s_i^l) = \min(s_i^r, s_i^l)$ for method C. Then, conditions (7.25)-(7.26) directly follow from Eq. (7.29) by eliminating β_i. $\qquad\square$

The achievable throughput region $\mathcal{A}$ under random access involves linear constraints on throughput rates for any given set of fixed transmission probabilities $\underline{p}$. It can be verified that the achievable throughput region $\mathcal{A}$ under random access is optimized by Method C and has the upper bounds given by the achievable throughput conditions under scheduled access.

7.8.3 Throughput Optimization in Random Access

First, we consider the maximization of λ_Σ. Although the constraints on $\mathcal{A}$ are independent of the transmission schedules $\underline{t}$ under scheduled access, they are non-linear functions of the transmission probabilities $\underline{p}$ under random access. The Logarithmic Barrier Method [92] can be used to solve the resulting problem of linear optimization with non-linear constraints. The achievable throughput rates per source-destination pair are depicted in Figure 7.10 for broadcast communication. Method C optimizes the throughput performance in terms of maximizing λ_Σ for both network coding and plain routing operations, whereas method A has the worst throughput performance, which, however, approaches that of methods B and C, as the number of nodes n increases.
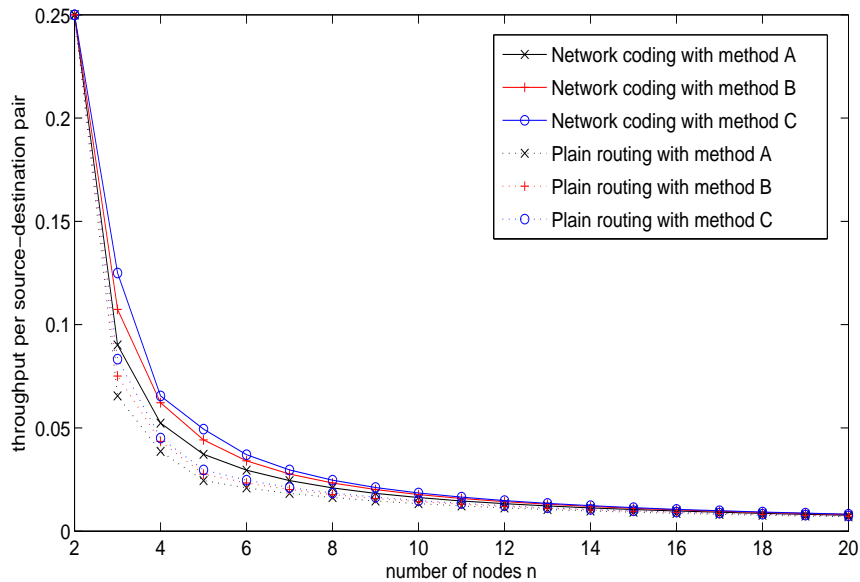
Figure 7.10: Achievable throughput rates per source-destination pair under broadcast communication and random access with the optimal throughput rate of λ_Σ.

Next, we consider the problem of maximizing $\lambda_{\min}$ for broadcast communication under method C. For network coding, the resulting conditions on the minimum throughput rate are given by $\lambda_{\min} < p(1-\beta)(1-p)^2$ and $(n-2)\lambda_{\min} \leq p\beta(1-p)^2$ for source and relay packets, respectively, with common values of $p_i = p$ and $\beta_i = \beta$ for $i \in N$. The combined condition is $\lambda_{\min} < \frac{p(1-p)^2}{n-1}$. The conditions for source and relay packets approach each other for sufficiently large n and the right hand side of the resulting inequality is maximized by $p = \frac{1}{3}$ to the value of $\frac{4}{27(n-1)}$. For plain routing, the conditions on the minimum throughput rate are given by $\lambda_{\min} < p(1-\beta)(1-p)^2$ and $(n-1)\lambda_{\min} < p\beta(1-p)^2$ for source and relay packets, respectively. The combined condition is $\lambda_{\min} < \frac{p(1-p)^2}{n}$ and the right hand side of this inequality is maximized by $p = \frac{1}{3}$ to the value of $\frac{4}{27n}$. As n increases, the maximum value of $\lambda_{\min}$ approaches $\frac{4}{9}$ of the throughput value of $\lambda_{\min}$ that is achievable under scheduled access under network coding or plain routing.

7.9 Non-Cooperative Network Coding Operation

Non-cooperative operation has been studied before in the context of plain routing for general network topologies [53, 90] and cooperation mechanisms have been presented to simulate packet forwarding by intermediate relay nodes [93, 94, 95]. We looked at the single-receiver MAC operation for non-cooperative transmitters in Chapters 3 and 4, and extended the analysis to the cross-layer design of plain routing and MAC in Chapter 5. In this section, we extend non-cooperative operation to network coding strategies (or plain routing strategies as a special case) in

conjunction with MAC.

We assume that each node i is rational and has the selfish objective of maximizing its own utility function u_i. The strategy space of node i is the set $\lambda_i^j, j \in N$, where λ_i^j is defined as the rate at which node i transmits packets of node j such that $\Lambda_i^r = \sum_{j \in N_i^r} \lambda_i^j$ and $\Lambda_i^l = \sum_{j \in N_i^l} \lambda_i^j$. The utility of node $i \in N$ is defined as function of $\lambda_j^i, j \in N$, to represent the total multicast throughput $|M_i|\lambda_i$, transmission energy cost $E_{t,i}$ and processing energy cost $E_{p,i}$ as

$$u_i(\underline{\lambda}_i, \underline{\lambda}_{-i}) = |M_i|\lambda_i - E_{t,i}(\underline{\lambda}) - E_{p,i}(\underline{\lambda}), \tag{7.30}$$

where $\underline{\lambda}_i = \{\lambda_i^j\}_{j \in N}$ and $\underline{\lambda}_{-i} = \{\underline{\lambda}_j, j \in N - \{i\}\}$. The components of u_i are $\lambda_i = \min(\lambda_i^i, \{\lambda_j^i\}_{j \in R_i})$, where R_i is the set of relay nodes for i, $E_{t,i}(\underline{\lambda}) = \mathcal{E}_t(\lambda_i^i + \max(\Lambda_i^r, \Lambda_i^l))$ for network coding, $E_{t,i}(\underline{\lambda}) = \mathcal{E}_t(\lambda_i^i + \Lambda_i^r + \Lambda_i^l)$ for plain routing, $E_{p,i}(\underline{\lambda}) = \sum_{k \in \{-1,0,1\}} \mathcal{E}_c \min(\Lambda_{i+k}^r, \Lambda_{i+k}^l) + \mathcal{E}_f(\max(\Lambda_i^r, \Lambda_i^l) - \min(\Lambda_i^r, \Lambda_i^l))$ for network coding, and $E_{p,i}(\underline{\lambda}) = \mathcal{E}_f(\Lambda_i^r + \Lambda_i^l)$ for plain routing. We define the best response function of node i as $\underline{b}_i(\underline{\lambda}_{-i}) = \text{argmax}_{\underline{\lambda}_i} u_i(\underline{\lambda}_i, \underline{\lambda}_{-i})$. Then, $\underline{\lambda}_i^*, i \in N$, is a Nash equilibrium strategy (such that no node can unilaterally improve its utility, if the strategies of other nodes remain the same), if and only if $\underline{\lambda}_i^* \in \underline{b}_i(\underline{\lambda}_{-i}^*)$ for all $i \in N$.

Theorem 7.9.1 *(a) For zero energy costs $\mathcal{E}_t$, $\mathcal{E}_c$ and $\mathcal{E}_f$, any achievable set of rates $\lambda_i, i \in N$, results in a non-cooperative Nash equilibrium strategy.*

(b) For $\mathcal{E}_t > 0$, $\mathcal{E}_c > 0$ or $\mathcal{E}_f > 0$, the unique non-cooperative Nash equilibrium strategies are $\lambda_i^j = 0$ for any $i \in N$, $j \in N - \{i\}$, and $\lambda_i^i = 0$, if $|R_i| > 0$ or $|R_i| = 0$ and $\mathcal{E}_t > |M_i|$. Otherwise, $\lambda_i^i = \frac{1}{2}$ for $n = 2$ and $\frac{1}{3}$ for $n \geq 3$.

Proof: (a) For zero energy costs $\mathcal{E}_t$, $\mathcal{E}_c$ and $\mathcal{E}_f$, the best response of node i given

any $\underline{\lambda}_{-i}$ is to set $\lambda_i^i \geq \lambda_j^i$ for $j \in R_i$. However, λ_i is upper bounded by λ_j^i for $j \in R_i$ and therefore the throughput rates $\lambda_j \in \mathcal{A}$ directly impose the throughput rate λ_i that cannot be improved individually by node i. Therefore, any achievable set of rates $\lambda_i, i \in N$, results in a non-cooperative Nash equilibrium strategy.

(b) For non-zero energy costs $\mathcal{E}_t$, $\mathcal{E}_c$ or $\mathcal{E}_f$, the best response of node $i \in N$ given any $\underline{\lambda}_{-i}$ is $\lambda_i^j = 0$ for $j \in N - \{i\}$ to minimize the energy costs $E_{t,i}(\underline{\lambda})$ and $E_{p,i}(\underline{\lambda})$. If the destination group M_i of node i only includes one-hop neighbor nodes (i.e. $|R_i| = 0$) and if the throughput reward is greater than total energy cost (i.e $|M_i| > \mathcal{E}_t$), the optimal strategy λ_i^i of node i is $\frac{1}{2}$ for $n = 2$ and $\frac{1}{3}$ for $n \geq 3$ to maximize the throughput λ_i and the utility u_i. Otherwise, the best strategy of node i is $\lambda_i^j = 0$ for any node j to minimize the energy costs and maximize u_i such that $\lambda_i = 0$. $\quad\square$

Since λ_Σ and $\lambda_{\min}$ are zero in non-cooperative equilibrium with multihop communication, we need to stimulate the cooperation of nodes. For non-zero costs, we can define a dependency graph of nodes with edges from i to j, if there exist a route where i is a relay node and j is a source [96]. The Tit-for-Tat cooperation mechanism that relies on nodes to mimic each other's strategies imposes $\lambda_i^k = 0$, if there exists no cycle in dependency graph including i and k. However, this results in $\lambda_{\min} = 0$ and can only achieve suboptimal values of λ_Σ.

7.9.1 Reward-Based Cooperation Stimulation

We introduce a reward-based cooperation mechanism [97] to improve the non-cooperative equilibrium. Each node i pays reward r to each relay node for the

throughput unit carried to motivate packet relaying. When a node i computes its equilibrium strategies, it must consider that the other nodes will respond by paying or requesting rewards for packet forwarding purposes. Thus, the problem faced by node i is not that of optimizing u_i with respect to $\underline{\lambda}_i$ considering fixed $\underline{\lambda}_{-i}$. Node i should know a priori that any relay node j is ready to pay $r\lambda_i^j$ to node $i \in R_j$ and node i must pay $r\lambda_j^i$ to node $j \in R_i$. The utility u_i has two additional components of $r\sum_{j \in N_i^r \cup N_i^l} \lambda_i^j$ (the reward for relaying packets of other nodes) and $r\sum_{j \in R_i} \lambda_j^i$ (the cost charged by nodes for relaying packets of node i). For network coding, we assume for simplicity that each coding node pays $\mathcal{E}_c$ to neighbors for any coded packet such that the coding nodes undertake all decoding costs and no effective cost is incurred for packet decoding at their neighbor nodes. Next, we evaluate the dependence of the non-cooperative strategies on the value of r.

Lemma 7.9.1 *The unique non-cooperative Nash equilibrium strategy of node i with rewarding is $\lambda_i^j = 0$ for all j, if $r > \frac{|M_i| - \mathcal{E}_t}{|R_i|}$ for all $i \in N$, or if $r < \min(\frac{\mathcal{E}_t + 3\mathcal{E}_c}{2}, \mathcal{E}_t + \mathcal{E}_f)$ for network coding, or if $r < \mathcal{E}_t + \mathcal{E}_f$ for plain routing.*

Proof: Node i achieves the total throughput $|M_i|\lambda_i$ at the expense of energy cost $\mathcal{E}_t\lambda_i$ and relaying cost $r|R_i|\lambda_i$. If $r > \frac{|M_i| - \mathcal{E}_t}{|R_i|}$, u_i is a decreasing function of λ_i and node i does not have any incentive to transmit source packets, i.e. $\lambda_i = 0$. Node i cannot unilaterally improve the utility u_i for fixed $\underline{\lambda}_{-i}$ and sets $\lambda_i^j = 0$ for any other node j to minimize the energy costs. For plain routing, any node i receives reward $r\lambda$ for relaying packets with rate λ at the expense of energy cost $(\mathcal{E}_t + \mathcal{E}_f)\lambda$. For fixed $\underline{\lambda}_{-i}$, relaying packets will decrease the utility u_i and therefore $\lambda_i^j = 0$, if $r < \mathcal{E}_t + \mathcal{E}_f$.

For network coding, any node i receives reward $2r\lambda$ or $r\lambda$ for relaying packets of two nodes or one node with rate λ at the expense of energy cost $(\mathcal{E}_t + 3\mathcal{E}_c)\lambda$ or $(\mathcal{E}_t + \mathcal{E}_f)\lambda$. For fixed $\underline{\lambda}_{-i}$, relaying packets will decrease the utility u_i and the best strategy is $\lambda_i^j = 0$, if $r < \min(\frac{\mathcal{E}_t + 3\mathcal{E}_c}{2}, \mathcal{E}_t + \mathcal{E}_f)$. $\qquad\square$

Next, we show that the non-cooperative operation with rewarding can reach the cooperative equilibrium performance under the assumption of scheduled access and saturated queues.

Theorem 7.9.2 *For broadcast communication with network coding, the unique non-cooperative equilibrium strategies with rewarding coincide with the cooperative equilibrium strategies with the optimal value of λ_Σ, if $\frac{n-1-\mathcal{E}_t}{n-2} > r > \max(\frac{n-1+3\mathcal{E}_c}{n-1}, \frac{\mathcal{E}_t + 3\mathcal{E}_c}{2})$.*

Proof: For network coding, the cooperative equilibrium strategies with optimal λ_Σ are achievable, if (a) node $i \in \{1, n\}$ with $|M_i| = n - 1$ and $|R_i| = n - 2$ prefers transmitting source packets, i.e. the utility of transmitting source packets, namely $|M_i| - r|R_i| - \mathcal{E}_t$, should be greater than the utility of staying idle, namely 0, such that $r < \frac{n-1-\mathcal{E}_t}{n-2}$, (b) node $i \in N - \{1, n\}$ with $|R_i| = n - 3$ prefers relaying packets over transmitting source packets and staying idle, i.e. the utilities of transmitting source packets and staying idle, namely $|M_i| - r|R_i| - \mathcal{E}_t$ and 0, should be smaller than the utility of relaying packets by network coding or plain routing, namely $2r - \mathcal{E}_t - 3\mathcal{E}_c$ or $r - \mathcal{E}_t - \mathcal{E}_f$, such that $r > \max(\frac{n-1+3\mathcal{E}_c}{n-1}, \frac{\mathcal{E}_t + 3\mathcal{E}_c}{2})$. $\qquad\square$

For network coding, we can achieve the Pareto optimal network operation in terms of λ_Σ, which is not possible for plain routing with $\mathcal{E}_t > 0$ or $\mathcal{E}_f > 0$. However, the cooperative and non-cooperative equilibrium values of $\lambda_{\min}$ are equal for both

network coding and plain routing.

Theorem 7.9.3 *For broadcast communication, the unique non-cooperative Nash equilibrium strategies with rewarding coincide with the cooperative equilibrium strategies with the optimal value of $\lambda_{\min}$, if $r < \min(\frac{n-1-\mathcal{E}_t}{n-2}, \frac{n-1+3\mathcal{E}_c}{n-1})$ under network coding, and if $r < \frac{n-1-\mathcal{E}_t}{n-2}$ under plain routing.*

Proof: For network coding or plain routing, the cooperative equilibrium strategies with optimal $\lambda_{\min}$ are achievable, if (a) every node i has incentive to transmit source packets, i.e. $\lambda_i^i(|M_i| - r|R_i| - \mathcal{E}_t) > 0$, where $|M_i| = n - 1$ and $|R_i| = n - 2$ for $i \in \{1, 2\}$ and $|R_i| = n - 3$ for $i \in N - \{1, n\}$, such that $r < \frac{|M_i| - \mathcal{E}_t}{|R_i|}$, (b) node $i \in N - \{1, n\}$ with $|R_i| = n - 3$ prefers transmitting source packets over relaying packets, i.e. $\lambda_i^i(|M_i| - r|R_i| - \mathcal{E}_t) > \lambda_i^i(2r - \mathcal{E}_t - 3\mathcal{E}_c)$ or $\lambda_i^i(r - \mathcal{E}_t - \mathcal{E}_f)$ for network coding or plain routing. Thus, each node can individually improve the throughput to the maximum common throughput $\lambda_{\min}$. $\qquad\square$

The next question is whether nodes should prefer network coding or plain routing.

Corollary 7.9.1 *For $r < \min(\frac{n-1-\mathcal{E}_t}{n-2}, \frac{n-1+3\mathcal{E}_c}{n-1})$, the non-cooperative strategies with rewarding use network coding for broadcast communication, if $3\mathcal{E}_c < \mathcal{E}_t + 2\mathcal{E}_f$.*

Proof: If we assume $r < \min(\frac{n-1-\mathcal{E}_t}{n-2}, \frac{n-1+3\mathcal{E}_c}{n-1})$, the non-cooperative equilibrium strategies with rewarding can maximize the value of $\lambda_{\min}$ for network coding, as shown in Theorem 7.9.3. The utility of node i is $\lambda_i(n - 1 - \mathcal{E}_t \max(|N_i^r|, |N_i^l|) - 3\mathcal{E}_c \min(|N_i^r|, |N_i^l|) - \mathcal{E}_f(\max(|N_i^r|, |N_i^l| - \min(|N_i^r|, |N_i^l|)))$. For plain routing, the utility of node i is $\lambda_i(n - 1 - \mathcal{E}_t(|N_i^r| + |N_i^l|) - \mathcal{E}_f(|N_i^r| + |N_i^l|))$. If $3\mathcal{E}_c < \mathcal{E}_t + 2\mathcal{E}_f$, selfish

nodes prefer network coding operations that reduce the energy costs and also improve the individual throughput rates. $\qquad\square$

Next, we allow the packet queues to empty, and evaluate the dependence of the non-cooperative equilibrium strategies 1 and 2 on the value of r.

Theorem 7.9.4 *For possibly emptying queues, node i follows strategy 1, if $r > \frac{|M_i|+3\mathcal{E}_c}{|R_i|+2}$ under network coding or if $r > \frac{|M_i|+\mathcal{E}_f}{|R_i|+1}$ under plain routing, and follows strategy 2 otherwise.*

Proof: Node i prefers strategy 1, if the utility of relaying packets by network coding and plain routing, namely $2r - \mathcal{E}_t - 3\mathcal{E}_c$ or $r - \mathcal{E}_t - \mathcal{E}_f$, is greater than the utility of transmitting source packets, namely $|M_i| - r|R_i| - \mathcal{E}_t$, i.e. if $r > \frac{|M_i|+3\mathcal{E}_c}{|R_i|+2}$ under network coding or if $r > \frac{|M_i|+\mathcal{E}_f}{|R_i|+1}$ under plain routing. On the other hand, each node prefers strategy 2, if transmitting source packets increases the utility compared to relaying packets. $\qquad\square$

7.9.2 Non-Cooperative Random Access

For random access, we assume saturated queues. The possible actions of a node i are waiting (W) or transmitting (T_i^j) a packet of node j, i.e. the action space of node i is $A_i = \{W, \{T_i^j\}_{j \in N_i^r \cup N_i^l}\}$. Let β_i^j denote the probability that the transmission of node i carries a packet of node j, where $0 \le \beta_i^j \le 1$, $\sum_{j \in i \cup N_i^r \cup N_i^l} \beta_i^j = 1$ for plain routing, or $\beta_i^i + \max(\sum_{j \in N_i^r} \beta_i^j, \sum_{j \in N_i^l} \beta_i^j) = 1$ for network coding. The mixed strategies of nodes are the probability distributions over the actions such that mixed strategy of node i is $\sigma_i = \{p_i, \beta_i^j, j \in N\}$. We define $\underline{\sigma} = \{\sigma_i, i \in N\}$,

$\underline{\sigma}_{-i} = \{\sigma_j, j \in N - \{i\}\}$ and $u_i(A|\underline{\sigma}_{-i})$ as the utility of node i provided that node i plays action $A \in A_i$ given $\underline{\sigma}_{-i}$. The mixed Nash equilibrium strategies are $\underline{\sigma}^*$, if for each node i the expected utility $u_i(A|\underline{\sigma}^*_{-i})$ to every action $A \in A_i$ (in support of σ_i^*) is the same for any given $\underline{\sigma}^*_{-i}$. For reward-based cooperation stimulation, the unique non-cooperative Nash equilibrium strategy of each node i is $p_i = 0$, if $r > \frac{|M_i| - \mathcal{E}_t}{|R_i|}$ for all $i \in N$, or if $r < \min(\frac{\mathcal{E}_t + 3\mathcal{E}_c}{2}, \mathcal{E}_t + \mathcal{E}_f)$ under network coding, or if $r < \mathcal{E}_t + \mathcal{E}_f$ under plain routing. The choice of the source packet transmission methods A, B or C is reflected in $\gamma_i(s_i^r, s_i^l)$ for any node i.

Theorem 7.9.5 *Define $v_{i,j} = s_i^r$ for $i < j$, $v_{i,j} = s_i^l$ for $i > j$ and $v_{i,j} = \gamma_i(s_i^r, s_i^l)$ for $i = j$. For $j \in N_i^r \cup N_i^l$, the non-cooperative Nash equilibrium strategies in random access satisfy*

$$(|M_i| - r|R_i|)v_{i,i} = \mathcal{E}_t, \beta_i^j = \frac{p_j \beta_j^j v_{j,j}}{p_i v_{i,j}}, i \in R_j, \text{for network coding or plain routing,}$$

$$\max(2r - 3\mathcal{E}_c, r - \mathcal{E}_f)v_{i,j} = \mathcal{E}_t, \beta_i^i = 1 - \max(\sum_{j \in N_i^r} \beta_i^j, \sum_{j \in N_i^l} \beta_i^j) \text{ for network coding,}$$

$$(r - \mathcal{E}_f)v_{i,j} = \mathcal{E}_t, \beta_i^i = 1 - \sum_{j \in N_i^r \cup N_i^l} \beta_i^j \text{ for plain routing.}$$

Proof: The non-cooperative Nash equilibrium strategies of any node i follow from the equalities of the conditional utilities such that $u_i(W|\underline{\sigma}_{-i}) = u_i(T_i^j|\underline{\sigma}_{-i})$ for all $j \in N_i^r \cup N_i^l$, where $u_i(W|\underline{\sigma}_{-i}) = 0$, $u_i(T_i^i|\underline{\sigma}_{-i}) = -\mathcal{E}_t + (|M_i| - r|R_i|)v_{i,i}$ (under the constraint that condition $p_i \beta_i^j v_{i,j} \le p_j \beta_j^j v_{j,j}$ for $i \in R_j$ is satisfied with equality such that the rate at which packets are relayed is optimized to the rate at which packets are generated) and $u_i(T_i^j|\underline{\sigma}_{-i}) = -\mathcal{E}_t + \max(2r - 3\mathcal{E}_c, r - \mathcal{E}_f)v_{i,j}$ for network coding or $u_i(T_i^j|\underline{\sigma}_{-i}) = -\mathcal{E}_t + (r - \mathcal{E}_f)v_{i,j}$ for plain routing. □

7.10 Summary and Conclusions

We illustrated the complex interactions between MAC and network layer variables in a simple wireless network topology. We focused on throughput regions, as well as simple (scalar) throughput measures under both stable and unstable buffer conditions. We evaluated the advantages of network coding over plain routing under either scheduled or random access and we also considered the energy efficiency aspect. The main result of our analysis is that even for the simple tandem topology, there are serious and complex trade-offs involved regarding the choices of MAC, queue management, and routing/coding.

The analysis needs to be extended to more general topologies, e.g. two-dimensional grid networks. A dynamic programming argument has been used in [98] to decompose a triangular lattice network into star and tandem subnetworks, and general network codes have been derived by combining network codes for simpler subnetworks. Thus the ingredients for such an extension are available. A similar idea can be applied to the stability analysis of network coding, where the local stability conditions for smaller subnetworks (as derived in this chapter for tandem subnetworks) can be combined to derive the global stability conditions for more general network topologies.

Furthermore, packet transmissions need not be completely reliable and may also reach beyond the one-hop neighbor nodes. In addition, the multicast traffic scenarios can be modified to allow anycast communication, under which packets from each source node can be transmitted to different destination nodes at different rates.

This requires an extension of the concept of the stability and throughput regions by allowing not only different source nodes but also distinguishing the throughput rates for different destination nodes in the same multicast group. The ultimate goal is to analyze network coding for arbitrary packet traffic models and general network topologies. The model could be further extended to incorporate the case of finite energy supplies that result in finite node and network lifetime. Finally, the performance objectives should incorporate the packet delay properties under stable operation. In this context, we also need to consider the practical case of limited queue capacities and evaluate the buffer overflow probabilities (in lieu of the stability conditions for the case of infinite buffer capacities). We continue in Chapter 8 with a closer look at the additional stability benefits of network coding for single-hop multicast communication.

Chapter 8

Effects of Network Coding on Queueing Stability in Wireless Access

The stability benefits of network coding are not limited to multi-hop communication problem that we discussed in Chapter 7. Network coding can also improve the stable throughput rates for single-hop multicast communication. In this chapter, we specify the stable throughput region for broadcast systems with one or two source nodes transmitting packets to multiple receiver nodes over independent channels with probabilistic reception. We show that the plain retransmission policy is suboptimal and that the stable operation is optimized by coded retransmissions with finite packet delay and limited complexity. First, we consider a linear random coding scheme and point at the benefits of coding in terms of stable throughput at the expense of additional processing costs. Then, we introduce a simple dynamic coding scheme to prove the equivalence of the queueing stability region and the maximum throughput region for random access systems with multi-packet reception channels. We also discuss the relationship between the maximum achievable throughput and stability regions and the general capacity region. Finally, we explore the maximum stable throughput region for unicast traffic of packets addressed to different receivers and combine the results with broadcast packet traffic.

8.1 Introduction

The maximum throughput and stability properties have been extensively studied for random access systems with multiple transmitters and a single receiver [12]. The maximum stable throughput and capacity regions are equivalent for the case of two source nodes randomly transmitting packets to a common receiver over collision channels [99]. The stability problem in random access has been extended in [21] to probabilistic multi-packet reception. The conjecture has been made in [22] that both regions coincide for general multi-packet reception channels. The extension to multicast communication with multiple receivers has been introduced in Chapter 7 through the application of network coding for the case of saturated queues. It has been proved in [100] that the queueing stability region differs from the capacity region for compound random access systems in which two source nodes randomly send packets to two receivers. The stability analysis has been based on the simple policy of plain packet retransmissions with feedback from receivers to transmitters on the channel outcome [101].

In this chapter, we show that the plain retransmission policy is suboptimal in terms of the maximum stable throughput rates and it is possible to optimize the stability region to the maximum achievable throughput region by employing a simple (network) coding scheme (with packet retransmissions) that differs from the forward error correction by (channel) coding and can operate with finite packet delay and possibly emptying queues. The proposed scheme is different from the rateless codes (such as LT codes [102] or Raptor codes [103]) that are used to transmit blocks of

packets without concern of delay buildup.

The maximum throughput region of tandem networks has been derived in Chapter 7 using the idea of network coding and the plain retransmission policies have been shown to be suboptimal for random access under the assumption of saturated queues. If we consider only one-hop transmissions of source packets (rather than combining them with relay packet traffic), then this result also applies to the case of non-saturated, i.e. possibly emptying, packet queues.

In this chapter, we derive the stability region for random access of two source nodes transmitting packets to two receiver nodes. We prove that the stability and maximum achievable throughput regions coincide through the use of coded packet retransmissions based on feedback from receivers to transmitters. This result deviates from the previous approaches based on plain retransmission policies that are strictly suboptimal and cannot achieve the maximum stable throughput rates in broadcast systems with one or two transmitter nodes (namely over the equivalent broadcast erasure channels between one or two transmitter nodes and two receivers).

First, we allow the transmissions of random linear packet combinations and evaluate the performance gains compared to repetition-based transmission policies. Network coding improves the stable throughput and transmission energy costs at the expense of additional processing costs for coding and decoding purposes and additional delay of packets in queues. We evaluate the stable throughput performance depending on the field size for linear network coding operations and the length of blocks over which the queue content is combined.

Then, we formulate a simple dynamic network coding scheme that is based

on the instantaneous queue content and can operate with finite delay and possibly emptying queues. The feedback information needs to be exchanged between transmitter and receiver nodes for retransmission policies. On the other hand, channel coding can achieve the broadcast capacity without feedback [104] but only with infinite packet delay. Thus, simple network coding on packet level with feedback can optimize the throughput region in stable operation (with finite packet delay and computational complexity) and can approach the capacity region derived in [100].

The chapter is organized as follows. In section 8.2, we introduce the single-source broadcast system model. In section 8.3, we look at the retransmission policies. We extend the analysis to random network coding policies in section 8.4. We generalize the ideas of coded retransmissions in section 8.5 and derive the optimal network coding policy to achieve the maximum stable throughput rate. We discuss the feedback, overhead and complexity properties of retransmission policies with and without coding. Then, we extend the results to random access of two transmitters in section 8.6 and prove that the stability region and the maximum achievable throughput region coincide. Then, we derive the stable throughput region for unicast traffic in section 8.7. We draw conclusions in section 8.8.

8.2   Single-Source Broadcast System Model

One transmitter node randomly generates messages to be encoded into packets. We assume that packets arrive with rate λ to be transmitted to both receivers 1 and 2 over two independent channels with probabilistic reception q_1 and q_2, as

shown in Figure 8.1. Only a packet received by both receivers contributes to the achievable throughput. The system is considered stable, if the length of packet queue Q at transmitter node is asymptotically finite, i.e. the packet delay is finite. We allow immediate and correct feedback on the channel outcome from the receivers to the transmitter node. The system is equivalent to a broadcast erasure channel.
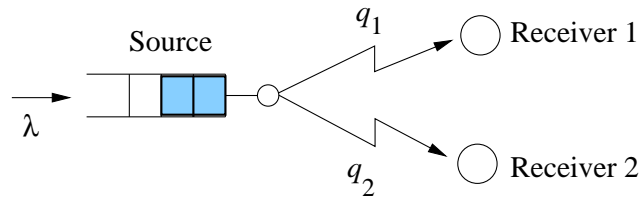

Figure 8.1: System model with one transmitter and two receivers.

The outcome of each independent broadcast channel is a success or failure, namely erasure. The erased (i.e. not successfully received) packets need to be retransmitted. Consider two broadcast cuts that separate the source node from each of the receivers. The packets can be successfully received by receiver $i = 1, 2$ with maximum throughput rate q_i. Then, the maximum throughput rate for common packet traffic is given by $\min(q_1, q_2)$.

If we assume that packets correspond to codeword symbols, the information-theoretic capacity C of the resulting broadcast erasure channel is $\min(q_1, q_2)$. This follows from the upper bound on the mutual information over two broadcast cuts that separate the transmitter node from each of the two receivers [75] and can be regarded as a special case of the capacity results for two source nodes in [100]. Channel coding results in reliable packet communication with rate R up to $\min(q_1, q_2)$ in the absence of feedback but only under the assumptions of saturated queues (i.e.

always availability of information symbols to be transmitted) and infinite packet delay.

8.3 Suboptimal Retransmission Policies

In simple transmission policy (STP), a packet transmission is considered to be successful, if the packet is successfully received by both receivers 1 and 2 at the same time slot. Otherwise, the packet is retransmitted. Transmitter node receives immediate feedback from receivers on the status on whether the transmitted packet is successfully received by both receivers or not. The feedback is transmitted on error-free feedback channels based on scheduled access. Only reception of a packet by all intended receivers contributes to the throughput. The service rate of the packet queue at transmitter node is $\mu = q_1 q_2$. The resulting stability condition is given by $\lambda < q_1 q_2$.

In plain retransmission policy (PRP), any packet failed at any of the two receivers is retransmitted until it is successfully received by both receivers. This plain retransmission policy (PRP) has been considered before in [100, 101] and [105, 36, 106] for broadcast communication. Immediate and accurate channel feedback is necessary from receivers to transmitter node to inform whether a packet has been successfully received or not. Section 8.4 shows that PRP is suboptimal (unless q_1 or q_2 is equal to 0 or 1) for random access in tandem networks with saturated queues. In section 8.5, we show that a coded retransmission policy is necessary for the case of possibly emptying queues to reach the information-theoretic capacity.

The service rate of PRP has been derived in [100] for broadcast systems with two source nodes using a Markov chain model for packet retransmissions. If we specialize it to the single-source case, the resulting service rate of the packet queue Q at transmitter node is given by

$$\mu = \frac{q_1 q_2 (q_1 + q_2 - q_1 q_2)}{(q_1 + q_2)(q_1 + q_2 - q_1 q_2) - q_1 q_2} \ . \tag{8.1}$$

Alternatively, we can derive the service rate, as proposed in [91] for rateless communication with arbitrary number of receiver nodes. We define $\mu = \frac{1}{E[T]}$, where T is the service time. Let T_i define the number of time slots until the packet under transmission is successfully received by receiver $i = 1, 2$. The random variable T_i has distribution of geometric nature with parameter q_i. Since we have $T = \max(T_1, T_2)$, the expected value of T is computed as

$$E[T] = 1 + \sum_{t=1}^{\infty} \left(1 - \prod_{i=1}^{2} \sum_{\tau=1}^{t} (1 - q_i)^{\tau-1} q_i \right) \ . \tag{8.2}$$

The value of $\mu = \frac{1}{E[T]}$ is equivalent to Eq. (8.1). As a result, the stability condition is given by

$$\lambda < \frac{q_1 q_2 (q_1 + q_2 - q_1 q_2)}{(q_1 + q_2)(q_1 + q_2 - q_1 q_2) - q_1 q_2} \ . \tag{8.3}$$

The rate of packets that can be delivered to both receivers satisfies $\lambda < \min(q_1, q_2)$. In section 8.4, we will show that this upper bound can be asymptotically approached by random network coding. For $q_1 = q_2 = q = 0.5$, Figure 8.2 depicts the service rates, namely boundaries of the stable throughput rates under STP and PRP, and the upper bound on the stable throughput rate. The service rate μ of PRP is strictly smaller than $\min(q_1, q_2)$ unless $q_i = 0$ or 1, $i = 1, 2$. We

will show in section 8.5 how to achieve the stable throughput rate arbitrarily close to $\min(q_1, q_2)$ for any given pair of fixed packet reception probabilities q_1 and q_2.
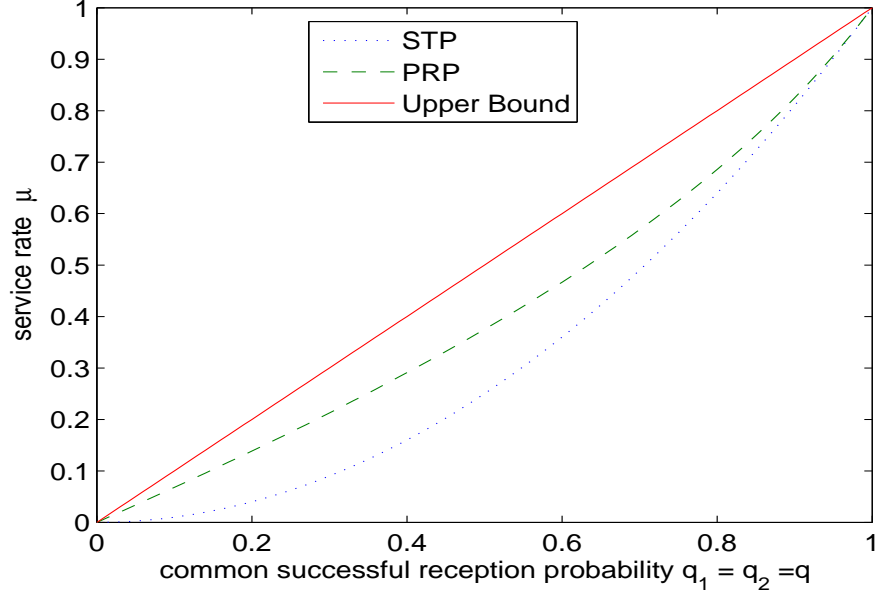

Figure 8.2: Service rates under STP and PRP and upper bound on the service rate.

8.4 Random Network Coding Policy

In PRP, source node transmits one packet at a time, until the packet is successfully received by all receivers. We can generalize this idea by accumulating packets into groups of fixed size K and transmitting a randomly generated linear combination of K packets at a time, until all these K packets are successfully decoded by all receivers.

To understand the throughput benefits of coding, consider the case in which packets A_1 and A_2 must be delivered to both receivers 1 and 2. We assume $\mathcal{F}_2$ as the field for (network) coding operations such that the bit-sum $A_1 + A_1$ of two packets

A_1 and A_2 is a modulo-2 vector addition of the corresponding vectors of each packet. If the transmission of packet A_1 is only received by receiver 1 and the transmission of packet A_2 is only received by receiver 2, source can transmit packet $A_1 + A_2$ (instead of separately transmitting packets A_1 and A_2 as in PRP). Thus, receiver 1 and 2 can decode A_2 and A_1 by computing $A_1 + (A_1 + A_2)$ and $A_2 + (A_1 + A_2)$, respectively, i.e. it is possible to deliver both packets in a single time slot, if the transmission is successful at both receivers. For $K > 1$, transmitter node randomly combines the first available K packets in queue by linear mapping with coefficients from a finite field $\mathcal{F}_r$. Consider the mth group of K packets. The jth packet transmitted from packet group m is given by $A_m(j) = \sum_{i=1}^{K} A_{(m-1)K+i}\, \alpha_{m,i}(j)$, where the coefficients $\alpha_{m,i}(j)$, $i = 1, ..., K$, are randomly and independently chosen from the finite field $\mathcal{F}_r$ and form a coding vector for any transmitted packet j in packet group m. Whenever transmitter node is informed over feedback channels that all receivers have decoded K packets, it removes them simultaneously from the queue.

8.4.1 Stable Throughput Properties

Let L denote the sufficient number of successful packet transmissions to decode K packets at each receiver. After successfully receiving the lth packet, each receiver forms the $l \times K$ matrix $M(l, K)$ with each row representing the coding vector of each packet received so far. The group of K packets is successfully decoded, if the matrix $M(l, K)$ has rank of at least K. We denote this event by $E_{l,K}$. The probability distribution of $E_{l,K}$ is computed as follows. The first column of $M(l, K)$

247

must contain at least one non-zero entry, which occurs with probability $1 - \frac{1}{r^l}$. The second column must contain at least one non-zero element and should be linearly independent of the first column, i.e. should not be one of the $r - 1$ possibilities of the first column multiplied by any non-zero symbol $\alpha = 1, ..., r - 1$. This occurs with probability $1 - \frac{1}{r^{l-1}}$. Proceeding this way for the first K columns, we can ensure the linear independence of each column $j = 1, ..., K$ of matrix $M(l, K)$ from the previous $j - 1$ columns such that the probability distribution of $E_{l,K}$ is given by

$$P(E_{l,K}) = \prod_{k=l+1-K}^{l} \left(1 - \frac{1}{r^k}\right). \tag{8.4}$$

It is sufficient to receive L packets for decoding K packets, if matrix $M(L - 1, K)$ has rank $K - 1$ and matrix $M(L, K)$ has rank K, since each transmission can increase the rank of matrix $M(l, K)$ at most one for any l. For coding field size r and block length K, let $P_{K,r}(l)$ denote the probability of $L = l$, namely the probability that at least l successful packet transmissions are sufficient to decode K packets, and let $E_{l,K}^k$ denote the event that $M(l, K)$ has rank k such that we have

$$P_{K,r}(l) = P(E_{l-1,K}^{K-1}) \cdot P(E_{l,K}^{K} | E_{l-1,K}^{K-1}) \tag{8.5}$$

for $l \geq K$. By counting arguments, we obtain

$$P(E_{l,K}^{K} | E_{l-1,K}^{K-1}) = 1 - \frac{1}{r}, \tag{8.6}$$

since a successfully received packet is innovative with probability $1 - \frac{1}{r}$. We can iteratively compute $P(E_{l-1,K}^{K-1})$ as $p_{l-1,K}(K)$ from combinatorial arguments, where

$p_{l-1,K}(k)$, $k = 1, ..., K$, evolves as

$$p_{l-1,K}(1) = \frac{1}{r^{l-K}}, \tag{8.7}$$

$$p_{l-1,K}(k) = \frac{1}{r^{l-1-K+k}} \prod_{m=0}^{k-2} \left(1 - \frac{1}{r^{l-1-K+k-m}}\right) + \left(1 - \frac{1}{r^{l-1-K+k}}\right) p_{l-1,K}(k-1)$$

for $k = 2, ..., K$. By combining Eqs. (8.5)-(8.7), we can obtain the distribution of L, namely $P_{K,r}(l)$, as

$$P_{K,r}(l) = \left(1 - \frac{1}{r}\right) \sum_{i=1}^{K} \left(\frac{1}{r^{l-1-K+i}} \prod_{m=0}^{i-2} \left(1 - \frac{1}{r^{l-1-K+i-m}}\right) \prod_{j=i+1}^{K} \left(1 - \frac{1}{r^{l-1-K+j}}\right)\right) \tag{8.8}$$

for $l \geq K$. The distribution $P_{K,r}(l)$ also follows from the definition of

$$P_{K,r}(l) = P(E_{l,K}^{K}) - P(E_{l-1,K}^{K}) \tag{8.9}$$

such that $P_{K,r}(l)$ is simply given by

$$P_{K,r}(l) = \frac{1}{r^{l-K}} \left(1 - \frac{1}{r^{K}}\right) \prod_{k=l+1-K}^{l-1} \left(1 - \frac{1}{r^{k}}\right). \tag{8.10}$$

We define $T_i(K, r)$ as the number time slots required by receiver $i = 1, ..., n$ to decode K packets and we define

$$T(K, r) = \max_{i=1,...,n} T_i(K, r) \tag{8.11}$$

as the total number of time slots required by all receivers to decode K packets. Provided that receivers need at least l successful packet transmissions to decode K packets, the conditional expected value of $T(K, r)$ can be written as

$$E[T(K, r)|L = l] = l + \sum_{t=l}^{\infty} \left(1 - \prod_{i=1}^{n} \sum_{\tau=l}^{t} \binom{\tau-1}{l-1}(1-q_i)^{\tau-l} q_i^{l}\right). \tag{8.12}$$

As a result, the expected service time to deliver K packets to all receivers is computed as

$$E[T(K,r)] = \sum_{l=K}^{\infty} P_{K,r}(l) \, E[T(K,r)|L=l] \tag{8.13}$$

by combining Eqs. (8.10) and (8.12). The service rate μ as function of K and r is given by $\frac{K}{E[T(K,r)]}$, since K packets are decoded by all receivers (and cleared from the transmitter queue) over $E[T(K,r)]$ time slots on the average. As a result, the stable service rate is given by

$$\mu = \frac{K}{E[T(K,r)]}. \tag{8.14}$$

For $n = 2$ and $q_i = q$, $i = 1, 2$, we depict in Figure 8.3 the service rate μ for different values of K and r, and we depict in Figure 8.4 the minimum code length K such that RCP can improve the stable throughput rate of PRP for different values of r. Network coding offers more throughput benefits for smaller successful packet reception probabilities.

The stable throughput of RCP increases with the coding field size r. PRP and RCP become equivalent for $K = 1$, as r goes to infinity, i.e. we have $\mu = \lim_{r \to \infty} \frac{1}{E[T(1,r)]}$ for PRP. The value of L approaches to K in probability, as r increases. As computed in [91] for rateless communication, the expected service time, as r goes to infinity, is given by

$$E[T(K)] = K + \sum_{t=K}^{\infty} \left(1 - \prod_{i=1}^{n} \sum_{\tau=K}^{t} \binom{\tau-1}{K-1} (1-q_i)^{\tau-K} q_i^K \right). \tag{8.15}$$

For $n = 2$, in Figure 8.5 shows μ, as r goes to infinity, and we have

$$\lim_{K \to \infty, \, r \to \infty} \frac{K}{E[T(K,r)]} = \min_{i=1,...,n} q_i \tag{8.16}$$
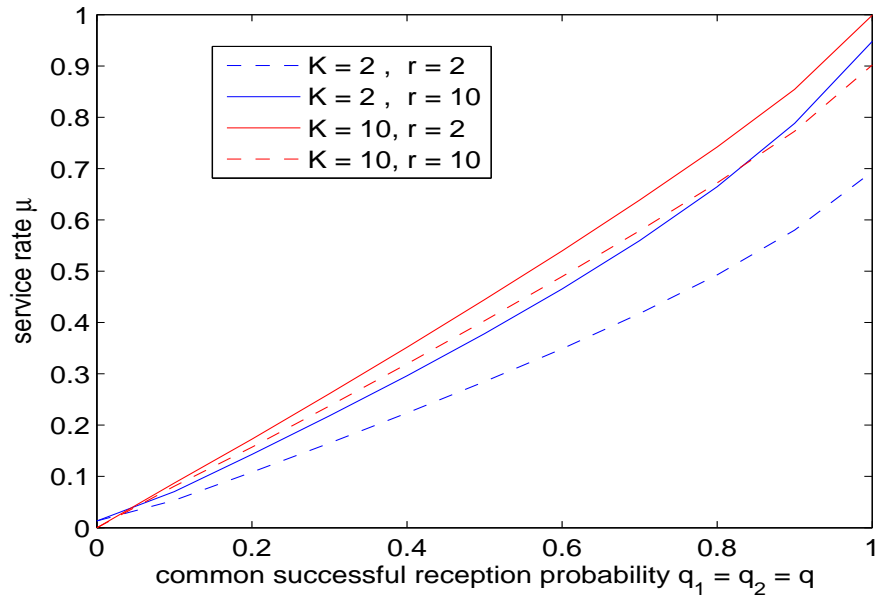
Figure 8.3: Service rate of RCP as function of common successful packet reception probability $q_i = q$, $i = 1, 2$, for different values of K and r.

such that RCP asymptotically approaches the stable throughput upper bound, as both r and K go to infinity. For $q_i = 0.5$, $i = 1, ..., n$, we evaluate the service rates in Figure 8.6 as function of n. Note that μ becomes less sensitive to any increase in n, as K and r increase.

8.4.2   Transmission and Processing Energy Costs

Each packet is transmitted with energy cost ϵ_t. The transmission energy cost per packet successfully decoded by all receivers is $\mathcal{E} = \frac{\epsilon_t}{\mu}$ for stable throughput rate μ. For $n = 2$ and $\epsilon_t = 1$, we evaluate in Figure 8.7 the transmission energy cost as function of the common successful reception probability q for different values of K, as r goes to infinity. For RCP, the processing cost at transmitter node is
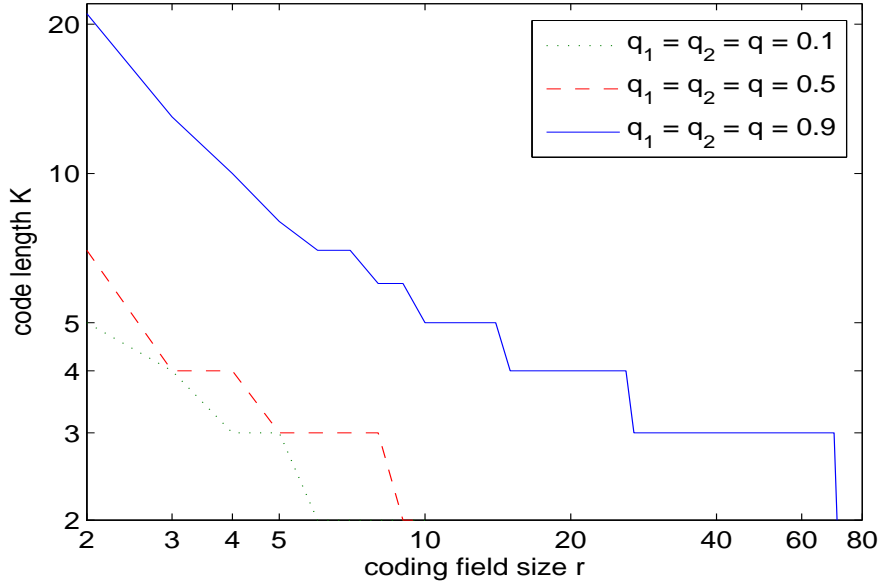
Figure 8.4: The minimum code length K as function of r such that RCP improves the stable throughput rate of PRP for common successful packet reception probability $q_i = q$, $i = 1, 2$.

$(K-1)E[L(K,r)]\epsilon_c$, where $E[L(K,r)]$ is the expected number of coding operations and ϵ_c is the coding cost for each symbol addition. Each receiver node performs Gaussian elimination to decode K packets such that the processing cost at each receiver is in the order of $K^3\epsilon_d$. As a result, the total energy cost for K packets is in the order of $E[T(K,r)]\epsilon_t + KE[L(K,r)]\epsilon_c + nK^3\epsilon_d$. Since both $E(T(K,r))$ and $E(L(K,r))$ scale linearly with K, the quantity $\mathcal{E}$ increases quadratically with K.

8.4.3 Extensions to Compound Random Access of Two Source Nodes

Assume that two sources randomly transmit packets addressed to all of n receivers, as shown in Figure 8.8. Let p_i denote the transmission probability of
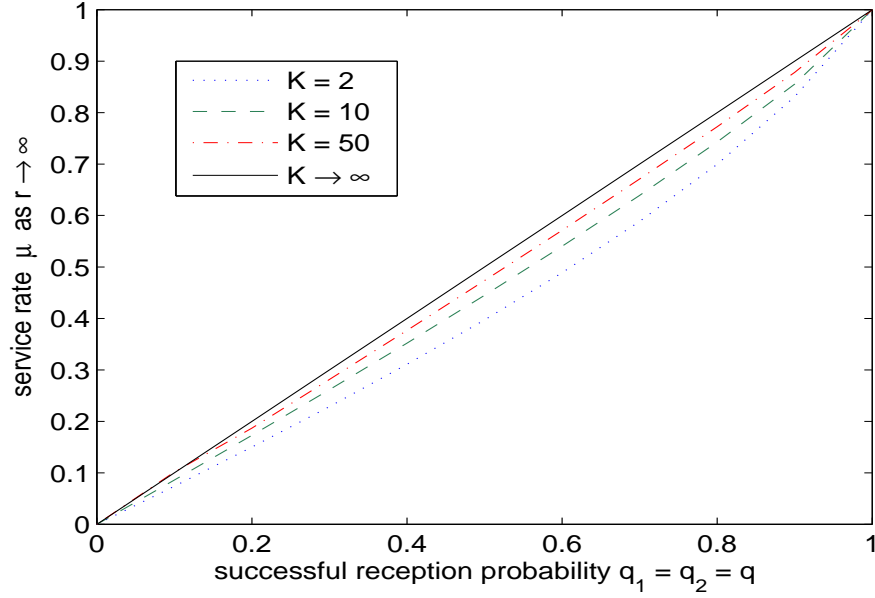
Figure 8.5: Service rate $\lim_{r \to \infty} \mu$ of RCP as function of common successful packet reception probability $q_i = q$, $i = 1, 2$, for different values of K.

source $i = 1, 2$. Let $\mu_{i,b}$ or $\mu_{i,e}$ denote the service rate of source i depending on whether the other node has a packet to transmit or not, i.e. $\mu_{1,e} = [\mu_{1,b}]_{p_2=0}$ and $\mu_{2,e} = [\mu_{2,b}]_{p_1=0}$. We assume multi-packet reception channels and denote $q_{i|T}^{(j)}$ as the probability of successful reception of packet of source i by receiver j, if set T of nodes transmit packets in the same time slot. We follow the dominating systems argument [12] to derive the stability region $\mathcal{S}$ as the set of arrival rates λ_1 and λ_2 at sources 1 and 2:

$$\lambda_1 < \frac{\lambda_2}{\mu_{2,b}}\mu_{1,b} + \left(1 - \frac{\lambda_2}{\mu_{2,b}}\right)\mu_{1,e}, \quad \lambda_2 \leq \mu_{2,b} \ , \tag{8.17}$$

$$\lambda_2 < \frac{\lambda_1}{\mu_{1,b}}\mu_{2,b} + \left(1 - \frac{\lambda_1}{\mu_{1,b}}\right)\mu_{2,e}, \quad \lambda_1 \leq \mu_{1,b} \ , \tag{8.18}$$

as given in [101]. We evaluate the stable throughput rates λ_1 and λ_2 on the boundary of $\mu_{1,b}$ and $\mu_{2,b}$. For RCP, each node attempts to deliver blocks of K randomly
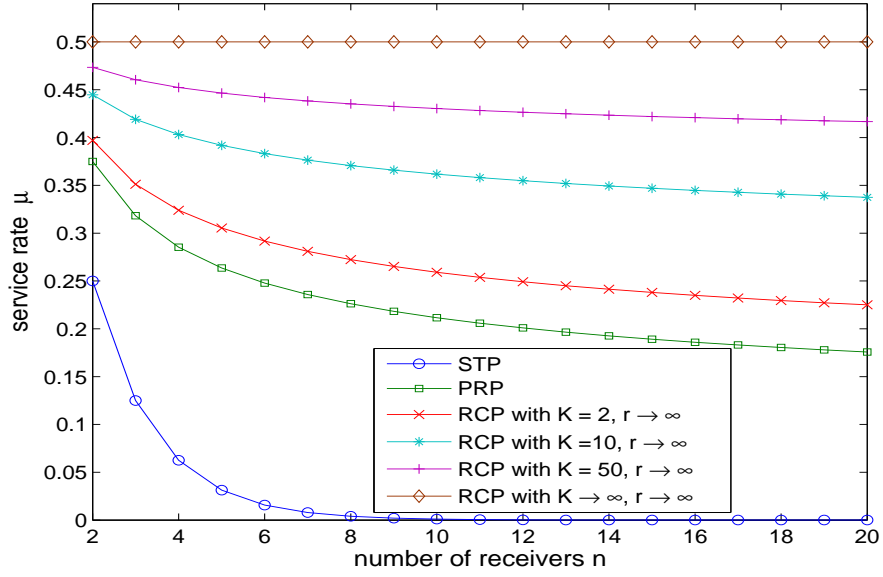
Figure 8.6: Service rates as function of the number of receivers n for $q_i = 0.5$, $i = 1, ..., n$.

combined packets. We denote by $T^i_j(K, r)$ the number of time slots required by receiver $j = 1, ..., n$ to decode K packets from source $i = 1, 2$. Let $T^i(K, r) = \max_{j=1,...,n} T^i_j(K, r)$ denote the total number of time slots required by all receivers to decode K packets from source i. If l successful packet transmissions are required by all receivers to decode K packets from source i, the conditional expected value of $T^i(K, r)$ is

$$E[T^i(K, r)|L = l] = l + \sum_{t=l}^{\infty} \left(1 - \prod_{j=1}^{n} \sum_{\tau=l}^{t} \binom{\tau - 1}{l - 1} (1 - \mathcal{P}_i^{(j)})^{\tau - l} (\mathcal{P}_i^{(j)})^l \right), \quad (8.19)$$

where $\mathcal{P}_1^{(j)} = p_1(p_2 q_{1|1,2}^{(j)} + (1 - p_2) q_{1|1}^{(j)})$ and $\mathcal{P}_2^{(j)} = p_2(p_1 q_{2|1,2}^{(j)} + (1 - p_1) q_{2|2}^{(j)})$ are obtained, if we condition the success probabilities depending on whether the other node transmits or not. This holds only if the packets of different transmitters are not cross-coded with each other. The expected service time to deliver K packets

254
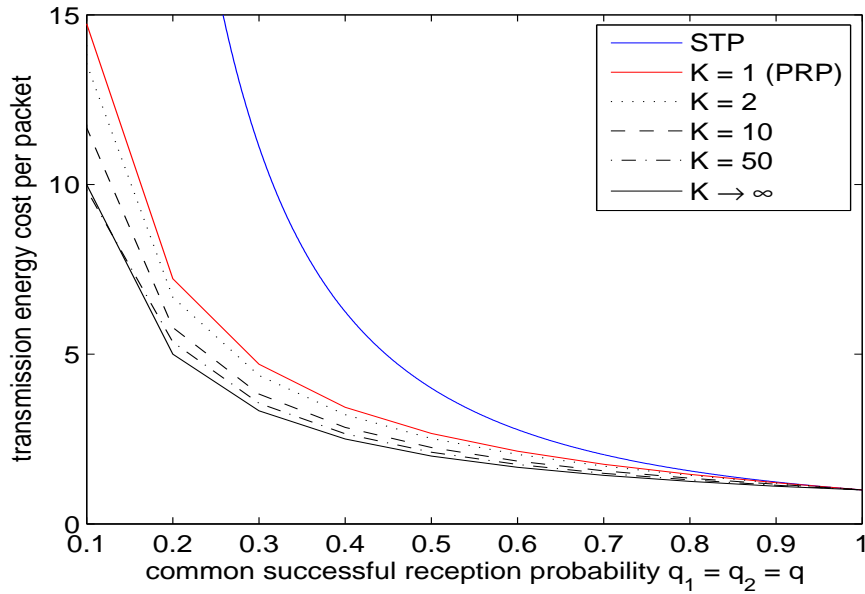
Figure 8.7: The transmission energy cost as function of common successful packet reception probability $q_i = q$, $i = 1, 2$, for different values of K.

from source $i = 1, 2$ to all receivers is given by

$$E[T^i(K,r)] = \sum_{l=K}^{\infty} P_{K,r}(l) \, E[T^i(K,r)|L^i = l], \qquad (8.20)$$

where L^i is the minimum number of randomly combined packets sufficient for all receivers to decode K packets of source node i. The events $L^i = l$, $i = 1, 2$, are independent for both sources (since packets of different sources are not coded with

Figure 8.8: System model with two transmitters and n receivers.

255

each other) and have the distribution $P_{K,r}(l)$ given by Eq. (8.10). Finally, the service rate $\mu_{i,b}$ can be computed as $\frac{K}{E[T^i(K,r)]}$ by combining Eqs. (8.10), (8.19) and (8.20). If we take 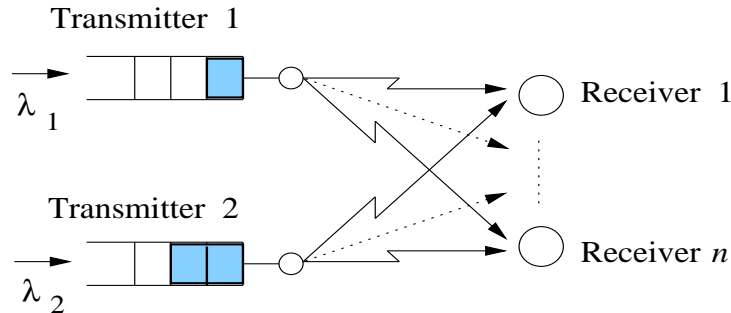the envelope over all $0 \leq p_i \leq 1$, $i = 1, 2$, we obtain the stability region for random access of two sources and n receivers. For different transmission policies STP, PRP and RCP, Figure 8.9 depicts the stability region for strong multi-packet reception channels with $q_{1|1}^{(1)} = q_{1|1}^{(2)} = q_{2|2}^{(1)} = q_{2|2}^{(2)} = 0.9$, $q_{1|12}^{(1)} = q_{1|12}^{(2)} = q_{2|12}^{(1)} = q_{2|12}^{(2)} = 0.6$.

The extension to arbitrary number of sources can be realized by following the conjecture [22] that the maximum throughput and stability regions coincide in random access for saturated and possibly emptying packet queues (under additional assumptions that multi-packet reception channels are standard and stationary queue distributions follow a sensitivity monotonicity property). Let S denote the set of source nodes and R_i denote the receiver set of source node $i \in S$. For PRP and RCP, the service rate μ_i of source node i is given by $\prod_{j \in R_i} \mathcal{P}_i^{(j)}$ and $\frac{K}{E[T^i(K,r)]}$, respectively, where $\mathcal{P}_i^{(j)} = p_i \sum_{T:i \in T \subseteq S} q_{i|T}^{(j)} \prod_{t \in T - \{i\}} p_t \prod_{s \notin T} (1 - p_s)$ for $i \in S$ and $j \in R_i$ under the assumption of saturated queues. The expected service time $E[T^i(K,r)]$ follows from Eqs. (8.10), (8.19) and (8.20). As K and r go to infinity, $\frac{K}{E[T^i(K,r)]}$ approaches $\min_{j \in R_i} \mathcal{P}_i^{(j)}$, whereas the service rate of PRP is given by $(\lim_{r \to \infty} E[T^i(1,r)])^{-1}$.

8.5 Optimal Coded Retransmission Policy for the Single Source Case

We consider coded retransmission policy (CRP) that retransmits linearly combined packets depending on the instantaneous queue content and operates with finite
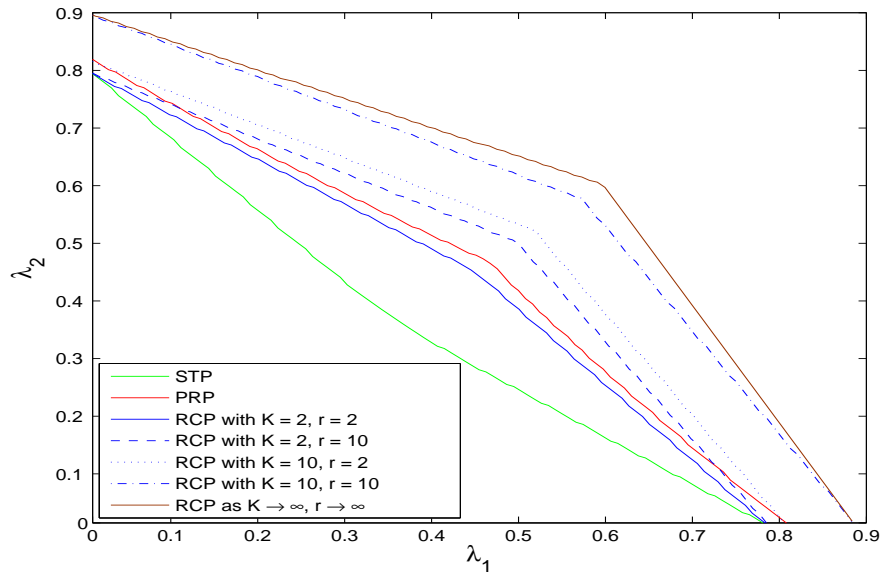
Figure 8.9: Stability region for different retransmission policies with and without coding.

delay and complexity.

The idea is based on network coding to combine the packet transmissions addressed to different receivers and has been used in Chapter 7 for scheduled access with possibly emptying queues and for random access with saturated queues in tandem networks with multiple source nodes. Random network coding has been proposed before in [91] for broadcasting blocks of packets in rateless communication. Random network coding for reliable communication [68, 107] is desirable because of simplicity but requires infinite block length and alphabet size to optimize the throughput rates. Instead, we consider random packet generation process with possibly emptying packet queue and consider dynamic but deterministic coding policies based on the instantaneous queue content.

Based on feedback, transmitter node puts each packet to a virtual queue depending on which receiver has successfully received that packet. We assume three virtual queues. All packets arrive at virtual queue $Q_{1,2}$ and are transmitted in a first-come-first-served fashion. If the packet has been successfully received by both receivers, the packet leaves the system. If the packet has not been received by any of two receiver nodes, the packet remains in queue $Q_{1,2}$. Packets that have been received by receiver 1 but not by receiver 2 enter virtual queue Q_1. Similarly, packets that have been received by receiver 2 but not by receiver 1 enter virtual queue Q_2. Then, packets leave queue Q_1 or Q_2 and consequently leave the system, if they have been received by receiver 2 or 1, respectively. We consider only stationary network operation, in which the queue distributions reach steady state. The queue Q at transmitter node is stable, if and only if the virtual queues $Q_{1,2}$, Q_1 and Q_2 are all stable. The virtual queue structure is illustrated in Figure 8.10.
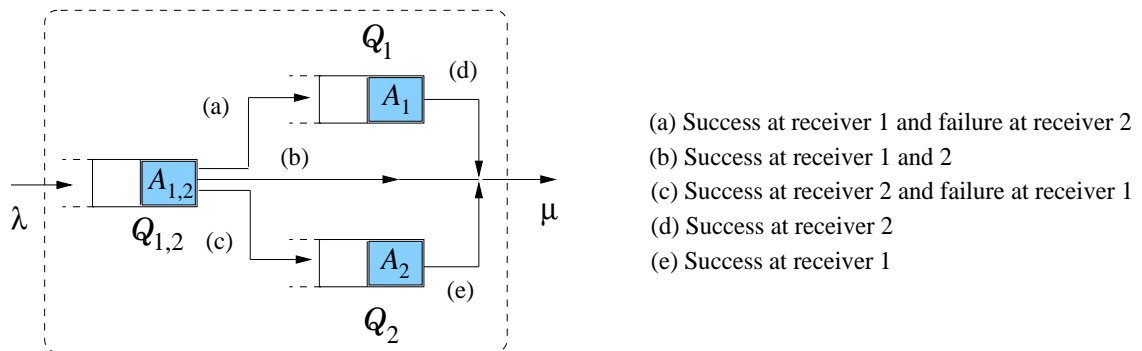

Figure 8.10: Virtual packet queues $Q_{1,2}$, Q_1 and Q_2 at transmitter node.

A packet is transmitted from queue $Q_{1,2}$, if it contains at least one packet. Otherwise, if queues Q_1 and Q_2 contain at least one packet, they are combined by linear coding before transmission. If A_1 and A_2 are the first available packets in

258

queues Q_1 and Q_2 and queue $Q_{1,2}$ is empty, the packet $A_1 + A_2$ is transmitted. If it is successfully received by receiver 1 (with probability q_1) or by receiver 2 (with probability q_2), it leaves queue Q_2 or Q_1. In other words, it is possible to deliver one packet to each receiver in a single transmission. However, this cannot be realized by PRP that needs two successful transmissions to deliver packets A_1 and A_2 to receiver 1 and 2, respectively. If only one of queues Q_1 and Q_2 contains packets, the packet is transmitted without coding (provided that queue $Q_{1,2}$ is empty). The arrival rate of queue $Q_{1,2}$ is λ. Since the packet leaves $Q_{1,2}$, unless the packet is received by both receivers, the service rate is $1 - (1 - q_1)(1 - q_2)$. According to Loynes Theorem [34], the stability condition for queue $Q_{1,2}$ is given by

$$\lambda < 1 - (1 - q_1)(1 - q_2). \tag{8.21}$$

If queue $Q_{1,2}$ is stable, the departure rate is λ. As a result, packets arrive at queues Q_1 and Q_2 with rates $\frac{q_1(1-q_2)}{1-(1-q_1)(1-q_2)}$ and $\frac{q_2(1-q_1)}{1-(1-q_1)(1-q_2)}$, respectively. Packets are transmitted from queues Q_1 and Q_2, if queue $Q_{1,2}$ is idle. According to Little's result [9], this occurs with probability $1 - \frac{\lambda}{1-(1-q_1)(1-q_2)}$, since queue $Q_{1,2}$ has the arrival rate λ and service rate $1 - (1 - q_1)(1 - q_2)$.

The service rates of queues Q_1 and Q_2 are $(1 - \frac{\lambda}{1-(1-q_1)(1-q_2)})q_2$ and $(1 - \frac{\lambda}{1-(1-q_1)(1-q_2)})q_1$, respectively, since queues Q_1 and Q_2 receive service simultaneously (due to the proposed coding operation) provided that queue $Q_{1,2}$ is empty and packets that are transmitted from queues Q_1 and Q_2 are successfully received by receiver 2 and 1 with probability q_2 and q_1, respectively. The resulting stability

259

conditions for queues Q_1 and Q_2 are

$$\frac{\lambda q_1(1-q_2)}{1-(1-q_1)(1-q_2)} < \left(1 - \frac{\lambda}{1-(1-q_1)(1-q_2)}\right)q_2, \qquad (8.22)$$

$$\frac{\lambda q_2(1-q_1)}{1-(1-q_1)(1-q_2)} < \left(1 - \frac{\lambda}{1-(1-q_1)(1-q_2)}\right)q_1. \qquad (8.23)$$

If we combine conditions (8.21)-(8.23), queues $Q_{1,2}$, Q_1 and Q_2 (and therefore queue Q) are stable, if

$$\lambda < \min(q_1, q_2). \qquad (8.24)$$

As a result, the maximum throughput rate bound (for the case of saturated queues) and the stability condition (for the case of possibly emptying packet queues) coincide through the application of CRP. On the other hand, PRP is strictly suboptimal for broadcast systems, as the comparison of conditions (8.3) and (8.24) reveals. For $q_1 = q_2 = q$, Figure 8.11 depicts the service rates under coded and plain retransmission policies.

8.5.1   Feedback, Packet Overhead and Complexity Issues

All policies with or without coding need feedback information to be exchanged between transmitter and receiver nodes. STP needs only $\lceil \log_r 2 \rceil$ control symbols to inform receivers whether the transmitted packet is new or it is a retransmission. On the other hand, PRP needs $\lceil \log_r(2^n-1) \rceil$ control symbols to inform receivers which of the 2^n-1 non-empty subsets of n receivers the transmitted packet is intended for. In RCP, transmitter needs to encode $\lceil \log_r K \rceil$ control symbols into the overhead of each packet to inform receivers how K packets are encoded into the transmitted packet.

260
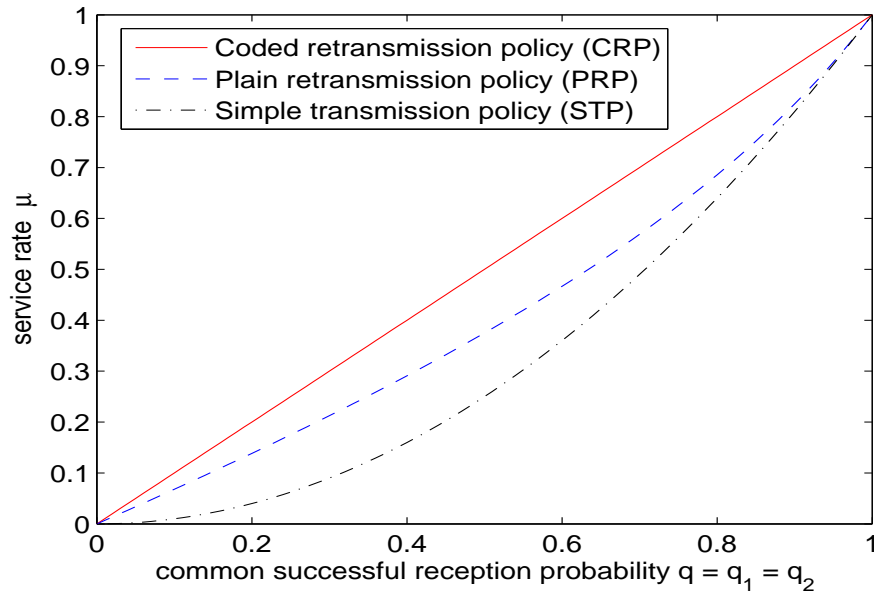
Figure 8.11: Service rates under coded and plain retransmission policies.

In CRP, transmitter nodes need to inform receivers the intended destination(s) of each packet and whether the packet is coded or not. However, packets can contain arbitrarily large number of information symbols so that we can assume that the overhead information is negligible.

Both PRP and CRP require the transmitter node to keep virtual queues depending on which receiver has successfully received the transmitted packets. This is equivalent to the memory requirement of PRP to track which receiver has successfully received each packet. As a result, all policies with and without coding are similar in terms of overhead and computational complexity (and the difference diminishes as the number of information symbols encoded into packets increases). The main disadvantage of RCP and CRP is that packets can be received in the mixed order of transmission. However, this can be easily handled through the use

of packet ID's. The resulting service rates are scaled down by the feedback information (provided that the information and control packet transmissions share the same channel).

8.6 Optimal Coded Retransmission Policy for Two Source Nodes

Assume that two transmitter nodes randomly transmit packets addressed to both receiver nodes, as shown in Figure 8.8. Let p_i denote the transmission probability of transmitter node $i = 1, 2$. Let $\mu_{i,b}$ or $\mu_{i,e}$ denote the service rate of node i depending on whether the other transmitter node has a packet or not, i.e. $\mu_{1,e} = [\mu_{1,b}]_{p2=0}$ and $\mu_{2,e} = [\mu_{2,b}]_{p1=0}$. We assume multi-packet reception (MPR) channels and denote $q_{i|K}^{(j)}$ as the probability of successful reception of packet of transmitter node i by receiver node j, if the set K of nodes transmit packets.

Under the assumption that packets correspond to either an information symbol or an idle symbol, the capacity region $\mathcal{C}$ has been derived in [100] as the closure of rates (R_1, R_2) such that

$$R_1 \leq p_1 \min(p_2 q_{1|1,2}^{(1)} + (1 - p_2)q_{1|1}^{(1)}, p_2 q_{1|1,2}^{(2)} + (1 - p_2)q_{1|1}^{(2)}), \qquad (8.25)$$

$$R_2 \leq p_2 \min(p_1 q_{2|1,2}^{(1)} + (1 - p_1)q_{2|2}^{(1)}, p_1 q_{2|1,2}^{(2)} + (1 - p_1)q_{2|2}^{(2)}). \qquad (8.26)$$

These capacity results coincide with the maximum achievable throughput bounds for the case in which we consider arbitrary information packets to be transmitted and saturated packet queues, i.e. the maximum achievable throughput rates achieve the bounds of conditions (8.25)-(8.26). First, consider source node 1. The maximum throughput rate λ_1^i over the cut that separates source node 1 from re-

ceiver $i = 1, 2$ is given by $p_1(p_2 q_{1|1,2}^{(i)} + (1 - p_2)q_{1|1}^{(i)})$, since receiver i receives packet of source node 1, if source node 1 transmits packets (with probability p_1) and the transmission is successfully received by receiver i with probability $q_{1|1,2}^{(i)}$, if source node 2 also transmits (with probability p_2) or with probability $q_{1|1}^{(i)}$, if source node 2 does not transmit (with probability $1 - p_2$). The maximum throughput rate for source node 1 is $\min_{i=1,2} \lambda_1^i$. The same argument also holds for source node 2.

For fixed transmission and reception probabilities, we follow the dominating systems argument to derive the stability region $\mathcal{S}$ as the set of arrival rates λ_1 and λ_2 satisfying conditions (8.17)-(8.18).

The stability region is illustrated in Figure 8.12. We consider the stable throughput rates λ_1 and λ_2 on the boundary of $\mu_{1,b}$ and $\mu_{2,b}$ and show that the optimal values of $\mu_{1,b}$ and $\mu_{2,b}$ coincide with the upper bounds on the achievable rates R_1 and R_2, respectively.
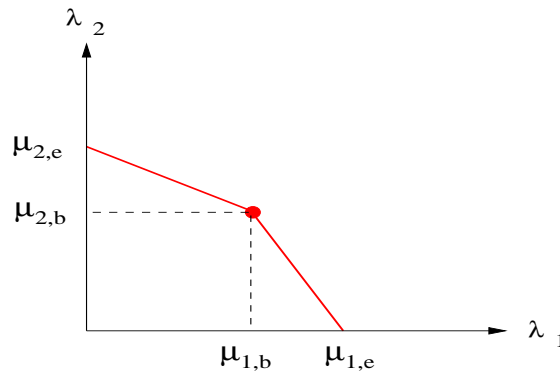

Figure 8.12: Broadcast stability region in random access.

First, we assume that transmitter node 2 has packets and derive the stability conditions for virtual queues Q_1, Q_2 and $Q_{1,2}$ at transmitter node 1. For queue $Q_{1,2}$, the arrival rate is λ_1 and the service rate is $1 - p_1(p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) -$

$(1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)}))$, since a packet of transmitter node 1 leaves queue $Q_{1,2}$, unless it is not successfully received by both receivers, which occurs with probability $(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)})$, if transmitter node 2 transmits a packet (with probability p_2), or occurs with probability $(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)})$, if transmitter node 2 does not transmit a packet (with probability $1 - p_2$). If queue $Q_{1,2}$ at transmitter node 1 is stable, queues Q_1 and Q_2 at transmitter node 1 have the arrival rates

$$\frac{\lambda_1 p_1 (p_2 q_{1|1,2}^{(1)}(1 - q_{1|1,2}^{(2)}) + (1 - p_2) q_{1|1}^{(1)}(1 - q_{1|1}^{(2)}))}{p_1(1 - p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)}))}, \quad (8.27)$$

$$\frac{\lambda_1 p_1 (p_2 q_{1|1,2}^{(2)}(1 - q_{1|1,2}^{(1)}) + (1 - p_2) q_{1|1}^{(2)}(1 - q_{1|1}^{(1)}))}{p_1(1 - p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)}))}, \quad (8.28)$$

respectively. Queues Q_1 and Q_2 receive service with rates $p_1(p_2 q_{1|1,2}^{(2)} + (1 - p_2) q_{1|1}^{(2)})$ and $p_1(p_2 q_{1|1,2}^{(1)} + (1 - p_2) q_{1|1}^{(1)})$, respectively, if queue $Q_{1,2}$ at transmitter node 1 is idle (with probability $1 - \frac{\lambda_1}{1 - p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)})}$ according to Little's result [9]). The resulting stability conditions for transmitter node 1 are given by

$$\lambda_1 < 1 - p_1(p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)})) \quad (8.29)$$

for queue $Q_{1,2}$ at transmitter node 1,

$$\frac{\lambda_1 p_1 (p_2 q_{1|1,2}^{(1)}(1 - q_{1|1,2}^{(2)}) + (1 - p_2) q_{1|1}^{(1)}(1 - q_{1|1}^{(2)}))}{p_1(1 - p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)}))}$$
$$< p_1(p_2 q_{1|1,2}^{(2)} + (1 - p_2) q_{1|1}^{(2)}) \quad (8.30)$$
$$\cdot \left(1 - \frac{\lambda_1}{p_1(1 - p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)}))}\right)$$

for queue Q_1 at transmitter node 1, and

$$\frac{\lambda_1 p_1 (p_2 q_{1|1,2}^{(2)}(1 - q_{1|1,2}^{(1)}) + (1 - p_2) q_{1|1}^{(2)}(1 - q_{1|1}^{(1)}))}{p_1(1 - p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)}))}$$

$$< p_1 (p_2 q_{1|1,2}^{(1)} + (1 - p_2) q_{1|1}^{(1)}) \qquad (8.31)$$

$$\cdot \left(1 - \frac{\lambda_1}{p_1(1 - p_2(1 - q_{1|1,2}^{(1)})(1 - q_{1|1,2}^{(2)}) - (1 - p_2)(1 - q_{1|1}^{(1)})(1 - q_{1|1}^{(2)}))}\right)$$

for queue Q_2 at transmitter node 1, respectively. We obtain the stability condition $\lambda_1 < \mu_{1,b}$, where the service rate is $\mu_{1,b} = p_1 \min(p_2 q_{1|1,2}^{(1)} + (1 - p_2) q_{1|1}^{(1)}, p_2 q_{1|1,2}^{(2)} + (1 - p_2) q_{1|1}^{(2)})$. From the symmetry, we obtain the service rate $\mu_{2,b} = p_2 \min(p_1 q_{2|1,2}^{(1)} + (1 - p_1) q_{2|2}^{(1)}, p_1 q_{2|1,2}^{(2)} + (1 - p_1) q_{2|2}^{(2)})$. Hence, the stability region $\mathcal{S}$ is the region of arrival rates λ_1 and λ_2 such that

$$\lambda_1 < p_1 \min(p_2 q_{1|1,2}^{(1)} + (1 - p_2) q_{1|1}^{(1)}, p_2 q_{1|1,2}^{(2)} + (1 - p_2) q_{1|1}^{(2)}), \qquad (8.32)$$

$$\lambda_2 < p_2 \min(p_1 q_{2|1,2}^{(1)} + (1 - p_1) q_{2|2}^{(1)}, p_1 q_{2|1,2}^{(2)} + (1 - p_1) q_{2|2}^{(2)}). \qquad (8.33)$$

As a result, the comparison of conditions (8.25)-(8.26) and (8.32)-(8.33) reveals that the queueing stability and maximum throughput (and capacity) regions coincide for compound random access systems (except for boundary).

Next, we present numerical results for $q_{1|1}^{(1)} = q_{1|1}^{(2)} = q_{2|2}^{(1)} = q_{2|2}^{(2)} = 0.9$. Figure 8.13 depicts the stability region for strong MPR channels with $q_{1|12}^{(1)} = q_{1|12}^{(2)} = q_{2|12}^{(1)} = q_{2|12}^{(2)} = 0.6$ and weak MPR channels with $q_{1|12}^{(1)} = q_{1|12}^{(2)} = q_{2|12}^{(1)} = q_{2|12}^{(2)} = 0.3$.

8.7 Extensions to Unicast Communication

We consider the single source case with packets addressed to only one of the receiver nodes. Packets addressed to receiver 1 and 2 only arrive with rates $\lambda^{(1)}$ and
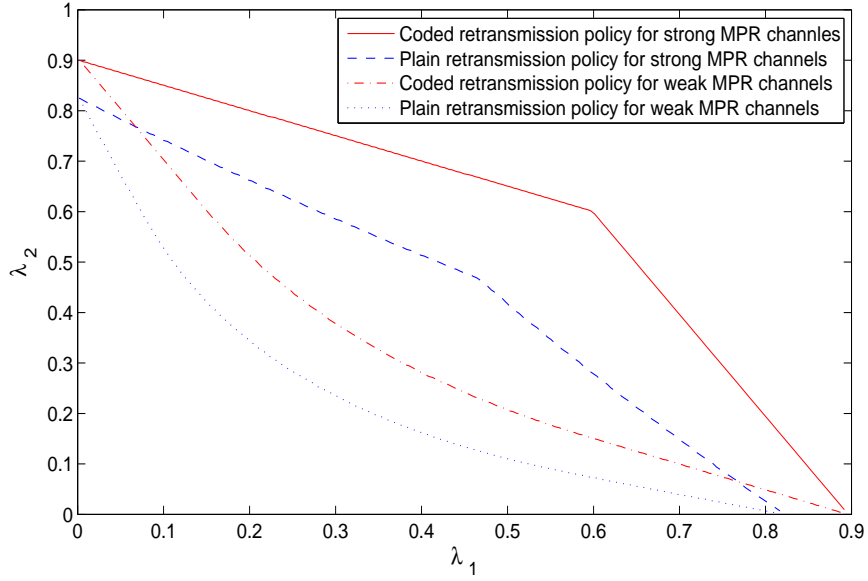
Figure 8.13: Stability region for strong and weak MPR channels.

$\lambda^{(2)}$, respectively. We assume the same structure of virtual queues as in section 8.5. All packets arrive at queue $Q_{1,2}$. Packets that are addressed to receiver node 2 but have been received only by receiver node 1 enter queue Q_1. Similarly, packets that are addressed to receiver node 1 but have been received only by receiver node 2 enter queue Q_2. Transmitter node first attempts to transmit a packet from queue $Q_{1,2}$. If queue $Q_{1,2}$ is empty, transmitter node attempts to transmit coded packets one from each of the queues Q_1 and Q_2. The resulting stability conditions for queues $Q_{1,2}$, Q_1 and Q_2 are given by

$$\lambda^{(1)} + \lambda^{(2)} < 1 - (1 - q_1)(1 - q_2), \tag{8.34}$$

$$\frac{\lambda^{(1)} q_2 (1 - q_1)}{1 - (1 - q_1)(1 - q_2)} < \left(1 - \frac{\lambda^{(1)} + \lambda^{(2)}}{1 - (1 - q_1)(1 - q_2)}\right) q_1, \tag{8.35}$$

$$\frac{\lambda^{(2)} q_1 (1 - q_2)}{1 - (1 - q_1)(1 - q_2)} < \left(1 - \frac{\lambda^{(1)} + \lambda^{(2)}}{1 - (1 - q_1)(1 - q_2)}\right) q_2. \tag{8.36}$$

266

As a result, the achievable stable throughput rates $\lambda^{(1)}$ and $\lambda^{(2)}$ satisfy

$$\frac{\lambda^{(1)}}{q_1} + \frac{\lambda^{(2)}}{1 - (1 - q_1)(1 - q_2)} < 1, \tag{8.37}$$

$$\frac{\lambda^{(1)}}{1 - (1 - q_1)(1 - q_2)} + \frac{\lambda^{(2)}}{q_2} < 1. \tag{8.38}$$

Next, we explore the capacity region for unicast packet traffic. Let $R^{(i)}$ be the rates of packets addressed to receiver $i = 1, 2$. Time sharing has been shown in [108] to be optimal for broadcast erasure channel with independent messages to be transmitted to receiver nodes. The capacity region for sending independent information to two receivers is described by the closure of rates $R^{(1)}$ and $R^{(2)}$ such that

$$\frac{R^{(1)}}{q_1} + \frac{R^{(2)}}{q_2} \leq 1. \tag{8.39}$$

However, this region is contained in the region of the achievable stable rates given by conditions (8.37)-(8.38). The reason is that the capacity region is derived without use of feedback from the receivers to the transmitter node. If we allow feedback, the capacity region should be expanded to include the stability region. An outer bound on the capacity region is given by

$$R^{(i)} \leq q_i, \quad i = 1, 2, \tag{8.40}$$

$$R^{(1)} + R^{(2)} \leq 1 - (1 - q_1)(1 - q_2). \tag{8.41}$$

Condition (8.40) follows from the upper bound on the mutual information over the cut that separates the transmitter node from the receiver $i = 1, 2$ only. On the other hand, the probability of successful transmission to at least one receiver is given

by $1 - (1 - q_1)(1 - q_2)$ such that the sum of rates $R^{(1)}$ and $R^{(2)}$ for both receivers cannot exceed $1 - (1 - q_1)(1 - q_2)$, as stated in condition (8.41). This outer bound holds independent of whether we allow feedback from receivers to transmitter or not. For fixed reception probabilities q_1 and q_2, Figure 8.14 depicts the stability region and inner and outer bounds on the capacity region. Future work should study the capacity region for multiple unicast communication with feedback.
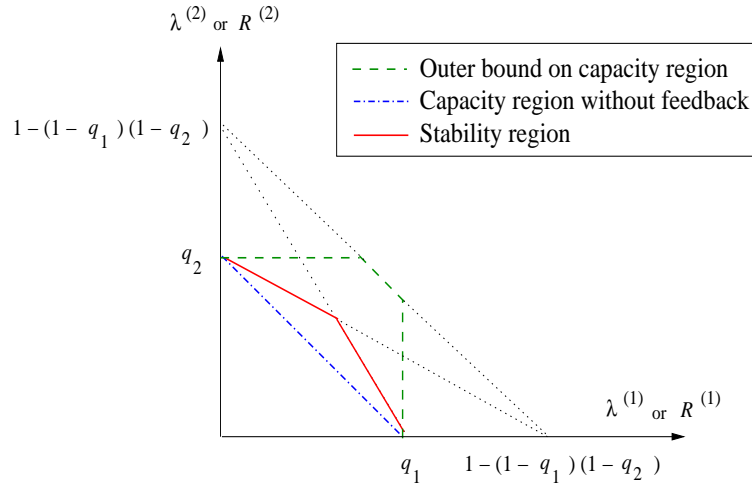
Figure 8.14: Outer bound on capacity region, capacity region without feedback and stability region for multiple unicast communication.

We can also combine broadcast and unicast communication. For that purpose, we assume that the transmitter node generates packets that are addressed to both receivers with rate $\lambda^{(1,2)}$ and generates packets that are addressed to only receiver 1 or 2 with rate $\lambda^{(1)}$ or $\lambda^{(2)}$. The resulting stability conditions for queues $Q_{1,2}$, Q_1 and Q_2 are given by

$$\lambda^{(1)} + \lambda^{(2)} + \lambda^{(1,2)} < 1 - (1 - q_1)(1 - q_2), \tag{8.42}$$

$$\frac{(\lambda^{(1)} + \lambda^{(1,2)})q_2(1 - q_1)}{1 - (1 - q_1)(1 - q_2)} < \left(1 - \frac{\lambda^{(1)} + \lambda^{(2)} + \lambda^{(1,2)}}{1 - (1 - q_1)(1 - q_2)}\right) q_1, \tag{8.43}$$

268

$$\frac{(\lambda^{(2)} + \lambda^{(1,2)})q_1(1 - q_2)}{1 - (1 - q_1)(1 - q_2)} < \left(1 - \frac{\lambda^{(1)} + \lambda^{(2)} + \lambda^{(1,2)}}{1 - (1 - q_1)(1 - q_2)}\right) q_2. \qquad (8.44)$$
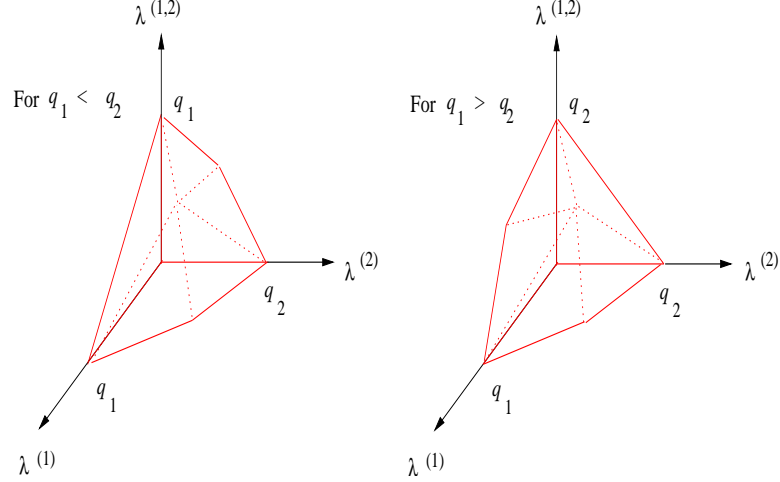
Figure 8.15: Stability region for combined traffic of multiple unicast and broadcast communication.

For fixed reception probabilities q_1 and q_2, the resulting region of stable throughput rates is depicted in Figure 8.15 separately for the cases of $q_1 < q_2$ and $q_2 < q_1$.

8.8   Summary and Conclusions

We addressed the effects of network coding on single-hop multicast communication over erasure channels with probabilistic reception. We derived the maximum stable throughput rates for one or two transmitter nodes broadcasting packets to two receivers. We proposed a coded retransmission policy to optimize the stable operation in random access and proved the equivalence of the stability and maximum throughput regions. We also discussed the implications on the capacity region.

Finally, we extended the results to multi-packet reception channels and multiple uni-cast communication demands. Future work should extend the results to arbitrary number of transmitter and receiver nodes with correlated channels and correlated packet generation processes at source nodes. In this context, the results in this chapter can be used in the random access model for the multi-hop network coding analysis in Chapter 7. We can evaluate the cut capacities and find the bounds on network capacity (under the model of probabilistic packet reception). Once the model is extended to arbitrary number of transmitters, the network capacity results would also follow for models with general interference effects in addition to the probabilistic errors.

Chapter 9

Conclusions

This dissertation addressed the cross-layer design at the medium access control and network layers in ad hoc wireless networks with and without user cooperation. We synthesized different approaches from the diverse fields of information theory, coding, queueing theory, game theory and optimization to answer the fundamental design questions of distributed control, stability, non-cooperative operation, throughput and energy efficiency. The results improve our understanding of the fundamental communication limits and clarify complex trade-offs in wireless network design.

- **Medium Access Control (MAC) in Ad Hoc Wireless Networks**

We introduced a two-step time division mechanism for throughput and energy-efficient resource allocation in stable operation. As the first step, we separately activated nodes as transmitters and receivers. For each activated receiver group, Group TDMA operation in the second step creates time orthogonality between transmissions to different receiver nodes. We formulated a linear optimization problem of maximizing the stable throughput rates and minimizing the transmission energy consumption in energy-limited systems. We also presented distributed implementation based on graph coloring. This approach prevents the unreliable feedback problem and offers a stable MAC operation. Group TDMA can rely on any arbi-

trary single-receiver MAC protocol to coordinate transmissions addressed to each receiver node.

- **Non-cooperative MAC and Joint Design with Routing**

As the network size grows, it is more difficult to coordinate nodes in full cooperation. Alternatively, we can allow a non-cooperative network operation with the individual performance objectives of throughput, energy and delay efficiency. First, we formulated stochastic games for single-receiver random access systems. We looked at the throughput and stability properties in non-cooperative random access of selfish transmitters with the individual packet queues. In addition, we updated the game model by incorporating malicious transmitters with the additional objectives of jamming the transmissions of the other selfish transmitters. Alternatively, we also considered a repeated game model of choosing transmissions probabilities for random access of backlogged packets. We further allowed nodes to distribute probabilities among different transmission powers and rates with extended strategy space. In addition, we evaluated the interactions between selfish and malicious transmitters for power-controlled MAC with SINR-based channel model. We also proposed adaptive best-response update mechanisms for distributed operation and compared the results with the cooperative strategies.

For mult-hop communication, we extended the game model to the joint design of MAC and routing over a simple relay network. In particular, we let nodes choose between transmitting packets directly to the destination and using the relay node to forward packets. In this context, we evaluated the cooperative and non-cooperative strategies, and introduced a reward-based cooperation stimulation mechanism to

272

improve the selfish network operation. The results underlined the importance of the cross-layer design for the non-cooperative network operation.

- **Cross-Layer Design of MAC and Network Coding**

Plain routing limits nodes to act as forwarding switches. Network coding is a new networking paradigm that allows nodes to code over the received packets along the path from sources to multiple destinations. We presented a cross-layer optimization framework for wireless network coding and MAC. We specified the improvement of throughput and energy efficiency at the expense of additional processing energy costs, packet overhead and delay. In addition, we showed how to design wireless network codes and discussed their properties in conjunction with the underlying MAC protocols. We also outlined distributed implementation methods based on subtree decomposition for joint MAC and network coding. The approach uses a single throughput criterion and it is based on the assumption of saturated queues without risk of packet underflow or concern about delay build-up. We also considered the problem of optimizing the region of throughput rates of multiple source nodes. We used the example of linear network topology with multiple source nodes to derive the maximum achievable throughput region for the case of saturated queues.

- **Effects of Wireless Network Coding on Queueing Stability**

We followed a queueing-theoretic approach to specify the maximum stable throughput region for the case of packet underflow. We also evaluated the transmission and energy costs as functions of the achievable or stable throughput rates for the separate cases of scheduled access and random access at the MAC layer.

Finally, we specified the effects of non-cooperative node behavior on the joint operation of wireless network coding and MAC. In this context, we introduced reward-based cooperation stimulation mechanisms to improve the non-cooperative network performance. We also considered the stability problem in a single-hop multicast communication system with independent probabilistic channels. We used queueing-theoretic arguments to show that a simple network coding scheme can improve the stable throughput rate to the Max-flow Min-cut bound. We further extended the model to include two source nodes randomly transmitting packets to multiple receiver nodes and specified the stability region of random-access systems with and without coding. We introduced random and deterministic network coding schemes to improve the stability properties of plain retransmission schemes.

- **Future Research Directions**

It is a fundamental problem to characterize the communication limits and trade-offs in wireless networks. There are numerous areas of investigation that naturally spring from the results of this dissertation.

The non-cooperative framework for autonomous network operation allows selfish and malicious users with objectives and incentives reflected in the utility functions. Multiple performance concerns such as energy expenditures require a general non-zero sum game formulation to represent the interactions between selfish and malicious users. A more general theory for hybrid (cooperative and non-cooperative) game formulations is needed to reflect the random nature in wireless networks. The current game formulations are based on complete information. Partial information and uncertainty can be incorporated in the context of Bayesian games. This

line of work has strong connections to the formulation of reputation or trust among users. In autonomous operation, users may form their beliefs on each other's actions, strategies and utilities. The control information exchange and the related packet overhead would have significant effects on how to establish reputation among each other, and open up new research questions. The formulation of malicious users combines two diverse research areas of network utility optimization and network security. The level of security needs to be further calibrated with respect to the achievable throughput rates (subject to the energy and delay constraints) provided that we formulate the security concerns as penalty.

Another fundamental extension arises from the network coding results. Wireless signals arrive at nodes as a weighted sum embedded in background noise and the corresponding incoming packets have a random nature. The network coding problem needs to be defined as a joint coding and detection problem. Two approaches are possible: (1) For any given network coding solutions, we can let nodes detect the linear combinations of packets (rather than detecting them individually and then combining them to code or decode). In this case, coding and decoding operations would be formulated as the noisy detection problem. (2) We can let nodes filter the incoming signals to produce output signals without detecting or decoding packets at the intermediate nodes. In this case, only destination nodes perform the detection operation. The performance objective of this modified network coding problem can be defined as minimizing the probability of error rather than maximizing the achievable capacity. A general anycast communication paradigm (with arbitrary communication demands with arbitrary source-destination pairs) is needed to fully

characterize the theoretical foundations of network coding.

Our general network model is based on omnidirectional transmissions. As an extension, it is important to evaluate the effects of directional antennas on the maximum stable throughput rates and to derive the scaling throughput laws depending on the directional antenna properties, as the network size grows [109]. The current network capacity studies [10] do not fully take into account the additional constraints such as energy consumption and packet delay. The ultimate goal is to study the relationship between information-theoretic capacity and stable throughput results achievable for limited energy resources and finite delay. This requires optimization of a region of throughput rates subject to energy and stability constraints. In this context, the necessary packet overhead and complexity can be considered as costs in the resulting optimization problem.

Bibliography

[1] S. Shakkottai, T. S. Rappaport and P. Karlsson, "Cross-layer Design for Wireless Networks," *IEEE Communications Magazine*, vol 41, no. 10, pp. 1-14, June 2003.

[2] H. Zimmermann, "OSI Reference Model: The ISO Model of Architecture for Open Systems Interconnection," *IEEE Transactions on Communications*, vol. 28, no. 4, pp. 425-432, Apr. 1980.

[3] A.B. MacKenzie and S.B. Wicker, "Game Theory and the Design of Self-Configuring, Adaptive Wireless Networks," *IEEE Communications Magazine*, pp.126-131, Nov. 2001.

[4] E. Arikan, "Some Complexity Results about Packet Radio Networks," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 681-685, July 1984.

[5] A. Ephremides and T. Truong, "Scheduling Broadcasts in Multihop Radio Networks, *IEEE Transactions on Communications*, vol. 38, no. 4, pp. 456-460, Apr. 1990.

[6] Y. E. Sagduyu and A. Ephremides, "Energy-Efficient MAC in Ad Hoc Networks Inspired by Conflict Resolution Concepts," *Elsevier ADHOC Journal*, vol. 1, no. 1, pp. 65-86, July 2003.

[7] G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "Multiple-Access for Multiple Destinations in Ad-hoc Networks," in *Proc. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, INRIA Sophia-Antipolis, France, Mar. 2003.

[8] G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "Collision-resolution Algorithms for Multiple Destinations in Wireless Networks," in *Proc. Conference on Information Sciences and Systems*, Baltimore, MD, Mar. 2003.

[9] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, Englewood Cliffs, NJ, 1992.

[10] P. Gupta and P.R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, vol. 46, no.2, pp. 388-404, Mar. 2000.

[11] W. Szpankowski, "A Multiqueue Problem: Bounds and Approximations," *Adv. Appl. Probab.*, pp. 498-515, vol. 26, 1994.

[12] R. Rao and A. Ephremides, "On the Stability of Interacting Queues in a Multi-Access System," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 918-930, Sept. 1988.

[13] W. Luo and A. Ephremides, "Stability of N Interacting Queues in Random-Access Systems," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1579-1587, July. 1999.

[14] S. Ramanathan and E. L. Lloyd,"Scheduling Algorithms for Multihop Radio Networks," *IEEE/ACM Transactions on Networking*, vol. 1, no. 2, pp. 166-177, Apr. 1993.

[15] I. Holyer, The NP-Completeness of Edge Colorings," *SIAM Journal on Computing*, vol. 10, pp. 718-720, 1981.

[16] Y. Jin and G. Kesidis, "Equilibria of a Non-cooperative Game for Heterogeneous Users of an ALOHA Network," *IEEE Communication Letters*, vol. 6, no 7, pp. 282-284, 2002.

[17] A.B. MacKenzie and S.B. Wicker, "Selfish Users in ALOHA: A Game Theoretic Approach," in *Proc. Fall 2001 IEEE Vehicular Technology Conference*, Atlantic City, NJ, Oct. 2001.

[18] A. B. MacKenzie and S. B. Wicker, "Stability of Slotted ALOHA with Multi-packet Reception and Selfish Users," in *Proc. IEEE INFOCOM*, San Francisco, CA, Apr. 2003.

[19] E. Altman, R. El-Azouzi and T. Jimenez, "Slotted ALOHA as a Stochastic Game with Partial Information," in *Proc. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, INRIA Sophia-Antipolis, France, Mar. 2003.

[20] V. Anantharam, "Stability Region of the Finite-User Slotted ALOHA Protocol," *IEEE Transactions on Information Theory*, vol. 37, no. 3, pp. 535-540, May 1991.

[21] V. Naware, G. Mergen, and L. Tong, "Stability and Delay of Finite-User Slotted ALOHA with Multipacket Reception," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2636-2656, July 2005.

[22] J. Luo and A. Ephremides, "On the Throughput, Capacity, and Stability Regions of Random Multiple Access," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2593-2607, June 2006.

[23] P. Marbach and R. Pang, "Transmission Costs, Selfish Nodes, and Protocol Design," in *Proc. WiOpt'05: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, Riva del Granda, Italy, Apr 2005.

[24] Y. E. Sagduyu and A. Ephremides, "Power Control and Rate Adaptation as Stochastic Games for Random Access," in *Proc. 42nd IEEE Conference on Decision and Control*, Maui, HI, Dec. 2003.

[25] Y. E. Sagduyu and A. Ephremides, "A Game-Theoretic Analysis of Denial of Service Attacks in Wireless Random Access," in *Proc. International Symposium on Optimization and Modeling in Mobile, Ad-Hoc and Wireless Networks, WiOpt'07*, Limassol, Cyprus, Apr. 2007.

[26] Y. E. Sagduyu and A. Ephremides, "A Game-Theoretic Look at Throughput and Stability in Random Access," in *Proc. IEEE Military Communications Conference MILCOM 2006*, Washington DC, Oct. 2006.

[27] J.-W. Lee, M. Chiang, and A. R. Calderbank, "Utility-Optimal Medium Access Control: Reverse and Forward Engineering," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.

[28] H. Inaltekin and S. Wicker, "A One-shot Random Access Game for Wireless Networks," in *Proc. 2005 International Conference on Wireless Networks, Communications and Mobile Computing*, pp. 940-945, Maui, HI, June 2005.

[29] D. Wang, C. Comaniciu, and U. Tureli, "Cooperation and Fairness for Slotted ALOHA", *Special Issue of Springer Wireless Personal Communications on Cooperation in Wireless Networks*, 2006.

[30] J. Zander, "Jamming in Slotted ALOHA Multihop Packet Radio Networks," *IEEE Transactions on Communications*, vol. 39, no. 10, pp. 1525-1531, Oct. 1991.

[31] Y. Liu, C. Comaniciu, and H. Man, "A Bayesian Game Approach for Intrusion Detection in Wireless Ad Hoc Networks," in *Proc. GameNets (Workshop on Game Theory for Networks)*, Pisa, Italy, Oct. 2006.

[32] V. Gupta, S. V. Krishnamurthy, and M. Faloutsos, "Denial of Service Attacks at the MAC Layer in Wireless Ad Hoc Networks," in *Proc. IEEE Military Communications Conference MILCOM 2002*, Anaheim, CA, Oct. 2002.

[33] S. Radosavac, A. A. Cardenas, J. S. Baras, and G. V. Moustakides, "Detecting IEEE 802.11 MAC Layer Misbehavior in Ad Hoc Networks: Robust Strategies against Individual and Colluding Attackers," *Journal of Computer Security*, vol. 15, no. 1, pp. 103-128, Jan. 2007.

[34] R. Loynes, "The Stability of a Queue with Non-interdependent Inter-arrival and Service Times," *Proc. Camb. Philos. Soc.*, vol. 58, pp. 497-520, 1962.

[35] J. S. Shamma and G. Arslan, "Dynamic Fictitious Play, Dynamic Gradient Play, and Distributed Convergence to Nash Equilibria," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 312-327, Mar. 2005.

[36] Y. E. Sagduyu and A. Ephremides, "On Broadcast Stability Region in Random Access through Network Coding," in *Proc. Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 2006.

[37] I. Cidon, H.Kodesh and M. Sidi, "Erasure, Capture, and Random Power Level Selection in Multiple-Access Systems," *IEEE Transactions on Communications*, vol. 36, no 3, pp. 263-271, Mar. 1988.

[38] E. Altman, D. Barman, A. Benslimane and R. E. Azouzi, "Slotted ALOHA with Priorities and Random Power," in *Proc. IFIP Networking Conference*, Ontario, Canada, May 2005.

[39] N. Feng, N. B. Mandayam and D. J. Goodman, "Joint Power and Rate Optimization for Wireless Data Services Based on Utility Functions," in *Proc. Conference on Information Sciences and Systems*, Baltimore, MD, Mar. 1999.

[40] Y. E. Sagduyu and A. Ephremides, "A Game-Theoretic Look at Simple Relay Channel," ACM/Kluwer Journal of Wireless Networks, vol. 12, no. 5, pp. 545-560, Oct. 2006.

[41] G. J. Foschini and Z. Miljanic, "A Simple Distributed Autonomous Power Control Algorithm and its Convergence," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 641-646, Nov. 1993.

[42] J. Zander, "Performance of Optimum Transmitter Power Control in Cellular Radio Systems," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 1, pp. 57-62, Feb. 1992.

[43] R. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341-1347, Sep. 1995.

[44] H. Ji and C. Huang, "Non-cooperative Uplink Power Control in Cellular Radio Systems," *ACM/Kluwer Journal of Wireless Networks*, vol. 4, no. 3, pp. 233-240, Apr. 1998.

[45] C. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient Power Control via Pricing in Wireless Data Networks," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 291-303, Feb. 2002.

[46] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An Energy-Efficient Approach to Power Control and Receiver Design in Wireless Data Networks," *IEEE Transactions on Communications*, vol. 53, no. 11, pp. 1885-1894, Nov. 2005.

[47] R. J. La, and V. Anantharam, "A Game-theoretic Look at the Gaussian Multi-access Channel," *DIMACS Workshop on Network Information Theory*, Rutgers University, Piscataway, NJ, USA, Mar. 2003.

[48] F. Meshkati, H. V. Poor, and S. C. Schwartz, "A Non-cooperative Power Control Game in Delay-Constrained Wireless Multiple-Access Networks," in *Proc. 2005 IEEE International Symposium on Information Theory*, Adelaide, Australia, Sept. 2005.

[49] Y. E. Sagduyu, A. Ephremides, "SINR-Based MAC Games for Selfish and Malicious Users," in *Proc. Information Theory and Applications Workshop 2007*, San Diego, CA, Jan. 2007.

[50] T. Alpcan, T. Basar, R. Srikant and E. Altman, "CDMA Uplink Power Control as a Noncooperative Game," *ACM/Kluwer Journal of Wireless Networks*, vol. 8, no. 6, pp. 659-670, 2002.

[51] M. Xiao, N. B. Shroff, and E. K. P. Chong, "A Utility-Based Power-Control Scheme in Wireless Cellular Systems," *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 210-221, Apr. 2003.

[52] T. Cover and A. El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572-584, Sept. 1979.

[53] A. Orda, R. Rom and N. Shimkin, "Competitive Routing in Multi-user Environments," *IEEE/ACM Transactions on Networking*, vol. 1, no. 5, pp. 510-521, Oct. 1993.

[54] A. Urpi, M. Bonuccelli, and S. Giordano, "Modelling Cooperation in Mobile Ad Hoc Networks: A Formal Description of Selfishness," in *Proc. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, INRIA Sophia-Antipolis, France, Mar. 2003.

[55] V. Srinivasan, P. Nuggehalli, C. F. Chiasserini, and R. R. Rao, "Cooperation in Wireless Ad Hoc Networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, Apr. 2003.

[56] L. Buttyan and J. P. Hubaux, "Stimulating Cooperation in Self-organizing Mobile Ad Hoc Networks," *ACM Journal for Mobile Networks (MONET)*, vol. 8, no. 5, Oct. 2003.

[57] M. L. Littman, "Markov Games as a Framework for Multi-agent Reinforcement Learning," in *Proc. Eleventh International Conference on Machine Learning*, New Brunswick, pp. 157-163, 1994 .

[58] J. Hu and M. P. Wellman, "Nash Q-Learning for General-Sum Stochastic Games," *Journal of Machine Learning Research*, vol. 4, pp. 1039-1069, Nov. 2003.

[59] M. Bowling and M. Veloso, "Rational and Convergent Learning in Stochastic Games," in *Proc. Seventeenth International Joint Conference on Artificial Intelligence*, Seattle, WA, pp. 1021-1026, Aug. 2001.

[60] R. Ahlswede, N. Cai, S. -Y. R. Li, and R. W. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204-1216, July 2000.

[61] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A Random Linear Network Coding Approach to Multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413-4430, Oct. 2006.

[62] T. Ho, B. Leong, M. Medard, R. Koetter, Y. Chang, and M. Effros, "On the Utility of Network Coding in Dynamic Environments," in *Proc. International Workshop on Wireless Ad-Hoc Networks (IWWAN)*, Oulu, Finland, June 2004.

[63] T. Ho, M. Medard, J. Shi, M. Effros and D. R. Karger, "On Randomized Network Coding", in *Proc. 41st Allerton Annual Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2003.

[64] A. Ramamoorthy, J. Shi, and R. D. Wesel, "On the Capacity of Network Coding for Random Networks," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2878-2885, Aug. 2005.

[65] D. S. Lun, M. Medard, T. Ho, and R. Koetter, "Network Coding with a Cost Criterion," in *Proc. 2004 Internatonal Symposium on Information Theory and its Applications (ISITA)*, Parma, Italy, Oct. 2004.

[66] D. S. Lun, N. Ratnakar, R. Koetter , M. Medard , E. Ahmed , and H. Lee, "Achieving Minimum-Cost Multicast: A Decentralized Approach Based on Network Coding," in *Proc. IEEE INFOCOM*, vol. 3, pp. 16907-1617, Miami, FL, Mar. 2005.

[67] D. S. Lun, N. Ratnakar, M. Medard, R. Koetter, D. R. Karger, T. Ho, E. Ahmed, and F. Zhao, "Minimum-Cost Multicast over Coded Packet Networks," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2608-2623, June 2006.

[68] D. S. Lun, M. Medard, and M. Effros. "On Coding for Reliable Communication over Packet Networks," in *Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2004.

[69] J. Widmer, C. Fragouli, and J.-Y. Le Boudec, "Energy-Efficient Broadcasting in Wireless Ad-Hoc Networks," in *Proc. First Workshop on Network Coding, Theory, and Applications*, Riva del Garda, Italy, Apr. 2005.

[70] Y. Wu, P. A. Chou, Q. Zhang, K. Jain, W. Zhu, and S.-Y. Kung, "Network Planning in Wireless Ad Hoc Networks: A Crosslayer Approach," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 136-150, Jan. 2005.

[71] Y. Wu, P. A. Chou, and S.-Y. Kung, "Minimum-Energy Multicast in Mobile Ad Hoc Networks Using Network Coding," *IEEE Transactions on Communications*, vol. 53, no. 11, pp. 1906-1918, Nov. 2005.

[72] C. Fragouli and E. Soljanin, "Information Flow Decomposition for Network Coding," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 829-848, Mar. 2006.

[73] Y. E. Sagduyu and A. Ephremides, "The Problem of Medium Access Control in Wireless Sensor Networks," *IEEE Wireless Communications*, vol. 11. no. 6, pp. 44-53, Dec. 2004.

[74] A. F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of Wireless Erasure Networks," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 789-804, Mar. 2006.

[75] T. Cover and J. Thomas, *Elements of Information Theory*, New York: John Wiley and Sons, 1991.

[76] S. -Y. R. Li, R. W. Yeung, and N. Cai, "Linear Network Coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 1204-1216, Feb. 2003.

[77] R. Koetter and M. Medard, "An Algebraic Approach to Network Coding," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782-795, Oct. 2003.

[78] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. M. G. M. Tolhuizen, "Polynomial Time Algorithms for Multicast Network Code Construction," in *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1973-982, June. 2005.

[79] C. Fragouli, E. Soljanin, and A. Shokrollahi, "Network Coding as a Coloring Problem," in *Proc. Conference on Information Sciences and Systems*, Princeton, NJ, Mar. 2004.

[80] D. Brelaz, "New Methods to Color the Vertices of a Graph," *Communications of the ACM*, vol. 22, no. 4, pp. 251-256, Apr. 1979.

[81] C. Fragouli, E. Soljanin, "Decentralized Network Coding," in *Proc. 2004 IEEE Information Theory Workshop*, San Antonio, Oct., 2004.

[82] A. Ephremides and B. Hajek, "Information Theory and Communication Networks: An Unconsummated Union," *IEEE Transactions on Inform. Theory*, vol. 44, no. 6, pp. 2416-2434, Oct. 1998.

[83] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource Allocation and Cross-Layer Control in Wireless Networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1-144, 2006.

[84] L. Tassiulas and A. Ephremides, "Stability Properties of Constrained Queueing Systems and Scheduling for Maximum Throughput in Multihop Radio Networks", *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936-1949, Dec. 1992.

[85] Y. E. Sagduyu and A. Ephremides, "On Joint MAC and Network Coding in Wireless Ad Hoc Networks," to appear in *IEEE Transactions on Information Theory*, Oct. 2007.

[86] T. Ho and H. Viswanathan, "Dynamic Algorithms for Multicast with Intra-session Network Coding", in *Proc. 43rd Allerton Annual Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2005.

[87] C. Chekuri, C. Fragouli, and E. Soljanin "On Average Throughput and Alphabet Size in Network Coding," in *Proc. IEEE International Symposium on Information Theory*, pp. 1593-1597, Adelaide, Australia, Sept. 2005.

[88] Y. Wu, P. Chou, S.-Y. Kung, "Information Exchange in Wireless Networks with Network Coding and Physical-layer Broadcast," in *Proc. Conference on Informations Sciences and Systems*, Baltimore, MD, Mar. 2005.

[89] S. Katti, D. Katabi, W. Hu, H. Rahul, and M. Medard, "The Importance of Being Opportunistic: Practical Network Coding for Wireless Environments," in *Proc. 43rd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2005.

[90] T. Roughgarden and E. Tardos, "How Bad is Selfish Routing?," *Journal of the ACM*, vol. 49, no.2, pp. 236-259, Mar. 2002.

[91] A. Eryilmaz, A. Ozdaglar, and M. Medard, "On Delay Performance Gains from Network Coding," in *Proc. Conference on Informations Sciences and Systems*, Princeton, NJ, Mar. 2006.

[92] S. Wright and J. Nocedal, *Numerical Optimization*, Springer, 1999.

[93] P. Michiardi and R. Molva, "A Game Theoretical Approach to Evaluate Cooperation Enforcement Mechanisms in Mobile Ad Hoc Networks", in *Proc. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, INRIA Sophia-Antipolis, France, Mar. 2003.

[94] J. Crowcroft, R. Gibbens, F. Kelly and S. Ostring, "Modelling Incentives for Collaboration in Mobile Ad Hoc Networks", in *Proc. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, INRIA Sophia-Antipolis, France, Mar. 2003.

[95] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks", in *Proc. ACM International Conference on Mobile Computing and Networking*, pp. 255-265, Boston, MA, 2000.

[96] M. Felegyhazi, L. Buttyan, and J. P. Hubaux, "Nash Equilibria of Packet Forwarding Strategies in Wireless Ad Hoc Networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 4, pp. 1-14, Apr. 2006.

[97] Y. Qiu and P. Marbach, "Bandwidth Allocation in Wireless Ad Hoc Networks: A Price-based Approach," in *Proc. IEEE INFOCOM*, San Francisco, CA, Mar. 2003.

[98] M. Effros, T. Ho, and S. Kim, "A Tiling Approach to Network Code Design for Wireless Networks," in *Proc. 2006 IEEE Information Theory Workshop*, Punta del Este, Uruguay, Mar. 2006.

[99] J. L. Massey and P. Mathys, "The Collision Channel without Feedback," *IEEE Transactions on Information Theory*, vol. 31, no. 2, pp. 192-204, Mar. 1985.

[100] B. Shrader and A. Ephremides, "The Capacity of the Asynchronous Compound Multiple Access Channel and Results for Random Access Systems," in *Proc. IEEE International Symposium on Information Theory*, Seattle, WA, July. 2006.

[101] B. Shrader and A. Ephremides, "Broadcast Stability in Random Access," in *Proc. IEEE International Symposium on Information Theory*, Adelaide, Australia, Sept. 2005.

[102] M. Luby, "LT-Codes," in *Proc. ACM Symposium on Foundations of Computer Science (FOCS)*, 2002.

[103] A. Shokrollahi, "Raptor Codes," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551-2567, June 2006.

[104] T. M. Cover, "Comments on Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2524-2530, Oct. 1998.

[105] Y. E. Sagduyu and A. Ephremides, "Some Optimization Trade-offs in Wireless Network Coding," in *Proc. Conference on Information Sciences and Systems*, Princeton, NJ, Mar. 2006.

[106] Y. E. Sagduyu and A. Ephremides, "Network Coding in Wireless Queueing Networks: Tandem Network Case," in *Proc. IEEE International Symposium on Information Theory*, Seattle, WA, July 2006.

[107] D. S. Lun, P. Pakzad, C. Fragouli, M. Medard, and R. Koetter, "An Analysis of Finite-Memory Random Linear Coding on Packet Streams," in *Proc. 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, Apr. 2006.

[108] M. Effros, M. Medard, T. Ho, S. Ray, D. Karger, and R. Koetter, "Linear Network Codes: A Unified Framework for Source Channel, and Network Coding", *DIMACS Workshop on Network Information Theory*, Rutgers University, Piscataway, NJ, Mar. 2003.

[109] Y. E. Sagduyu and A. Ephremides, "On the Capacity Bounds of Wireless Networks with Directional Antennas," in *Proc. Conference on Information Sciences and Systems CISS 2004*, Princeton, NJ, March 2004.