Opportunities for Network Coding: To Wait or Not to Wait

Yu-Pin Hsu*, Navid Abedini†, Natarajan Gautam‡, Alex Sprintson*, and Srinivas Shakkottai*
*Department of Electrical & Computer Engineering, Texas A&M University
†Qualcomm Flarion Technologies Inc.
‡Department of Industrial & Systems Engineering, Texas A&M University
yupinhsu@tamu.edu, navida@qti.qualcomm.com, {gautam, spalex, sshakkot}@tamu.edu

***Abstract*—It has been well established that wireless network coding can significantly improve the efficiency of multi-hop wireless networks. However, in a stochastic environment some of the packets might not have coding pairs, which limits the number of available coding opportunities. In this context, an important decision is whether to delay packet transmission in hope that a coding pair will be available in the future or transmit a packet without coding. The paper addresses this problem by establishing a stochastic dynamic framework whose objective is to minimize a long-run average cost. We identify an optimal control policy that minimizes the costs due to a combination of transmissions and packet delays. We show that the optimal policy would be stationary, deterministic, and threshold type based on queue lengths. Our analytical approach is applicable for many cases of interest such as time-varying ON/OFF channels. We further substantiate our results with simulation experiments for more generalized settings.**

I. Introduction

In recent years, there has been a growing interest in the applications of network coding techniques in wireless networks. It was shown that network coding can result in significant improvements in the performance in terms of delay and transmission count. For example, consider a wireless network coding scheme depicted in Fig. 1(a). Here, wireless nodes 1 and 2 need to exchange packets x_1 and x_2 through a relay node (node 3). A simple *store-and-forward* approach needs four transmissions. In contrast, the network coding solution uses a *store-code-and-forward* approach in which the two packets x_1 and x_2 are combined by means of a bitwise XOR operation at the relay and are broadcast to nodes 1 and 2 simultaneously. Nodes 1 and 2 can then decode the packets they need from the coded packet and the packets available at these nodes in the beginning of data exchange.

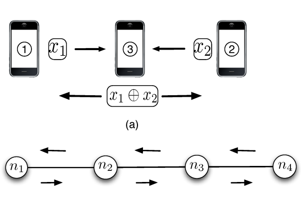

Fig. 1. (a) Wireless Network Coding (b) Reverse carpooling.

An earlier version of this paper was presented at the ISIT, 2011. This material is based upon work partially supported by the NSF under grants CNS-0954153, CNS-0963818, CNS-0904520 and CNS-1149458, by the AFOSR under contract No. FA9550-13-1-0008 and the DTRA under grant HDTRA1-13-1-0030.

Effros et al. [1] introduced the strategy of reverse carpooling that allows two opposite information flows share bandwidth along a shared path. Fig. 1(b) shows an example of two connections, from n_1 to n_4 and from n_4 to n_1 that share a common path (n_1, n_2, n_3, n_4). The wireless network coding approach results in a significant (up to 50%) reduction in the number of transmissions for two connections that use reverse carpooling. In particular, once the first connection is established, the second connection (of the same rate) can be established in the opposite direction with little additional cost.

In this paper, we focus on the design and analysis of scheduling protocols that exploit the fundamental trade-off between the number of transmissions and delay in the reverse carpooling schemes. In particular, to cater to delay-sensitive applications, the network must be aware that savings achieved by coding may be offset by delays incurred in waiting for such opportunities. Accordingly, we design delay-aware controllers that use local information to decide whether or not to wait for a coding opportunity, or to go ahead with an uncoded transmission. By sending uncoded packets we do not take advantage of network coding, resulting in a penalty in terms of transmission count, and, as a result, energy-inefficiency. However, by waiting for a coding opportunity, we might be able to achieve energy efficiency at the cost of a delay increase.

Consider a relay node that transmits packets between two of its adjacent nodes with flows in opposite directions, as depicted in Fig. 2. The relay maintains two queues, q_1 and q_2, such that q_1 and q_2 store packets that need to be delivered to node 2 and node 1, respectively. If both queues are not empty, then it can relay two packets from both queues by performing an XOR operation. However, what should the relay do if one of the queues has packets to transmit, while the other queue is empty? Should the relay wait for a coding opportunity or just transmit a packet from a non-empty queue without coding? This is the fundamental question we seek to answer.

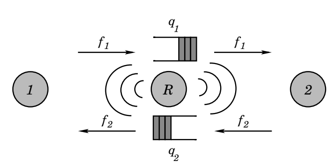

Fig. 2. 3-Node Relay Network.

A. Related Work

Network coding research was initiated by the seminal work of Ahlswede et al. [2] and since then attracted major interest from the research community. Network coding technique for wireless networks has been considered by Katti et al. [3]. They propose an architecture, referred to as COPE, which contains a special network coding layer between the IP and MAC layers. In [4], an opportunistic routing protocol is proposed, referred to as MORE, that randomly mixes packets that belong to the same flow before forwarding them to the next hop. Sagduyu et al. [5] characterize the capacity region for the similar broadcast networks with erasure channels. In addition, several works, e.g., [6–11], investigate the scheduling and/or routing problems in the network coding enabled networks. Sagduyu and Ephremides [6] focus on the network coding in the tandem networks and formulate related cross-layer optimization problems, while Khreishah et al. [7] devise a joint coding-scheduling-rate controller when the pairwise intersession network coding is allowed. Reddy et al. [8] have showed how to design coding-aware routing controllers that would maximize coding opportunities in multihop networks. References [9] and [10] attempt to schedule the network coding between multiple-session flows. Xi and Yeh [11] propose a distributed algorithm that minimizes the transmission cost of a multicast session.

References [12–14] analyze the similar trade-off between power consumption and packet delays from different perspectives. Ciftcioglu et al. [12] propose a threshold policy using the Lyapunov technique. The threshold policy in [12] is an approximate solution with some performance guarantees. Nguyen and Yang [13] present a basic Markov decision process (MDP) framework for the problem at hand. Huang et al. [14] analyze the performance of the transport protocols over meshed networks as well as several implementation issues. In contrast, we focus on the detailed theoretical analysis of the problem at hand, present a provably optimal control policy, and identify its structure.

In this paper, we consider a stochastic arrival process and address the decision problem of whether or not a packet should wait for a coding opportunity. Our objective is therefore to study the delicate trade-off between the energy consumption and the queueing delay when network coding is an option. We use the MDP framework to model this problem and formulate a stochastic dynamic program that determines the optimal control actions in various states. While there exists a large body of literature on the analysis of MDPs (see, e.g., [15–18]), there is no clear methodology to find optimal policies for the problems that possess the proprieties of infinite horizon, average cost optimization, and have a countably infinite state space. Indeed, reference [18] remarks that it is difficult to analyze and obtain optimal policies for such problems. The works in [19–22] contribute to the analysis of MDPs with countably infinite state space. Moreover, reference [23] that surveys the recent results on the monotonic structure of optimal policy, states that while one dimensional MDP with convex cost functions has been extensively studied, limited models for multi-dimensional spaces are dealt with due to the correlations between dimensions. In many high-dimension cases, one usually directly investigates the properties of the cost function. As we will see later, this paper poses precisely such a problem, and showing the properties of optimal solution is one of our main contributions.

B. Main Results

We first consider the case illustrated in Fig. 2, in which we have a single relay node with two queues that contain packets traversing in opposite directions. We assume that time is slotted, and the relay can transmit at most one packet via noiseless broadcast channels during each time slot. We also assume that the arrivals into each queue are independent and identically distributed. Each transmission by the relay incurs a cost, and similarly, each time slot when a packet waits in the queue incurs a certain cost. Our goal is to minimize the weighted sum of the transmission and waiting costs.

We can think of the system state as the two queue lengths. We find that the optimal policy is a simple *queue-length threshold* policy with one threshold for each queue at the relay, and whose action is simply: if a coding opportunity exists, code and transmit; else transmit a packet if the threshold for that queue is reached. We then show how to find the optimal thresholds.

We examine three general models afterward. In the first model, the service capacity of the relay is not restricted to one packet per time slot. Then, if the relay can serve a batch of packets, we find that the optimal controller is of the threshold type for one queue, when the queue length of the other queue is fixed. Secondly, we study an arrival process with memory, i.e., *Markov modulated* arrival process. Here, we discover that the optimal policy has multiple thresholds. Finally, we extend our results for time-varying channels.

We then perform a numerical study of a number of policies that are based on waiting time and queue length, waiting time only, as well as the optimal deterministic queue-length threshold policy to indicate the potential of our approach. We also evaluate the performance of a deterministic queue length based policy in the line network topology via simulations.

Contributions. Our contributions can be summarized as follows. We consider the problem of delay versus coding efficiency trade-off, as well as formulate it as an MDP problem and obtain the structure of the optimal policy. It turns out that the optimal policy does not use the waiting time information. Moreover, we prove that the optimal policy is stationary and based on the queue-length threshold, and therefore is easy to implement. While it is easy to analyze MDPs that have a finite number of states, or involve a discounted total cost optimization with a single communicating class, our problem does not possess any of these properties. Hence, although our policy is simple, the proof is extremely intricate. Furthermore, our policy and proof techniques can be extended to other scenarios such as batched service and Markov-modulated arrival process.

II. System Overview

A. System model

Our first focus is on the case of a single relay node of interest, which has the potential for network coding packets from flows in opposing directions. Consider Fig. 2 again. We assume that there is a flow f_1 that goes from node 1 to 2 and another flow f_2 from node 2 to 1, both of which are through the relay under consideration. The packets from both flows are stored at separate queues, q_1 and q_2, at relay node R.

For clarity of presentation, we assume a simple time division multiple access (TDMA) scheme, however or results are easy to generalize to more involved settings. We assume that time is divided into slots and each slot is further divided into three mini-slots. In each slot, each node is allowed to transmit in its assigned mini-slot: node 1 uses the first mini-slot and node 2 uses the second mini-slot, while the last mini-slot in a slot is used by the relay. In particular, the time period between transmission opportunities for the relay is precisely one slot. Our model is consistent with the scheduled and time synchronized scheme such as LTE. Moreover, we use slot as the unit of packet delays. We assume if a packet is transmitted in the same slot when it arrived at the relay, its latency is zero.

The number of arrivals between consecutive slots to both flows is assumed to be independent of each other and also independent and identically distributed (i.i.d.) over time, with the random variables $\mathcal{A}_i$ for $i = 1, 2$ respectively. In each slot, n packets arrive at q_i with the probability $\mathbb{P}(\mathcal{A}_i = n) = p_n^{(i)}$ for $n \in \mathbb{N} \cup \{0\}$. Afterward, the relay gets an opportunity to transmit. Initially we assume that the relay can transmit a maximum of one packet in each time slot.

B. Markov Decision Process Model

We use a Markov decision process (MDP) model to develop a strategy for the relay to decide its best course of action at every transmission opportunity. For $i = 1, 2$ and $t = 0, 1, 2, \cdots$, let $Q_t^{(i)}$ be the number of packets in q_i at the t^{th} time slot just before an opportunity to transmit. Let a_t be the action chosen at the end of the t^{th} time slot with $a_t = 0$ implying the action is to do nothing and $a_t = 1$ implying the action is to transmit. Clearly, if $Q_t^{(1)} + Q_t^{(2)} = 0$, then $a_t = 0$ because that is the only feasible action. Also, if $Q_t^{(1)} Q_t^{(2)} > 0$, then $a_t = 1$ because the best option is to transmit as a coded XOR packet as it reduces both the number of transmissions as well as latency. However, when exactly one of $Q_t^{(1)}$ and $Q_t^{(2)}$ is non-zero, it is unclear what the best action is.

To develop a strategy for that, we first define the costs for latency and transmission. Let C_T be the cost for transmitting a packet and C_H be the cost of holding a packet for a length of time equal to one slot. The power for transmitting a packet is much higher than the processing energy for network coding because of the simple XOR operation. We therefore ignore the effect of the processing cost. However, to include the processing cost is a small extension and will not change the analytical approach. Hence we assume that the cost of transmitting a coded packet is the same as that of a uncoded packet.

We define the MDP$\{(Q_t, a_t), t \geq 0\}$ where $Q_t = (Q_t^{(1)}, Q_t^{(2)})$ is the state of the system and a_t is the control action chosen by the relay at the t^{th} slot. The state space (i.e., all possible values of Q_t) is the set $\{(i, j) : i = 0, 1, \cdots ; j = 0, 1, \cdots\}$.

Let $C(Q_t, a_t)$ be the *immediate cost* if action a_t is taken at time t when the system is in state $Q_t = (Q_t^{(1)}, Q_t^{(2)})$. Therefore,

$$C(Q_t, a_t) = C_H([Q_t^{(1)} - a_t]^+ + [Q_t^{(2)} - a_t]^+) + C_T a_t, \quad (1)$$

where $[x]^+ = \max(x, 0)$.

C. Average-optimal policy

A *policy* θ specifies the decisions at all decision epoch, i.e., $\theta = \{a_0, a_1, \cdots\}$. A policy is *history dependent* if a_t depends on $a_0, \cdots a_{t-1}$ and $Q_0 \cdots, Q_t$, while that is *Markov* if a_t only depends on Q_t. A policy is *stationary* if $a_{t_1} = a_{t_2}$ when $Q_{t_1} = Q_{t_2}$ for some t_1, t_2. In general, a policy belongs to one of the following sets [15]:

- Π^{HR}: a set of randomized history dependent policies;
- Π^{MR}: a set of randomized Markov policies;
- Π^{SR}: a set of randomized stationary policies;
- Π^{SD}: a set of deterministic stationary policies.

The long-run average cost for some policy $\theta \in \Pi^{\text{HR}}$ is given by

$$V(\theta) = \lim_{K\to\infty} \frac{1}{K+1} \mathbb{E}_\theta \left[\sum_{t=0}^{K} C(Q_t, a_t) | Q_0 = (0, 0) \right], \quad (2)$$

where $\mathbb{E}_\theta$ is the expectation operator taken for the system under policy θ. We consider our initial state to be an empty system, since if we view our system as an ad-hoc network with some initial energy, then the initial state of all queue would be zero to begin with.

Our goal is to characterize and obtain the *average-optimal policy*, i.e., the policy that minimizes $V(\theta)$. It is not hard to see (as shown in [15]) that

$$\Pi^{\text{SD}} \subset \Pi^{\text{SR}} \subset \Pi^{\text{MR}} \subset \Pi^{\text{HR}}.$$

As in [15, 17] there might not exist a SR or SD policy that is optimal, in what regime does the average-optimal policy lie?

We first describe the probability law for our MDP and then in subsequent sections develop a methodology to obtain the average-optimal policy. For the MDP$\{(Q_t, a_t), t \geq 0\}$, let $P_{a_t}(Q_t, Q_{t+1})$ be the transition probability from state Q_t to Q_{t+1} associated with action $a_t \in \{0, 1\}$. Then the probability law can be derived as $P_0((i, j), (k, l)) = p_{k-i}^{(1)} p_{l-j}^{(2)}$ for all $k \geq i$ and $l \geq j$; otherwise, $P_0((i, j), (k, l)) = 0$. Also, $P_1((i, j), (k, l)) = p_{k-[i-1]^+}^{(1)} p_{l-[j-1]^+}^{(2)}$ for all $k \geq [i-1]^+$ and $l \geq [j-1]^+$; otherwise, $P_1((i, j), (k, l)) = 0$.

A list of important notation used in this paper is summarized in Table I.

D. Waiting time information

Intuition tells us that if a packet has not been waiting for a long time then perhaps it could afford to wait a little more, but if a packet has waited for long, it might be better to just transmit it. That seems logical considering that we try our best to code but we cannot wait too long because it hurts in terms of holding costs. It is easy to keep track of waiting time information using time-stamps on packets when they are issued. Let $T^{(i)}$ be the arrival time of i^{th} packet and $\mathcal{D}_\theta^{(i)}$ be its delay (i.e., the waiting time before it is transmitted) while policy θ is applied. We also denote by $\mathcal{T}_{t,\theta}$ the number of transmissions by time t under policy θ. Then Eq. (2) can be

TABLE I
NOTATION TABLE

$\mathcal{A}_i$	Random variable that represents the number of packets that arrives at q_i for each time slot
$p_n^{(i)}$	Probability that n packets arrive at q_i, i.e., $\mathbb{P}(\mathcal{A}_i = n)$
$Q_t^{(i)}$	The number of packets in q_i at time t
Q_t	System state, i.e., $(Q_t^{(1)}, Q_t^{(2)})$
a_t	Action chosen by relay at time t
C_T	Cost of transmitting one packet
C_H	Cost of holding a packet for one time slot
$C(Q_t, a_t)$	Immediate cost if action a_t is taken at time t when the system is in state Q_t
$V(\theta)$	Time average cost under the policy θ
$P_{a_t}(Q_t, Q_{t+1})$	Transition probability from state Q_t to Q_{t+1} when action a_t is chosen
$V_\alpha(i, j, \theta)$	Total expected discounted cost under the policy θ when the initial state is (i, j)
$V_\alpha(i, j)$	Minimum total expected discounted cost when the initial state is (i, j), i.e., $\min_\theta V_\alpha(i, j, \theta)$
$v_\alpha(i, j)$	Difference of the minimum total expected discounted cost between the states (i, j) and $(0, 0)$, i.e., $V_\alpha(i, j) - V_\alpha(0, 0)$
$V_{\alpha,n}(i, j)$	Iterative definition for the optimality equation of $V_\alpha(i, j)$
$\mathcal{V}_\alpha(i, j, a)$	$V_\alpha(i, j) = \min_{a\in\{0,1\}} \mathcal{V}_\alpha(i, j, a)$, which is the optimality equation of $V_\alpha(i, j)$
$\Delta\mathcal{V}(i, j)$	$\mathcal{V}_\alpha(i, j, 1) - \mathcal{V}_\alpha(i, j, 0)$

written as

$$V(\theta) = \lim_{K\to\infty} \frac{1}{K+1} \mathbb{E}_\theta \left[\sum_{i:T^{(i)}\le K} C_H \mathcal{D}_\theta^{(i)} + C_T \mathcal{T}_{K,\theta} \right]. \quad (3)$$

Would we be making better decisions by also keeping track of waiting times of each packet? We can answer this question by applying [15, Theorem 5.5.3].

Proposition 1.

(i) *For the MDP$\{(Q_t, a_t), t \ge 0\}$, if there exists a randomized history dependent policy that is average-optimal then there exists a randomized Markov policy $\theta^* \in \Pi^{MR}$ that minimizes $V(\theta)$.*

(ii) *Further, one cannot find a policy which also uses waiting time information that would yield a better solution than $V(\theta^*)$.*

E. Remarks

To inform nodes 1 and 2 whether the transmitted packet is coded or not, we can just put one bit in front of each packet, where 0 for a uncoded packet and 1 for a coded packet. See [3] for more implementation issues.

In Sections III and IV, we prove that there exists an optimal policy that is stationary, deterministic, and queue-length threshold for the system model of this section. The result will be generalized in Section VII.

- In Subsection VII-A, we consider the batched service, where more than one packet can be served for each time.
- In Subsection VII-B, instead of i.i.d. arrivals, we consider the Markov-modulated arrival process.
- In Subsection VII-C, we consider time-varying channels.

III. STRUCTURE OF THE AVERAGE-OPTIMAL POLICY - STATIONARY AND DETERMINISTIC PROPERTY

In the previous section, we showed that there exists an average-optimal policy that does not include the waiting time in the state of the system. Next, we focus on queue length based and randomized Markov policies, as well as determine the structure of the average-optimal policy. In this section, we will show that there exists an average-optimal policy that is stationary and deterministic.

We begin by considering the infinite horizon α-discounted cost case, where $0 < \alpha < 1$, which we then tie to the average cost case. This method is typically used in the MDP literature (e.g., [22]), where the conditions for the structure of the average-optimal policy usually rely on the results of the infinite horizon α-discounted cost case. For our MDP$\{(Q_t, a_t), t \ge 0\}$, the total expected discounted cost incurred by a policy $\theta \in \Pi^{\text{HR}}$ is

$$V_\alpha(i, j, \theta) = \lim_{K\to\infty} \mathbb{E}_\theta \left[\sum_{t=0}^{K} \alpha^t C(Q_t, a_t) | Q_0 = (i, j) \right]. \quad (4)$$

In addition, we define $V_\alpha(i, j) = \min_\theta V_\alpha(i, j, \theta)$ as well as $v_\alpha(i, j) = V_\alpha(i, j) - V_\alpha(0, 0)$. Define the *$\alpha$-optimal policy* as the policy θ that minimizes $V_\alpha(i, j, \theta)$.

A. Preliminary results

In this subsection, we introduce the important properties of $V_\alpha(i, j)$, which are mostly based on the literature [22]. We first show that $V_\alpha(i, j)$ is finite (Proposition 2) and then introduce the *optimality equation* of $V_\alpha(i, j)$ (Lemma 3).

Proposition 2. *If $\mathbb{E}[\mathcal{A}_i] < \infty$ for $i = 1, 2$, then $V_\alpha(i, j) < \infty$ for every state (i, j) and α.*

Proof: Let $\tilde{\theta}$ be a stationary policy of waiting (i.e., $a_t = 0$ for all t) in each time slot. By definition of optimality, $V_\alpha(i, j) \le V_\alpha(i, j, \tilde{\theta})$. Hence, if $V_\alpha(i, j, \tilde{\theta}) < \infty$, then $V_\alpha(i, j) < \infty$. Note that

$$\begin{aligned} V_\alpha(i, j, \tilde{\theta}) &= \lim_{K\to\infty} \mathbb{E}_{\tilde{\theta}} \Big[\sum_{t=0}^{K} \alpha^t C(Q_t, a_t) | Q_0 = (i, j) \Big] \\ &= \sum_{t=0}^{\infty} \alpha^t C_H \left(i + j + t\mathbb{E}[\mathcal{A}_1 + \mathcal{A}_2] \right) \\ &= \frac{C_H(i + j)}{1 - \alpha} + \frac{\alpha C_H}{(1 - \alpha)^2} \mathbb{E}[\mathcal{A}_1 + \mathcal{A}_2] < \infty. \end{aligned}$$

■

The next lemma follows from Propositions 1 in [22] and the fact that $V_\alpha(i, j)$ is finite (by Proposition 2).

Lemma 3 ([22], Proposition 1)**.** *If $\mathbb{E}[\mathcal{A}_i] < \infty$ for $i = 1, 2$, then the optimal expected discounted cost $V_\alpha(i, j)$ satisfies the* *following optimality equation:*

$$V_\alpha(i,j) = \min_{a\in\{0,1\}}[C_H([i-a]^+ + [j-a]^+) + C_T a + \alpha\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_a\big((i,j),(k,l)\big)V_\alpha(k,l)]. \quad (5)$$

Moreover, the stationary policy that realizes the minimum of right hand side of (5) will be an α-optimal policy.

We define $V_{\alpha,0}(i,j) = 0$ and for $n \geq 0$,

$$V_{\alpha,n+1}(i,j) = \min_{a\in\{0,1\}}[C_H([i-a]^+ + [j-a]^+) + C_T a + \alpha\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_a\big((i,j),(k,l)\big)V_{\alpha,n}(k,l)]. \quad (6)$$

Lemma 4 below follows from Proposition 3 in [22].

Lemma 4 ([22], Proposition 3)**.** *$V_{\alpha,n}(i,j) \to V_\alpha(i,j)$ as $n \to \infty$ for every i, j, and α.*

Eq. (6) will be helpful for identifying the properties of $V_\alpha(i,j)$, e.g., to prove that $V_\alpha(i,j)$ is a non-decreasing function.

Lemma 5. *$V_\alpha(i,j)$ is a non-decreasing function with respect to (w.r.t.) i for fixed j, and vice versa.*

Proof: The proof is by induction on n in Eq. (6). The result clearly holds for $V_{\alpha,0}(i,j)$. Now, assume that $V_{\alpha,n}(i,j)$ is non-decreasing. First, note that $C_H([i-a]^+ + [j-a]^+) + C_T a$ is a non-decreasing function of i and j (since C_H is non-negative). Next, we note that

$$\alpha\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_a\big((i,j),(k,l)\big)V_{\alpha,n}(k,l)$$
$$= \alpha\sum_{r=0}^{\infty}\sum_{s=0}^{\infty} p_r^{(1)}p_s^{(2)}V_{\alpha,n}([i-a]^+ + r, [j-a]^+ + s),$$

which is also a non-decreasing function in i and j separately due to the inductive assumption. Since the sum and the minimum (in Eq. (6)) of non-decreasing functions are a non-decreasing function, we conclude that $V_{\alpha,n+1}(i,j)$ is a non-decreasing function as well. ■

B. Main result

Using the α-discounted cost and the optimality equation, we show that the MDP defined in this paper has an average-optimal policy that is stationary and deterministic.

Theorem 6. *For the MDP$\{(Q_t, a_t), t \geq 0\}$, there exists a stationary and deterministic policy θ^* that minimizes $V(\theta)$ if $\mathbb{E}[\mathcal{A}_i^2] < \infty$ and $\mathbb{E}[\mathcal{A}_i] < 1$ for $i = 1, 2$.*

Proof: See Appendix A. ■

According to Borkar [24], it is possible to find the randomized policy that is closed to the average-optimal by applying linear programming methods for an MDP of a very generic setting, where randomized stationary policies are average-optimal. However, since the average-optimal policy has further been shown in Theorem 6 to be deterministic, in the next section we investigate the structural properties of the average-optimal policy and using a Markov-chain based enumeration to find the average-optimal polity that would be deterministic stationary.

IV. Structure of the Average-Optimal Policy - Threshold Based

Now that we know the average-optimal policy is stationary and deterministic, the question is how do we find it? If we know that the average-optimal policy satisfies the structural properties, then it is possible to search through the space of stationary deterministic policies and obtain the optimal one. We will study the α-optimal policy first and then discuss how to correlate it with the average-optimal policy. Before investigating the general i.i.d. arrival model, we study a special case, namely Bernoulli process. Our objective is to determine the α-optimal policy for the Bernoulli arrival process.

Lemma 7. *For the i.i.d. Bernoulli arrival process and the system starting with the empty queues, the α-optimal policy is of threshold type. In particular, there exist optimal thresholds $L^*_{\alpha,1}$ and $L^*_{\alpha,2}$ so that the optimal deterministic action in state $(i,0)$ is to wait if $i \leq L^*_{\alpha,1}$, and to transmit without coding if $i > L^*_{\alpha,1}$; while in state $(0,j)$ is to wait if $j \leq L^*_{\alpha,2}$, and to transmit without coding if $j > L^*_{\alpha,2}$.*

Proof: We define

$$\mathcal{V}_\alpha(i,0,a) = C_H([i-a]^+) + C_T a + \alpha\sum_{k,l} P_a\big((i,0),(k,l)\big)V_\alpha(k,l).$$

According to Eq. (5),

$$V_\alpha(i,0) = \min_{a\in\{0,1\}} \mathcal{V}_\alpha(i,0,a).$$

Let $L^*_{\alpha,1} = \min\{i \in \mathbb{N}\cup\{0\} : \mathcal{V}_\alpha(i,0,1) > \mathcal{V}_\alpha(i,0,0)\} - 1$. Then the α-optimal policy is $a_t = 0$ for the states $(i,0)$ with $i \leq L^*_{\alpha,1}$, and $a_t = 1$ for the state $(L^*_{\alpha,1}+1, 0)$. However, the system starts with empty queues; as such, the states $(i,0)$ for $i > L^*_{\alpha,1}+1$ are not accessible as $(L^*_{\alpha,1}+1,0)$ only transits to $(L^*_{\alpha,1},0)$, $(L^*_{\alpha,1}+1,0)$, $(L^*_{\alpha,1},1)$, and $(L^*_{\alpha,1}+1,1)$. Hence, we do not need to define the policy of the states $(i,0)$ for $i > L^*_{\alpha,1}+1$. The similar argument is applicable for the states $(0,j)$. Consequently, there exists a policy of threshold type that is α-optimal. ■

Here we are providing an intuition of the threshold policy. If a packet is transmitted immediately without coding, the system cost increases significantly due to a large transmission cost. To wait at present for a coding opportunity in the future incurs a smaller waiting cost. Therefore, the packet might be delayed until the delay cost cannot be compensated by the saving from coding. An optimal policy might be as follows: to increase as time goes the probability to transmit the packet. Moreover, we have shown in Section III that there is an optimal policy that is stationary and deterministic; as such the optimal policy could be threshold type.

A. General i.i.d. arrival process

For the i.i.d. Bernoulli arrival process, we have just shown that the α-optimal policy is threshold based. Our next objective is to extend this result to any i.i.d. arrival process. We define that $\mathcal{V}_\alpha(i,j,a) = C_H([i-a]^+ + [j-a]^+) + C_T \cdot a + \alpha\mathbb{E}[V_\alpha([i-a]^+ + \mathcal{A}_1, [j-a]^+ + \mathcal{A}_2)]$. Moreover, let $\mathcal{V}_{\alpha,n}(i,j,a) = C_H([i-a]^+ + [j-a]^+) + C_T a + \alpha\mathbb{E}[V_{\alpha,n}([i-a]^+ + \mathcal{A}_1, [j-a]^+ + \mathcal{A}_2)]$. Then Eq. (5) can be written as $V_\alpha(i,j) = \min_{a\in\{0,1\}} \mathcal{V}_\alpha(i,j,a)$, while Eq.

(6) can be written as $V_{\alpha,n+1}(i,j) = \min_{a\in\{0,1\}} \mathcal{V}_{\alpha,n}(i,j,a)$. For every discount factor α, we want to show that there exists an α-optimal policy that is of threshold type. To be precise, let the α-optimal policy for the first dimension be $a^*_{\alpha,i} = \min\{a' \in \arg\min_{a\in\{0,1\}} \mathcal{V}_\alpha(i,0,a)\}$,[1] and we will show that $a^*_{\alpha,i}$ is non-decreasing as i increases, and so is the second dimension. We start with a number of definitions.

Definition 8 ([23], Submodularity)**.** A function $f : (\mathbb{N}\cup\{0\})^2 \to \mathbb{R}$ is submodular if for all $i,j \in \mathbb{N}\cup\{0\}$

$$f(i,j) + f(i+1,j+1) \le f(i+1,j) + f(i,j+1).$$

Definition 9 ($\mathcal{K}$-Convexity)**.** A function $f : (\mathbb{N}\cup\{0\})^2 \to \mathbb{R}$ is $\mathcal{K}$-convex (where $\mathcal{K}\in\mathbb{N}$) if for every $i,j\in\mathbb{N}\cup\{0\}$

$$f(i+\mathcal{K},j) - f(i,j) \le f(i+\mathcal{K}+1,j) - f(i+1,j);$$
$$f(i,j+\mathcal{K}) - f(i,j) \le f(i,j+\mathcal{K}+1) - f(i,j+1).$$

Definition 10 ($\mathcal{K}$-Subconvexity)**.** A function $f : (\mathbb{N}\cup\{0\})^2 \to \mathbb{R}$ is $\mathcal{K}$-subconvex (where $\mathcal{K}\in\mathbb{N}$) if for all $i,j\in\mathbb{N}\cup\{0\}$

$$f(i+\mathcal{K},j+\mathcal{K}) - f(i,j) \le f(i+\mathcal{K}+1,j+\mathcal{K}) - f(i+1,j);$$
$$f(i+\mathcal{K},j+\mathcal{K}) - f(i,j) \le f(i+\mathcal{K},j+\mathcal{K}+1) - f(i,j+1).$$

Remark 11. If a function $f : (\mathbb{N}\cup\{0\})^2 \to \mathbb{R}$ is submodular and $\mathcal{K}$-subconvex, then it is $\mathcal{K}$-convex, and for every $r\in\mathbb{N}$ with $1\le r<\mathcal{K}$,

$$f(i+\mathcal{K},j+r) - f(i,j) \le f(i+\mathcal{K}+1,j+r) - f(i+1,j);$$
$$f(i+r,j+\mathcal{K}) - f(i,j) \le f(i+r,j+\mathcal{K}+1) - f(i,j+1).$$

For simplicity, we will ignore $\mathcal{K}$ in definitions 9 and 10 when $\mathcal{K}=1$. We will show in Subsection IV-C that $V_\alpha(i,j)$ is non-decreasing, submodular, and subconvex, that result in the threshold base of α-optimal policy. Note that the definition of $\mathcal{K}$-Convexity (Definition 9) is dimension-wise, which is different from the definition of convexity for the continuous function in two dimensions.

B. Proof overview

Before the technical proofs in Subsection IV-C, in this subsection, we overview why submodularity and subconvexity of $V_\alpha(i,j)$ lead to the α-optimality of the threshold based policy.

- *We claim that to show that α-optimal policy is monotonic w.r.t. state $(i,0)$, it suffices to show that $\mathcal{V}_\alpha(i,0,1) - \mathcal{V}_\alpha(i,0,0)$ is a non-increasing function w.r.t. i:* Suppose that $\mathcal{V}_\alpha(i,0,1)-\mathcal{V}_\alpha(i,0,0)$ is non-increasing, i.e., $\mathcal{V}_\alpha(i+1,0,1) - \mathcal{V}_\alpha(i+1,0,0) \le \mathcal{V}_\alpha(i,0,1) - \mathcal{V}_\alpha(i,0,0)$. If the α-optimal policy for state $(i,0)$ is $a^*_{\alpha,i}=1$, i.e., $\mathcal{V}_\alpha(i,0,1) - \mathcal{V}_\alpha(i,0,0) \le 0$, then the α-optimal policy for state $(i+1,0)$ is also $a^*_{\alpha,i+1}=1$ according to $\mathcal{V}_\alpha(i+1,0,1) - \mathcal{V}_\alpha(i+1,0,0) \le \mathcal{V}_\alpha(i,0,1) - \mathcal{V}_\alpha(i,0,0) \le 0$. Similarly, if the α-optimal policy for state $(i+1,0)$ is $a^*_{\alpha,i+1}=0$ then the α-optimal policy for state $(i,0)$ is $a^*_{\alpha,i}=0$. Hence, the α-optimal policy is monotonic in i.
- *We claim that to prove that $\mathcal{V}_\alpha(i,0,1)-\mathcal{V}_\alpha(i,0,0)$ is non-increasing, it is sufficient to show that $V_\alpha(i,j)$ is convex:* When $i\ge1$, the claim is true since
$$\begin{aligned}&\mathcal{V}_\alpha(i,0,1) - \mathcal{V}_\alpha(i,0,0)\\ =&C_T - C_H + \alpha\mathbb{E}[V_\alpha(i-1+\mathcal{A}_1,\mathcal{A}_2) - V_\alpha(i+\mathcal{A}_1,\mathcal{A}_2)].\end{aligned}$$
- *Similarly, to show that α-optimal policy of state (i,j) is monotonic w.r.t. i for fixed j and vice versa, it suffices to show that $V_\alpha(i,j)$ is subconvex:* When $i,j\ge1$, we observe that
$$\begin{aligned}&\mathcal{V}_\alpha(i,j,1) - \mathcal{V}_\alpha(i,j,0)\\ =&C_t - 2C_h + \alpha\mathbb{E}[V_\alpha(i-1+\mathcal{A}_1,j-1+\mathcal{A}_2)-\\ &\qquad V_\alpha(i+\mathcal{A}_1,j+\mathcal{A}_2)].\end{aligned}$$
- *We claim that $V_\alpha(i,j)$ is submodular:* We intend to prove the convexity and subconvexity of $V_\alpha(i,j)$ by induction, which will require the relation between $V_\alpha(i,j)+V_\alpha(i+1,j+1)$ and $V_\alpha(i+1,j)+V_\alpha(i,j+1)$. There will be two choices: (i) $V_\alpha(i,j)+V_\alpha(i+1,j+1) \le V_\alpha(i+1,j)+V_\alpha(i,j+1)$, or (ii) $V_\alpha(i,j)+V_\alpha(i+1,j+1) \ge V_\alpha(i+1,j)+V_\alpha(i,j+1)$. First, We might assume that $V_\alpha(i,j)$ satisfies (i). Then (i) and the subconvexity of $V_\alpha(i,j)$ implies the convexity of $V_\alpha(i,j)$. In the contrary, the convexity of $V_\alpha(i,j)$ and (ii) lead to the subconvexity of $V_\alpha(i,j)$. In other words, both choices are possible since they do not violate the convexity and subconvexity of $V_\alpha(i,j)$. However, we are going to argue that the choice (ii) is wrong as follows. Suppose that the actions of α-optimal policy for the states (i,j), $(i+1,j)$, $(i,j+1)$, $(i+1,j+1)$ are $0,0,1,1$ respectively. If the choice (ii) is true, then when $i\ge1$, we have
$$\begin{aligned}&C_H(i+j) + \mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1,j+\mathcal{A}_2)]+\\ &C_T + C_H(i+j) + \mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1,j+\mathcal{A}_2)]\\ \ge&C_H(i+1+j) + \mathbb{E}[V_{\alpha,n}(i+1+\mathcal{A}_1,j+\mathcal{A}_2)]+\\ &C_T + C_H(i-1+j) + \mathbb{E}[V_{\alpha,n}(i-1+\mathcal{A}_1,j+\mathcal{A}_2)].\end{aligned}$$
By simplifying the above inequality, we can observe the contradiction to the fact that $V_{\alpha,n}(i,j)$ is convex. Therefore, $V_\alpha(i,j)$ is submodular.

Based on the above discussion, we understand that if we show $V_\alpha(i,j)$ is submodular and subconvex, then the α-optimal policy of state (i,j) is non-decreasing separately in the direction of i and j (i.e., threshold type). Next, we briefly discuss how Lemmas 12-15 and Theorem 16 in the next subsection work together. Theorem 16 states that the α-optimal policy is of threshold type, while the proof is based on an induction on n in Eq. (6). First, when $n=0$ we observe that $V_{\alpha,0}(i,j)$ is non-decreasing, submodular, and subconvex. Second, based on Lemma 12 and Corollary 13, $\min\{a'\in\arg\min_{a\in\{0,1\}}\mathcal{V}_{\alpha,0}(i,j,a)\}$ is non-decreasing w.r.t. i for fixed j, and vice versa. Third, according to Lemmas 5, 14, and 15, we know that $V_{\alpha,1}(i,j)$ is non-decreasing, submodular, and subconvex. Therefore, as n goes to infinity, we conclude that $V_\alpha(i,j)$ is non-decreasing, submodular, and subconvex, as well as $\min\{a'\in\arg\min_{a\in\{0,1\}}\mathcal{V}_\alpha(i,j,a)\}$ is non-decreasing w.r.t. i for fixed j, and vice versa.

C. Main results

Lemma 12. *Given $0<\alpha<1$ and $n\in\mathbb{N}\cup\{0\}$. If $V_{\alpha,n}(i,j)$ is non-decreasing, submodular, and subconvex, then $\mathcal{V}_{\alpha,n}(i,j,a)$*

[1] This notation also used in [15] combines two operations: First we let $\Lambda = \{a\in\{0,1\} : \min\mathcal{V}_{\alpha,n}(i,0,a)\}$, and then do $\min\Lambda$. In other words, we choose $a=0$ when both $a=0$ and $a=1$ result in the same $\mathcal{V}_{\alpha,n}(i,j,a)$.

is submodular for i and a when j is fixed, and so is for j and a when i is fixed.

Proof: See Appendix B. ∎

Submodularity of $\mathcal{V}_{\alpha,n}(i,j,a)$ implies the monotonicity of the optimal minimizing policy [15, Lemma 4.7.1] as described in the following Corollary. This property will simplify the proofs of Lemmas 14 and 15.

Corollary 13. *Given $0<\alpha<1$ and $n\in\mathbb{N}\cup\{0\}$. If $V_{\alpha,n}(i,j)$ is non-decreasing, submodular, and subconvex, then $\min\{a'\in \arg\min_{a\in\{0,1\}}\mathcal{V}_{\alpha,n}(i,j,a)\}$ is non-decreasing w.r.t. i for fixed j, and vice versa.*

Lemma 14. *Given $0<\alpha<1$ and $n\in\mathbb{N}\cup\{0\}$. If $V_{\alpha,n}(i,j)$ is non-decreasing, submodular, and subconvex, then $V_{\alpha,n+1}(i,j)$ is submodular.*

Proof: See Appendix C. ∎

Lemma 15. *Given $0<\alpha<1$ and $n\in\mathbb{N}\cup\{0\}$. If $V_{\alpha,n}(i,j)$ is non-decreasing, submodular, and subconvex, then $V_{\alpha,n+1}(i,j)$ is subconvex.*

Proof: See Appendix D. ∎

Based on the properties of $V_\alpha(i,j)$, we are ready to state the optimality of the threshold type policy in terms of the total expected discounted cost.

Theorem 16. *For the MDP$\{(Q_t,a_t),t\geq 0\}$ with any i.i.d. arrival processes to both queues, there exists an α-optimal policy that is of threshold type. Given $Q_t^{(2)}$, the α-optimal policy is monotone w.r.t. $Q_t^{(1)}$, and vice versa.*

Proof: We prove by induction. $V_{\alpha,0}(i,j)=0$ is non-decreasing, submodular, and subconvex, that leads to the non-decreasing $\min\{a'\in\arg\min_{a\in\{0,1\}}\mathcal{V}_{\alpha,0}(i,j,a)\}$ based on Corollary 13. These properties propagate as n goes to infinity according to lemmas 5, 14, 15, and Corollary 13. ∎

Thus far, the α-optimal policy is characterized. A useful relation between the average-optimal policy and the α-optimal policy is described in the following lemma.

Lemma 17 ([22], Lemma and Theorem (i))**.** *Consider MDP$\{(Q_t,a_t),t\geq 0\}$. Let $\{\alpha_n\}$ converging to 1 be any sequence of discount factors associated with the α-optimal policy $\{\theta_{\alpha_n}(i,j)\}$. There exists a subsequence $\{\beta_n\}$ and a stationary policy $\theta^*(i,j)$ that is the limit point of $\{\theta_{\beta_n}(i,j)\}$. If the three conditions in Lemma 27 are satisfied, $\theta^*(i,j)$ is the average-optimal policy for Eq. (2).*

Theorem 18. *Consider any i.i.d. arrival processes to both queues. For the MDP$\{(Q_t,a_t),t\geq 0\}$, the average-optimal policy is of threshold type. There exist the optimal thresholds L_1^* and L_2^* so that the optimal deterministic action in states $(i,0)$ is to wait if $i\leq L_1^*$, and to transmit without coding if $i>L_1^*$; while in state $(0,j)$ is to wait if $j\leq L_2^*$, and to transmit without coding if $j>L_2^*$.*

Proof: Let $(\tilde{i},0)$ be any state which average-optimal policy is to transmit, i.e., $\theta^*(\tilde{i},0)=1$ in Lemma 17. Since there is a sequence of discount factors $\{\beta_n\}$ such that $\theta_{\beta_n}(i,j)\to\theta^*(i,j)$, then there exists $N>0$ so that $\theta_{\beta_n}(\tilde{i},0)=1$ for all $n\geq N$. Due to the monotonicity of α-optimal policy in Theorem 16, $\theta_{\beta_n}(i,0)=1$ for all $i\geq\tilde{i}$ and $n\geq N$. Therefore, $\theta^*(i,0)=1$ for all $i\geq\tilde{i}$. To conclude, the average-optimal policy is of threshold type. ∎

V. Obtaining the Optimal Deterministic Stationary Policy

We have shown in the previous sections that the average-optimal policy is stationary, deterministic and of threshold type, so we only need to consider the subset of deterministic stationary policies. Given the thresholds of the both queues, the MDP is reduced to a Markov chain. The next step is to find the optimal threshold. First note that the condition $\mathbb{E}[\mathcal{A}_i]<1$ might not be sufficient for the stability of the queues since the threshold based policy leads to an average service rate lower than 1 packet per time slot. In the following theorem, we claim that the conditions $\mathbb{E}[\mathcal{A}_i^2]<\infty$ and $\mathbb{E}[\mathcal{A}_i]<1$ for $i=1,2$ are enough for the stability of the queues.

Theorem 19. *For the MDP$\{(Q_t,a_t),t\geq 0\}$ with $\mathbb{E}[\mathcal{A}_i^2]<\infty$ and $\mathbb{E}[\mathcal{A}_i]<1$ for $i=1,2$. The Markov chain obtained by applying the stationary and deterministic threshold policy to the MDP is positive recurrent, i.e., the stationary distribution exists.*

Proof: (Sketch) The proof is based on Foster-Lyapunov theorem [25] associated with the Lyapunov function $\mathcal{L}(x,y)=x^2+y^2$. ∎

We realize that if $\mathbb{E}[\mathcal{A}_i^2]<\infty$ and $\mathbb{E}[\mathcal{A}_i]<1$ for $i=1,2$, then there exists a stationary threshold type policy that is average-optimal and can be obtained from the reduced Markov chain. The following theorem gives an example of how to compute the optimal thresholds.

Theorem 20. *Consider the Bernoulli arrival process. The optimal thresholds L_1^* and L_2^* are*

$$(L_1^*,L_2^*)=\arg\min_{L_1,L_2} C_T\mathcal{T}(L_1,L_2)+C_H\mathcal{H}(L_1,L_2),$$

where

$$\begin{aligned}\mathcal{T}(L_1,L_2) &= p_1^{(1)}p_1^{(2)}\pi_{0,0}+p_1^{(2)}\sum_{i=1}^{L_1}\pi_{i,0}+p_1^{(1)}\sum_{j=1}^{L_2}\pi_{0,j}+\\ &\quad p_1^{(1)}p_0^{(2)}\pi_{L_1,0}+p_0^{(1)}p_1^{(2)}\pi_{0,L_2};\\ \mathcal{H}(L_1,L_2) &= \sum_{i=1}^{L_1}i\pi_{i,0}+\sum_{j=1}^{L_2}j\pi_{0,j},\end{aligned}$$

for which

$$\begin{aligned}\pi_{0,0} &= \frac{1}{\left(\frac{1-\zeta^{L_1+1}}{1-\zeta}\right)+\left(\frac{1-1/\zeta^{L_2+1}}{1-1/\zeta}\right)-1};\\ \pi_{i,0} &= \zeta^i\pi_{0,0};\\ \pi_{0,j} &= \pi_{0,0}/\zeta^j;\\ \zeta &= \frac{p_1^{(1)}p_0^{(2)}}{p_0^{(1)}p_1^{(2)}}.\end{aligned}$$

Proof: Let $Y_t^{(i)}$ be the number of type i packets at the t^{th} slot *after* transmission. It is crucial to note that this observation time is different from when the MDP is observed. Then the bivariate stochastic process $\{(Y_t^{(1)},Y_t^{(2)}),t\geq 0\}$ is a discrete-time Markov chain which state space is smaller than the original MDP, i.e., $(0,0)$, $(1,0)$, $(2,0)$, $\cdots$, $(L_1,0)$, $(0,1)$, $(0,2)$, $\cdots$, $(0,L_2)$. Define ζ as a parameter such that

$$\zeta = \frac{p_1^{(1)} p_0^{(2)}}{p_0^{(1)} p_1^{(2)}}.$$

Then, the balance equations for $0 < i \leq L_1$ and $0 < j \leq L_2$ are:

$$\begin{aligned} \pi_{i,0} &= \zeta \pi_{i-1,0}; \\ \zeta \pi_{0,j} &= \pi_{0,j-1}. \end{aligned}$$

Since $\pi_{0,0} + \sum_{i,j} \pi_{i,0} + \pi_{0,j} = 1$, we have

$$\pi_{0,0} = \frac{1}{\left(\frac{1-\zeta^{L_1+1}}{1-\zeta}\right) + \left(\frac{1-1/\zeta^{L_2+1}}{1-1/\zeta}\right) - 1}.$$

The expected number of transmissions per slot is

$$\begin{aligned} \mathcal{T}(L_1, L_2) &= p_1^{(1)} p_1^{(2)} \pi_{0,0} + p_1^{(2)} \sum_{i=1}^{L_1} \pi_{i,0} + p_1^{(1)} \sum_{j=1}^{L_2} \pi_{0,j} + \\ & \quad p_1^{(1)} p_0^{(2)} \pi_{L_1,0} + p_0^{(1)} p_1^{(2)} \pi_{0,L_2}. \end{aligned}$$

The average number of packets in the system at the beginning of each slot is

$$\mathcal{H}(L_1, L_2) = \sum_{i=1}^{L_1} i \pi_{i,0} + \sum_{j=1}^{L_2} j \pi_{0,j}.$$

Thus upon minimizing we get the optimal thresholds L_1^* and L_2^*. ■

Whenever $C_H > 0$, it is relatively straightforward to obtain L_1^* and L_2^*. Since it costs C_T to transmit a packet and C_H for a packet to wait for a slot, it would be better to transmit a packet than make a packet wait for more than C_T/C_H slots. Thus L_1^* and L_2^* would always be less than C_T/C_H. Hence, by completely enumerating between 0 and C_T/C_H for both L_1 and L_2, we can obtain L_1^* and L_2^*. One could perhaps find faster techniques than complete enumeration, but it certainly serves the purpose.

Subsequently, we study a special case, $p_1^{(1)} = p_1^{(2)} \triangleq p$, in Theorem 20. Then $L_1 = L_2 \triangleq L$ as both arrival processes are identical. It can be calculated that $\zeta = 1$ and $\pi_{i,j} = 1/(2L+1)$ for all i, j, and

$$\mathcal{T}(L) = \frac{2pL + 2p - p^2}{2L+1};$$

$$\mathcal{H}(L) = \frac{L^2 + L}{2L+1}.$$

Define $\nu = C_T/C_H$. The optimal threshold is

$$L^*(p,\nu) = \arg\min_L \frac{\nu(2pL + 2p - p^2) + L + L^2}{2L+1}.$$

By taking the derivative, we obtain that $L^* = 0$ if $\nu < 1/(2p - 2p^2)$ and otherwise,

$$L^*(p,\nu) = \frac{-1 + \sqrt{1 - 2(1 - 2\nu p + 2\nu p^2)}}{2}.$$

We can observe that $L^*(p,\nu)$ is a concave function w.r.t. p. Given ν fixed, $L^*(1/2,\nu) = (\sqrt{\nu - 1} - 1)/2$ is the largest optimal threshold among various values of p. When $p < 1/2$, the optimal-threshold decreases as there is a relatively lower probability for packets in one queue to wait for a coding pair in another queue. When $p > 1/2$, there will be a coding pair already in the relay node with a higher probability, and therefore the optimal-threshold also decreases. Moreover, $L^*(1/2,\nu) = \mathcal{O}(\sqrt{\nu})$, so the maximum optimal threshold grows with the square root of ν, but not linearly. When p is very small, $L^*(p,\nu) = \mathcal{O}(\sqrt{\nu p})$ grows slower than $L^*(1/2,\nu)$. Figure 3 depicts the optimal threshold $L^*(p,\nu)$ for various values of arrival rate p, and $\nu = C_T/C_H$.

VI. Numerical Studies

In this section we present several numerical results to compare the performance of different policies in the single relay setting as well as in a line network. We analyzed the following policies:

1) Opportunistic Coding (OC): this policy does not waste any opportunities for transmitting the packets. That is when a packet arrives, coding is performed if a coding opportunity exists, otherwise transmission takes place immediately.
2) Queue-length based threshold (QLT): this stationary deterministic policy applies the thresholds, proposed by Theorem 20, on the queue lengths.
3) Queue-length-plus-Waiting-time-based (QL+WT) thresholds: this is a history dependent policy which takes into account the waiting time of the packets in the queues as well as the queue lengths. That is a packet will be transmitted (without coding), if the queue length hits the threshold or the head-of-queue packet has been waiting for at least some pre-determined amount of time. The optimal waiting-time thresholds are found using exhaustive search through stochastic simulations for the given arrival distributions.
4) Waiting-time (WT) based threshold: this is another history dependent policy that *only* considers the waiting times of the packets, in order to schedule the transmissions. The optimum waiting times of the packets are found through exhaustive search.

We simulate these policies on two different cases: (i) the single relay network with Bernoulli arrivals (Fig. 4, 5, and 6) and (ii) a line network with 4 nodes, in which the sources are Bernoulli (Fig. 7, and 8). Note that in case (ii), since the departures from one queue determine the arrivals into the other queue, the arrival processes are significantly different from Bernoulli. As expected, for the single relay network, the QLT policy has the optimal performance and the QL+WT policy does not have any advantage.

Moreover, there are results (see [26]) that indicate that the independent arrivals model is accurate under heavy traffic for multi-hop networks. Hence, our characterization of the optimal policy does have value in a more general case. Our simulation results indicate that QLT policy also exhibits a near optimal performance for the line network. We also observe, from the simulation results for the waiting-time-based policy, that making decisions based on waiting time alone leads to a suboptimal performance. In all experiments, the opportunistic policy has the worst possible performance.

The results are intriguing as they suggest that achieving a near-perfect trade-off between waiting and transmission costs is possible using simple policies; moreover, coupled with optimal network-coding aware routing policies like the one in

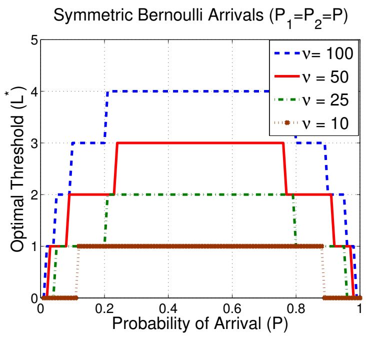

Fig. 3. Optimal queue length threshold for a single relay with symmetric Bernoulli arrivals.

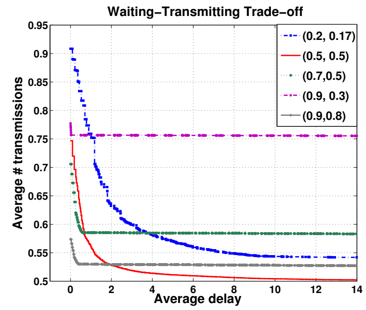

Fig. 4. Trade-off between average delay and number of transmissions in a single relay using queue-length based threshold (QLT) policy for different Bernoulli arrival rates (p_1, p_2).

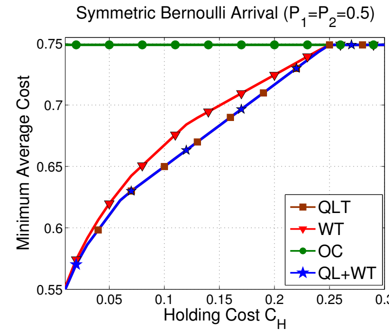

Fig. 5. Comparison of the minimum average cost (per slot) in a single relay with Bernoulli arrival rates $(0.5, 0.5)$, for different policies, where the costs are normalized by the transmission cost.

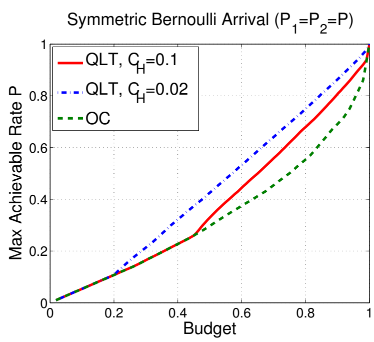

Fig. 6. Achievable arrival rate versus average budget (per slot) in a single relay with Bernoulli arrivals, using opportunistic coding (OC) and QLT policies, where the costs are normalized by the transmission cost.

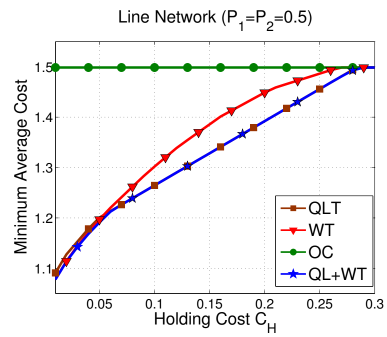

Fig. 7. Comparison of different policies in a line network with two intermediate nodes and two Bernoulli flows with mean arrival rates $(0.5, 0.5)$.

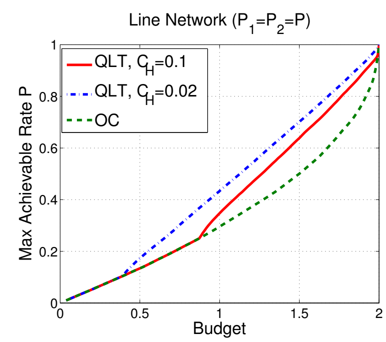

Fig. 8. Achievable arrival rate versus average budget (per slot) in a line network with two intermediate nodes and two Bernoulli flows.

our earlier work [8], have the potential to exploit the positive externalities that network coding offers.

VII. Extensions

We have known that the average-optimal policy is stationary and threshold based for the i.i.d. arrival process and the perfect channels with at most one packet served per time slot. Three more general models are discussed here. We focus on the characterization of the optimality equation, which results in the structure of the average-optimal policy.

A. Batched service

Assume that the relay R can serve a group of packets with the size of $\mathcal{M}$ at end of the time slot. At the end of every time slot, relay R decides to transmit, $a_t = 1$, or to wait $a_t = 0$. The holding cost per unit time for a packet is C_H, while C_T is the cost to transmit a batched packet. Then the immediate cost is

$$\begin{aligned} & C^{(\mathcal{M})}(Q_t, a_t) \\ = \; & C_H([Q_t^{(1)} - a_t\mathcal{M}]^+ + [Q_t^{(2)} - a_t\mathcal{M}]^+) + C_T a_t. \end{aligned}$$

We also want to find the optimal policy θ^* that minimizes the long-run average cost $V^{(\mathcal{M})}(\theta)$, called $\mathcal{M}$-MDP$\{(Q_t, a_t), t \geq 0\}$ problem,

$$V^{(\mathcal{M})}(\theta) = \lim_{K\to\infty} \frac{1}{K+1}\mathbb{E}_\theta\left[\sum_{t=0}^{K} C^{(\mathcal{M})}(Q_t, a_t)|Q_0 = (0,0)\right].$$

Notice that the best policy might not just transmit when both queues are non-empty. When $\mathcal{M} > 1$, R might also want to wait even if $Q_t^{(1)}Q_t^{(2)} > 0$ because the batched service of size less than $\mathcal{M}$ has the same transmission cost C_T. The optimality equation of the expected α-discounted cost is revised as

$$\begin{aligned} V_\alpha^{(\mathcal{M})}(i,j) = \min_{a\in\{0,1\}} \Big[& C_H([i - a\mathcal{M}]^+ + [j - a\mathcal{M}]^+) + C_T a + \\ & \mathbb{E}[V_\alpha^{(\mathcal{M})}([i - a\mathcal{M}]^+ + \mathcal{A}_1, [j - a\mathcal{M}]^+ + \mathcal{A}_2)]\Big]. \end{aligned}$$

We can get the following results.

Theorem 21. *Given α and $\mathcal{M}$, $V_\alpha^{(\mathcal{M})}(i,j)$ is non-decreasing, submodular, and $\mathcal{M}$-subconvex. Moreover, there is an α-optimal policy that is of threshold type. Fixed j, the α-optimal policy is monotone w.r.t. i, and vice versa.*

Theorem 22. *Consider any i.i.d. arrival processes to both queues. For the $\mathcal{M}$-MDP$\{(Q_t, a_t), t \geq 0\}$, the average-optimal policy is of threshold type. Given $j = \tilde{j}$ fixed, there exists the optimal threshold $L_{\tilde{j}}^*$ such that the optimal stationary and deterministic policy in state $(i, \tilde{j})$ is to wait if $i \leq L_{\tilde{j}}^*$, and* *to transmit if $i > L_j^*$. Similar argument holds for the other queue.*

B. Markov-Modulated Arrival Process

While the i.i.d. arrival process is examined so far, a specific arrival process with memory is studied here, i.e., Markov-modulated arrival process (MMAP). The service capacity of R is focused on $\mathcal{M} = 1$ packet. Let $\mathcal{N}^{(i)} = \{0, 1, \cdots, N^{(i)}\}$ be the state space of MMAP at node i, with the transition probability $p_{k,l}^{(i)}$ where $k, l \in \mathcal{N}^{(i)}$. Then the number of packets generated by the node i at time t is $\mathcal{N}_t^{(i)} \in \mathcal{N}^{(i)}$. Then the decision of R is made based on the observation of $(Q_t^{(1)}, Q_t^{(2)}, \mathcal{N}_t^{(1)}, \mathcal{N}_t^{(2)})$. Similarly, the objective is to find the optimal policy that minimizes the long-run average cost, named MMAP-MDP$\{((Q_t^{(1)}, Q_t^{(2)}, \mathcal{N}_t^{(1)}, \mathcal{N}_t^{(2)}), a_t) : t \geq 0\}$ problem. The optimality equation of the expected α-discounted cost becomes

$$\begin{aligned}
&V_\alpha^{\mathrm{MMAP}}(i, j, n_1, n_2) \\
=& \min_{a \in \{0,1\}} [C_H([i-a]^+ + [j-a]^+) + C_T a + \\
&\alpha \sum_{k=0}^{N^{(1)}} \sum_{l=0}^{N^{(2)}} p_{n_1,k}^{(1)} p_{n_2,l}^{(2)} V_\alpha^{\mathrm{MMAP}}([i-a]^+ + k, [j-a]^+ + l, k, l)].
\end{aligned}$$

Then we conclude the following results.

Theorem 23. *Given $n_1 \in \mathcal{N}^{(1)}$ and $n_2 \in \mathcal{N}^{(2)}$, $V_\alpha^{MMAP}(i, j, n_1, n_2)$ is non-decreasing, submodular, and subconvex w.r.t. i and j. Moreover, there is an α-optimal policy that is of threshold type. Fixed n_1 and n_2, the α-optimal policy is monotone w.r.t. i when j is fixed, and vice versa.*

Theorem 24. *Consider any MMAP arrival process. For the MMAP-MDP$\{((Q_t^{(1)}, Q_t^{(2)}, \mathcal{N}_t^{(1)}, \mathcal{N}_t^{(2)}), a_t) : t \geq 0\}$, the average-optimal policy is of multiple thresholds type. There exists a set of optimal thresholds $\{L_{1,n_1,n_2}^*\}$ and $\{L_{2,n_1,n_2}^*\}$, where $n_1 \in \mathcal{N}^{(1)}$ and $n_2 \in \mathcal{N}^{(2)}$, so that the optimal stationary decision in states $(i, 0, n_1, n_2)$ is to wait if $i \leq L_{1,n_1,n_2}^*$, and to transmit without coding if $i > L_{1,n_1,n_2}^*$; while in state $(0, j, n_1, n_2)$ is to wait if $j \leq L_{2,n_1,n_2}^*$, and to transmit without coding if $j > L_{2,n_1,n_2}^*$.*

C. Time-varying channel

In this subsection, we examine the scenario in which the relay transmits packets over time-varying ON/OFF channels, while we assume that the arrivals are i.i.d. and the relay can serve at most one packet for each time slot. Let $S_t = (S_t^{(1)}, S_t^{(2)})$ be the channel state at time t, where $S_t^{(i)} \in \{0(\mathrm{OFF}), 1(\mathrm{ON})\}$ indicates the channel condition from the relay to node i. We assume that the channel states are i.i.d. over time. Moreover, when $S_t^{(i)} = 1$, to transmit a packet from the relay to node i takes the cost of $C_T^{(i)}$. Then the immediate cost $C(Q_t, S_t, a_t)$ is

$$\begin{aligned}
C(Q_t, S_t, a_t) =& C_H([i - a_t S_t^{(1)}]^+ + [j - a_t S_t^{(2)}]^+) + \\
& \max(a_t S_t^{(1)} C_T^{(1)}, a_t S_t^{(2)} C_T^{(2)}).
\end{aligned}$$

The objective is also to find the optimal policy that minimizes the long-run average cost. The optimality equation of the expected α-discounted cost becomes

$$\begin{aligned}
&V_\alpha(i, j, s_1, s_2) \\
=& \min_{a \in \{0,1\}} \big[C_H([i - as_1]^+ + [j - as_2]^+) + \\
& \max(as_1 C_T^{(1)}, as_2 C_T^{(2)}) + \\
& \alpha \mathbb{E}[V_\alpha([i - as_1]^+ + \mathcal{A}_1, [j - as_2]^+ + \mathcal{A}_2), S_t^{(1)}, S_t^{(2)}]\big].
\end{aligned}$$

Then we conclude the following results.

Theorem 25. *$V_\alpha(i, j, 1, 1)$ is non-decreasing, submodular, and subconvex. $V_\alpha(i, j, 1, 0)$ is convex in i for any fixed j and $V_\alpha(i, j, 0, 1)$ is convex in j for any fixed i. Moreover, there is an α-optimal policy that is of threshold type. For each channel state, the α-optimal policy is monotone in i for fixed j, and vice versa.*

Theorem 26. *Consider any i.i.d. arrivals to both queues and time-varying ON/OFF channels. The average-optimal policy is of threshold type. For state (s_1, s_2), there exist the optimal thresholds L_{1,s_1,s_2}^* and L_{2,s_1,s_2}^* so that the optimal deterministic action in states $(i, 0)$ is to wait if $i \leq L_{1,s_1,s_2}^*$, and to transmit without coding if $i > L_{1,s_1,s_2}^*$; while in state $(0, j)$ is to wait if $j \leq L_{2,s_1,s_2}^*$, and to transmit without coding if $j > L_{2,s_1,s_2}^*$.*

VIII. Conclusion

In this paper we investigate the delicate trade-off between waiting and transmitting using network coding. We started with the idea of exploring the whole space of history dependent policies, but showed step-by-step how we could move to simpler regimes, finally culminating in a stationary deterministic queue-length threshold based policy. The policy is attractive because its simplicity enables us to characterize the thresholds completely, and we can easily illustrate its performance on multiple networks. We showed by simulation how the performance of the policy is optimal in the Bernoulli arrival scenario, and how it also does well in other situations such as for line networks. Moreover, our policy can be applied for real-time applications. In our work, we explicitly model the cost of packet delays; as such, we can compute the probability of meeting the deadline, and then tune our holding cost so that the probability is met.

IX. Acknowledgement

The authors are grateful to Daren B. H. Cline, Vivek Borkar, and Rajesh Sundaresan for the useful discussions.

Appendix A
Proof of Theorem 6

The next two lemmas, which can be proven via the similar arguments in [22], specify the conditions for the existence of the optimal stationary and deterministic policy.

Lemma 27 ([22], Theorem (i))**.** *There exists a stationary and deterministic policy that is average-optimal for the MDP$\{(Q_t, a_t), t \geq 0\}$ if the following conditions are satisfied:*

(i) *$V_\alpha(i, j)$ is finite for all i, j, and α;*
(ii) *There exists a nonnegative N such that $v_\alpha(i, j) \geq -N$ for all i, j, and α;*

(iii) *There exists a nonnegative $M_{i,j}$ such that $v_\alpha(i,j) \leq M_{i,j}$ for every i, j, and α. Moreover, for each state (i,j) there is an action $a(i,j)$ such that $\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_{a(i,j)}\big((i,j),(k,l)\big)M_{k,l} < \infty$.*

Lemma 28 ([22], Proposition 5)**.** *Assume there exists a stationary policy θ inducing an irreducible and ergodic Markov chain with the following properties: there exists a nonnegative function $F(i,j)$ and a finite nonempty subset $G \subseteq (\mathbb{N}\cup\{0\})^2$ such that for $(i,j) \in (\mathbb{N}\cup\{0\})^2 - G$ it holds that*

$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_{a(\theta)}((i,j),(k,l))F(k,l) - F(i,j) \leq -C((i,j),a(\theta)), \tag{7}$$

where $a(\theta)$ is the action when the policy θ is applied. Moreover, for $(i,j) \in G$ it holds that

$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_{a(\theta)}((i,j),(k,l))F(k,l) < \infty.$$

Then, the condition (iii) in Lemma 27 holds.

proof to Theorem 6 : As described earlier it is sufficient to show that the three conditions in Lemma 27 are satisfied. Proposition 2 implies that the condition (i) holds, while the condition (ii) is satisfied due to Lemma 5 (i.e., $N = 0$ in Lemma 27). We denote by $\tilde{\theta}$ the stationary policy of transmitting at each time slot. We use this policy for each of the three cases described below and show that condition (iii) of Lemma 27 holds.

Case (i): $p_0^{(i)} + p_1^{(i)} < 1$ for $i = 1,2$, i.e., the probability that two or more packets arrive for each time slot is non-zero. This policy $\tilde{\theta}$ results in an irreducible and ergodic Markov chain, and therefore Lemma 28 can be applied. Let $F(i,j) = B(i^2+j^2)$ for some positive B. Then, for all states $(i,j) \in (\mathbb{N}\cup\{0\})^2 - \{(0,0),(0,1),(1,0)\}$, it holds that

$$\begin{aligned}
&\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_{a(\tilde{\theta})}\left((i,j),(k,l)\right)\left[F(k,l)-F(i,j)\right]\\
=&\sum_{r=0}^{\infty}\sum_{s=0}^{\infty} P_1\left((i,j),([i-1]^+ + r,[j-1]^+ + s)\right)\cdot\\
&\qquad\left[F([i-1]^+ + r,[j-1]^+ + s)-F(i,j)\right]\\
=&\sum_{r=0}^{\infty}\sum_{s=0}^{\infty} p_r^{(1)}p_s^{(2)}B\big[2i(r-1)+(r-1)^2+\\
&\qquad 2j(s-1)+(s-1)^2\big]\\
=&2B\Big(i(\mathbb{E}[\mathcal{A}_1]-1)+j(\mathbb{E}[\mathcal{A}_2]-1)\Big)+\\
&B\Big(\mathbb{E}[(\mathcal{A}_1-1)^2]+\mathbb{E}[(\mathcal{A}_2-1)^2]\Big).
\end{aligned}$$

Note that $\mathbb{E}[\mathcal{A}_i] < 1$, hence $2B(\mathbb{E}[\mathcal{A}_i]-1) < -C_H$ for sufficiently large B. Moreover, since $\mathbb{E}[\mathcal{A}_i^2] < \infty$, it holds that

$$\begin{aligned}
&\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_{a(\tilde{\theta})}\left((i,j),(k,l)\right)\left[F(k,l)-F(i,j)\right]\\
\leq\ & -C((i,j),a(\tilde{\theta})),
\end{aligned}$$

when i,j are large enough, where

$$C((i,j),a(\tilde{\theta})) = C_H([i-1]^+ + [j-1]^+) + C_T.$$

We observe that there exists a finite set G that contains states $\{(0,0),(0,1),(1,0)\}$ such that Eq. (7) is satisfied for $(i,j) \in (\mathbb{N}\cup\{0\})^2 - G$. Then, for $(i,j) \in G$, it holds that

$$\begin{aligned}
&\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} P_{a(\tilde{\theta})}\left((i,j),(k,l)\right)F(k,l)\\
=&B\sum_{r=0}^{\infty}\sum_{s=0}^{\infty} p_r^{(1)}p_s^{(2)}\left[([i-1]^+ + r)^2+([j-1]^+ + s)^2\right]\\
=&B\Big\{(i-1)^2+2[i-1]^+\mathbb{E}[\mathcal{A}_1]+\mathbb{E}[\mathcal{A}_1^2]+\\
&(j-1)^2+2[j-1]^+\mathbb{E}[\mathcal{A}_2]+\mathbb{E}[\mathcal{A}_2^2]\Big\} < \infty.
\end{aligned}$$

Therefore, the condition of Lemma 28 is satisfied, which implies, in turn, that condition (iii) in Lemma 27 is satisfied as well.

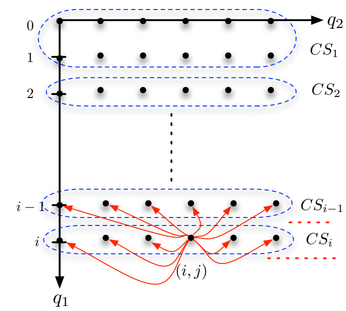

Fig. 9. Case (ii) in the proof of Theorem 6: state (i,j) can only transit to the states in the CS_i and CS_{i-1}.

Case (ii): $p_0^{(1)} + p_1^{(1)} = 1$ and $p_0^{(2)} + p_1^{(2)} < 1$. Note that $\tilde{\theta}$ results in a reducible Markov chain. That is, there are several communicating classes [27]. We define the classes $\text{CS}_1 = \{(a,b) : a = 0,1 \text{ and } b \in \mathbb{N}\cup\{0\}\}$ and $\text{CS}_i = \{(a,b) : a = i, b \in \mathbb{N}\cup\{0\}\}$ for $i \geq 2$, as shown in Fig. 9. Then each CS_i is a communicating class under the policy $\tilde{\theta}$. The states in CS_1 are positive-recurrent, and each CS_i for $i \geq 2$ is a transient class (see [27]).

For $i \geq 2$, let $\overline{C}_{i,j}$ be the expected cost of the first passage from state (i,j) (in class CS_i) to a state in class CS_{i-1}. Moreover, we denote the expected cost of a first passage from state (i,j) to (k,l) by $\overline{C}_{(i,j),(k,l)}$. Let $T_0 = \min\{t \geq 1 : (Q_t^{(1)},Q_t^{(2)}) = (0,0)\}$ and for $i \geq 1$, $T_i = \min\{t \geq 1 : Q_t^{(1)} = i\}$. Then we can express the expected cost of the first passage from state (i,j) to $(0,0)$ as follows.

$$\overline{C}_{(i,j),(0,0)} = \overline{C}_{i,j} + \sum_{k=1}^{i-2}\overline{C}_{i-k,Q_{T_{i-k}}^{(2)}} + \overline{C}_{(1,Q_{T_1}^{(2)}),(0,0)}.$$

Note that state (i,j) has the probability of $p_0^{(1)}$ to escape to class CS_{i-1} and $p_1^{(1)}$ to remain in class CS_i. By considering all the possible paths, we compute $\overline{C}_{i,j}$ as follows.

$$\begin{aligned}
\overline{C}_{i,j} =&\mathbb{E}\left[\sum_{k=0}^{\infty}(p_1^{(1)})^k p_0^{(1)}\sum_{t=0}^{k}C((i,Q_t^{(2)}),1)|(Q_0^{(1)},Q_0^{(2)})=(i,j)\right]\\
=&p_0^{(1)}\mathbb{E}\left[\sum_{t=0}^{\infty}C((i,Q_t^{(2)}),1)\sum_{k=t}^{\infty}(p_1^{(1)})^k|(Q_0^{(1)},Q_0^{(2)})=(i,j)\right]\\
=&\mathbb{E}\left[\sum_{t=0}^{\infty}(p_1^{(1)})^t C((i,Q_t^{(2)}),1)|(Q_0^{(1)},Q_0^{(2)})=(i,j)\right],
\end{aligned}$$

where $C((i,Q_t^{(2)}),1) = C_T + C_H([i-1]^+ + [Q_t^{(2)}-1]^+)$. Following the similar argument to the proof of Proposition 2, we conclude that $\overline{C}_{i,j} < \infty$. Moreover, Proposition 4 in [22] implies that $\overline{C}_{(1,j),(0,0)} < \infty$ for any j, where the intuition is that the expected traveling time from state $(1,j)$ to $(0,0)$ is finite due to the positive recurrence of CS_1. Therefore, we conclude that $\overline{C}_{(i,j),(0,0)} < \infty$.

Let $\hat{\theta}$, be a policy that always transmits until time slot T_0 after which the α-optimal policy is employed. Then, $V_\alpha(i,j)$ can be bounded by

$$\begin{aligned} V_\alpha(i,j) \le \quad & \mathbb{E}_{\hat{\theta}}\left[\sum_{t=0}^{T_0-1} \alpha^t C(Q_t,a_t)|Q_0=(i,j)\right] + \\ & \mathbb{E}_{\hat{\theta}}\left[\sum_{t=T_0}^{\infty} \alpha^t C(Q_t,a_t)|Q_{T_0}=(0,0)\right] \\ \le \quad & \overline{C}_{(i,j),(0,0)} + V_\alpha(0,0). \end{aligned}$$

Then the condition (iii) of Lemma 27 is satisfied by choosing $M_{i,j} = \overline{C}_{(i,j),(0,0)}$. In particular, it holds that $v_\alpha(i,j) = V_\alpha(i,j) - V_\alpha(0,0) \le M_{i,j}$ and $M_{i,j} < \infty$. Moreover, $\sum_{k=0}^\infty \sum_{l=0}^\infty P_1\big((i,j),(k,l)\big) M_{k,l} = \sum_{k=0}^\infty \sum_{l=0}^\infty P_1\big((i,j),(k,l)\big)\overline{C}_{(k,l),(0,0)} \le \overline{C}_{(i,j),(0,0)} < \infty$.

Case (iii): $p_0^{(i)} + p_1^{(i)} = 1$ for $i = 1,2$, i.e., Bernoulli arrivals to both queues. Note that in this case $\tilde{\theta}$ also results in a reducible Markov chain. The proof is similar to case (ii); we can define $M_{i,j} = \overline{C}_{(i,j),(0,0)}$, and show that $\overline{C}_{(i,j),(0,0)}$ is finite for this case. ∎

APPENDIX B
PROOF TO LEMMA 12

Proof: We define $\Delta\mathcal{V}_{\alpha,n}(i,j) = \mathcal{V}_{\alpha,n}(i,j,1) - \mathcal{V}_{\alpha,n}(i,j,0)$. We claim that $\Delta\mathcal{V}_{\alpha,n}(i,j)$ is non-increasing, i.e., $\Delta\mathcal{V}_{\alpha,n}(i,j)$ is a non-increasing function w.r.t. i while j is fixed, and vice versa (we will focus on the former part). Notice that

$$\begin{aligned} \Delta\mathcal{V}_{\alpha,n}(i,j) \quad = \quad & C_H([i-1]^+ + (j-1)^+) + C_T + \\ & \alpha\mathbb{E}[V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)] - \\ & C_H(i+j) - \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)]. \end{aligned}$$

To be precise, when $i \ge 1$,

$$\begin{aligned} \Delta\mathcal{V}_{\alpha,n}(i,j) =& C_T - 2C_H + \alpha\mathbb{E}[V_{\alpha,n}(i-1+\mathcal{A}_1, j-1+\mathcal{A}_2) - \\ & V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)] \text{ for } j \ge 1; \qquad (8) \\ \Delta\mathcal{V}_{\alpha,n}(i,j) =& C_T - C_H + \alpha\mathbb{E}[V_{\alpha,n}(i-1+\mathcal{A}_1, \mathcal{A}_2) - \\ & V_{\alpha,n}(i+\mathcal{A}_1, \mathcal{A}_2)] \text{ for } j = 0. \qquad (9) \end{aligned}$$

Because of the subconvexity of $V_{\alpha,n}(i,j)$ in Eq. (8), when $i \ge 1$ and $j \ge 1$, $\Delta\mathcal{V}_{\alpha,n}(i,j)$ does not increase as i increases. The same is for $i \ge 1$ and $j = 0$ in Eq. (9) due to the convexity of $V_{\alpha,n}(i,j)$.

We proceed to establish the boundary conditions. When $j \ge 1$,

$$\begin{aligned} \Delta\mathcal{V}_{\alpha,n}(1,j) \quad = \quad & C_T - 2C_H + \alpha\mathbb{E}[V_{\alpha,n}(\mathcal{A}_1, j-1+\mathcal{A}_2) - \\ & V_{\alpha,n}(1+\mathcal{A}_1, j+\mathcal{A}_2)]; \\ \Delta\mathcal{V}_{\alpha,n}(0,j) \quad = \quad & C_T - C_H + \alpha\mathbb{E}[V_{\alpha,n}(\mathcal{A}_1, j-1+\mathcal{A}_2) - \\ & V_{\alpha,n}(\mathcal{A}_1, j+\mathcal{A}_2)]. \end{aligned}$$

Note that $\mathbb{E}[V_{\alpha,n}(1+\mathcal{A}_1, j+\mathcal{A}_2)] \ge \mathbb{E}[V_{\alpha,n}(\mathcal{A}_1, j+\mathcal{A}_2)]$ according to non-decreasing $V_{\alpha,n}(i,j)$ and then $\Delta\mathcal{V}_{\alpha,n}(1,j) \le \Delta\mathcal{V}_{\alpha,n}(0,j)$ when $j \ge 1$. Finally, when $j = 0$ we have

$$\begin{aligned} \Delta\mathcal{V}_{\alpha,n}(1,0) \quad = \quad & C_T - C_H + \alpha\mathbb{E}[V_{\alpha,n}(\mathcal{A}_1, \mathcal{A}_2) - \\ & V_{\alpha,n}(1+\mathcal{A}_1, \mathcal{A}_2)]; \\ \Delta\mathcal{V}_{\alpha,n}(0,0) \quad = \quad & C_T. \end{aligned}$$

Here, $\Delta\mathcal{V}_{\alpha,n}(1,0) \le \Delta\mathcal{V}_{\alpha,n}(0,0)$ since $\mathbb{E}[V_{\alpha,n}(\mathcal{A}_1,\mathcal{A}_2) - V_{\alpha,n}(1+\mathcal{A}_1,\mathcal{A}_2)] \le 0$ as $V_{\alpha,n}(i,j)$ is non-decreasing. Consequently, $\Delta\mathcal{V}_{\alpha,n}(i,j)$ is a non-increasing function w.r.t. i while j is fixed. ∎

APPENDIX C
PROOF TO LEMMA 14

Proof: We intend to show that $V_{\alpha,n+1}(i+1,j+1) - V_{\alpha,n+1}(i+1,j) \le V_{\alpha,n+1}(i,j+1) - V_{\alpha,n+1}(i,j)$ for all $i,j \in \mathbb{N}\cup\{0\}$. According to Corollary 13, only 6 cases of $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i,j+1}, a^*_{i+1,j+1})$ are considered, where $a^*_{i,j} = \min\{a' \in \arg\min_{a\in\{0,1\}} \mathcal{V}_{\alpha,n}(i,j,a)\}$.

Case (i): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i,j+1}, a^*_{i+1,j+1}) = (1,1,1,1)$, we claim that

$$\begin{aligned} & \mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - V_{\alpha,n}(i+\mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)] \\ \le \quad & \mathbb{E}[V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, j+\mathcal{A}_2) - \\ & V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)]. \end{aligned}$$

When $i,j \ne 0$, it is true according to submodularity of $V_{\alpha,n}(i,j)$. Otherwise, both sides of the inequality are 0.

Case (ii): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i,j+1}, a^*_{i+1,j+1}) = (0,0,0,0)$, we claim that

$$\begin{aligned} & \mathbb{E}[V_{\alpha,n}(i+1+\mathcal{A}_1, j+1+\mathcal{A}_2) - V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2)] \\ \le & \mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+1+\mathcal{A}_2) - V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)]. \end{aligned}$$

This is obvious from the submodularity of $V_{\alpha,n}(i,j)$.

Case (iii): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i,j+1}, a^*_{i+1,j+1}) = (0,0,0,1)$, we claim that

$$\begin{aligned} & C_T - C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ & \qquad V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2)] \\ \le & C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+1+\mathcal{A}_2) - \\ & \qquad V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)]. \end{aligned}$$

From the submodularity of $V_{\alpha,n}(i,j)$, it is obtained that

$$\begin{aligned} & V_{\alpha,n}(i,j) - V_{\alpha,n}(i+1,j) + V_{\alpha,n}(i,j) - V_{\alpha,n}(i,j+1) \\ \le & V_{\alpha,n}(i,j) - V_{\alpha,n}(i+1,j) + V_{\alpha,n}(i+1,j) - \\ & V_{\alpha,n}(i+1,j+1) \\ = & V_{\alpha,n}(i,j) - V_{\alpha,n}(i+1,j+1). \end{aligned}$$

Since $a^*_{i+1,j+1} = 1$, we have $\Delta\mathcal{V}_{\alpha,n}(i+1,j+1) \le 0$, i.e.,

$$\begin{aligned} & C_T - 2C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ & \qquad V_{\alpha,n}(i+1+\mathcal{A}_1, j+1+\mathcal{A}_2)] \le 0. \end{aligned}$$

The claim follows from the following equation:

$$\begin{aligned} & C_T - 2C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ & \qquad V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2) + \\ & \qquad V_n(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ & \qquad V_{\alpha,n}(i+\mathcal{A}_1, j+1+\mathcal{A}_2)] \\ \le & \Delta\mathcal{V}_{\alpha,n}(i+1,j+1) \le 0. \end{aligned}$$

Case (iv): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i,j+1}, a^*_{i+1,j+1}) = (0,0,1,1)$, we claim that

$$\begin{aligned} &-C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2)] \\ \leq\ & C_H([i-1]^+ - i) + \alpha\mathbb{E}[V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)] \end{aligned}$$

When $i \neq 0$, it is satisfied because $V_{\alpha,n}(i,j)$ is convex. Otherwise, it is true since $V_{\alpha,n}(i,j)$ is non-decreasing.

Case (v): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i,j+1}, a^*_{i+1,j+1}) = (0,1,0,1)$, we claim that

$$\begin{aligned} &C_H(j-[j-1]^+) + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+\mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)] \\ \leq & C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+1+\mathcal{A}_2) - V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)]. \end{aligned}$$

When $j \neq 0$, it holds since $V_{\alpha,n}(i,j)$ is convex. It is true for other cases because of the non-decreasing $V_{\alpha,n}(i,j)$.

Case (vi): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i,j+1}, a^*_{i+1,j+1}) = (0,1,1,1)$, we claim that

$$\begin{aligned} &C_H(j-[j-1]^+) + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+\mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)] \\ \leq & C_T + C_H([i-1]^+ - i) + \alpha\mathbb{E}[V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)]. \end{aligned}$$

Based on the submodularity of $V_{\alpha,n}(i,j)$, we have

$$\begin{aligned} &V_{\alpha,n}([i-1]^+, j) - V_{\alpha,n}(i,j) + V_{\alpha,n}(i,[j-1]^+) - V_{\alpha,n}(i,j) \\ \geq & V_{\alpha,n}([i-1]^+, [j-1]^+) - V_{\alpha,n}(i,[j-1]^+) + \\ &V_{\alpha,n}(i,[j-1]^+) - V_{\alpha,n}(i,j) \\ = & V_{\alpha,n}([i-1]^+, [j-1]^+) - V_{\alpha,n}(i,j). \end{aligned}$$

It is noted that $a^*_{i,j} = 0$ and hence $\Delta\mathcal{V}_{\alpha,n}(i,j) \geq 0$, i.e.,

$$\begin{aligned} &C_T + C_H([i-1]^+ + [j-1]^+ - i - j) + \\ &\alpha\mathbb{E}[V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, [j-1]^+ + \mathcal{A}_1) - \\ &V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_1)] \geq 0. \end{aligned}$$

Therefore, it can be concluded that

$$\begin{aligned} &C_T + C_H([i-1]^+ + [j-1]^+ - i - j) + \\ &\alpha\mathbb{E}[V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, j+\mathcal{A}_2) - \\ &\quad V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) + V_{\alpha,n}(i+\mathcal{A}_1, [j-1]^+ + \mathcal{A}_2) - \\ &\quad V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)] \\ \geq & \Delta\mathcal{V}_{\alpha,n}(i,j) \geq 0. \end{aligned}$$

■

Appendix D
Proof to Lemma 15

Proof: We want to show that $V_{\alpha,n+1}(i+1, j+1) - V_{\alpha,n+1}(i,j) \leq V_{\alpha,n+1}(i+2, j+1) - V_{\alpha,n+1}(i+1, j)$ for all i and j. There will be 5 cases of $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i+1,j+1}, a^*_{i+2,j+1})$ that need to be considered.

Case (i): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i+1,j+1}, a^*_{i+2,j+1}) = (1,1,1,1)$, we claim that

$$\begin{aligned} &C_H(i-[i-1]^+) + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}([i-1]^+ + \mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)] \\ \leq & C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+\mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)]. \end{aligned}$$

When $i,j \neq 0$, it is true according to the subconvexity of $V_{\alpha,n}(i,j)$. The argument is satisfied for $i = 0, j \neq 0$ due to the the non-decreasing $V_{\alpha,n}(i,j)$, and for the case $i \neq 0, j = 0$ due to the convexity of $V_{\alpha,n}(i,j)$. Otherwise, it holds according to the non-decreasing property.

Case (ii): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i+1,j+1}, a^*_{i+2,j+1}) = (0,0,0,0)$, we claim that

$$\begin{aligned} &\mathbb{E}[V_{\alpha,n}(i+1+\mathcal{A}_1, j+1+\mathcal{A}_2) - V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)] \\ \leq & \mathbb{E}[V_{\alpha,n}(i+2+\mathcal{A}_1, j+1+\mathcal{A}_2) - \\ &\quad V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2)]. \end{aligned}$$

The above results from the subconvexity of $V_{\alpha,n}(i,j)$.

Case (iii): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i+1,j+1}, a^*_{i+2,j+1}) = (0,0,0,1)$, we claim that

$$\begin{aligned} &2C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+1+\mathcal{A}_1, j+1+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2)] \leq C_T. \end{aligned}$$

Since $a^*_{i+1,j+1} = 0$, we have $\Delta\mathcal{V}_{\alpha,n}(i+1,j+1) \geq 0$, i.e.,

$$\begin{aligned} &C_T - 2C_H + \alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, j+\mathcal{A}_2) - \\ &\qquad V_{\alpha,n}(i+1+\mathcal{A}_1, j+1+\mathcal{A}_2)] \geq 0. \end{aligned}$$

Hence the claim is verified.

Case (iv): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i+1,j+1}, a^*_{i+2,j+1}) = (0,0,1,1)$, it is trivial since the both $V_{\alpha,n+1}(i+1,j+1) - V_{\alpha,n+1}(i,j)$ and $V_{\alpha,n+1}(i+2,j+1) - V_{\alpha,n+1}(i+1,j)$ are zeros.

Case (v): if $(a^*_{i,j}, a^*_{i+1,j}, a^*_{i+1,j+1}, a^*_{i+2,j+1}) = (0,1,1,1)$, we claim that

$$\begin{aligned} C_T \leq\ & C_H(1+j-[j-1]^+) + \\ &\alpha\mathbb{E}[V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2) - \\ &\quad V_{\alpha,n}(i+\mathcal{A}_1, [j-1]^+ + \mathcal{A}_2)]. \end{aligned}$$

Notice that $a^*_{i+1,j} = 1$, so $\Delta\mathcal{V}_{\alpha,n}(i+1,j) \leq 0$, i.e.,

$$\begin{aligned} &C_T - C_H(1+j-[j-1]^+) + \\ &\alpha\mathbb{E}[V_{\alpha,n}(i+\mathcal{A}_1, [j-1]^+ + \mathcal{A}_2) - \\ &\quad V_{\alpha,n}(i+1+\mathcal{A}_1, j+\mathcal{A}_2)] \leq 0. \end{aligned}$$

■

References

[1] M. Effros, T. Ho, and S. Kim, "A Tiling Approach to Network Code Design for Wireless Networks," *Proc of IEEE ITW*, pp. 62–66, 2006.

[2] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network Information Flow," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, 2000.

[3] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "XORs in the Air: Practical Wireless Network Coding," *IEEE/ACM Trans. Netw.*, vol. 16, pp. 497–510, 2008.

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

[5] Y. E. Sagduyu, L. Georgiadis, L. Tassiulas, and A. Ephremides, "Capacity and Stable Throughput Regions for the Broadcast Erasure Channel With Feedback: An Unusual Union," *IEEE Trans. Inf. Theory*, pp. 2841–2862, 2013.

[6] Y. E. Sagduyu and A. Ephremides, "Cross-Layer Optimization of MAC and Network Coding in Wireless Queueing Tandem Networks," *IEEE Trans. Inf. Theory*, vol. 54, pp. 554–571, 2008.

[7] A. Khreishah, C. C. Wang, and N. Shroff, "Cross-Layer Optimization for Wireless Multihop Networks with Pairwise Intersession Network Coding," *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 606–621, 2009.

[8] V. Reddy, S. Shakkottai, A. Sprintson, and N. Gautam, "Multipath Wireless Network Coding: A Population Game Perspective," *Proc of IEEE INFOCOM*, 2010.
[9] A. Eryilmaz, D. S. Lun, and B. T. Swapna, "Control of Multi-Hop Communication Networks for Inter-Session Network Coding," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1092–1110, 2011.
[10] T. Ho and H. Viswanathan, "Dynamic Algorithms for Multicast With Intra-Session Network Coding," *IEEE Trans. Inf. Theory*, vol. 55, pp. 797–815, 2009.
[11] Y. Xi and E. M. Yeh, "Distributed Algorithms for Minimum Cost Multicast With Network Coding," *IEEE/ACM Trans. Netw.*, vol. 18, pp. 379–392, 2010.
[12] E. C. Ciftcioglu, Y. E. Sagduyu, R. A. Berry, and A. Yener, "Cost-Delay Tradeoffs for Two-Way Relay Networks," *IEEE Trans. Wireless Commun.*, pp. 4100–4109, 2011.
[13] D. N. T. Nguyen and X. Yang, "Multimedia Wireless Transmission with Network Coding," *In Proceedings of IEEE Packet Video*, pp. 326–335, 2007.
[14] Y. Huang, M. Ghaderi, D. Towsley, and W. Gong, "TCP Performance in Coded Wireless Mesh Networks," *Proc. of IEEE SECON*, pp. 179–187, 2008.
[15] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. The MIT Press, 1994.
[16] S. M. Ross, *Introduction to Stochastic Dynamic Programming*. Academic Press, 1994.
[17] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. Wiley-Interscience, 1998.
[18] D. P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall, 1987.
[19] V. S. Borkar, "Control of Markov Chains with Long-Run Average Cost Criterion: The Dynamic Programming Equations," *SIAM Journal on Control and Optimization*, pp. 642–657, 1989.
[20] R. Cavazos-Cadena and L. I. Sennott, "Comparing Recent Assumptions for the Existence of Average Optimal Stationary Policies," *Operations Research Letters*, pp. 33–37, 1992.
[21] M. Schal, "Average Optimality in Dynamic Programming with General State Space," *Mathematics of Operations Research*, pp. 163–172, 1993.
[22] L. I. Sennott, "Average Cost Optimal Stationary Policies in Infinite State Markov Decision Processes with Unbounded Costs," *Operations Research*, vol. 37, pp. 626–633, 1989.
[23] G. Koole, *Monotonicity in Markov Reward and Decision Chains: Theory and Applications*. Now Publishers Inc, 2007.
[24] V. S. Borkar, "Convex analytic methods in markov decision processes," in *Handbook of Markov Decision Processes*, E. Feinberg and A. Shwartz, Eds. Springer US, 2002, pp. 347–375.
[25] S. Meyn, *Control Techniques for Complex Networks*. Cambridge University Press, 2007.
[26] W. N. Kang, F. P. Kelly, N. H. Lee, and R. J. Williams, "Product Form Stationary Distributions for Diffusion Approximations to a Flow-Level Model Operating under a Proportional Fair Sharing Policy," *ACM SIGMETRICS Performance Evaluation Review*, pp. 36–38, 2007.
[27] S. I. Resnick, *Adventures in Stochastic Processes*. Birkhauser, 1992.

Yu-Pin Hsu received the BS and MS degrees from the National Chiao Tung University, Taiwan, in 2005 and 2007 respectively, and his PhD degree from the Texas A&M University, USA, in 2014. He was with the Center for Information and Communications Technology in the National Chiao Tung University, as a research assistant between July 2007 and August 2009. Since October 2014, he has been a joint postdoctoral fellow with the Massachusetts Institute of Technology and Singapore University of Technology and Design. His current research interests center around theoretical foundations of communication networks with the focus on algorithmic and control-theoretic aspects.

Navid Abedini received his Bachelor of Science degree in electrical engineering from Sharif University of Technology, Iran, in 2008, and his Ph.D. in electrical and computer engineering from Texas A&M University in 2012. He is currently a Senior Systems Engineer at Qualcomm Flarion Technologies Inc., where he is working on the next wireless communications technologies like LTE. He has received the 2010 Capocelli prize for the best student paper at IEEE Data Compression Conference. His research interests include design and analysis of wireless communication networks (like device to device networks), information and queueing theory.

Natarajan Gautam (SM '13) is a Professor in the Department of Industrial and Systems Engineering at Texas A&M University with a courtesy appointment in the Department of Electrical and Computer Engineering. Prior to joining Texas A&M University in 2005, he was on the Industrial Engineering faculty at Penn State University for eight years. He received his M.S. and Ph.D. in Operations Research from the University of North Carolina at Chapel Hill, and his B.Tech. from Indian Institute of Technology, Madras. His research interests are in the areas of modeling, analysis and performance evaluation of stochastic systems with special emphasis on optimization and control in computer, communication and information systems. He is an Associate Editor for the INFORMS Journal on Computing, IIE Transactions, and OMEGA.

Alex Sprintson (S'00-M'03) is an Associate Professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station. From 2003 to 2005, he was a Postdoctoral Research Fellow with the California Institute of Technology, Pasadena. His research interests lie in the general area of communication networks with a focus on network coding and software defined networks.

Dr. Sprintson received the Wolf Award for Distinguished Ph.D.students, the Viterbi Postdoctoral Fellowship, and the NSF CAREER award. Currently, he serves as an associate editor of the IEEE Transactions on Wireless Communications. He has been a member of the Technical Program Committee for the IEEE Infocom 2006–2015.

Srinivas Shakkottai (S '00-M'08) received his Bachelor of Engineering degree in electronics and communication engineering from the Bangalore University, India, in 2001 and his M.S. and Ph.D degrees from the University of Illinois at Urbana-Champaign in 2003 and 2007, respectively, both in electrical engineering. He was Postdoctoral Scholar at Stanford University until December 2007, and is currently an Assistant Professor in the Department of Electrical and Computer Engineering at Texas A&M University. He has received the Defense Threat reduction Agency Young Investigator Award (2009) and the NSF Career Award (2012) as well as research awards from Cisco (2008) and Google (2010).

His research interests include content distribution systems, pricing approaches to network resource allocation, game theory, congestion control, and the measurement and analysis of Internet data.