Scheduling Algorithms for Throughput Maximization in Time-Varying Networks with Reconfiguration Delays

by

Güner Dinçer Çelik

B.Sc., Electrical and Electronics Engineering
Middle East Technical University, Ankara, TURKEY (2005)

S.M., Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge, MA (2007)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2012

Author .....
Department of Electrical Engineering and Computer Science
August 2, 2012

Certified by .....
Eytan Modiano
Professor
Thesis Supervisor

Accepted by .....
Leslie A. Kolodziejski
Chair, Department Committee on Graduate Students

Scheduling Algorithms for Throughput Maximization in Time-Varying Networks with Reconfiguration Delays

by

Güner Dinçer Çelik

Submitted to the Department of Electrical Engineering and Computer Science
on August 2, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

We consider the control of possibly time-varying wireless networks under reconfiguration delays. Reconfiguration delay is the time it takes to switch network resources from one subset of nodes to another and it is a widespread phenomenon observed in many practical systems. Optimal control of networks has been studied to a great extent in the literature, however, the significant effects of reconfiguration delays received limited attention. Moreover, simultaneous presence of time-varying channels and reconfiguration delays has never been considered and we show that it impacts the system fundamentally.

We first consider a Delay Tolerant Network model where data messages arriving randomly in time and space are collected by mobile collectors. In this setting reconfiguration delays correspond to travel times of collectors. We utilize a combination of *wireless transmission* and *controlled mobility* to improve the system delay scaling with load ρ from $\Theta(\frac{1}{(1-\rho)^2})$ to $\Theta(\frac{1}{1-\rho})$, where the former is the delay for the corresponding system without wireless transmission. We propose control algorithms that stabilize the system whenever possible and have optimal delay scaling.

Next, we consider a general queuing network model under reconfiguration delays and interference constraints which includes wireless, satellite and optical networks as special cases. We characterize the impacts of reconfiguration delays on system stability and delay, and propose scheduling algorithms that persist with service schedules for durations of time based on queue lengths to minimize negative impacts of reconfiguration delays. These algorithms provide throughput-optimality *without requiring knowledge of arrival rates* since they dynamically adapt inter-switching durations to stochastic arrivals.

Finally, we present optimal scheduling under time-varying channels and reconfiguration delays, which is the main contribution of this thesis. We show that under the simultaneous presence of these two phenomenon network stability region shrinks, previously suggested policies are unstable, and new algorithmic approaches are necessary. We propose techniques based on state-action frequencies of Markov Decision Process theory to characterize the network stability region and propose throughput-optimal algorithms. The state-action frequency technique is applicable to a broad class of systems with or without

reconfiguration delays, and provides *a new framework for characterizing network stability region and developing throughput-optimal scheduling policies.*

Thesis Supervisor: Eytan Modiano
Title: Professor

Acknowledgments

I am profoundly thankful to my advisor, Prof. Eytan Modiano, for his support for this work and for his invaluable advise that helped me continue my research in the right direction. He has been more than an advisor to me, he has been a friend whose insightful guidance and positive mood have never changed during my Ph.D. I would like to thank Long Le, Sem Borst and Phil Whiting for their invaluable contributions and suggestions during our research collaborations at Bell Labs and MIT. I thank my thesis committee members Prof. Devavrat Shah, Prof. Emilio Frazzoli, and Prof. Munther Dahleh for their support and guidance throughout my Ph.D.

I would like to thank my mother, my father, my sister and her husband, and my friends for their endless moral support throughout my career. I have always felt your love and friendship in my hearth. I would also like to thank my current office mates Seb, Matt, Greg, Nathan, Marzieh, Kyu, Mihalis, and alumni of our research group Anand, Gil, Krishna, Andrew, Murtaza, Jun, and Li-Wei for their friendly mood, for the technical discussions we have had, and for the fun atmosphere that they helped create in the office, which I have enjoyed being a part of throughout my Ph.D.

Finally, this work was supported by NSF grants CNS-0626781 and CNS-0915988, by ARO Muri grant number W911NF-08-1-0238, and by NSF ITR grant CCR-0325401.

Contents

1 Introduction **15**

1.1 Related Work 18

1.2 Contributions 20

1.2.1 Data Collection via Mobile Servers in DTNs 21

1.2.2 Scheduling in Networks with Reconfiguration Delays 23

1.2.3 Scheduling in Networks with Time-Varying Channels and Reconfiguration Delay 24

2 System Description and Assumptions **29**

3 Data Collection via Mobile Servers in DTNs **35**

3.0.4 Related Work 37

3.0.5 Contributions and Outline 41

3.1 The Single Collector Case 42

3.1.1 Model 43

3.1.2 Stability 45

3.1.3 Lower Bound On Delay 48

3.1.4 Collector Policies 51

3.2 Multiple collectors - Interference-Free Networks 57

3.3 Multiple Collectors - Systems with Interference Constraints 62

3.3.1 Framed-Max-Weight Policy 65

3.4 Concluding Remarks 68

4 Scheduling in Queueing Networks with Reconfiguration Delays **97**

4.1 Related Work . . . 102

4.2 Organization . . . 104

4.3 System Description and Preliminaries . . . 104

4.4 Motivation . . . 105

4.5 Throughput-Optimal Algorithms . . . 107

4.5.1 Sufficient Stability Conditions . . . 108

4.5.2 Hysteresis Based (HB) Algorithms . . . 111

4.5.3 Switching Curve Based (SCB) Algorithms . . . 112

4.6 Simulations . . . 114

4.7 Related Work . . . 119

4.8 Concluding Remarks . . . 121

5 Dynamic Server Allocation over Time Varying Channels with Switching Delay 151

5.0.1 Related Work . . . 154

5.0.2 Main Contribution and Organization . . . 155

5.1 Two-Queue System . . . 156

5.1.1 System Model . . . 156

5.1.2 Motivation: Channels Without Memory . . . 158

5.1.3 Channels With Memory - Stability Region . . . 160

5.1.4 Frame Based Dynamic Control (FBDC) Policy . . . 172

5.1.5 Myopic Control Policies . . . 178

5.2 General System . . . 185

5.2.1 Model . . . 185

5.2.2 Stability Region . . . 186

5.2.3 Myopic Policy for the Saturated System . . . 193

5.2.4 Frame-Based Dynamic Control Policy . . . 197

5.2.5 Myopic Control Policies . . . 201

5.3 Numerical Results . . . 203

5.4 Concluding Remarks . . . 208

6 Scheduling in Networks with Time-Varying Channels and Reconfiguration Delays **251**

6.1 Main Contribution and Organization 254

6.2 Model 254

6.3 Memoryless Channels 257

6.3.1 Stability Region 257

6.3.2 Variable Frame Based Max-Weight (VFMW) Algorithm 258

6.3.3 Simulation Results - Memoryless Channels 261

6.4 Channels With Memory 263

6.4.1 Stability Region 264

6.4.2 Frame Based Dynamic Control Policy 270

6.4.3 Myopic Control Policies 272

6.4.4 Simulation Results - Channels with Memory 273

6.5 Concluding Remarks 276

7 Conclusion and Future Work **297**

7.1 Future Directions 299

List of Figures

2-1 System model. A single-hop wireless network with interference constraints, time-varying channels and reconfiguration delays. 31

2-2 An example 4x3 satellite network. Ground stations are subject to time-varying channels C_1, C_2, C_3, C_4 and the servers are subject to T_r slot reconfiguration (switchover) delay. Server 2 is forced to be idle due to interference constraints. 32

3-1 The system model for the case of a single collector. The collector adjusts its position in order to collect randomly arriving messages via wireless communication. The circles with radius r^* represent the communication range and the dashed line segments represent the collector's path. 36

3-2 The partitioning of the network region into square subregions of side $\sqrt{2}r^*$. The circle with radius r^* represents the communication range and the dashed lines represent the collector's path. 54

3-3 Delay lower bound vs. network load using different communication ranges for $A = 200$, $v = 1$ and $s = 1$. 57

3-4 Delay in the Partitioning policy vs the delay lower bound for $r^* = 2.2$, $\beta = 2$ and $\alpha = 4$. Case-1: Dominant travel time ($A = 800$, $v = 1$, $s = 2$). Case-2: Comparable travel and reception times ($A = 60$, $v = 10$, $s = 2$). 58

3-5 The network model. Red regions are the exclusion zones for the servers currently in service. Two servers are forced to be inactive since the messages in their vicinity are in the exclusion zones of other servers. 64

4-1 System model. A single-hop wireless network with interference constraints and reconfiguration delays. 99

4-2 System model with $N = 4$ queues and $M = 3$ servers. Server 2 is forced to be idle due to interference constraints. 99

4-3 The total average queue size for (a) the Max-Weight policy, (b) the SCB policy and (c) the VFMW policy for 2 queues and $T_r = 2$ slots. 115

4-4 Delay (total average queue size) vs the throughput under the Max-Weight, MWFF and the VFMW policies for 2 queues and $T_r = 20$ slots. 118

4-5 Delay (total average queue size) vs the throughput under the VFMW, HB, and the SCB policies for the 4-queues and 3-servers system and $T_r = 10$ slots. 118

5-1 System model. Parallel queues with randomly varying connectivity processes $C_1(t), C_2(t), ..., C_N(t)$ and $t_s = 1$ slot switching time. 154

5-2 Markov modulated ON/OFF channel process. We have $p_{10} + p_{01} < 1$ ($\epsilon < 0.5$) for positive correlation. 157

5-3 Stability region under memoryless (i.i.d.) channels and channels with memory (Markovian with $\epsilon < 0.5$) with and without switching delay. 159

5-4 Stability region under channels with memory, with and without switching delay for (a) $\epsilon = 0.25 < \epsilon_c$ and (b) $\epsilon = 0.40 \geq \epsilon_c$. 171

5-5 Stability region for 3 parallel queues for $p_{10} = p_{01} = 0.3$. 191

5-6 Stability region outer bound for 3 parallel queues for $p_{10} = p_{01} = 0.3$. . . . 192

5-7 The total average queue size for (a) the FBDC policy and (b) the Myopic policy for $T = 25$ and $\epsilon = 0.25$. 205

5-8 The total average queue size for (a) the FBDC policy and (b) the Myopic Policy implemented without the use of frames (i.e., for $T = 1$) and $\epsilon = 0.40$. 206

5-9 Delay vs Sum-throughput for the FBDC, the OLM, and the Max-Weight policies implemented without the use of frames (i.e., for $T = 1$) for $N = 2$ queues and $\epsilon = 0.40$. 207

5-10 Delay vs Sum-throughput for the FBDC, the OLM, and the Max-Weight policies implemented without the use of frames (i.e., for $T = 1$) for $N = 3$ queues and $\epsilon = 0.30$. 208

6-1 System model. A single-hop wireless network with interference constraints, time-varying channels and reconfiguration delays. 252

6-2 An example 4x3 satellite network. Ground stations are subject to time-varying channels $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$, $\mathbf{C}_4$ and the servers are subject to T_r slot reconfiguration (switchover) delay. Server 2 is forced to be idle due to interference constraints. 253

6-3 Markov modulated ON/OFF channel process. The case of $p_{10} + p_{01} < 1$ provides positive correlation. 256

6-4 Delay vs throughput under the VFMW, MW, and the FFMW policies. . . . 262

6-5 Total average queue size vs the sum-throughput under the FBDC policy with frame length 10, the OLM, and the MW policies, for $p_{10} = p_{01} = 0.25$ (upper figure), and for $p_{10} = p_{01} = 0.10$ (lower figure). 274

6-6 Total average queue size vs the sum-throughput under the FBDC policy implemented without frames, the OLM, and the MW policies for $p_{10} = p_{01} = 0.30$. 276

Chapter 1

Introduction

We consider the impact of reconfiguration delays on network scheduling in this thesis. Reconfiguration delay is the duration of time required for the network to switch its resources from a subset of users to another subset, and it occurs in many telecommunication applications such as wireless, satellite, optical or delay tolerant networks (DTNs) [3], [20], [81], [126]. Optimal scheduling problem for communication networks subject to time-varying link gains and constraints on simultaneous usage of network resources has been a very active field (e.g., [43, 50, 75, 82, 86, 87, 99, 103, 109, 110, 122, 125]). However, the significant effects of server switching delays or the time to reconfigure schedules have been largely ignored. In satellite networks where multiple mechanically steered antennas are servicing ground stations, the reconfiguration delay can be due to the time for a satellite antenna to steer from one station to another. This delay can be around 10ms [20], [112], which is of the same order as the time to transmit multiple packets. Similarly, in optical communication systems, laser tuning delay for different wavelengths can take significant time (μs-ms) [26], [81]. For an optical transceiver transmitting a packet of length 10000 bits at 10 Gb/s, even relatively small reconfiguration delays on the order of micro seconds correspond to the time to transmit multiple packets [81]. Large switching delays occur in Delay

Tolerant Networks (DTNs) where mobile servers (e.g., unmanned aerial vehicles (UAV)) are used as data collectors from sensors in a field [90], [117]. In wireless networks, delays for electronic beamforming or channel switching that occurs in oscillators can be more than $200\mu s$ [3], [20], [112], [126]. Worse yet, such small delay is impossible to achieve due to delays incurred during different processing tasks such as channel estimation, signal to interference ratio calculation, and power control at the physical layer [3], [58], and stopping and restarting the interrupt service routines of various drivers in upper layers [3], [95], which cause switching delays on the order of milliseconds [95].

We consider several network models such as DTNs, wireless uplinks/downlinks, single-hop networks, or optical networks and study the impact of reconfiguration delays on throughput and delay performance of such networks. We model networks as stochastic and possibly time-varying systems and use queuing theoretical analysis to study their performance. Under such a network model, messages (or packets) arrive to network nodes according to a stochastic process and are transmitted to their destinations over network links. These transmissions need to be scheduled because simultaneous transmissions may interfere with each other. Sets of links that can be activated simultaneously without interfering with each other are called feasible schedules or interference constraints of the system, and we assume that such schedules are made available to the network scheduler (controller). Such constraints may be due to, for instance, Signal to Interference and Noise Ratio (SINR) requirements at the receiver nodes [69]. Furthermore, wireless links may have gains that vary over time. Such time variations may be due to, for instance, fading or shadowing effects [86], [110]. In this thesis, we model such time variations in link gains as stochastic processes.

Reconfiguration delay is a relatively new component of this network model that received limited attention in the literature. Moreover, networks subjected to the simultaneous

presence of reconfiguration delays and time-varying channels have not been considered in the literature. Such networks are significantly harder to analyze and we show that the simultaneous presence of time-varying channels and reconfiguration delays significantly reduces the system stability region and changes the structure of optimal policies.

1.1 Related Work

In this section we give a brief overview of related work as it pertains to the overall thesis. In addition, within each chapter, we discuss previous work specifically related to that chapter in detail. Optimal control of queuing systems and communication networks has been a very active research topic over the past two decades (e.g., [43, 50, 75, 82, 86, 87, 103, 109, 110, 122, 125]). In the seminal paper [109], Tassiulas and Ephremides characterized the stability region of multihop wireless networks and proposed the throughput-optimal Max-Weight scheduling algorithm. In [110], the same authors considered a parallel queuing system with randomly varying connectivity where they characterized the stability region of the system explicitly and proved the throughput-optimality of the Longest-Connected-Queue scheduling policy. These results were later extended to joint power allocation and routing in wireless networks in [86, 87] and optimal scheduling for switches in [99, 103]. More recently, suboptimal distributed scheduling algorithms with throughput guarantees were studied in [28, 69, 75, 122], while [43, 82] developed distributed algorithms that achieve throughput-optimality (see [50], [84] for a detailed review). The effect of delayed channel state information was considered in [60, 91, 125] which showed that the stability region is reduced and that a policy similar to the Max-Weight algorithm is throughput-optimal. These works do not consider switching delays.

Switching delay has been considered in Polling models in the Queuing Theory com-

munity. Data collection by a server from a finite number of queues has been thoroughly analyzed under Polling models in Queuing Theory literature (e.g., [8], [21], [22], [24], [45], [51], [72], [77], [124]). Stability of Polling systems under Exhaustive, Gated or Limited service disciplines under different routing models was studied in [8], [45], and [51]. Steady-state queue length distribution and various delay properties were considered in [21], [22], [24] and [124], and optimal server routing and various dominance relationships were analyzed in [72] and [77]. These works do not consider time-varying channel gains in the system. A detailed survey of the works in this field can be found in [105] or [114].

The problem of optimally servicing job requests by mobile servers in a 2-dimensional space, related to the data collection problem in DTNs, has been extensively studied in the Vehicle Routing Problems (VRPs) literature (e.g., [7], [12], [16], [46], [80], [107], [115], [116]). The common example of a VRP is the Euclidean Traveling Salesman Problem (TSP) in which a single server is to visit each member of a fixed set of locations on the plane such that the total travel cost is minimized. Several extensions of TSP have been considered in the literature such as random demand arrivals and the use of multiple servers [16], [17], [46], [117]. Of particular relevance to the work on DTNs in this thesis is the Dynamic Traveling Repairman Problem (DTRP) due to Bertsimas and Van Ryzin [16], [17], [18]. DTRP is a stochastic and dynamic VRP in which vehicles are to serve demands that arrive randomly in time and space. Fundamental lower bounds on delay were established and several vehicle routing policies were analyzed for DTRP for a single server in [16], for multiple servers in [17], and for general demand and interarrival time distributions in [18]. Later, [115], [116], [117] generalized the DTRP model to analyze Dynamic Pickup and Delivery Problem (DPDP), where fundamental lower bounds on delay were established and several dynamic pick up and delivery policies were analyzed.

Furthermore, [117] considered the wireless DPDP problem for the first time, where the simultaneous use of controlled mobility and wireless transmission was studied for a simple two-queue network. The system was analyzed from a physical layer perspective and the stability properties of system as well as trade-offs between throughput and delay were studied. The first chapter of this thesis considers simpler wireless communication models as compared to the wireless DPDP problem in [117] and generalizes the DTRP model to wireless networks where the demands are data messages to be transmitted to a collector that can receive messages from a distance using wireless transmission.

1.2 Contributions

The main contribution of this thesis is solving the scheduling problem in networks subject to three fundamental features which are *reconfiguration delays*, *time-varying channels* and *interference constraints*. We develop different queueing models tailored for analysis of different network models such as DTNs, wireless uplinks/downlinks, single-hop networks, or optical networks and analyze the impact of reconfiguration delays on throughput and delay performance of such networks.

1.2.1 Data Collection via Mobile Servers in DTNs

We first consider a DTN model where messages arriving randomly in time at random locations in the network are collected by mobile receivers (collectors). The collectors are responsible for receiving these messages through *wireless communication* by dynamically adjusting their position in the network using *controlled mobility*. In such a setting, switching delays are in the form of travel times of mobile collectors. Our goal is to utilize the combination of *wireless transmission* and *controlled mobility* effectively to improve the

throughput and delay performance in such networks. First, we consider a system with a single collector. We show that the necessary and sufficient stability condition for such a system is given by $\rho < 1$ where ρ is the expected system load. We derive lower bounds for the expected message waiting time in the system, and develop policies that are stable for all loads $\rho < 1$ and that have asymptotically optimal delay scaling. We show that the combination of mobility and wireless transmission results in a delay scaling of $\Theta(\frac{1}{1-\rho})$ with the system load ρ, in contrast to the $\Theta(\frac{1}{(1-\rho)^2})$ delay scaling in the corresponding system without wireless transmission, where the collector must visit each message location. We propose dynamic routing and data collection policies based on the Traveling Salesman Problem with Neighborhoods (TSPN) and on partitioning the network region into subregions. We show that these policies stabilize the system whenever possible, and have optimal delay scaling.

Next, we consider the system with multiple collectors. In the case where simultaneous transmissions to different collectors do not interfere with each other, we show that both the stability condition and the delay scaling extend from the single collector case. Moreover, generalizations of the single collector policies, i.e., partitioning the network region into subregions, one for each collector, and then applying a single collector policy in each subregion, achieves optimal delay scaling. In the case where simultaneous transmissions to different collectors interfere with each other, we consider a simple grid mobility model and formulate a dynamic scheduling problem. In this case, the switching delay is given by the time to reshuffle the positions of mobile servers on the grid. We characterize the stability region of the system and show that a frame-based version of the well-known Max-Weight policy stabilizes the system using the arrival rate information. The topic of data collection via mobile servers in DTNs is addressed in Chapter 3.

1.2.2 Scheduling in Networks with Reconfiguration Delays

Motivated by the DTN scheduling problem under interference, we consider a general queuing model in Chapter 4 and study the scheduling problem for a large class of networks under interference constraints and reconfiguration delays, where reconfiguration delay in this case denotes the duration of time required to switch between feasible schedules. The general queuing model we consider in this part of the thesis is applicable to wireless networks without channel variations (i.e., without multipath fading, shadowing etc.), optical networks, or satellite networks. Under zero reconfiguration delay it is well-known that the Max-Weight scheduling algorithm is throughput-optimal without requiring knowledge of arrival rates. However, we show that this property of the Max-Weight policy no longer holds when there is a nonzero reconfiguration delay, as the policy makes schedule switching decisions frequently, incurring large throughput losses during switching. Therefore, algorithms that dynamically arrange scheduling service intervals as functions of the system state and the reconfiguration delay are necessary.

We propose scheduling algorithms that persist with the current service schedule for a duration of time determined as a function of the queue sizes. We comprehensively analyze several such methods for choosing the length of service intervals.We propose the Variable Frame-Based Max-Weight (VFMW) algorithms that are implemented over frames, and that employ the Max-Weight schedule corresponding to the beginning of the frame during an interval of duration set as a deterministic function of the queue lengths at the beginning of the frame. Next, we consider algorithms that persist with the current schedule by giving an additive *bias* to the weight of the current schedule, which are functions of queue lengths in the system. We call such algorithms the *Switching Curve Based* (SCB) algorithms. We show that the bias sizes that are subliner functions of the *current* queue lengths lead to good

throughput and delay properties. The SCB policies provide more adaptability to bursts in arrivals by allowing the system to track the queue dynamics more closely as compared to the VFMW policies. This is because the SCB algorithms do not set a fixed frame-size at the beginning of each frame, but they rather make switching decisions based on the current queue states. The VFMW and the SCB algorithms provide throughput-optimality *without requiring knowledge of the arrival rates* since they dynamically adapt the inter-switching duration to the stochastic arrivals using queue sizes.

1.2.3 Scheduling in Networks with Time-Varying Channels and Reconfiguration Delay

Next, we move onto the main contribution of this thesis in chapters 5 and 6; namely, the optimal network scheduling problem under simultaneous presence of time-varying channels and switching delays, which has not been considered previously in the literature. We show that, as compared to the scheduling problem in networks subject to the presence of only one of these phenomenons, the network stability region shrinks, and optimal algorithms change; namely, previously suggested policies such as Max-Weight or Exhaustive scheduling are unstable. We propose techniques based on Markov Decision Processes (MDPs) in order to characterize the stability region of such networks and propose throughput-optimal algorithms with good delay properties.

We first consider the problem of dynamic server allocation over parallel queues subject to *time-varying channels* and *server switching delay* between the queues in Chapter 5. At each time slot the server dynamically decides to stay with the current queue, or switch to another queue based on the connectivity and the queue length information. We show that, as compared to systems without switching delay, the simultaneous presence of time-

varying channels (or randomly varying connectivity) and switching delay can significantly reduce the system stability region depending on the memory in the channel processes. Furthermore, we show that throughput-optimal policies take a very different structure from the celebrated Max-Weight algorithm or its variants.

In the first part of the chapter, we consider a system of two parallel queues, and develop an approach that explicitly characterizes the stability region of the system using *state-action frequencies* which are stationary solutions to a Markov Decision Process (MDP) formulation. We then develop a frame-based dynamic control (FBDC) policy, based on the state-action frequencies, and show that it is throughput-optimal asymptotically in the frame length. The FBDC policy is applicable to a broad class of network control systems and provides *a new framework for developing throughput-optimal network control policies* using state-action frequencies. Furthermore, we develop simple *Myopic policies* that provably achieve more than 90% of the stability region.

In the second part of the chapter, we extend our results to systems with an arbitrary finite number of queues. In particular, we show that the stability region characterization in terms of state-action frequencies and the throughput-optimality of the FBDC policy follow for the general case. Furthermore, we characterize an outer bound on the stability region and an upper bound on sum-throughput and show that a simple Myopic policy can achieve this sum-throughput upper-bound in the corresponding saturated (fully-backlogged) system.

Finally, in Chapter 6, we consider the most general setting and study the optimal scheduling problem for single-hop networks subjected to *time-varying channels*, *reconfiguration delays*, and *interference constraints*. We model the network by a graph consisting of nodes, links, and a set of link interference constraints, where based on the current network state, the controller decides either to stay with the current link-service configuration

or switch to another service configuration at the cost of idling during schedule reconfiguration.

We first consider the case of memoryless (i.i.d.) channel processes where we characterize the stability region in closed form as the convex hull of feasible activation vectors weighted by the *average* channel gain of each link. This result shows that, in the presence of reconfiguration delays, it is not possible to take advantage of the diversity in time-varying channels because the i.i.d. channel processes refresh during each reconfiguration interval. Moreover, we show that the class of Variable Frame-based Max-Weight (VFMW) algorithms considered in Chapter 4, adjusted to make scheduling decisions based on average channel gains, stabilizes the system by keeping the current schedule over a frame of duration that is a sublinear function of the queue lengths.

Next, we consider Markov modulated channel processes with memory. We generalize the state-action frequency framework of Chapter 5 to arbitrary single-hop networks and characterize the stability region of the system. We show that the stability region enlarges with the memory in the channel processes, which is in contrast to the case of no reconfiguration delays [50], [86], [110]. Furthermore, we extend the frame-based dynamic control (FBDC) policy to single-hop networks and show that it achieves the full stability region. To our knowledge, this is the first throughout-optimal scheduling algorithm for wireless networks with time-varying channels and reconfiguration delays.

Chapter 2

System Description and Assumptions

Throughout this thesis we will model wireless networks as queueing systems, where data packets that arrive at network nodes are to be transmitted to their destinations using the network medium. For the Delay Tolerant Network (DTN) model in Chapter 3, we consider mobile servers that are responsible for gathering messages that arrive randomly in time at randomly distributed geographical locations in the network region. The messages are transmitted when a collector is within their communication range and depart the system upon successful transmission. The queueing model in Chapter 3 is a continuous-time model and it is described in Chapter 3 in more detail. In the following, we describe the common features of the models used in chapters 4, 5, and 6, and leave details specific to each chapter to be explained in the corresponding chapter.

For chapters 4, 5, and 6, we consider the impact of reconfiguration delays on network models given by a graph structure $\mathcal{G}(\mathcal{N}, \mathcal{L})$ of nodes $\mathcal{N}$ and links $\ell \in \mathcal{L} \doteq \{1, 2, ..., L\}$, where $L \doteq |\mathcal{L}|$. Examples of networks that fall under the model studied in this thesis include single-hop wireless networks as shown in Fig. 2-1, satellite networks or uplink/downlinks as shown in Fig. 2-2, or input queued switches, and optical networks. [99]. For such systems, we consider discrete-time (slotted) systems, where the time index t takes

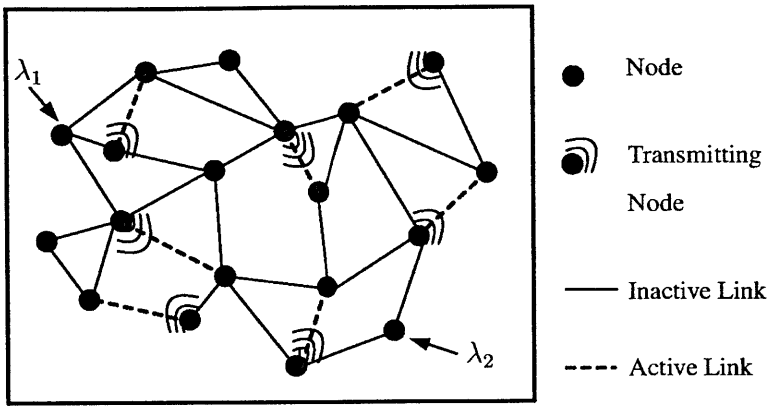


Figure 2-1: System model. A single-hop wireless network with interference constraints, time-varying channels and reconfiguration delays.

values in the set $\{0, 1, 2, ...\}$. Data packets arriving at each link ℓ are to be transmitted to their single-hop destinations, where we refer to the packets waiting for service at link ℓ as queue ℓ. An integer number of data packets can arrive at or depart from the queues corresponding to each link during each time slot. Let $A_\ell(t)$, denote the number of arrivals to queue ℓ at time slot t. We assume that the processes $A_\ell(t)$ takes nonnegative integer values, i.e., $A_\ell(t) \in \mathbb{N}_0$ where $\mathbb{N}_0 \doteq \{0\} \cup \mathbb{N}$. We also assume that the processes $A_\ell(t)$ are independent of each other and are i.i.d. over time with $\mathbb{P}\{A_\ell(t) = 0\} > 0$, $\mathbb{E}[A_\ell(t)^2] \leq (A_{\max})^2$, $\mathbb{E}[A_\ell(t)] = \lambda_\ell \leq A_{\max}, \forall \ell \in \mathcal{L}$. Let $\mathbf{Q}(t) = \{Q_1(t), ..., Q_L(t)\}$ denote the queue sizes at the links at the beginning of time slot t. Each link ℓ may be subject to a time-varying channel process denoted by $C_\ell(t)$ that takes values in a set $\mathcal{C} = \{0, \mu_{\min}, ..., \mu_{\max}\}$, where $C_\ell(t)$ corresponds to the number of packets that can be served from queue ℓ at time t. Note that $C_\ell(t) = 1$, $\forall \ell \in \mathcal{L}$, for the model in Chapter 4 where we consider time-invariant networks.

Let T_r denote the system reconfiguration delay, namely, it takes T_r time slots for the system to change a schedule, during which all the servers are necessarily idle. The set of

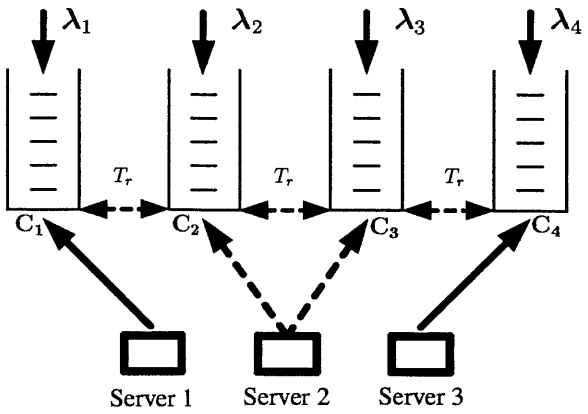


Figure 2-2: An example 4x3 satellite network. Ground stations are subject to time-varying channels $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$, $\mathbf{C}_4$ and the servers are subject to T_r slot reconfiguration (switchover) delay. Server 2 is forced to be idle due to interference constraints.

all schedules in the system, $\mathcal{I}$, is given by the set of *feasible* finite integer valued activation vectors $\mathbf{I} = (\mathbf{I}_\ell)_{\ell=1,\ldots,L}$. The set $\mathcal{I}$ is assumed to be arbitrary, except we assume that there is a uniform departure rate bound, $\mu_{\max}$, over all schedules and queues: $\mathbf{I}_\ell \leq \mu_{\max}$, $\forall \ell, \mathbf{I}$. If the activation vector $\mathbf{I}(t)$ is used at time slot t, then $\min\{C_\ell(t)\mathbf{I}_\ell(t), Q_\ell(t)\}$ packets depart from queue ℓ [1]. We include the vectors dominated by the feasible activation vectors, as well as the zero vector $\mathbf{I} = \mathbf{0}$ in $\mathcal{I}$, where the activation vector $\mathbf{I}(t)$ is equal to $\mathbf{0}$ for all time slots during which the system is undergoing reconfiguration. Let $\mathbb{H}(T) = [\mathbf{Q}(t)]|_{t=0}^{T} \cup [\mathbf{I}(t)]|_{t=0}^{T-1} \cup [\mathbf{C}(t)]|_{t=0}^{T}$ denote the full history of the system until time t and let $\Upsilon(\mathcal{I})$ denote the set of all probability distributions on the set of all actions $\mathcal{I}$. A control policy π is a mapping from $\mathbb{H}(t)$ to $\Upsilon(\mathcal{I})$ [86], [94].

The availability of a schedule is determined by the interference constraints in the system, which are assumed to be arbitrary. For instance, in a wireless mesh network as shown in Fig. 2-1, the set $\mathcal{I}$ can be determined according to the well-studied k-hop interference

[1] $\mathbf{I}(t)$ takes binary values for time-varying networks, and integer values for time-invariant networks (i.e., $C_\ell(t) = 1$, $\forall \ell, t$).

model [50]. Alternatively, for a satellite network of N queues and M servers where there are a possible $L = NM$ links as shown in Fig. 2-2, the set $\mathcal{I}$ can be the set of all binary vectors of dimension NM with at most M nonzero elements such that no two active servers interfere with each other [32]. For an NxN input-queued optical switch, the set $\mathcal{I}$ can be the set of all matchings [99]. Finally, for the single-server parallel queueing system considered in Chapter 5, $\mathcal{I}$ are unit vectors of length equal to the number of queues in the system. We assume that the system is initially empty and that the arrivals take place after the departures in any given time slot. Under this model, the queue sizes evolve according to the following expression.

$$Q_\ell(t+1) = \max\{Q_\ell(t) - \mathbf{I}_\ell(t)C_\ell(t), 0\} + A_\ell(t), \forall \ell \in \mathcal{L}. \tag{2.1}$$

Throughout this thesis we explore the impact of reconfiguration delays on stability and delay performance of networks.

Definition 1 (Stability [84], [86]) The system is stable if

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \sum_{\ell\in\mathcal{L}} \mathbb{E}[Q_\ell(\tau)] < \infty.$$

For the case of integer valued arrival and departure processes, this stability criterion implies the existence of a stationary distribution for the queue size Markov chain with bounded first moments [84].

Definition 2 (Stability Region [84], [86]) The stability region $\mathbf{\Lambda}$ is the closure of the set of all arrival rate vectors $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_L)$ such that there exists a control algorithm that stabilizes the system.

A policy is said to be throughput-optimal if it stabilizes the system for all input rates strictly inside $\mathbf{\Lambda}$. The δ-stripped stability region is defined for some $\delta > 0$ as $\mathbf{\Lambda}^{\delta} \doteq \left\{\boldsymbol{\lambda} | (\lambda_1 + \delta, ..., \lambda_L + \delta) \in \mathbf{\Lambda}\right\}$. A policy is said to achieve γ-fraction of $\mathbf{\Lambda}$, if it stabilizes the system for all input rates inside $\gamma\mathbf{\Lambda}$, where $\gamma = 1$ for a throughput-optimal policy.

Finally, throughout the thesis, we represent vectors, matrices, and sets with bold letters and we explicitly state when a matrix is utilized. We use the following notation for the inner product of two L-dimensional vectors: $\mathbf{u} \cdot \mathbf{v} \doteq \sum_{\ell=1}^{L} \mathbf{u}_\ell \mathbf{v}_\ell$.

Chapter 3

Data Collection via Mobile Servers in DTNs

There has been a significant amount of interest in performance analysis of DTNs or mobility assisted wireless networks in the last decade (e.g., [54], [78], [90], [100] [101], [117]). Typically, throughput and delay performance of networks have been analyzed where nodes moving according to a random mobility model are utilized for relaying data (e.g., [54], [48]). More recently, networks deploying nodes with controlled mobility have been considered focusing primarily on route design and ignoring the communication aspect of the problem (e.g., [27], [49], [59], [78], [101], [117]). In this chapter we explore the use of *controlled mobility* and *wireless transmission* in order to improve the throughput and delay performance of such networks. We consider a dynamic vehicle routing problem where vehicles (collectors) use a combination of physical movement and wireless reception to receive randomly arriving data messages.

Our model consists of collectors that are responsible for gathering messages that arrive randomly in time at uniformly distributed geographical locations. The messages are transmitted when a collector is within their communication range and depart the system upon

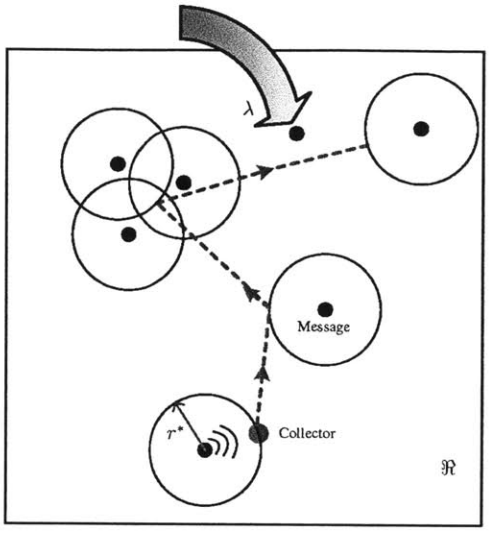


Figure 3-1: The system model for the case of a single collector. The collector adjusts its position in order to collect randomly arriving messages via wireless communication. The circles with radius r^* represent the communication range and the dashed line segments represent the collector's path.

successful transmission. Collectors adjust their positions in order to successfully receive these messages as shown in Fig. 3-1 for the case of a single collector. This setup is particularly applicable to networks, such as DTNs, deployed in a large area so that mobile elements are necessary to provide connectivity between spatially separated entities in the network [59], [78], [101]. Moreover, this model can be used to analyze the delay performance of a densely deployed sensor network where mobile base stations collect data from a large number of sensors inside the network [61], [101]. Our model is also applicable to a network where Unmanned Aerial Vehicles (UAVs) are used as data harvesting devices in a battlefield environment [90], [59], and to networks where data rate is relatively low so that data transmission time is comparable to the collector's travel time, for instance, in underwater sensor networks [4].

3.0.4 Related Work

Vehicle Routing Problems (VRPs) have been extensively studied in the past (e.g., [7], [12], [16], [46], [80], [107], [115], [116]). The common example of a VRP is the Euclidean Traveling Salesman Problem (TSP) in which a single server is to visit each member of a fixed set of locations on the plane such that the total travel cost is minimized. Several extensions of TSP have been considered in the past such as stochastic demand arrivals and the use of multiple servers [16], [17], [46], [117]. In particular, in the TSP with neighborhoods (TSPN) problem, a server is to visit a neighborhood of each demand location [12], [80], which can be used to model a mobile collector receiving messages from a communication distance. A more detailed review of the literature in this field can be found in [68], [80] and [117].

Of particular relevance to us among the VRPs is the Dynamic Traveling Repairman Problem (DTRP) due to Bertsimas and Van Ryzin [16], [17], [18]. DTRP is a stochastic and dynamic VRP in which a vehicle is to serve demands that arrive randomly in time and space. Fundamental lower bounds on delay were established and several vehicle routing policies were analyzed for DTRP for a single server in [16], for multiple servers in [17], and for general demand and interarrival time distributions in [18]. Later, [115], [116], [117] considered the Dynamic Pickup and Delivery Problem (DPDP) as a generalization of the DTRP problem, where fundamental lower bounds on delay were established and several dynamic pick up and delivery policies were analyzed. The wireless DPDP problem was considered for the first time in the last chapter of [117], where the simultaneous use of controlled mobility and wireless transmission was studied for a simple two-queue network. A communication model based on a physical layer perspective was utilized and the stability properties of the system as well as trade-offs between throughput and delay were studied.

This chapter considers a simpler wireless communication model as compared to the wireless DPDP problem in [117] and applies the DTRP model to wireless networks where the demands are data messages to be transmitted to a collector which can receive messages from a distance using wireless transmission[1].

Models similar to the DTRP problem have been considered in the past. Altman and Levy studied spatial polling models in two or higher-dimensional spaces in [6], [7], which are closely related to the work in this chapter. In [7], they considered the *Queuing in Space* model which is similar to the DTRP model, however, it is more general in that it allows stochastic arrivals to queue up at particular locations in the Euclidean plane. They showed that $\rho < 1$ is a necessary and sufficient condition for stability and that a greedy version of the cyclic and Globally Gated policy, which serves the nearest message among the messages in the current cycle, stabilizes the system. Similar to DTRP, the delay scaling with load is shown to be $\Theta(1/(1-\rho)^2)$. This model was later extended to general independent arrival processes with only first-moment conditions on service and walking times in [6]. Tassiulas considered a similar model in [107], where they proposed adaptive routing policies that achieve maximum throughput independent of the statistical parameters of the system. Similar to DTRP, in these works the collector must visit each message location without the possibility of wireless transmission. Continuous polling systems where wireless servers used as data relays on a one-dimensional space (such as a circle) were considered in [62] and [63], and expected waiting time and workload in the system were analyzed. In [61], the same authors considered the use of message ferries as data relays on periodic routes determined offline.

In a system where multiple mobile nodes with controlled mobility and communication

[1]We refer to the DTRP model as the system without wireless transmission since in this model the collector needs to be at the message location in order to be able to serve it.

capability relay the messages of static nodes, [100] derived a lower bound on node travel times. Message sources and destinations were modeled as static nodes in [100] and queuing aspects were not considered. A mobile server harvesting data from two spatial queues in a wireless network was considered in [90] where the stability region of the system was characterized using a fluid model approximation.

Another closely related body of literature lies in the area of utilizing mobile elements that can control their mobility to collect sensor data in Delay Tolerant Networks (DTN) (e.g., [27, 101, 128]). Route selection (e.g., [101]), scheduling or dynamic mobility control (e.g., [27], [128]) algorithms were proposed to maximize network lifetime, to provide connectivity or to minimize delay. These works focus primarily on mobility and usually consider particular policies for the mobile element. The work in this chapter is the first attempt to develop fundamental bounds on delay in a network where a collector is to gather data messages randomly arriving in time and space using *wireless communication* and *controlled mobility*.

3.0.5 Contributions and Outline

This work is the first attempt to develop fundamental lower and upper bounds on delay in a system where a collector is to gather data messages randomly arriving in time and space using *wireless communication* and *controlled mobility*. We first consider a system with a single collector and extend the results of [7] and [16] to the communication setting. In particular, we show that $\rho < 1$ is the necessary and sufficient stability condition where ρ is the system load. We derive lower bounds on delay and develop algorithms that are asymptotically within a constant factor of the lower bounds. We show that the combination of mobility and wireless transmission results in a delay scaling of $\Theta(1/(1-\rho))$. This is in

sharp contrast to the $\Theta(1/(1-\rho)^2)$ delay scaling in the system where the collector visits each message location analyzed in [7], [16].

Next, we consider the system with multiple collectors under the assumption that simultaneous transmissions to different collectors do not interfere with each other. We show that the necessary and sufficient stability condition is still given by $\rho < 1$, where ρ is the load on multiple collectors. We develop lower bounds on delay and generalize the single-collector policies, analyzed in the first part, to the multiple-collectors case. Finally, we study a multiple-collector system subject to interference constraints on simultaneous transmissions to different collectors. We consider a simplified grid mobility model and formulate a scheduling problem and characterize the stability region of the system in terms of interference constraints. We show that a frame-based version of the Max-Weight scheduling policy can stabilize the system whenever it is feasible to do so at all.

This chapter is organized as follows. In Section 3.1 we consider the single-collector case. We present the model in Section 3.1.1, and characterize the necessary and sufficient stability condition in Section 3.1.2. We derive a delay lower bound in Section 3.1.3, and analyze single-collector policies in Section 3.1.4. In Section 3.2, we extend the results for the single-collector to systems with multiple collectors whose transmissions do not interfere with each other. Finally, in Section 3.3 we present preliminary results for the system with interference constraints on simultaneous transmissions.

3.1 The Single Collector Case

In this section we consider the case of a single collector and develop fundamental insights into the problem. We extend the stability and the delay results of [7] and [16], established for the system where the collector visits each message location, to systems with wireless

transmission capability. We show that the combination of mobility and wireless transmission results in a delay scaling of $\Theta(\frac{1}{1-\rho})$ with the system load ρ, which is in contrast to the $\Theta(\frac{1}{(1-\rho)^2})$ delay scaling in the corresponding system without wireless transmission in [7] and [16].

3.1.1 Model

Consider a square region $\mathcal{R}$ of area A and messages arriving into $\mathcal{R}$ according to a Poisson process (in time) of intensity λ. Upon arrival the messages are distributed independently and uniformly in $\mathcal{R}$ and they are to be gathered by a collector via wireless reception. An arriving message is transmitted to the collector when the collector comes within the *communication range* of the message location and grants access for the message's transmission. Therefore, there is no interference power from the neighboring nodes during message receptions.

We assume a disk model [36], [55] for determining successful message receptions. Let r^* be the *communication range* of the collector. Under the disk model, a transmission can be received only if it is within a disk of radius r^* around the collector. Note that this model is similar to the Signal to Noise Ratio (SNR) packet reception model [36], [54], [55], under which a transmission is successfully decoded at the collector if its received SNR is above a threshold. Under this model, if the location of the next message to be received is within r^*, the collector stops and attempts to receive the message. Otherwise, the collector travels towards the message location until it is within a distance r^* from the message. Under the disk model, transmissions are assumed to be at a constant rate taking a fixed amount of time denoted by s.

The collector travels from the current message reception point to the next message re-

ception point at a constant speed v. We assume that at a given time the collector knows the locations and the arrival times of messages that arrived before this time. The knowledge of the service locations is a standard assumption in vehicle routing literature [7], [16], [46], [80]. Location information can be obtained from GPS devices or Inertial Measurement Units (IMUs), and distributed using a low-rate, but long-range, control channel. In the context of sensor networks, location information can also be obtained via distributed localization schemes using wide-band signalling [121].

Let $N(t)$ denote the total number of messages in the system at time t. The system is said to be *stable* under a given control policy π if the number of messages in the system $N(t)$ converges in distribution to a stationary process with a finite mean. Let $\rho = \lambda s$ denote the load arriving into the system per unit time. For stable systems, ρ denotes the fraction of time the collector spends receiving messages. *The stability region* $\mathbf{\Lambda}$ is the set of all loads ρ such that there exists a control algorithm that stabilizes the system. A policy is said to be *stabilizing* if it stabilizes the system for all loads strictly inside $\mathbf{\Lambda}$.

We define T_i as the time between the arrival of message i and its successful reception. T_i has three components: $W_{d,i}$, the waiting time due to collector's travel distance from the time message i arrives until it gets served, $W_{s,i}$, the waiting time due to the reception times of messages served from the time message i arrives until it gets served, and s, reception time of the message. The total waiting time of message i is given by $W_i = W_{d,i} + W_{s,i}$, hence $W_i = T_i - s$. The expected waiting time W is defined as $W \doteq \lim_{i\to\infty} \mathbb{E}[W_i]$ whenever the limit exists. T, W, W_d and W_s are defined similarly, and $T = W_d + W_s + s$, whenever the limits exist. Finally, T^* is defined to be the optimal system time which is given by the policy that minimizes T.

3.1.2 Stability

Next we show that $\rho < 1$ is a necessary and sufficient condition for stability of the system. Note that this condition is also necessary and sufficient for stability of the corresponding system without wireless transmission, as shown in [7], as well as for a G/G/1 queue [65]. We first lower bound the number of messages in the system by that in the equivalent system in which travel times are zero (i.e., $v = \infty$). This idea was used in [7] to establish a necessary stability condition for the corresponding system without wireless transmission.

Lemma 1 *A necessary condition for stability is* $\rho < 1$.

The proof of this lemma can be found in in Appendix A. It is based on an induction argument that the total number of messages in the system dominates that in the corresponding infinite-speed system, i.e., the *M/D/1* queue, for which the stability condition is $\rho < 1$.

Next, we show that $\rho < 1$ is a sufficient condition for stability of the system under a policy based on Euclidean TSP with neighborhoods (TSPN). TSPN is a generalization of TSP in which the server is to visit a neighborhood of each demand location via the shortest path [12], [80], for which polynomial-time $(1+\epsilon)$-approximation algorithms parameterized by $\epsilon > 0$ has been developed [80]. In our case, the neighborhoods are disks of radius r^* around each message location.

Under the TSPN policy, the collector performs a cyclic service of the messages present in the system starting and ending the cycle at the center of the network region. Let time t_k be the time that the collector returns to the center for the kth time, where $t_0 \doteq 0$. Suppose the system is initially empty at time t_0. The TSPN policy is described in detail in Algorithm 1. Let the total number of messages waiting for service at time t_k, $N(t_k)$, be the system state at time t_k. Note that $\{N(t_k)\};\ k \in \mathbb{N}$ is an irreducible Markov chain on countable state space $\mathbb{N}$.

Algorithm 1 *TSPN Policy*

1: Wait at the center of $\mathcal{R}$ until the first message arrival, move to serve this message and return to the center.

2: If the system is empty at time $t_k, k = 1, 2, ...$, repeat the above process.

3: If there are messages waiting for service at time $t_k, k = 1, 2, ...$, compute the TSPN tour through all the messages that are present in the system at time t_k, receive these messages in that tour and return to the center. Repeat 2 and 3.

Theorem 1 *The system is stable under the TSPN policy for all loads* $\rho < 1$.

The proof is given in Appendix B. It follows techniques similar to those of [7] and [24] to first establish a bound on the waiting time of an arbitrary message. Then the proof utilizes the stationary version of Little's law to establish the finiteness of the expected number of messages in the system.

Theorem 1 establishes that $\rho < 1$ is also sufficient for stability. The travel time does not affect the stability region of the system as expected. Note that for the analysis above, we assumed that the computation time of the TSPN tour is negligible as compared to the travel time of the collector. In a real-world scenario, having to wait for the computation can potentially affect the stability region. For instance, if the computation time takes ϵ-fraction of the expected cycle duration, then the TSPN policy cannot stabilize the system for arrival rates in the outer ϵ-strip of the stability region. Note that the Partitioning policy proposed in Section 3.1.4 is a simpler policy in that it does not require the knowledge of the message locations or arrival times, nor it needs to compute a tour for each cycle. The advantage of the TSPN policy is that it leads to shorter travel times in each cycle resulting in delay savings. Finally, simple greedy and cyclic policies based on receiving the closest message in the current cycle were considered in [7] and [71]. These policies do not need any tour

computation, however, the analysis of such policies in the context of wireless transmission does not appear to be tractable.

While the wireless transmission capability does not enlarge the stability region, it fundamentally affects the delay scaling in the system as we show in the next section.

3.1.3 Lower Bound On Delay

For wireless networks with a small area or very good channel quality such that $r^* \geq \sqrt{A/2}$, the collector can receive messages from the center of the network region. In that case we have an *M/D/1* queue and the associated queuing delay is given by the P-K formula as $W = \lambda s^2/(2(1-\rho))$. However, when $r^* < \sqrt{A/2}$, the collector has to move in order to receive some of the messages. In this case the reception time s is still a constant, however, the travel time per message is a random variable. Next we provide a delay lower bound, similar to a lower bound in [16], with the added complexity of communication capability in our system.

Theorem 2 *The optimal expected message waiting time in steady-state T^* is lower bounded by*

$$T^* \geq \frac{\mathbb{E}[\max(0, ||U|| - r^*)]}{v(1-\rho)} + \frac{\lambda s^2}{2(1-\rho)} + s. \tag{3.1}$$

where U is a random variable that has a uniform distribution over the network region $\mathcal{R}$, and $||U||$ is the distance of U to the center of $\mathcal{R}$.

Note that the $\mathbb{E}[\max(0, ||U|| - r^*)]$ term can be further lower bounded by $\mathbb{E}[||U||] - r^*$, where $\mathbb{E}[||U||] = 0.383\sqrt{A}$ [16]. **Proof:** As outlined in Section 5.1.1, the expected delay

of a message in steady state has three components:

$$T = W_d + W_s + s. \tag{3.2}$$

A lower bound on W_d is found as follows: Note that $W_d.v$ is the expected *distance* the collector moves during the waiting time of a message. This distance is at least as large as the average distance between the location of the message and the collector's location at the time of the message's arrival less the reception distance r^*. The location of an arrival is determined according to the uniform distribution over the network region, while the collector's location distribution is in general unknown as it depends on the collector's policy. We can lower bound W_d by characterizing the expected distance between a uniform arrival and the best a priori location in the network that minimizes the expected distance to a uniform arrival. Namely, we are after the location ν that minimizes $\mathbb{E}[||U - \nu||]$ where U is a uniformly distributed random variable. The location ν that solves this optimization is called the *median* of the region and in our case the median is the center of the square shaped network region. Thus, we obtain the following bound:

$$W_d \geq \frac{\mathbb{E}[\max(0, ||U|| - r^*)]}{v}. \tag{3.3}$$

Let N be the expected number of messages served in a waiting time and let R be the average residual service time. Due to the PASTA property of Poisson arrivals [13, p. 171] a given arrival in steady state observes the steady state occupancy distribution. Therefore, the average residual time observed by an arrival is also R, and it is given by $\lambda s^2/2$, which gives [16]

$$W_s = sN + R. \tag{3.4}$$

When the system is stable and in steady state, the expected number of messages served during a waiting time is equal to the expected number of arrivals during a waiting time, which in turn is equal to the expected number of messages in the system in steady state [16]. To see this, note that since the future arrivals are independent of the current number of messages in the system under Poisson arrivals, the steady state occupancy distribution observed by a Poisson arrival is the same as the time-stationary distribution of the number of messages in the system [13, pp. 172]. Furthermore, since the messages are served one at a time, every state $N(t) = n$ is visited infinitely often in a stable system, and the steady state occupancy distribution observed by a departing customer is also equal to the occupancy distribution observed by an arriving customer [47, pp. 173]. Therefore, the expected number of messages served during a waiting time, N, must be equal to the expected number of messages that arrive during a waiting time, which is also equal to the expected number of messages in the system. Finally, the last quantity is given by the steady state version of Little's law to be $N = \lambda W = \lambda(W_d + W_s)$ [16], [117]. Substituting this in (3.4) we obtain

$$W_s = s\lambda(W_d + W_s) + \frac{\lambda s^2}{2}.$$

This implies

$$W_s = \frac{\rho}{1-\rho} W_d + \frac{\lambda s^2}{2(1-\rho)}. \tag{3.5}$$

Substituting (3.3) and (3.5) in (3.2) yields (3.1). □

Theorem 2 shows that, in addition to the expected waiting time of an *M/G/1* queue $\lambda s^2/(2(1-\rho))$, the queueing delay has another component dependent on the travel time of the collector.

3.1.4 Collector Policies

We derive upper bounds on expected delay by analyzing policies for the collector. From (3.19) it can be shown that the expected number of messages in steady state under the TSPN policy is $O(1/(1-\rho))$. Threfore, the TSPN policy has optimal delay scaling. We consider the First Come First Serve (FCFS) and the Partitioning policies that are much simpler than the TSPN policy, and have good delay properties. In particular, the FCFS policy is delay-optimal at light loads and the Partitioning policy has delay performance that is very close to the lower bound when the travel and reception times are comparable.

First Come First Serve (FCFS) Policy

A straightforward policy is the FCFS policy where the messages are served in the order of their arrival times. A version of the FCFS policy, call FCFS′, where the collector has to return to the center of the network region after each message reception was shown to be delay optimal at light loads for the DTRP problem [16], i.e., $T_{\text{FCFS}'} \to T^*$ as $\rho \to 0$. This is because the center of the network region is the location that minimizes the expected distance to a uniformly distributed arrival. Since in our system we can do at least as well as the DTRP by setting $r^* = 0$, FCFS′ is delay optimal also for our system at light loads. Furthermore, the FCFS′ policy is not stable for all loads $\rho < 1$, namely, there exists a value $\hat{\rho}$ such that the system is unstable under FCFS′ policy for all $\rho > \hat{\rho}$. This is because under the FCFS′ policy, the average per-message travel component of the service time is fixed, which makes the average arrival rate greater than the average service rate as ρ increases. Therefore, it is better for a policy to serve more messages in the same "neighborhood" in order to reduce the amount of time spent on mobility.

Partitioning Policy

Next we propose a policy based on partitioning the network region into subregions and the collector performing a cyclic service of the subregions. This policy is an adaptation of the Partitioning policy of [16] to the case of a system with wireless transmission and it implements a cyclic polling discipline with exhaustive service. An adaptive version of the Partitioning policy of [16] was considered in [107], where the subregion sizes are determined adaptively based on the number of service demands in the system, and the system is stabilized without requiring the knowledge of the arrival rates. We explicitly derive the delay expression for this policy and show that it scales with the load as $O(\frac{1}{1-\rho})$.

We divide the network region into ($\sqrt{2}r^*$ x $\sqrt{2}r^*$) squares as shown in Fig. 3-2. This choice ensures that every location in the square is within the communication distance r^* of the center of the square. The number of subregions in such a partitioning is given by[2] $n_s = A/(2(r^*)^2)$. The partitioning in Fig. 3-2 represents the case of $n_s = 16$ subregions. The collector services the subregions in a cyclic order, as shown in Fig. 3-2, by receiving the messages in each subregion from its center using an FCFS order. The messages within each subregion are served exhaustively, i.e., all the messages in a subregion are served before moving to the next subregion. The collector then serves the messages in the next subregion exhaustively using FCFS order and repeats this process. The distance traveled by the collector between each subregion is a constant equal to $\sqrt{2}r^*$. It is easy to verify that the Partitioning policy behaves as a multiuser *M/G/1* system with reservations and exhaustive service (see [13, p. 198]), where the n_s subregions correspond to users and the travel time

[2]Note that such a partitioning requires $\sqrt{n_s} = \sqrt{A/(2(r^*)^2)}$ to be an integer. This may not hold for a given area A and a particular choice of r^*. In that case one can partition the region using the largest reception distance $\underline{r}^* < r^*$ such that this integer condition is satisfied.

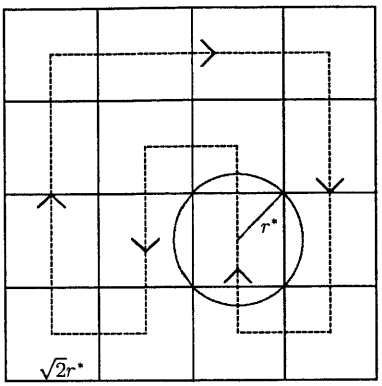


Figure 3-2: The partitioning of the network region into square subregions of side $\sqrt{2}r^*$. The circle with radius r^* represents the communication range and the dashed lines represent the collector's path.

between the subregions corresponds to the reservation interval. Using the delay expression for the multiuser *M/G/1* queue with reservations and exhaustive service in [13, p. 200] we obtain,

$$T_{part} = \frac{\lambda s^2}{2(1-\rho)} + \frac{n_s - \rho}{2v(1-\rho)}\sqrt{2}r^* + s. \tag{3.6}$$

Combining this result with (3.1) and noting that the above expression is finite for all loads $\rho < 1$, we have established the following observation.

Observation 1 *The expected message waiting time in steady state scales as* $\Theta(\frac{1}{1-\rho})$ *with the load* ρ*, and the Partitioning policy is stable for all* $\rho < 1$.

Despite the travel component of the service time, we can achieve $\Theta(\frac{1}{1-\rho})$ delay as in the *M/G/1* queue. This is the fundamental difference between this system and the corresponding system where wireless transmission is not used, where the delay scaling with load is $\Theta(\frac{1}{(1-\rho)^2})$ [7], [16]. This difference can be explained intuitively as follows. Denote by N the average number of departures in a waiting time. The W_s expression as a function of W_d in (3.5) implies that N can be lower bounded by $\frac{\lambda W_d}{1-\rho}$. For the system in [16], the minimum per-message distance traveled in the high load regime scales as $\Omega(\frac{\sqrt{A}}{\sqrt{N}})$ [16]. This is due to

the fact that the nearest neighbor distance among N uniformly distributed points on a square of area A is $\Omega(\frac{\sqrt{A}}{\sqrt{N}})$. Therefore, for this system we have $W_d \approx N\Omega(\frac{\sqrt{A}}{\sqrt{N}}) \approx \Omega(\sqrt{NA})$ which gives $N \approx \Omega(\frac{\lambda^2 A}{(1-\rho)^2})$. In contrast, with the wireless reception capability, the collector does not need to move for messages that are inside a disk of radius r^* around it. Since a finite (constant) number of such disks cover the network region, W_d can be upper bounded by a constant independent of the system load.

It is interesting to note that [39] considered the case where messages were transmitted to the collector according to a random access scheme, i.e., transmissions occur with probability p in each time slot. There the delay scaling of $\Omega(\frac{1}{(1-\rho)^2})$ was observed, which is similar to the system without wireless transmission. The reason for this is that in order to have successful transmissions under the random access interference of neighboring nodes, the reception distance should be of the same order as the nearest neighbor distances [39], [54].

Numerical Results-Single Collector

We present numerical results corresponding to the analysis in the previous sections. We lower bound the delay expression in (3.1) using $\mathbb{E}[\max(0, ||U|| - r^*)] \geq \mathbb{E}[||U||] - r^*$, where $[||U||] = 0.383\sqrt{A}$ is the expected distance of a uniform arrival to the center of square region of area A [16]. Fig. 3-3 shows the delay lower bound as a function of the network load for increased values of the communication range r^* [3]. As the communication range increases, the message delay decreases as expected. For heavy loads, the delay in the system is significantly less than the delay in the corresponding system without wireless transmission in [16], demonstrating the difference in the delay scaling between the two

[3]For the delay plot of the system without wireless transmission, the point that is not smooth arises since the plot is the maximum of two delay lower bounds proposed in [16].

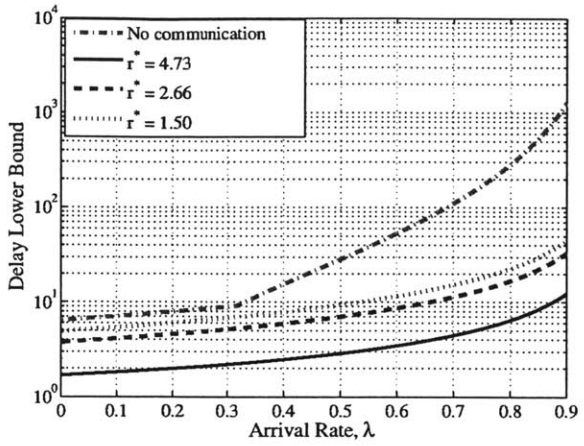


Figure 3-3: Delay lower bound vs. network load using different communication ranges for $A = 200$, $v = 1$ and $s = 1$.

systems. For light loads and small communication ranges, the delay performance of the wireless network tends to the delay performance of [16].

Fig. 3-4 compares the delay in the Partitioning Policy to the delay lower bound for two different cases. When the travel time dominates the reception time, the delay in the Partitioning policy is about 10 times the delay lower bound. For a more balanced case, i.e., when the reception time is comparable to the travel time, the delay ratio drops to 2.4.

3.2 Multiple collectors - Interference-Free Networks

The analysis in the previous section can be extended to a system with $m > 1$ identical collectors that do not interfere with one another. This can be done, for example, by partitioning the network region into m subregions and performing independent single-collector policies within each subregion. We call the class of such policies the *network partitioning policies*. For this class of policies, the interference-free assumption is satisfied if transmissions in different subregions use different frequency bands.

The main difference in analysis as compared to the single-collector case is that we

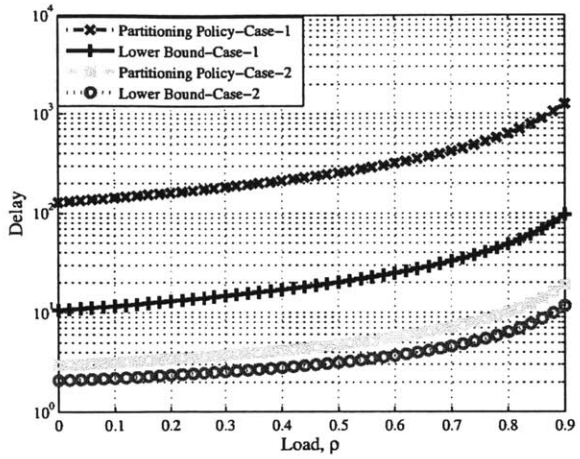


Figure 3-4: Delay in the Partitioning policy vs the delay lower bound for $r^* = 2.2$, $\beta = 2$ and $\alpha = 4$. Case-1: Dominant travel time ($A = 800$, $v = 1$, $s = 2$). Case-2: Comparable travel and reception times ($A = 60$, $v = 10$, $s = 2$).

utilize the m-median problem in order to bound the travel times of the collectors and a load balancing argument in order to derive a delay lower bound. We show that similar to the single-collector case, the stability region is the set of loads such that $\rho < 1$ and the delay scaling is $\Theta(1/(1-\rho))$, where $\rho = \lambda s/m$.

A necessary condition for stability of the multi-collector system is given by $\rho = \lambda s/m < 1$. This is because the number of messages in the system considered stochastically dominates that in the corresponding system with zero travel times (i.e., an M/D/m queue, a queue with Poisson arrivals, constant service time and m servers) similar to Section 3.1.2.

Lemma 2 A necessary condition for the stability is $\rho = \lambda s/m < 1$. Furthermore, the optimal steady state time average delay T_m^* is lower bounded by

$$T_m^* \geq \frac{\lambda s^2}{2m^2(1-\rho)} - \frac{m-1}{m}\frac{s^2}{2s} + s, \tag{3.7}$$

where $\rho = \lambda s/m$ is the system load.

The proof of this lemma can be found in Appendix C. For showing the sufficiency of the stability condition $\rho < 1$, we can generalize the single-collector TSPN policy to the case of multiple collectors through network partitioning, and obtain that the system under the generalized TSPN policy is stable for all loads $\rho < 1$.

Next, we derive a lower bound on W_d, the average waiting time due to the collectors' travel, using a result from [56] for the m-median problem.

Lemma 3

$$W_d \geq \frac{\max\left(0, \frac{2}{3}\sqrt{\frac{A}{m\pi}} - r^*\right)}{v}. \tag{3.8}$$

Proof: Let Ω be *any* set of points in $\Re$ with $|\Omega| = m$. Let U be a uniformly distributed location in $\Re$ independent of Ω and define $Z^* \triangleq \min_{\nu \in \Omega} \parallel U - \nu \parallel$. Let the random variable Y be the distance from the center of a disk of area A/m to a uniformly distributed point within the disk. It is shown in [56] that,

$$\mathbb{E}[f(Z^*)] \geq \mathbb{E}[f(Y)] \tag{3.9}$$

for any nondecreasing function $f(.)$. Using this result we obtain $\mathbb{E}[\max(0, Z^* - r^*)] \geq \mathbb{E}[\max(0, Y - r^*)]$. Note that W_d can be lower bounded by the expected distance of an arrival whose location is uniformly distributed on the network region to the closest collector at the time of arrival less r^*. Because the travel distance is nonnegative, we have

$$W_d \geq \mathbb{E}[\max(0, Y - r^*)]/v \geq \max(0, \mathbb{E}[Y] - r^*)/v,$$

where the second bound is due to Jensen's inequality. Substituting $\mathbb{E}[Y] = \frac{2}{3}\sqrt{\frac{A}{m\pi}}$ into the above expression completes the proof. □

Intuitively the best a priori placement of m points in $\Re$ in order to minimize the distance of a uniformly distributed point in the region to the closest of these points is to cover the region with m disjoint disks of area A/m and place the points at the centers of the disks. Such a partitioning of the region is not possible, however, using this idea we can lower bound the expected distance as in (3.9).

When W_d is lower bounded by a constant, a simple convexity argument shows that the equal area partitioning of the network region minimizes the resulting delay expression over all area partitionings. Using this result, and similar steps to the proof of Theorem 2, yields the following lower bound on average delay for the class of network partitioning policies:

Lemma 4 For the class of network partitioning policies, the optimal steady state time average delay T_m^* is lower bounded by

$$T^* \geq \frac{\max\left(0, \frac{2}{3}\sqrt{\frac{A}{m\pi}} - r^*\right)}{v(1-\rho)} + s. \tag{3.10}$$

where $\rho = \lambda s/m$ is the system load.

This lemma is proved in Appendix D. This lower bound is more useful than (3.7) since it takes travel time into account.

Finally, by partitioning the region into m subregions and then applying the single-collector Partitioning Policy in each subregion shows that the average delay of this generalized Partitioning Policy is given by the average delay of the single-collector Partitioning Policy applied to a system with arrival rate λ/m and area A/m:

$$T_{part} = \frac{\lambda s^2}{2m(1-\rho)} + \frac{\frac{A}{2m(r^*)^2} - \rho}{2v(1-\rho)}\sqrt{2}r^* + s. \tag{3.11}$$

From (3.10) and (3.11), we have that the delay scaling in the system is $\Theta(\frac{1}{1-\rho})$, in contrast to the $\Theta(\frac{1}{(1-\rho)^2})$ delay scaling for the multi-collector DTRP [17].

3.3 Multiple Collectors - Systems with Interference Constraints

In this section we consider systems in which simultaneous transmissions to different collectors interfere with each other. The problem is to dynamically determine a subset of collectors to route and schedule for transmission based on the present collector configuration and the number of messages in the system. The objective is to minimize the expected message waiting time in the system. This is a joint scheduling and euclidian vehicle routing problem which has not been considered previously. Here we obtain preliminary results for this problem by emphasizing the scheduling aspect through fairly general interference constraints and simplifying the mobility aspect by discretizing the collectors' motion. We characterize the stability region of the system in terms of interference constraints and show that a frame-based version of the Max-Weight scheduling policy [25], [109], can stabilize the system whenever it is feasible to do so at all.

First we explain the mobility and the interference models. We assume that time is slotted, $t = 0, 1, 2, ...$, where the slot length is equal to one message transmission time s. The collectors are confined to move on a grid $\mathcal{G}$ of $(\frac{r^*}{\sqrt{2}} \mathrm{x} \frac{r^*}{\sqrt{2}})$ squares, i.e., $K \doteq \frac{A}{(r^*)^2/2}$ square cells of diameter r^* as shown in Fig. 3-5. Assuming a fixed ordering of the K cells, where each cell $i = 1, 2, ..., K$, receives Poisson arrivals with rate $\lambda_i \doteq \lambda/K$. Let $A_i(t)$ denote the number of messages that arrive into cell i at time slot t. The expected *load* entering cell i per time slot is given by $\rho_i = \lambda_i s$. Let $N_i(t)$ be the number of messages in cell i at the beginning of time slot t, let $\mathbf{N}(t) \doteq [N_1(t), ..., N_K(t)]$ denote the vector of

queue sizes, and let $N(t) \doteq N_1(t) + ... + N_K(t)$ denote the total number of messages in the system at time slot t.

Definition 3 (Cell Interference Model) Given a collector that is at the common corner of multiple adjacent cells, a transmission to the collector from one of the cells is successfully received if there is no other transmission within any adjacent cell.

The Cell Interference Model essentially creates an *exclusion region* of up to 4 cells around a collector receiving a message. Similar interference models have been considered in the literature. For example, the Protocol Model considered in [55], [100], assumes successful transmission if a disk region around the receiver has no other transmission. We characterize the interference constraints of the system in terms of *activation vectors*. We call a cell *active* if at least 1 message in the cell is scheduled for transmission, and we assume that each cell $k = 1, 2, ..., K$ is associated with exactly one pick up location on the grid $\mathcal{G}$ [4]. For instance, the pick up location for each cell could be the upper left corner of the cell as shown in Fig. 3-5. Therefore, specifying the set of cells to activate also specifies the locations of the collectors. A feasible activation vector $\mathbf{I} \in \mathcal{I}$ is one under which transmissions from a set of active cells do not interfere with each other, where $\mathcal{I}$ is the set of all feasible activation vectors. The set $\mathcal{I}$ consists of K-dimensional vectors of at most m nonzero entries, where $\mathbf{I}_k = 1$ if cell k is active under $\mathbf{I}$, and $\mathbf{I}_k = 0$ otherwise. Note that we include the zero vector $\mathbf{I} = \mathbf{0}$ in $\mathcal{I}$ for convenience.

Let T_r denote the maximum *reconfiguration time*, i.e., the number of time slots required for a collector to move from the lower right corner of $\mathcal{G}$ to it's upper left corner. The *corresponding system with infinite speed*, i.e., $T_r = 0$, is a parallel queueing

[4] Of course such an assumption may reduce the stability region. Here we make this assumption in order to present preliminary results for the general problem.

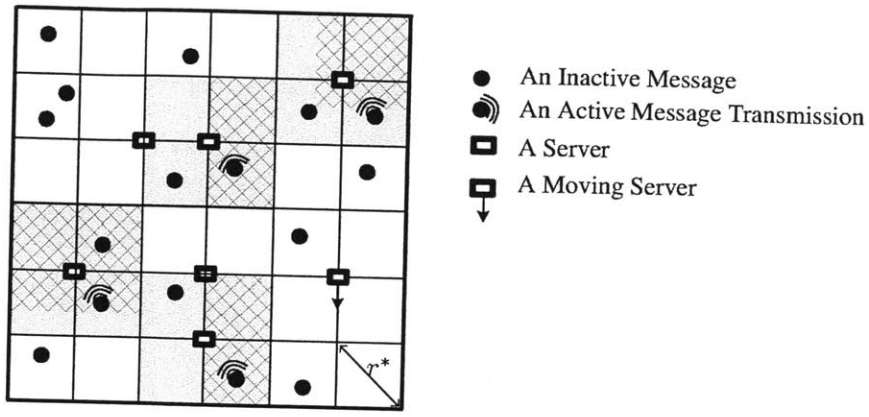


Figure 3-5: The network model. Red regions are the exclusion zones for the servers currently in service. Two servers are forced to be inactive since the messages in their vicinity are in the exclusion zones of other servers.

system with multiple servers and interference constraints, which is a special case of [86] or [109]. When $T_r = 0$, the stability region of this system, $\mathbf{\Lambda}^0$, consists of all load vectors $\boldsymbol{\rho} = [\rho_1, \rho_2, ..., \rho_K] = s[\lambda_1, \lambda_2, ..., \lambda_K]$ in the convex hull of $\mathcal{I}$ [25]:

$$\mathbf{\Lambda}^0 = \{\boldsymbol{\rho} | \boldsymbol{\rho} \in Conv\{\mathcal{I}\}\}. \tag{3.12}$$

When $T_r > 0$, we have a significantly different system for which previously proposed algorithms are not stabilizing. Since we lose service opportunities during the reconfiguration times we have

$$\mathbf{\Lambda} \subseteq \mathbf{\Lambda}^0. \tag{3.13}$$

We will show that $\mathbf{\Lambda} = \mathbf{\Lambda}^0$. The celebrated Max-Weight scheduling algorithm was introduced in [109] and was shown to stabilize the system for all $\boldsymbol{\rho} \in \mathbf{\Lambda}^0$ when $T_r = 0$. Specifically, the Max-Weight policy activates the set of users in $\mathbf{I}^*(t)$ where

$$\mathbf{I}^*(t) = \arg\max_{\mathbf{I} \in \mathcal{I}} \mathbf{N}(t) \cdot \mathbf{I}, \tag{3.14}$$

where $\mathbf{a} \cdot \mathbf{b} \doteq a_1 b_1 + ... + a_K b_K$.

3.3.1 Framed-Max-Weight Policy

For systems with nonzero reconfiguration times, the Max-Weight policy is not stabilizing [25]. The intuitive reason behind this is that the Max-Weight policy makes frequent reconfiguration decisions, resulting in throughput loss during reconfiguration intervals. A frame based version of the Max-Weight policy where the same schedule is used throughout the frame incurs less throughput loss. A similar frame-based approach was used in [25] in the context of optical networks. We show that the Framed-Max-Weight (FMW) policy defined below stabilizes the system considered in this section. We prove this result using a quadratic Lyapunov drift technique.

Under the FMW policy, time is divided into intervals of length T slots. The FMW policy employs the activation vector corresponding to the Max-Weight configuration at the beginning of each frame for $T - T_r$ slots, where the first T_r slots of each frame are reserved for the servers to travel to their assigned locations. The policy requires $T > T_r/\epsilon(\rho)$ where $\epsilon(\rho)$ is determined by solving the Linear Program below [25].

$$\begin{aligned} \epsilon(\boldsymbol{\rho}) &\triangleq \max\left(1 - \sum_{\mathbf{I}\in\mathcal{I}} \alpha_{\mathbf{I}}\right) \\ \text{subject to} \quad \rho_i &= \lambda_i s \leq \sum_{\mathbf{I}\in\mathcal{I}} \alpha_{\mathbf{I}} \mathbf{I}_i, \; i \in \{1, ..., K\} \\ \sum_{\mathbf{I}\in\mathcal{I}} \alpha_{\mathbf{I}} &\leq 1, \text{ and } \alpha_{\mathbf{I}} \geq 0, \; \forall \mathbf{I} \in \mathcal{I}. \end{aligned} \tag{3.15}$$

Note that $\epsilon(\rho)$ is a measure of distance of the load vector to the boundary of the stability region [25]. The FMW policy is described in Algorithm 2 below.

Algorithm 2 *Framed-Max-Weight Policy*

1: Assuming the system is at the jth frame, find

$$\mathbf{I}^* = \arg\max_{\mathbf{I}\in\mathcal{I}} \mathbf{N}(jT)\cdot\mathbf{I}. \tag{3.16}$$

2: Reconfigure the collectors to their new locations for the next T_r slots.

3: Apply the activation vector $\mathbf{I}^*$ for $T - T_r$ slots.

Theorem 3 *For any $\boldsymbol{\rho} = [\rho_1, ..., \rho_K]$ strictly inside $\boldsymbol{\Lambda_0}$, the FMW policy stabilizes the system as long as $T > T_r/\epsilon(\rho)$.*

The proof is given in Appendix E. The reason that the FMW policy stabilizes the system is that as the load approaches the boundary of the stability region, the policy employs maximum-weight schedules over longer frames, decreasing the fraction of time spent on reconfiguration. The proof in Appendix E is based on a quadratic Lyapunov drift argument over frames of duration T. The proof establishes that the T-step expected drift of the queue lengths satisfies

$$\mathbb{E}\left[L(\mathbf{N}(jT+T)) - L(\mathbf{N}(jT))\middle|\mathbf{N}(jT)\right] \leq KBT^2 - \frac{2T}{K}(\epsilon - \frac{T_r}{T})\sum_i N_i(jT), \tag{3.17}$$

where $B = 1 + \frac{\lambda s}{K} + \frac{\lambda^2 s^2}{K^2}$ is a constant. From (3.17) we see that the drift becomes negative when the queue size is sufficiently large. Stability follows from this condition, similar to the proof of Theorem 1. Combining Theorem 3 with (3.13), we have the following Corollary.

Corollary 1

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda_0}.$$

Similar to the case of collectors whose transmissions do not interfere, the stability region is not affected by the collector travel times.

3.4 Concluding Remarks

In this chapter we considered the use of dynamic vehicle routing in order to improve the throughput and delay performance of wireless networks where messages arriving randomly in time and space are gathered by mobile collectors via wireless communications. For the case of a single collector, we characterized the stability region of this system. We developed a fundamental lower bound on expected message waiting time as well as matching upper bounds. For the case of multiple collectors whose communications do not interfere with each other, we extended the stability and delay scaling results of the single collector case. Our results show that combining controlled mobility and wireless transmission results in $\Theta(\frac{1}{1-\rho})$ delay scaling with load ρ. This is the fundamental difference between our system and the system without wireless transmission (DTRP) analyzed in [16] and [17] where the delay scaling with the load is $\Theta(\frac{1}{(1-\rho)^2})$. Finally, for the the case where simultaneous transmissions to different collectors interfere with each other, we formulated a scheduling problem and characterized the stability region of the system in terms of interference constraints. We show that a frame-based version of the Max-Weight policy is stabilizing asymptotically in the frame length.

We have utilized a simple wireless communication model based on a communication range. Possible future directions include more sophisticated communication and interference models that take into account the signal to interference and noise ratio (SINR).

Appendix A - Proof of Lemma 1

We first show that the unfinished work and the delay experienced by a message in the system stochastically dominates that in the equivalent system with zero travel times for the

collector.

Lemma 5 The steady state time average delay in the system is at least as big as the delay in the equivalent system in which travel times are considered to be zero (i.e., $v = \infty$).

Proof: Consider the summation of per-message reception and travel times, s and d_i, as the total service requirement of a message in each system. Since d_i is zero for all i in the infinite velocity system and since the reception times are constant equal to s for both systems, the total service requirement of each message in our system is deterministically greater than that of the same message in the infinite velocity system. Let $D_1, D_2, ..., D_n$ and $D_1^{'}, D_2^{'}, ..., D_n^{'}$ be the departure instants of the 1^{st}, 2^{nd} and similarly the n^{th} message in the original and the infinite velocity system respectively. Similarly let $A_1, A_2, ..., A_n$ be the arrival times of the 1^{st}, 2^{nd} and the n^{th} message in both systems. We will use induction to prove that $D_i \geq D_i^{'}$ for all i. Since the service requirement of each message is smaller in the infinite velocity system, we have $D_1 \geq D_1^{'}$. Assuming we have $D_n \geq D_n^{'}$, we need to show that $D_{n+1} \geq D_{n+1}^{'}$.

$$A_{n+1} \leq D_{n+1} - s, \tag{3.18}$$

hence the $n + 1^{th}$ message is available before the time $D_{n+1} - s$. We also have

$$D_n^{'} \leq D_n \leq D_{n+1} - s.$$

The first inequality is due to the induction hypothesis and the second inequality is because we need at least s amount of time between the n^{th} and $n + 1^{th}$ transmissions. Hence there is at least one collector available in the infinite velocity system before the time $D_{n+1} - s$. Combining this with (3.18) proves the induction. Now let $D(t)$ and $D^{'}(t)$ be the total number of departures by time t in our system and the infinite velocity system respectively.

Similarly let $N(t)$ and $N'(t)$ be the total number of messages in the two systems at time t. Finally let $A(t)$ be the total number of arrivals by time t in both systems. We have $A(t) = N(t) - D(t) = N'(t) - D'(t)$. From the above induction we have $D(t) \leq D'(t)$ and therefore

$$N(t) \geq N'(t).$$

Since this is true at all times, we have that the time average number of customers in the system is greater than that in the infinite velocity system. Finally using Little's law proves the lemma. □

Since in our case the infinite velocity system behaves as an M/G/1 queue (an M/G/1 queue is a queue with Poisson arrivals, general i.i.d. service times and 1 server and an M/D/1 queue has constant service times), the average waiting time in this system is given by the Pollaczek-Khinchin (P-K) formula for M/G/1 queues [13, p. 189], i.e., $\lambda s^2/(2(1-\lambda s))$. Furthermore, a direct consequence of this lemma is that a necessary condition for stability in the infinite speed system is also necessary for our system. It is well-known that the necessary (and sufficient) condition for stability in the M/G/1 systems is given by $\rho < 1$ (see e.g., [13] or [50]).

Appendix B - Proof of Theorem 1

Proof: Let $N(t)$ denote the number of messages in the system at time t, and let W_j denote the delay experienced by the jth message. Recall the definition of time t_k, $k \geq 1$, the time at which the collector returns to the center of the network region for the kth time, where $t_0 \doteq 0$. Let $N(t_k)$ denote the total number of messages waiting for service at time t_k. We will denote $N(t_k)$ by N_k for notational simplicity. The duration of time between t_{k-1} and t_k is called the kth cycle, and is denoted by C_k, $k \in \mathbb{Z}_+$. Note that $\{N_k : \ k \in \mathbb{N}\}$ is

an irreducible Markov chain on countable state space $\mathbb{N}$, termed the Cycle Markov chain. Given the system state N_k at time t_k, we find the TSPN tour of length L_k through the N_k neighborhoods[5]. We prove Theorem 1 by establishing the following properties:

1. We first prove that the discrete-time Markov chain $\{N_k\}$ is positive recurrent and has a steady state distribution with a finite first moment.

2. Using this steady state distribution, we derive bounds on the first and second moments of the cycle duration, as well as the residual and past cycle durations under the TSPN policy.

3. Next, we show that the message delays $\{W_j : j \in \mathbb{Z}_+\}$ and the queue length process $\{N(t) : t \geq 0\}$ form positive recurrent regenerative processes, and therefore, converge in distribution to stationary processes.

4. Using the bounds on the residual and the past cycle times, we show that the stationary process of message delays has a finite expectation.

5. Finally, we utilize the stationary version of Little's law to show that the stationary process of number of messages in the system has finite expectation.

Cycle Markov Chain $\{N_k\}$

First, we will use the Foster-Lyapunov criterion to show that the Markov chain described by the states N_k is positive recurrent. We use the linear Lyapunov function $V(N_k) = sN_k$, the total load served during the kth cycle. Note that $V(0) = 0$, $S_k = \{x : V(x) \leq B\}$

[5] L_k can be upper bounded by a constant L for all N_k. This is because the collector does not have to move for messages within its communication range, and a finite number of such disks can cover the network region for any $r^* > 0$. The collector then can serve the messages in each disk from its center incurring a tour of constant length L. An example of such a tour is shown in Fig. 3-2.

is a bounded set for all finite B and $V(.)$ is a non-decreasing function. Since the arrival process is Poisson, the expected number of arrivals during a cycle can be upper bounded as follows:

$$\mathbb{E}[N_{k+1}|N_k] \leq \lambda(L/v + sN_k). \tag{3.19}$$

Hence, we obtain the following drift expression for the load during a cycle.

$$\mathbb{E}[sN_{k+1} - sN_k|N_k] \leq \rho L/v - (1-\rho)sN_k. \tag{3.20}$$

Since $\rho < 1$, there exist a $\delta > 0$ such that $\rho + \delta < 1$:

$$\mathbb{E}[sN_{k+1} - sN_k|N_k] \leq \rho L/v - \delta sN_k.$$

Fix $\epsilon \in (0, \delta)$. A simple derivation shows that when N_k is outside the finite and bounded set $S = \{N \in \mathbb{N} : N \leq \frac{\rho L/v + \epsilon}{s(\delta - \epsilon)}\}$ the drift expression is given by

$$\mathbb{E}[sN_{k+1} - sN_k|N_k] \leq -\epsilon(1 + sN_k).$$

For $N_k \in S$, using $\rho < 1 - \delta$ and the definition of the set S, we have from (3.19),

$$\mathbb{E}[sN_{k+1}|N_k] \leq \rho L/v + (1-\delta)sN_k < \rho L/v + \frac{(1-\delta)(\rho L/v + \epsilon)}{(\delta - \epsilon)} < \infty.$$

Moreover, since the state space is countable, the set S is finite, and since the states in the Markov chain $\{N_k\}$ have nonzero probability of self transition, the Markov chain is strongly aperiodic. Therefore, all the conditions of Lemma 4.2 in [7] are satisfied (by the choice of the function $g(N_k) = 1 + sN_k$), and we have that the Markov chain $\{N_k\}$ is positive (Harris) recurrent, N_k has a steady state distribution, where we let the random

variable N^c denote this steady state distribution. Moreover, $\mathbb{E}[N_k]$ converges to $\mathbb{E}[N^c]$, and the expected number of messages in steady state, $\mathbb{E}[N^c]$, is finite [7].

Moments of Cycle Duration

Next, we derive bounds on the first and second moments of the cycle duration, and the expected residual and past cycle durations. These bounds will be necessary in order to obtain an upper bound for the expected message delay and the number of messages in the system. We will prove the finiteness of the expected number of messages in the system by first establishing that the expected message delay in the system is finite, and then utilizing the stationary version of Little's law. The analysis in this section is similar to that in [7]. Let C_k denote the duration of the kth cycle, $C_k = sN_k + L_k/v$. The location distributions of messages in different cycles are independent and uniformly distributed and the TSPN policy obtains the travel paths, L_k, using a stationary algorithm [80]. Therefore, C_k is a function of N_k and the location distribution of these N_k messages. Note that the lengths of the travel paths L_k are uniformly bounded from above by L for all $k \in \mathbb{Z}_+$. Let $\mathbb{E}^0$ denote expectation at the time corresponding to the beginning of a cycle, in steady state. We let N^c and C denote the steady state versions of N_k and C_k. Taking the expectation of (3.19) with respect to the steady state distribution at the beginning of the cycles we have

$$\mathbb{E}^0[N^c] \leq \frac{\lambda L}{v} + \rho \mathbb{E}^0[N^c],$$

which implies that

$$\mathbb{E}^0[N^c] \leq \frac{\lambda L}{v(1-\rho)}. \tag{3.21}$$

Using the bound on the cycle time $C_k = sN_k + L_k/v \leq sN_k + L/v$ we have[6].

$$\mathbb{E}^0[C] \leq \frac{\rho L}{v(1-\rho)} + \frac{L}{v}. \tag{3.22}$$

In order to lower bound the expected cycle duration, we lower bound the expected travel distance per cycle. This distance is at least as large as the expected distance between a uniformly distributed point (message location) in the network region and the center of the region less r^*. For a square shaped region of area A, this distance can be lower bounded by $\underline{d} \doteq 0.383\sqrt{A} - r^*$ [16]. Therefore, we have

$$\mathbb{E}[N_{k+1}|N_k] \geq \lambda(\underline{d}/v + sN_k).$$

Upon taking expectations we have

$$\mathbb{E}^0[N^c] \geq \frac{\lambda \underline{d}}{v(1-\rho)}, \tag{3.23}$$

and

$$\mathbb{E}^0[C] \geq \frac{\underline{d}}{v(1-\rho)}. \tag{3.24}$$

Next, we characterize the second moment of the cycle duration. Let T_s denote the time it takes to serve Poisson arrivals arriving in a time interval of random duration D. If the interarrival times, service times s, and the duration of time D are independent, the second

[6]Note that letting $A(t_1, t_2)$ denote the number of Poisson arrivals in the time interval (t_1, t_2), we have $A(t_1, t_2) = A(t_2 - t_1)$, and $N_{k+1} = A(C_k)$. Taking expectations gives $\mathbb{E}[N_{k+1}] = \lambda \mathbb{E}[C_k]$. Finally, taking the limit as $k \to \infty$ yields $\mathbb{E}^0[C] = \mathbb{E}^0[N^c]/\lambda$, which gives the same relationship as (3.22).

moment of T_s is given by [64, pp. 238] or [7, pp. 1107]

$$\mathbb{E}[T_s^2] = \lambda^2 \mathbb{E}[s]^2 \mathbb{E}[D^2] + \lambda \mathbb{E}[s^2]\mathbb{E}[D].$$

This result can be applied to the workload in our system, where the workload for the $(k+1)$th cycle in our system is given by sN_{k+1}, and the random duration of interest is the kth cycle duration C_k. The reason we can use the result from [64] in our system is that the duration of the kth cycle is a function of the arrivals in the previous cycle, and it is independent of the interarrival times during the kth cycle. Therefore,

$$\mathbb{E}[s^2 N_{k+1}^2] = \lambda^2 s^2 \mathbb{E}[C_k^2] + \lambda s^2 \mathbb{E}[C_k]. \tag{3.25}$$

Thus, we have

$$\begin{aligned}
\mathbb{E}^0[s^2(N^c)^2] &\leq \rho^2 \mathbb{E}^0[(sN^c + L/v)^2] + \lambda s^2 \left(\frac{\rho L}{v(1-\rho)} + \frac{L}{v}\right) \\
&\leq \rho^2 \left(\mathbb{E}^0[s^2(N^c)^2] + \frac{2sL}{v}\mathbb{E}^0[N^c] + \frac{L^2}{v^2}\right) + \lambda s^2 \left(\frac{\rho L}{v(1-\rho)} + \frac{L}{v}\right),
\end{aligned}$$

which upon utilizing the upper bounds on the first moments of N^c and C in (3.21) and (3.22) gives

$$\begin{aligned}
\mathbb{E}^0[s^2(N^c)^2] &\leq \frac{\frac{2\rho^3 L^2}{v^2(1-\rho)} + \frac{\rho^2 L^2}{v^2} + \frac{\rho^2 sL}{v(1-\rho)} + \frac{\rho sL}{v}}{1-\rho^2} \\
&= \frac{\frac{\rho^2 L^2}{v^2}\left(\frac{1+\rho}{1-\rho}\right) + \frac{\rho sL}{v}\left(\frac{1}{1-\rho}\right)}{1-\rho^2} \doteq \overline{N^c},
\end{aligned}$$

where we let $\overline{N^c}$ denote the finite constant on the right hand side. Using the bound on $\mathbb{E}^0[(N^c)^2]$, we can upper bound the second moment of the cycle duration easily as follows:

$$\begin{aligned}\mathbb{E}^0[C^2] &\leq \mathbb{E}^0[(sN^c + L/v)^2] = \mathbb{E}^0[s^2(N^c)^2 + 2sLN^c/v + L^2/v^2] \\ &\leq \overline{N^c} + \frac{2sL}{v}\frac{\lambda L}{v(1-\rho)} + \frac{L^2}{v^2}. \end{aligned} \tag{3.26}$$

Expected Waiting Time

Next, we bound the expected waiting time in order to bound the expected number of messages in the system via Little's law. For this, we first establish that the delay process $\{W_j : j \in \mathbb{Z}_+\}$, and the queue length process $\{N(t) : t \geq 0\}$ converge to stationary processes.

Lemma 6 *The processes* $\{W_j : j \in \mathbb{Z}_+\}$ *and* $\{N(t) : t \geq 0\}$ *form positive recurrent regenerative processes under the TSPN policy.*

The proof is given at the end of the proof of Theorem 1. It establishes that the times when an arrival finds an empty system with the collector at the center of the network region constitute regeneration epochs for the system. Because the regeneration processes $\{N(t) : t \geq 0\}$ and $\{W_j : j \in \mathbb{Z}_+\}$ are positive recurrent and their regeneration periods are aperiodic, the sequences of message delays converge in distribution to a (customer)-stationary process, denoted by $\tilde{W}$, and the queue length process $\{N(t) : t \geq 0\}$ converge in distribution to a time-stationary process, denoted by $\tilde{N}$, see [102]. Now, we derive a bound on the expected waiting time, $\mathbb{E}[\tilde{W}]$, according to the stationary delay distribution. This bound is derived in a similar way to [7] or [24]. The delay of an arbitrary message is upper bounded by the sum of the residual cycle time C_R, plus the duration of the next cycle C_N. Note that the cycle during which the arrival occurs is a-typical and has expected

duration $\mathbb{E}[C_R] + \mathbb{E}[C_P]$, where $\mathbb{E}[C_R]$ and $\mathbb{E}[C_P]$ denote the expected residual and past cycle times and are given by $\mathbb{E}^0[C^2]/2\mathbb{E}^0[C]$ [7], [24]. Therefore,

$$\mathbb{E}[\tilde{W}] \leq \mathbb{E}[C_R] + \mathbb{E}[C_N] \leq \frac{\mathbb{E}^0[C^2]}{2\mathbb{E}^0[C]} + \mathbb{E}[C_N]. \tag{3.27}$$

Note that C_N is also a-typical and equal to the sum of the travel time plus the amount of workload that arrived during the previous cycle. Therefore, we have [7],

$$\mathbb{E}[C_N] \leq \rho(\mathbb{E}[C_P] + \mathbb{E}[C_R]) + \frac{L}{v}. \tag{3.28}$$

Finally, combining (3.28) with the expression for the expected residual time, $\mathbb{E}[C_R] = \mathbb{E}^0[C^2]/2\mathbb{E}^0[C]$, we have from (3.27),

$$\mathbb{E}[\tilde{W}] \leq (\rho + \frac{1}{2})\frac{\mathbb{E}^0[C^2]}{2\mathbb{E}^0[C]} + \frac{L}{v} < \infty$$

where the last inequality holds due to (3.24) and (3.26),

Finally, the stationary version of Little's law gives a relationship between the first moment of the time-stationary process $\tilde{N}$, and the first moment of the customer-stationary process $\tilde{W}$: We have

$$\mathbb{E}[\tilde{N}] = \lambda\mathbb{E}[\tilde{W}] < \infty.$$

This establishes the stability of the TSPN policy for any load $\rho < 1$. □

Lemma 6 Proof: Let the arrival time of the jth message be $\tilde{t}_j$, and its delay W_j. We consider the Markov chain $\{N_k : k \in \mathbb{N}\}$ at the beginning of cycles which is positive recurrent, and therefore, hits the empty state infinitely often. Consecutive epochs and times

at which an arrival finds the collector at the center of an empty system (i.e., start of a cycle) constitute an embedded renewal process for both processes $\{W_j : j \in \mathbb{Z}_+\}$, and $\{N(t) : t \geq 0\}$. Namely, let the sequence $\{\ell_n : n \in \mathbb{Z}_+\}$ denote the sequence of arrivals that find an empty system with the collector at the center. Because the arrival and the service processes are stationary, the discrete sequence $\{\ell_n : n \in \mathbb{Z}_+\}$ serve as an embedded renewal process for the delay process $\{W_n : n \in \mathbb{Z}_+\}$, and the continuous times $\tilde{t}_{\ell_n}$ serve as one for the queue length process $\{N(t) : t \geq 0\}$. More precisely, we have that the process $\{N(\tilde{t}_{\ell_1} + t) : t \geq 0\}$ is independent of $\{N(t) : t < \tilde{t}_{\ell_1}\}$ and of $\tilde{t}_{\ell_1}$, and the process $\{N(\tilde{t}_{\ell_1} + t) : t \geq 0\}$ is stochastically identical to $\{N(t) : t \geq 0\}$. Similarly, the process $\{W_{\ell_1+n} : n \in \mathbb{Z}_+\}$ is independent of $\{W_n : n < \ell_1\}$ and of ℓ_1, and the process $\{W_{\ell_1+n} : n \in \mathbb{Z}_+\}$ is stochastically identical to $\{W_n : n \in \mathbb{Z}_+\}$, see [102].

Next, we show that these renewal processes are positive recurrent. Namely, we show that the expectation of the interrenewal periods, $\{\tilde{t}_{\ell_n} - \tilde{t}_{\ell_{n-1}} : n \in \mathbb{Z}_+\}$, are finite. Let T_r be the duration of the rth renewal period, where the sequence $\{T_r : r \in \mathbb{Z}_+\}$ is i.i.d., and we need to show that $\mathbb{E}[T_1] < \infty$. Let m_0 be the mean recurrence time of the empty state (i.e., the state $N_k = 0$) in the Markov chain $\{N_k\}$, which is finite since the Markov chain is positive recurrent. Note that m_0 also denotes the expected number of cycles between renewals. Given K let $M(K)$ be the number of renewals that have taken place up to and including cycle K. Since the last renewal might have taken place before cycle K, we have

$$\frac{\sum_{k=1}^{K} C_k}{\sum_{r=1}^{M(K)} T_r} \geq 1. \tag{3.29}$$

Furthermore, we have from Strong Law of Large Numbers (SLLN)

$$\lim_{K\to\infty} \frac{M(K)}{K} = \frac{1}{m_0}, \ \ a.s. \tag{3.30}$$

The extended version of the Strong Law of Large Numbers (SLLN) for nonnegative valued random variables states that if the expectation of the random variables involved is infinite, then their average converges to infinity, see for example [97, pp. 370]. Now, applying the extended version of the SLLN to T_r we have,

$$\lim_{K\to\infty} \frac{1}{M(K)} \sum_{r=1}^{M(K)} T_r = \mathbb{E}[T_1], \ a.s. \tag{3.31}$$

Note that we will establish that the above expectation is indeed finite. We utilize the upper bound on the cycle times $C_k \leq sN_k + L/v$ in (3.29) to have

$$\frac{\sum_{k=1}^{K} \left(sN_k + \frac{L}{v}\right)}{\sum_{r=1}^{M(K)} T_r} \geq 1. \tag{3.32}$$

Since the Markov chain $\{N_k : \ k \in \mathbb{N}\}$ is ergodic, we have

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} N_k = \mathbb{E}^0[N^c], \ \ a.s. \tag{3.33}$$

Finally, rewriting (3.32), taking the limit as K tends to infinity, and applying (3.30), (3.31),

and (3.33), we have

$$\lim_{K\to\infty} \frac{K}{M(K)} \frac{\frac{1}{K}\sum_{k=1}^{K} sN_k + \frac{L}{v}}{\frac{1}{M(K)}\sum_{r=1}^{M(K)} T_r} = m_0 \frac{s\mathbb{E}^0[N^c] + \frac{L}{v}}{\mathbb{E}[T_1]} \geq 1,$$

which implies that

$$\mathbb{E}[T_1] \leq m_0\Big(s\mathbb{E}^0[N^c] + \frac{L}{v}\Big) = m_0\Big(\frac{\rho L}{v(1-\rho)} + \frac{L}{v}\Big) < \infty,$$

where we used (3.21) for the last inequality. This establishes the fact that the regenerative processes $\{W_j : j \in \mathbb{Z}_+\}$ and $\{N(t) : t \geq 0\}$ are positive recurrent. □

Appendix C - Proof of Lemma 2

Proof: The proof is similar to the proof of Lemma 1 in Appendix A. First consider the following lemma.

Lemma 7 The steady state time average delay in the system is at least as big as the delay in the equivalent system in which travel times are considered to be zero (i.e., $v = \infty$).

Proof: The proof is similar to the proof of Lemma 5. Consider the summation of per-message reception and travel times, s and d_i, as the total service requirement of a message in each system. Since d_i is zero for all i in the infinite velocity system and since the reception times are constant equal to s for both systems, the total service requirement of each message in our system is deterministically greater than that of the same message in the infinite velocity system. Let $D_1, D_2, ..., D_n$ be the departure instants of the first, second and similarly the n^{th} message in the system. Similarly let $D_1', D_2', ..., D_n'$ be the departure instants of the first, second and similarly the n^{th} message in the infinite velocity system.

Similarly let $A_1, A_2, ..., A_n$ be the arrival times of the first second and similarly the n^{th} message in both systems. We will use *complete induction* to prove that $D_i \geq D_i'$ for all i. Since the service requirement of each message is less in the infinite velocity system, we have $D_1 \geq D_1'$. Assume we have $D_i \geq D_i'$ for all $i \leq n$. We need to show that $D_{n+1} \geq D_{n+1}'$ in order to complete the complete induction. We have

$$A_{n+1} \leq D_{n+1} - s, \tag{3.34}$$

hence the $n + 1^{th}$ message is available at time $D_{n+1} - s$. We also have

$$D_{n+1-m}' \leq D_{n+1-m} \leq D_{n+1} - s.$$

The first inequality is due to the complete induction hypothesis and the second inequality is due the fact that the m^{th} last departure before the $n + 1^{th}$ departure has to occur before the time $D_{n+1} - s$. Hence there is at least one collector available in the infinite velocity system before the time $D_{n+1} - s$. Combining this with (3.34) proves the complete induction. Now let $D(t)$ and $D'(t)$ be the total number of departures by time t in our system and the infinite velocity system respectively. Similarly let $N(t)$ and $N'(t)$ be the total number of messages in the two systems at time t. Finally let $A(t)$ be the total number of arrivals by time t in both systems. We have $N(t) = A(t) - D(t)$ and $N'(t) = A(t) - D'(t)$. From the above induction we have $D(t) \leq D'(t)$ and therefore $N(t) \geq N'(t)$. Since this is true at all times, we have that the time average number of customers in the system is greater than that in the infinite velocity system. Finally using Little's law proves the lemma. □

When the travel time is considered to be zero, the system becomes an M/D/m queue (a queue with Poisson arrivals, constant service time and m servers). Therefore we can bound

T_m^* using bounds for general G/G/m systems. In particular, the waiting time $W_{G/G/m}$ in a G/G/m queue with service time s is bounded below by [65, p. 48]

$$W_{G/G/m} \geq \hat{W} - \frac{m-1}{m}\frac{s^2}{2s}, \tag{3.35}$$

where $\hat{W}$ is the waiting time in a single server system with the same arrivals as in the G/G/m queue and service time s/m. Since in our case the infinite velocity system behaves as an M/D/m system, $\hat{W}$ has an exact expression given by the P-K formula: $\hat{W} = \lambda s^2/(2m^2(1-\rho))$ where $\rho = \lambda s/m$. Substituting this in (3.35) and using Lemma 7 we have (3.7). □

Appendix D - Proof of Lemma 4

Proof: Here we use an approach similar to the proof of Theorem 2. We divide the average delay T into three components:

$$T = W_d + W_s + s. \tag{3.36}$$

We utilize the lower bound proposed in Lemma 3 for W_d. We now derive a lower bound on W_s. Let $R^1, R^2, ..., R^m$ be the network partitioning with areas $A^1, A^2, ..., A^m$ respectively ($\sum_{j=1}^m A^j = A$). Consider the message receptions in steady state that are received by collector j eventually. Let λ^j be the fraction of the arrival rate served by collector j. Due to the uniform distribution of the message locations we have

$$\frac{\lambda^j}{\lambda} = \frac{A^j}{A}.$$

Let N^j be the average number of message receptions for which the messages that are served by collector j waits in steady state. Similarly let W_s^j and W_d^j be the average waiting times for messages served by collector j due to the time spent on message receptions and collector j's travel respectively. Using (3.4) and lower bounding the residual time by zero we have

$$W_s^j \geq sN^j.$$

Using Little's law ($N^j = \lambda^j(W_s^j + W_d^j)$) similar to the derivation of (3.5) we have

$$W_s^j \geq \frac{\lambda^j s}{1-\lambda^j s} W_d^j \tag{3.37}$$

The fraction of messages served by collector j is A^j/A. Therefore, we can write W_s as

$$\begin{aligned} W_s &= \sum_{j=1}^{m} \frac{A^j}{A} W_s^j \\ &\geq \sum_{j=1}^{m} \frac{A^j}{A} \frac{\lambda^j s}{1-\lambda^j s} W_d^j. \end{aligned} \tag{3.38}$$

For a given region R^j with area A^j, W_d^j is lower bounded by (similar to the derivation of (3.3)) the distance of a uniform arrival to the median of the region less r^*.

$$\begin{aligned} W_d^j &\geq \frac{\mathrm{E}[\max(0, ||U-\nu|| - r^*)]}{v} \\ &\geq \frac{\max(0, \mathrm{E}[||U-\nu||] - r^*)}{v}, \end{aligned} \tag{3.39}$$

where ν is the median of R^j and $||U - \nu||$ is the distance of U, a uniformly distributed location inside R^j, to ν. The inequality in (3.39) is due to Jensen's inequality for convex functions. A disk shaped region yields the minimum expected distance of a uniform arrival to the median of the region. Using this we further lower bound W_d by noting that for a disk shaped region of area A_j, E$[||U - \nu||]$ is just the expected distance of a uniform arrival to the center of the disk given by $\frac{2}{3}\sqrt{\frac{A^j}{\pi}}$. Hence

$$W_d^j \geq \frac{\max(0, \frac{2}{3}\sqrt{\frac{A^j}{\pi}} - r^*)}{v} = \frac{\max(0, c_1\sqrt{A^j} - r^*)}{v}, \tag{3.40}$$

where $c_1 = \frac{2}{3\sqrt{\pi}} = 0.376$. Letting $f(A^j) = \frac{\lambda \frac{A^j}{A} s}{1 - \lambda \frac{A^j}{A} s}$, which is a convex and increasing function of A^j, we rewrite (3.38) as

$$W_s \geq \sum_{j=1}^{m} \frac{f(A^j)}{vA} A^j \max(0, c_1\sqrt{A^j} - r^*). \tag{3.41}$$

Next we will show that the function $f(A^j)A^j\max(0, c_1\sqrt{A^j} - r^*)$ is a convex function of A^j via the two lemmas below.

Lemma 8 Let $f(.)$ and $g(.)$ be two convex and increasing functions (possibly nonlinear) defined on $[0, A]$. The function $h(.) = f.g(.)$ is also convex and increasing on its domain $[0, A]$x$[0, A]$.

Proof: Clearly h is increasing. Let x and y be two points in the domain of h and let $\alpha \in (0,1)$ be a real number.

$$\begin{aligned} h(\alpha x + (1-\alpha)y) &= f(\alpha x + (1-\alpha)y)g(\alpha x + (1-\alpha)y) \\ &\leq (\alpha f(x) + (1-\alpha)f(y)). \\ &\quad .(\alpha g(x) + (1-\alpha)g(y)) \\ &= \alpha^2 f(x)g(x) + (1-\alpha)^2 f(y)g(y) \\ &\quad +\alpha(1-\alpha)f(x)g(y) + \alpha(1-\alpha)f(y)g(x)), \end{aligned}$$

where the inequality is due to the convexity of f and g. We add and subtract $\alpha f(x)g(x)$ and after some algebra obtain

$$\begin{aligned} h(\alpha x + (1-\alpha)y) &\leq \alpha h(x) + (1-\alpha)h(y) \\ &\quad +\alpha(1-\alpha)(f(x) - f(y))(g(y) - g(x)) \\ &\leq \alpha h(x) + (1-\alpha)h(y), \end{aligned}$$

where the last inequality is due to the fact that f and g are increasing functions. □

Lemma 9 $h(x) = x\max(0, c_1\sqrt{x} - c_2)$ is a convex and increasing function of x.

Proof: It is clear that $h(x)$ is an increasing function of x. Let x and y be two points in the domain of h and let $\alpha \in (0,1)$ be a real number.

$$
\begin{aligned}
h(\alpha x+(1-\alpha)y) = & \\
& = (\alpha x + (1-\alpha)y)\max(0, c_1\sqrt{\alpha x + (1-\alpha)y} - c_2) \\
& = \max(0, c_1(\alpha x + (1-\alpha)y)^{\frac{3}{2}} - c_2(\alpha x + (1-\alpha)y)) \\
& \leq \max\big(0, c_1(\alpha x^{\frac{3}{2}} + (1-\alpha)y^{\frac{3}{2}}) - c_2(\alpha x + (1-\alpha)y)\big) \\
& = \max\big(0, \alpha x(c_1\sqrt{x} - c_2) + (1-\alpha)y(c_1\sqrt{y} - c_2)\big) \\
& \leq \max\big(0, \alpha x(c_1\sqrt{x} - c_2)\big) + \max\big(0, (1-\alpha)y(c_1\sqrt{y} - c_2)\big) \\
& = \alpha h(x) + (1-\alpha)h(y),
\end{aligned}
$$

where the first inequality is due to the convexity of the function $x^{\frac{3}{2}}$. □

Letting $g(A^j) \doteq f(A^j)A^j\max(0, c_1\sqrt{A^j} - r^*)$, we have from the lemmas 8 and 9 that the function $g(A^j)$ is convex. Now rewriting (3.41) we have

$$
W_s \geq (\frac{m}{vA})\frac{1}{m}\sum_{j=1}^{m} g(A^j).
$$

Using the convexity of the function $g(A^j)$ we have

$$
\begin{aligned}
W_s \;&\geq\; (\frac{m}{vA})g\Big(\frac{\sum_{j=1}^{m} A^j}{m}\Big) = \frac{m}{vA}g(\frac{A}{m}) \\
&= \frac{\frac{\lambda s}{m}}{1-\frac{\lambda s}{m}}\frac{\max(0, c_1\sqrt{A/m} - r^*)}{v} = \frac{\rho}{1-\rho}\frac{\max\big(c_1\sqrt{\frac{A}{m}} - r^*\big)}{v}. \qquad (3.42)
\end{aligned}
$$

The above analysis essentially implies that the W_s expression in (3.41) is minimized by the *equitable partitioning* of the network region. Finally combining (3.36), (3.8) and (3.42) we obtain (3.10). □

Appendix E - Proof of Theorem 3

We prove Theorem 3 for a broader class of arrival processes. We assume that each cell i has an arrival process $A_i(t)$ that is i.i.d. over time and satisfies $\mathbb{E}[A_i(t)^2] \leq A_{\max}^2$ independent of the number of messages in the system, which is satisfied if the overall arrival process into the system is Poisson. Note that we have $\mathbb{E}[A_i(t)] = \lambda_i s$ independent of the number of messages in the system. Let t_k, $k = 0, 1, ...,$ be the first time slot of the kth frame. Let $D_i(t)$, $t \in \{t_k + T_r, t_{k+1} - 1\}$, be 1 if cell i is scheduled to be active during the kth frame and zero otherwise. Note that $D_i(t)$ is the *service opportunity* given to cell i at time slot t and not the actual departure process. Let $N_i(t)$ be the number of messages in cell i at the beginning of time slot t. We assume that arrivals take place at the end of time slots. We have the following queue evolution relation.

$$N_i(t+1) = \max\{N_i(t) - D_i(t), 0\} + A_i(t).$$

Similarly, the following T-step queue evolution expression holds:

$$N_i(t_k + T) \leq \max\left\{N_i(t_k) - \sum_{\tau=0}^{T-1} D_i(t_k + \tau), 0\right\} + \sum_{\tau=0}^{T-1} A_i(t_k + \tau).$$

The inequality is due to the fact that cell i might become empty and that some arrivals depart during the frame. Squaring both sides we have,

$$\begin{aligned}(N_i(t_k+T))^2-(N_i(t_k))^2 &\leq \Big(\sum_{\tau=0}^{T-1}D_i(t_k+\tau)\Big)^2+\Big(\sum_{\tau=0}^{T-1}A_i(t_k+\tau)\Big)^2 \\ &- 2N_i(t_k)\Big(\sum_{\tau=0}^{T-1}D_i(t_k+\tau)-\sum_{\tau=0}^{T-1}A_i(t_k+\tau)\Big). \quad (3.43)\end{aligned}$$

Define the quadratic Lyapunov function

$$L(\mathbf{N}(t_k))=\sum_{i=1}^{K}N_i^2(t_k),$$

and the T-step conditional Lyapunov drift

$$\Delta_T(t_k)\triangleq\mathbb{E}\left\{L(\mathbf{N}(t_k+T))-L(\mathbf{N}(t_k))\big|\mathbf{N}(t_k)\right\}.$$

Summing (3.43) over the queues, taking conditional expectation, using $D_i(t)\leq 1$ for all time slots t, $\mathbb{E}\{A_i(t)^2\}\leq A_{\max}^2$ and $\mathbb{E}\{A_i(t_1)A_i(t_2)\}\leq\sqrt{\mathbb{E}\{A_i(t_1)\}^2\mathbb{E}\{A_i(t_2)\}^2}\leq A_{\max}^2$ for all t_1 and t_2 we have

$$\begin{aligned}\Delta_T(t_k) &\leq KBT^2+2\mathbb{E}\left\{\sum_i N_i(t_k)\sum_{\tau=0}^{T-1}\left[A_i(t_k+\tau)-D_i(t_k+\tau)\right]\big|\mathbf{N}(t_k)\right\} \\ &= KBT^2+2T\sum_i N_i(t_k)\lambda_i s-2\sum_i N_i(t_k)\mathbb{E}\left\{\sum_{\tau=0}^{T-1}D_i(t_k+\tau)\big|\mathbf{N}(t_k)\right\}\end{aligned}$$

where $B = 1 + A_{\max}^2$ is a constant. Note that $D_i(t+\tau) = 0, \forall i \in \{1, ..., K\}$ for $\tau \in \{0, 1, ..., T_r - 1\}$ since the system is idle for the first T_r slots of the frame under the FMW policy. Therefore,

$$\Delta_T(t_k) \leq NBT^2 + 2T \sum_i N_i(t_k)\lambda_i s - 2 \sum_i \sum_{\tau=T_r}^{T-1} N_i(t_k)\mathbb{E}\left\{D_i(t_k+\tau)\middle|\mathbf{N}(t_k)\right\}$$

Now using the fact that for any load vector $\boldsymbol{\rho} = \boldsymbol{\lambda} s$ that is strictly inside $\boldsymbol{\Lambda}^0$, there exist real numbers $\alpha_1, ..., \alpha_{|\mathcal{I}|}$ such that $\alpha_j > 0, \forall j \in 1, ..., |\mathcal{I}|$, $\sum_{j=1}^{|\mathcal{I}|} \alpha_j = 1 - \epsilon$ for some $\epsilon > 0$ and

$$\boldsymbol{\rho} = \sum_{j=1}^{|\mathcal{I}|} \alpha_j \mathbf{I}^j,$$

where $\mathbf{I}^j$ is a K-dimensional vector in $\mathcal{I}$. Over the time interval $[t + T_r, t + T - 1]$, the FMW policy applies the activation vector that has the property

$$\mathbf{I}^*(t_k) = \arg\max_{\mathbf{I} \in \mathcal{I}} \mathbf{N}(t_k) \cdot \mathbf{I}. \tag{3.44}$$

Therefore $\sum_i N_i(t_k) D_i(t_k + \tau) = \mathbf{N}(t_k) \cdot \mathbf{I}^*(t_k)$. Hence, we have

$$\begin{aligned}
\Delta_T(t_k) &\leq KBT^2 + 2T\mathbf{N}(t_k).\left(\sum_{j=1}^{|\mathcal{I}|}\alpha_j\mathbf{I}^j\right) - 2T(1-\frac{T_r}{T})\mathbf{N}(t_k).\mathbf{I}^*(t_k) \\
&= KBT^2 - 2T\sum_{j=1}^{|\mathcal{I}|}\alpha_j(\mathbf{N}(t_k).\mathbf{I}^*(t_k) - \mathbf{N}(t_k).\mathbf{I}^j) \\
&\quad - 2T(1-\sum_{j=1}^{|\mathcal{I}|}\alpha_j)\mathbf{N}(t_k).\mathbf{I}^*(t_k) + 2T_r\mathbf{N}(t_k).\mathbf{I}^*(t_k) \\
&\leq KBT^2 - 2T\epsilon\mathbf{N}(t_k).\mathbf{I}^*(t_k) + 2T_r\mathbf{N}(t_k).\mathbf{I}^*(t_k) \\
&= KBT^2 - 2T(\epsilon - \frac{T_r}{T})\mathbf{N}(t_k).\mathbf{I}^*(t_k). \qquad (3.45)
\end{aligned}$$

Note that we have $\mathbf{N}(t_k).\mathbf{I}^*(t_k) \geq \frac{1}{K}\sum_i N_i(t_k)$ since the maximum weight schedule has more weight than the average. Therefore, for $T > \frac{T_r}{\epsilon}$ we have

$$\Delta_T(t_k) \leq KBT^2 - 2T(\epsilon - \frac{T_r}{T})\frac{1}{K}\sum_i N_i(t_k). \qquad (3.46)$$

Therefore, the T-step conditional Lyapunov drift is negative if $T > \frac{T_r}{\epsilon}$ and if the queue sizes are outside a bounded set. Therefore, the stability at the frame boundaries follows from Lemma 4.2 in [7] due to a similar reasoning to the proof of Theorem 1. This implies the stability of the system since the frame length T is a constant.

Chapter 4

Scheduling in Queueing Networks with Reconfiguration Delays

In the previous chapter, we first studied the use of controlled mobility and wireless transmission in order to improve the throughput and delay performance of DTNs under the assumption that simultaneous transmissions to different collectors do not interfere with each other. Then, for DTNs under interference, we simplified the mobility of the servers to a grid and formulated a scheduling problem subject to switching delays, which were given by the time to reshuffle collectors' positions on the grid. We showed that a Max-Weight policy implemented over frames of fixed duration can stabilize such systems if the arrival rate information is available. Motivated by this scheduling problem, in this chapter we consider a general queueing model for networks with interference constraints and reconfiguration delays, where interference constraints are given by feasible schedules, and reconfiguration delay is defined as the duration of time required for one (feasible) service schedule to be dropped and a distinct service schedule to be adopted in the network. The network model we consider includes single-hop networks as shown in Fig. 4-1, wireless uplinks/downlinks or satellite networks as shown in Fig. 4-2, or optical networks [81], [99] as special cases.

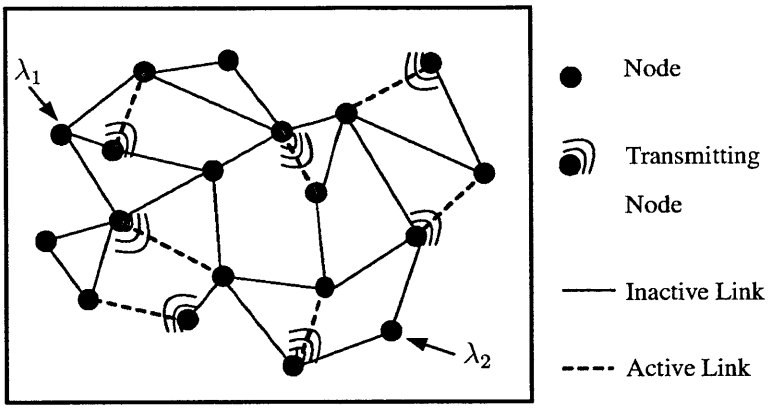


Figure 4-1: System model. A single-hop wireless network with interference constraints and reconfiguration delays.

Dynamic scheduling of such stochastic networks with interference constraints has been a very active field (e.g., [43], [75], [82], [86], [87], [125], [103], [109], [110], [122]). However, the significant effects of server switching delays or the time to reconfigure schedules have been largely ignored. Reconfiguration delay is a widespread phenomenon that is observed in many practical telecommunication systems such as satellite networks [20], [112], optical communication systems [26], [81], or wireless networks, [3], [20], [112], [126]. In this Chapter, we study the impact of reconfiguration delays on throughput and delay performance of such networks, and on throughput-optimal policies.

We consider communication networks in the absence of time-varying channels in this chapter, which can model wired (static) networks, time-invariant wireless or satellite networks, or optical networks. The time-invariance condition is relaxed in the subsequent two chapters, where we show that the simultaneous presence of time-varying channels and re-

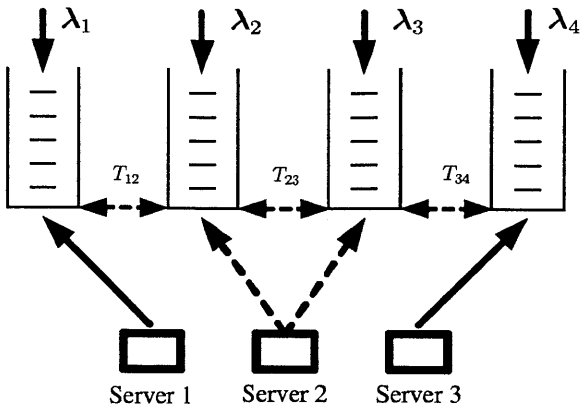


Figure 4-2: System model with $N = 4$ queues and $M = 3$ servers. Server 2 is forced to be idle due to interference constraints.

configuration delays results in fundamentally different system characteristics and requires a separate treatment. It is known that, in the absence of time-varying channels, switching delay does not reduce the stability region [26], [32]. However, as we show in Section 4.4, the celebrated Max-Weight policy and its variations do not achieve maximum throughput as they reconfigure the schedule too often, incurring large throughput losses during reconfiguration. In order to overcome the negative effects of reconfiguration delays, [26] considered a frame based scheme which persists with a Max-Weight schedule during a frame of fixed duration. This scheme, i.e., the Framed-Max-Weight policy considered at the end of Chapter 3, was shown to be throughput-optimal *when the arrival rate vector was known in advance.*

When the arrival rate information is not available, algorithms that dynamically arrange the scheduling service intervals in order to account for the switching delay are necessary. We propose scheduling policies that keep the active schedule for a duration of time determined as a function of the queue lengths. We consider several such methods for choosing service intervals, such as service intervals that are deterministic or randomized functions

of queue lengths. We first consider a frame-based policy, the Variable Frame Max-Weight (VFMW) policy, which operates over frames of length based on the queue sizes. The VFMW policy applies the Max-Weight schedule corresponding to the beginning of the frame during an interval of length that is a deterministic and sublinear function of the queue sizes at the beginning of the frame, and maximizes the throughput in the system *without requiring knowledge of the arrival rates.*

Next, we consider more adaptive classes of policies that have the capability of making switching decision instantly in case of a large batch of arrivals. We propose the Hysteresis Based (HB) policies that give a bias based on queue sizes to the weight of the current schedule at each switching instant, and wait for this bias to diminish before the next switching decision. Next, we consider *the Switching Curve Based* (SCB) policies that make switching decisions based on appropriately designed switching curves as functions of the queue lengths. We show that the bias sizes and the switching curves that are sublinear functions of the *current* queue lengths lead to good throughput and delay properties. These classes of policies are more adaptive to bursts in arrivals since they track the queue dynamics more closely as compared to the VFMW policies. This is because the HB and the SCB policies do not set a fixed frame-size at the beginning of each frame, but rather make switching decisions based on the current queue lengths. We develop sufficient conditions on the first and second moments of service intervals that guarantee stability *without requiring knowledge of the arrival rates*, and show that these policies satisfy these conditions. The reason the VFMW, the HB and the SCB policies do not require the knowledge of arrival rates is that they dynamically adapt the interswitching duration to the stochastic arrivals using queue lengths. Numerical results regarding the application of the HB, the SCB and the VFMW policies demonstrate that these policies outperform the well-known Max-Weight

policy, and also that the SCB policies have a better delay performance as compared to the VFMW and the HB algorithms.

4.1 Related Work

Switching delay has been considered in Polling models in Queuing Theory community usually in the context of a single server serving a finite number of queues (e.g., [8], [72], [77], [114]). Steady-state queue length distribution of various Polling models with and without switchover delays was analyzed in [21], where the impact of the switchover delay was studied, and a method for computing moments of message waiting time was proposed. Stability of Polling systems under various service disciplines was studied under cyclic routing in [8], [51], and under state-dependent routing in [45]. In this context, [24] and [124] characterized mean waiting times for Exhaustive, Gated and Globally Gated service disciplines, while [22] derived the pseudo conservation law for mean waiting times. Finally, optimal server routing and various dominance relationships were analyzed in [72] and [77]. An extensive review of related work in the context of Polling Systems can be found in [105] or [114]. These works either consider a system with a single server or multiple servers that do not interfere with each other, and usually analyze the performance of algorithms that do not rely on queue length information, such as the Exhaustive, (Globally) Gated, or Limited-service disciplines.

Tassiulas and Ephremides characterized the stability region of wireless networks with interference constraints in [109] and proved throughput-optimality for the Max-Weight scheduling policy that works without requiring arrival rate information. These results were later extended to joint power allocation and routing in wireless networks in [86,87], and optimal scheduling for switches in [99, 103]. More recently, suboptimal distributed schedul-

ing algorithms with throughput guarantees were studied in [28, 69, 75, 122], while [43, 82] developed distributed algorithms that achieve throughput-optimality (see [50], [84] for a detailed review). These models do not consider the server switching delays and they fail to provide stability when there are switching delays.

Of particular relevance to our work in the literature are the works in [26] by Brzezinsky and Modiano, and [108], [111] by Tassiulas and Papavassiliou, which consider perturbations such as switching delays in a network setting. The models and the assumptions considered in these works are significantly different from this chapter. While [26] assumes the knowledge of arrival rates, [108] considers a deterministic setting where servers or queues do not interfere. Furthermore, [111] proposes a policy similar to one of the policies proposed in this chapter, however, the system considered in [111], i.e., a single server serving multiple queues with asynchronous transmission opportunities in the absence of switching delays, is significantly different from our model. After we present the results in this chapter, we discuss the similarities and the differences between these works and the work in this chapter in more detail.

4.2 Organization

This chapter is organized as follows. In Section 4.3 we describe the system considered in more detail. We give a motivating example in Section 4.4 showing that the Max-Weight policy is not throughput-optimal for systems with non-zero reconfiguration delays, and build insight into the properties of stable policies. We introduce sufficient conditions for throughput-optimality of a policy and show that the class of VFMW and the SCB algorithms satisfy such conditions in Section 4.5.1. Finally, we present simulation results regarding the application of the VFMW, SCB, algorithms to different network models in

Section 4.6.

4.3 System Description and Preliminaries

Consider a single-hop wireless network given by a graph structure $\mathcal{G}(\mathcal{N}, \mathcal{L})$ of nodes $\mathcal{N}$ and links $\ell \in \mathcal{L} \doteq \{1, 2, ..., L\}$, where $L \doteq |\mathcal{L}|$. This model can also be used to analyze parallel queuing systems as shown in Fig. 4-2. In addition to the modeling assumptions in Chapter 2, we model the interference constraints in the system by the set of all possible activation vectors (schedules), $\mathcal{I} = \{\mathbf{I}^1, ..., \mathbf{I}^{|\mathcal{I}|}\}$, where vectors $\mathbf{I} \in \mathcal{I}$ take non-negative integer values. Namely, if the activation vector $\mathbf{I}(t) = (\mathbf{I}_\ell(t))_{\ell=1,...,L}$ is used at time slot t, then $\min\{\mathbf{I}_\ell(t), Q_\ell(t)\}$ packets depart from queue ℓ. We assume that the queues are initially empty and that the arrivals take place at the end of time slots. Under this model, the queue sizes evolve according to the following expression.

$$Q_\ell(t+1) = \max\{Q_\ell(t) - \mathbf{I}_\ell(t), 0\} + A_\ell(t), \forall \ell \in \mathcal{L}. \tag{4.1}$$

Let $\mathbb{H}(T) = [\mathbf{Q}(t)]|_{t=0}^{T} \cup [\mathbf{I}(t)]|_{t=0}^{T-1}$ denote the full history of the system until time t and let $\Upsilon(\mathcal{I})$ denote the set of all probability distributions on the set of all actions $\mathcal{I}$. A control policy π is a mapping from $\mathbb{H}(t)$ to $\Upsilon(\mathcal{I})$ [86], [94].

When $T_r = 0$, the stability region of this system, $\mathbf{\Lambda}^0$, consists of all arrival rate vectors $\boldsymbol{\lambda}$ in the convex hull of the vectors in $\mathcal{I}$ [86], [109], i.e., $\mathbf{\Lambda}^0 = \{\boldsymbol{\lambda} | \boldsymbol{\lambda} \in Conv\{\mathcal{I}\}\}$. The Max-Weight algorithm, which applies the activation vector $\mathbf{I}^*(t) = \arg\max_{\mathbf{I} \in \mathcal{I}} \mathbf{Q}(t) \cdot \mathbf{I}$ in each time slot, is throughput-optimal for $T_r = 0$. As we show in the next section, this property of the Max-Weight algorithm no longer holds when $T_r > 0$. Because we lose service opportunities during the reconfiguration times, the stability region of our system

satisfies $\Lambda \subseteq \Lambda^0$. We will establish that $\Lambda = \Lambda^0$.

4.4 Motivation

In this section we first show the instability of the Max-Weight policy in systems with non-zero reconfiguration delays. Then, using the insight from this analysis, we give intuition regarding the properties of throughput-optimal algorithms.

We show the instability of the Max-Weight policy for a 2-queue and single server system with i.i.d. Bernoulli arrivals with probability $p < 1/2$. The set of available activation vectors is $\mathcal{I} = \{(1,0),(0,1)\}$, and the switching delay is $T_r = 1$ slot. The stability region of this simple system is $\{p|p < 1/2\}$. The Max-Weight policy decides to switch when the boundary $Q_1 = Q_2$ is crossed where the prior decision is maintained on the boundary. By construction, the queue lengths are confined to satisfy $|Q_1 - Q_2| \leq 3$, and there are an infinite number of service switches almost surely (a.s.).

Lemma 10 Max-Weight policy is not throughput-optimal. Furthermore, there exists an arrival rate $\hat{p} < 0.5$ such that both queues grow to infinity a.s. for all $p > \hat{p}$.

The proof is given in Appendix A. For the system considered in this example, the Max-Weight policy decides to switch the server every time the $Q_1(t) = Q_2(t)$ line is crossed (i.e., the 45 degree line in the first quadrant), spending a significant fraction of the server's time during reconfiguration. A better policy should keep the current schedule for an extended period of time, and should minimize the fraction of time spent on switching. Moreover, this average fraction of time spent on switching should be decreasing as the arrival rates are get close to the boundary of the stability region. Therefore, the question that must be addressed is how to dynamically arrange the interswitching times. In the next sections, we

discuss 2 major classes of algorithms that decide on the next switching instant based on queue sizes achieve the objectives listed above.

4.5 Throughput-Optimal Algorithms

In this section we present a general method of obtaining throughput-optimal algorithms termed the Generalized Max-Weight Algorithms (GMW). These algorithms all have in common the property that they apply the Max-Weight schedule corresponding to the beginning of the current service interval for a certain duration of time. They differ in the method utilized in determining the next switching epoch, termed the *stopping rule*. This choice of the duration of the service interval is crucial as it determines stability and delay performance of the GMW algorithms. A precise description of the GMW policies is given below. Assume that the current switching epoch is t_k, and the stopping rule utilized is $\mathcal{E}$.

STRUCTURE of GMW POLICIES

1: Find the Max-Weight activation vector at time t_k, $\mathbf{I}^*(t_k)$:

$$\mathbf{I}^*(t_k) = \arg\max_{\mathbf{I}\in\mathcal{I}} \mathbf{Q}(t_k)\cdot\mathbf{I}$$

2: If $\mathbf{I}^*(t_k) \neq \mathbf{I}^*(t_{k-1})$, then let the system reconfigure for T_r slots.

3: Serve queues in $\mathbf{I}^*(t_k)$ until the stopping rule $\mathcal{E}$ is realized.

4: Repeat above for the next service interval starting at t_{k+1}.

We first present sufficient conditions which lead to stability if they are satisfied by a given control algorithm. Next, we give examples of classes of control algorithms that satisfy such conditions, such as the Variable Frame-Based Max-Weight algorithms (VFMW) that fixes the frame sizes at the beginning of the service intervals. Next, we present the class of Switching Curve Based (SCB) policies that are more adaptive to bursts in arrivals and show their throughput-optimality.

4.5.1 Sufficient Stability Conditions

Let χ_k be the length of the $k+1$th service period, i.e., $\chi_k = t_{k+1} - t_k$. In the following, a function $F(\cdot)$ is called *sublinear* if it is a satisfies $F(\cdot) > 0$ and

$$\lim_{y \to \infty} \frac{F(y)}{y} = 0.$$

Theorem 4 Suppose the following conditions are satisfied under a given GMW-type control policy: There exists a compact set $\mathcal{C}$ such that whenever $\mathbf{Q}(t_k)$ is outside $\mathcal{C}$, we have

$$i) \quad \mathbb{E}[\chi_k^2 | \mathbf{Q}(t_k)] \leq T_r^2 + c_1 F(S(\mathbf{Q}(t_k)))^2 \tag{4.2}$$

$$ii) \quad \mathbb{E}[\chi_k | \mathbf{Q}(t_k)] \geq c_2(1 - \delta(\mathbf{Q}(t_k)))F(S(\mathbf{Q}(t_k))) \tag{4.3}$$

where c_1 and c_2 are constants, $\delta(\cdot)$ is a decreasing function of $S(\mathbf{Q}(t_k))$, and $F(\cdot)$ is a monotonically increasing sublinear function. Then the system is stable for all arrival rates $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^0$ without requiring knowledge of $\boldsymbol{\lambda}$.

The proof of Theorem 4 is given in Appendix B. An example for function F is $(\sum_\ell Q_\ell(t_k))^\alpha$ for a fixed $\alpha \in (0,1)$. The proof is based on establishing a negative drift over the switching epochs t_k using a quadratic Lyapunov function, and then utilizing this result to establish stability of the overall system. A given policy prescribes a stopping time that marks the end of the current service interval. The proof is novel in that it performs a drift analysis over a period of random duration determined by the stopping rule as a function of the queue sizes. The basic intuition behind the proof is that if the queue sizes are large, under stable policies, total accumulated negative drift of the queue sizes over this random service interval must overcome the total positive drift accumulated during reconfiguration. Note that choosing the function $F(.)$ as a *sublinear* function of the queue sizes is critical. This is because GMW algorithms use the Max-Weight schedule corresponding to the beginning of the service interval, which "loses weight" as the interval goes on. Therefore, one

needs to make sure that the system is not subjected to this "light-weight" schedule for too long. In particular, service intervals sublinear in queue sizes work, however, those that are linear in queue sizes do not guarantee stability.

Next we give examples of classes of policies that satisfy the conditions of Theorem 4.

Variable Frame Based Max-Weight (VFMW) Algorithm

The VFMW policy operates over frames whose lengths are determined solely as a function of the queue sizes at the beginning of the frame, $\mathbf{Q}(t_k)$. The VFMW policy applies the Max-Weight schedule $\mathbf{I}^*(t_k)$ for an interval of duration $F(S(\mathbf{Q}(t_k)))$ slots where $F(.) > 0$ is a monotonically increasing sublinear function. The VFMW policy is defined in detail in Algorithm 3.

Algorithm 3 VFMW ALGORITHM

1: Apply the GMW structure with the stopping rule $\mathcal{E}_{VFMW}$ defined through the associated service period χ_k:

$$\chi_k = T_r + F(S(\mathbf{Q}(t_k))),$$

where $F(\cdot) > 0$ is a monotonically increasing sublinear function.

Theorem 5 The system is stable under the VFMW policy for all arrival rates $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^0$ without requiring knowledge of $\boldsymbol{\lambda}$.

Proof: The stopping rule $\mathcal{E}_{VFMW}$ is a deterministic function of $S(\mathbf{Q}(t_k))$ for the VFMW policy. Namely, given $\mathbf{Q}(t_k)$, the duration of the frame is fixed at $\chi_k \doteq T_r + F(S(\mathbf{Q}(t_k)))$. Therefore, the stability of the VFMW policy follows from Theorem 4. □

This establishes that there exists a stabilizing policy for all $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_0$. Therefore, we have the following corollary:

Corollary 2 $\mathbf{\Lambda} = \mathbf{\Lambda}^0$.

Under the VFMW policy the frequency of service reconfiguration is small when the queue sizes are large, limiting the fraction of time spent to switching. Note that this frequency should not be too small otherwise the system becomes unstable as it is subjected to a bad schedule for an extended period of time. Indeed, frame sizes linear in queue lengths do not guarantee stability in our framework. When the queue sizes are small, the VFMW policy gives frequent reconfiguration decisions, becoming more adaptive.

Below we present the HB and the SCB policies which do not fix the duration of the current service interval, but determine the next switching time according to a stopping rule based on the current queue sizes. This gives the HB and the SCB policies the ability to switch the service schedule instantly, making it more responsive to changes in queue sizes.

4.5.2 Hysteresis Based (HB) Algorithms

The HB policies calculate a built in bias for the current schedule at each switching instant, and wait for this bias to diminish before the next switching decision. The HB policy is defined in detail in Algorithm 4. We let $T_k \doteq T_r + \tau$ denote the associated stopping time. Note that the $\mathcal{L}_1$ norm $||\mathbf{Q}||$

Algorithm 4 HB ALGORITHM

1: Apply the GMW policy structure with the following stopping rule $\mathcal{E}_{HB}$.

Stop when the following inequality is achieved

$$||\mathbf{Q}(t_k + T_r + \tau) - \mathbf{Q}(t_k)|| \geq \xi_{HB} F(S(\mathbf{Q}(t_k))),$$

where $0 < \xi_{HB} < 1$ is a fixed constant to be chosen shortly, and $F(.) > 0$ is the monotonically increasing sublinear function.

for a vector $\mathbf{Q} = [Q_1, ..., Q_N]'$ of N elements is given by $\sum_{\ell=1}^{N} |Q_\ell|$.

Theorem 6 HB algorithms stabilize the system for all arrival rates $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^0$ without requiring knowledge of $\boldsymbol{\lambda}$.

The proof is given in Appendix C. It establishes that T_k satisfies the first and second moment conditions of Theorem 4. This is because the first and second moments of the time until we observe a change in queues lengths of size $F(S(\mathbf{Q}(t_k)))$ is proportional to $F(S(\mathbf{Q}(t_k)))$ and $F(S(\mathbf{Q}(t_k)))^2$.

Next, we introduce the SCB policies that make switching decisions based on appropriately designed switching curves as functions of the queue lengths.

4.5.3 Switching Curve Based (SCB) Algorithms

The SCB policy is defined in detail in Algorithm 5. In Appendix D, we show that the SCB

Algorithm 5 SCB ALGORITHM

1: Apply the GMW policy structure with the following stopping rule $\mathcal{E}_{SCB}$.

Stop the current service interval if there exists a schedule $\mathbf{I}(t)$ that satisfies

$$\mathbf{Q}(t) \cdot \mathbf{I}(t) \geq F(S(\mathbf{Q}(t))) + \mathbf{Q}(t) \cdot \mathbf{I}^*(t_k), \tag{4.4}$$

$F(.) > 0$ is a monotonically increasing sublinear function.

algorithm provides stability *at the decision times (frame boundaries)*, for all arrival rates $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^0$. Namely, the embedded Markov chain $\mathbf{Q}(t_k)$, $k \in \{0, 1, 2, ...\}$ is positive recurrent, and has a finite expectation in steady-state. However, we do not have the full stability for the SCB policy. The technical difficulty associated with this is that the SCB policy requires a competing schedule $\mathbf{I}(t)$ to pass the weight of the current schedule $\mathbf{I}^*(t_k)$ by a certain margin. Because the arrival processes are unbounded, the *the last arrivals that cause the stopping event* $\mathcal{E}_{SCB}$ can have arbitrarily large values. In particular, expectations of these *stopping arrivals* may be infinite for certain arrival processes. Therefore, our analysis does not apply. On the other hand, if the arrival processes are bounded, i.e.,

if $A_\ell(t) \leq A_{\max} < \infty$, $\forall \ell$, $\forall t$, then it can be shown, via an analysis very similar to that for the HB algorithm in Appendix C, that χ_{SCB} satisfies the sufficient stability conditions of Theorem 4.

The technique used in Appendix D to show the stability of the SCB policy at the frame boundaries for general arrival processes is novel. It first shows that $\chi_{SCB}(w) \geq \chi_{HB}(w)$ for all sample paths w for an appropriate choice of the constant ξ_{HB}. Next, using the definition of SCB policy in 4.4, it shows that for queue lengths outside a compact set, the one-step drift of the queue lengths is negative with respect to the quadratic Lyapunov function. Therefore, during the time interval between the stopping events $\mathcal{E}_{HB}$ and $\mathcal{E}_{SCB}$, the SCB policy keeps accumulating more negative drift until the switching curve $\mathcal{E}_{SCB}$ is hit, leading to stability at the decision times.

The SCB policies employ a schedule whose weight is *sufficiently* close to the current Max-Weight schedule. The stopping rules $\mathcal{E}_{HB}$ and $\mathcal{E}_{SCB}$ guarantee this property by making sure that the weight of the currently active schedule is at most $F(S(\mathbf{Q}(t)))$ away from the Max-Weight schedule. Since $F(S(\mathbf{Q}(t)))$ is an increasing function of $S(\mathbf{Q}(t))$, the SCB policies do not switch too frequently unlike the ordinary Max-Weight policy. Furthermore, since $F(S(\mathbf{Q}(t)))$ is a sublinear function of $S(\mathbf{Q}(t_k))$, the system is not subjected to a possibly light-weight schedule for too long.

The main advantage of the SCB policies as compared to the VFMW policies is that the SCB policies are more responsive (or adaptive) to bursts in arrivals. This is because the VFMW policies determine each frame size at the beginning of the frame and stick to the schedule until the end of the frame. Whereas the SCB policies can switch the schedule instantly if there is large burst of arrivals to a queue that is not currently receiving any service.

The VFMW and the SCB algorithms have much lower computational complexity than the ordinary Max-Weight algorithm since they perform scheduling computation only once per frame. As first suggested in [89], by using out-of-date queue length information, algorithms that operate over frames can be implemented to perform the computation of the next schedule during the current frame, without reducing the stability region. Therefore, letting C_{MW} denote the computational complexity of the ordinary Max-Weight algorithm per time slot, the VFMW or the SCB algorithms

have $C_{MW}/\mathbb{E}[\chi]$ computational complexity per time slot, where $\mathbb{E}[\chi]$ is the steady-state expected frame length.

4.6 Simulations

We performed simulation experiments that determine average queue occupancy values for the Switching Curve Based (SCB), Hysteresis Based (HB), and the Variable Frame Based Max-Weight (VFMW) policies, as well as the ordinary Max-Weight policy and the Max-Weight policy with fixed frame sizes (MWFF). The average queue occupancy of queue ℓ over T_s slots is given by $\frac{1}{T_s}\sum_{t=1}^{T_s} Q_\ell(t)$ and the sublinear function $F(\cdot)$ for the SCB, HB, and the VFMW policies is chosen to be $F(S(\mathbf{Q}(t))) = (\sum_\ell Q_\ell(t))^{0.7}$. Through Little's law, the long-run packet-average delay in the system is equal to the time-average number of packets divided by the total arrival rate into the system.

We considered a network of 4 links and 3 servers as shown in Fig. 4-2, where servers 1 and 3 are dedicated to links (queues) 1 and 4 respectively, and server 2 is shared between queues 2 and 3. This system can also model an appropriate single-hop wireless network. Due to interference constraints, no two links that are "adjacent" to each other can be activated simultaneously, namely, the set of feasible activations are given by $\mathbf{I}^1 = [1010]$, $\mathbf{I}^2 = [0101]$, and $\mathbf{I}^3 = [1001]$. The stability region for this network is given by $\mathbf{\Lambda} = Conv\{\mathbf{I}^0, \mathbf{I}^1, \mathbf{I}^2, \mathbf{I}^3\}$, where $\mathbf{I}^0 = [0000]$, and the sum throughput bound is given by $\sum_{\ell=1}^4 \lambda_\ell \leq 2$. We also consider the 2-queue network described in Section 4.4, except that we have different Bernoulli arrivals to each queue and that the switchover delay T_r is taken to be greater than 1 in the experiments. For the two-queue system, the stability region is given by $\mathbf{\Lambda} = \{(\lambda_1, \lambda_2) | \lambda_1 + \lambda_2 \leq 1, \lambda_1, \lambda_2 \geq 0\}$.

For the two-queue network of Section 4.4, Fig. 4-3 compares the stability regions of the SCB, VFMW and the Max-Weight policies when $T_r = 2$. This experiment can model single-hop optical networks where the reconfiguration time T_r is usually relatively small. Fig. 4-3 (a) confirms that the plain Max-Weight policy is not throughput-optimal, and the points corresponding to the sudden

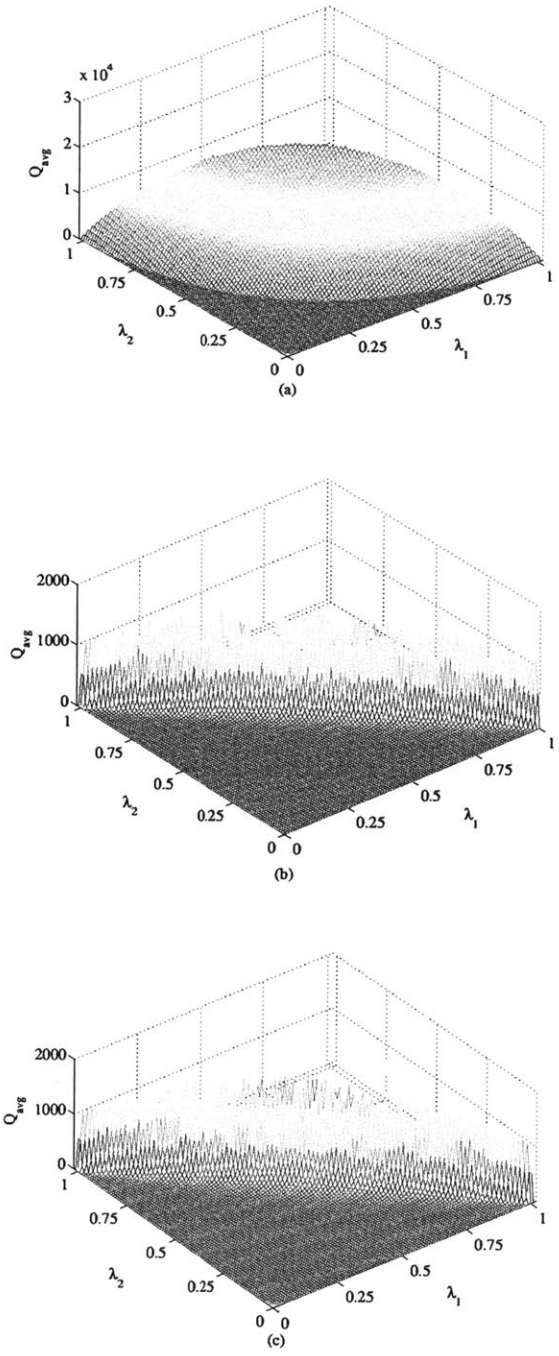


Figure 4-3: The total average queue size for (a) the Max-Weight policy, (b) the SCB policy and (c) the VFMW policy for 2 queues and $T_r = 2$ slots.

jump in the plot represent the boundary of the region stabilized by the Max-Weight policy. Figures 4-3 (b) and (c) show that the SCB and the VFMW policies have bounded queue sizes for all arrival rates inside the stability region $\mathbf{\Lambda}$. Note that the HB policy applied to this network yields a similar plot, and therefore it is not presented here.

Fig. 4-4 presents the delay as a function of throughput for the VFMW, Max-Weight and the MWFF (with frame sizes $T = 30$ and $T = 80$) policies along the main diagonal line. The switchover delay in this experiment, $T_r = 20$ slots, is relatively large, which could represent a DTN application such as mobile elements gathering data from sensors in a field. Fig. 4-4 confirms that the VFMW policy is throughput-optimal for this system and that the system quickly becomes unstable under the Max-Weight policy as the arrival rate is increased. In Fig. 4-4, the MWFF policy with frame length $T = 30$ has similar delay performance to the Max-Weight policy for small arrival rates, however, under this policy the system becomes unstable around $\lambda_1 = \lambda_2 = 0.2$. Increasing the frame length improves the stability region of the MWFF policy at the expense of delay performance for small arrival rates. As opposed to fixed frame lengths, the VFMW policy dynamically adapts the frame length as a function of the queue states and stabilizes the system whenever possible, while providing a delay performance that is similar to that of the Max-Weight policy for small arrival rates.

Fig. 4-5 presents the delay as a function of throughput for the SCB, HB, and the VFMW policies along the main diagonal line. The switching delay in this experiment is $T_r = 20$ slots and the maximum sum-throughput is 2. The SCB policy has 22% maximum, 13% on average for higher loads ($\sum_\ell \lambda_\ell > 0.2$), and 3% on average overall delay saving as compared to the VFMW policy. As compared to the HB policy, the SCB policy has 85% maximum, 77% on average for higher loads ($\sum_\ell \lambda_\ell > 0.2$), and 63% on average delay saving. The SCB policy has a better delay performance than the VFMW policy since the SCB policy is more adaptive to changes in the queue lengths as compared to the VFMW policy. This is because the SCB policy is able to switch the service schedule whenever the switching condition is satisfied, whereas, the VFMW policy keeps

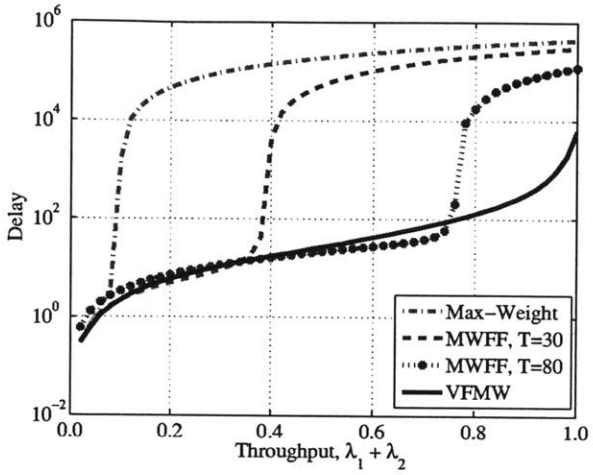


Figure 4-4: Delay (total average queue size) vs the throughput under the Max-Weight, MWFF and the VFMW policies for 2 queues and $T_r = 20$ slots.

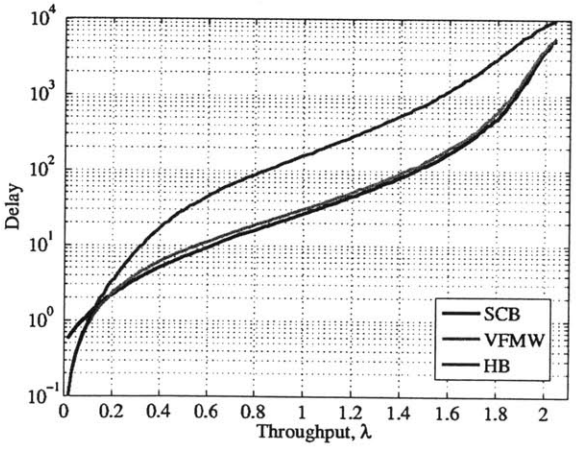


Figure 4-5: Delay (total average queue size) vs the throughput under the VFMW, HB, and the SCB policies for the 4-queues and 3-servers system and $T_r = 10$ slots.

the current schedule until the end of the predetermined frame duration.

4.7 Related Work

In this section we discuss papers that are of particular relevance to our work. A frame based scheme for optical networks which persists with a Max-Weight schedule during a frame of fixed duration was considered in [26]. This scheme was shown to be throughput-optimal *when the arrival rate vector was known in advance.* The fluid limit of the system was considered in [26] and throughput-optimality was established under rate stability[1]. As shown in Fig. 4-4, when the arrival rate information is not available, such schemes can be significantly outperformed by the VFMW policy in terms of both throughput and delay.

A single server allocation problem over N parallel queues without switchover delay in a continuous time system was considered in [108], where the service of a queue can only be initiated at certain connectivity instances that are asynchronous for different users. The server can serve only a single queue at a given time when the connectivity instance of the queue arrives, and the arrival times of these connectivity instances are either deterministic and periodic or according to a Poisson process. At any time t, the future connectivity instances are available for scheduling decisions and the service time is fixed at τ seconds. As a result, the connectivity instances that arrive within the same τ-second service times conflict with each other. An algorithm similar to the VFMW algorithm was proposed in [108] in order to stably schedule this system. Even though the idea of keeping a schedule for a duration of time that is sublinear in queue lengths stabilizes the system in [108], this system is significantly different than the model considered in this chapter.

A scheduling problem in a multihop network of multiple queues and servers with nonzero switchover delay was considered in [111]. There are no interference constraints for servers or

[1]A queue of length $Q_\ell(t)$ at time t is rate stable if $\lim_{t\to\infty} Q_\ell(t)/t = 0$. This is a weaker notion of stability as compared to strong stability in Definition 1 which implies bounded first moments of a stationary measure.

queues in [111]. Namely, each server can switch independently of other servers, and all servers can be active simultaneously. Hence, there is no global scheduling in [111], but a local scheduling that is separate for each server, simplifying the scheduling aspect of the problem. Furthermore, the system considered in [111] is deterministic in that arrivals and service processes do not have random components. In order to deal with the adverse effects of switching delays, [111] arranges switching frequencies separately for each server by giving a multiplicative bias to the queues in service based on queue sizes. The system considered in this chapter is significantly different from the system in [111] in that we analyze networks with arbitrary interference constraints between the servers and the queues. Hence our results, analysis, and protocols depend on interference constraints in the system. Moreover, we consider stochastic arrival processes which make the analysis more challenging for networks under nonzero reconfiguration delays. Furthermore, we consider various different classes of throughput-optimal policies such as policies that choose service intervals that are deterministic or randomized, and sublinear or linear functions of queue sizes.

4.8 Concluding Remarks

We investigated the scheduling problem for networks under arbitrary interference constraints and reconfiguration delays. We showed that the Max-Weight scheduling algorithm is not throughput-optimal for such systems and we first developed sufficient conditions on the expected inter-switching time that leads to stability. We discussed the class of Variable Frame Based Max-Weight (VFMW), Hysteresis Based (HB) and and Switching Curve Based (SCB) algorithms that satisfy these conditions and that provide throughput-optimality[2] without requiring the knowledge of the arrival rates.

The VFMW algorithms persist with the Max-Weight schedule during an interval of duration dependent on the queue sizes, which dynamically adapts the frame sizes to stochastic arrivals and provides a reasonable delay performance in addition to stability. The HB and the SCB policies are more adaptive to bursty arrivals in that these policies do not have a fixed frame-size at the beginning

[2]SCB algorithms are throughput optimal for bounded arrival processes.

of each frame, but they switch based on the instantaneous queue states.

Possibly future directions include developing low-complexity distributed algorithms with large throughput guarantees and joint scheduling and routing algorithms in multihop networks with interference constraints and switchover delays.

Appendix A - Proof of Lemma 10

Let t_k denote the epoch immediately after k^{th} reconfiguration and let $M_k \doteq Q_1(t_k) + Q_2(t_k), k = 0, 1, 2, \cdots$, where $M_0 = 0$ by definition. Consider the total drift along the $Q_1(t) = Q_2(t)$ line (the main diagonal) and suppose without loss of generality that queue 1 is being served. We call slot t a *freeze slot* if the queues retain their state in the next slot, i.e., $\mathbf{Q}(t+1) = \mathbf{Q}(t)$. Note that all other state changes are called *progressive* since they bring the queue states closer to the main diagonal. A freeze slot happens with probability $p(1-p)$ and hence the number of freeze slots until a progressive state is $\sim Geom(1-p(1-p))$ with mean $\frac{(p(1-p))}{1-p+p^2}$. Furthermore the drift along the main diagonal per service slot is, $-(1-2p)$, which is toward the origin for $p < 1/2$. Observe that at most 4 progressive state changes will lead to $Q_2 > Q_1$, where for small queue sizes smaller number of progressive steps may cause switching. It follows that the expected accumulated drift during a service interval is bounded from above by,

$$4(1-2p) \times \left(\frac{1}{1-p+p^2)}\right).$$

During the 1 slot reconfiguration the drift away from the origin is $2p$. Therefore, the expected drift along the main diagonal must be positive provided,

$$2p - 4(1-2p) \times \left(1 + \frac{p(1-p)}{1-p+p^2}\right) > 0, \tag{4.5}$$

which implies $p > 0.42049\cdots$. We have thus shown that there is a $p < 1/2$ such that the sequence M_k satisfies

$$\mathbb{E}\left[M_{k+1}|M_k\right] \geq M_k + \eta,$$

where $\eta > 0$ is a fixed constant. We now show that $M_k \to \infty$ a.s. which in turn implies $Q_\ell(t) \to \infty$. Define Z_k via $M_{k+1} = M_k + Z_k$, $k = 0, 1, 2, \cdots$ and observe that $Z_k \leq 6$. Define $R_k \doteq \frac{1}{M_k+1}$, and observe

$$\begin{aligned}\mathbb{E}\left[\frac{1}{M_{k+1}+1}|\mathcal{F}_k\right] &= \frac{1}{M_k+1} - \mathbb{E}\left[\frac{Z_k}{(M_k+1)(M_k+1+Z_k)}\right] \\ &\leq \frac{1}{M_k+1} - \frac{\eta}{(M_k+1)\,(M_k+1+6)},\end{aligned}$$

from which it follows that R_k is a nonnegative supermartingale where $\mathcal{F}_k$ is an appropriate sigma algebra at time slot t_k. Hence $\lim_k R_k \leq 1$ exists a.s. and so does $\lim_k M_k$. Finally

$$\mathbb{E}\left[R_\infty\right] \leq \liminf_k \mathbb{E}\left[R_k\right] \leq 1 - \sum_{k=0}^{\infty} \mathbb{E}\left[\frac{\eta}{(M_k+1)(M_k+7)}\right],$$

which implies

$$\begin{aligned}0 &= \liminf_k \mathbb{E}\left[\frac{\eta}{(M_k+1)\,(M_k+7)}\right] \\ &\geq \liminf_k \mathbb{E}\left[\frac{\eta}{(M_k+7)}^2\right]\end{aligned}$$

This shows that $\liminf_k (M_k+7)^{-1} = 0$, a.s. from Fatou's lemma and since actually there is a limit we have $\lim_k M_k = \infty$ a.s. But this implies $\lim_k Q_\ell(t_k) \infty$ a.s., since $|Q_1 - Q_2| \leq 3$. Moreover if t satisfies, $t_k \leq t < t_{k+1}$ then $|Q_\ell(t) - Q_\ell(t_k)| \leq 4, \ell = 1, 2$ so that $Q_\ell(t) \to \infty$ a.s. Since $\infty = \mathbb{E}\left[\lim_t Q_\ell(t)\right] \leq \liminf_t \mathbb{E}\left[Q_\ell(t)\right]$ so the stability criterion does not hold.

Appendix B - Proof of Theorem 4

Consider the following χ_k-step queue evolution expression:

$$Q_\ell(t_k+\chi_k) \leq \max\left\{Q_\ell(t_k) - \sum_{\tau=0}^{\chi_k-1} \mathbf{I}_\ell(t_k+\tau), 0\right\} + \sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau).$$

To see this, note that if $\sum_{\tau=0}^{\chi_k-1} \mathbf{I}_\ell(t_k+\tau)$, the total service opportunity given to queue ℓ during the k^{th} frame, is smaller than $Q_\ell(t_k)$, then we have an equality. Otherwise, the first term is 0 and we have an inequality. This is because some of the arrivals during the frame might depart before the end of the frame. We first prove stability at the frame boundaries. Squaring both sides, using $\max(0,x)^2 \leq x^2, \forall x \in \Re$, and $\mathbf{I}_\ell(t) \leq \mu_{\max}, \forall t$ we have

$$\begin{aligned} Q_\ell(t_k+\chi_k)^2 - Q_\ell(t_k)^2 \leq \chi_k^2 \mu_{\max}^2 + \Big(\sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big)^2 \\ - 2Q_\ell(t_k)\Big(\sum_{\tau=0}^{\chi_k-1} \mathbf{I}_\ell(t_k+\tau) - \sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big). \end{aligned} \tag{4.6}$$

Define the quadratic Lyapunov function $L(\mathbf{Q}(t)) = \sum_{\ell=1}^{L} Q_\ell^2(t)$, and the χ_k-step conditional Lyapunov drift

$$\Delta_{\chi_k}(t_k) \triangleq \mathbb{E}\left[L(\mathbf{Q}(t_k+\chi_k)) - L(\mathbf{Q}(t_k)) \middle| \mathbf{Q}(t_k)\right].$$

We sum (4.6) over the queues, take the conditional expectation and use Wald's equality to have the following χ_k-step drift expression

$$\begin{aligned}\Delta_{\chi_k}(t_k) \quad &\leq \quad L\mu_{\max}^2\mathbb{E}[\chi_k^2|\mathbf{Q}(t_k)]+\sum_{\ell}\mathbb{E}\left[\Big(\sum_{\tau=0}^{\chi_k-1}A_\ell(t_k+\tau)\Big)^2|\mathbf{Q}(t_k)\right] \\ &+ \quad 2\mathbb{E}[\chi_k|\mathbf{Q}(t_k)]\sum_{\ell}Q_\ell(t_k)\lambda_\ell-2\sum_{\ell}Q_\ell(t_k)\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1}\mathbf{I}_\ell(t_k+\tau)\big|\mathbf{Q}(t_k)\right] \qquad (4.7)\end{aligned}$$

where $B = A_{\max}^2 + \mu_{\max}^2$ is a constant and we used the fact that the arrival processes are i.i.d. over time, independent of the queue lengths. Now we apply the following facts:

- $\mathbb{E}[\chi_k^2|\mathbf{Q}(t_k)] \leq T_r^2 + c_1F(S(\mathbf{Q}(t_k)))^2$

- $\mathbb{E}[\chi_k|\mathbf{Q}(t_k)] \geq c_2(1-\delta(\mathbf{Q}(t_k)))F(S(\mathbf{Q}(t_k)))$

- Recalling that the system is idle for the first T_r slots of the frame, we have

  $\sum_\ell Q_\ell(t_k)\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1}\mathbf{I}_\ell(t_k+\tau)\big|\mathbf{Q}(t_k)\right]=(\mathbb{E}[\chi_k|\mathcal{Q}(t_k)]-T_r)\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k)$

- for any arrival rate vector $\boldsymbol{\lambda}$ that is strictly inside $\boldsymbol{\Lambda}^0$, there exist real numbers $\beta^1,...,\beta^{|\mathcal{I}|}$ such that $\beta^j \geq 0, \forall j \in 1,...,|\mathcal{I}|$, $\sum_{j=1}^{|\mathcal{I}|}\beta^j = 1-\epsilon$ for some $\epsilon > 0$ and $\boldsymbol{\lambda} = \sum_{j=1}^{|\mathcal{I}|}\beta^j\mathbf{I}^j$ [26].

Using these results in (4.7) we have

$$\begin{aligned}
\Delta_{\chi_k}(t_k) &\leq L\mu_{\max}^2(T_r^2 + F(S(\mathbf{Q}(t_k))^2)) + \sum_{\ell}\mathbb{E}\left[\Big(\sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big)^2 |\mathbf{Q}(t_k)\right] \\
&+ 2\mathbb{E}[\chi_k|\mathbf{Q}(t_k)]\mathbf{Q}(t_k)\cdot\sum_{j=1}^{|\mathcal{I}|}\beta^j\mathbf{I}^j - 2(\mathbb{E}[\chi_k|\mathbf{Q}(t_k)]-\tau)\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k) \\
&\leq L\mu_{\max}^2(T_r^2 + F(S(\mathbf{Q}(t_k))^2)) + \sum_{\ell}\mathbb{E}\left[\Big(\sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big)^2 |\mathbf{Q}(t_k)\right] \\
&- 2c_2(1-\delta)F(S(\mathbf{Q}(t_k)))\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k)(\epsilon) + 2T_r\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k), \qquad (4.8)
\end{aligned}$$

where in the last inequality we used the fact that $\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k) \geq \mathbf{Q}(t_k)\cdot\mathbf{I}, \forall\mathbf{I}\in\mathcal{I}$. We use the following lemma in [93] to bound the second term in (4.8):

Lemma 11 Let $\{Y_i\}$ be a sequence of independent variables $\mathbb{E}[Y_i]=\lambda_i$, $\mathrm{Var}(Y_i)=\sigma_i^2<\infty$. Let T be a stopping time $T\in\sigma(Y_1,...,Y_n,...)$, and $\tilde{T}$ be an integer-valued random variable independent of the Ys but having the same distribution as T. Let $S_T = Y_1+Y_2+...+Y_T$. Then,

$$\mathbb{E}[S_T^2] \leq 2\mathbb{E}\left[S_{\tilde{T}}^2\right] = 2\mathbb{E}\left[\sum_{i=1}^{T}\sigma_i^2\right] + 2\mathbb{E}\left[\left(\sum_{i=1}^{T}\lambda_i\right)^2\right]. \qquad (4.9)$$

Applying this result to the second term in (4.8), we have

$$\begin{aligned}
\mathbb{E}\left[\Big(\sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big)^2\Big|\mathbf{Q}(t_k)\right] &\leq 2\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1} A_{\max}^2 - \lambda_\ell^2\,\Big|\mathbf{Q}(t_k)\right] + 2\mathbb{E}\left[\left(\sum_{\tau_1=0}^{\chi_k-1}\lambda_\ell\right)^2\Big|\mathbf{Q}(t_k)\right] \\
&= 2(A_{\max}^2-\lambda_\ell^2)\mathbb{E}[\chi_k|\mathbf{Q}(t_k)] + 2\lambda_\ell^2\mathbb{E}[\chi_k^2|\mathbf{Q}(t_k)]
\end{aligned}$$

Using $\mathbb{E}[\chi_k^2|\mathbf{Q}(t_k)] \leq T_r^2 + c_1 F(S(\mathbf{Q}(t_k)))^2$, we have

$$\mathbb{E}\left[\Big(\sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big)^2 \Big| \mathbf{Q}(t_k)\right] \leq c_3 F(S(\mathbf{Q}(t_k)))^2, \tag{4.10}$$

where c_3 is an appropriate constant. Using this result in (4.8) we have

$$\begin{aligned}\Delta_{\chi_k}(t_k) &\leq L\mu_{\max}^2 T_r^2 + (L\mu_{\max}^2 + c_3)F(S(\mathbf{Q}(t_k)))^2 \\ &- 2F(S(\mathbf{Q}(t_k))\left(c_2(1-\delta)\epsilon - \frac{T_r}{F(S(\mathbf{Q}(t_k)))}\right)\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k)).\end{aligned}$$

Since $F(S(\mathbf{Q}(t_k)))$ is a monotonically increasing function of $F(S(\mathbf{Q}(t_k)))$, and $\delta(S(\mathbf{Q}(t_k)))$ is a monotonically decreasing function of $S(\mathbf{Q}(t_k))$, there exists a constant c_4 such that, if $S(\mathbf{Q}(t_k)) > c_4$, $c_2(1-\delta)\epsilon - \frac{T_r}{F(S(\mathbf{Q}(t_k)))} > \delta_1 > 0$:

$$\Delta_{\chi_k}(t_k) \leq L\mu_{\max}^2 T_r^2 + (L\mu_{\max}^2 + c_3)F(S(\mathbf{Q}(t_k)))^2 - 2\delta_1 F(S(\mathbf{Q}(t_k))\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k).$$

Hence, for $S(\mathbf{Q}(t_k)) > c_4$, we use $\mathbf{Q}(t_k)\cdot\mathbf{I}^*(t_k) \geq \frac{1}{N}\sum_\ell Q_\ell(t_k)$ to have

$$\Delta_{\chi_k}(t_k) \leq L\mu_{\max}^2 T_r + (L\mu_{\max}^2 + c_3)F(S(\mathbf{Q}(t_k)))^2 - \frac{2\delta_1}{L}F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k)).$$

Since $F(.)$ is a sublinear function of $S(\mathbf{Q}(t_k))$, there exists a constant c such that

$$\Delta_{\chi_k}(t_k) \leq c - \delta_2 F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k)), \tag{4.11}$$

where $\delta_2 \triangleq \delta_1/L$. Taking expectations with respect to $\mathbf{Q}(t_k)$, we have

$$\mathbb{E}[L(\mathbf{Q}(t_{k+1}))] - \mathbb{E}[L(\mathbf{Q}(t_k))] \leq c - \delta_2 \mathbb{E}\left[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))\right]. \tag{4.12}$$

Writing a similar expression over the frame boundaries $t_k, k \in \{0, 1, 2, ..., K\}$, summing them and telescoping these expressions leads to

$$L(\mathbf{Q}(t_K)) - L(\mathbf{Q}(0)) \leq Kc - \delta_2 \sum_{k=0}^{K-1} \mathbb{E}[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))]\,.$$

Using $L(\mathbf{Q}(t_K)) \geq 0$ and $L(\mathbf{Q}(0)) = 0$, we have

$$\frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}\left[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))\right] \leq \frac{c}{\delta_2} < \infty.$$

This implies that

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}\left[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))\right] \leq \frac{c}{\delta_2} < \infty. \tag{4.13}$$

Furthermore, we have

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \sum_{\ell} \mathbb{E}[Q_\ell(t_k)] \leq \frac{c}{\delta_2} < \infty. \tag{4.14}$$

This establishes stability (as defined in Definition 1) at the frame boundaries $t_k, k \in \{0, 1, 2, ...\}$.

Now, we have for all frames $k \in \{0, 1, 2, ...\}$,

$$\begin{aligned}
\sum_{\tau=0}^{\chi_k-1} \sum_{\ell} Q_\ell(t_k+\tau) &\leq \sum_{\tau=0}^{\chi_k-1} \sum_{\ell} \left(Q_\ell(t_k) + \sum_{\tau_1=0}^{\chi_k-1} A_\ell(t_k+\tau_1) \right) \\
&\leq \sum_{\ell} \left(\chi_k Q_\ell(t_k) + \chi_k \sum_{\tau_1=0}^{\chi_k-1} A_\ell(t_k+\tau_1) \right) \\
&\leq \sum_{\ell} \left(\chi_k Q_\ell(t_k) + \chi_k \sum_{\tau_1=0}^{\chi_k-1} (1 + A_\ell(t_k+\tau_1)) \right) \\
&\leq \sum_{\ell} \left(\chi_k Q_\ell(t_k) + \left(\sum_{\tau_1=0}^{\chi_k-1} (1 + A_\ell(t_k+\tau_1)) \right)^2 \right)
\end{aligned}$$

Taking conditional expectation we have

$$\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1} \sum_{\ell} Q_\ell(t_k+\tau) | \mathbf{Q}(t_k) \right] \leq \mathbb{E}\left[\chi_k | \mathbf{Q}(t_k)\right] \sum_{\ell} Q_\ell(t_k) + \sum_{\ell} \mathbb{E}\left[\left(\sum_{\tau_1=0}^{\chi_k-1} (1 + A_\ell(t_k+\tau_1)) \right)^2 \Big| \mathbf{Q}(t_k) \right]$$

where we used the fact that arrival processes are i.i.d. and independent of the queue lengths. Applying Lemma 11 to the second term on the right hand side

$$\begin{aligned}
\mathbb{E}\left[\Big(\sum_{\tau=0}^{\chi_k-1} (1 + A_\ell(t_k+\tau)) \Big)^2 \Big| \mathbf{Q}(t_k) \right] &\leq 2\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1} (1 + 2\lambda_\ell + A_{\max}^2 - \lambda_\ell^2) \Big| \mathbf{Q}(t_k) \right] \\
&+ 2\mathbb{E}\left[\left(\sum_{\tau_1=0}^{\chi_k-1} (1+\lambda_\ell) \right)^2 \Big| \mathbf{Q}(t_k) \right] \\
&= 2(A_{\max}^2 + 1 + 2\lambda_\ell - \lambda_\ell^2)\mathbb{E}[\chi_k | \mathbf{Q}(t_k)] \\
&+ 2(1+\lambda_\ell^2)\mathbb{E}[\chi_k^2 | \mathbf{Q}(t_k)]. \qquad (4.15)
\end{aligned}$$

Applying this result and using $\mathbb{E}[\chi_k^2|\mathbf{Q}(t_k)] \leq T_\tau^2 + c_1 F(S(\mathbf{Q}(t_k)))^2$, we have

$$\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1}\sum_{\ell} Q_\ell(t_k+\tau)|\mathbf{Q}(t_k)\right] \leq c_5\left(S(\mathbf{Q}(t_k)) + F(S(\mathbf{Q}(t_k)))S(\mathbf{Q}(t_k)) + F(S(\mathbf{Q}(t_k)))^2\right),$$

where c_5 is an appropriate positive constant. Taking expectation w.r.t. $\mathbf{Q}(t_k)$ we have

$$\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1}\sum_{\ell} Q_\ell(t_k+\tau)\right] \leq c_5\left(\mathbb{E}\left[S(\mathbf{Q}(t_k))\right] + \mathbb{E}\left[F(S(\mathbf{Q}(t_k)))S(\mathbf{Q}(t_k))\right] + \mathbb{E}\left[F(S(\mathbf{Q}(t_k)))^2\right]\right),$$

Now, for any given large T, let K_T be the number of frames up to and including T. We have

$$\sum_{t=0}^{T-1}\sum_{\ell}\mathbb{E}[Q_\ell(t)] \leq \sum_{k=0}^{K_T-1} 3c_5\mathbb{E}\left[F(S(\mathbf{Q}(t_k)))S(\mathbf{Q}(t_k))\right]$$

Using renewal theory arguments, we establish in Lemma 12 that K_T/T converges to a constant as T tends to infinity. Therefore, dividing both sides by T, using the fact that $T > K_T$ for any T, taking the $\limsup_T$ of both sides, using (4.13) and $0 < \alpha < 1$, we have

$$\limsup_{T\to\infty} \frac{K_T}{T}\frac{1}{K_T}\sum_{t=0}^{T-1}\sum_{\ell=1}^{L}\mathbb{E}\left[Q_\ell(t)\right] < \infty. \tag{4.16}$$

Therefore, the system is stable.

Lemma 12 We have that $\lim_{T\to\infty} K_T/T$ is a finite constant. Additionally,

- the queues are all empty infinitely often (i.o.) with finite mean recurrence time slots
- the queue sizes follow an irreducible and positive recurrent Markov chain
- the stationary measure of the Markov chain has bounded first moments.

Proof: We first show the results for the queue process at the frame boundaries. Using this, we establish the results everywhere. The queue processes at the frame boundaries $(Q_\ell(t_k))_{k\geq 0}$ evolve according to a discrete time Markov chain on countable state space $\mathbb{N}_0$. We first prove the positive recurrence of the $(Q_\ell(t_k))_{k\geq 0}$ Markov chain. From (4.14), the Lyapunov function is a non-negative supermartingale outside some compact set containing the origin, call $\mathcal{C}$. Therefore, $\mathcal{C}$ is visited *i.o.* and hence $(Q_\ell(t_k))_{k\geq 0}$ is an irreducible and positive recurrent Markov chain. Therefore, there is a stationary distribution, $\mathbf{Q}^f$, such that $\mathbf{Q}(t_k)$ converges in distribution to $\mathbf{Q}^f$. From (4.14) we see that

$$\liminf \mathbb{E}[Q_\ell(t_k)] \leq \frac{c}{\delta} < \infty. \tag{4.17}$$

By Theorem 5.3 of Billingsley [19], we have

$$\mathbb{E}[Q_\ell^f] \leq \liminf \mathbb{E}[Q_\ell(t_k)] \leq \frac{c}{\delta} < \infty, \tag{4.18}$$

therefore, the stationary distribution has finite first moments. Since there is a positive probability of no arrivals to each queue and that the arrival processes to each queue are independent, the queues will empty and stay empty after a bounded number of frames with positive probability. Hence, the all queues empty state at the start of the frame is in the Markov chain. Therefore, all queues are empty i.o. over frames and the process of all-empty queue sizes at the frame boundaries is a renewal process with finite mean. This establishes the theorem at the frame boundaries.

Now we show that the mean number of slots between two renewal frames is also finite. The above argument shows that the mean number of frames between all queues empty has finite mean in frames. Define the total duration of the first K frames in slots to be,

$$S_K = \sum_{k=1}^{K} \chi_k.$$

It follows from (4.13) that

$$\limsup_{K\to\infty} \frac{1}{K}\mathbb{E}[S_K] < \frac{c}{\delta_2} < \infty. \tag{4.19}$$

Next, define X_n to be the duration of the nth renewal period in slots. X_1 is the total time in slots to the first frame boundary after time 0, with all queues empty. Let $s \triangleq \mathbb{E}[X_1]$, and note for now that s may be infinite. Finally let $m_0 < \infty$ be the finite expected number of frames between all queue empty frames. Given K, a number of frames, define $M(K)$ to be the number of renewals, which have taken place up to and including frame K. We have

$$\frac{S_K}{\sum_{n=1}^{M(K)} X_n} \geq 1. \tag{4.20}$$

$M(K)$ increases with K, and by the SLLN and the definition of $M(K)$ we have

$$\lim_{K\to\infty} \frac{M(K)}{K} = \frac{1}{m_0} \ \ a.s.$$

Consider the following extended version of the SLLN which argues that if the mean of the variables involved is infinite, the time average is also infinite: Suppose that $Y_n \geq 0$, $n = 1, 2, ...$, are i.i.d. and such that $\mathbb{E}[Y_1] = \infty$. Define, $Y_n^K \doteq \min\{Y_n, K\}$ for $K > 0$. It follows that $m_K \doteq \mathbb{E}[Y_1^K] \to \mathbb{E}[Y_1] = \infty$, by the Monotone Convergence Theorem. Also,

$$m_K = \frac{1}{N}\sum_{n=1}^{N} Y_n^K \leq \liminf_N \frac{1}{N}\sum_{n=1}^{N} Y_n, \ \ a.s.$$

Hence, it follows that

$$\lim_N \frac{1}{N}\sum_{n=1}^{N} Y_n = \infty, \ \ a.s.$$

Now, applying the extended version of the SLLN to X_n we have

$$\lim_{K\to\infty} \frac{1}{M(K)} \sum_{n=1}^{M(K)} X_n = s \leq \infty \ \ a.s.$$

Rewriting (4.20), we have

$$\frac{K}{M(K)} \frac{\frac{S_K}{K}}{\frac{1}{M(K)} \sum_{n=1}^{M(K)} X_n} \geq 1.$$

It follows that when we take a sequence $K_m \to \infty$ such that we converge to the lim inf of the numerator, we have

$$m_0 \frac{\liminf_K \frac{S_K}{K}}{s} \geq 1 \ \ a.s.$$

Taking expectations and applying Fatou's lemma we have

$$\mathbb{E}[\liminf \frac{S_K}{K}] \leq \liminf \frac{1}{K}\mathbb{E}[S_K] \leq \limsup \frac{1}{K}\mathbb{E}[S_K] < \frac{c}{\delta_2} < \infty,$$

which follows from (4.19). Hence,

$$s \leq mc/\delta_2 < \infty.$$

This establishes that K_T/T converges to the finite constant m_0/s as T tends to infinity. Since s is finite, using the fact that the queue sizes at the frame boundaries are stable and they empty i.o., the queues within the frame boundaries empty with a frequency at least $1/s$. Therefore, the mean time in slots between all queues empty for the full process is finite.

Note that the service opportunity given to a queue depends on the queue sizes at the beginning of the frame. Therefore, the queue sizes that are within the frame boundaries do not follow a Markov chain. However, if we define a new process with state $H(t) \triangleq (\mathbf{Q}(t), \mathbf{Q}(t_k), E(t))$, where $E(t)$ is the *age* in the current frame (i.e., $t - t_k < \chi_k$), then the new process $\mathbf{H}(t)$ is a Markov chain. This

irreducible Markov chain hits the state $\mathbf{H}(t) = (\mathbf{0}, \mathbf{0}, 0)$ i.o. due to above arguments. Therefore, it is positive recurrent and it has a steady state distribution $\mathbf{H} = (\mathbf{Q}, \mathbf{Q^f}, E)$.

Finally, using (4.14) and a similar argument that lead to (4.18) gives

$$\mathbb{E}[E] < \infty,$$

$$\mathbb{E}[Q_\ell] \leq \liminf \mathbb{E}[Q_\ell(t)] \leq \frac{c}{\delta} < \infty.$$

Therefore, the first moment of the stationary distribution of the Markov chain $\mathbf{H}(t)$ is finite. □

Appendix C - Proof of Theorem 6

The event that

$$||\mathbf{Q}(t_k + T_r + \tau) - \mathbf{Q}(t_k)|| \geq \xi_{HB} F(S(\mathbf{Q}(t_k))),$$

happens if a single queue size *changes* by an amount $M \doteq \xi_{HB} F(S(\mathbf{Q}(t_k)))$. Since we utilize the Max-Weight schedule w.r.t. to the queue lengths at time slot t_k, as we show next, there always exists a queue, say queue ℓ, with negative drift ($\mathbf{I}_\ell > \lambda_\ell$) and length larger than $\xi_{HB} F(S(\mathbf{Q}(t_k)))$ for some $0 < \xi_{HB} F(S(\mathbf{Q}(t_k))) < 1$. For $\boldsymbol{\lambda}$ strictly inside $\boldsymbol{\Lambda}$, there exists and ϵ such that $(1+\epsilon)\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$. Therefore, the Max-Weight schedule $\mathbf{I}^*(t_k) = [\mathbf{I}_1^*, ..., \mathbf{I}_L^*]$ satisfies

$$\sum_\ell Q_\ell(t_k)\mathbf{I}_\ell^* \geq \sum_\ell Q_\ell(t_k)\lambda_\ell + \epsilon \sum_\ell Q_\ell(t_k).$$

Therefore,

$$\sum_\ell Q_\ell(t_k)(\mathbf{I}_\ell^* - \lambda_\ell) \geq \epsilon \sum_\ell Q_\ell(t_k).$$

Since some queues might have negative contributions to the sum on the left hand side, we have

$$\sum_{\ell:\mathbf{I}_\ell^*>\lambda_\ell} Q_\ell(t_k)(\mathbf{I}_\ell^* - \lambda_\ell) \geq \epsilon \sum_\ell Q_\ell(t_k).$$

Let ℓ^* be the queue with the maximum contribution to the sum on the left hand side. We have

$$Q_\ell^*(t_k) \geq \frac{\epsilon}{D_{\ell^*}^* - \lambda_{\ell^*}} \sum_\ell Q_\ell(t_k),$$

where we let $\xi_{HB} \doteq \frac{\epsilon}{D_{\ell^*}^* - \lambda_{\ell^*}}$. We will use the time until $Q_{\ell^*}(t_k)$ decreases by M in order to upper bound χ_{HB}.

Lemma 13 (Upper Bound on Second Moment of χ_{HB})

$$\mathbb{E}[\chi_{HB}^2 | \mathbf{Q}(t_k)] \leq T_r^2 + c_1 F(S(\mathbf{Q}(t_k)))^2. \tag{4.21}$$

where c_1 is a constant.

Proof:

Let Y be the random variable denoting the time until $Q_{\ell^*}(t_k)$ decreases by $2M$. Note that we work with the decrease amount of $2M$ since the length of queue ℓ^* might increase during reconfiguration. We have

$$\chi_{HB} \leq_{st} Y.$$

In each time slot after reconfiguration, the $Q_{\ell^*}(t)$ decreases by D_{ℓ^*}. Therefore, the random variable Y satisfies

$$Y = \left\lceil \frac{2M}{D_{\ell^*}} \right\rceil + \sum_{\tau=0}^{\lceil Y/D_{\ell^*} \rceil} A_{\ell^*}(t_k + \tau).$$

This expression holds, since considering queue ℓ^* in isolation, the time until $Q_{\ell^*}(t_k)$ decreases by M is at least $\lceil 2M/D_{\ell^*} \rceil$ plus all the *descendants* of each of the $\lceil 2M/D_{\ell^*} \rceil$ new arrivals. To see this,

consider queue ℓ^* in isolation and note that since $D_{\ell^*} > \lambda_{\ell^*}$, queue ℓ^* gets empty with probability 1. Therefore, queue ℓ^* must decrease by M. Using the fact that the service rate is D_{ℓ^*}, we have the above equality for Y. Let $\hat{M} \doteq \lceil 2M/D_{\ell^*} \rceil$. Considering queue ℓ^* in isolation, the time to decrease by $\hat{M}$ can be expressed as the summation of $\hat{M}$ busy periods:

$$Y = \sum_{i=1}^{\hat{M}} B_i,$$

where B_i are i.i.d. random variables denoting *the time to decrease by D_{ℓ^*} for an isolated queue of i.i.d. arrival process $A_2(t)$ and constant service rate D_{ℓ^*}*. B_i have finite first and second moments since we have $\mathbb{E}[A_2(t)^2] \leq A_{\max}^2$ for all t. Let $\mathbb{E}[B_1^2|\mathbf{Q}(t_k)] \leq B^2$ which implies $\mathbb{E}[B_1|\mathbf{Q}(t_k)] \leq B$, where B is a constant independent of $\mathbf{Q}(t_k)$. Finally, conditional on $\mathbf{Q}(t_k)$, $\hat{M}$ is independent of the future arrivals $(A_2(t_k + \tau))_{\tau \geq 0}$. Therefore,

$$\mathbb{E}[\chi_{HB}^2|\mathbf{Q}(t_k)] \leq \mathbb{E}[Y^2|\mathbf{Q}(t_k)] = \mathbb{E}[\hat{M}|\mathbf{Q}(t_k)]\mathbb{E}[B_1^2|\mathbf{Q}(t_k)] + \mathbb{E}[\hat{M}(\hat{M}-1)|\mathbf{Q}(t_k)]\mathbb{E}[B_1|\mathbf{Q}(t_k)]^2,$$

where we used the independence of the random variables B_k in the last inequality. Simplifying,

$$\begin{aligned}
\mathbb{E}[\chi_{HB}^2|\mathbf{Q}(t_k)] &\leq \left\lceil \frac{2\xi_{HB}F(S(\mathbf{Q}(t_k)))}{D_{\ell^*}} \right\rceil B^2 + B^2 \left\lceil \frac{2\xi_{HB}F(S(\mathbf{Q}(t_k)))}{D_{\ell^*}} \right\rceil^2 \\
&\leq T_r^2 + c_1 F(S(\mathbf{Q}(t_k)))^2,
\end{aligned}$$

where c_1 is an appropriate constant. □

Lemma 14 (Lower Bound on First Moment of χ_{HB})

$$\mathbb{E}[\chi_{HB}|\mathbf{Q}(t_k)] \geq c_2(1 - \delta(\mathbf{Q}(t_k)))F(S(\mathbf{Q}(t_k))), \tag{4.22}$$

where δ is a decreasing function of $S(\mathbf{Q}(t_k))$.

Proof:

First consider the case where switching occurs right after the current reconfiguration interval, i.e., at time slot $t_k + T_r$. This event will be denoted by $\mathcal{A}_e$:

$$\mathcal{A}_e = \left\{ \omega : \sum_{\tau=0}^{T_r-1} \left[\sum_{\ell=1}^{L} A_\ell(t_k+\tau) \right] > \xi_{HB} F(S(\mathbf{Q}(t_k))) \right\}.$$

By Markov's inequality, the strong Markov property, and independence of the arrivals following switching,

$$\mathbb{P}\{\mathcal{A}_e\} \leq \frac{LT_r \sum_\ell \lambda_\ell}{\xi_{HB} F(S(\mathbf{Q}(t_k)))}$$

Letting $\delta \doteq LT_r \sum_\ell \lambda_\ell / \xi_{HB} F(S(\mathbf{Q}(t_k)))$, δ a decreasing function of $S(\mathbf{Q}(t_k))$, and it can be arbitrarily small for $\mathbf{Q}(t_k)$ outside a large enough compact set.

Now consider the case where switching does not occur immediately after reconfiguration. To obtain a lower bound we may consider the arrivals and departures separately and add their changes. This leads to an earlier stopping time, $T_E(w) < \chi_{HB}(w)$, strictly finite such that,

$$\sum_{\tau=0}^{T_E-1} \sum_{\ell} A(t_k+\tau) + \sum_{\tau=T_r}^{T_E-1} \sum_{\ell} D(t_k+\tau) \geq \xi_{HB} F(S(\mathbf{Q}(t_k)))$$

Note that since we are obtaining a lower bound, we may as well work with virtual as opposed to actual departures for the time during reconfiguration and for queues that empty. Applying optional stopping we find that [19],

$$\mathbb{E}\left[\sum_{\tau=0}^{T_E-1} \sum_{\ell} \left(A_\ell(\tau) + \mathbf{I}_\ell(\tau) \right) \right] = \mathbb{E}[T_E] \left(\sum_\ell \lambda_\ell + (\mathbb{E}[T_E] - T_r) \sum_\ell \mathbf{I}_\ell \right)$$

Therefore,

$$\mathbb{E}\left[\chi_{HB}\right] \geq \mathbb{E}\left[T_E\right] \geq \frac{\xi_{HB} F(S(\mathbf{Q}(t_k)))}{(\sum_\ell \lambda_\ell + \mathbf{I}_\ell)},$$

where the inequality follows from the fact that at stopping the process must exceed $\xi_{HB} F(S(\mathbf{Q}(t_k)))$ by definition. This is the required lower bound on $\mathbb{E}\left[\chi_{HB}|\mathbf{Q}(t_k)\right]$.

Now conditioning on $\chi_{HB} > T_r$ and $\chi_{HB} = T_r$ we have

$$\begin{aligned}\mathbb{E}[\chi_{HB}|\mathbf{Q}(t_k)] &= \mathbb{P}\{\chi_{HB} > T_r|\mathbf{Q}(t_k)\}\mathbb{E}[\chi_{HB}|\mathbf{Q}(t_k), \chi_{HB} > T_r] \\ &+ \mathbb{P}\{\chi_{HB} = T_r|\mathbf{Q}(t_k)\}\mathbb{E}[\chi_{HB}|\mathbf{Q}(t_k), \chi_{HB} = T_r].\end{aligned}$$

Finally, we obtain

$$\mathbb{E}[\chi_{HB}|\mathbf{Q}(t_k)] \geq (1-\delta)\frac{\xi_{HB} F(S(\mathbf{Q}(t_k)))}{(\sum_\ell \lambda_\ell + \mathbf{I}_\ell)},$$

where δ is a decreasing function of $S(\mathbf{Q}(t_k))$. □

The stability of the Bias-Based policies now follow from lemmas 13, 14 and Theorem 4.

Appendix D - Stability of the SCB Algorithm at Decision Epochs

Let χ_{SCB} be the stopping time associated with the stopping rule $\mathcal{E}_{SCB}$ in (4.4). We define $\mathcal{E}_{HB}$ such that $\chi_{HB}(w) < \chi_{SCB}(w)$ for all sample paths w, and we show that stability of the HB policy at the frame boundaries (frame-stability) implies the frame stability of the SCB policy using the fact that the schedule currently employed under the SCB policy is at most $F(S(\mathbf{Q}(t)))$ away from the

current Max-Weight schedule. Specifically, define the quadratic Lyapunov function:

$$L(\mathbf{Q}(t)) \doteq \sum_{i=\ell}^{L} Q_\ell^2(t) \tag{4.23}$$

Define the single-step drift:

$$\Delta_1(t) \doteq \mathbb{E}\left[L(\mathbf{Q}(t+1)) - L(\mathbf{Q}(t))\middle|\mathbf{Q}(t)\right].$$

We first show that for $\mathbf{Q}(t_k)$ outside a compact set, the single-step drift for the SCB policy is negative.

1-Step Drift

The following 1-step queue evolution expression holds for the SCB policy whenever the system is not in a reconfiguration interval.

$$Q_\ell(t+1) \leq \max\{Q_\ell(t) - \mathbf{I}_\ell(t), 0\} + A_\ell(t). \tag{4.24}$$

Squaring both sides, using $\max(0,x)^2 \leq x^2, \forall x \in \Re$, $\mathbf{I}_\ell(t) \leq \mu_{\max}, \forall t$, and $\max(Q_\ell(t) - \mathbf{I}_\ell(t), 0) \leq Q_\ell(t)$ we have

$$Q_\ell(t+1)^2 - Q_\ell(t)^2 \leq 1 + A_\ell(t)^2 - 2Q_\ell(t)(\mathbf{I}_\ell(t) - A_\ell(t)) \tag{4.25}$$

Summing over the queues and taking conditional expectations we have

$$\Delta_1(t) \leq LB + 2\sum_\ell Q_\ell(t)\lambda_\ell - 2\sum_\ell Q_\ell(t)\mathbb{E}\left[\mathbf{I}_\ell(t)\middle|\mathbf{Q}(t)\right]$$

where $B \doteq 1 + A_{\max}^2$. Now, for the SCB policy, the weight of the current schedule is at most

$F(S(\mathbf{Q}(t)))$ away from the weight of the Max-Weight schedule at time t, $\mathbf{I}^*(t)$

$$\sum_{\ell} Q_\ell(t)\mathbb{E}[\mathbf{I}_\ell(t)\big|\mathbf{Q}(t)] > \mathbf{Q}(t)\cdot\mathbf{I}^*(t) - F(S(\mathbf{Q}(t))). \tag{4.26}$$

For any arrival rate vector $\boldsymbol{\lambda}$ that is strictly inside $\boldsymbol{\Lambda}$, there exist real numbers $\alpha^1, ..., \alpha^{|\mathcal{I}|}$ such that $\alpha^j \geq 0,\ j \in 1, ..., |\mathcal{I}|, \sum_j \alpha^j = 1-\epsilon$ for some $\epsilon > 0$ and $\boldsymbol{\lambda} = \sum_j \alpha^j \mathbf{I}^j$ Therefore,

$$\begin{aligned}
\Delta_1(t) &\leq LB + 2\mathbf{Q}(t)\cdot(\sum_j \alpha^j\mathbf{I}^j) - 2\Big(\mathbf{Q}(t)\cdot\mathbf{I}^*(t) - F(S(\mathbf{Q}(t)))\Big) \\
&= L\Big(B + F(S(\mathbf{Q}(t)))\Big) + 2\mathbf{Q}(t)\cdot\mathbf{I}^*(t)(1-\epsilon) - 2\mathbf{Q}(t)\cdot\mathbf{I}^*(t) \\
&= L\Big(B + F(S(\mathbf{Q}(t)))\Big) - 2\epsilon\mathbf{Q}(t)\cdot\mathbf{I}^*(t) \\
&\leq L\Big(B + F(S(\mathbf{Q}(t)))\Big) - \frac{2\epsilon}{L}\sum_{\ell} Q_\ell(t) \\
&\leq -\frac{\epsilon}{L}\sum_{\ell} Q_\ell(t)
\end{aligned}$$

as long as $\mathbf{Q}(t)$ is outside a compact set $\mathcal{C}$. We can further bound right hand side by

$$\Delta_1(t) \leq -\frac{\epsilon}{L}\sum_{\ell\notin\mathbf{I}^*(t_k)} Q_\ell(t_k),$$

as long as $\mathbf{Q}(t)$ is outside a compact set $\mathcal{C}$, where for the last inequality we used the fact that the queues that do not belong to the current schedule can only increase their queue lengths. Therefore, the 1-step drift is negative during the whole frame (except reconfiguration interval) as long as the queue sizes are outside a compact set, denoted by $\mathcal{C}$.

We now show that the stopping time χ_{HB} happens before the stopping time χ_{SCB} if the queue lengths are outside a compact set. The intuition behind this result is that the stopping rule for χ_{SCB},

i.e., the event that the weight of some schedule becomes larger than the weight of the current schedule by an amount $F(S(\mathbf{Q}(t_k)))$, necessitates a *change* in the queue sizes of the order $F(S(\mathbf{Q}(t_k)))$.

Lemma 15 We have $\chi_{HB}(w) \leq \chi_{SCB}(w)$ for all sample paths w if $\mathbf{Q}(t_k)$ is outside a compact set $\mathcal{C}$.

This lemma is proved in Appendix E. As an example, if $F(||\mathbf{Q}(t_k)||)$ is chosen to be $\sqrt{||\mathbf{Q}(t_k)||}$, then the reverse condition that $\chi_{HB}(w) > \chi_{SCB}(w)$ leads to

$$(1 - (L\mu_{\max}\xi_{HB})^2)||\mathbf{Q}(t_k)|| < \xi_{HB}\sqrt{||\mathbf{Q}(t_k)||},$$

which is violated if $||\mathbf{Q}(t_k)|| > \xi_{HB}^2/(1 - (L\mu_{\max}\xi_{HB})^2)^2$.

Lemma 16 (Accumulation Lemma) Let $L_t \geq 0, \ \ t = 0, 1, 2, \cdots$ be a non-negative supermartingale with respect to the filtration $\mathcal{F}_t, \ \ t = 0, 1, 2, \cdots$, satisfying,

$$\mathbb{E}\left[L_t | \mathcal{F}_{t-1}\right] = L_{t-1} - \delta_t$$

where $\delta_t \geq 0$ is a nonnegative *previsible decrement* with (necessarily) finite expectation. Further let T be a stopping time which is finite a.s. then

$$\mathbb{E}\left[L_0\right] - \mathbb{E}\left[L_T\right] \geq \mathbb{E}\left[\sum_{t=1}^{T} \delta_t\right].$$

This lemma is proved in Appendix F. This result has a clear interpretation in gambling terms: δ_t represents a loss to the gambler which is known in advance. Lemma 16 then states that under any rule to stop play, the expected loss will be at least the sum of the expected previsible losses.

We apply this lemma to the $L(\mathbf{Q}(t))$ process for the SCB policy where $\delta_t \doteq \frac{\epsilon}{L}\sum_{\ell \notin \mathbf{I}^*(t_k)} Q_\ell(t_k)$. Considering the two stopping times for the HB and the SCB policies, χ_{HB} and χ_{SCB}, such that

$\chi_{HB}(w) \leq \chi_{SCB}(w)$, applying the Strong Accumulation Lemma starting from time $\chi_{HB}(w)$ instead of time 0 until the time $\chi_{SCB}(w)$ yields

$$\mathbb{E}\left[L(\mathbf{Q}(t_k + \chi_{SCB}))\right] \leq \mathbb{E}\left[L(\mathbf{Q}(t_k + \chi_{HB}))\right], \tag{4.27}$$

whenever $\mathbf{Q}(t_k)$ is outside the compact set $\mathcal{C}$ of Lemma 15. We established in Appendix C that χ_{HB} satisfies the conditions of Theorem 4. Therefore, the following expression, similar to (4.12) in the proof of Theorem 4, holds for χ_{HB}:

$$\mathbb{E}[L(\mathbf{Q}(t_k + \chi_{HB}))] - \mathbb{E}[L(\mathbf{Q}(t_k))] \leq C - \delta_2 \mathbb{E}\left[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))\right].$$

Therefore, (4.27) implies that

$$\mathbb{E}[L(\mathbf{Q}(t_k + \chi_{SCB}))] - \mathbb{E}[L(\mathbf{Q}(t_k))] \leq C - \delta_2 \mathbb{E}\left[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))\right].$$

Writing a similar expression over the frame boundaries $t_k, k \in \{0, 1, 2, ..., K\}$, summing them and telescoping these expressions leads to

$$L(\mathbf{Q}(t_K)) - L(\mathbf{Q}(0)) \leq KC - \delta_2 \sum_{k=0}^{K-1} \mathbb{E}[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))].$$

Using $L(\mathbf{Q}(t_K)) \geq 0$ and $L(\mathbf{Q}(\mathbf{0})) = 0$ we have

$$\frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}\left[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))\right] \leq \frac{C}{\delta_2} < \infty.$$

This implies that

$$\limsup_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}\left[F(S(\mathbf{Q}(t_k))S(\mathbf{Q}(t_k))\right] \leq \frac{C}{\delta_2} < \infty. \tag{4.28}$$

Furthermore, we have

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \sum_{\ell} \mathbb{E}[Q_\ell(t_k)] \leq \frac{C}{\delta_2} < \infty. \tag{4.29}$$

This establishes stability of the SCB policy at the frame boundaries $t_k, k \in \{0, 1, 2, ...\}$.

Appendix E - Proof of Lemma 15

Let $D_\ell(t)$ be the actual number of packets departing from queue ℓ (as opposed to service opportunities) at time slot t, and let $\mathbf{A}(t)$ and $\mathbf{D}(t)$ denote the vector of arrivals and departures at times slot t respectively. Fix a sample path w. We suppose that the stopping time χ_{SCB} has happened, and show that χ_{HB} must also have happened. Our supposition implies that

$$\begin{aligned}
\Phi &\doteq \mathbf{Q}(t) \cdot \mathbf{I}^*(t) - \mathbf{Q}(t) \cdot \mathbf{I}^*(t_k) \\
&= (\mathbf{Q}(t_k) + \sum_{\tau=0}^{\chi_{SCB}-1} (\mathbf{A}(t_k+\tau) - \mathbf{D}(t_k+\tau))) \cdot (\mathbf{I}^*(t) - \mathbf{I}^*(t_k)) \\
&\geq F(S(\mathbf{Q}(t))) \qquad (4.30)
\end{aligned}$$

Using $\mathbf{Q}(t) \cdot \mathbf{I}^*(t) \leq \mathbf{Q}(t) \cdot \mathbf{I}^*(t_k)$, we have from Schwartz's Inequality

$$\Phi \leq || \sum_{\tau=0}^{\chi_{SCB}-1} (\mathbf{A}(t_k+\tau) - \mathbf{D}(t_k+\tau))|| L\mu_{\max}.$$

Assume for contradiction that χ_{HB} has not happened. Then, we have

$$|| \sum_{\tau=0}^{\chi_{SCB}-1} (\mathbf{A}(t_k+\tau) - \mathbf{D}(t_k+\tau))|| \leq \xi_{HB} F(S(\mathbf{Q}(t_k))). \tag{4.31}$$

That is, because of our assumption that χ_{HB} has not happened, the stopping rule for χ_{HB} cannot have been satisfied at χ_{SCB}. Therefore,

$$\Phi \leq L\mu_{\max}\xi_{HB}F(S(\mathbf{Q}(t_k))). \tag{4.32}$$

Furthermore, from (4.30) we have

$$\begin{aligned}
\Phi &\geq F\left(||\mathbf{Q}(t_k)|| + ||\sum_{\tau=0}^{\chi_{SCB}-1}(\mathbf{A}(t_k+\tau) - \mathbf{D}(t_k+\tau))||\right) \\
&\geq F\left(||\mathbf{Q}(t_k)|| - ||\sum_{\tau=0}^{\chi_{SCB}-1}(\mathbf{A}(t_k+\tau) - \mathbf{D}(t_k+\tau))||\right) \\
&\geq F\Big(||\mathbf{Q}(t_k)|| - \xi_{HB}F(S(\mathbf{Q}(t_k)))\Big),
\end{aligned} \tag{4.33}$$

where we used (4.31) for the last inequality. Combining (4.32) and (4.33) we have

$$L\mu_{\max}\xi_{HB}F(||\mathbf{Q}(t_k)||) \geq F\Big(||\mathbf{Q}(t_k)|| - \xi_{HB}F(||\mathbf{Q}(t_k)||)\Big). \tag{4.34}$$

We choose ξ_{HB} such that $L\mu_{\max}\xi_{HB} < 1$. Then, for sufficiently large $||\mathbf{Q}(t_k)||$, (4.34) is violated, and we have a contradiction.

Appendix F - Proof of Lemma 16

Define $M_0 := L_0$

$$M_t \doteq L_t + \sum_{\tau=1}^{t}\delta_\tau$$

then $M_t \geq 0$ is a nonnegative supermartingale. This follows as

$$\begin{aligned}
\mathbb{E}\left[M_t|\mathcal{F}_{t-1}\right] &= \mathbb{E}\left[L_t|\mathcal{F}_{t-1}\right] + \mathbb{E}\left[\sum_{\tau=1}^{t} \delta_\tau|\mathcal{F}_{t-1}\right], \quad t \geq 1 \\
&\leq L_{t-1} + \sum_{\tau=1}^{t-1} \delta_\tau \\
&\leq M_{t-1}
\end{aligned}$$

where the sum is understood to be 0, if $t = 1$.

Since this is the case, as a consequence of Fatou's lemma, it follows that,

$$\mathbb{E}\left[M_0\right] - \mathbb{E}\left[M_T\right] \geq 0$$

which upon substitution implies

$$\mathbb{E}\left[L_0\right] - \mathbb{E}\left[L_T\right] \geq \mathbb{E}\left[\sum_{\tau=1}^{T} \delta_\tau\right].$$

Chapter 5

Dynamic Server Allocation over Time Varying Channels with Switching Delay

In the previous chapter, we studied the impact of switching delays for networks with static channels, i.e., networks with constant channel gains over time. Time variation in channel gains is a common property in wireless networks due to phenomenon such as multipath fading, shadowing etc. The topic of this chapter is to study the effect of server switching delays on throughput and delay performance of wireless uplinks and downlinks subject to time-varying channel gains as shown in Fig. 5-1. As compared to the previous chapters, the combination of time-varying channels and switching delays results in fundamental changes in system stability and calls for new scheduling algorithms. We show that the stability region changes as a function of the memory in the channel processes, and it is significantly reduced as compared to systems subject to solely either switching delay or time-varying channels. Furthermore, we show that throughput-optimal policies take a very different structure from previously proposed network scheduling algorithms such as the celebrated Max-Weight or Exhaustive policies.

More specifically, we consider a dynamic server allocation problem over parallel queues with *time-varying channels* and *server switching delay* between the queues. At each time slot, the server decides either to stay with the current queue or switch to another queue based on the current con-

nectivity and the queue length information. In the first part of the chapter we consider a two-queue system and develop fundamental insights for the problem. We first consider the case of memoryless (i.i.d.) channels where we characterize the stability region explicitly and show that simple Exhaustive type policies that ignore the current queue size and channel state information are throughput-optimal.

Next, we consider the Gilbert-Elliot channel model which is a commonly used model to abstract physical channels with memory [1], [52]. We develop a new methodology to characterize the stability region of the system using *state-action frequencies* which are steady-state solutions to an MDP formulation for the corresponding saturated system, and characterize the stability region explicitly in terms of the channel parameters. Using this state-action frequency approach, we develop a frame-based dynamic control (FBDC) policy and show that it is throughput-optimal asymptotically in the frame length. The FBDC policy is the only known policy to stabilize systems with randomly varying connectivity and switchover delay and it utilizes the state-action frequencies of the MDP formulation in a dynamic queuing system. Moreover, we develop a simple 1-Lookahead Myopic policy that provably achieves at least 90% of the stability region, and myopic policies with 2 and 3 lookahead that achieve more than 94% and 96% of the stability region respectively. Finally, we present simulation results suggesting that the myopic policies may be throughput-optimal and more delay efficient than the FBDC policy.

In the second part of the chapter we consider the model with arbitrary finite number of parallel queues. For memoryless (i.i.d.) channel processes, we explicitly characterize the stability region and the throughput-optimal policy. For channels with memory, we show that the stability region characterization in terms of state-action frequencies extends to the general case and establish a tight outer bound on the stability region and an upper bound on the sum-throughput explicitly in terms of the connectivity parameters. We quantify the *switching loss* in sum-throughput as compared to the system with no switchover delays and show that simple myopic policies achieve the sum-throughput upper bound in the corresponding saturated system. We also show that the throughput-optimality

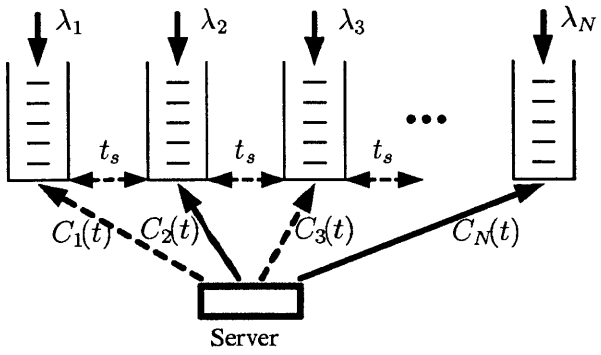


Figure 5-1: System model. Parallel queues with randomly varying connectivity processes $C_1(t), C_2(t), ..., C_N(t)$ and $t_s = 1$ slot switching time.

of the FBDC policy extend to the general case. In fact, the FBDC policy provides *a new framework for achieving throughput-optimal network control* by applying the state-action frequencies of the corresponding saturated system over frames in the dynamic queueing system. The FBDC policy is applicable to a broad class of systems whose corresponding saturated model is Markovian with a weakly communicating and finite state space, for example, systems with arbitrary switchover delays (i.e., systems that take any finite number of time slots for switching the server from a given queue to another queue) and general Markov modulated channel processes. Moreover, the framework of the FBDC policy can be utilized to achieve throughput-optimality in systems without switchover delay, for instance, in classical network control problems such as those considered in [87], [99], [110], [125].

5.0.1 Related Work

In the seminal paper [110], Tassiulas and Ephremides considered a parallel queuing system with randomly varying connectivity where they characterized the stability region of the system explicitly and proved the throughput-optimality of the Max-Weight (or the Longest-Connected-Queue First) scheduling policy. These results were later extended to more general systems in [86] and [87]. The effect of delayed channel state information was considered in [60, 91, 125], which showed that the stability region is reduced and that a policy similar to the Max-Weight algorithm is throughput-

optimal.

Perhaps the closest problem to ours is that of dynamic server allocation over parallel channels with randomly varying connectivity and limited channel sensing that has been investigated in [1, 73, 127] under the Gilbert-Elliot channel model. The saturated system was considered and the optimality of a myopic policy was established for a single server and two channels in [127], for arbitrary number of channels in [1], and for arbitrary number of channels and non-interfering servers in [2]. The problem of maximizing the throughput in the network while meeting average delay constraints for a small subset of users was considered in [85]. The average delay constraints were turned into penalty functions in [85] and the the theory of Stochastic Shortest Path problems, which is used for solving Dynamic Programs with certain special structures, was utilized to minimize the resulting drift+penalty terms. Finally, a partially observable Markov decision process (POMDP) model was used in [29] to analyze dynamic multichannel access in cognitive radio systems. These existing works do not consider the server switching delays.

Switching delay has been considered in Polling models in queuing theory community (e.g., [8, 72, 77, 114]), however, randomly varying connectivity was not considered since it may not arise in classical Polling applications. A detailed survey of the works in this field can be found in [114].

5.0.2 Main Contribution and Organization

The main contribution of this chapter is solving the scheduling problem in parallel queues with *time-varying channels* and *server switching delays* for the first time. For this, we provide a new framework for solving network control problems via characterizing the stability region in terms of state-action frequencies and achieving throughput-optimality by utilizing the state-action frequencies over frames.

This chapter is organized as follows. We consider the two-queue system in Section 5.1 where we characterize the stability region together with the throughput-optimal policy for memoryless channels. We develop the state-action frequency framework in Section 5.1.3 for channels with memory and use it to explicitly characterize the system stability region. We prove the throughput-

optimality of the FBDC policy in Section 5.1.4 and analyze simple myopic policies in Section 5.1.5. We extend our results to the general case in Section 5.2 where we also develop outer bounds on the stability region and an upper bound on the sum-throughput achieved by a simple Myopic policy. We present simulation results in Section 5.3 and conclude in Section 5.4.

5.1 Two-Queue System

5.1.1 System Model

Consider two parallel queues with time varying channels and one server receiving data packets from the queues. Time is slotted into unit-length time slots equal to one packet transmission time; $t \in \{0, 1, 2, ...\}$. It takes one slot for the server to switch from one queue to the other, and $m(t)$ denotes the queue at which the server is present at slot t. Let the i.i.d. stochastic process $A_i(t)$ with average arrival rate λ_i denote the number of packets arriving to queue i at time slot t, where $\mathbb{E}[A_i^2(t)] \leq A_{\max}^2$, $i \in \{1, 2\}$. Let $\mathbf{C}(t) = (C_1(t), C_2(t))$ be the channel (connectivity) process at time slot t, where $C_i(t) = 0$ for the OFF state (disconnected) and $C_i(t) = 1$ for the ON state (connected). We assume that the processes $A_1(t), A_2(t), C_1(t)$ and $C_2(t)$ are independent.

The process $C_i(t), i \in \{1, 2\}$, is assumed to form the two-state Markov chain with transition probabilities p_{10} and p_{01} as shown in Fig. 5-2, i.e., the Gilbert-Elliot channel model [1], [52], [73], [127], [131]. The Gilbert-Elliot Channel model has been commonly used in modeling and analysis of wireless channels with memory [1], [73], [119], [127], [131]. For ease of exposition, we present the analysis in this section for the symmetric Gilbert-Elliot channel model, i.e., $p_{10} = p_{01} = \epsilon$, and we state the corresponding results for the non-symmetric case in Appendix D. The steady state probability of each channel state is equal to 0.5 in the symmetric Gilbert-Elliot channel model. Moreover, for $\epsilon = 0.5$, $C_i(t) = 1$, w.p. 0.5, independently and identically distributed (i.i.d.) at each time slot. We refer to this case as the *memoryless channels* case.

Let $\mathbf{Q}(t) = (Q_1(t), Q_2(t))$ be the queue lengths at time slot t. We assume that $\mathbf{Q}(t)$ and $\mathbf{C}(t)$

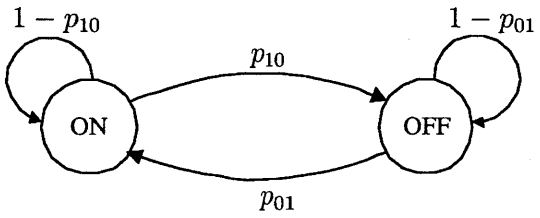


Figure 5-2: Markov modulated ON/OFF channel process. We have $p_{10} + p_{01} < 1$ ($\epsilon < 0.5$) for positive correlation.

are known to the server at the beginning of each time slot. Let $a_t \in \{0, 1\}$ denote the action taken at the beginning of slot t, where $a_t = 1$ if the server stays with the current queue and $a_t = 0$ if it switches to the other queue. One packet is successfully received from queue i at time slot t, if $m(t) = i, a_t = 1$ and $C_i(t) = 1$.

In the following, we start by explicitly characterizing the stability region for both memoryless channels and channels with memory and show that channel memory can be exploited to enlarge the stability region significantly.

5.1.2 Motivation: Channels Without Memory

In this section we assume that $\epsilon = 0.5$ so that the channel processes are i.i.d. over time. The stability region of the corresponding system with no-switching time was established in [110]: $\lambda_1, \lambda_2 \in [0, 0.5]$ and $\lambda_1 + \lambda_2 \leq 0.75$. Note that when the switching time is zero, the stability region is the same for both i.i.d. and Markovian channels, which is a special case of the results in [86]. However, when the switching time is non-zero, the stability region is reduced considerably:

Theorem 7 The stability region of the system with i.i.d. channels and one-slot switching delay is given by,

$$\mathbf{\Lambda} = \{(\lambda_1, \lambda_2) | \lambda_1 + \lambda_2 \leq 0.5, \lambda_1, \lambda_2 \geq 0\}. \tag{5.1}$$

In addition, the simple Exhaustive (Gated) policy is throughput-optimal.

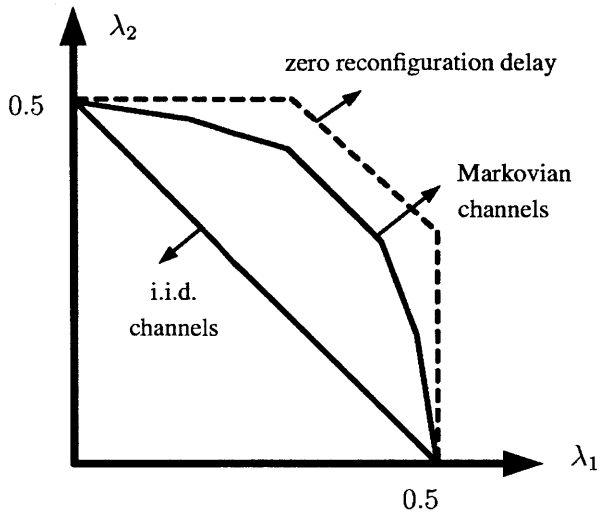


Figure 5-3: Stability region under memoryless (i.i.d.) channels and channels with memory (Markovian with $\epsilon < 0.5$) with and without switching delay.

The proof is given in Appendix A for a more general system. The basic idea behind the proof is that as soon as the server switches to queue i under some policy, the time to the ON state is a geometric random variable with mean 2 slots, independent of the policy. Therefore, a necessary condition for stability is given by the stability condition for a system without switching times and i.i.d. service times with geometric distribution of mean 2 slots as given by (5.1). The fact that the simple Gated policy is throughput-optimal follows from the observation that as the arrival rates are close to the boundary of the stability region, the fraction of time the server spends receiving packets dominates the fraction of time spent on switching [114].

As depicted in Fig. 5-3, the stability region of the system is considerably reduced for nonzero switching delay. Note that for systems in which channels are always connected, the stability region is given by $\lambda_1 + \lambda_2 \leq 1, \lambda_1, \lambda_2 \geq 0$ and is *not* affected by the switching delay [114]. Therefore, *it is the combination of switching delay and random connectivity that result in fundamental changes in system stability.*

Remark 1 *As shown in Appendix A, the results in this subsection can easily be generalized to the*

case of non-symmetric Gilbert-Elliot channels with arbitrary switching delays. For a system of 2 queues with arbitrary switching delays and i.i.d. channels with probabilities $p_i, i \in \{1,2\}$, $\mathbf{\Lambda}$ is the set of all $\boldsymbol{\lambda} \geq 0$ such that $\lambda_1/p_1 + \lambda_2/p_2 \leq 1$. Moreover, simple Exhaustive (Gated) policy is throughput-optimal.

When channel processes have memory, it is clear that one can achieve better throughput region than the i.i.d. channels case if the channels are positively correlated over time. This is because we can exploit the channel diversity when the channel states stay the same with high probability. In the following, we show that indeed the throughput region approaches the throughput region of no switching time case in in [110] as the channels become more correlated over time. Note that the throughput region in [110] is the same for both i.i.d. and Markovian channels under the condition that probability of ON state for the $i.i.d.$ channels is the same as the steady state probability of ON state for the two state Markovian channels. This fact can be derived as a special case of the seminal work of Neely in [86].

5.1.3 Channels With Memory - Stability Region

When switching times are non-zero, the memory in the channel can be exploited to improve the stability region considerably. Moreover, as $\epsilon \rightarrow 0$, the stability region tends to that achieved by the system with no-switching time and for $0 < \epsilon < 0.5$ it lies between the stability regions corresponding to the two extreme cases $\epsilon = 0.5$ and $\epsilon \rightarrow 0$ as shown in Fig. 5-3.

We start by analyzing the corresponding system with saturated queues, i.e., both queues are always non-empty. Let $\mathbf{\Lambda}_s$ denote the set of all time average expected departure rates that can be obtained from the two queues in the saturated system under all possible policies that are possibly history dependent, randomized and non-stationary. We will show that $\mathbf{\Lambda} = \mathbf{\Lambda}_s$. We prove the necessary stability conditions in the following Lemma and establish sufficiency in the next section.

Lemma 17 We have that

$$\mathbf{\Lambda} \subseteq \mathbf{\Lambda}_s.$$

This lemma is proved in Appendix B.

Now, we establish the region $\mathbf{\Lambda}_s$ by formulating the system dynamics as a Markov Decision Process (MDP). Let $\mathbf{s}_t = (m(t), C_1(t), C_2(t)) \in S$ denote the system state at time t where $\mathcal{S}$ is the set of all states. Also, let $a_t \in \mathcal{A} = \{0, 1\}$ denote the action taken at time slot t where $\mathcal{A}$ is the set of all actions at each state. Let $\mathbb{H}(t) = [\mathbf{s}_\tau]|_{\tau=0}^{t} \cup [a_\tau]|_{\tau=0}^{t-1}$ denote the full history of the system state until time t and let $\Upsilon(\mathcal{A})$ denote the set of all probability distributions on $\mathcal{A}$. For the saturated system, a policy is a mapping from the set of all possible past histories to $\Upsilon(\mathcal{A})$ [10], [79]. A policy is said to be *stationary* if, given a particular state, it applies the same decision rule in all stages and under a stationary policy, the process $\{\mathbf{s}_t; t \in \mathbb{N} \cup \{0\}\}$ forms a Markov chain. In each time slot t, the server observes the current state $\mathbf{s}_t$ and chooses an action a_t. Then the next state j is realized according to the transition probabilities $\mathbf{P}(j|\mathbf{s}, a)$, which depend on the random channel processes. Now, we define the reward functions as follows:

$$\overline{r}_1(\mathbf{s}, a) \doteq 1 \text{ if } \mathbf{s} = (1,1,1) \text{ or } \mathbf{s} = (1,1,0), \text{ and } a = 1 \tag{5.2}$$

$$\overline{r}_2(\mathbf{s}, a) \doteq 1 \text{ if } \mathbf{s} = (2,1,1) \text{ or } \mathbf{s} = (2,0,1), \text{ and } a = 1, \tag{5.3}$$

and $\overline{r}_1(\mathbf{s}, a) = \overline{r}_2(\mathbf{s}, a) \doteq 0$ otherwise. That is, a reward is obtained when the server stays at an ON channel. We are interested in the set of all possible time average expected departure rates, therefore, given some $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$, we define the system reward at time t by $\overline{r}(\mathbf{s}, a) \doteq \alpha_1 \overline{r}_1(\mathbf{s}, a) + \alpha_2 \overline{r}_2(\mathbf{s}, a)$. The average reward of policy π is defined as follows:

$$r^\pi \doteq \limsup_{K \to \infty} \frac{1}{K} E\Big\{ \sum_{t=1}^{K} \overline{r}(\mathbf{s}_t, a_t^\pi) \Big\}.$$

Given some $\alpha_1, \alpha_2 \geq 0$, we are interested in the policy that achieves the maximum time average expected reward $r^* \doteq \max_\pi r^\pi$. This optimization problem is a discrete time MDP characterized by the state transition probabilities $\mathbf{P}(j|\mathbf{s}, a)$ with 8 states and 2 actions per state. Furthermore,

any given pair of states are accessible from each other (i.e., there exists a positive probability path between the states) under some stationary-deterministic policy. Therefore this MDP belongs to the class of *Weakly Communicating* MDPs[1] [94].

The State-Action Frequency Approach

For Weakly Communicating MDPs with finite state and action spaces and bounded rewards, there exists an optimal stationary-deterministic policy, given as a solution to standard Bellman's equation, with optimal average reward independent of the initial state [94, Theorem 8.4.5]. This is because if a stationary policy has a nonconstant gain over initial states, one can construct another stationary policy with constant gain which dominates the former policy, which is possible since there exists a positive probability path between any two recurrent states under some stationary policy [79]. The *state-action frequency* approach, or the *Dual Linear Program (LP)* approach, given below provides a systematic and intuitive framework to solve such average cost MDPs, and it can be derived using Bellman's equation and the monotonicity property of Dynamic Programs [Section 8.8] [94]:

$$\text{Maximize} \quad \sum_{\mathbf{s}\in\mathcal{S}}\sum_{a\in\mathcal{A}} \overline{r}(\mathbf{s},a)\mathbf{x}(\mathbf{s},a) \tag{5.4}$$

subject to the balance equations

$$\mathbf{x}(\mathbf{s};1)+\mathbf{x}(\mathbf{s};0) = \sum_{\mathbf{s}'\in\mathcal{S}}\sum_{a\in\mathcal{A}} \mathbf{P}(\mathbf{s}|\mathbf{s}',a)\mathbf{x}(\mathbf{s}',a),\ \forall\, \mathbf{s}\in S, \tag{5.5}$$

the normalization condition

$$\sum_{\mathbf{s}\in\mathcal{S}} \mathbf{x}(\mathbf{s};1)+\mathbf{x}(\mathbf{s};0) = 1, \tag{5.6}$$

[1] In fact, other than the trivial suboptimal policy π_s that decides to stay with the current queue in all states, all stationary deterministic policies are unichain, namely, they have a single recurrent class regardless of the initial state. Hence, when π_s is excluded, we have a Unichain MDP.

and the nonnegativity constraints

$$\mathbf{x}(\mathbf{s}, a) \geq 0, \ \mathbf{s} \in \mathcal{S}, a \in \mathcal{A}. \tag{5.7}$$

The feasible region of this LP constitutes a polytope called the *state-action polytope* $\mathbf{X}$ and the elements of this polytope $\mathbf{x} \in \mathbf{X}$ are called state-action frequency vectors. Clearly, $\mathbf{X}$ is a convex, bounded and closed set. Note that $\mathbf{x}(\mathbf{s}; 1)$ can be interpreted as the stationary probability that action *stay* is taken at state $\mathbf{s}$. More precisely, a point $\mathbf{x} \in \mathbf{X}$ corresponds to a stationary randomized policy that takes action $a \in \{0, 1\}$ at state $\mathbf{s}$ w.p.

$$\mathbf{P}(\textit{action } a \text{ at state } \mathbf{s}) = \frac{\mathbf{x}(\mathbf{s}, a)}{\mathbf{x}(\mathbf{s}; 1) + \mathbf{x}(\mathbf{s}; 0)}, a \in A, \mathbf{s} \in S_x, \tag{5.8}$$

where S_x is the set of recurrent states given by $S_x \equiv \{\mathbf{s} \in S : \mathbf{x}(\mathbf{s}; 1) + \mathbf{x}(\mathbf{s}; 0) > 0\}$, and actions are arbitrary for transient states $\mathbf{s} \in S/S_x$ [79], [94].

Next we argue that the empirical state-action frequencies corresponding to any given policy (possibly randomized, non-stationary, or non-Markovian) lies in the state-action polytope $\mathbf{X}$. This ensures us that the optimal solution to the dual LP in (5.4) is over possibly non-stationary and history-dependent policies. In the following we give the precise definition and the properties of the set of empirical state-action frequencies. We define the *empirical* state-action frequencies $\hat{x}^T(\mathbf{s}, a)$ as

$$\hat{x}^T(\mathbf{s}, a) \doteq \frac{1}{T} \sum_{t=1}^{T} I_{\{\mathbf{s}_t = \mathbf{s}, a_t = a\}}, \tag{5.9}$$

where I_E is the indicator function of an event E, i.e., $I_E = 1$ if E occurs and $I_E = 0$ otherwise. Given a policy π, let P^π be the state-transition probabilities under the policy π and $\phi = (\phi_s)$ an initial state distribution with $\sum_{\mathbf{s} \in \mathcal{S}} \phi_{\mathbf{s}} = 1$. We let $x^T_{\pi,\phi}(\mathbf{s}, a)$ be the *expected empirical* state-action

frequencies under policy π and initial state distribution ϕ:

$$x^T_{\pi,\phi}(\mathrm{s},a) \doteq \mathbb{E}^{\pi,\phi}\left[\hat{x}^T(\mathrm{s},a)\right]$$

$$= \frac{1}{T}\sum_{t=1}^{T}\sum_{\mathrm{s}'\in\mathcal{S}}\phi_{\mathrm{s}'}P^{\pi}(\mathrm{s}_t=\mathrm{s},a_t=a|\mathrm{s}_0=\mathrm{s}').$$

We let $\mathbf{x}_{\pi,\phi} \in \Upsilon(\mathcal{S}\times\mathcal{A})$ (as in [79], [94]) be the limiting expected state-action frequency vector, if it exists, starting from an initial state distribution ϕ, under a general policy π (possibly randomized, non-stationary, or non-Markovian):

$$x_{\pi,\phi}(\mathrm{s},a) = \lim_{T\to\infty} x^T_{\pi,\phi}(\mathrm{s},a). \tag{5.10}$$

Let the set of all limit points be defined by

$$\mathbf{X}^{\phi}_{\Pi} \doteq \{x\in\Upsilon(\mathcal{S}\times\mathcal{A}) : \text{ there exists a policy } \pi \text{ s.t.}$$

$$\text{the limit in (5.10) exists and } \mathbf{x}=\mathbf{x}_{\pi,\phi}\}.$$

Similarly let $X^{\phi}_{\Pi'}$ denote the set of all limit points of a particular class of policies Π', starting from an initial state distribution ϕ. We let Π_{SD} denote the set of all stationary-deterministic policies and we let $co(\Xi)$ denote the closed convex hull of set Ξ. The following theorem establishes the equivalency between the set of all achievable limiting state-action frequencies and the state-action polytope:

Theorem 8 *[94, Theorem 8.9.3], [79, Theorem 3.1].* For any initial state distribution ϕ

$$co(\mathbf{X}^{\phi}_{\Pi_{SD}}) = \mathbf{X}^{\phi}_{\Pi} = \mathbf{X}.$$

We have $co(\mathbf{X}^{\phi}_{\Pi_{SD}}) \subseteq \mathbf{X}^{\phi}_{\Pi}$ since convex combinations of vectors in $\mathbf{X}^{\phi}_{\Pi_{SD}}$ correspond to limiting expected state-action frequencies for stationary-randomized policies, which can also be obtained

by time-sharing between stationary-deterministic policies. The inverse relation $co(\mathbf{X}^{\phi}_{\Pi_{SD}}) \supseteq \mathbf{X}^{\phi}_{\Pi}$ holds since for weakly communicating MDPs, there exists a stationary-deterministic optimal policy independent of the initial state distribution. Next, for any stationary-deterministic policy, the underlying Markov chain is stationary and therefore the limits $\mathbf{x}_{\pi,\phi}$ exists and satisfies the constraints (5.5), (5.6) and (5.7) of the polytope $\mathbf{X}$. Using $co(\mathbf{X}^{\phi}_{\Pi_{SD}}) = \mathbf{X}^{\phi}_{\Pi}$ and the convexity of $\mathbf{X}$ establishes $\mathbf{X}^{\phi}_{\Pi} \subseteq \mathbf{X}$. Furthermore, via (5.8), every $\mathbf{x} \in \mathbf{X}$ corresponds to a stationary-randomized policy for which the limits $\mathbf{x}_{\pi,\phi}$ exists, establishing $\mathbf{X}^{\phi}_{\Pi} \supseteq \mathbf{X}$.

Letting $\text{ext}(\mathbf{X})$ denote the set of extreme (corner) points of $\mathbf{X}$, an immediate corollary to Theorem 8 is as follows:

Corollary 3 *[79], [94].* For any initial state distribution ϕ

$$ext(\mathbf{X}) = \mathbf{X}^{\phi}_{\Pi_{SD}}.$$

The intuition behind this corollary is that if $\mathbf{x}$ is a corner point of $\mathbf{X}$, it cannot be expressed as a convex combination of any two other elements in $\mathbf{X}$, therefore, for each state s only one action has a nonzero probability.

Finally, we have that under any policy the probability of a large distance between the empirical expected state-action frequency vectors and the state-action polytope $\mathbf{X}$ decays exponentially fast in time. This result is similar to the mixing time of an underlying Markov chain to its steady state and we utilize such convergence results within the Lyapunov drift analysis for the dynamic queuing system in Section 5.1.4.

The Rate Polytope Λ_s

Using the theory on state-action polytopes in the previous section, we characterize the set of all achievable time-average expected rates in the saturated system, Λ_s. The following linear transfor-

mation of the state-action polytope $\mathbf{X}$ defines the 2 dimensional *rate polytope* [79]:

$$\mathbf{\Lambda}_s = \Big\{(r_1, r_2)\Big| \quad r_1 = \sum_{\mathbf{s}\in\mathcal{S}}\sum_{a\in\mathcal{A}} \mathbf{x}(\mathbf{s}, a)\overline{r}_1(\mathbf{s}, a)$$

$$r_2 = \sum_{\mathbf{s}\in\mathcal{S}}\sum_{a\in\mathcal{A}} \mathbf{x}(\mathbf{s}, a)\overline{r}_2(\mathbf{s}, a),\ \mathbf{x}\in\mathbf{X}\Big\},$$

where $\overline{r}_1(\mathbf{s}, a)$ and $\overline{r}_2(\mathbf{s}, a)$ are the reward functions defined in (5.2) and (5.3). This polytope is the set of all time average expected departure rate pairs that can be obtained in the saturated system, i.e., it is the rate region $\mathbf{\Lambda}_s$. An explicit way of characterizing $\mathbf{\Lambda}_s$ is given in Algorithm 6.

Algorithm 6 *Stability Region Characterization*

1: Given $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$ solve the following Linear Program (LP)

$$\begin{aligned} \max. \quad & \alpha_1 r_1(\mathbf{x}) + \alpha_2 r_2(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x}\in\mathbf{X}. \end{aligned} \tag{5.11}$$

2: For a given α_2/α_1 ratio, there exists an optimal solution (r_1^*, r_2^*) of the LP in (5.11) at a corner point of $\mathbf{\Lambda_s}$. Find all possible corner points and take their convex combination.

The fundamental theorem of Linear Programming guarantees that an optimal solution of the LP in (5.11) lies at a corner (extreme) point of the polytope $\mathbf{X}$ [15]. Furthermore, the one-to-one correspondence between the extreme points of the polytope $\mathbf{X}$ and stationary-deterministic polices stated in Corollary 3 is useful for finding the solutions of the above LP for all possible α_2/α_1 ratios. Namely, there are a total of 2^8 stationary-deterministic policies since we have 8 states and 2 actions per state and finding the rate pairs corresponding to these 256 stationary-deterministic policies and taking their convex combination gives $\mathbf{\Lambda}_s$. Fortunately, we do not have to go through this tedious procedure. The fact that at a vertex of (5.11) either $\mathbf{x}(\mathbf{s}; 1)$ or $\mathbf{x}(\mathbf{s}; 0)$ has to be zero for each $\mathbf{s}\in\mathbf{S}$ provides a useful guideline for analytically solving this LP. The following theorem characterizes the stability region explicitly. It shows that the stability region enlarges as the channel has more memory

and that there is a critical value of the channel correlation parameter given by $\epsilon_c \doteq 1 - \sqrt{2}/2$ at which the structure of the stability region changes.

Theorem 9 The rate region $\mathbf{\Lambda}_s$ is the set of all rates $r_1 \geq 0$, $r_2 \geq 0$ that for $\epsilon < \epsilon_c$ satisfy

$$\begin{aligned}
\epsilon r_1 + (1-\epsilon)^2 r_2 &\leq \frac{(1-\epsilon)^2}{2} \\
(1-\epsilon) r_1 + (1+\epsilon-\epsilon^2) r_2 &\leq \frac{3}{4} - \frac{\epsilon}{2} \\
r_1 + r_2 &\leq \frac{3}{4} - \frac{\epsilon}{2} \\
(1+\epsilon-\epsilon^2) r_1 + (1-\epsilon) r_2 &\leq \frac{3}{4} - \frac{\epsilon}{2} \\
(1-\epsilon)^2 r_1 + \epsilon r_2 &\leq \frac{(1-\epsilon)^2}{2},
\end{aligned}$$

and for $\epsilon \geq \epsilon_c$ satisfy

$$\begin{aligned}
r_1 + (1-\epsilon)(3-2\epsilon) r_2 &\leq \frac{(1-\epsilon)(3-2\epsilon)}{2} \\
r_1 + r_2 &\leq \frac{3}{4} - \frac{\epsilon}{2} \\
(1-\epsilon)(3-2\epsilon) r_1 + r_2 &\leq \frac{(1-\epsilon)(3-2\epsilon)}{2}.
\end{aligned}$$

The proof of the theorem is given in Appendix C and it is based on solving the LP in (5.11) for all weights α_1 and α_2 to find the corner points of $\mathbf{\Lambda}_s$, and then applying Algorithm 6. The following observation follows from Theorem 9.

Observation 2 The maximum achievable sum-rate in the saturated system is given by

$$r_1 + r_2 = \frac{3}{4} - \frac{\epsilon}{2}.$$

Note that $r_1 + r_2 \leq \frac{3}{4}$ is the boundary of the stability region for the system without switching delay analyzed in [110], where the probability that at least 1 channel is in ON state is $3/4$. Therefore, $\epsilon/2$ is the *throughput loss due to the 1 slot switching delay*. This throughput loss corresponds to the probability that the server is at a queue with an OFF state when the other queue is in an ON state.

The stability regions for the two ranges of ϵ are displayed in Fig. 5-4 (a) and (b). As $\epsilon \to 0.5$, the stability region converges to that of the i.i.d. channels with ON probability equal to 0.5. In this regime, knowledge of the current channel state is of no value. As $\epsilon \to 0$ the stability region converges to that for the system with no-switching time in [110]. In this regime, the channels are likely to stay the same for many consecutive time slots, therefore, the effect of switching delay is negligible.

The rate region $\mathbf{\Lambda}_s$ for the case of non-symmetric Gilbert-Elliot channels is given in Appendix D.

Remark 2 *The stability region characterization in terms of state-action frequencies is general. For instance,* this technique can be used to establish the stability regions of systems with more than two queues, arbitrary switching times, and more complicated Markovian channel processes. *Of course, explicit characterization as in Theorem 9 may not always be possible.*

5.1.4 Frame Based Dynamic Control (FBDC) Policy

We propose a frame-based dynamic control (FBDC) policy inspired by the characterization of the stability region in terms of state-action frequencies and prove that it is throughput-optimal asymptotically in the frame length. The motivation behind the FBDC policy is that a policy π^* that achieves the optimization in (5.11) for given weights α_1 and α_2 for the saturated system should achieve a *good* performance in the original system when the queue sizes Q_1 and Q_2 are used as weights. This is because first, the policy π^* will lead to similar average departure rates in both systems for sufficiently high queue sizes, and second, the usage of queue sizes as weights creates self adjusting policies that capture the dynamic changes due to stochastic arrivals similar to Max-Weight schedul-

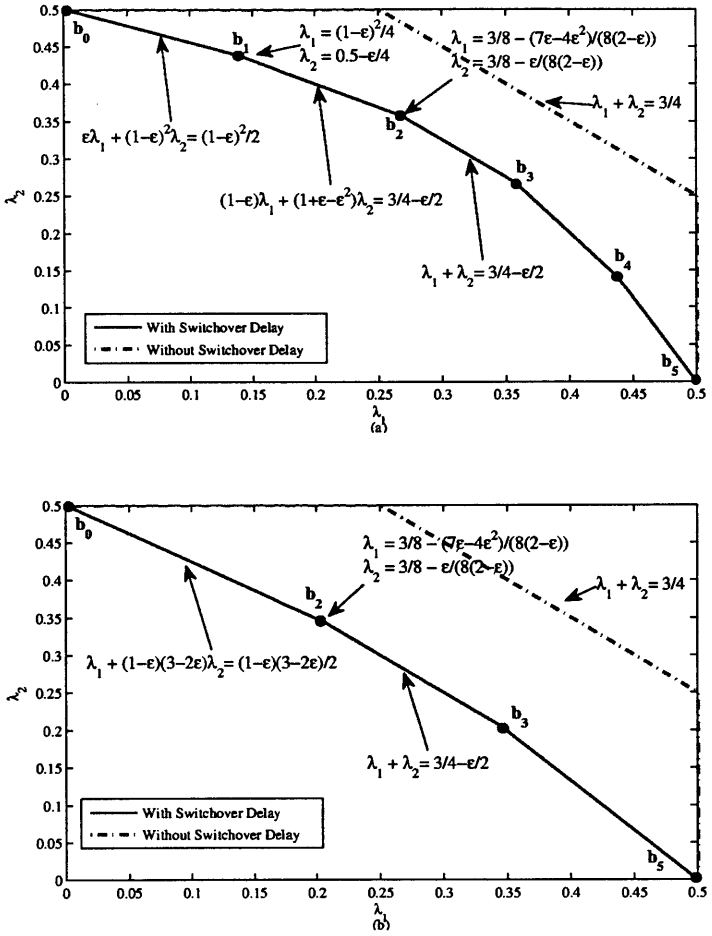


Figure 5-4: Stability region under channels with memory, with and without switching delay for (a) $\epsilon = 0.25 < \epsilon_c$ and (b) $\epsilon = 0.40 \geq \epsilon_c$.

ing in [109]. Specifically, we divide the time into equal-size intervals of T slots and let $Q_1(jT)$ and $Q_2(jT)$ be the queue lengths at the beginning of the jth interval. We find the deterministic policy that optimally solves (5.11) when $Q_1(jT)$ and $Q_2(jT)$ are used as weights and then apply this policy in each time slot of the frame. The FBDC policy is described in Algorithm 7 in details.

Algorithm 7 FRAME BASED DYNAMIC CONTROL (FBDC) POLICY

1: Find the policy π^* that optimally solves the following LP

$$\begin{aligned} \max._{\{r_1,r_2\}} \quad & Q_1(jT)r_1 + Q_2(jT)r_2 \\ \text{subject to} \quad & (r_1, r_2) \in \mathbf{\Lambda}_s \end{aligned} \tag{5.12}$$

where $\mathbf{\Lambda}_s$ is the rate polytope derived in Section 5.1.3.

2: Apply π^* in each time slot of the frame.

There exists an optimal solution (r_1^*, r_2^*) of the LP in (5.12) that is a corner point of $\mathbf{\Lambda}_s$ [15] and the policy π^* that corresponds to this point is a stationary-deterministic policy by Corollary 3.

Theorem 10 For any $\delta > 0$, there exists a large enough frame length T such that the FBDC policy stabilizes the system for all arrival rates within the δ-stripped stability region $\mathbf{\Lambda}_s^\delta = \mathbf{\Lambda}_s - \delta\mathbf{1}$.

An immediate corollary to this theorem is as follows:

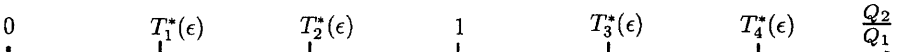
Corollary 4 The FBDC policy is throughput-optimal asymptotically in the frame length.

The proof of Theorem 10 follows from the proof of the FBDC policy for a more general system given in Appendix E. It performs a drift analysis using the standard quadratic Lyapunov function. However, it is novel in utilizing the state-action frequency framework of MDP theory within the Lyapunov drift arguments. The basic idea is that, for sufficiently large queue lengths, when the optimal policy solving (5.12), π^*, is applied over a sufficiently long frame of T slots, the average output rates of both the actual system and the corresponding saturated system converge to $\mathbf{r}^*$. For the saturated system, the probability of a large difference between empirical and steady state rates

decreases exponentially fast in T [79], similar to the convergence of a positive recurrent Markov chain to its steady state. Therefore, for sufficiently large queue lengths, the difference between the empirical rates in the actual system and r^* also decreases with T. This ultimately results in a negative Lyapunov drift when $\boldsymbol{\lambda}$ is inside the $\delta(T)$-stripped stability region since from (5.12) we have $Q_1(jT)r_1^* + Q_2(jT)r_2^* > Q_1(jT)\lambda_1 + Q_2(jT)\lambda_2$.

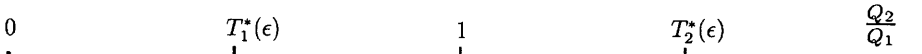
The FBDC policy is easy to implement since it does not require the arrival rate information for stabilizing the system for arrival rates in $\boldsymbol{\Lambda} - \delta(T)\mathbf{1}$, and it does not require the solution of the LP (5.12) for each frame. Instead, one can solve the LP (5.12) for all possible (Q_1, Q_2) pairs only *once* in advance and create a mapping from (Q_1, Q_2) pairs to the corners of the stability region. Then, this mapping can be used to find the corresponding optimal saturated-system policy to be applied during each frame. Solving the LP in (5.12) for all possible (Q_1, Q_2) pairs is possible because first, the solution of the LP will be one of the corner points of the stability region in Fig. 5-4, and second, the weights (Q_1, Q_2), which are the inputs to the LP, determine which corner point is optimal. The theory of Linear Programming suggests that the solution to the LP in (5.12) depends only on the relative value of the weights (Q_1, Q_2) with respect to each other. Namely, changing the queue size ratio Q_2/Q_1 varies the slope of the objective function of the LP in (5.12), and the value of this slope Q_2/Q_1 with respect to the slopes of the lines in the stability region in Fig. 5-4 determine which corner point the FBDC policy operates on. These mappings from the queue size ratios to the corners of the stability region are shown in Table 5.1 for the case of $\epsilon < \epsilon_c$ and in Table 5.2 for the case of $\epsilon \geq \epsilon_c$. The corresponding mappings for the FBDC policy for the case of non-symmetric Gilbert-Elliot channels are shown in Appendix D. Given these tables, one no longer needs to solve the LP (5.12) for each frame, but just has to perform a simple table look-up to determine the optimal policy to use in each frame.

In the next subsection we provide an upper bound to the long-run packet-average delay under the FBDC policy, which is linear in T. This suggest that the packet delay increases with increasing frame lengths as expected. However, such increases are at most linear in T. Note that the FBDC



corner b_5	corner b_4	corner b_3	corner b_2	corner b_1	corner b_0
(1,1,1): stay	(1,1,1): stay	(1,1,1): stay	(1,1,1): stay	(1,1,1): switch	(1,1,1): switch
(1,1,0): stay	(1,1,0): stay	(1,1,0): stay	(1,1,0): stay	(1,1,0): stay	(1,1,0): switch
(1,0,1): stay	(1,0,1): switch	(1,0,1): switch	(1,0,1): switch	(1,0,1): switch	(1,0,1): switch
(1,0,0): stay	(1,0,0): stay	(1,0,0): stay	(1,0,0): switch	(1,0,0): switch	(1,0,0): switch
(2,1,1): switch	(2,1,1): switch	(2,1,1): stay	(2,1,1): stay	(2,1,1): stay	(2,1,1): stay
(2,1,0): switch	(2,1,0): switch	(2,1,0): switch	(2,1,0): switch	(2,1,0): switch	(2,1,0): stay
(2,0,1): switch	(2,0,1): stay	(2,0,1): stay	(2,0,1): stay	(2,0,1): stay	(2,0,1): stay
(2,0,0): switch	(2,0,0): switch	(2,0,0): switch	(2,0,0): stay	(2,0,0): stay	(2,0,0): stay

Table 5.1: FBDC policy mapping from the queue sizes to the corners of $\mathbf{\Lambda_s}$, $b_0, b_1, b_2, b_3, b_4, b_5$ shown in Fig. 5-4 (a), for $\epsilon < \epsilon_c$. For each state $\mathbf{s} = (m(t), C_1(t), C_2(t))$ the optimal action is specified. The thresholds on Q_2/Q_1 are $0, T_1^* = \epsilon/(1-\epsilon)^2, T_2^* = (1-\epsilon)/(1+\epsilon-\epsilon^2), 1, T_3^* = (1+\epsilon-\epsilon^2)/(1-\epsilon), T_4^* = (1-\epsilon)^2/\epsilon$.



corner b_3	corner b_2	corner b_1	corner b_0
(1,1,1): stay	(1,1,1): stay	(1,1,1): stay	(1,1,1): switch
(1,1,0): stay	(1,1,0): stay	(1,1,0): stay	(1,1,0): switch
(1,0,1): stay	(1,0,1): switch	(1,0,1): switch	(1,0,1): switch
(1,0,0): stay	(1,0,0): stay	(1,0,0): switch	(1,0,0): switch
(2,1,1): switch	(2,1,1): stay	(2,1,1): stay	(2,1,1): stay
(2,1,0): switch	(2,1,0): switch	(2,1,0): switch	(2,1,0): stay
(2,0,1): switch	(2,0,1): stay	(2,0,1): stay	(2,0,1): stay
(2,0,0): switch	(2,0,0): switch	(2,0,0): stay	(2,0,0): stay

Table 5.2: FBDC policy mapping from the queue sizes to the corners of $\mathbf{\Lambda_s}$, b_0, b_1, b_2, b_3 shown in Fig. 5-4 (b), for $\epsilon \geq \epsilon_c$. For each state $\mathbf{s} = (m(t), C_1(t), C_2(t))$ the optimal action is specified. The thresholds on Q_2/Q_1 are $0, T_1^* = 1/((1-\epsilon)(3-2\epsilon)), 1, T_2^* = (1-\epsilon)(3-2\epsilon)$.

policy can also be implemented without any frames by setting $T = 1$, i.e., by solving the LP in Algorithm 7 in each time slot. The simulation results in Section 5.3 suggest that the FBDC policy implemented without frames has a similar throughput performance to the original FBDC policy. This is because for large queue lengths, the optimal solution of the LP in (5.12) depends on the queue length ratios, and hence, the policy π^* that solves the LP optimally does not change fast when the queue lengths get large. When the policy is implemented without the use of frames, it becomes more adaptive to dynamic changes in the queue lengths, which results in a better delay performance than the frame-based implementations.

Delay Upper Bound

The delay upper bound in this section is easily derived once the stability of the FBDC algorithm is established. The stability proof utilizes the following quadratic Lyapunov function

$$L(\mathbf{Q}(t)) = \sum_{i=1}^{2} Q_i^2(t),$$

which represents a quadratic measure of the total load in the system at time slot t. Let t_k denote the time slots at the frame boundaries, $k = 0, 1, ...$, and define the T-step conditional drift

$$\Delta_T(t_k) \triangleq \mathbb{E}\left[L(\mathbf{Q}(t_k + T)) - L(\mathbf{Q}(t_k))\middle|\mathbf{Q}(t_k)\right],$$

The following drift expression follows from the stability analysis in Appendix E:

$$\frac{\Delta_T(t_k)}{2T} \leq BT - \Big(\sum_i Q_i(t_k)\Big)\xi,$$

where $B = 1 + A_{\max}^2$, $\boldsymbol{\lambda}$ is strictly inside the *δ-stripped* stability region $\boldsymbol{\Lambda} - \delta\mathbf{1}$, and $\xi > 0$ represents a measure of the distance of $\boldsymbol{\lambda}$ to the boundary of $\boldsymbol{\Lambda} - \delta\mathbf{1}$. Taking expectations with respect to $\mathbf{Q}(t_k)$,

writing a similar expression over the frame boundaries $t_k, k \in \{0, 1, 2, ..., K\}$, summing them and telescoping these expressions lead to

$$\mathbb{E}[L(\mathbf{Q}(t_K))] - \mathbb{E}[L(\mathbf{Q}(0))] \leq 2KBT^2 - 2\xi T \sum_{k=0}^{K-1} \mathbb{E}\left[\sum_i Q_i(t_k)\right].$$

Using $L(\mathbf{Q}(t_K)) \geq 0$ and $L(\mathbf{Q}(0)) = 0$, we have

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \sum_i \mathbb{E}[Q_i(t_k)] \leq \frac{BT}{\xi}.$$

For $t \in (t_k, t_{k+1})$ we have $Q_i(t) \leq Q_i(t_k) + \sum_{\tau=0}^{T-1} A_i(t_k+\tau)$. Therefore, $\mathbb{E}[Q_i(t)] \leq \mathbb{E}[Q_i(t_k)] + T\lambda_i \leq \mathbb{E}[Q_i(t_k)] + TA_{\max}$. Therefore, for $T_K \doteq KT$ we have

$$\limsup_{T_K\to\infty} \frac{1}{KT} \sum_{t=0}^{KT-1} \sum_i \mathbb{E}[Q_i(t)]$$

$$\leq \limsup_{K\to\infty} \frac{1}{TK} \sum_{k=0}^{K-1} \sum_i T\mathbb{E}[Q_i(t_k)] + T^2 A_{\max} \leq \frac{(B + A_{\max}\xi)T}{\xi}.$$

Dividing by the total arrival rate into the system $\sum_i \lambda_i$ and applying Little's law, the average delay is upper bounded by an expression that is linear in the frame length T.

In the next section we consider Myopic policies that do not require the solution of an LP and that are able to stabilize the network for arrival rates within over 90% of the stability region. Simulation results in Section 5.3 suggest that the Myopic policies may in fact achieve the full stability region while providing better delay performance than the FBDC policy for most arrival rates.

5.1.5 Myopic Control Policies

We investigate the performance of simple *Myopic* policies that make scheduling/switching decisions according to weight functions that are products of the queue lengths and the channel predictions for a small number of slots into the future. We refer to a Myopic policy considering k future time slots as the k-Lookahead Myopic policy. We implement these policies over frames of length T time slots where during the jth frame, the queue lengths at the beginning of the frame, $Q_1(jT)$ and $Q_2(jT)$, are used for weight calculations during the frame. Specifically, in the One-Lookahead Myopic policy, assuming that the server is with queue 1 at some $t \in \{jT, ..., j(T+1)-1\}$, the weight of queue 1 is the product of $Q_1(jT)$ and the summation of the current state of the channel process C_1 and the probability that C_1 will be in the ON state at $t+1$. The weight of queue 2 is calculated similarly, however, the current state of the channel process C_2 is not included in the weight since queue 2 is not available to the server in the current time slot. The detailed description of the One-Lookahead Myopic policy is given in Algorithm 8 below.

Algorithm 8 ONE-LOOKAHEAD MYOPIC POLICY

1: Assuming that the server is currently with queue 1 and the system is at the jth frame, calculate the following weights in each time slot of the current frame;

$$W_1(t) = Q_1(jT)\Big(C_1(t) + \mathbb{E}\big[C_1(t+1)|C_1(t)\big]\Big)$$

$$W_2(t) = Q_2(jT)\mathbb{E}\big[C_2(t+1)|C_2(t)\big]. \tag{5.13}$$

2: If $W_1(t) \geq W_2(t)$ stay with queue 1, otherwise, switch to the other queue. A similar rule applies for queue 2.

Next, we establish a lower bound on the stability region of the One-Lookahead Myopic Policy by comparing its drift over a frame to the drift of the FBDC policy.

Theorem 11 The One-Lookahead Myopic policy achieves at least γ-fraction of the stability region $\mathbf{\Lambda_s}$ asymptotically in T where $\gamma \geq 90\%$.

The proof is constructive and will be establish in various steps in the following. The basic idea behind the proof is that the One-Lookahead Myopic (OLM) policy produces a mapping from the set of queue sizes to the stationary deterministic policies corresponding to the corners of the stability region. This mapping is similar to that of the FBDC policy, however, the thresholds on the queue size ratios Q_2/Q_1 are determined according to (5.13):

Mapping from queue sizes to actions. Case-1: $\epsilon < \epsilon_c$

For $\epsilon < \epsilon_c$, there are 6 corners in the stability region denoted by $b_0, b_1, ..., b_5$ where b_0 is $(0, 0.5)$ and b_5 is $(0.5, 0)$ as shown in Fig. 5-4 (a). We derive conditions on Q_2/Q_1 such that the OLM policy chooses the stationary deterministic decisions that correspond to a given corner point.

Corner b_0:

Optimal actions are to stay at queue 2 for every channel condition. Therefore, the server chooses queue 2 even when the channel state is $C_1(t), C_2(t) = (1, 0)$. Therefore, using (5.13), for the Myopic policy to take the deterministic actions corresponding to b_0 we need

$$Q_1.(1-\epsilon) < Q_2.(\epsilon) \Rightarrow \frac{Q_2}{Q_1} > \frac{1-\epsilon}{\epsilon}.$$

This means that if we apply the Myopic policy with coefficients Q_1, Q_2 such that $Q_2/Q_1 > (1-\epsilon)/\epsilon$, then the system output rate will be driven towards the corner point b_0 (both in the saturated system or in the actual system with large enough arrival rates).

Corner b_1:

The optimal actions for the corner point b_1 are as follows: At queue 1, for the channel state 10:stay, for the channel states 11, 01 and 00: switch. At queue 2, for the channel state 10: switch, for the channel states 11, 01 and 00: stay. The most limiting conditions are 11 at queue 1 and 10 at queue

2. Therefore we need, $Q_1(2-\epsilon) < Q_2(1-\epsilon)$ and $Q_1(1-\epsilon) > Q_2\epsilon$. Combining these we have

$$\frac{2-\epsilon}{1-\epsilon} < \frac{Q_2}{Q_1} < \frac{1-\epsilon}{\epsilon}.$$

Note that the condition $\epsilon < \epsilon_c = 1 - \sqrt{2}/2$ implies that $\frac{1-\epsilon}{\epsilon} > \frac{2-\epsilon}{1-\epsilon}$.

Corner b_2:

The optimal actions for the corner point b_1 are as follows: At queue 1, for the channel state 10 and 11:stay, for the channel states 01 and 00: switch. At queue 2, for the channel states 10: switch, for the channel states 11, 01 and 00: stay. The most limiting conditions are 11 at queue 1 and 00. Therefore we need, $Q_1(2-\epsilon) > Q_2(1-\epsilon)$ and $Q_1 < Q_2$. Combining these we have

$$1 < \frac{Q_2}{Q_1} < \frac{2-\epsilon}{1-\epsilon}.$$

The conditions for the rest of the corners are symmetric and can be found similarly to obtain the mapping in Fig. 5.3.

Mapping from queue sizes to actions. Case-2: $\epsilon \geq \epsilon_c$

In this case there are 4 corner points in the throughput region. We enumerate these corners as b_0, b_2, b_3, b_5 where b_0 is $(0, 0.5)$ and b_5 is $(0.5, 0)$.

Corner b_0:

The analysis is the same as the b_0 analysis in the previous case and we obtain that for the Myopic policy to take the deterministic actions corresponding to b_0 we need

$$\frac{Q_2}{Q_1} > \frac{1-\epsilon}{\epsilon}.$$

Corner b_2:

This is the same corner point as in the previous case corresponding to the same deterministic policy: At queue 1, for the channel state 10 and 11:stay, for the channel states 01 and 00: switch. At queue 2, for the channel states 10: switch, for the channel states 11, 01 and 00: stay. The most limiting conditions are 10 at queue 2 (since $\epsilon \geq \epsilon_c$ we have $\frac{1-\epsilon}{\epsilon} < \frac{2-\epsilon}{1-\epsilon}$) and 00. Therefore we need, $Q_1(1-\epsilon) > Q_2\epsilon$ and $Q_1 < Q_2$. Combining these we have

$$1 < \frac{Q_2}{Q_1} < \frac{1-\epsilon}{\epsilon}.$$

The conditions for the rest of the corners are symmetric and can be found similarly to obtain the mapping in Fig. 5.4 for $\epsilon \geq \epsilon_c$.

The conditions for the corners b_2 and b_3 are symmetric, completing the mapping from the queue sizes to the corners of $\mathbf{\Lambda_s}$ for $\epsilon \geq \epsilon_c$ shown in Table 5.4. This mapping is in general different from the corresponding mapping of the FBDC policy in Table 5.2. Therefore, for a given ratio of the queue sizes Q_2/Q_1, the FBDC and the OLM policies *may* apply different stationary deterministic policies corresponding to different corner points of $\mathbf{\Lambda_s}$, denoted by $\mathbf{r}^*$ and $\hat{\mathbf{r}}$ respectively. The shaded intervals of Q_2/Q_1 in Table 5.4 are the intervals in which the OLM and the FBDC policies apply different policies. A similar mapping can be obtained for the OLM policy for $\epsilon < \epsilon_c$. The corresponding mapping for the OLM policy for the case of non-symmetric Gilbert-Elliot channels is given in Appendix D.

The following lemma is proved in Appendix F and completes the proof by establishing the 90% bound on the weighted average departure rate of the OLM policy w.r.t. to that of the FBDC policy.

Lemma 18 We have that

$$\Psi \doteq \frac{\sum_i Q_i(t)\hat{r}_i}{\sum_i Q_i(t)r_i^*} \geq 90\%. \tag{5.14}$$

Furthermore, $\Psi \geq 90\%$ is a sufficient condition for the OLM policy to achieve at least 90% of $\mathbf{\Lambda_s}$

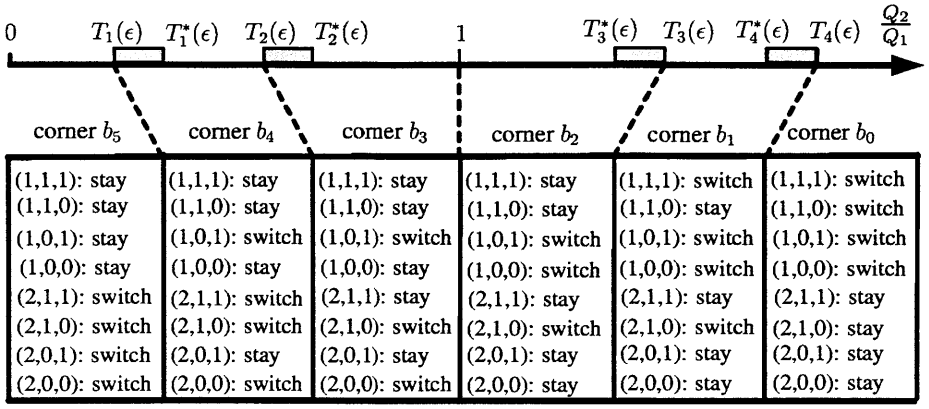


Table 5.3: One-Lookahead Myopic policy mapping from the queue sizes to the corners of $\mathbf{\Lambda_s}$, $b_0, b_1, b_2, b_3, b_4, b_5$ shown in Fig. 5-4 (a), for $\epsilon < \epsilon_c$. For each state $\mathbf{s} = (m(t), C_1(t), C_2(t))$ the optimal action is specified. The thresholds on Q_2/Q_1 are $0, T_1 = \epsilon/(1-\epsilon), T_2 = (1-\epsilon)/(2-\epsilon), 1, T_3 = (2-\epsilon)/(1-\epsilon), T_4 = (1-\epsilon)/\epsilon$. The corresponding thresholds for the FBDC policy are $0, T_1^*, T_2^*, 1, T_3^*, T_4^*$. For example, corner b_2 is chosen in the FBDC policy if $1 \leq Q_2/Q_1 < T_3^*$, whereas in the OLM policy if $1 \leq Q_2/Q_1 < T_3$.

asymptotically in T.

A similar analysis shows that the *Two-Lookahead Myopic Policy* achieves at least 94% of $\mathbf{\Lambda}_s$, while the *Three-Lookahead Myopic Policy* achieves at least 96% of $\mathbf{\Lambda}_s$. The k-Lookahead Myopic Policy is the same as before except that the following weight functions are used for scheduling decisions: Assuming the server is with queue 1 at time slot t,

$W_1(t) = Q_1(jT)\big(C_1(t) + \sum_{\tau=1}^{k} \mathbb{E}[C_1(t+\tau)|C_1(t)]\big)$ and $W_2(t) = Q_2(jT)\sum_{\tau=1}^{k} \mathbb{E}[C_2(t+\tau)|C_2(t)]$.

5.2 General System

In this section we extend the results developed in the previous section to the general case of an arbitrary number of queues in the system.

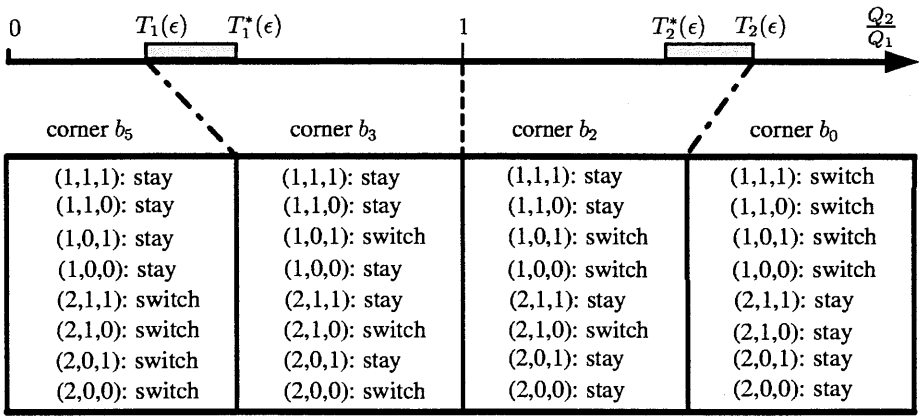


Table 5.4: One-Lookahead Myopic policy mapping from the queue sizes to the corners of $\mathbf{\Lambda_s}$, b_0, b_1, b_2, b_3 shown in Fig. 5-4 (b), for $\epsilon \geq \epsilon_c$. For each state $\mathrm{s} = (m(t), C_1(t), C_2(t))$ the optimal action is specified. The thresholds on Q_2/Q_1 are $0, T_1 = \epsilon/(1-\epsilon), 1, T_2 = (1-\epsilon)/\epsilon$. The corresponding thresholds for the FBDC policy are $0, T_1^*, 1, T_2^*$. For example, corner b_1 is chosen in the FBDC policy if $1 \leq Q_2/Q_1 < T_2^*$, whereas in the OLM policy if $1 \leq Q_2/Q_1 < T_2$.

5.2.1 Model

Consider the same model as in Section 5.1.1 with $N > 1$ queues for some $N \in \mathbb{N}$ as shown in Fig. 5-1. Let the i.i.d. process $A_i(t)$ with arrival rate λ_i denote the number of arrivals to queue i at time slot t, where $\mathbb{E}[A_i^2(t)] \leq A_{\max}^2$, $i \in \{1, 2, ..., N\}$. Let $C_i(t)$ be the channel (connectivity) process of queue i, $i \in \{1, 2, ..., N\}$, that forms the two-state Markov chain with transition probabilities p_{01} and p_{10} as shown in Fig. 5-2. We assume that the processes $A_i(t), i \in \{1, ..., N\}$ and $C_i(t), i \in \{1, ..., N\}$ are independent. It takes one slot for the server to switch from one queue to the other, and $m(t) \in \{1, ..., N\}$ denotes the queue at which the server is present at slot t. Let $\mathrm{s}_t = (m(t), C_1(t), ..., C_N(t)) \in \mathcal{S}$ denote the state of the corresponding saturated system at time t where $\mathcal{S}$ is the set of all states. The action $a(t)$ in each time slot is to choose the queue at which the server will be present in the next time slot, i.e., $a_t \in \{1, ..., N\} \doteq \mathcal{A}$ where $\mathcal{A}$ is the set of all actions at each state.

5.2.2 Stability Region

In this section we characterize the stability region of the general system under non-symmetric channel models[2]. For the case of i.i.d. channel processes we explicitly characterize the stability region and the throughput-optimal policy. For Markovian channel models, we extend the stability region characterization in terms of state-action frequencies to the general system. Furthermore, we develop a tight outer bound on the stability region using an upper bound on the sum-throughput and show that a simple myopic policy achieves this upper bound for the corresponding saturated system.

A dynamic server allocation problem over parallel channels with randomly varying connectivity and limited channel sensing has been investigated in [1, 2, 127] under the Gilbert-Elliot channel model. The goal in [1, 2, 127] is to maximize the sum-rate for the saturated system, where it is proved that a myopic policy is optimal. In this section we prove that a myopic policy is sum-rate optimal under the Gilbert-Elliot channel model and 1-slot server switching delay. Furthermore, our goal is to characterize the set of all achievable rates, i.e., the stability region, together with a throughput-optimal scheduling algorithm for the dynamic queuing system.

Memoryless Channels

The results established in Section 5.1.2 for the case of i.i.d. connectivity processes can easily be extended to the system of N queues with non-symmetric i.i.d. channels as the same intuition applies for the general case. We state this result in the following theorem whose proof can be found in Appendix A.

Theorem 12 For a system of N queues with *arbitrary* switching times and i.i.d. channels with

[2]For Markovian (Gilbert-Elliot) channels, we preserve the symmetry of the channel processes across the queues.

probabilities $p_i, i \in \{1, ..., N\}$, the stability region $\mathbf{\Lambda}$ is given by

$$\mathbf{\Lambda} = \left\{ \boldsymbol{\lambda} \geq \mathbf{0} \middle| \sum_{i=1}^{N} \frac{\lambda_i}{p_i} \leq 1 \right\}.$$

In addition, the simple Exhaustive (Gated) policy is throughput-optimal.

As for the case of two queues, the simultaneous presence of randomly varying connectivity and the switching delay significantly reduces the stability region as compared to the corresponding system without switching delay analyzed in [110]. Furthermore, when the channel processes are memoryless, no policy can take advantage of the channel diversity as the simple queue-blind Exhaustive-type policies are throughput-optimal.

In the next section, we show that, similar to the case of two queues, the memory in the channel improves the stability region of the general system.

Channels With Memory

Similar to Section 5.1.3, we start by establishing the rate region $\mathbf{\Lambda}_s$ by formulating an MDP for rate maximization in the corresponding saturated system. The reward functions in this case are given as follows:

$$\overline{r}_i(\mathrm{s}, a) \doteq 1 \text{ if } m = i, C_i = 1, \text{ and } a = i,\ i = 1, ..., N, \tag{5.15}$$

and $\overline{r}_i(\mathrm{s}, a) \doteq 0$ otherwise, where m denotes the queue at which the server is present. That is, one reward is obtained when the server stays at a queue with an ON channel. Given some $\alpha_i \geq 0$, $i \in \{1, ..., N\}$, $\sum_i \alpha_i = 1$, we define the system reward at time t as

$$\overline{r}(\mathrm{s}, a) \doteq \sum_{i=1}^{N} \alpha_i \overline{r}_i(\mathrm{s}, a).$$

The average reward of policy π is defined as

$$r^{\pi} \doteq \lim_{K\to\infty} \frac{1}{K} E\Big\{ \sum_{t=1}^{K} \overline{r}(\mathbf{s}_t, a_t^{\pi}) \Big\}.$$

Therefore, the problem of maximizing the time average expected reward over all policies, $r^* \doteq \max_{\pi} r^{\pi}$, is a discrete time MDP characterized by the state transition probabilities $\mathbf{P}(\mathbf{s}'|\mathbf{s}, a)$ with $N2^N$ states and N possible actions per state. Furthermore, similar to the two-queue system, there exists a positive probability path between any given pair of states under some stationary-deterministic policy. Therefore, this MDP belongs to the class of *Weakly Communicating* MDPs [94] for which there exists a stationary-deterministic optimal policy independent of the initial state [94]. The *state-action polytope*, $\mathbf{X}$ is the set of state-action frequency vectors $\mathbf{x}$ that satisfy the balance equations

$$\sum_{a\in\mathcal{A}} \mathbf{x}(\mathbf{s}, a) = \sum_{\mathbf{s}'\in\mathcal{S}} \sum_{a\in\mathcal{A}} \mathbf{P}(\mathbf{s}|\mathbf{s}', a)\mathbf{x}(\mathbf{s}', a), \ \forall\, \mathbf{s} \in \mathcal{S}, \tag{5.16}$$

the normalization condition

$$\sum_{\mathbf{s}\in\mathcal{S}} \sum_{a\in\mathcal{A}} \mathbf{x}(\mathbf{s}, a) = 1,$$

and the nonnegativity constraints

$$\mathbf{x}(\mathbf{s}, a) \geq 0, \text{ for } \mathbf{s} \in \mathcal{S}, a \in \mathcal{A},$$

where the transition probabilities $\mathbf{P}(\mathbf{s}|\mathbf{s}', a)$ are functions of the channel parameters p_{10} and p_{01}. The following linear transformation of the state-action polytope $\mathbf{X}$ defines the *rate polytope* $\mathbf{\Lambda}_s$,

namely, the set of all time average expected rate pairs that can be obtained in the saturated system.

$$\mathbf{\Lambda}_s = \left\{ \mathbf{r} | r_i = \sum_{\mathbf{s}\in\mathcal{S}} \sum_{a\in\mathcal{A}} \mathbf{x}(\mathbf{s},a)\overline{r}_i(\mathbf{s},a), \mathbf{x} \in \mathbf{X}, i \in \{1,2,...,N\} \right\},$$

where the reward functions $\overline{r}_i(\mathbf{s},a), i \in \{1,...,N\}$, are defined in (5.15). Algorithm 9 gives an alternative characterization of the rate region $\mathbf{\Lambda}_s$.

Algorithm 9 *Stability Region Characterization*

1: Given $\alpha_1,...,\alpha_N \geq 0$, $\sum_i \alpha_i = 1$, solve the following LP

$$\max_{\mathbf{x}} . \quad \sum_{i=1}^{N} \alpha_i r_i(\mathbf{x})$$

$$\text{subject to} \quad \mathbf{x} \in \mathbf{X}. \tag{5.17}$$

2: There exists an optimal solution $(r_1^*,...,r_N^*)$ of this LP that lies at a corner point of $\mathbf{\Lambda_s}$. Find all possible corner points and take their convex combination.

Similar to the two-queue case, the fundamental theorem of Linear Programming guarantees existence of an optimal solution to (5.17) at a corner point of the polytope $\mathbf{X}$ [15]. We will establish in the next section that the rate region, $\mathbf{\Lambda}_s$ is in fact achievable in the dynamic queueing system, which will imply that $\mathbf{\Lambda} = \mathbf{\Lambda}_s$. For the case of 3 queues, Fig. 5-5 shows the stability region Λ. As expected, the stability region is significantly reduced as compared to the corresponding system with zero switching delays analyzed in [110].

Analytical Outer Bound For The Stability Region

In this section we first derive an upper bound to the sum-throughput in the saturated system and then use it to characterize an outer bound to the rate region $\mathbf{\Lambda}_s$. Let $C_0^{(N)} \doteq \frac{p_{10}^N}{(p_{10}+p_{01})^N}$ denote the probability that all channels are in OFF state in steady state.

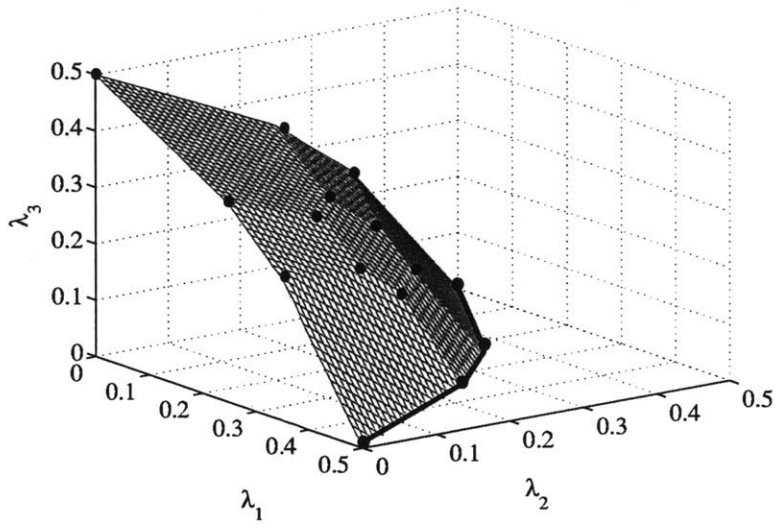


Figure 5-5: Stability region for 3 parallel queues for $p_{10} = p_{01} = 0.3$.

Lemma 19 An upper bound on the sum-rate in the saturated system is given by

$$\sum_{i=1}^{N} r_i \leq 1 - C_0^{(N)} - \left(p_{10}(1 - C_0^{(N)}) - p_{01}C_0^{(N)}\right). \tag{5.18}$$

The proof is given in Appendix G. In the next section we propose a simple myopic policy for the saturated system that achieves this upper bound. Similar to the case of two-queues, the surface $\sum_{i=1}^{N} r_i \leq 1 - C_0^{(N)}$ is one of the boundaries of the stability region for the system without switching delay analyzed in [110], where the probability that at least 1 channel is in ON state in steady state is $1 - C_0^{(N)}$. Therefore, $p_{10}(1 - C_0^{(N)}) - p_{01}C_0^{(N)}$ is *the throughput loss due to 1 slot switching delay* in our system. The analysis of the myopic policy in the next section shows that this throughput loss due to switching delay corresponds to the probability that the server is at a queue with OFF state when at least one other queue is in ON state. For the case of $N = 3$ queues, the sum-throughput upper bound in Lemma 19 is the hexagonal region at the center of the plot in Fig. 5-5.

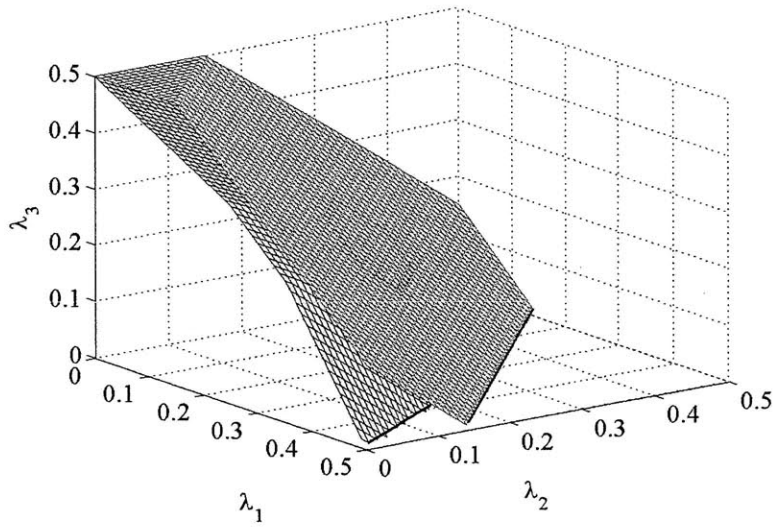


Figure 5-6: Stability region outer bound for 3 parallel queues for $p_{10} = p_{01} = 0.3$.

Because any convex combination of $r_i, i \in \{1, ..., N\}$, must lie under the sum-rate surface, (5.18) is in fact an outer bound on the whole rate region $\mathbf{\Lambda}_s$. Furthermore, no queue can achieve a time average expected rate that is greater than the steady state probability that the corresponding queue is in ON state, i.e., $p_{01}/(p_{10}+p_{01})$. Therefore, the intersection of these $N+1$ surfaces in the N dimensional space constitutes an outer bound for the rate-region $\mathbf{\Lambda}_s$. Note that this outer bound is tight in that the sum-rate surface of the maximum rate region $\mathbf{\Lambda}_s$, as well as the corner points $p_{01}/(p_{10}+p_{01})$ coincide with the outer bound. This outer bound with respect to the rate region are displayed in Fig. 5-6 for the case of $N = 3$ nodes.

5.2.3 Myopic Policy for the Saturated System

We show in this section that a simple and intuitive policy, termed the Greedy Myopic (GM) policy, achieves the sum rate maximization for the saturated system. This policy is a greedy policy in that under the policy, if the current queue is available to serve, the server serves it. Otherwise, the server switches to a queue with ON channel state, if such a queue exists. The policy is described in Algorithm 10. Recall that $m(t)$ denotes the queue the server is present at time slot t.

Algorithm 10 *Greedy Myopic Policy*

1: For all time slots t, if $C_{m(t)}(t) = 1$, serve queue $m(t)$.

2: Otherwise, if $\exists j \in \{1, ..., N\}, j \neq m(t)$, such that $C_j(t) = 1$, among the queues that have ON channel state, switch to the queue with the smallest index in a cyclic order starting from queue $m(t)$.

The cyclic switching order under the GM policy is as follows: If the server is at queue i and the decision is to switch, then the server switches to queue j, where for $i = N, j = \arg\min_{j\in\{1,...,N-1\}}(C_j(t) = 1)$ and for $i \neq N$ if $\exists j \in \{i+1, ..., N\}$, such that $C_j(t) = 1$, we have $j = \arg\min_{j\in\{i+1,...,N\}}(C_j(t) = 1)$, if not, then $j = \arg\min_{j\in\{1,...,i-1\}}(C_j(t) = 1)$.

Theorem 13 The GM policy achieves the sum-rate upper bound.

Proof: Given a fixed decision rule at each state, the system state forms a finite state space, irreducible and positive recurrent Markov chain. Therefore, under the GM policy, the system state converges to a steady state distribution. We partition the total probability space into three disjoint events:

E_1: the event that all the channels are in OFF state,

E_2: the event that at least 1 channel is in ON state and the server is at a queue with ON state

E_3: the event that at least 1 channel is in ON state and the server is at a queue with OFF state

Since these events are disjoint we have,

$$1 = \mathbb{P}(E_1) + \mathbb{P}(E_2) + \mathbb{P}(E_3).$$

We have $\mathbb{P}(E_1) = C_0^{(N)}$ by definition. Since the GM policy decides to serve the current queue if it is in ON state, $\mathbb{P}(E_2)$ gives the sum throughput $\sum_i r_i$ for the GM policy. Therefore, we have that under the GM policy

$$\sum_i r_i = 1 - C_0^{(N)} - \mathbb{P}(E_3).$$

We show that $\mathbb{P}(E_3) = p_{10}(1 - C_0^{(N)}) - p_{01}C_0^{(N)}$. Consider a time slot t in steady state and let $\kappa(t)$ be the number of channels with ON states at time slot t and let $E_0(t)$ be the event that the server is at a queue with OFF state at time slot t. We have

$$\begin{aligned}\mathbb{P}(E_3) &= \mathbb{P}(E_0(t) \text{ and } 1 \leq \kappa(t) \leq N-1) = \mathbb{P}(E_0(t) \text{ and } \kappa(t) \geq 1) \\ &= \mathbb{P}(E_0(t) \text{ and } \kappa(t) \geq 1 | \kappa(t-1) \geq 1)\mathbb{P}(\kappa(t-1) \geq 1) \\ &+ \mathbb{P}(E_0(t) \text{ and } \kappa(t) \geq 1 | \kappa(t-1) = 0)\mathbb{P}(\kappa(t-1) = 0).\end{aligned}$$

Since t is a time slot in steady state, we have that

$\mathbb{P}(\kappa(t-1) = 0) = C_0^{(N)}$. Therefore, $\mathbb{P}(E_3)$ is given by

$$\begin{aligned}&\mathbb{P}\left(\kappa(t) \geq 1 | E_0(t), \kappa(t-1) \geq 1\right)\mathbb{P}(E_0(t) | \kappa(t-1) \geq 1)(1 - C_0^{(N)}) \\ &+ \mathbb{P}\left(\kappa(t) \geq 1 | E_0(t), \kappa(t-1) = 0\right)\mathbb{P}(E_0(t) | \kappa(t-1) = 0)C_0^{(N)}.\end{aligned}$$

We have $\mathbb{P}(E_0(t) | \kappa(t-1) \geq 1) = p_{10}$ since the GM policy chooses a queue with ON state if there is such a queue and $\mathbb{P}(E_0(t) | \kappa(t-1) = 0)$ is the probability that the queue chosen by the GM policy

keeps its OFF channel state, given by $1-p_{01}$.

$$\begin{aligned}\mathbb{P}(E_3)=&(1-\mathbb{P}(\kappa(t)=0|E_0(t),\kappa(t-1)\geq 1))p_{10}(1-C_0^{(N)})\\&+(1-\mathbb{P}(\kappa(t)=0|E_0(t),\kappa(t-1)=0))(1-p_{01})C_0^{(N)}\\\\=&p_{10}(1-C_0^{(N)})\left(1-\frac{\mathbb{P}(\kappa(t)=0,E_0(t)|\kappa(t-1)\geq 1)}{\mathbb{P}(E_0(t)|\kappa(t-1)\geq 1)}\right)\\&+(1-p_{01})C_0^{(N)}\left(1-\frac{\mathbb{P}(\kappa(t)=0,E_0(t)|\kappa(t-1)=0)}{\mathbb{P}(E_0(t)|\kappa(t-1)=0)}\right)\\=&p_{10}(1-C_0^{(N)})\left(1-\frac{\mathbb{P}(\kappa(t)=0|\kappa(t-1)\geq 1)}{p_{10}}\right)\\&+(1-p_{01})C_0^{(N)}\left(1-\frac{\mathbb{P}(\kappa(t)=0|\kappa(t-1)=0)}{1-p_{01}}\right).\end{aligned}$$

We have that $\mathbb{P}(\kappa(t)=0|\kappa(t-1)\geq 1)$ is given by

$$\frac{\mathbb{P}(\kappa(t)=0)-\mathbb{P}(\kappa(t)=0|\kappa(t-1)=0)\mathbb{P}(\kappa(t-1)=0)}{\mathbb{P}(\kappa(t-1)\geq 1)},$$

which is equivalent to $(C_0^{(N)}-(1-p_{01})^N C_0^{(N)})/(1-C_0^{(N)})$. Therefore, $\mathbb{P}(E_3)$ is given by

$$\begin{aligned}\mathbb{P}(E_3)=&p_{10}(1-C_0^{(N)})\left(1-\frac{C_0^{(N)}-(1-p_{01})^N C_0^{(N)}}{p_{10}(1-C_0^{(N)})}\right)\\&+(1-p_{01})C_0^{(N)}\left(1-\frac{(1-p_{01})^N}{1-p_{01}}\right)\\=&p_{10}(1-C_0^{(N)})-p_{01}C_0^{(N)}.\end{aligned}$$

As mentioned in the previous section, $\mathbb{P}(E_3)$ is the throughput loss due to switching as it represents the fraction of time the server is at a queue with OFF state when there are queues with ON state in

the system. □

5.2.4 Frame-Based Dynamic Control Policy

In this section we generalize the FBDC policy to the general system and show that it is throughput-optimal asymptotically in the frame length for the general case. The FBDC algorithm for the general system is very similar to the FBDC algorithm described for two queues in Section 5.1.4. Specifically, the time is divided into equal-size intervals of T slots. We find the stationary-deterministic policy that optimally solves (5.17) for the saturated system when $Q_1(jT), ..., Q_N(jT)$ are used as weights and then apply this policy in each time slot of the frame in the actual system. The FBDC policy is described in Algorithm 11 in details.

Algorithm 11 FRAME BASED DYNAMIC CONTROL (FBDC) POLICY

1: Find the optimal solution to the following LP

$$\begin{aligned} &\max._{\{\mathbf{r}\}} \ \sum_{i=1}^{N} Q_i(jT) r_i \\ &\text{subject to} \quad \mathbf{r} = (r_1, ..., r_N) \in \mathbf{\Lambda}_s \end{aligned} \tag{5.19}$$

where $\mathbf{\Lambda}_s$ is the rate region for the saturated system.

2: The optimal solution $(r_1^*, ..., r_N^*)$ in step 1 is a corner point of $\mathbf{\Lambda}_s$ that corresponds to a stationary-deterministic policy denoted by π^*. Apply π^* in each time slot of the frame.

Theorem 14 For any $\delta > 0$, there exists a large enough frame length T such that the FBDC policy stabilizes the system for all arrival rates within the δ-stripped stability region $\mathbf{\Lambda}_s^\delta = \mathbf{\Lambda}_s - \delta \mathbf{1}$.

The proof is very similar to the proof of Theorem 10 and is omitted. The theorem establishes the asymptotic throughput-optimality of the FBDC policy for the general system.

Remark 3 *The FBDC policy provides* a new framework for developing throughput-optimal policies for network control. *Namely, given any queuing system whose corresponding saturated system is*

Markovian with finite state and action spaces, throughput-optimality is achieved by solving an LP in order to find the stationary MDP solution of the corresponding saturated system and applying this solution over a frame in the actual system. In particular, the FBDC policy can stabilize systems with arbitrary switching times and more complicated Markov modulated channel structures. *The FBDC policy can also be used to achieve throughput-optimality for classical network control problems such as the parallel queueing systems in [87], [110], scheduling in switches in [99] or scheduling under delayed channel state information [125].*

Similar to the delay analysis in Section 5.1.4 for the two-queue system, a delay upper bound that is linear in the frame length T can be obtained for the FBDC policy for the the general system. Moreover, the FBDC policy for the general system can also be implemented without any frames by setting $T = 1$, i.e., by solving the LP in Algorithm 11 in each time slot. The simulation results regarding such implementations suggest that the FBDC policy implemented without frames has a similar throughput performance and an improved delay performance as compared to the original FBDC policy.

Discussion

For systems with switching delay, it is well-known that the celebrated Max-Weight scheduling policy is not throughput-optimal [26]. In the absence of randomly varying connectivity, variable frame based generalizations of the Max-Weight policy are throughput-optimal [32]. However, when the switching delay and randomly varying connectivity are simultaneously present in the system, the FBDC policy is the only policy to achieve throughput-optimality and it has a significantly different structure from the Max-Weight policy.

The FBDC policy for a fixed frame length T does not require the arrival rate information for stabilizing the system for arrival rates in $\mathbf{\Lambda} - \delta(T)\mathbf{1}$, however, it requires the knowledge of the channel connectivity parameters p_{10}, p_{01}. To deal with this problem one can estimate the channel parameters periodically and use these estimates to solve the LP in (5.19). This approach, of course,

incurs a throughput loss depending on how large the estimation error is.

As mentioned in Remark 3, the FBDC policy can stabilize a large class of network control problems whose corresponding saturated system is weakly communicating Markovian with a finite state and action spaces. However, one caveat of the FBDC policy is that the state space of the LP that needs to be solved increases exponentially with the number of links in the system. The celebrated Max-Weight policy (which is not stabilizing for the system considered here) has linear complexity for the single server system considered in this chapter. However, for general multi-server systems with N servers or for a single hop network with N interfering links, the Max-Weight policy has to solve a maximum-independent set problem over all links at each time slot, which is a hard problem whose state space is also exponential in the number of links N. The FBDC policy on the other hand, only has to solve an LP, for which there are standard solvers available such as CPLEX. Furthermore, the FBDC policy has to solve the LP once per frame, whereas the the Max-Weight policy performs maximum-independent set computation each time slot. If the frame length for the FBDC policy is chosen to be bigger than the computational complexity of the LP in (5.19), then the per-slot computational complexity of the algorithm is reduced to $O(1)$. Such a frame-based implementation is also possible for the Max-Weight policy to reduce its complexity to $O(1)$ per time slot. On the other hand, the shortcoming of such an approach for both policies is the increase in delay as a result of the larger frame length. This outlines a tradeoff between complexity and delay, whereby tuning a reduction in complexity by adjusting the frame length comes at the expense of delay.

The celebrated Max-Weight policy was first introduced in [109] for multi-hop networks and, despite its exponential complexity in number of links, it provided a useful structure for designing queue length based scheduling algorithms. Later, this structure suggested by the Max-Weight policy lead to suboptimal but low-complexity algorithms, as well as distributed implementations of the Max-Weight policy for certain systems (see e.g., Greedy-Maximal network scheduling in [122]). Our aim in proposing the FBDC policy and the state-action frequency framework for network scheduling is to give a structure for throughput-optimal algorithms for systems with time-varying

channels and switching delays, and hopefully to provide insight into designing scalable algorithms that can stabilize such systems. The Myopic control policies we discuss in the next section constitute a first approach towards characterizing the structure of some more scalable algorithms.

5.2.5 Myopic Control Policies

In this section, we generalize Myopic policies that we introduced for the two-queue system in Section 5.1 to the general system. Myopic policies make scheduling decisions based on queue lengths and simple channel predictions into the future. We present an implementation of these policies over frames of length T time slots where during the jth frame, the queue lengths at the beginning of the frame, $Q_1(jT), ..., Q_N(jT)$, are used for weight calculations during the frame. We describe the One-Lookahead Myopic (OLM) policy for the general system in Algorithm 12.

Algorithm 12 ONE-LOOKAHEAD MYOPIC POLICY

1: Assuming that the server is currently with queue 1 and the system is at the jth frame, calculate the following weights in each time slot of the current frame;

$$W_1(t) = Q_1(jT)\Big(C_1(t) + \mathbb{E}\big[C_1(t+1)|C_1(t)\big]\Big)$$

$$W_i(t) = Q_i(jT)\mathbb{E}\big[C_i(t+1)|C_i(t)\big]. \tag{5.20}$$

2: If $W_1(t) \geq W_i(t)$, $\forall i \in \{2, ..., N\}$, then stay with queue 1. Otherwise, switch to a queue that achieves

$$\max_i Q_i(jT)\mathbb{E}\big[C_i(t+1)|C_i(t)\big].$$

A similar rule applies when the server is at other queues.

The technique used for the case of two queues for analyzing the stability region achieved by the OLM policy is extremely cumbersome to generalize to the general system with N queues. Therefore, for the general system, we have investigated the performance of the OLM policy in

simulations. The simulation results in Section 5.3 suggest that the OLM policy may achieve the full stability region while providing a better delay performance as compared to the FBDC policy.

Similar to the FBDC policy, the Myopic policies can be implemented without the use of frames by setting $T = 1$, i.e., by utilizing the current queue lengths for updating the decision rules every time slot. This could potentially lead to more delay-efficient policies that are more adaptive to dynamic changes in queue lengths. We elaborate on this via the numerical results in the next section.

Similar to the system with two queues, the k-Lookahead Myopic Policy is the same as before except that the following weight functions are used for scheduling decisions: Assuming the server is with queue 1 at time slot t,

$W_1(t) = Q_1(jT)\big(C_1(t) + \sum_{\tau=1}^{k} \mathbb{E}[C_1(t+\tau)|C_1(t)]\big)$ and $W_i(t) = Q_i(jT)\sum_{\tau=1}^{k} \mathbb{E}[C_i(t+\tau)|C_i(t)]$, $i \in \{2, ..., N\}$.

These policies have very low complexity and they are simpler to implement as compared to the FBDC policy.

5.3 Numerical Results

We performed simulation experiments that present average queue occupancy results for the FBDC, the One-Lookahead Myopic (OLM) and the Max-Weight (MW) policies for systems with $N = 2$ or $N = 3$ queues. We first verified that in the simulation results for the FBDC policy, queue sizes grow unbounded only for arrival rates outside the stability region, and then performed experiments for the One-Lookahead Myopic (OLM) policy. In all the reported results, we have $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ with 0.01 increments. For each point at the boundary of $\boldsymbol{\Lambda}$, we simulated one point outside the stability region. Furthermore, for each data point, the arrival processes were i.i.d., the channel processes were Markovian as in Fig. 5-2 and the simulation length was $T_s = 100,000$ slots.

Fig. 5-7 (a) presents the total average queue size, $\mathbf{Q}_{avg} \doteq \sum_{t=1}^{T_s}(Q_1(t) + Q_2(t))/T_s$, under the FBDC policy for $N = 2$ queues, $\epsilon = 0.25 < \epsilon_c$, and a frame size of $T = 25$ slots. The boundary of the stability region is shown by (red) lines on the two dimensional $\lambda_1 - \lambda_2$ plane. We observe

that the average queue sizes are small for all $(\lambda_1, \lambda_2) \in \mathbf{\Lambda}_s$ and the big jumps in queue sizes occur for points outside $\mathbf{\Lambda}$. Fig. 5-7 (b) presents the performance of the OLM policy with $T = 25$ slot frames for the same system. The simulation results suggest that there is no appreciable difference between the stability regions of the FBDC and the OLM policies. Note that the total average queue size is proportional to long-run packet-average delay in the system through Little's law. For these two figures, the average delay under the OLM policy is less than that under the FBDC policy for 81% of all arrival rates considered.

Next, we implemented the FBDC and the OLM policies without the use of any frames (i.e., for $T = 1$). When there are no frames, the FBDC policy solves the LP in Algorithm 7 in each time slot, and the OLM policy utilizes the queue length information in the *current* time slot for the weight calculations in (5.13). Fig. 5-8 (a) and (b) present the total average queue size under the FBDC and the OLM policies for $N = 2$ queues, $T = 1$, and $\epsilon = 0.40 > \epsilon_c$. Similar to the frame based implementations, we observe that the average queue sizes are small for all $(\lambda_1, \lambda_2) \in \mathbf{\Lambda}$ for both policies and the big jumps in queue sizes occur for points outside $\mathbf{\Lambda}$, which suggests that the the non-frame based implementation of the FBDC and the OLM policies may achieve the full stability region. The reason why the FBDC and the OLM policies provide stability without the use of frames is because for large queue lengths, the corner point that these policies choose to apply depend completely on the queue length ratios, and hence, the choice of corner points and the associated saturated-system policies utilized in the FBDC and the OLM policies do not change fast when the queue lengths get large. Furthermore, the no-frame implementations of these policies are more adaptive to dynamic changes in the queue sizes as compared to implementations with large frames.

For the same system (i.e., $N = 2$ queues and $\epsilon = 0.40 > \epsilon_c$), Fig. 5-9 presents the long-run packet average delay as a function of the sum-throughput $\lambda_1 + \lambda_2$ along the main diagonal line (i.e., $\lambda_1 = \lambda_2$). We compare the delay performance of the FBDC and the OLM policies with $T = 1$, and the Max-Weight policy which, in each time slot t, chooses the queue that achieves $\max_i Q_i(t)C_i(t)$.

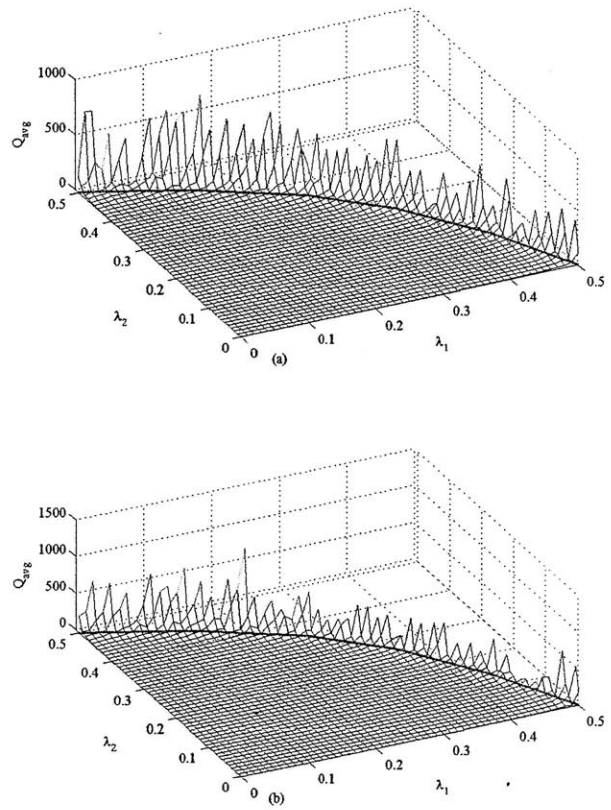


Figure 5-7: The total average queue size for (a) the FBDC policy and (b) the Myopic policy for $T = 25$ and $\epsilon = 0.25$.

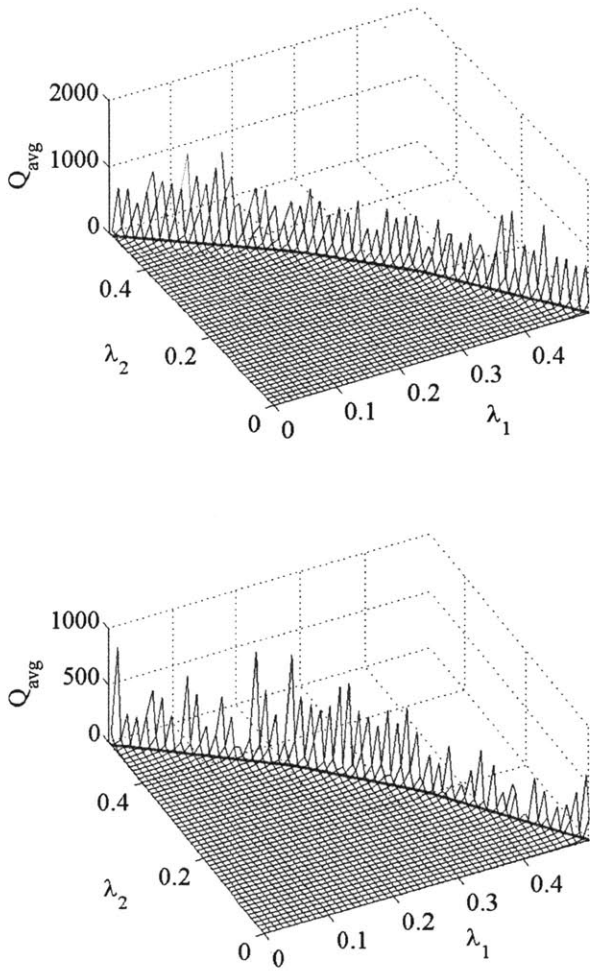


Figure 5-8: The total average queue size for (a) the FBDC policy and (b) the Myopic Policy implemented without the use of frames (i.e., for $T = 1$) and $\epsilon = 0.40$.

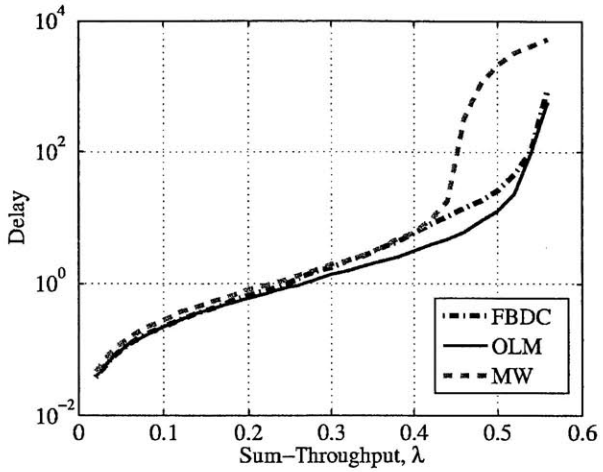


Figure 5-9: Delay vs Sum-throughput for the FBDC, the OLM, and the Max-Weight policies implemented without the use of frames (i.e., for $T = 1$) for $N = 2$ queues and $\epsilon = 0.40$.

The maximum sum-throughput is $0.75 - \epsilon/2 = 0.55$ as suggested by Theorem 9. Fig. 5-9 shows that while FBDC and the OLM policies stabilize the system for all $\lambda_1 + \lambda_2 < 0.55$, the system becomes unstable under the Max-Weight policy around $\lambda_1 + \lambda_2 = 0.45$. This result also confirms that the OLM policy has a much better delay performance than the FBDC and the Max-Weight policies.

For $N = 3$ queues and $\epsilon = 0.30$, Fig. 5-10 presents the long-run packet-average delay as a function of the sum-throughput $\sum_i \lambda_i$ along the main diagonal line (i.e., $\lambda_1 = \lambda_2 = \lambda_3$). The maximum sum-throughput is $1 - 0.5^N - \epsilon * (1 - 0.5^N) + \epsilon * 0.5^N = 0.65$ $1 - \epsilon - 0.5^N(1 - 2\epsilon) = 0.65$ as suggested by Lemma 19. Similar to the previous case, Fig. 5-10 shows that FBDC and the OLM policies stabilize the system for all $\sum_i \lambda_i < 0.65$, the system becomes unstable under the Max-Weight policy around $\sum_i \lambda_i = 0.48$. This result also confirm that the OLM policy has a delay performance than the other two policies.

The delay results in this section show that the OLM policy is not only simpler to implement as compared to the FBDC policy, but it can also be more delay efficient.

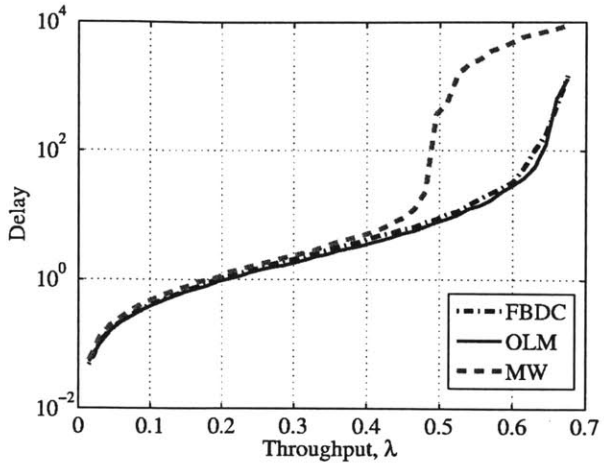


Figure 5-10: Delay vs Sum-throughput for the FBDC, the OLM, and the Max-Weight policies implemented without the use of frames (i.e., for $T = 1$) for $N = 3$ queues and $\epsilon = 0.30$.

5.4 Concluding Remarks

We investigated the dynamic server allocation problem with *time-varying channels* and *server switching time*. For the case of two queues, we analytically characterized the stability region of the system using state-action frequencies that are stationary solutions to an MDP formulation for the corresponding saturated system. We developed the throughput-optimal FBDC policy. We also developed simple *Myopic Policies* that achieve a large fraction of the stability region. We extended the stability region characterization in terms of state-action frequencies and the throughput-optimality of the FBDC policy to the general system with arbitrary number of queues. We characterized tight analytical outer bounds on the stability region using an upper bound on the sum-rate and showed that a simple greedy-myopic policy achieves this sum-rate bound. The stability region characterization in terms of the state-action frequencies of the saturated system and the throughput-optimality of the FBDC policy hold for systems with arbitrary switching times and general Markovian channels. Furthermore, the FBDC policy provides a new framework for developing throughput-optimal policies for network control as this policy can be used to stabilize a large class of other network control problems. Possible future directions include explicitly characterizing the stability region of general system models with multiple-slot switching times and general Markov modulated channels,

and developing Myopic policies with large throughput guarantees. .

Appendix A - Proof of Theorem 7

We prove Theorem 7 for a more general system with N-queues and travel time between queue-i and queue-j given by D_{ij} slots. We call the term $\sum_{i=1}^{N} \lambda_i/p_i$ the system load and denote it by ρ since it is the rate with which the work is entering the system in the form of service slots.

Necessity

We prove that a necessary condition for the stability of any policy is $\rho = \sum_{i=1}^{N} \lambda_i/p_i < 1$. **Proof:** Since queues have memoryless channels, for any received packet, as soon as the server switches to some queue i, the expected time to ON state is $1/p_i$. Namely, the time to ON state is a geometric random variable with parameter p_i, and hence, $1/p_i$ is the "service time per packet" for queue i. In a multiuser single-server system *with or without switching delays*, with i.i.d. arrivals whose average arrival rates are $\lambda_i, i \in \{1, ..., N\}$, and i.i.d. service times independent of arrivals with mean $1/p_i, i \in \{1, ..., N\}$, a necessary condition for stability is given by the system load, ρ, less than 1 [23]. □ To futher ellaborate on this, consider the polling system with zero switching times, i.i.d. arrivals of mean λ_i and i.i.d. service times of mean $1/p_i$. The throughput region of this system is an upperbound on the throughput region of the corresponding system with nonzero switching times. This is because for the same sample path of arrival and service processes, the system with zero switching time can achieve exactly the same departure process as the system with nonzero switching times by making the server idle when necessary. A necessary condition for the stability of the former system is $\rho = \lambda_1/p_1 + \lambda_N/p_N + ... + \lambda_1/p_N < 1$, (e.g., [118], [23]). Finally, note that this necessary stability condition can also be derived by utilizing the state-action frequency approach of Section 5.1.3 for the system with i.i.d. connectivity processes.

Sufficiency **Proof:** Under the Gated cyclic policy, we have a Polling system with i.i.d. arrivals with

mean $\lambda_i, i \in \{1, ..., N\}$, i.i.d. service times independent of arrivals with mean $1/p_i, i \in \{1, ..., N\}$, and finite and constant switching delays. It is shown in [23] that the Gated cyclic policy results in an ergodic system if $\rho = \sum_{i=1}^{N} \frac{\lambda_i}{p_i} < 1$, the expected per-message waiting times in steady-state are finite, and they satisfy a pseudo-conversation law. Through Little's law ([47, pp. 139] or [7, pp. 1109]), this implies that the expected number of packets in the system in steady state is finite, which in turn implies that the system is stable. □

Appendix B - Proof of Lemma 17

Given a policy π for the dynamic queueing system specifying the switch and stay actions based possibly on observed channel and queue state information, consider the saturated system with *the same sample path of channel realizations* for $t \in \{0, 1, 2, ...\}$ and *the same set of actions* as policy π at each time slot $t \in \{0, 1, 2, ...\}$. Let this policy for the saturated system be π'. Let $D_i(t), i \in \{1, 2\}$ be total number departures by time t from queue-i in the original system under policy π and let $D_i'(t), i \in \{1, 2\}$ be the corresponding quantity for the saturated system under policy π'. It is clear that $\lim_{t\to\infty}(D_1(t) + D_2(t))/t \leq 1$, where the same statement also holds for the limit of $D_i'(t), i \in \{1, 2\}$. Since some of the ON channel states are wasted in the original system due to empty queues, we have

$$D_1(t) \leq D_1'(t), \quad \text{and}, \quad D_2(t) \leq D_2'(t). \tag{5.21}$$

Therefore, the time average expectation of $D_i(t), i \in \{1, 2\}$ is also less than or equal to the time average expectation of $D_i'(t), i \in \{1, 2\}$. This completes the proof since (5.21) holds under any policy π for the original system.

Appendix C - Proof of Theorem 9

We enumerate the states as follows:

$$\begin{array}{llll} s=(1,1,1)\equiv 1, & s=(1,1,0)\equiv 2, & s=(1,0,1)\equiv 3, & s=(1,0,0)\equiv 4, \\ s=(2,1,1)\equiv 5 & s=(2,1,0)\equiv 6 & s=(2,0,1)\equiv 7 & s=(2,0,0)\equiv 8. \end{array} \tag{5.22}$$

We rewrite the balance equations in (5.11) in more details.

$$\begin{aligned} x(1;1)+x(1;0) &= (1-\epsilon)^2\big(x(1;1)+x(5;0)\big)+\epsilon(1-\epsilon)\big(x(2;1)+x(6;0)\big) \\ &\quad +\epsilon(1-\epsilon)\big(x(3;1)+x(7;0)\big)+\epsilon^2\big(x(4;1)+x(8;0)\big) \end{aligned} \tag{5.23}$$

$$\begin{aligned} x(2;1)+x(2;0) &= \epsilon(1-\epsilon)\big(x(1;1)+x(5;0)\big)+(1-\epsilon)^2\big(x(2;1)+x(6;0)\big) \\ &\quad +\epsilon^2\big(x(3;1)+x(7;0)\big)+\epsilon(1-\epsilon)\big(x(4;1)+x(8;0)\big) \end{aligned} \tag{5.24}$$

$$\cdots$$

$$\begin{aligned} x(5;1)+x(5;0) &= (1-\epsilon)^2\big(x(5;1)+x(1;0)\big)+\ \epsilon(1-\epsilon)\big(x(6;1)+x(2;0)\big) \\ &\quad +\epsilon(1-\epsilon)\big(x(7;1)+x(3;0)\big)+\epsilon^2\big(x(8;1)+x(4;0)\big) \end{aligned} \tag{5.25}$$

$$\begin{aligned} x(7;1)+x(7;0) &= \epsilon(1-\epsilon)\big(x(5;1)+x(1;0)\big)+\epsilon^2\big(x(6;1)+x(2;0)\big) \\ &\quad +(1-\epsilon)^2\big(x(7;1)+x(3;0)\big)+\epsilon(1-\epsilon)\big(x(8;1)+x(4;0)\big) \end{aligned} \tag{5.26}$$

$$\cdots$$

The following equations hold the channel state pairs (C_1, C_2).

$$x(1;1) + x(1;0) + x(5;1) + x(5;0) = 1/4 \tag{5.27}$$

$$x(2;1) + x(2;0) + x(6;1) + x(6;0) = 1/4 \tag{5.28}$$

$$x(3;1) + x(3;0) + x(7;1) + x(7;0) = 1/4 \tag{5.29}$$

$$x(4;1) + x(4;0) + x(8;1) + x(8;0) = 1/4. \tag{5.30}$$

Let $u_1 = (x(1;1) + x(2;1)$ and $u_2 = (x(5;1) + x(7;1))$. Summing up (5.23) with (5.24) and (5.25) with (5.26) we have

$$\epsilon u_1 = -\big(x(1;0) + x(2;0)\big) + \epsilon\big(x(3;1) + x(4;1)\big) + \epsilon\big(x(7;0) + x(8;0)\big) + (1-\epsilon)\big(x(5;0) + x(6;0)\big)$$

$$\epsilon u_2 = -\big(x(5;0) + x(7;0)\big) + \epsilon\big(x(6;1) + x(8;1)\big) + \epsilon\big(x(2;0) + x(4;0)\big) + (1-\epsilon)\big(x(1;0) + x(3;0)\big)$$

Rearranging and using (5.27)-(5.30) we have

$$u_1 = \frac{1-\epsilon}{2} + \epsilon\big(x(3;1) + x(4;1) + x(7;0) + x(8;0)\big) - (2-\epsilon)\big(x(1;0) + x(2;0)\big)$$
$$- \quad (1-\epsilon)\big(x(5;1) + x(6;1)\big) \tag{5.31}$$

$$u_2 = \frac{2-\epsilon}{4} + \epsilon\big(x(2;0) - x(4;1) + x(6;1) - x(8;0)\big) - (2-\epsilon)\big(x(5;0) + x(7;0)\big)$$
$$- \quad (1-\epsilon)\big(x(1;1) + x(3;1)\big) \tag{5.32}$$

Using (5.25) in (5.31) and (5.23) in (5.32) we have

$$u_1 = \frac{1-\epsilon}{2} + \epsilon\big(x(3;1) + x(4;1) + x(7;0) + x(8;0)\big) - \frac{\epsilon(1-\epsilon)}{2-\epsilon}\big(x(4;0) + x(8;1)\big)$$
$$- \frac{(1-\epsilon)(3-2\epsilon)}{2-\epsilon}x(6;1) + \frac{1-\epsilon}{\epsilon(2-\epsilon)}x(5;0) - \frac{1+\epsilon-\epsilon^2}{\epsilon(2-\epsilon)}x(1;0)$$
$$- \frac{(1-\epsilon)^2}{2-\epsilon}\big(x(3;0) + x(7;1)\big) - \Big(2-\epsilon+\frac{(1-\epsilon)^2}{2-\epsilon}\Big)x(2;0) \tag{5.33}$$

$$u_2 = \frac{2-\epsilon}{4} + \epsilon\big(x(2;0) + x(6;1)\big) - \Big(\epsilon+\frac{\epsilon(1-\epsilon)}{2-\epsilon}\Big)\big(x(4;1) + x(8;0)\big)$$
$$- \frac{(1-\epsilon)(3-2\epsilon)}{2-\epsilon}x(3;1) + \frac{1-\epsilon}{\epsilon(2-\epsilon)}x(1;0) - \frac{1+\epsilon-\epsilon^2}{\epsilon(2-\epsilon)}x(5;0)$$
$$- \frac{(1-\epsilon)^2}{2-\epsilon}\big(x(2;1) + x(6;0)\big) - \Big(2-\epsilon+\frac{(1-\epsilon)^2}{2-\epsilon}\Big)x(7;0). \tag{5.34}$$

Using (5.29) and (5.30) in (5.33) and (5.28) in (5.34) we have

$$u_1 = \frac{(1-\epsilon)(3-2\epsilon)}{4(2-\epsilon)} + \Big(\epsilon+\frac{\epsilon(1-\epsilon)}{2-\epsilon}\Big)\big(x(4;1) + x(8;0)\big) + \frac{1}{2-\epsilon}\big(x(3;1) + x(7;0)\big)$$
$$- \frac{(1-\epsilon)(3-2\epsilon)}{2-\epsilon}x(6;1) + \frac{1-\epsilon}{\epsilon(2-\epsilon)}x(5;0) - \frac{1+\epsilon-\epsilon^2}{\epsilon(2-\epsilon)}x(1;0) - \Big(2-\epsilon+\frac{(1-\epsilon)^2}{2-\epsilon}\Big)x(2;0)$$

(5.35)

$$u_2 = \frac{3-2\epsilon}{4(2-\epsilon)} - \Big(\epsilon+\frac{\epsilon(1-\epsilon)}{2-\epsilon}\Big)\big(x(4;1) + x(8;0)\big) + \frac{1}{2-\epsilon}\big(x(2;0) + x(6;1)\big)$$
$$- \frac{(1-\epsilon)(3-2\epsilon)}{2-\epsilon}x(3;1) + \frac{1-\epsilon}{\epsilon(2-\epsilon)}x(1;0) - \frac{1+\epsilon-\epsilon^2}{\epsilon(2-\epsilon)}x(5;0) - \Big(2-\epsilon+\frac{(1-\epsilon)^2}{2-\epsilon}\Big)x(7;0).$$

(5.36)

Consider the LP objective function $\alpha_1\big(x(1;1)+x(2;1)\big)+\alpha_2\big(x(5;1)+x(7;1)\big)$, and note that the solution to this LP is a stationary deterministic policy for any given α_1 and α_2. This means that, for any state s either $x(s;1)$ or $x(s;0)$ has to be zero. In order to maximize $\alpha_1\big(x(1;1)+x(2;1)\big)+\alpha_2\big(x(5;1)+x(7;1)\big)$ we need

$$
\begin{aligned}
x(7;0)=0 \quad &\text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{1}{(2-\epsilon)^2+(1-\epsilon)^2},\\
x(3;1)=0 \quad &\text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{1}{(1-\epsilon)(3-2\epsilon)},\\
x(5;0)=0 \quad &\text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{1-\epsilon}{1+\epsilon-\epsilon^2},\\
x(8;0)=x(4;1)=0 \quad &\text{if} \quad \frac{\alpha_2}{\alpha_1} \geq 1,\\
x(6;0)=0 \quad &\text{if} \quad \frac{\alpha_2}{\alpha_1} \geq (1-\epsilon)(3-2\epsilon),\\
x(1;1)=0 \quad &\text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{1+\epsilon-\epsilon^2}{1-\epsilon}.
\end{aligned}
$$

Note that we have

$$
\begin{aligned}
(2-\epsilon)^2+(1-\epsilon)^2 &\geq (1-\epsilon)(3-2\epsilon) \geq 1\\
(2-\epsilon)^2+(1-\epsilon)^2 &\geq \frac{1+\epsilon-\epsilon^2}{1-\epsilon} \geq 1
\end{aligned}
$$

holding for all $\epsilon \in [0, 0.5]$. Consider the following two cases:

Case-1: $\epsilon > \epsilon_c = 1 - \sqrt{2}/2$

In this case we have $(1-\epsilon)(3-2\epsilon) < (1+\epsilon-\epsilon^2)/(1-\epsilon)$. This means that we have the following optimal policies depending on the value of α_2/α_1.

$1 \leq \frac{\alpha_2}{\alpha_1} \leq (1-\epsilon)(3-2\epsilon)$:

$$\begin{aligned} &\text{@queue } 1 : (1,1,1) : \text{stay}, \quad (1,1,0) : \text{stay}, \quad (1,0,1) : \text{switch}, \quad (1,0,0) : \text{switch}. \\ &\text{@queue } 2 : (2,1,1) : \text{stay}, \quad (2,1,0) : \text{switch}, \quad (2,0,1) : \text{stay}, \quad (2,0,0) : \text{stay}. \end{aligned}$$

Substituting zero values for the state action pairs that are not chosen into (5.35) and (5.36), it can be seen that this policy achieves the rate pair

$$r_1 = \frac{(1-\epsilon)(3-2\epsilon)}{4(2-\epsilon)}, \quad r_2 = \frac{3-2\epsilon}{4(2-\epsilon)}.$$

$\frac{\alpha_2}{\alpha_1} > (1-\epsilon)(3-2\epsilon)$:

$$\text{@queue } 2 : (2,1,1) : \text{stay}, \quad (2,1,0) : \text{stay}, \quad (2,0,1) : \text{stay}, \quad (2,0,0) : \text{stay}.$$

In this case it is optimal to stay at queue 2 for all channel conditions. The decisions at queue 1 are to switch to queue 2. Namely, it is sufficient that at least one state corresponding to server being at queue 1 to take a switch decision, which is the case for $\alpha_2/\alpha_1 \geq ((1-\epsilon)(3-2\epsilon))$, since $x(3;1) = 0$ if $\alpha_2/\alpha_1 \geq 1/((1-\epsilon)(3-2\epsilon))$. Since the policy decides to always stay at queue 2, it achieves the rate pair

$$r_1 = 0, \quad r_2 = 0.5.$$

Note that the case for $\alpha_2/\alpha_1 < 1$ is symmetric and can be obtained similarly.

Case-2: $\epsilon < \epsilon_c = 1 - \sqrt{2}/2$

In this case we have $(1-\epsilon)(3-2\epsilon) > (1+\epsilon-\epsilon^2)/(1-\epsilon)$. This means that before the state $x(6;0)$ becomes zero, namely for $(1+\epsilon-\epsilon^2)/(1-\epsilon) < \alpha_2/\alpha_1 < (1-\epsilon)(3-2\epsilon)$, having $x(1;1) = 0$ is optimal. This means that there is one more corner point of the rate region for $\epsilon < \epsilon_c$. We have the following optimal policies.

$1 \leq \frac{\alpha_2}{\alpha_1} \leq \frac{1+\epsilon-\epsilon^2}{1-\epsilon}$:

@queue 1 : $(1,1,1)$: stay, $(1,1,0)$: stay, $(1,0,1)$: switch, $(1,0,0)$: switch.
@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: switch, $(2,0,1)$: stay, $(2,0,0)$: stay.

This policy is the same policy as in the previous case and it achieves the rate pair

$$r_1 = \frac{(1-\epsilon)(3-2\epsilon)}{4(2-\epsilon)}, \quad r_2 = \frac{3-2\epsilon}{4(2-\epsilon)}.$$

$\frac{\alpha_2}{\alpha_1} > \frac{1+\epsilon-\epsilon^2}{1-\epsilon}$:

We have the following deterministic actions.

@queue 1 : $(1,1,1)$: switch, $(1,1,0)$: ?, $(1,0,1)$: switch, $(1,0,0)$: switch.
@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: ?, $(2,0,1)$: stay, $(2,0,0)$: stay.

In order to find the final threshold on α_2/α_1, we substitute the above deterministic decisions in

(5.24), (5.25) and (5.26). Utilizing also (5.27), (5.28), (5.29) and (5.30) we obtain

$$x(2;1) = \frac{(1-\epsilon)^2}{4} - (1-\epsilon)^2 x(6;1) \tag{5.37}$$

$$x(5;1) + x(7;1) = \frac{2-\epsilon}{4} + \epsilon x(6;1) \tag{5.38}$$

The previous threshold on α_2/α_1 for $x(6;0)$ to be zero, i.e., $(1-\epsilon)(3-2\epsilon)$, is valid for the case where $x(1;0) = 0$. Other decisions staying the same, when $x(1;0)$ is positive and $x(1;1) = 0$, r_2 increases and r_1 decreases. Therefore the threshold on α_2/α_1 for $x(6;0)$ to be zero changes, in particular, a simple derivation shows that it becomes $\alpha_2/\alpha_1 > (1-\epsilon)^2/\epsilon$. This gives the following two regions:

$\frac{1+\epsilon-\epsilon^2}{1-\epsilon} \leq \frac{\alpha_2}{\alpha_1} \leq \frac{(1-\epsilon)^2}{\epsilon}$:

The optimal policy is

@queue 1 : $(1,1,1)$: switch, $(1,1,0)$: stay, $(1,0,1)$: switch, $(1,0,0)$: switch.
@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: switch, $(2,0,1)$: stay, $(2,0,0)$: stay.

From (5.37) and (5.38) it is easy to see that this policy achieves

$$r_1 = \frac{(1-\epsilon)^2}{4}, \quad r_2 = \frac{2-\epsilon}{4}.$$

$\frac{\alpha_2}{\alpha_1} > \frac{(1-\epsilon)^2}{\epsilon}$:

The optimal decisions at queue 1 are to switch to queue 2. This policy achieves

$$@\text{queue } 2 : (2,1,1) : \text{stay}, \quad (2,1,0) : \text{stay}, \quad (2,0,1) : \text{stay}, \quad (2,0,0) : \text{stay}.$$

This policy achives

$$r_1 = 0, \qquad r_2 = 0.5.$$

Similar to Case-1, the case $\alpha_2/\alpha_1 < 1$ is symmetric and can be solved similarly.

Thus we have characterized the corner point of the stability region for the two regions of ϵ. Using these corner points, it is easy to derive the expressions for the lines connecting these corner points, which are given in Theorem 9.

Appendix D - Generalization to Non-symmetric Gilbert-Elliot Channels

In the following, we state results analogous to the results established in Section 5.1 for symmetric Gilbert-Elliot channels to the case of non-symmetric Gilbert-Elliot channel model as given in Fig. 5-2.

Theorem 15 When the connectivity processes follow the non-symmetric Gilbert-Elliot channel

model, the rate region $\mathbf{\Lambda}_s$ is the set of all rates $r_1 \geq 0, r_2 \geq 0$ that for $p_{01} < \frac{(1-p_{10})^2}{2-p_{10}}$ satisfy

$$\begin{aligned}
p_{01}r_1 + h_1 r_2 &\leq h_1 \frac{p_{01}}{p_{01}+p_{10}} \\
(1-p_{10})r_1 + h_2 r_2 &\leq 1 - \frac{p_{10}^2}{(p_{10}+p_{01})^2} - \frac{p_{10}p_{01}}{p_{01}+p_{10}} \\
r_1 + r_2 &\leq 1 - \frac{p_{10}^2}{(p_{10}+p_{01})^2} - \frac{p_{10}p_{01}}{p_{01}+p_{10}} \\
h_2 r_1 + (1-p_{10})r_2 &\leq 1 - \frac{p_{10}^2}{(p_{10}+p_{01})^2} - \frac{p_{10}p_{01}}{p_{01}+p_{10}} \\
h_1 r_1 + p_{01} r_2 &\leq h_1 \frac{p_{01}}{p_{01}+p_{10}},
\end{aligned}$$

where $h_1 = (1-p_{01})(1-p_{10})$, $h_2 = 1 + p_{10} - p_{10}^2$, and for $p_{01} \geq \frac{(1-p_{10})^2}{2-p_{10}}$ satisfy

$$\begin{aligned}
p_{01}r_1 + h_3 r_2 &\leq h_3 \frac{p_{01}}{p_{10}+p_{01}} \\
r_1 + r_2 &\leq 1 - \frac{p_{10}^2}{(p_{10}+p_{01})^2} - \frac{p_{10}p_{01}}{p_{01}+p_{10}} \\
h_3 r_1 + p_{01} r_2 &\leq h_3 \frac{p_{01}}{p_{10}+p_{01}},
\end{aligned}$$

where $h_3 = (1-p_{10})(p_{10} + (p_{10}+p_{01})(1-p_{10}))$.

Proof: We enumerate the states as follows:

$$\begin{array}{llll}
s = (1,1,1) \equiv 1, & s = (1,1,0) \equiv 2, & s = (1,0,1) \equiv 3, & s = (1,0,0) \equiv 4, \\
s = (2,1,1) \equiv 5 & s = (2,1,0) \equiv 6 & s = (2,0,1) \equiv 7 & s = (2,0,0) \equiv 8.
\end{array} \tag{5.39}$$

We rewrite the balance equations in (5.11) in more details.

$$x(1;1)+x(1;0)=(1-p_{10})^2\big(x(1;1)+x(5;0)\big)+p_{01}(1-p_{10})\big(x(2;1)+x(6;0)\big)$$
$$+p_{01}(1-p_{10})\big(x(3;1)+x(7;0)\big)+p_{01}^2\big(x(4;1)+x(8;0)\big) \tag{5.40}$$

$$x(2;1)+x(2;0)=p_{10}(1-p_{10})\big(x(1;1)+x(5;0)\big)+(1-p_{10})(1-p_{01})\big(x(2;1)+x(6;0)\big)$$
$$+p_{01}p_{10}\big(x(3;1)+x(7;0)\big)+p_{01}(1-p_{01})\big(x(4;1)+x(8;0)\big) \tag{5.41}$$

$$\dots$$

$$x(5;1)+x(5;0)=(1-p_{10})^2\big(x(5;1)+x(1;0)\big)+\ p_{01}(1-p_{10})\big(x(6;1)+x(2;0)\big)$$
$$+p_{01}(1-p_{10})\big(x(7;1)+x(3;0)\big)+p_{01}^2\big(x(8;1)+x(4;0)\big) \tag{5.42}$$

$$x(7;1)+x(7;0)=p_{10}(1-p_{10})\big(x(5;1)+x(1;0)\big)+p_{10}p_{01}\big(x(6;1)+x(2;0)\big)$$
$$+(1-p_{01})(1-p_{10})\big(x(7;1)+x(3;0)\big)+p_{01}(1-p_{01})\big(x(8;1)+x(4;0)\big) \tag{5.43}$$

$$\dots$$

The following equations hold the channel state pairs (C_1, C_2).

$$x(1;1)+x(1;0)+x(5;1)+x(5;0)=\frac{p_{01}^2}{(p_{01}+p_{10})^2} \tag{5.44}$$

$$x(2;1)+x(2;0)+x(6;1)+x(6;0)=\frac{p_{01}p_{10}}{(p_{01}+p_{10})^2} \tag{5.45}$$

$$x(3;1)+x(3;0)+x(7;1)+x(7;0)=\frac{p_{01}p_{10}}{(p_{01}+p_{10})^2} \tag{5.46}$$

$$x(4;1)+x(4;0)+x(8;1)+x(8;0)=\frac{p_{10}^2}{(p_{01}+p_{10})^2}. \tag{5.47}$$

Let $u_1 = (x(1;1)+x(2;1)$ and $u_2 = (x(5;1)+x(7;1))$. Summing up (5.40) with (5.41) and

(5.42) with (5.43) we have

$$p_{10}u_1 = -\big(x(1;0)+x(2;0)\big) + p_{01}\big(x(3;1)+x(4;1)\big)$$

$$+ \quad p_{01}\big(x(7;0)+x(8;0)\big) + (1-p_{10})\big(x(5;0)+x(6;0)\big)$$

$$p_{10}u_2 = -\big(x(5;0)+x(7;0)\big) + p_{01}\big(x(6;1)+x(8;1)\big)$$

$$+ \quad p_{01}\big(x(2;0)+x(4;0)\big) + (1-p_{10})\big(x(1;0)+x(3;0)\big)$$

Rearranging and using (5.44)-(5.47) we have

$$u_1 = \frac{p_{01}(1-p_{10})}{p_{01}+p_{10}} + p_{01}\big(x(3;1)+x(4;1)+x(7;0)+x(8;0)\big)$$

$$- \quad (2-p_{10})\big(x(1;0)+x(2;0)\big) - (1-p_{10})\big(x(5;1)+x(6;1)\big) \qquad (5.48)$$

$$u_2 = \frac{p_{01}(p_{01}+p_{10}-p_{01}p_{10})}{(p_{01}+p_{10})^2} + p_{01}\big(x(2;0)-x(4;1)+x(6;1)-x(8;0)\big)$$

$$- \quad (2-p_{10})\big(x(5;0)+x(7;0)\big) - (1-p_{10})\big(x(1;1)+x(3;1)\big). \qquad (5.49)$$

Using (5.42) in (5.48) and (5.40) in (5.49) we have

$$u_1=\frac{p_{01}(1-p_{10})}{p_{01}+p_{10}}+p_{01}\big(x(3;1)+x(4;1)+x(7;0)+x(8;0)\big)-\frac{p_{01}^2(1-p_{10})}{p_{10}(2-p_{10})}\big(x(4;0)+x(8;1)\big)$$

$$-\ (1-p_{10})(1+\frac{p_{01}(1-p_{10})}{p_{10}(2-p_{10})})x(6;1)+\frac{1-p_{10}}{p_{10}(2-p_{10})}x(5;0)-\frac{1+p_{10}-p_{10}^2}{p_{10}(2-p_{10})}x(1;0)$$

$$-\ \frac{p_{01}(1-p_{10})^2}{p_{10}(2-p_{10})}\big(x(3;0)+x(7;1)\big)-\Big(2-p_{10}+\frac{p_{01}(1-p_{10})^2}{p_{10}(2-p_{10})}\Big)x(2;0) \tag{5.50}$$

$$u_2=\frac{p_{01}(p_{01}+p_{10}-p_{01}p_{10})}{(p_{01}+p_{10})^2}+p_{10}\big(x(2;0)+x(6;1)\big)-\Big(p_{01}+\frac{p_{01}^2(1-p_{10})}{p_{10}(2-p_{10})}\Big)\big(x(4;1)+x(8;0)\big)$$

$$-\ (1-p_{10})(1+\frac{p_{01}(1-p_{10})}{p_{10}(2-p_{10})})x(3;1)+\frac{1-p_{10}}{p_{10}(2-p_{10})}x(1;0)-\frac{1+p_{10}-p_{10}^2}{p_{10}(2-p_{10})}x(5;0)$$

$$-\ \frac{p_{01}(1-p_{10})^2}{p_{10}(2-p_{10})}\big(x(2;1)+x(6;0)\big)-\Big(2-p_{10}+\frac{p_{01}(1-p_{10})^2}{p_{10}(2-p_{10})}\Big)x(7;0). \tag{5.51}$$

Using (5.46) and (5.47) in (5.50) and (5.45) in (5.51) we have

$$u_1 = \frac{p_{01}(1-p_{10})}{p_{01}+p_{10}} - \frac{p_{01}^2(1-p_{10})}{(2-p_{10})(p_{01}+p_{10})^2} + \frac{p_{01}^2(1-p_{10})}{p_{10}(2-p_{10})}\big(x(4;1)+x(8;0)\big)$$

$$+ \frac{p_{01}}{p_{10}(2-p_{10})}\big(x(3;1)+x(7;0)\big) - (1-p_{10})(1+\frac{p_{01}(1-p_{10})}{p_{10}(2-p_{10})})x(6;1) + \frac{1-p_{10}}{p_{10}(2-p_{10})}x(5;0)$$

$$- \frac{1+p_{10}-p_{10}^2}{p_{10}(2-p_{10})}x(1;0) - \Big(2-p_{10}+\frac{p_{01}(1-p_{10})^2}{p_{10}(2-p_{10})}\Big)x(2;0) \tag{5.52}$$

$$u_2 = \frac{p_{01}(p_{01}+p_{10}-p_{01}p_{10})}{(p_{01}+p_{10})^2} - \frac{p_{01}^2(1-p_{10})^2}{(2-p_{10})(p_{01}+p_{10})^2} - \Big(p_{01}+\frac{p_{01}^2(1-p_{10})}{p_{10}(2-p_{10})}\Big)\big(x(4;1)+x(8;0)\big)$$

$$+ \frac{p_{01}}{p_{10}(2-p_{10})}\big(x(2;0)+x(6;1)\big) - (1-p_{10})(1+\frac{p_{01}(1-p_{10})}{p_{10}(2-p_{10})})x(3;1) + \frac{1-p_{10}}{p_{10}(2-p_{10})}x(1;0)$$

$$- \frac{1+p_{10}-p_{10}^2}{p_{10}(2-p_{10})}x(5;0) - \Big(2-p_{10}+\frac{p_{01}(1-p_{10})^2}{p_{10}(2-p_{10})}\Big)x(7;0). \tag{5.53}$$

Consider the LP objective function $\alpha_1\big(x(1;1)+x(2;1)\big)+\alpha_2\big(x(5;1)+x(7;1)\big)$, and note that the solution to this LP is a stationary deterministic policy for any given α_1 and α_2. This means that, for any state s either $x(s;1)$ or $x(s;0)$ has to be zero. In order to maximize $\alpha_1\big(x(1;1)+x(2;1)\big)+\alpha_2\big(x(5;1)+x(7;1)\big)$ we need

$$x(7;0) = 0 \quad \text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{p_{01}}{p_{10}(2-p_{10})^2 + p_{01}(1-p_{10})^2},$$

$$x(3;1) = 0 \quad \text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{p_{01}}{(1-p_{10})(p_{10}(2-p_{10}) + p_{01}(1-p_{10}))},$$

$$x(5;0) = 0 \quad \text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{1-p_{10}}{1+p_{10}-p_{10}^2},$$

$$x(8;0) = x(4;1) = 0 \quad \text{if} \quad \frac{\alpha_2}{\alpha_1} \geq 1,$$

$$x(6;0) = 0 \quad \text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{(1-p_{10})(p_{10}(2-p_{10}) + p_{01}(1-p_{10}))}{p_{01}},$$

$$x(1;1) = 0 \quad \text{if} \quad \frac{\alpha_2}{\alpha_1} \geq \frac{1+p_{10}-p_{10}^2}{1-p_{10}}.$$

Note that we have that

$$\frac{p_{10}(2-p_{10})^2 + p_{01}(1-p_{10})^2}{p_{01}} \geq \frac{1+p_{10}-p_{10}^2}{1-p_{10}} \geq 1$$

$$\frac{p_{10}(2-p_{10})^2 + p_{01}(1-p_{10})^2}{p_{01}} \geq \frac{(1-p_{10})(p_{10}(2-p_{10}) + p_{01}(1-p_{10}))}{p_{01}} \geq 1,$$

whenever $p_{10} + p_{10} < 1$ (the condition for positive correlation). Consider the following two cases:

Case-1: $p_{01} \geq \frac{(1-p_{10})^2}{2-p_{10}}$

In this case we have $\frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}} \leq \frac{1+p_{10}-p_{10}^2}{1-p_{10}}$. This means that we have the following optimal policies depending on the value of α_2/α_1.

$1 \leq \frac{\alpha_2}{\alpha_1} \leq \frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}}$:

@queue 1 : $(1,1,1)$: stay, $(1,1,0)$: stay, $(1,0,1)$: switch, $(1,0,0)$: switch.
@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: switch, $(2,0,1)$: stay, $(2,0,0)$: stay.

Substituting the above zero variables into (5.52) and (5.53), it can be seen that this policy achieves the rate pair

$$r_1 = \frac{1-p_{10}}{2-p_{10}} \frac{p_{01}p_{10} + p_{01}(1-p_{10})(p_{01}+p_{10})}{(p_{01}+p_{10})^2}, \quad r_2 = \frac{1}{2-p_{10}} \frac{p_{01}p_{10} + p_{01}(1-p_{10})(p_{01}+p_{10})}{(p_{01}+p_{10})^2}.$$

$\frac{\alpha_2}{\alpha_1} > \frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}}$:

@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: stay, $(2,0,1)$: stay, $(2,0,0)$: stay.

In this case it is optimal to stay at queue 2 for all channel conditions. The decisions at queue 1 are to switch to queue 2. Namely, it is sufficient that at least one state corresponding to server being at queue 1 to take a switch decision, which is the case for $\alpha_2/\alpha_1 \geq \frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}}$, since $x(3;1) = 0$ if $\alpha_2/\alpha_1 \geq (\frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}})^{-1}$. Since the policy decides to always stay at queue 2, it achieves the rate pair

$$r_1 = 0, \quad r_2 = \frac{p_{01}}{p_{01}+p_{10}}.$$

Note that the case for $\alpha_2/\alpha_1 < 1$ is symmetric and can be obtained similarly.

Case-2: $p_{01} < \frac{(1-p_{10})^2}{2-p_{10}}$

In this case we have $\frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}} > \frac{1+p_{10}-p_{10}^2}{1-p_{10}}$. This means that before the state $x(6;0)$ becomes zero, namely for $\frac{1+p_{10}-p_{10}^2}{1-p_{10}} < \alpha_2/\alpha_1 < \frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}}$, having $x(1;1) = 0$ is optimal. This means that there is one more corner point to the rate region for $p_{01} < \frac{(1-p_{10})^2}{2-p_{10}}$. We have the following optimal policies.

$1 \leq \frac{\alpha_2}{\alpha_1} \leq \frac{1+p_{10}-p_{10}^2}{1-p_{10}}$:

@queue 1 : $(1,1,1)$: stay, $(1,1,0)$: stay, $(1,0,1)$: switch, $(1,0,0)$: switch.
@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: switch, $(2,0,1)$: stay, $(2,0,0)$: stay.

This policy is the same policy as in the previous case and it achieves the rate pair

$$r_1 = \frac{1-p_{10}}{2-p_{10}} \frac{p_{01}p_{10}+p_{01}(1-p_{10})(p_{01}+p_{10})}{(p_{01}+p_{10})^2}, \quad r_2 = \frac{1}{2-p_{10}} \frac{p_{01}p_{10}+p_{01}(1-p_{10})(p_{01}+p_{10})}{(p_{01}+p_{10})^2}.$$

$\frac{\alpha_2}{\alpha_1} > \frac{1+p_{10}-p_{10}^2}{1-p_{10}}$:

We have the following deterministic actions.

@queue 1 : $(1,1,1)$: switch, $(1,1,0)$: ?, $(1,0,1)$: switch, $(1,0,0)$: switch.
@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: ?, $(2,0,1)$: stay, $(2,0,0)$: stay.

In order to find the final threshold on α_2/α_1, we substitute the above deterministic decisions in (5.41), (5.42) and (5.43). Utilizing also (5.44), (5.45), (5.46) and (5.47) we obtain

$$x(2;1) = \frac{(1-\epsilon)^2}{4} - (1-\epsilon)^2 x(6;1) \tag{5.54}$$

$$x(5;1) + x(7;1) = \frac{p_{01}(p_{10} - p_{10}p_{01} + p_{01})}{(p_{01}+p_{10})^2} + p_{01}x(6;1) \tag{5.55}$$

The previous threshold on α_2/α_1 for $x(6;0)$ to be zero, i.e., $\frac{(1-p_{10})(p_{10}(2-p_{10})+p_{01}(1-p_{10}))}{p_{01}}$, is valid for the case where $x(1;0) = 0$. Other decisions staying the same, when $x(1;0)$ is positive and $x(1;1) = 0$, r_2 increases and r_1 decreases. Therefore the threshold on α_2/α_1 for $x(6;0)$ to be zero changes, in particular, a simple derivation shows that it becomes $\alpha_2/\alpha_1 > (1-p_{10})(1-p_{01})/p_{01}$. This gives the following two regions:

$\frac{1+p_{10}-p_{10}^2}{1-p_{10}} \leq \frac{\alpha_2}{\alpha_1} \leq (1-p_{10})(1-p_{01})/p_{01}$:

The optimal policy is

@queue 1 : $(1,1,1)$: switch, $(1,1,0)$: stay, $(1,0,1)$: switch, $(1,0,0)$: switch.
@queue 2 : $(2,1,1)$: stay, $(2,1,0)$: switch, $(2,0,1)$: stay, $(2,0,0)$: stay.

From (5.54) and (5.55) it is easy to see that this policy achieves

$$r_1 = (1-p_{10})(1-p_{01})\frac{p_{01}p_{10}}{(p_{01}+p_{10})^2}, \qquad r_2 = \frac{p_{01}(p_{10} - p_{10}p_{01} + p_{01})}{(p_{01}+p_{10})^2}.$$

$\frac{\alpha_2}{\alpha_1} > (1-p_{10})(1-p_{01})/p_{01}$:

The optimal policy is

$$@\text{queue } 2 : (2,1,1) : \text{stay}, \quad (2,1,0) : \text{stay}, \quad (2,0,1) : \text{stay}, \quad (2,0,0) : \text{stay}.$$

The optimal decisions at queue 1 are to switch to queue 2. This policy achieves

$$r_1 = 0, \qquad r_2 = \frac{p_{01}}{p_{01} + p_{10}}.$$

Similar to Case-1, the case $\alpha_2/\alpha_1 < 1$ is symmetric and can be solved similarly.

Thus we have characterized the corner point of the stability region for the two regions of p_{01} and p_{10}. Using these corner points, it is easy to derive the expressions for the lines connecting these corner points, which are given in Theorem 15. □

Closely examining the upper bound on sum-rate $r_1 + r_2$, the term $1 - p_{10}^2/(p_{10} + p_{01})^2$ is the steady state probability that at least one channel is in ON state. This is the maximum achievable sum-rate value for the system with zero switching delay studied in [110]. Therefore, the term $\frac{p_{10}p_{01}}{p_{01}+p_{10}}$ is exactly the loss due to switching delay. It can be shown that, under a sum-rate-optimal policy, this term is equal to the steady state probability that server is at a queue with an OFF channel state when the other queue is at an ON channel state.

The FBDC policy is asymptotically throughput-optimal under the non-symmetric Gilbert-Elliot channel model. This is straightforward as the FBDC policy only needs to solve the LP in Algorithm 7 for a given Markovian state transition structure, and the non-symmetric Gilbert-Elliot channel model leads to a Markovian state transition structure. For the non-symmetric Gilbert-Elliot channels case, the mappings from the queue sizes to the corner points of the rate region used by the FBDC policy, analogues to the mappings in tables 5.2, and 5.1 can be obtained from the slopes of the lines forming the boundary of the stability region. Furthermore, an analysis very similar to the one in Section 5.1.5 gives the corresponding mapping for the One-Lookahead Myopic (OLM)

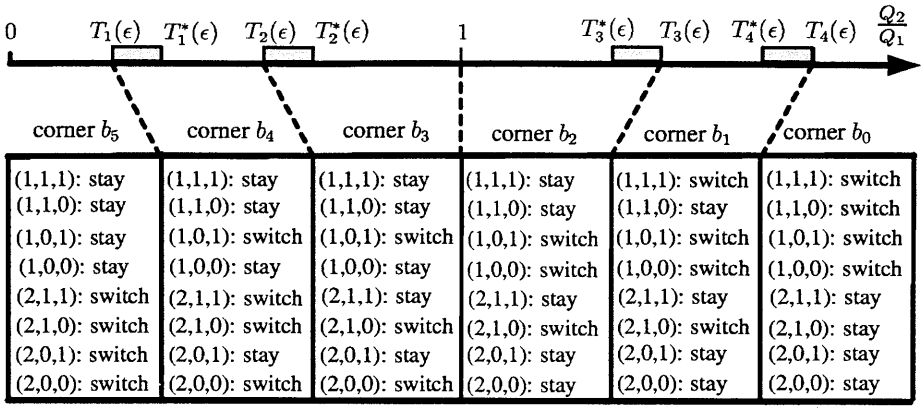


Table 5.5: Mapping from the queue sizes to the corners of $\mathbf{\Lambda_s}$, $b_0, b_1, b_2, b_3, b_4, b_5$, for $p_{01} < \frac{(1-p_{10})^2}{2-p_{10}}$. For each state $\mathbf{s} = (m(t), C_1(t), C_2(t))$ the optimal action is specified. The thresholds on Q_2/Q_1 for the FBDC policy are $0, T_1^* = p_{01}/((1-p_{01})(1-p_{10})), T_2^* = (1-p_{10})/(1+p_{10}-p_{10}^2), 1, T_3^* = 1/T_2^*, T_4^* = 1/T_1^*$ and for the OLM policy are $0, T_1 = p_{01}/(1-p_{10}), T_2 = (1-p_{10})/(2-p_{01}), 1, T_3 = 1/T_2, T_4 = 1/T_1$. For example corner b_1 is chosen in the FBDC policy if $1 \leq Q_2/Q_1 < T_3^*$, whereas in the OLM policy if $1 \leq Q_2/Q_1 < T_3$.

policy. These mappings are shown in Table 5.5 for the case of $p_{01} < \frac{(1-p_{10})^2}{2-p_{10}}$, and in Table 5.6 for the case of $p_{01} \geq \frac{(1-p_{10})^2}{2-p_{10}}$.

Appendix E - Proof of Theorem 14

Let t_k be the first slot of the kth frame where $t_{k+1} = t_k + T$. Let $D_i(t)$ be the *service opportunity* given to queue i at time slot t, where $D_i(t)$ is equal to 1 if queue i is scheduled at time slot t (regardless of whether queue i is empty or not) and zero otherwise. We have the following queue evolution relation:

$$Q_i(t+1) = \max(Q_i(t) - D_i(t), 0) + A_i(t). \tag{5.56}$$

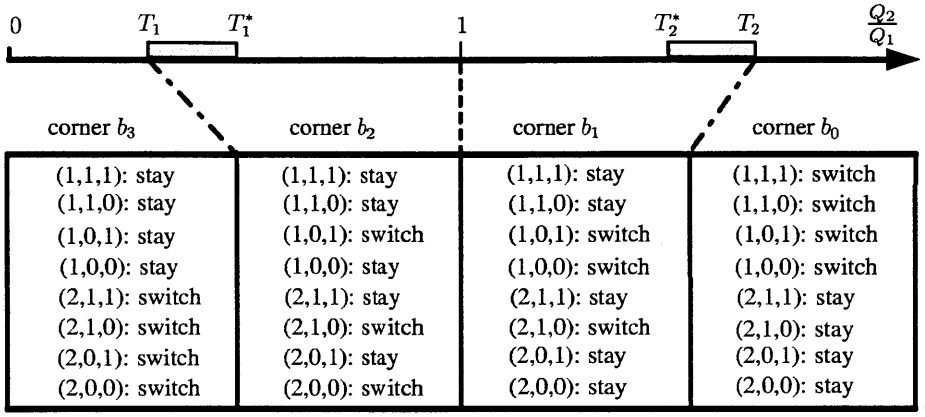


Table 5.6: Mapping from the queue sizes to the corners of $\mathbf{\Lambda_s}$, b_0, b_1, b_2, b_3, for $p_{01} \geq \frac{(1-p_{10})^2}{2-p_{10}}$. For each state $\mathbf{s} = (m(t), C_1(t), C_2(t))$ the optimal action is specified. The thresholds on Q_2/Q_1 for the FBDC policy are $0, T_1^* = p_{01}/((1-p_{10})(p_{10}+(p_{10}+p_{01})(1-p_{10}))), 1, T_2^* = (1-p_{10})(p_{10}+(p_{10}+p_{01})(1-p_{10}))/p_{01}$ and for the OLM policy are $0, T_1 = p_{01}/(1-p_{10}), 1, T_2 = (1-p_{10})/p_{01}$. For example corner b_1 is chosen in the FBDC policy if $1 \leq Q_2/Q_1 < T_2^*$, whereas in the OLM policy if $1 \leq Q_2/Q_1 < T_2$.

Similarly, the following T-step queue evolution relation holds:

$$Q_i(t_k+T) \leq \max\left\{Q_i(t_k) - \sum_{\tau=0}^{T-1} D_i(t_k+\tau), 0\right\} + \sum_{\tau=0}^{T-1} A_i(t_k+\tau), \tag{5.57}$$

where $\sum_{\tau=0}^{T-1} D_i(t_k+\tau)$ is the total *service opportunity* given to queue i during the k^{th} frame. To see this, note that if $\sum_{\tau=0}^{T-1} D_i(t_k+\tau)$, the total *service opportunity* given to queue i during the k^{th} frame, is smaller than $Q_i(t_k)$, then we have an equality. Otherwise, the first term is 0 and we have an inequality. This is because some of the arrivals during the frame might depart before the end of the frame. Note that $\sum_{\tau=0}^{T-1} D_i(t_k+\tau)$ denotes the link i departures that would happen in the corresponding *saturated system* if we were to apply the *same* switching decisions over T time slots in the corresponding saturated system. We first prove stability at the frame boundaries. Define the

quadratic Lyapunov function

$$L(\mathbf{Q}(t)) = \sum_{i=1}^{N} Q_i^2(t),$$

which represents a quadratic measure of the total load in the system at time slot t. Define the T-step conditional drift

$$\Delta_T(t_k) \triangleq \mathbb{E}\left[L(\mathbf{Q}(t_k+T)) - L(\mathbf{Q}(t_k))\middle|\mathbf{Q}(t_k)\right],$$

where the conditional expectation is over the randomness in arrivals and possibly the scheduling decisions. Squaring both sides of (5.57), using $\max(0,x)^2 \leq x^2, \forall x \in \mathbb{N} \cup \{0\}$, and $D_i(t) \leq 1, \forall t$ we have

$$Q_i(t_k+T)^2 - Q_i(t_k)^2 \leq T^2 + \Big(\sum_{\tau=0}^{T-1} A_i(t_k+\tau)\Big)^2$$

$$-2Q_i(t_k)\Big(\sum_{\tau=0}^{T-1} D_i(t_k+\tau) - \sum_{\tau=0}^{T-1} A_i(t_k+\tau)\Big). \tag{5.58}$$

Summing (5.58) over the queues, using $\mathbb{E}[A_i(t)^2] \leq A_{\max}^2$ and $\mathbb{E}[A_i(t_1)A_i(t_2)] \leq \sqrt{\mathbb{E}[A_i(t_1)]^2\mathbb{E}[A_i(t_2)]^2} \leq A_{\max}^2$ for all time slots t_1 and t_2, we can easily derive the following T-step conditional Lyapunov drift

$$\Delta_T(t_k) \leq NBT^2 + 2T\sum_{i=1}^{N} Q_i(t_k)\lambda_i$$

$$-2\sum_{\ell} Q_i(t_k)\mathbb{E}\left[\sum_{\tau=0}^{T-1} D_i(t_k+\tau)\middle|\mathbf{Q}(t_k)\right],$$

where $B \doteq 1 + A_{\max}^2$. Recall the definition of the reward functions $\overline{r}_i(\mathbf{s}_t, \mathbf{a}_t)$, $i \in \{1, ..., N\}$, in (5.15) and let $\overline{r}_i(\mathbf{s}_t, \mathbf{a}_t)$ be the reward function associated with applying policy $\boldsymbol{\pi}^*$ given in the

definition of the FBDC policy in Algorithm 11 to the saturated system. Let $\overline{r}_i(t)$ denote $\overline{r}_i(\mathbf{s}_t, \mathbf{a}_t)$ for notational simplicity, $i \in \{1, ..., N\}$. Note again that $r_i(t)$ is equal to $D_i(t)$, since $D_i(t)$ is the *service opportunity* given to link i at time slot t. Now let $\mathbf{r}^* = (r_i^*)_i$ be the infinite horizon average rate associated with policy π^*. Let $\mathbf{x}^*$ be the optimal vector of state-action frequencies corresponding to π^*. Define the time-average empirical reward from queue i in the saturated system, $\hat{r}_{T,i}(t_k)$, $i \in \{1, ..., N\}$ by

$$\hat{r}_{T,i}(t_k) \doteq \frac{1}{T} \sum_{\tau=0}^{T-1} \overline{r}_i(t_k + \tau).$$

Similarly, define the time average empirical state-action frequency vector $\hat{\mathbf{x}}_T(t_k; \mathbf{s}, \mathbf{a})$.

$$\hat{\mathbf{x}}_T(t_k; \mathbf{s}, \mathbf{a}) \doteq \frac{1}{T} \sum_{\tau=t_k}^{t_k+T-1} I_{\{\mathbf{s}_\tau = \mathbf{s}, \mathbf{a}_\tau = \mathbf{a}\}},$$

where I_E is the indicator function of an event E, i.e., $I_E = 1$ if E occurs and $I_E = 0$ otherwise. Using the definition of the reward functions in (5.15), we have that

$$\hat{r}_{T,i}(t_k) = \sum_{\mathbf{s} \in \mathcal{S}} \sum_{\mathbf{a} \in \mathcal{A}} \overline{r}_i(\mathbf{s}, \mathbf{a}) \hat{\mathbf{x}}_T(t_k; \mathbf{s}, \mathbf{a}),\ i \in \{1, ..., N\},$$

and $\hat{\mathbf{r}}_T(t_k) = (\hat{r}_{T,1}(t_k))_i$. Similarly, we have

$$r_i^* = \sum_{\mathbf{s} \in \mathcal{S}} \sum_{\mathbf{a} \in \mathcal{A}} \overline{r}_i(\mathbf{s}, \mathbf{a}) \mathbf{x}^*(\mathbf{s}, \mathbf{a}),\ i \in \{1, ..., N\}.$$

Again utilizing Lemma 4.1 in [79], we have that for every choice of initial state distribution, there exists constants c_1 and c_2 such that

$$\mathbb{P}(||\hat{\mathbf{r}}_T(t_k) - \mathbf{r}^*|| \geq \delta_1) \leq c_1 e^{-c_2 \delta_1^2 T},\ \forall T \geq 1, \forall \delta_1 > 0. \tag{5.59}$$

Furthermore, convergence of $\hat{\mathbf{r}}_T(t_k)$ to $\mathbf{r}^*$ is w.p. 1. Now let $R_T(t_k) \doteq \sum_i Q_i(t_k)\hat{r}_{T,i}(t_k)$ and $R^*(t_k) \doteq \sum_i Q_i(t_k)r_i^*$. We rewrite the drift expression:

$$\begin{aligned}\frac{\Delta_T(t_k)}{2T} &\leq \frac{NBT}{2} + \sum_i Q_i(t_k)\lambda_i - \mathbb{E}\left[R_T(t_k)\middle|\mathbf{Q}(t_k)\right] \\ &= \frac{NBT}{2} + \sum_i Q_i(t_k)\lambda_i - \sum_i Q_i(t_k)r_i^* \\ &\quad + \mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k)\right]. \end{aligned} \tag{5.60}$$

Now we bound the last term. For all $\delta_2 > 0$ we have

$$\begin{aligned}&\mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k)\right] = \\ &= \mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k), R^*(t_k) - R_T(t_k) \geq \delta_2||\mathbf{Q}(t_k)||\right] \\ &\quad . \mathbb{P}\left(R^*(t_k) - R_T(t_k) \geq \delta_2||\mathbf{Q}(\mathbf{t_k})|| \,\middle|\mathbf{Q}(t_k)\right) \\ &\quad + \mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k), R^*(t_k) - R_T(t_k) < \delta_2||\mathbf{Q}(t_k)||\right] \\ &\quad . \mathbb{P}\left(R^*(t_k) - R_T(t_k) < \delta_2||\mathbf{Q}(\mathbf{t_k})|| \,\middle|\mathbf{Q}(t_k)\right) \\ &\leq \left(\sum_i Q_i(t_k)\right)\mathbb{P}\left(|R^*(t_k) - R_T(t_k)| \geq \delta_2||\mathbf{Q}(t_k)|| \,\middle|\mathbf{Q}(t_k)\right) \\ &\quad + \delta_2||\mathbf{Q}(t_k)||, \end{aligned} \tag{5.61}$$

where we bound the first expectation by $\sum_i Q_i(t_k)$ by using $||\mathbf{r}^*|| < 1$, the second expectation by $\delta_2||\mathbf{Q}(t_k)||$ and the second probability by 1. By Schwartz inequality we have

$$\begin{aligned}&\mathbb{P}\left(|R^*(t_k) - R_T(t_k)| \geq \delta_2||\mathbf{Q}(\mathbf{t_k})|| \,\middle|\mathbf{Q}(t_k)\right) \\ &\leq \mathbb{P}\left(||\mathbf{r}^* - \hat{\mathbf{r}}_T(t_k)|| \geq \delta_2\middle|\mathbf{Q}(t_k)\right). \end{aligned} \tag{5.62}$$

Using (5.59) and (5.62) in (5.61), we have

$$\mathbb{E}[R^*(t_k) - R_T(t_k)|\mathbf{Q}(t_k)] \leq \Big(\sum_i Q_i(t_k)\Big) c_1 e^{-c_2\delta_2^2 T} + \delta_2 ||\mathbf{Q}(t_k)||.$$

Hence, using $||\mathbf{Q}(t_k)|| \leq \sum_i Q_i(t_k)$, we bound (5.60) as

$$\frac{\Delta_T(t_k)}{2T} \leq \frac{NBT}{2} + \sum_i Q_i(t_k)\lambda_i - \sum_i Q_i(t_k) r_i^* + \Big(\sum_i Q_i(t_k)\Big)\Big(c_1 e^{-c_3\delta_2^2 T} + \delta_2\Big).$$

Therefore, calling $\delta \doteq c_1 e^{-c_3\delta_2^2 T} + \delta_2$, we have

$$\frac{\Delta_T(t_k)}{2T} \leq \frac{NBT}{2} + \sum_i Q_i(t_k)\lambda_i - \sum_i Q_i(t_k) r_i^* + \delta \sum_i Q_i(t_k). \tag{5.63}$$

Now for $\boldsymbol{\lambda}$ strictly inside the *δ-stripped* stability region $\boldsymbol{\Lambda}_s^\delta$, there exist a small $\xi > 0$ such that $\boldsymbol{\lambda} + \xi.\mathbf{1} = \mathbf{r} - \delta\mathbf{1}$, for some $\mathbf{r} \in \boldsymbol{\Lambda}_s$. Utilizing this and the fact that $\sum_i Q_i(t)(r_i - r_i^*) \leq 0$ by definition of the FBDC policy in Algorithm 7, we have,

$$\frac{\Delta_T(t_k)}{2T} \leq \frac{NBT}{2} - \Big(\sum_i Q_i(t_k)\Big)\xi. \tag{5.64}$$

Therefore, the queue sizes have negative drift when $\sum_i Q_i(t_k)$ is larger than $\frac{NBT}{2\xi}$. This establishes stability of the queue sizes at the frame boundaries $t = kT$, $k = \{0, 1, 2, ...\}$ for $\boldsymbol{\lambda}$ within the *δ-stripped* stability region $\boldsymbol{\Lambda}_s^\delta$ (see e.g., [86, Theorem 3]). To see this; note that taking expectation of both sides of (5.64) with respect to $\mathbf{Q}(t_k)$, writing a similar expression over the frame boundaries

$t_k, k \in \{0, 1, 2, ..., K\}$, summing them and telescoping these expressions lead to

$$L(\mathbf{Q}(t_K)) - L(\mathbf{Q}(0)) \leq KNBT^2 - 2\xi T \sum_{k=0}^{K-1} \mathbb{E}\left[\sum_i Q_i(t_k)\right].$$

Using $L(\mathbf{Q}(t_K)) \geq 0$ and $L(\mathbf{Q}(0)) = 0$, we have

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \sum_i \mathbb{E}[Q_i(t_k)] \leq \frac{NBT}{2\xi}.$$

For $t \in (t_k, t_{k+1})$ we have $Q_i(t) \leq Q_i(t_k) + \sum_{\tau=0}^{T-1} A_i(t_k + \tau)$. Therefore, $\mathbb{E}[Q_i(t)] \leq \mathbb{E}[Q_i(t_k)] + T\lambda_i \leq \mathbb{E}[Q_i(t_k)] + TA_{\max}$. Therefore, for $T_K \doteq KT$ we have

$$\limsup_{T_K\to\infty} \frac{1}{KT} \sum_{t=0}^{KT-1} \sum_i \mathbb{E}[Q_i(t)]$$

$$\leq \limsup_{K\to\infty} \frac{1}{TK} \sum_{k=0}^{K-1} \sum_i T\mathbb{E}[Q_i(t_k)] + T^2 A_{\max} \leq \frac{(NB + 2A_{\max}\xi)T}{2\xi}.$$

This proves the stability of the overall system.

Finally, $\delta = c_1 e^{-c_3 \delta_2^2 T} + \delta_2$ for any $\delta_2 > 0$. Therefore, choosing δ_2 appropriately (for example, $\delta_2 = T^{-0.5+\delta_3}$ for some small $\delta_3 > 0$), we have that $\delta(T)$ is a decreasing function of T. Therefore, for any $\delta > 0$, we can find T such that the hypothesis of the theorem holds.

Appendix F - Proof of Lemma 18

We first establish that

$$\Psi = \frac{\sum_i Q_i(t)\hat{r}_i}{\sum_i Q_i(t) r_i^*} \geq 0.90.$$

Considering the mappings in tables 5.3 and 5.4, for the regions of ϵ where the OLM policy and the optimal policy "choose" the same corner point, we have $\Psi = 1$. In the following we analyze the ratio Ψ in the regions where the two policies choose different corner points, which we call "discrepant" regions. We will use Q_1 and Q_2 instead of $Q_1(t)$ and $Q_2(t)$ for notational simplicity. We first consider the case $Q_2 > Q_1$, and divide the proof into separate cases for different regions of ϵ values.

Weighted Departure-Rate Ratio Analysis, Case 1: $\epsilon < \epsilon_c$

Note that the following inequality always holds: $\frac{2-\epsilon}{1-\epsilon} > \frac{(1+\epsilon-\epsilon^2)}{(1-\epsilon)}$. However, we have $\frac{2-\epsilon}{1-\epsilon} = \frac{(1-\epsilon)^2}{\epsilon}$ for $\epsilon = \epsilon_t \doteq 0.245$ for the case of $\epsilon < \epsilon_c = 0.293$.

Case 1.1: $\epsilon < \epsilon_t$

For this case we have $\frac{2-\epsilon}{1-\epsilon} < \frac{(1-\epsilon)^2}{\epsilon}$.

Discrepant Region 1: $\frac{(1-\epsilon)^2}{\epsilon} < \frac{Q_2}{Q_1} < \frac{1-\epsilon}{\epsilon}$

In this case the OLM policy chooses the corner point b_1 whereas the optimal policy chooses the corner point b_0. Therefore,

$$\Psi = \frac{Q_1\left(\frac{(1-\epsilon)^2}{4}\right) + Q_2\left(\frac{1}{2} - \frac{\epsilon}{4}\right)}{Q_2\frac{1}{2}} \geq 1 - \frac{\epsilon}{2} + \frac{(1-\epsilon)^2}{2}\frac{\epsilon}{1-\epsilon}$$

$$= 1 - \frac{\epsilon^2}{2} \geq 0.9700.$$

Discrepant Region 2: $\frac{(1+\epsilon-\epsilon^2)}{1-\epsilon} < \frac{Q_2}{Q_1} < \frac{2-\epsilon}{1-\epsilon}$

In this case the OLM policy chooses the corner point b_2 whereas the optimal policy chooses the

corner point b_1. Therefore,

$$\Psi = \frac{Q_1\left(\frac{3}{8} - \frac{\epsilon}{2} + \frac{\epsilon}{8(2-\epsilon)}\right) + Q_2\left(\frac{3}{8} - \frac{\epsilon}{8(2-\epsilon)}\right)}{Q_1\left(\frac{(1-\epsilon)^2}{4}\right) + Q_2\left(\frac{1}{2} - \frac{\epsilon}{4}\right)}$$

$$= \frac{\frac{3}{8} - \frac{\epsilon}{2} + \frac{\epsilon}{8(2-\epsilon)} + \frac{Q_2}{Q_1}\left(\frac{3}{8} - \frac{\epsilon}{8(2-\epsilon)}\right)}{\frac{(1-\epsilon)^2}{4} + \frac{Q_2}{Q_1}\left(\frac{1}{2} - \frac{\epsilon}{4}\right)} \geq 0.9002.$$

This is a minimization of a function of two variables for all possible ϵ values in the interval $0 \leq \epsilon \leq \epsilon_t$, and the ratio $\frac{Q_2}{Q_1}$ in the interval $\frac{(1+\epsilon-\epsilon^2)}{1-\epsilon} < \frac{Q_2}{Q_1} < \frac{2-\epsilon}{1-\epsilon}$.

CASE 1.2: $\epsilon_t < \epsilon < \epsilon_c$

For this case we have $\frac{2-\epsilon}{1-\epsilon} \geq \frac{(1-\epsilon)^2}{\epsilon}$

Discrepant Region 1: $\frac{(2-\epsilon)}{(1-\epsilon)} < \frac{Q_2}{Q_1} < \frac{1-\epsilon}{\epsilon}$

In this case the OLM policy chooses the corner point b_1 whereas the optimal policy chooses the corner point b_0. Therefore,

$$\Psi = \frac{Q_1\left(\frac{(1-\epsilon)^2}{4}\right) + Q_2\left(\frac{1}{2} - \frac{\epsilon}{4}\right)}{Q_2\frac{1}{2}} \geq 1 - \frac{\epsilon}{2} + \frac{(1-\epsilon)^2}{2}\frac{\epsilon}{1-\epsilon}$$

$$= 1 - \frac{\epsilon^2}{2} \geq 0.9500.$$

Discrepant Region 2: $\frac{(1-\epsilon)^2}{\epsilon} < \frac{Q_2}{Q_1} < \frac{2-\epsilon}{1-\epsilon}$

In this case the OLM policy chooses the corner point b_2 whereas the optimal policy chooses the

corner point b_0. Therefore,

$$\Psi = \frac{Q_1\left(\frac{3}{8} - \frac{\epsilon}{2} + \frac{\epsilon}{8(2-\epsilon)}\right) + Q_2\left(\frac{3}{8} - \frac{\epsilon}{8(2-\epsilon)}\right)}{Q_2 \frac{1}{2}}$$

$$\geq \left(\frac{1-\epsilon}{2-\epsilon}\right)\left(\frac{3}{4} - \epsilon + \frac{\epsilon}{4(2-\epsilon)}\right) + \frac{3}{4} - \frac{\epsilon}{4(2-\epsilon)} \geq 0.9150.$$

Discrepant Region 3: $\frac{(1+\epsilon-\epsilon^2)}{1-\epsilon} < \frac{Q_2}{Q_1} < \frac{(1-\epsilon)^2}{\epsilon}$

In this case the OLM policy chooses the corner point b_2 whereas the optimal policy chooses the corner point b_1. Therefore,

$$\Psi = \frac{Q_1\left(\frac{3}{8} - \frac{\epsilon}{2} + \frac{\epsilon}{8(2-\epsilon)}\right) + Q_2\left(\frac{3}{8} - \frac{\epsilon}{8(2-\epsilon)}\right)}{Q_1\left(\frac{(1-\epsilon)^2}{4}\right) + Q_2\left(\frac{1}{2} - \frac{\epsilon}{4}\right)}$$

$$\geq \frac{\frac{3}{8} - \frac{\epsilon}{2} + \frac{\epsilon}{8(2-\epsilon)} + \frac{Q_2}{Q_1}\left(\frac{3}{8} - \frac{\epsilon}{8(2-\epsilon)}\right)}{\frac{(1-\epsilon)^2}{4} + \frac{Q_2}{Q_1}\left(\frac{1}{2} - \frac{\epsilon}{4}\right)} \geq 0.9474.$$

Due to symmetry, the same bounds on Ψ applies for $Q_2 < Q_1$.

Weighted Departure-Rate Ratio Analysis, Case 2: $\epsilon \geq \epsilon_c$

For the case where $\epsilon \geq \epsilon_c$, we have $(1-\epsilon)(3-2\epsilon) \leq (1-\epsilon)/\epsilon$ and $\frac{1-\epsilon}{\epsilon} < \frac{2-\epsilon}{1-\epsilon}$. Therefore, the only discrepant region between the FBDC and the OLM policies for $Q_2 > Q_1$ is given by $(1-\epsilon)(3-2\epsilon) \leq \frac{Q_2}{Q_1} < \frac{1-\epsilon}{\epsilon}$, where for this interval the OLM policy chooses the corner point b_1, whereas the FBDC policy chooses the corner point b_0.

Discrepant Region 1: $(1-\epsilon)(3-2\epsilon) < \frac{Q_2}{Q_1} < \frac{1-\epsilon}{\epsilon}$

In this case the OLM policy chooses the corner point b_1 whereas the optimal policy chooses the

corner point b_0. Therefore,

$$\Psi = \frac{Q_1\left(\frac{3}{8} - \frac{\epsilon}{2} + \frac{\epsilon}{8(2-\epsilon)}\right) + Q_2\left(\frac{3}{8} - \frac{\epsilon}{8(2-\epsilon)}\right)}{Q_2\frac{1}{2}}$$

$$\geq \left(\frac{\epsilon}{1-\epsilon}\right)\left(\frac{3}{4} - \epsilon + \frac{\epsilon}{4(2-\epsilon)}\right) + \frac{3}{4} - \frac{\epsilon}{4(2-\epsilon)} \geq 0.914.$$

Due to symmetry, the same bound on Ψ applies for $Q_2 < Q_1$. Combining all the cases, for all $\epsilon \in [0, 0.5]$, we have that $\Psi \geq 0.90$ for all possible Q_1 and Q_2.

Now the following drift expression for the OLM policy can be derived similarly to the derivation of (5.63) used in the stability proof for the FBDC policy in Appendix E:

$$\frac{\Delta_T(t_k)}{2T} \leq BT + \sum_i Q_i(t_k)\lambda_i - \sum_i Q_i(t_k)\hat{r}_i + \delta_4 \sum_i Q_i(t_k),$$

where $\delta_4(T)$ is a decreasing function of T. Using (5.14)

$$\frac{\Delta_T(t_k)}{2T} \leq BT + \sum_i Q_i(t_k)\lambda_i - 0.9\sum_i Q_i(t_k)r_i^* + \delta_4 \sum_i Q_i(t_k).$$

Using an argument similar to that for (5.64) we have that for (λ_1, λ_2) strictly inside the 0.9 fraction of the δ_4-*stripped* stability region, there exist a small $\xi > 0$ such that $(\lambda_1, \lambda_2) + (\xi, \xi) = 0.9(r_1, r_2) - (\delta_4, \delta_4)$, for some $\mathbf{r} = (r_1, r_2) \in \mathbf{\Lambda_s}$. Substituting this expression for (λ_1, λ_2) and using $\sum_i Q_i(t)(r - r_i^*) \leq 0$ we have,

$$\frac{\Delta_T(t_k)}{T} \leq (B+K)T + 0.9\sum_i Q_i(t_k)(r - r_i^*)$$

$$-\left(\sum_i Q_i(t_k)\right)\delta_4 - \left(\sum_i Q_i(t_k)\right)\xi + \left(\sum_i Q_i(t_k)\right)\delta_4.$$

After cancelations we have,

$$\frac{\Delta_T(t_k)}{2T} \leq BT - \Big(\sum_i Q_i(t_k)\Big)\xi.$$

Therefore, using an argument similar to the stability proof for the FBDC policy in Appendix E, the system is stable for arrival rates within at least the 0.9 fraction of δ_4-*stripped* stability region, where $\delta_4(T)$ is a decreasing function of T.

Appendix G - Proof of Lemma 19

We follow similar steps to the stability region derivation for the case of two queues. In order to obtain an expression for $r_i, i \in \{1, ..., N\}$, we sum the 2^{N-1} equations in (5.16) for which the server location m is i and the channel process of queue i, C_i, is 1. This gives for all $i \in \{1, ..., N\}$

$$p_{10} r_i = - \sum_{\substack{s:\, m=i \\ C_i=1}} \sum_{a \neq i} \mathbf{x}(s,a) + p_{01} \sum_{\substack{s:\, m=i \\ C_i=0}} \mathbf{x}(s;i)$$

$$+ (1-p_{10}) \sum_{\substack{s:\, m\neq i \\ C_i=1}} \mathbf{x}(s;i) + p_{01} \sum_{\substack{s:\, m\neq i \\ C_i=0}} \mathbf{x}(s;i).$$

Summing r_i over all queues and using the normalization condition $\sum_s \sum_a \mathbf{x}(s,a) = 1$, we have

$$\begin{aligned}(p_{10}+p_{01})\sum_{i=1}^{N} r_i \;&=\; p_{01} \;-\; \sum_{i=1}^{N}\sum_{j\neq i}\sum_{\substack{s:\, m=i \\ C_i=1, C_j=0}} \mathbf{x}(s;j) \\ &-\; (p_{01}+p_{10})\sum_{i=1}^{N}\sum_{j\neq i}\sum_{\substack{s:\, m=i \\ C_i=1, C_j=1}} \mathbf{x}(s;j) \\ &+\; (1-p_{01}-p_{10})\sum_{i=1}^{N}\sum_{j\neq i}\sum_{\substack{s:\, m=i \\ C_i=0, C_j=1}} \mathbf{x}(s;j).\end{aligned}$$

From Corollary 3, there exists a stationary-deterministic policy π that solves this LP of maximizing $\sum_i r_i(\mathbf{x})$ over the state-action polytope $\mathbf{X}$. Therefore, under this policy π, at each state, at least one of the actions must have 0 state-action frequency. Therefore, in order to maximize the sum-rate, the terms that have negative contribution to the sum-rate must be zero:

$$(p_{10}+p_{01})\sum_{i=1}^{N} r_i = p_{01} + (1-p_{01}-p_{10})\sum_{i=1}^{N}\sum_{j\neq i}\sum_{\substack{s:\, m=i \\ C_i=0, C_j=1}} \mathbf{x}(s;j). \tag{5.65}$$

Similar to the two-queue case in Appendix A, we utilize the expressions resulting from the fact that the steady state probability of each channel state vector is known. For instance, for $C_0^{(N)} \doteq \mathbb{P}\big((C_1,...,C_N)=(0,...,0)\big) = \frac{p_{10}^N}{(p_{10}+p_{01})^N}$, we have

$$\sum_{i=1}^{N}\sum_{\substack{s:\, m=i \\ C_j=0, \forall j}}\sum_{a\in A} \mathbf{x}(i,a) = C_0^{(N)}.$$

Summing these expressions we obtain

$$\sum_{i=1}^{N} \sum_{j \neq i} \sum_{\substack{s:\, m=i \\ C_i=0, C_j=1}} \mathbf{x}(s;j) = 1 - C_0^{(N)} - \sum_{i=1}^{N} r_i.$$

Combining this expression with (5.65) we obtain

$$\sum_{i=1}^{N} r_i = 1 - C_0^{(N)} - \left(p_{10}(1 - C_0^{(N)}) - p_{01} C_0^{(N)}\right).$$

Chapter 6

Scheduling in Networks with Time-Varying Channels and Reconfiguration Delays

In the previous chapter, we studied the impact of a unit switching delay on stability of wireless uplinks/downlinks with a simple two-state time-varying channel. In this chapter, we consider the optimal scheduling problem for networks subject to arbitrary *time-varying channels*, *reconfiguration delays*, and *interference constraints*. We model the network by a graph consisting of nodes, links, and a set of link interference constraints. The network controller is to decide either to stay with the current link-service configuration or switch to another service configuration based on the channel process and the queue length information, where each decision to reconfigure leaves the network *idle* for an arbitrary but finite duration of time, corresponding to *the reconfiguration delay*. Our system model can be used to abstract single-hop wireless networks as shown in Fig. 6-1 or satellite networks with M servers and N ground stations as shown in Fig. 6-2. Our goal is to study the impact of reconfiguration delays on system stability and optimal algorithms. We show that the simultaneous presence of time-varying channels and reconfiguration delays significantly reduces the system stability region and changes the structure of optimal policies.

We first consider the case of memoryless (i.i.d.) channel processes where we characterize the

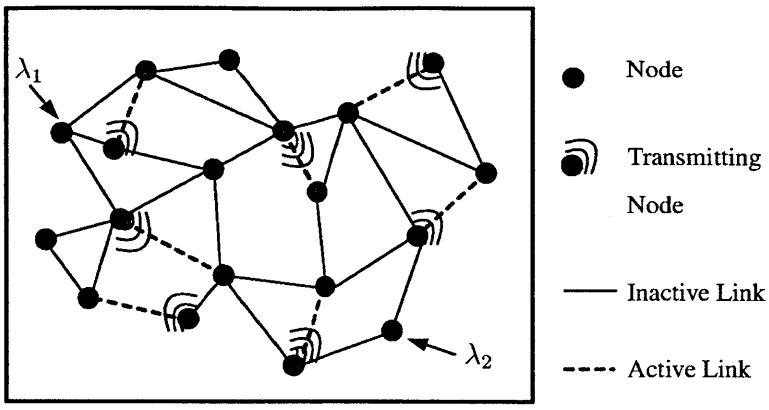


Figure 6-1: System model. A single-hop wireless network with interference constraints, time-varying channels and reconfiguration delays.

stability region in closed form as the convex hull of feasible activation vectors weighted by the *average* channel gain of each link. This result shows that in the presence of reconfiguration delays, it is not possible to opportunistically take advantage of the diversity in time-varying channels because the i.i.d. channel processes refresh during each reconfiguration interval. Moreover, we prove that a Variable Frame Max-Weight (VFMW) scheduling algorithm that sets frame durations as a function of the queue sizes and the *average* channel gains is throughput-optimal.

Next, we consider arbitrary Markov modulated channel processes with memory and characterize the stability region of the system using *state-action frequencies* which are stationary solutions to a Markov Decision Process (MDP) formulation. We show that the stability region enlarges with the memory in the channel processes, which is in contrast to the case of no reconfiguration delays [50], [86], [110]. Moreover, we generalize the Frame-Based Dynamic Control (FBDC) policy of Chapter 5 based on the state-action frequencies, and show that it is throughput-optimal asymptotically in the frame length. The FBDC policy is applicable to a broad class of network control systems, with or without reconfiguration delays, and provides *a new framework for network control by reducing stability region characterization and throughput-optimal algorithm development to solving Linear Programs (LP)* based on state-action frequencies. Finally, we consider simple Myopic policies that do not require the solution of an LP, and that have better delay properties as compared to the FBDC policy.

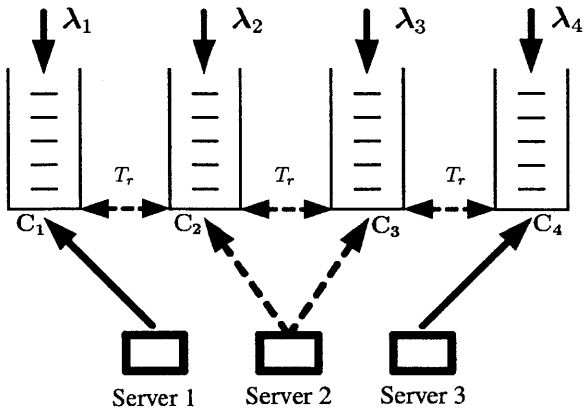


Figure 6-2: An example 4x3 satellite network. Ground stations are subject to time-varying channels $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$, $\mathbf{C}_4$ and the servers are subject to T_r slot reconfiguration (switchover) delay. Server 2 is forced to be idle due to interference constraints.

6.1 Main Contribution and Organization

The main contribution of this chapter is solving the scheduling problem in single-hop networks under arbitrary *reconfiguration delays*, *time-varying channels* and *interference constraints*. We introduce the system model in detail in Section 6.2. For systems with memoryless channel processes, we characterize the stability region and propose the class of throughput-optimal VFMW policies in Section 6.3. We develop the state-action frequency approach and characterize the stability region for systems with Markov modulated channels in Section 6.4.1. We develop the throughput-optimal FBDC policy in Section 6.4.2 and present simulation results in Section 6.4.4.

6.2 Model

Consider a single-hop wireless network given by a graph structure $\mathcal{G}(\mathcal{N}, \mathcal{L})$ of nodes $\mathcal{N}$ and links $\ell \in \mathcal{L} \doteq \{1, 2, ..., L\}$, where $L \doteq |\mathcal{L}|$. Data packets arriving at each link ℓ are to be transmitted to their single-hop destinations, where we refer to the packets waiting for service at link ℓ as queue ℓ. We consider a discrete-time (slotted) system where an integer number of data packets can arrive at or depart from the corresponding queue at each link during each time slot. In addition to the modeling

assumptions in Chapter 2, we assume that each link $\ell \in \mathcal{L}$ is subject to a time-varying channel process denoted by $C_\ell(t)$ that takes values in a set $\mathcal{C} = \{0, \mu_{\min}, ..., \mu_{\max}\}$ with $K \doteq |\mathcal{C}|$, where $C_\ell(t)$ corresponds to the number of packets that can be served from queue ℓ at time t. We consider both memoryless channel processes and Markovian channels with memory as defined below.

Definition 4 (Memoryless Channels) The channel process $\{C_\ell(t); t \geq 0\}$, $\ell \in \mathcal{L}$, takes independent and identically distributed (i.i.d.) values from the set $\mathcal{C}$ at each time slot t, according to a probability distribution for link ℓ, $\mathbb{P}^\ell$.

A simple example of a memoryless channel process is the Bernoulli process with 2-state i.i.d. ON-OFF channels.

Definition 5 (Channels With Memory) The channel process $\{C_\ell(t); t \geq 0\}$ forms the K-state irreducible and aperiodic Markov chain over the set $\mathcal{C}$, according to a transition probability distribution $\mathbb{P}^\ell(\cdot|j), j \in \mathcal{C}$.

The basic example of a Markovian channel process with memory is the commonly used Gilbert-Elliot channel model shown in Fig. 6-3. We let $\overline{C}_\ell$ denote the time-average channel quality of link $\ell, \ell \in \mathcal{L}$, defined by

$$\overline{C}_\ell \doteq \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} C_\ell(\tau). \tag{6.1}$$

The limit exists for both i.i.d. and Markovian channel processes and is equal to the corresponding ensemble (steady state) average with probability (w.p.) 1 due to the Strong Law of Large Numbers (SLLN) [47]. We assume that all the arrival and channel processes, $A_\ell(t)$, $C_\ell(t), \ell \in \mathcal{L}$, are independent.

Let T_r denote the system reconfiguration delay, namely, it takes T_r time slots for the system to change a schedule, during which all the servers are necessarily idle[1]. The set of all schedules in the

[1]Note that in a slotted system, even a minimal reconfiguration delay will lead to a loss of a slot due to synchronization issues

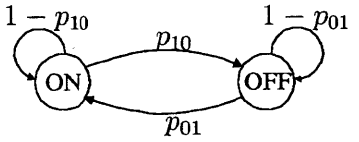


Figure 6-3: Markov modulated ON/OFF channel process. The case of $p_{10} + p_{01} < 1$ provides positive correlation.

system, $\mathcal{I}$, is given by the set of feasible *binary* activation vectors $\mathbf{I} = (\mathbf{I}_\ell)_{\ell=1,\ldots,L}$, $\mathbf{I}_\ell \in \{0,1\}$. If the activation vector $\mathbf{I}(t)$ is used at time slot t, then $\min\{C_\ell(t)\mathbf{I}_\ell(t), Q_\ell(t)\}$ packets depart from queue ℓ. We include the vectors dominated by the feasible activation vectors, as well as the zero vector $\mathbf{I} = \mathbf{0}$ in $\mathcal{I}$, where the activation vector $\mathbf{I}(t)$ is equal to $\mathbf{0}$ for all time slots at which the system is undergoing reconfiguration. A policy π is a mapping from the set of all possible queue length, channel process, and action histories $\mathbb{H}(T) = [\mathbf{Q}(t)]|_{t=0}^{T} \cup [\mathbf{I}(t)]|_{t=0}^{T-1} \cup [\mathbf{C}(t)]|_{t=0}^{T}$, to the set of all probability distributions on the set of all actions $\mathcal{I}$, $\Upsilon(\mathcal{I})$.

The availability of a schedule is determined by the interference constraints in the system, which are assumed to be arbitrary. For instance, in a wireless mesh network as shown in Fig. 6-1, the set $\mathcal{I}$ can be determined according to the well-studied k-hop interference model [50]. Alternatively, for a satellite network of N queues and M servers where there are a possible $L = NM$ links as shown in Fig. 6-2, the set $\mathcal{I}$ can be the set of all binary vectors of dimension NM with at most M nonzero elements such that no two active servers interfere with each other [32]. Finally, for an NxN input-queued optical switch, the set $\mathcal{I}$ can be the set of all matchings [99].

We say that an activation vector $\mathbf{I}$ is *ready to be activated in the current time slot* if the system does not need to reconfigure in order to activate $\mathbf{I}$, i.e., in such a case the servers that will be activated under $\mathbf{I}$ are *present* at their corresponding links at the beginning of the time slot. Finally, we assume that the queues are initially empty and that the arrivals take place after the departures in any given time slot. Under this model, the queue sizes evolve according to the following expression.

$$Q_\ell(t+1) = \max\{Q_\ell(t) - \mathbf{I}_\ell(t)C_\ell(t), 0\} + A_\ell(t), \forall \ell \in \mathcal{L}. \tag{6.2}$$

6.3 Memoryless Channels

6.3.1 Stability Region

We start by characterizing the system stability region for the case of memoryless channels.

Theorem 16 (Stability Region Λ - Memoryless Channels) The stability region $\mathbf{\Lambda}$ is given by

$$\mathbf{\Lambda}=\left\{\boldsymbol{\lambda}\left|\exists\,\boldsymbol{\alpha}\geq\mathbf{0},\sum_{\mathbf{I}\in\mathcal{I}}\alpha_{\mathbf{I}}\leq 1,\textit{such that }\lambda_\ell\leq\overline{C}_\ell\sum_{\mathbf{I}\in\mathcal{I}}\alpha_{\mathbf{I}}\mathbf{I}_\ell,\forall\ell\in\mathcal{L}\right.\right\}. \tag{6.3}$$

The necessity of the conditions in Theorem 16 is proved in Appendix A and the sufficiency of these conditions are proved in the next section where we show that a variable frame-based algorithm that keeps the current activation for a duration of time based on the current queue lengths is throughput-optimal. Theorem 16 shows that in the presence of reconfiguration delays, no policy can take advantage of the diversity in time-varying memoryless channels and achieve a greater rate than the average channel gain for each link. This is because the system cannot switch to another schedule instantly in order to opportunistically exploit better channel states of this schedule, but can switch only after (at least) one time slot and observe an average channel gain upon switching. This is in sharp contrast to the corresponding systems without reconfiguration delay considered in [87], [86] and the references therein, where throughput-optimal policies are able to take advantage of the diversity in i.i.d. channels by instantly and opportunistically switching schedules.

Theorem 16 also establishes that, as long as $T_r \geq 1$, the duration of the reconfiguration interval has no effect on the stability region of the system with memoryless channel processes. This is because for memoryless channels, giving infrequent reconfiguration decisions minimizes the fraction of time slots lost to reconfiguration. In fact, this is the intuition behind the throughput-optimal policy proposed in Section 6.3.2, which delays the reconfiguration decisions as a function of the queue lengths and the channel gains.

6.3.2 Variable Frame Based Max-Weight (VFMW) Algorithm

In this section we propose a throughput-optimal algorithm based on the following intuition: Given that no policy can take advantage of the diversity in channel processes, giving infrequent reconfiguration decisions minimizes throughput lost to reconfiguration. For networks with nonzero reconfiguration delays, in the absence of randomly varying connectivity, we proved in Chapter 4 that a variable-size frame-based Max-Weight algorithm which keeps the same schedule over a frame of duration based on the queue lengths is throughput-optimal. We show here that an adaptation of the algorithm in Chapter 4 that also takes into account *the average channel gains of time-varying links* is throughput-optimal for systems with memoryless channel processes. Specifically, let t_k be the first slot of the kth frame, let $\mathbf{Q}(t_k)$ be the queue lengths at t_k, and let $S(\mathbf{Q}(t_k)) \doteq \sum_i Q_i(t_k)$. The VFMW policy calculates the Max-Weight schedule with respect to $\mathbf{Q}(t_k)$ and $\overline{\mathbf{C}} \doteq (\overline{C}_1, ..., \overline{C}_L)$, and applies this schedule during the frame as defined in detail in Algorithm 13.

Algorithm 13 VFMW ALGORITHM WITH FRAME LENGTH $\chi_k = T_r + F(S(\mathbf{Q}(t_k)))$:

1: Find the Max-Weight schedule at time t_k, $\mathbf{I}^*(t_k)$, w.r.t. $\mathbf{Q}(t_k)$ weighted by the average channel gains $\overline{\mathbf{C}}$:

$$\mathbf{I}^*(t_k) = \arg\max_{\mathbf{I}\in\mathcal{I}} \sum_{\ell} \mathbf{I}_\ell \overline{C}_\ell Q_\ell(t_k)$$

2: If $\mathbf{I}^*(t_k) \neq \mathbf{I}^*(t_{k-1})$, then invoke reconfiguration for the next T_r slots.

3: Apply $\mathbf{I}^*(t_k)$ for an interval of duration $F(S(\mathbf{Q}(t_k)))$ slots where $\chi_k \doteq T_r + F(S(\mathbf{Q}(t_k)))$, $F(\cdot) > 0$ is a monotonically increasing sublinear function, i.e., $\lim_{y\to\infty} F(y)/y = 0$.

4: Repeat above for the next frame starting at $t_{k+1} = t_k + \chi_k$.

The VFMW algorithm sets the frame length as a suitably increasing sublinear function of the queue lengths, which dynamically adapts the frame duration to the stochastic arrivals. For instance, $\chi_k = T_r + (\sum_i Q_i(t_k))^\alpha$ with $\alpha \in (0,1)$ satisfies the criteria for the frame duration. Under the

VFMW policy, the frequency of service reconfiguration is small when the queue lengths are large, limiting the fraction of time spent to switching. Note that this frequency should not be too small otherwise the system becomes unstable as it is subjected to a bad schedule for an extended period of time. When the queue lengths are small, the VFMW policy gives frequent reconfiguration decisions, becoming more adaptive and providing good delay performance[2].

Theorem 17 The VFMW policy stabilizes the system for all arrival rates $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, without requiring knowledge of $\boldsymbol{\lambda}$.

An immediate corollary to this theorem is as follows:

Corollary 5 The conditions in (6.3) are sufficient for stability.

The proof of Theorem 17 is given in Appendix B and is presented using the frame length function $\chi_k = T_r + (\sum_i Q_i(t_k))^\alpha$ for a fixed $\alpha \in (0,1)$ for ease of exposition. It establishes the fact that the drift over the switching epochs, i.e., $\mathbb{E}\big[L(\mathbf{Q}(t_k+\chi_k)) - L(\mathbf{Q}(t_k))\big|\mathbf{Q}(t_k)\big]$, is negative using the quadratic Lyapunov function, $L(\mathbf{Q}(t)) = \sum_{\ell=1}^{L} Q_\ell^2(t)$. The basic intuition behind the proof is that if the queue sizes are large, the VFMW policy accumulates sufficient negative drift during the frame, which overcomes the cost accumulated during reconfiguration. Moreover, for large queue lengths, since the policy keeps the same schedule during the resulting long frames, we obtain the time-average channel gains in the system, as seen in the stability condition in (6.13). Note that choosing the frame length as a *sublinear* function of the queue sizes is critical. This is because the VFMW algorithm uses the Max-Weight schedule corresponding to the beginning of the frame, which "loses weight" as the frame goes on. Therefore, one needs to make sure that the system is not subjected to this "light-weight" schedule for too long. In particular, frame lengths sublinear in queue sizes work, however, frame lengths that are linear in queue sizes do not guarantee stability using the drift analysis approach in this chapter via

[2]Note that a similar policy was considered in [111] to stabilize a very different system, i.e., a system without reconfiguration delays but with asynchronous transmission opportunities.

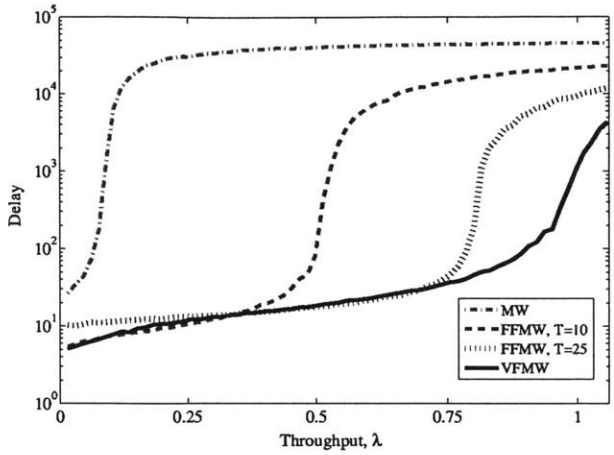


Figure 6-4: Delay vs throughput under the VFMW, MW, and the FFMW policies.

quadratic (or other Polynomial) Lyapunov functions. This is because under frame lengths linear in queue sizes, the drift in the queue lengths over the frames becomes positive under commonly used polynomial Lyapunov functions $L(\cdot)$.

In Section 6.4, we show that for channel processes with memory, delaying the reconfiguration decisions as in the VFMW algorithm does not work and more sophisticated algorithms are necessary in order to exploit the channel memory.

6.3.3 Simulation Results - Memoryless Channels

We performed simulation experiments that determine average queue occupancy values for the VFMW policy, the ordinary Max-Weight (MW) policy and the Max-Weight policy with fixed frame sizes (FFMW), where the MW policy "chooses" the schedule $\arg\max_{\mathbf{I}} \sum_\ell Q_\ell(t) C_\ell(t) \mathbf{I}_\ell$, and the FFMW policy applies the same activation vector as the VFMW policy over frames of constant duration. The average total queue occupancy over T_s slots is defined by $Q_{\text{avg}} \doteq \frac{1}{T_s} \sum_{t=1}^{T_s} \sum_{\ell \in \mathcal{L}} Q_\ell(t)$ and the frame length for the VFMW policy is chosen as $\chi_k = T_r + (\sum_i Q_i(t_k))^{0.9}$. Through Little's law, the average packet delay in the system is equal to the average queue size divided by the total arrival rate into the system. We considered a network of 4 links and 3 servers as shown in Fig. 6-2, where server 1 and 3 are dedicated to links (queues) 1 and 4 respectively, and server 2 is shared between

queues 2 and 3. This system can also model an appropriate single-hop wireless network as in Fig. 6-1. Due to interference constraints, no two links that are "adjacent" to each other can be activated simultaneously, namely, the set of feasible activations are given by $\mathbf{I}^1 = [1010]$, $\mathbf{I}^2 = [0101]$, and $\mathbf{I}^3 = [1001]$. For each data point, the simulation length was 100,000 slots, and the arrival and the channel processes were i.i.d. Bernoulli, with arrival rate λ, and probability of ON channel state equal to 0.5, respectively.

We simulated delay as a function of sum-throughput $\sum_\ell \lambda_\ell$ for $\boldsymbol{\lambda}$ along the line between the origin and the maximum sum-throughput point $\boldsymbol{\lambda}_{\max}$ given by $\arg\max_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \sum_\ell \lambda_\ell$, where from (6.3), $\boldsymbol{\lambda}_{\max}$ can be calculated to be [0.33 0.17 0.17 0.33] with $\sum_\ell \lambda_\ell = 1$. Note that maximum sum-throughput for the corresponding system with zero reconfiguration delay is about 1.44 [50], which shows the significant reduction in throughput due to the reconfiguration delay. Fig. 6-4 presents the delay as a function of sum-throughput for the VFMW, MW, and the FFMW (with frame sizes T= 10 and T= 25) policies, for T_r=5 slot reconfiguration delay. Fig. 6-4 confirms that as the arrival rates are increased, the system quickly becomes unstable under the MW policy and that the VFMW policy provides stability for all sum-rates less than 1. The FFMW policy has larger stability region than that of the MW policy, and increasing the frame length of the FFMW policy improves it's stability region at the expense of delay performance. The VFMW policy provides a good balance by dynamically adapting the frame length based on the queue sizes and stabilizes the system whenever possible, while providing a delay performance that is similar to that of a FFMW policy with a small frame length for small arrival rates.

6.4 Channels With Memory

In this section we establish the stability region of the system and propose a throughput-optimal dynamic control policy when the time-varying channels have memory. We generalize the framework of characterizing the stability region in terms of state-action frequencies that we introduced in Chapter 5 to wireless networks with reconfiguration delays, time-varying channels, and inter-

ference constraints. The state-action frequency approach is a general and unifying framework in that, for the simpler case of no-reconfiguration delay in the system, it provides the stability region characterizations of classical network control papers such as [87], [99], [110].

We show that *the stability region expands with memory in the channel processes*, in particular, it lies between the stability region for the case of i.i.d. channels and the stability region for the case of no reconfiguration delay. For classical network control systems such as [87], [86], [110], the memory in the channel processes does not affect the stability region [50]. Therefore, scheduling under reconfiguration delays and time-varying channels calls for new control algorithms that can improve their performance with increases in channel memory.

6.4.1 Stability Region

We start by analyzing the corresponding system with saturated queues, i.e., all queues are always non-empty. Let $\mathbf{\Lambda}_s$ denote the set of all time average expected departure rate vectors $\mathbf{r} = (r_1, ..., r_L)$ that can be obtained in the saturated system under all possible policies that are possibly history dependent, randomized, or non-stationary. We will show that the stability region $\mathbf{\Lambda}$ satisfies $\mathbf{\Lambda} = \mathbf{\Lambda}_s$. The following Lemma, which follows from Lemma 17 in Chapter 5, proposes an outer bound on the stability region.

Lemma 20 We have that $\mathbf{\Lambda} \subseteq \mathbf{\Lambda}_s$.

We show in the next section that the region $\mathbf{\Lambda}_s$ is indeed achievable.

We establish the region $\mathbf{\Lambda}_s$ by formulating the system dynamics as a Markov Decision Process (MDP).

MDP Formulation For Saturated System

For ease of exposition, we present the analysis for the case of a single slot reconfiguration delay, i.e., $T_r = 1$ [3]. For $T_r = 1$, let $\mathbf{s}_t = (\mathbf{I}(t), \mathbf{C}(t)) \in \mathcal{S}$ denote the system state at time t, where

[3]We demonstrate how to generalize the analysis to the case of $T_r > 1$ whenever appropriate.

$\mathbf{I}(t)$ is the schedule in use at time slot t, $\mathbf{C}(t)$ is the vector of channel processes at each link at time slot t, and $\mathcal{S}$ is the set of all states. Also, let $\mathbf{a}_t \in \mathcal{I}$ denote the action taken at time slot t, which determines the activation vector that will be available at the beginning of the next time slot. For $T_r > 1$ the state would have one more variable that counts the number of time slots since the last reconfiguration decision.

For the saturated system, a policy is a mapping from the set of all possible channel state and action histories to the set of all probability distributions on the set of all actions [10], [79], [94]. Namely, a policy prescribes the probability of any particular action for every given system history. A *stationary* policy is a policy that depends only on the current state, and under a stationary policy the process $\{\mathbf{s}_t; t \geq 0\}$ forms a Markov chain. In each time slot t, the server observes the current state $\mathbf{s}_t$ and chooses an action a_t. Then the next state j is realized according to the transition probabilities $\mathbb{P}(j|\mathbf{s}, \mathbf{a})$, which depend on the random channel processes. Let $\mathbf{I}_\ell$ be 1 if link ℓ is active under the activation vector $\mathbf{I}$, and 0 otherwise. Now, we define the reward for link ℓ as a function of the state $\mathbf{s}_t = (\mathbf{I}(t), \mathbf{C}(t))$ as follows:

$$\overline{r}_l(\mathbf{s}_t, \mathbf{a}_t) \doteq C_l(t), \text{ if } \mathbf{I}_\ell(t) = 1 \text{ and } \mathbf{a}_t \doteq \mathbf{I}(t+1) = \mathbf{I}(t), \tag{6.4}$$

and $\overline{r}_\ell(\mathbf{s}, \mathbf{a}) \doteq 0$ otherwise. That is, a reward of $C_\ell(t)$ is obtained if the controller decides to stay with the current schedule and if link ℓ is active under the current schedule. We are interested in the set of all possible time average expected departure rates. Therefore, given some $\alpha_\ell \geq 0$, $\ell \in \mathcal{L}$, we define the system reward at time t by the weighted sum-throughput $\overline{r}(\mathbf{s}_t, \mathbf{a}_t) \doteq \sum_{\ell \in \mathcal{L}} \alpha_\ell \overline{r}_\ell(\mathbf{s}_t, \mathbf{a}_t)$. The average reward of policy π is defined by

$$r^\pi \doteq \limsup_{K \to \infty} \frac{1}{T} E\Big[\sum_{t=1}^{T} \overline{r}(\mathbf{s}_t, \mathbf{a}_t^\pi)\Big]. \tag{6.5}$$

Given weights $\alpha_\ell \geq 0, \ell \in \mathcal{L}$, we are interested in the policy that achieves the maximum time average expected reward $r^* \doteq \max_\pi r^\pi$. This optimization problem is a discrete time MDP characterized by

the state transition probabilities $\mathbb{P}(.|\mathbf{s},\mathbf{a})$ with $K^L|\mathcal{I}|$ states and $|\mathcal{I}|$ actions per state, where K is the number of channel states. Furthermore, any given pair of states are accessible from each other (i.e., there is a positive probability path between the states) under some stationary deterministic policy. Therefore, this MDP belongs to the class of *Weakly Communicating* MDPs [94], for which there exists a stationary deterministic optimal policy independent of the initial state [94].

State-Action Frequency Approach

The *state-action frequency* approach, or the *Dual Linear Program (LP)* approach, given below provides a systematic and intuitive framework to solve such average cost MDPs [Section 8.8] [94]:

$$\max_{\mathbf{x}\in\mathbf{X}} \cdot \quad \sum_{\mathbf{s}\in\mathcal{S}}\sum_{\mathbf{a}\in\mathcal{I}} \overline{r}(\mathbf{s},\mathbf{a})\mathbf{x}(\mathbf{s},\mathbf{a}) \tag{6.6}$$

subject to the balance equations

$$\sum_{\mathbf{a}\in\mathcal{I}} \mathbf{x}(\mathbf{s},\mathbf{a}) = \sum_{\mathbf{s}'\in\mathcal{S}}\sum_{\mathbf{a}\in\mathcal{I}} \mathbb{P}(\mathbf{s}|\mathbf{s}',\mathbf{a})\mathbf{x}(\mathbf{s}',\mathbf{a}),\ \forall\, \mathbf{s}\in\mathcal{S}, \tag{6.7}$$

the normalization condition $\sum_{\mathbf{s}\in\mathcal{S}}\sum_{\mathbf{a}\in\mathcal{I}}\mathbf{x}(\mathbf{s},\mathbf{a}) = 1$, and the nonnegativity constraints $\mathbf{x}(\mathbf{s},\mathbf{a}) \geq 0$, for $\mathbf{s}\in\mathcal{S}, \mathbf{a}\in\mathcal{I}$, where the transition probabilities $\mathbb{P}(\mathbf{s}|\mathbf{s}',\mathbf{a})$ are functions of the channel transition probabilities. The feasible region of this LP constitutes a polytope called the *state-action polytope* $\mathbf{X}$ and the elements of this polytope $\mathbf{x}\in\mathbf{X}$ are called state-action frequency vectors. A component of a state-action frequency vector, $\mathbf{x}(\mathbf{s},\mathbf{a})$, corresponds to the probability that the system is at state $\mathbf{s}$ and action $\mathbf{a}$ is taken under the following stationary randomized policy: Action $\mathbf{a}$ is taken at state $\mathbf{s}$ w.p.

$$\mathbb{P}(\textit{action } a \text{ at state } \mathbf{s}) = \frac{\mathbf{x}(\mathbf{s},\mathbf{a})}{\sum_{\mathbf{a}'\in\mathcal{I}}\mathbf{x}(\mathbf{s},\mathbf{a}')}, \mathbf{a}\in\mathcal{I}, \mathbf{s}\in S_x, \tag{6.8}$$

where S_x is the set of recurrent states, i.e., S_x is the set of states with positive probability of occupancy in steady state given by [79, 94]

$$S_x \equiv \{\mathbf{s} \in S : \sum_{\mathbf{a} \in \mathcal{I}} \mathbf{x}(\mathbf{s}; \mathbf{a}) > 0\}.$$

If there is a transient state $\mathbf{s}'$, i.e., $\mathbf{s}' \in S/S_x$, then an action that leads the system to the set S_x is chosen at $\mathbf{s}'$. It can be shown that $\mathbf{X}$ is convex, bounded, and closed [79]. Furthermore, every point $\mathbf{x} \in \mathbf{X}$ can be achieved by a stationary randomized policy as in (6.8) [79], [94]. An inverse statement also holds, namely, the expected empirical state-action frequency vector of *any* policy lies in $\mathbf{X}$ regardless of the initial state distribution. The following lemma establishes the equivalence between the corners of the state-action polytope $\mathbf{X}$ and stationary deterministic policies [79], [94].

Lemma 21 The vertices of the LP in (6.6) have a one-to-one correspondence with stationary-deterministic policies.

The intuition behind this lemma is that if $\mathbf{x}$ is a corner point of $\mathbf{X}$, it cannot be expressed as a convex combination of any two other elements in $\mathbf{X}$, therefore, for each state $\mathbf{s}$ only one action has a nonzero steady-state probability of occurrence. Furthermore, the state-action frequencies of stationary randomized policies can be expressed as convex combinations of those for stationary deterministic policies.

The Rate Polytope $\mathbf{\Lambda}_s$

Using the theory of state-action polytopes in the previous section, we characterize the set of all achievable time-average expected rates in the saturated system, $\mathbf{\Lambda}_s$. The following linear transformation of the state-action polytope $\mathbf{X}$ defines the L dimensional *rate polytope* [79]:

$$\mathbf{\Lambda}_s = \left\{\mathbf{r} \middle| r_\ell = \sum_{\mathbf{s} \in \mathcal{S}} \sum_{\mathbf{a} \in \mathcal{I}} \overline{r}_\ell(\mathbf{s}, \mathbf{a}) \mathbf{x}(\mathbf{s}, \mathbf{a}), \ell \in \mathcal{L}\right\}, \tag{6.9}$$

where $\overline{r}_\ell$ is the reward function for link ℓ defined in (6.4). This polytope is the set of all time average expected departure rate pairs that can be obtained in the saturated system, i.e., it is the rate region $\mathbf{\Lambda}_s$. Furthermore, $\mathbf{\Lambda}_s$ is a linear transformation of $\mathbf{X}$. This is because the reward functions, $r_\ell, \ell \in \mathcal{L}$ in (6.9), are linear combinations of $\mathbf{x}(\mathbf{s},\mathbf{a})$, where, given a link ℓ, the coefficient of the linear combination $\overline{r}_\ell(\mathbf{s},\mathbf{a})$ is equal to C_ℓ if the action $\mathbf{a}$ is the same as the current schedule in the state $\mathbf{I}$, and if the lth component of $\mathbf{I}$ is 1 (see (6.4)). Therefore, corner points of $\mathbf{\Lambda}_s$ are also achieved by stationary deterministic policies. An explicit way of characterizing $\mathbf{\Lambda}_s$ is given in Algorithm 14. Note that (6.10) is an LP because $r_\ell(\mathbf{x}), \ell \in \mathcal{L}$, are linear functions of $\mathbf{x}(\mathbf{s},\mathbf{a})$ defined through (6.4),

Algorithm 14 *Stability Region Characterization*

1: Consider the following LP for some $\alpha_1, ..., \alpha_L \geq 0$

$$\max_{\mathbf{x}} . \quad \sum_{\ell=1}^{L} \alpha_\ell r_\ell(\mathbf{x})$$

$$\text{subject to} \quad \mathbf{x} \in \mathbf{X}. \tag{6.10}$$

2: There exists an optimal solution $(r_1^*, ..., r_L^*)$ for this LP that constitutes a corner point of $\mathbf{\Lambda_s}$ and hence of $\mathbf{X}$. Find all possible corner points by evaluating all basic-feasible solutions of this LP, which corresponds to the state-action frequencies of stationary deterministic policies, and take their convex combination.

where $\mathbf{s} = (\mathbf{I}, \mathbf{C})$. The fundamental theorem of Linear Programming guarantees the existence of an optimal solution to (6.10) at a corner point of the polytope $\mathbf{X}$ and hence of $\mathbf{\Lambda}_s$ [15]. We will establish in the next section that the rate region $\mathbf{\Lambda}_s$ is in fact achievable in the dynamic queueing system, which will imply that $\mathbf{\Lambda} = \mathbf{\Lambda}_s$. Furthermore, the one-to-one correspondence between the extreme points of the polytope $\mathbf{X}$ and stationary deterministic polices stated in Lemma 21 is useful for finding the solutions of the above LP. For instance, for the two-queue and single server system introduced in Chapter 5, this LP can be solved explicitly to derive the rate region $\mathbf{\Lambda}_s$. For more complicated systems, the LP in (6.10) can be solved numerically.

6.4.2 Frame Based Dynamic Control Policy

We extend the frame-based dynamic control (FBDC) policy introduced in Chapter 5 for the single server system to single-hop networks. Similar to in Chapter 5. We show that the FBDC policy, inspired by the state-action frequency approach, is throughput-optimal asymptotically in the frame length. The motivation behind the FBDC policy is similar to before: a policy π^* that achieves the optimization in (6.10) for given weights $\alpha_\ell, \ell \in \mathcal{L}$, for the saturated system should achieve a *good* performance in the original system when the queue sizes $\mathbf{Q}$ are used as weights. The FBDC policy is described in detail in Algorithm 15. The LP in (6.11) can be restated as $\max._{\{\mathbf{r}\}} \sum_{\ell=1}^{L} \overline{Q}_\ell(jT) r_\ell$

Algorithm 15 FRAME BASED DYNAMIC CONTROL POLICY

1: Find the policy π^* that optimally solves the following LP

$$\begin{aligned} \max_{\mathbf{x}} \quad & \sum_{\ell=1}^{L} \overline{Q}_\ell(jT) r_\ell(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \mathbf{X} \end{aligned} \tag{6.11}$$

2: Apply π^* in each time slot of the frame.

subject to $\mathbf{r} \in \mathbf{\Lambda}_s$. There exists an optimal solution $\mathbf{r}^*$ of the LP in (6.11) that is a corner point of $\mathbf{X}$ (and hence of $\mathbf{\Lambda}_s$) [15], and the policy π^* that corresponds to this point is a stationary deterministic policy by Lemma 21.

Theorem 18 For any $\delta > 0$, there exists a large enough frame length T such that the FBDC policy stabilizes the system for all arrival rates within the δ-stripped stability region $\mathbf{\Lambda}_s^\delta = \mathbf{\Lambda}_s - \delta\mathbf{1}$.

The proof of this theorem is given in Appendix C and it is similar to the proof of the FBDC policy for the single server system in Chapter 5. This theorem immediately implies that $\mathbf{\Lambda} = \mathbf{\Lambda_s}$.

Recall that the FBDC policy provides a new framework for developing throughput-optimal policies for network control. Namely, given any queuing system whose corresponding saturated system is Markovian with finite state and action spaces, throughput-optimality is achieved by solving an LP

in order to find the stationary MDP solution for the corresponding saturated system and applying this solution over frames in the actual system.

Similar to the analysis of the FBDC policy for the single server system in Chapter 5, we can derive a delay upper bound for the FBDC policy that is linear in the number of links L, and the frame length T. Moreover, the FBDC policy can also be implemented without any frames by setting $T = 1$, i.e., by solving the LP in Algorithm 15 in each time slot. The simulation results in Section 6.4.4 suggest that the FBDC policy implemented without frames has a similar throughput performance to the original FBDC policy. This is because for large queue lengths, the optimal solution of the LP in (6.11) depends on the queue length ratios, and hence, the policy π^* that solves the LP optimally does not change fast when the queue lengths are large. When the policy is implemented without the use of frames, it becomes more adaptive to dynamic changes in the queue lengths, which results in better delay performance than the frame-based implementations.

In the next section, we consider Myopic policies that do not require the solution of an LP. Simulation results in Section 6.4.4 suggest that the stability region achieved by the Myopic policies is close to the full stability region while delay performance of these policies are similar to that of the FBDC policy.

6.4.3 Myopic Control Policies

We investigate the performance of simple *Myopic* policies that make scheduling/switching decisions according to weight functions that are products of the queue lengths and the channel predictions for a small number of slots into the future. We refer to a Myopic policy considering k future time slots as the k-Lookahead Myopic policy. Specifically, in the One-Lookahead Myopic policy, assuming that the system is employing schedule $\mathbf{I}^j$ at some time slot t, $W_{\mathbf{I}^j}$, the weight of $\mathbf{I}^j$ is obtained as follows: If link ℓ is active under $\mathbf{I}^j$, the contribution of link ℓ to $W_{\mathbf{I}^j}(t)$ is the product of $Q_\ell(t)$ and the expectation of C_ℓ in the current and the next time slot. The weight of other schedules are calculated similarly except that we consider the expectation of channel processes only at time slot

$t+1$ because the system will idle at time t in order to switch to other schedules. We describe the One-Lookahead Myopic (OLM) policy in Algorithm 16.

Algorithm 16 ONE-LOOKAHEAD MYOPIC POLICY

1: Assuming that schedule $\mathbf{I}^j$ is currently being used at time slot t, calculate the following weights;

$$W_{\mathbf{I}^j}(t) = \sum_{\ell} \mathbf{I}^j_{\ell} \Big(C_{\ell}(t) + \mathbb{E}\big[C_{\ell}(t+1) | C_{\ell}(t)\big] \Big) Q_{\ell}(t)$$

$$W_{\mathbf{I}^i}(t) = \sum_{\ell} \mathbf{I}^i_{\ell} \mathbb{E}\big[C_{\ell}(t+1) | C_{\ell}(t)\big] Q_{\ell}(t), \ \forall i \neq j. \tag{6.12}$$

2: If $W_{\mathbf{I}^j}(t) \geq W_{\mathbf{I}^i}(t), \ i \neq j, \ i \in \mathcal{I}$, then stay with schedule $\mathbf{I}^j$. Otherwise, switch to a schedule with the maximum weight $W_{\mathbf{I}^i}(t)$.

We investigated the performance of the OLM policy in simulations. The simulation results in Section 6.4.4 suggest that the OLM policy may achieve the full stability region while providing a better delay performance as compared to the FBDC policy.

The k-Lookahead Myopic Policy is the same as before except that the following weight functions are used for scheduling decisions: Assuming that the system is currently employing schedule $\mathbf{I}^j$ at time slot t,

$W_{\mathbf{I}^j}(t) = \sum_{\ell} \mathbf{I}^j_{\ell} \big(C_{\ell}(t) + \sum_{\tau=1}^{k} \mathbb{E}[C_1(t+\tau) | C_1(t)] \big) Q_{\ell}(t)$ and

$W_{\mathbf{I}^i}(t) = \sum_{\ell} \mathbf{I}^i_{\ell} \big(\sum_{\tau=1}^{k} \mathbb{E}[C_{\ell}(t+\tau) | C_{\ell}(t)] \big) Q_{\ell}(t), \ \forall i \neq j.$

These policies have low complexity and they are simpler to implement as compared to the FBDC policy.

6.4.4 Simulation Results - Channels with Memory

We performed simulation experiments that determine average queue occupancy values for the FBDC, One-Lookahead Myopic (OLM), and the Max-Weight (MW) policies. We consider the same system

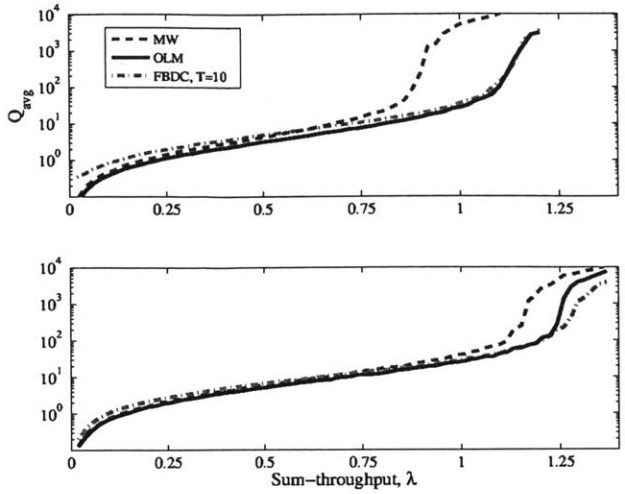


Figure 6-5: Total average queue size vs the sum-throughput under the FBDC policy with frame length 10, the OLM, and the MW policies, for $p_{10} = p_{01} = 0.25$ (upper figure), and for $p_{10} = p_{01} = 0.10$ (lower figure).

and the simulation model as in Section 6.3.3, except that we use the Gilbert-Elliot channel model in Fig. 6-3 and that the switching delay T_r is taken to be 1 slot.

We utilized three sets of transition probabilities, for $p \doteq p_{10} = p_{01}$; $p = 0.10$, 0.25, and 0.30. As for the case of i.i.d. channels considered in Section 6.3.3, the steady state probability of ON channel state for each queue is 0.5 in each of these cases. By numerically solving the LP in (6.10), the maximum sum-throughput can be calculated to be 1.11 for $p = 0.30$, 1.14 for $p = 0.25$, and 1.29 for $p = 0.10$ as we show in Table 6.1. While these values are significantly larger than the maximum sum-throughput of 1 for the case of i.i.d. channels, they are less than the sum-throughput of 1.44 for the corresponding system with zero reconfiguration delays, as expected. The enlargement in the stability region with channel memory is in sharp contrast to systems with zero reconfiguration delays for which the stability region only depends on the steady state behavior of the channel processes [50].

Fig. 6-5(upper figure) presents delay as a function of sum-throughput along the line between the origin and the maximum sum-throughput point for the FBDC policy with frame length 10 and the MW policy for $p_{10} = p_{01} = 0.25$. This figure shows that the system becomes unstable around

Table 6.1: Maximum sum-throughput under different values of channel memory and reconfiguration delay.

p	T_r	max. sum-throughput
0.50	0	1.44
0.50	1	1.00
0.30	1	1.11
0.25	1	1.14
0.10	1	1.29

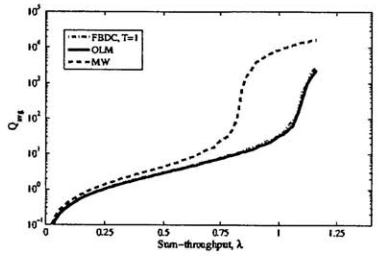


Figure 6-6: Total average queue size vs the sum-throughput under the FBDC policy implemented without frames, the OLM, and the MW policies for $p_{10} = p_{01} = 0.30$.

the sum-throughput value of 0.9 under the MW policy. Moreover, the FBDC policy with frame length 10, and the OLM policy have large queue lengths only for sum-throughputs greater than the maximum sum-throughput value of 1.14 and the throughput loss due to the fixed frame length of 10 appears to be negligible. Furthermore, the OLM policy provides a similar delay performance to the FBDC policy in Fig. 6-5. Fig. 6-5(lower figure) shows delay as a function of sum-throughput for the FBDC policy with frame length 10, the OLM, and the MW policies for $p_{10} = p_{01} = 0.10$. While confirming similar results to Fig. 6-5(upper), Fig. 6-5(lower) also shows that the stability region becomes larger with increasing channel memory.

Fig. 6-6 presents total average queue length as a function of sum-throughput along the line between the origin and the maximum sum-throughput point for the FBDC policy implemented without frames, the OLM, and the MW policies for $p_{10} = p_{01} = 0.30$. This figure shows that the system becomes unstable around the sum-throughput value of 0.84 under the MW policy. Moreover, the FBDC and the OLM policies have large queue lengths only for sum-throughputs greater than the value of 1.11. This figures suggests that the no-frame implementations of the FBDC and the OLM

policies provide a good throughput-delay performance. Furthermore, we observe that the difference in delay between the FBDC and the Max-Weight policies is wider for the no-frame implementation of the FBDC policy in Fig. 6-6, as compared to Fig. 6-5. This suggests that as the frame length is decreased, the delay performance of the FBDC policy improves.

6.5 Concluding Remarks

We investigated the optimal scheduling problem for systems with *reconfiguration delays*, *time-varying channels*, and *interference constraints*. We characterized the stability region of the system in closed form for the case of i.i.d. channel processes and proved that the Variable Frame Max-Weight algorithm that makes scheduling decisions based on the queue lengths and the *average channel gains* is throughput-optimal. For the case of Markovian channels with memory, we characterized the system stability region using state-action frequencies which are stationary solutions to an MDP formulation. We developed the FBDC policy based on the state-action frequencies and proved that it is throughput-optimal asymptotically in the frame length. Finally, we investigated the performance of low-complexity Myopic algorithms that appear to have a similar throughput-delay performance to that of the FBDC policy in simulations. The state-action frequency approach provides a new framework for stability region characterization and throughput-optimal policy development for general network control systems, with or without reconfiguration delays. Possible future directions include developing joint scheduling, routing and power control algorithms in multihop networks with time-varying channels and reconfiguration delays.

Appendix A - Proof of Theorem 16 - Necessity

Let Λ' denote the set at the right hand side of the equality in (6.3). The set Λ' can be rewritten as

$$\Lambda' = \left\{\boldsymbol{\lambda} | \mathbf{C}^{-1}\boldsymbol{\lambda} \in Conv\{\mathcal{I}\}\right\}, \tag{6.13}$$

where $\mathbf{C}$ is a diagonal matrix with the ℓth diagonal element equal to $\overline{C}_\ell$, $\ell \in \mathcal{L}$, and $Conv\{\mathcal{I}\}$ denotes the convex hull of the set of all activations given by

$$Conv\{\mathcal{I}\} \doteq \left\{ \sum_{\mathbf{I}\in\mathcal{I}} \alpha_{\mathbf{I}}\mathbf{I} \;\middle|\; \text{for all } \boldsymbol{\alpha} \geq \mathbf{0} \text{ such that } \sum_{\mathbf{I}\in\mathcal{I}} \alpha_{\mathbf{I}} = 1 \right\}.$$

We show that $\boldsymbol{\Lambda} \subseteq \boldsymbol{\Lambda}'$ by proving that for a fixed $\boldsymbol{\lambda} \notin \boldsymbol{\Lambda}'$, we have $\boldsymbol{\lambda} \notin \boldsymbol{\Lambda}$. Since $\boldsymbol{\lambda} \notin \boldsymbol{\Lambda}'$, and $\boldsymbol{\Lambda}'$ is closed, using the convex set separation theorem [14], there exist a hyperplane, namely a vector $\mathbf{h} \geq \mathbf{0}$ and a constant $b > 0$, such that

$$\boldsymbol{\lambda} \cdot \mathbf{h} > b > (\mathbf{C}\boldsymbol{\mu}) \cdot \mathbf{h}, \quad \forall \boldsymbol{\mu} \in Conv\{\mathcal{I}\}. \tag{6.14}$$

We prove the result by contradiction: Suppose that there exists a stabilizing policy π and consider the following queue evolution equation under π for a link with $\lambda_\ell > 0$, which follows from (6.2):

$$Q_\ell(t) \geq Q_\ell(0) + \sum_{\tau=0}^{t-1} A_\ell(\tau) - \sum_{\tau=0}^{t-1} \mathbf{I}_\ell(\tau) C_\ell(\tau), \forall \ell \in \mathcal{L}. \tag{6.15}$$

Let $\mu_\ell(t) \doteq \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbf{I}_\ell(\tau) C_\ell(\tau)$. The inequality in (6.15) is due to the fact that $t\mu_\ell(t)$ denotes the total *service opportunities* given to queue ℓ until time slot t, which may be larger than the actual number of departures due to the fact that queue ℓ may be empty. Let $\hat{\mathbf{I}}(t)$ be the activation vector that is *ready to be activated* at the beginning of time slot t. Note that the vectors $\hat{\mathbf{I}}(t)$ and $\mathbf{I}(t)$ are the same whenever $\hat{\mathbf{I}}(t)$ gets activated at time slot t, and they are different whenever t is the first time slot of a reconfiguration interval. Let the counting process $M_\ell(t)$ be the number of time slots between times 0 and t for which an activation vector $\hat{\mathbf{I}}$ that activates link ℓ is *ready to be activated* under policy π, and let $t_0, t_1, ..., t_{M_\ell(t)}$ be these time slots. By definition we have $\frac{M_\ell(t)}{t} = \frac{1}{t}\sum_{\tau=0}^{t-1} \hat{\mathbf{I}}_\ell(t)$. If $\lim_{t\to\infty} M_\ell(t) < B$ for some constant B, then the time-average service rate of link ℓ under policy

π is 0 and queue ℓ is unstable, which establishes the contradiction. Therefore, consider $M_\ell(t) \to \infty$ as $t \to \infty$. We have

$$\mu_\ell(t) = \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbf{I}_\ell(\tau) C_\ell(\tau) = \frac{1}{t}\sum_{\tau=t_0}^{t_{M_\ell(t)}} \hat{\mathbf{I}}_\ell(\tau) C_\ell(\tau),$$

$$\leq \frac{M_\ell(t)}{t}\frac{1}{M_\ell(t)}\sum_{\tau=t_0}^{t_{M_\ell(t)}} C_\ell(\tau). \tag{6.16}$$

Since $M_\ell(t) \to \infty$ in the infinite time horizon, the sequence $C_\ell(\tau), \tau \in \{t_0, t_1, t_2, ...\}$, observed at the time slots for which an activation vector $\hat{\mathbf{I}}$ that activates link ℓ is *ready to be activated*, forms an i.i.d. sequence. Therefore, we have from SLLN [47] that $\lim_{t\to\infty} \frac{1}{M_\ell(t)} \sum_{\tau=t_0}^{t_{M_\ell(t)}} C_\ell(\tau) = \overline{C}_\ell$. This implies that for any $\epsilon > 0$, there exist a time slot T such that for all $t > T$ we have $\frac{1}{M_\ell(t)} \sum_{\tau=t_0}^{t_{M_\ell(t)}} C_\ell(\tau) \leq (1+\epsilon)\overline{C}_\ell, \ell \in \mathcal{L}$. Therefore, we have that $\mu_\ell(t) \leq (1+\epsilon)\overline{C}_\ell M_\ell(t)/t, \ell \in \mathcal{L}$ for $t > T$. Using this in (6.15), we have for $t > T$

$$Q_\ell(t) \geq \sum_{\tau=0}^{t-1} A_\ell(\tau) - (1+\epsilon)\overline{C}_\ell M_\ell(t), \forall \ell \in \mathcal{L}. \tag{6.17}$$

Taking expectation of (6.17) with respect to the randomness in arrivals and possibly in policy π, we have for the vector of queue lengths $\mathbf{Q}(t)$ for $t > T$ that

$$\mathbb{E}\left[\mathbf{Q}(t)\right] \geq t\boldsymbol{\lambda} - t(1+\epsilon)\mathbf{C}\,\mathbb{E}\left[\frac{1}{t}\sum_{\tau=0}^{t-1}\hat{\mathbf{I}}(\tau)\right], \tag{6.18}$$

where we used the fact that $\frac{M_\ell(t)}{t} = \frac{1}{t}\sum_{\tau=0}^{t-1}\hat{\mathbf{I}}_\ell(\tau)$ and $\mathbf{C}$ is the diagonal matrix of entries $\overline{C}_\ell, \ell \in \mathcal{L}$. Letting $\hat{\boldsymbol{\mu}}(t) \doteq \mathbb{E}\left[\frac{1}{t}\sum_{\tau=0}^{t-1}\hat{\mathbf{I}}(\tau)\right]$ we have for $t > T$

$$\mathbb{E}\left[\mathbf{Q}(t)\cdot\mathbf{h}\right] \geq t\left(\boldsymbol{\lambda} - (1+\epsilon)(\mathbf{C}\hat{\boldsymbol{\mu}}(t))\right)\cdot\mathbf{h}, \tag{6.19}$$

where $\mathbf{h}$ is the nonnegative valued vector introduced in (6.14). Now for $\boldsymbol{\lambda} \notin \boldsymbol{\Lambda}'$, using (6.14) together with the fact that $\hat{\boldsymbol{\mu}}(t) \in Conv\{\mathcal{I}\}$ for all t [50], there exists $\delta > 0$ such that $(\boldsymbol{\lambda} - (1+\epsilon)(\mathbf{C}\hat{\boldsymbol{\mu}}(t)))\cdot \mathbf{h} > \delta$ for all t for some small ϵ (which can be chosen as a function of δ in (6.17)). Therefore, we have

$$\liminf_{t\to\infty} \mathbb{E}[\mathbf{Q}(t)] = \infty. \tag{6.20}$$

Using (6.20), it is easy to show that the system is unstable with respect to the stability Definition in (1), establishing the desired contradiction. To see this, let $V(\tau) \doteq \inf_{t\geq\tau}\sum_{\ell\in\mathcal{L}} Q_\ell(t)$ and notice that $V(\tau) \leq V(\tau+1)$ for all τ, and as $\tau \to \infty$, $V(\tau) \to \liminf_{t\to\infty}\sum_{\ell\in\mathcal{L}} Q_\ell(t) = \infty$. Therefore, for all $B > 0$, there exists a large time slot T such that $V(\tau) > B$ for all $\tau \geq T$. This implies that $\sum_{\ell\in\mathcal{L}} Q_\ell(t) > B$ for all $t \geq T$. Therefore,

$$\begin{aligned}\frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{\ell\in\mathcal{L}} Q_\ell(\tau) &= \frac{1}{t}\sum_{\tau=0}^{T-1}\sum_{\ell\in\mathcal{L}} Q_\ell(\tau) + \frac{1}{t}\sum_{\tau=T}^{t-1}\sum_{\ell\in\mathcal{L}} Q_\ell(\tau) \\ &\geq \frac{1}{t}\sum_{\tau=T}^{t-1}\sum_{\ell\in\mathcal{L}} Q_\ell(\tau) \geq \frac{t-T}{t}B \geq \frac{k-1}{k}B,\end{aligned}$$

for all $t \geq kT$. Therefore, for all $B > 0$, there exists a time slot T' such that $\frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{\ell\in\mathcal{L}} Q_\ell(\tau) \geq B$ for all $t \geq T'$, which shows that $\frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{\ell\in\mathcal{L}} Q_\ell(\tau)$ diverges to infinity.

Appendix B - Proof of Theorem 17

Consider the following χ_k-step queue evolution expression, similar to the 1-step queue evolution in (6.2). where χ_k is the length of the k^{th} frame length, fixed given $\mathbf{Q}(t_k)$.

$$Q_\ell(t_k+\chi_k) \leq \max\left\{ Q_\ell(t_k) - \sum_{\tau=0}^{\chi_k-1} C_\ell(t_k+\tau)\mathbf{I}_\ell(t_k+\tau)), 0 \right\} + \sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau). \tag{6.21}$$

To see this, note that if $\sum_{\tau=0}^{\chi_k-1} \mathbf{I}_\ell(t_k+\tau)C_\ell(t_k+\tau)$, the total *service opportunity* given to link ℓ during the k^{th} frame, is smaller than $Q_\ell(t_k)$, then we have an equality. Otherwise, the first term is 0 and we have an inequality. This is because some of the arrivals during the frame might depart before the end of the frame. Note that $C_\ell(t)\mathbf{I}_\ell(t)$ is not the actual number of departures from queue ℓ, but it is the service opportunity given to queue ℓ at time slot t under activation $\mathbf{I}(t)$. We first prove stability at the frame boundaries. Squaring both sides, using $\max(0,x)^2 \leq x^2, \forall x \in \mathbb{N} \cup \{0\}$, and $\mathbf{I}_\ell(t)C_\ell(t) \leq \mu_{\max}, \forall t$ we have

$$Q_\ell(t_k+\chi_k)^2 - Q_\ell(t_k)^2 \leq \chi_k^2\mu_{\max}^2 + \Big(\sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big)^2 - 2Q_\ell(t_k)\Big(\sum_{\tau=0}^{\chi_k-1}\mathbf{I}_\ell(t_k+\tau)C_\ell(t_k+\tau) - \sum_{\tau=0}^{\chi_k-1} A_\ell(t_k+\tau)\Big). \tag{6.22}$$

Define the quadratic Lyapunov function

$$L(\mathbf{Q}(t)) = \sum_{\ell=1}^{L} Q_\ell^2(t),$$

which represents a quadratic measure of the total load in the system at time slot t. Define the χ_k-step conditional drift

$$\Delta_{\chi_k}(t_k) \triangleq \mathbb{E}\left[L(\mathbf{Q}(t_k+\chi_k)) - L(\mathbf{Q}(t_k))\middle|\mathbf{Q}(t_k)\right],$$

where the conditional expectation is over the randomness in arrivals, channel processes, and possibly the scheduling decisions. For the VFMW algorithm, scheduling decisions are deterministic given the queue lengths at the beginning of the frames. Summing (6.22) over the links, taking conditional expectation, using the assumption $\mathbb{E}[A_\ell(t)^2] \leq A_{\max}^2, \forall t$ (which also implies $\mathbb{E}[A_\ell(t_1)A_\ell(t_2)] \leq \sqrt{\mathbb{E}[A_\ell(t_1)]^2\mathbb{E}[A_\ell(t_2)]^2} \leq A_{\max}^2$ for all t_1 and t_2), we have

$$\begin{aligned}\Delta_{\chi_k}(t_k) \leq & LB\chi_k^2 + 2\chi_k \sum_\ell Q_\ell(t_k)\lambda_\ell \\ & -2\sum_\ell Q_\ell(t_k)\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1}\mathbf{I}_\ell(t_k+\tau)C_\ell(t_k+\tau)\middle|\mathbf{Q}(t_k)\right],\end{aligned} \tag{6.23}$$

where $B \doteq A_{\max}^2 + \mu_{\max}^2$ and we used the fact that the arrival processes are i.i.d. over time, independent of the queue lengths. The VFMW policy makes scheduling decisions once per frame based on the queue sizes at the beginning of the frame. Therefore, given $\mathbf{Q}(t_k)$, the scheduling variables $\mathbf{I}_\ell(t) = \mathbf{I}_\ell$ are deterministic and same for all $t \in \{t_k + T_r, ..., t_{k+1} - 1\}$, and independent of $C_\ell(t)$, the value of the i.i.d. channel process at time t, for all time slots t. Using this and the fact that the system is idle for the first T_r slots of the frame due to reconfiguration, namely, $\mathbf{I}_\ell(t) = 0, t \in \{t_k, ..., t_k + T_r - 1\}, \forall \ell \in \mathcal{L}$, we have,

$$\begin{aligned}\Delta_{\chi_k}(t_k) \leq & \; LB\chi_k^2 + 2\chi_k \sum_\ell Q_\ell(t_k)\lambda_\ell \\ & - 2\sum_\ell Q_\ell(t_k)\,\mathbf{I}_\ell\mathbb{E}\left[\sum_{\tau=T_r}^{\chi_k-1} C_\ell(t_k+\tau)\middle|\mathbf{Q}(t_k)\right].\end{aligned}$$

We have $\mathbb{E}\left[\sum_{\tau=T_r}^{\chi_k-1} C_\ell(t_k+\tau)\middle|\mathbf{Q}(t_k)\right] = (\chi_k - T_r)\overline{C}_\ell$, because the channel processes are i.i.d. independent of the queue lengths for all times slots. Therefore, we have

$$\begin{aligned}\Delta_{\chi_k}(t_k) \;\leq\; & LB\chi_k^2 + 2\chi_k \sum_\ell Q_\ell(t_k)\lambda_\ell \\ & - \; 2(\chi_k - T_r)\sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}_\ell.\end{aligned}$$

Now, we have $\sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}_\ell = \sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}^*{}_\ell(t_k)$ by definition of the VFMW policy in Algorithm 13. From Theorem 16, we have that for any arrival rate vector $\boldsymbol{\lambda}$ strictly inside $\boldsymbol{\Lambda}$, $\mathbf{C}^{-1}\boldsymbol{\lambda} \doteq (\lambda_1/\overline{C}_1, ..., \lambda_L/\overline{C}_L)$ is dominated by some rate vector in the convex hull of the activation vectors $\mathcal{I}$, where $\mathbf{C}$ is a diagonal matrix with elements $\overline{C}_1, ..., \overline{C}_L$. Therefore, for any arrival rate vector $\boldsymbol{\lambda}$ that is strictly inside $\boldsymbol{\Lambda}$, there exist real numbers $\beta^1, ..., \beta^{|\mathcal{I}|}$ such that $\beta^j \geq 0, \forall j \in 1, ..., |\mathcal{I}|$, $\sum_{j=1}^{|\mathcal{I}|} \beta^j = 1 - \epsilon$ for some $\epsilon > 0$ and $\mathbf{C}^{-1}\boldsymbol{\lambda} = \sum_{j=1}^{|\mathcal{I}|} \beta^j \mathbf{I}^j, \forall \ell \in \mathcal{L}$ [26]. Therefore, we have,

$$\begin{aligned}\Delta_{\chi_k}(t_k) \;\leq\; & LB\chi_k^2 + 2\chi_k \mathbf{Q}(t_k).\left(\sum_{j=1}^{|\mathcal{I}|}\beta^j \sum_\ell \mathbf{I}_\ell^j \overline{C}_\ell\right) \\ =\; & LB\chi_k^2 + 2\chi_k \left(\sum_{j=1}^{|\mathcal{I}|}\beta^j \sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}_\ell^j\right) \\ & - \; 2(\chi_k - T_r)\sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}^*{}_\ell(t_k) \\ \leq\; & LB\chi_k^2 + 2\chi_k \left(\sum_{j=1}^{|\mathcal{I}|}\beta^j \sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}_\ell^*(t_k)\right) \\ & - \; 2(\chi_k - T_r)\sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}^*{}_\ell(t_k)\end{aligned}$$

where we used the fact that $\sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}^*{}_\ell(t_k) \geq \sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}_\ell, \forall \mathbf{I} \in \mathcal{I}$ by definition of the VFMW policy in Algorithm 13. Changing the order of the summation in the second term on the right hand side and using $\sum_{j=1}^{|\mathcal{I}|} \beta^j = 1 - \epsilon$ we have

$$\Delta_{\chi_k}(t_k) = \chi_k \left(LB\chi_k - 2\left(\epsilon - \frac{T_r}{\chi_k}\right) \sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}^*{}_\ell(t_k) \right),$$

If $\chi_k = T_r + (\sum_\ell Q_\ell(t_k))^\alpha \leq \frac{T_r}{\epsilon}$ then $\Delta_{\chi_k}(t_k) \leq C_0$ where C_0 is a constant. Otherwise, there exists a small $\delta_1 > 0$ such that $\epsilon - \frac{T_r}{\chi_k} > \delta_1$. Hence, for $\sum_\ell Q_\ell(t_k) > \frac{T_r^{1/\alpha}(1-\epsilon)^{1/\alpha}}{\epsilon^{1/\alpha}}$, we have

$$\Delta_{\chi_k}(t_k) \leq \chi_k^2 LB - 2\chi_k \delta_1 \frac{\overline{C}_{\min}}{L} \sum_\ell Q_\ell(t_k),$$

where we also used the fact that $\sum_\ell Q_\ell(t_k)\overline{C}_\ell \mathbf{I}^*{}_\ell(t_k) \geq \frac{\overline{C}_{\min}}{L} \sum_\ell Q_\ell(t_k)$ with $\overline{C}_{\min} \doteq \min\{\overline{C}_1, ..., \overline{C}_L\}$. Therefore, there exists a constant B_1 such that

$$\Delta_{\chi_k}(t_k) \leq B_1 - \delta \left(\sum_\ell Q_\ell(t_k) \right)^{1+\alpha}, \tag{6.24}$$

where $\delta \triangleq \delta_1 \overline{C}_{\min}/L$. Taking expectations with respect to $\mathbf{Q}(t_k)$, writing a similar expression over the frame boundaries $t_k, k \in \{0, 1, 2, ..., M\}$, summing them and telescoping these expressions leads to

$$L(\mathbf{Q}(t_M)) - L(\mathbf{Q}(0)) \leq MB_1 - \delta \sum_{k=0}^{M-1} \mathbb{E}\left[\left(\sum_\ell Q_\ell(t_k) \right)^{1+\alpha} \right].$$

Using $L(\mathbf{Q}(t_M)) \geq 0$ and $L(\mathbf{Q}(\mathbf{0})) = 0$ we have

$$\frac{1}{M}\sum_{k=0}^{M-1}\mathbb{E}\left[\left(\sum_{\ell}Q_{\ell}(t_k)\right)^{1+\alpha}\right] \leq \frac{B_1}{\delta} < \infty.$$

This implies that

$$\limsup_{M\to\infty}\frac{1}{M}\sum_{k=0}^{M-1}\mathbb{E}\left[\left(\sum_{\ell}Q_{\ell}(t_k)\right)^{1+\alpha}\right] \leq \frac{B_1}{\delta} < \infty. \tag{6.25}$$

This establishes stability (as defined in Definition 1) at the frame boundaries $t_k, k \in \{0, 1, 2, ...\}$.

Now, we have for all frames $k \in \{0, 1, 2, ...\}$,

$$\sum_{\tau=0}^{\chi_k-1}\sum_{i}Q_i(t_k+\tau) \leq \sum_{\tau=0}^{\chi_k-1}\sum_{i}\left(Q_i(t_k)+\sum_{\tau_1=0}^{\chi_k-1}A_i(t_k+\tau_1)\right).$$

Taking conditional expectation we have,

$$\sum_{\tau=0}^{\chi_k-1}\sum_{i}\mathbb{E}\left[Q_i(t_k+\tau)|\mathbf{Q}(t_k)\right] \leq \chi_k\sum_{i}Q_i(t_k)+\chi_k^2\sum_{i}\lambda_i,$$

where we used the fact that arrival processes are i.i.d. and independent of the queue lengths. Recalling $\chi_k = T_r + (\sum_i Q_i(t_k))^{\alpha}$ with $0 < \alpha < 1$ we have

$$\sum_{\tau=0}^{\chi_k-1}\sum_{i}\mathbb{E}\left[Q_i(t_k+\tau)|\mathbf{Q}(t_k)\right] \leq \left(\sum_{i}Q_i(t_k)\right)^{1+\alpha}+T_r\sum_{i}Q_i(t_k)$$
$$+\left(T_r+\left(\sum_{i}Q_i(t_k)\right)^{\alpha}\right)^2\sum_{i}\lambda_i.$$

Now, for any given large T, let K_T be the number of frames up to and including T. We have

$$\sum_{t=0}^{T-1}\sum_{i}\mathbb{E}[Q_i(t)] \leq \sum_{k=0}^{K_T-1}\mathbb{E}\left[\left(\sum_i Q_i(t_k)\right)^{1+\alpha} + T_r\sum_i Q_i(t_k)\right]$$

$$+\sum_{k=0}^{K_T-1}\mathbb{E}\left[\left(T_r+(\sum_i Q_i(t_k))^{\alpha}\right)^2\right]\sum_i \lambda_i.$$

Dividing both sides by T, using $T > K_T$ for any T, taking the $\limsup_T$ of both sides, using (6.25) and $0 < \alpha < 1$, we have

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\mathbb{E}\left[Q_i(t)\right] < \infty. \tag{6.26}$$

Therefore, the system is stable.

Appendix C - Proof of Theorem 18

Let t_k be the first slot of the kth frame where $t_{k+1} = t_k + T$. The following T-step queue evolution similar to (6.21) holds.

$$Q_\ell(t_k+T) \leq \max\left\{Q_\ell(t_k) - \sum_{\tau=0}^{T-1} C_\ell(t_k+\tau)\mathbf{I}_\ell(t_k+\tau), 0\right\}$$

$$+ \sum_{\tau=0}^{T-1} A_\ell(t_k+\tau). \tag{6.27}$$

Let $D_\ell(t) \doteq C_\ell(t)\mathbf{I}_\ell(t)$, where $\sum_{\tau=0}^{T-1} D_\ell(t_k+\tau)$ is the total *service opportunity* given to link ℓ during the k^{th} frame, similar to (6.21). Note that $\sum_{\tau=0}^{T-1} D_\ell(t_k+\tau)$ denotes the link ℓ departures that would happen in the corresponding *saturated system* if we were to apply the *same* reconfiguration decisions over T time slots in the corresponding saturated system. We first prove stability at the

frame boundaries. Define the quadratic Lyapunov function

$$L(\mathbf{Q}(t)) = \sum_{\ell=1}^{L} Q_\ell^2(t),$$

which represents a quadratic measure of the total load in the system at time slot t. Define the T-step conditional drift

$$\Delta_T(t_k) \triangleq \mathbb{E}\left[L(\mathbf{Q}(t_k+T)) - L(\mathbf{Q}(t_k))\middle|\mathbf{Q}(t_k)\right].$$

The following T-step drift expression follows similar to (6.23) in Appendix A:

$$\Delta_T(t) \leq LBT^2 + 2T\sum_{\ell} Q_\ell(t_k)\lambda_\ell$$

$$-2\sum_{\ell} Q_\ell(t_k)\mathbb{E}\left[\sum_{\tau=0}^{T-1} D_\ell(t_k+\tau)\middle|\mathbf{Q}(t_k)\right], \tag{6.28}$$

where $B \doteq A_{\max}^2 + \mu_{\max}^2$. Recall the definition of the reward functions $\overline{r}_\ell(\mathbf{s}_t, \mathbf{a}_t)$ in (6.4) and let $\overline{r}_\ell(\mathbf{s}_t, \mathbf{a}_t)$ be the reward function associated with applying policy π^* given in the definition of the FBDC policy in Algorithm 15 to the saturated system. Let $\overline{r}_\ell(t)$ denote $\overline{r}_\ell(\mathbf{s}_t, \mathbf{a}_t)$ for notational simplicity, $\ell \in \mathcal{L}$. Note that $r_\ell(t)$ is equal to $D_\ell(t)$, since $D_\ell(t)$ is the *service opportunity* given to link ℓ at time slot t. Now let $\mathbf{r}^* = (r_\ell^*)_\ell$ be the infinite horizon average rate associated with policy π^*. Let $\mathbf{x}^*$ be the optimal vector of state-action frequencies corresponding to π^*. Define the time-average empirical reward from queue ℓ in the saturated system, $\hat{r}_{T,\ell}(t_k)$, $\ell \in \mathcal{L}$ by

$$\hat{r}_{T,\ell}(t_k) \doteq \frac{1}{T}\sum_{\tau=0}^{T-1} \overline{r}_\ell(t_k+\tau).$$

Similarly, define the time average empirical state-action frequency vector $\hat{\mathbf{x}}_T(t_k; \mathbf{s}, \mathbf{a})$.

$$\hat{\mathbf{x}}_T(t_k; \mathbf{s}, \mathbf{a}) \doteq \frac{1}{T} \sum_{\tau=t_k}^{t_k+T-1} I_{\{\mathbf{s}_\tau=\mathbf{s}, \mathbf{a}_\tau=\mathbf{a}\}}, \tag{6.29}$$

where I_E is the indicator function of an event E, i.e., $I_E = 1$ if E occurs and $I_E = 0$ otherwise. Using the definition of the reward functions in (6.4), we have that

$$\hat{r}_{T,\ell}(t_k) = \sum_{\mathbf{s}\in\mathcal{S}} \sum_{\mathbf{a}\in\mathcal{I}} \overline{r}_\ell(\mathbf{s}, \mathbf{a}) \hat{\mathbf{x}}_T(t_k; \mathbf{s}, \mathbf{a}), \ \ell \in \mathcal{L},$$

and $\hat{\mathbf{r}}_T(t_k) = (\hat{r}_{T,1}(t_k))_\ell$. Similarly, we have

$$r^*_\ell = \sum_{\mathbf{s}\in\mathcal{S}} \sum_{\mathbf{a}\in\mathcal{I}} \overline{r}_\ell(\mathbf{s}, \mathbf{a}) \mathbf{x}^*(\mathbf{s}, \mathbf{a}), \ \ell \in \mathcal{L}.$$

Now we utilize the following key MDP theory result in Lemma 4.1 [79], which states that as T increases, $\hat{\mathbf{x}}_T(t_k) = (\hat{\mathbf{x}}_T(t_k; \mathbf{s}, \mathbf{a}))_{\mathbf{s},\mathbf{a}}$ converges to $\mathbf{x}^*$.

Lemma 22 For every choice of initial state distribution, there exists constants c_1 and c_2 such that

$$\mathbb{P}(||\hat{\mathbf{x}}_T(t_k) - \mathbf{x}^*|| \geq \delta_0) \leq c_1 e^{-c_2 \delta_0^2 T}, \quad \forall T \geq 1, \ \forall \delta_0 > 0.$$

Furthermore, convergence of $\hat{\mathbf{x}}_T(t_k)$ to $\mathbf{x}^*$ is with probability (w.p.) 1.

This result applies in our system because every extreme point $\mathbf{x}^*$ of $\mathbf{X}$ *can be* attained by a stationary and deterministic policy that has a single irreducible recurrent class in its underlying Markov chain [79], [94][4]. Due to the linear mapping from the state-action frequencies to the rewards, by Schwartz inequality, each component of $\hat{\mathbf{r}}_T(t_k)$ also converges to the corresponding component of

[4]Note that in general multiple stationary-deterministic policies can yield the same optimal reward vector $\mathbf{r}^*$. Among these, we choose the one that forms a Markov chain with a single recurrent class.

$\mathbf{r}^*$. Therefore, we have that for every choice of initial state distribution, there exists constants c_1 and c_2 such that

$$\mathbb{P}(||\hat{\mathbf{r}}_T(t_k) - \mathbf{r}^*|| \geq \delta_1) \leq c_1 e^{-c_2 \delta_1^2 T}, \ \forall T \geq 1, \forall \delta_1 > 0. \tag{6.30}$$

Furthermore, convergence of $\hat{\mathbf{r}}_T(t_k)$ to $\mathbf{r}^*$ is w.p. 1. Now let $R_T(t_k) \doteq \sum_i Q_\ell(t_k)\hat{r}_{T,\ell}(t_k)$ and $R^*(t_k) \doteq \sum_i Q_\ell(t_k) r^*_\ell$. We rewrite the drift expression:

$$\begin{aligned}
\frac{\Delta_T(t_k)}{2T} &\leq \frac{LBT}{2} + \sum_\ell Q_\ell(t_k)\lambda_i - \mathbb{E}\left[R_T(t_k)\middle|\mathbf{Q}(t_k)\right] \\
&= \frac{LBT}{2} + \sum_\ell Q_\ell(t_k)\lambda_i - \sum_\ell Q_\ell(t_k) r^*_\ell \\
&\quad + \ \mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k)\right].
\end{aligned} \tag{6.31}$$

Now we bound the last term. For all $\delta_2 > 0$ we have

$$\begin{aligned}
&\mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k)\right] = \\
&= \mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k), R^*(t_k) - R_T(t_k) \geq \delta_2||\mathbf{Q}(t_k)||\right] \\
&\quad . \mathbb{P}\left(R^*(t_k) - R_T(t_k) \geq \delta_2||\mathbf{Q}(\mathbf{t_k})|| \ \middle|\mathbf{Q}(t_k)\right) \\
&+ \mathbb{E}\left[R^*(t_k) - R_T(t_k)\middle|\mathbf{Q}(t_k), R^*(t_k) - R_T(t_k) < \delta_2||\mathbf{Q}(t_k)||\right] \\
&\quad . \mathbb{P}\left(R^*(t_k) - R_T(t_k) < \delta_2||\mathbf{Q}(\mathbf{t_k})|| \ \middle|\mathbf{Q}(t_k)\right) \\
&\leq \mu_{\max}\left(\sum_\ell Q_\ell(t_k)\right)\mathbb{P}\left(|R^*(t_k) - R_T(t_k)| \geq \delta_2||\mathbf{Q}(t_k)|| \ \middle|\mathbf{Q}(t_k)\right) \\
&\quad + \ \delta_2||\mathbf{Q}(t_k)||,
\end{aligned} \tag{6.32}$$

where we bound the first expectation by $\sum_\ell Q_\ell(t_k)$ by using $||\mathbf{r}^*|| < \mu_{\max}$, the second expectation by $\delta_2||\mathbf{Q}(t_k)||$ and the second probability by 1. By Schwartz inequality we have

$$\begin{aligned}
&\mathbb{P}\left(|R^*(t_k) - R_T(t_k)| \geq \delta_2||\mathbf{Q}(\mathbf{t_k})|| \,\middle|\, \mathbf{Q}(t_k)\right) \\
&\leq \mathbb{P}\left(||\mathbf{r}^* - \hat{\mathbf{r}}_T(t_k)|| \geq \delta_2 \middle| \mathbf{Q}(t_k)\right). \qquad (6.33)
\end{aligned}$$

Using (6.30) and (6.33) in (6.32), we have

$$\mathbb{E}[R^*(t_k) - R_T(t_k)|\mathbf{Q}(t_k)] \leq \Big(\sum_\ell Q_\ell(t_k)\Big) c_1 e^{-c_2\delta_2^2 T} + \delta_2||\mathbf{Q}(t_k)||.$$

Hence, using $||\mathbf{Q}(t_k)|| \leq \sum_\ell Q_\ell(t_k)$, we bound (6.31) as

$$\begin{aligned}
\frac{\Delta_T(t_k)}{2T} \leq \frac{LBT}{2} &+ \sum_\ell Q_\ell(t_k)\lambda_\ell - \sum_\ell Q_\ell(t_k) r_\ell^* \\
&+ \Big(\sum_\ell Q_\ell(t_k)\Big)\left(c_1 e^{-c_3\delta_2^2 T} + \delta_2\right).
\end{aligned}$$

Therefore, calling $\delta \doteq c_1 e^{-c_3\delta_2^2 T} + \delta_2$, we have

$$\frac{\Delta_T(t_k)}{2T} \leq \frac{LBT}{2} + \sum_\ell Q_\ell(t_k)\lambda_\ell - \sum_\ell Q_\ell(t_k) r_\ell^* + \delta \sum_\ell Q_\ell(t_k).$$

Now for $\boldsymbol{\lambda}$ strictly inside the *δ-stripped* stability region $\boldsymbol{\Lambda}_s^\delta$, there exist a small $\xi > 0$ such that $\boldsymbol{\lambda} + \xi.\mathbf{1} = \mathbf{r} - \delta\mathbf{1}$, for some $\mathbf{r} \in \boldsymbol{\Lambda}_s$. Utilizing this and the fact that $\sum_\ell Q_\ell(t)(r_\ell - r_\ell^*) \leq 0$ by definition of the FBDC policy in Algorithm 15, we have,

$$\frac{\Delta_T(t_k)}{2T} \leq \frac{LBT}{2} - \Big(\sum_\ell Q_\ell(t_k)\Big)\xi. \qquad (6.34)$$

Therefore, the queue sizes have negative drift when $\sum_\ell Q_\ell(t_k)$ is larger than $\frac{LBT}{2\xi}$. Following an argument similar to that for (5.64) in Appendix E of Chapter 5 (in the stability proof of the FBDC algorithm for the single server system), (6.34) establishes stability of the system for $\boldsymbol{\lambda}$ within the *δ-stripped* stability region $\boldsymbol{\Lambda}_s^\delta$. Furthermore, similar to the analysis in Appendix E of Chapter 5, for any $\delta > 0$, we can find T such that the hypothesis of the theorem holds.

Chapter 7

Conclusion and Future Work

We studied the impacts of reconfiguration delays on the control of stochastic networks in this thesis. We showed that reconfiguration delays impact network performance significantly in several ways. In particular, throughput-optimal policies take a very different structure from previously proposed network scheduling algorithms such as the celebrated Max-Weight or Exhaustive policies. We studied networks with time-varying channels and reconfiguration delays, and showed that the simultaneous presence of time varying channels and reconfiguration delays reduces the network stability region significantly, and in contrast to the previous work in the literature, the stability region is improved with increasing memory in the channel processes. We proposed new techniques based on the state-action frequency approach for solving Markov Decision Processes (MDPs) in order to characterize the stability regions of a large class of complex networks and to design throughput-optimal algorithms for such systems.

In the first chapter, we considered the use of controlled mobility and wireless transmission in order to improve the throughput and delay performance of a DTN model, where messages arriving randomly in time and space are gathered by mobile collectors via wireless communications. We developed lower bounds on expected delay as well as matching upper bounds. For interference free DTNs, we showed that the delay scales as $\Theta(\frac{1}{1-\rho})$ with the system load ρ. This is in sharp contrast to the $\Theta(\frac{1}{(1-\rho)^2})$ delay scaling in the corresponding system without wireless transmission. For

DTNs subject to interference of simultaneous transmissions, we simplified the collectors' mobility to a grid to formulate a scheduling problem and characterized the stability region of the system as well as a frame-based algorithm that stabilizes this system.

In the next chapter, we investigated the scheduling problem for networks subject to arbitrary interference constraints and reconfiguration delays. We developed sufficient conditions on the expected interswitching time that leads to stability. We discussed the Variable Frame Max-Weight (VFMW) and the Switching Curve Based (SCB) algorithms that satisfy these conditions and provide throughput-optimality without requiring the knowledge of the arrival rates. These algorithms persist with the Max-Weight schedule during an interval of duration dependent on the queue sizes, which dynamically adapts the interval duration to stochastic arrivals and provides a reasonable delay performance in addition to stability.

In the final two chapters, we considered the optimal scheduling problem for networks subject to time-varying channels and reconfiguration delays, which has not been considered previously in the literature. We first considered the dynamic server allocation problem over parallel queues with *time-varying channels* and *server switching delays*. We analytically characterized the stability region of the system for both i.i.d. and Markovian channel processes and explicitly derived the throughput loss due to switching delays. For the case of Markovian channels, in order to characterize the system stability region and the throughput-optimal FBDC policy, we developed the state-action frequency framework using an MDP formulation for the corresponding saturated system. We also developed simple Myopic Policies that achieve a large fraction of the stability region and provide a better delay performance as compared to the FBDC policy.

Next, we investigated the optimal scheduling problem for networks with arbitrary time-varying channels, reconfiguration delays, and interference constraints. For the case of i.i.d. channel processes, we characterized the stability region of the system in closed form and proved that a Variable Frame Max-Weight algorithm that makes scheduling decisions based on average channel gains is throughput-optimal. For the case of Markovian channels with memory, we generalized the state-action frequency framework to arbitrary single-hop networks, and characterized the system stability

region and the throughput-optimal FBDC policy. The state-action frequency approach provides a new framework for stability region characterization and throughput-optimal policy development for general network control systems, with or without reconfiguration delays.

7.1 Future Directions

We utilized a simple wireless communication model based on a communication range in Chapter 3 and assumed that each collector can receive a single transmission at a time in this chapter. A possible future direction for this work is to consider random access models for transmissions together with more sophisticated communication models that take into account the signal to interference and noise ratio (SINR). Another possible future direction is to relax the grid assumption for the movements of collectors and analyze the resulting joint routing and scheduling problem on the continuous 2-dimensional plane.

We considered centralized algorithms for scheduling in networks subject to reconfiguration delays and interference constraints in Chapter 4. A possible future direction is to develop low-complexity distributed joint scheduling and routing algorithms for multihop networks with interference constraints and nonzero reconfiguration delays.

Finally, developing Myopic policies with low computation complexity and large throughput guarantees can be attractive alternatives to the state-action frequency approach developed in chapters 5 and 6, and developing joint scheduling and routing policies for multihop networks with time-varying channels, reconfiguration delays, and interference constraints is an open future direction.

Bibliography

[1] S. Ahmad, L. Mingyan, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multichannel opportunistic access," *IEEE Trans. Infor. Theory*, vol. 55, no. 9, pp. 4040-4050, Sept. 2009.

[2] S. Ahmad and M. Liu, "Multi-channel opportunistic access: A case of restless bandits with multiple plays," In *Proc. Allerton'09*, Oct. 2009.

[3] I. F. Akyildiz and X. Wang, "Wireless mesh networks", *Wiley*, 2009.

[4] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Ad Hoc Networks (Elsevier)*, vol. 3, no. 3, pp. 257-279, Mar. 2005.

[5] E. Altman, "Constrained Markov decision processes", *Chapman & Hall*, London, 1999.

[6] E. Altman and S. Foss, "Polling on a space with general arrival and service time distribution", *Oper. Res. Let.*, vol. 20, no. 4, pp. 187-194, May 1997.

[7] E. Altman and H. Levy, "Queueing in space", *Adv. Appl. Prob.*, vol. 26, no. 4, pp. 1095-1116, Dec. 1994.

[8] E. Altman, P. Konstantopoulos, and Z. Liu, "Stability, monotonicity and invariant quantities in general polling systems," *Queuing Sys.*, vol. 11, pp. 35-57, 1992.

[9] E. Altman and H. J. Kushner, "Control of polling in presense of vacations in heavy traffic with applications to satellite and mobile radio systems," *SIAM J. on Control and Opt.*, vol. 41, pp. 217-252, 2002.

[10] E. Altman and A. Shwartz, "Markov decision problems and state-action frequencies," *SIAM J. on Control and Opt.*, vol. 29, no 4, pp. 786-809, Jul. 1991.

[11] A. Al-Hanbali, A. A. Kherani, R. Groenevelt, P. Nain, and E. Altman, "Impact of mobility on the performance of relaying in ad hoc networks," In *Proc. IEEE INFOCOM'06*, Apr. 2006.

[12] E. M. Arkin and R. Hassin, "Approximation algorithms for the geometric covering salesman problem," *Discrete Appl. Math.*, vol. 55, pp. 197-218, 1994.

[13] D. Bertsekas and R. Gallager, "Data networks," Prentice Hall, 1992.

[14] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, "Convex analysis and optimization," *Boston: Athena Scientific*, 2003.

[15] D. Bertsimas and J. Tsitsiklis, "Introduction to Linear Optimization", *Athena Scientific*, 1997.

[16] D. J. Bertsimas and G. Van Ryzin, "A stochastic and dynamic vehicle routing problem in the Euclidean plane," *Oper. Res.*, vol. 39, pp. 601-615, 1990.

[17] D. J. Bertsimas and G. Van Ryzin, "Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles," *Oper. Res.*, vol. 41, pp. 60-76, 1993.

[18] D. J. Bertsimas and G. Van Ryzin, "Stochastic and dynamic vehicle routing with general demand and interarrival time distributions," *Adv. App. Prob.*, vol. 20, pp. 947-978, 1993.

[19] P. Billingsley, "Convergence of Probability Measures", Wiley, 1968.

[20] L. Blake and M. Long, "Antennas: Fundamentals, Design, Measurement," *SciTech*, 2009.

[21] S. C. Borst and O. J. Boxma, "Polling models with and without switchover times", *Oper. Res.*, vol. 45, no. 4, pp. 536-543, Aug. 1997.

[22] O. J. Boxma, "Workloads and waiting times in single-server systems with multiple customer classes", *Queueing Syst.*, vol. 5, no. 1-3, pp. 185-214, Nov. 1989.

[23] O.J. Boxma, W.P. Groenendijk, and J.A. Weststrate, "A Pseudoconservation law for service systems with a polling table", *IEEE Trans. Commun.*, vol. 38, no. 10, pp. 18651870, 1990.

[24] O. J. Boxma, H. Levy, and U. Yechiali, "Cyclic reservation schemes for efficient operation of multiple-queue single-server systems", *Ann. Oper. Res.*, vol. 35, no. 14, pp. 187-208, 1992.

[25] A. Brzezinski, "Scheduling algorithms for throughput maximization in data networks," Ph.D. thesis, MIT, 2007.

[26] A. Brzezinski and E. Modiano, "Dynamic reconfiguration and routing algorithms for IP-over-WDM networks with stochastic traffic," *IEEE Journal of Lightwave Tech.*, vol. 23, no. 10, pp. 3188-3205, Oct. 2005

[27] B. Burns, O. Brock, and B. N. Levine, "MV routing and capacity building in disruption tolerant networks," In *Proc. IEEE INFOCOM'05*, Mar. 2005.

[28] P. Chaporkar, K. Kar, and S. Sarkar, "Throughput guarantees through maximal scheduling in wireless networks," In *Proc. Allerton'05*, Sept. 2005.

[29] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Infor. Theory*, vol. 54, no. 5, pp. 20532071, May 2008.

[30] E. G. Coffman and E. N. Gilbert, "Polling and greedy servers on line", *Queueing Sys.* vol. 2, pp. 115-145, 1987

[31] E. G. Coffman and A. Stolyar, "Continuous polling on graphs", *Pro.* vol. 2, pp. 115-145, 1987

[32] G. D. Çelik, S. Borst, P. Whiting, and E. Modiano, "Variable frame based Max-Weight algorithms for networks with switchover delay," In *Proc. IEEE ISIT'11*, Aug. 2011.

[33] G. D. Çelik, L. B. Le, and E. Modiano, "Scheduling in parallel queues with randomly varying connectivity and switchover delay," In *Proc. IEEE INFOCOM'11 (Mini Conference)*, Apr. 2011.

[34] G. D. Çelik and E. Modiano, "Scheduling in n etworks with time-varying channels and reconfiguration delay," In *Proc. IEEE INFOCOM'12*, Mar. 2012.

[35] G. D. Çelik and E. Modiano, "Dynamic vehicle routing for data gathering in wireless networks," In *Proc. IEEE CDC'10*, Dec. 2010.

[36] G. D. Çelik, G. Zussman, W. Khan, and E. Modiano, " MAC for networks with multipacket reception capability and spatially distributed nodes", *IEEE Trans. Mob. Comput.*, vol. 9, no. 2, pp. 226–240, Feb. 2010.

[37] G. D. Çelik, G. Zussman, W. Khan and E. Modiano, "MAC for networks with multipacket reception capability and spatially distributed nodes", *In Proc. IEEE INFOCOM'08*, Apr. 2008.

[38] G. D. Celik, L. B. Le, and E. Modiano, "Scheduling in parallel queues with randomly varying connectivity and switchover delay," *Technical Report*, MIT, Jul. 2010 [Online]. Available: http://web.mit.edu/gcelik/Public

[39] G. D. Çelik and E. Modiano, "Random access wireless networks with controlled mobility", In *Proc. IFIP MEDHOCNET'09*, Jun. 2009.

[40] G. D. Çelik and E. Modiano, "Dynamic vehicle routing for data gathering in wireless networks", *Technical Report, LIDS, MIT*, Jul. 2009.

[41] A. Dumitrescu and J. Mitchell, "Approximation algorithms for tsp with neighborhoods in the plane," In *Proc. ACM-SIAM SODA'01*, Jan. 2001.

[42] A. Dumitrescu and J. Mitchell, "Approximation algorithms for TSP with neighborhoods in the plane," *J. Algorithms*, vol. 48, pp. 135-159, 2003.

[43] A. Eryilmaz, A. Ozdaglar, and E. Modiano, "Polynomial complexity algorithms for full utilization of multi-hop wireless networks," In *Proc. IEEE INFOCOM'07*, May. 2007.

[44] W. Feller, "An introduction to probability theory and its applications," vol. 2, John Wiley & Sons, 2^{nd} Ed., 1971.

[45] S. Foss and G. Last, "Stability of Polling systems with Exhaustive service policies and state-dependent routing", *Ann. of Appl. Probab.*, vol. 6, no. 1, pp. 116-137, Feb. 1996.

[46] E. Frazzoli and F. Bullo, "Decentralized algorithms for vehicle routing in a stochastic time-varying environment," In *Proc. IEEE CDC'04*, Dec. 2004.

[47] R. G. Gallager, "Discrete stochastic processes," Kluwer, 1996, 2nd edition online: http://www.rle.mit.edu/rgallager/notes.htm.

[48] A. E. Gammal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," In *Proc. IEEE INFOCOM'04*, Mar. 2004.

[49] M. Garetto, P. Giaccone, and E. Leonardi, "On the capacity of ad hoc wireless networks under general node mobility," In *Proc. IEEE INFOCOM'07*, May 2007.

[50] L. Georgiadis, M. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," Now Publishers, 2006.

[51] L. Georgiadis and W. Szpankowski, "Stability of token passing rings," *Queueing Sys.*, vol. 11, no. 1-2, pp. 7-33, 1992.

[52] E. N. Gilbert, "Capacity of burst-noise channels," *Bell Syst. Tech. J.*, vol. 39, pp. 12531265, Sept. 1960.

[53] P. W. Glynn and D. Ormoneit, "Hoeffding's inequality for uniformly ergodic Markov chains," *Stat. and Poly. Letters*, vol. 56, pp. 143-146, 2002.

[54] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks ," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 125-137, Feb. 2003.

[55] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[56] M. Haimovich and T. L. Magnanti, "Extremum properties of of hexagonal partitioning and the uniform distribution in euclidian location," *SIAM J. Disc. Math.* 1, 50-64, 1988.

[57] A. Harel and A. Stulman, "Polling, greedy and horizon servers on a circle", *Oper. Res.*, no.43, pp. 177186, 1995.

[58] Hawwar et al., "3G UMTS wireless system physical layer: Baseband processing hardware implementation perspective," *IEEE Communications Magazine - Topics in Radio Communications*, vol. 44, no. 9, pp. 52-58, Sep. 2006.

[59] D. Jea, A. A. Somasundara, and M. B. Srivastava, "Multiple controlled mobile elements (data mules) for data collection in sensor networks," In *Proc. IEEE/ACM DCOSS'05*, Jun. 2008.

[60] K. Kar, X. Luo, and S. Sarkar, "Throughput-optimal scheduling in multichannel access point networks under infrequent channel measurements," In *Proc. IEEE Infocom'07*, May. 2007.

[61] V. Kavitha and E. Altman, "Queueing in space: design of message ferry routes in sensor networks," In *Proc. ITC'09*, Sep. 2009.

[62] V. Kavitha and E. Altman, "Analysis and design of message ferry routes in sensor networks using Polling models," In *Proc. WIOPT'10*, May 2010.

[63] V. Kavitha and E. Altman, "Continuous Polling models and application to ferry assisted WLAN" (online) *to appear in Ann. Oper. Res.*.

[64] A. Khamisy, E. Altman, and M. Sidi,"Pollings systems with synchronization constraints," *Ann. Oper. Res. Special Issue on Stochastic Modeling of Telecommunication Systems*, vol. 35, no. 3, pp. 231-267, 1992.

[65] L. Kleinrock, "Queueing systems: volume 2: computer applications," John Wiley & Sons, 1976.

[66] D. P. Kroese and V. Schmidt, "A continuous polling system with general service times," *Ann. Appl. Probab.*, vol. 2, no. 4, pp. 906-927, 1992.

[67] D. P. Kroese and V. Schmidt, "Single-server queues with spatially distributed arrivals", *Queueing Sys.* vol. 17, no. 1-2, pp. 317-345, 1994.

[68] E. L. Lawler, J.Lenstra, A. Rinnooy Kan, and D. Shmoys, " The Traveling salesman problem :a guided tour of combinatorial optimization," Wiley, 1985.

[69] L. B. Le, E. Modiano, C. Joo, and N. B. Shroff, "Longest-queue-first scheduling under SINR interference model," In *Proc. ACM MobiHoc'10*, Sept. 2010.

[70] J. K. Lenstra, A. H. G. Rinnooy Kan, "Complexity of vehicle routing and scheduling problems," *Networks*, vol. 11, no. 2, pp. 221-227, 1981.

[71] L. Leskelä and F. Unger, "Stability of a spatial polling system with greedy myopic service," In *ArXiv*, 0908.4585v4, Apr. 2010.

[72] H. Levy, M. Sidi, and O.L. Boxma, "Dominance relations in polling systems," *Queueing Systems*, vol. 6, pp. 155-172, Apr. 1990.

[73] C. Li and M. Neely, "On achievable network capacity and throughput-achieving policies over Markov ON/OFF channels," In *Proc. WiOpt'10*, Jun. 2010.

[74] X. Lin and N. B. Shroff, "The fundamental capacity-delay tradeoff in large mobile ad hoc networks," In *Proc. Mediterranean Ad Hoc Networking Workshop'04.*, Jun. 2004.

[75] X. Lin and N. B. Shroff, "The impact of imperfect scheduling on cross-layer rate control in wireless networks," In *Proc. IEEE Infocom'05*, Mar. 2005.

[76] X. Lin, G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks with Brownian mobility," *IEEE/ACM Trans. Netw.*, vol. 14, no. 6, pp. 2777-2784, Jun. 2006.

[77] Z. Liu, P. Nain, and D. Towsley, "On optimal polling policies," *Queuing Sys.*, vol. 11, pp. 59-83, Jul. 1992.

[78] J. Luo and J. P. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," In *Proc. IEEE INFOCOM'05*, Mar. 2005.

[79] S. Mannor and J. N. Tsitsiklis, "On the emperical state-action frequencies in Markov Decision Processes under general policies," *Mathematics of Operation Research*, vol. 30, no. 3, Aug. 2005.

[80] J. S. B. Mitchell, "A ptas for TSP with neighborhoods among fat regions in the plane," In *Proc. ACM-SIAM SODA'07*, Jan. 2007.

[81] E. Modiano and R. Barry, "A novel medium access control protocol for WDM-based LAN's and access networks using a Master/Slave scheduler," *IEEE J. Lightwave Tech.*, vol. 18, no. 4, pp. 461–468, Apr. 2000.

[82] E. Modiano, D. Shah and G. Zussman, "Maximizing throughput in wireless networks via Gossip," In *Proc. ACM SIGMETRICS/Performance'06*, June 2006.

[83] R. M. de Moraes, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "On mobility-capacity-delay trade-off in wireless ad hoc networks," In *Proc. IEEE MASCOTS'04*, Oct. 2004.

[84] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," Morgan and Claypool, 2010.

[85] M. J. Neely, "Stochastic optimization for Markov modulated networks with application to delay constrained wireless scheduling," In *Proc. IEEE CDC'09*, Dec. 2009.

[86] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 89–103, Jan. 2005.

[87] M. J. Neely, E. Modiano, and C. E. Rohrs, "Power allocation and routing in multi-beam satellites with time varying channels," *IEEE Trans. Netw.*, vol. 11, no. 1, pp. 138–152, Feb. 2003.

[88] M. J. Neely and E. Modiano, " Capacity and delay tradeoffs for ad hoc mobile networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1917-1937, Jun. 2005.

[89] M. J. Neely, E. Modiano, and C. E. Rohrs, "Tradeoffs in delay guarantees and computation complexity in nxn packet switches," In *Proc. CISS'02*, Mar. 2002.

[90] J. Le Ny, M. Dahleh, E. Feron, and E. Frazzoli, "Continuous path planning for a data harvesting mobile server," In *Proc. IEEE CDC'08*, Dec. 2008.

[91] A. Pantelidou, A. Ephremides, and A. Tits, "A cross-layer approach for stable throughput maximization under channel state uncertainty," *Wireless Networks*, vol. 15, no.5, pp. 555-569, Jul. 2009.

[92] M. Pavone, N. Bisnik, E. Frazzoli and V. Isler, "Decentralized vehicle routing in a stochastic and dynamic environment with customer impatience," In *Proc. RoboComm'07*, Oct. 2007.

[93] V. H. de la Pea, Z. Govindarajulu, "A note on second moment of a randomly stopped sum of independent variables," *Statistics and Probability Letters*, vol. 14, no. 4, pp. 275–281, Jul. 1992.

[94] M. Puterman, "Markov Decision Processes: Discrete Stochastic Dynamic Programming," *Wiley*, 2005.

[95] V. Raman and N. Vaidya , "SHORT: A static-hybrid approach for routing real time applications over multichannel, multihop wireless networks," In *Proc. WWIC'10*, Jun. 2006.

[96] L. Rojas-Nandayapa, S. Foss, D. P. Kroese, "Stability and performance of greedy server systems - a review and open problems," *Queuing Sys. Theory and Applic.*, vol. 68, no. 3-4, pp. 221-227, Aug. 2011.

[97] G. G. Roussas, "An introduction to measure-theoretic probability," *Gulf Professional Publishing*, 2005.

[98] W. Saad, Z. Han, T. Basar, M. Debbah, and H. Hjorungnes, "A selfish approach to coalition formation among unmanned air vehicles in wireless networks," In *Proc. GAMENETS'09*, May. 2009.

[99] D. Shah and D. J. Wischik, "Optimal scheduling algorithms for input-queued switches," In *Proc. IEEE Infocom'06*, Mar. 2006.

[100] V. Sharma, E. Frazzoli, and P. G. Voulgaris, "Delay in mobility-assisted constant-throughput wireless networks," In *Proc. IEEE CDC'05*, Dec. 2005.

[101] Y. Shi and Y. T. Hou, "Theoretical results on base station movement problem for sensor network," In *Proc. IEEE INFOCOM'08*, Apr. 2008.

[102] K. Sigman and R. Wolf, "A review of regenerative processes," *SIAM Review*, vol. 35, no. 2, pp. 269-288, Jun. 1993

[103] A. L. Stolyar, "Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic," *Annals of Appl. Prob.*, vol. 14, no. 1, pp. 1-53, 2004.

[104] H. Takagi, "Queueing analysis of polling models," *ACM Computing Surveys*, pp. 5-28, no. 1, Mar. 1988.

[105] H. Takagi, "Analysis of Polling Systems", *The MIT Press*, 1986.

[106] M. M. Tariq, M. Ammar, and E. Zegura, "Message ferry route design for sparse ad hoc networks with mobile nodes," In *Proc. ACM MobiHoc'06*, May 2006.

[107] L. Tassiulas, "Adaptive routing on the plane," *Oper. Res.*, vol. 44, no. 5, pp. 823-832, Oct. 1996.

[108] L. Tassiulas, "Adaptive back-pressure congestion control based on local information," *IEEE Trans. on Auto. Control*, vol. 40, no. 2, pp. 236-250, Feb. 1995.

[109] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936-1948, Dec. 1992.

[110] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Infor. Theory*, vol. 39, no. 2, pp. 466-478, Mar. 1993.

[111] L. Tassiulas, S. Papavassiliou, "Optimal anticipative scheduling with asynchronous transmission opportunities," *IEEE Trans. on Auto. Control*, vol. 40, no. 12, pp. 2052-2062, Dec. 1995.

[112] A. Tolkachev, V. Denisenko, A. Shishlov, and A. Shubov, "High gain antenna systems for millimeter wave radars with combined electronical and mechanical beam steering," In *Proc. IEEE Symp. Phased Array Sys. Tech.*, Oct. 2006.

[113] R. Vedantham, S. Kakumanu, S. Lakshmanan, and R. Sivakumar, "Component based channel assignment in single radio, multichannel ad hoc networks," In *Proc. MobiCom'06*, Sep. 2006.

[114] V. M. Vishnevskii and O. V. Semenova, "Mathematical methods to study the polling systems," *Auto. and Rem. Cont.*, vol. 67, no. 2, pp. 173-220, Feb. 2006.

[115] H. Waisanen, D. Shah, M. A. Dahleh, " Minimal delay in controlled mobile relay networks," In *Proc. IEEE CDC'06*, Dec. 2006.

[116] H. Waisanen, D. Shah, M. A. Dahleh, " Lower bounds for multi-stage vehicle routing," In *Proc. IEEE CDC'07*, Dec. 2007.

[117] H. Waisanen, "Control of mobile networks using dynamic vehicle routing," Ph.D. Thesis, MIT, 2007.

[118] J. Walrand, "Queuing Networks," Englewood Cliffs, NJ:Prentice Hall, 1988.

[119] H. Wang and P. Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," *IEEE Trans. Veh. Tech.*, vol. 45, no. 2, pp. 353-357, May 1996.

[120] W. Whitt, "A review of $L = \lambda W$ and extensions," *Queueing Sys.*, vol. 9, no. 3, pp. 235-268, Oct. 1991

[121] M. Z. Win, A. Conti, S. Mazuelas, Y. Shen, W. M. Gifford, and D. Dardari, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56-62, May 2011.

[122] X. Wu and R. Srikant, "Bounds on the capacity region of multi-hop wireless networks under distributed greedy scheduling," In *Proc. IEEE Infocom'06*, Mar. 2006.

[123] H. Xu, "Optimal policies for stochastic and dynamic vehicle routing problems," Ph.D. Thesis, MIT, 1994.

[124] U. Yechiali, "Analysis and control of Polling systems", In *Perform. Eval. Comput. and Commun. Sys.*, Berlin: Springer, no. 729, pp. 630-650, 1993.

[125] L. Ying, and S. Shakkottai, "On throughput-optimality with delayed network-state information," In *Proc. ITA'08*, Jan. 2008.

[126] M. Yun, Y. Zhou, A. Arora, and H. Choi, "Channel-assignment and scheduling in wireless mesh networks considering switching overhead", In *Proc. IEEE ICC'09*, Jun. 2009.

[127] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multichannel opportunistic access: Structure, optimality, and performance," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 54315440, Dec. 2008.

[128] M. Zhao, M. Ma, and Y. Yang, "Mobile data gathering with space-division multiple access in wireless sensor networks," In *Proc. IEEE INFOCOM'08*, Apr. 2008.

[129] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," In *Proc. ACM MobiHoc'04*, May 2004.

[130] ——, "Controlling the mobility of multiple data transport ferries in a delay-tolerant network," In *Proc. IEEE INFOCOM'05*, Mar. 2005.

[131] M. Zorzi, R. Rao, and L. Milstein, "On the accuracy of a first-order Markov model for data transmission on fading channels," In *Proc. ICUPC'95*, 1995.