

Throughput Optimal Decentralized Scheduling of Multi-Hop Networks with End-to-End Deadline Constraints: Unreliable Links

Rahul Singh *Member, IEEE*, and P. R. Kumar *Fellow, IEEE*,

Abstract—We consider multi-hop networks serving multiple flows in which packets not delivered to their destination nodes by their deadlines are dropped from the network. The throughput of packets that are delivered within their end-to-end deadlines is called the timely-throughput. We address the design of policies for routing and scheduling packets that optimize any specified weighted average of the timely-throughputs of several flows, under nodal power constraints.

We provide a new approach which directly yields an optimal distributed scheduling policy that attains any desired maximal timely-throughput vector (i.e., any point on the Pareto frontier) under average-power constraints on the nodes. This completely distributed and tractable solution structure arises from a novel intrinsically stochastic decomposition of the Lagrangian of the constrained network-wide Markov Decision Process rather than of the fluid model.

The derived policies are highly decentralized in several ways. All decisions regarding a packet's transmission scheduling, transmit power level, and routing, are based solely on the age of the packet, not requiring any knowledge of network state or queue lengths at any of the nodes. Global coordination is achieved through a "price" for energy usage paid by a packet each time that its transmission is attempted at a node. This price decouples packets entirely from each other. It is different from that used to derive the backpressure policy where price corresponds to queue lengths.

The complexity of calculating the prices is tractable, being related only to the number of nodes multiplied by the relative deadline bound, and is considerably smaller than the number of network states which is exponentially large. Prices can be determined offline and stored.

If nodes have peak-power constraints instead of average-power constraints, then the decentralized policy obtained by truncation is near-optimal with respect to the timely-throughput as link capacities increase in a proportional way.

Index Terms—Communication Networks, Wireless Networks, Delay Guarantees in Networks, Scheduling Networks, Quality of Service.

I. INTRODUCTION

The past quarter century has seen the pioneering work of Tassioulas and Ephremides [1], Lin and Shroff

[2], Lin, Shroff and Srikant [3], and Neely, Modiano and Rohrs [4] on Max-Weight and backpressure based scheduling policies for communication networks that are provably throughput optimal, attaining any desired maximal throughput vector on the Pareto frontier of the feasible throughput region. In this paper we develop completely decentralized and tractable scheduling policies that achieve any desired maximal throughput vector of packets that meet specified hard end-to-end relative deadlines for packets under average-power constraints on nodes.

To see why this is a challenge, one may consider the situation depicted in Figure 1. Suppose that Node i needs to decide whether to serve the packet of Flow 1 or the packet of Flow 2. Packets of Flow 1 are experiencing downstream congestion, in contrast to packets of Flow 2 which face no downstream congestion. Thus Node i is better off serving the packet of Flow 2 rather than Flow 1 since the latter would anyway get delayed and not make it to its destination on time. Therefore network state is useful information. The central contribution of this paper shows just how to obtain an optimal distributed scheduling policy when nodes face average-power constraints. When nodes face peak-power constraints we provide a distributed approximately optimal policy that approaches optimality in a precisely quantifiable way.

The main contribution is the design of completely decentralized optimal routing and scheduling policies which can attain any desired maximal throughput vector of packets that are delivered end-to-end by their stipulated deadlines, when nodes have average-power constraints. That is, the policies are precisely optimal with respect to the throughput delivered under hard end-to-end delay constraints. These results addressing per-packet hard delay bounds are obtained by considering a decomposition of the Lagrangian of the constrained network-wide Markov Decision Process that is intrinsically stochastic and different from a fluid-based analysis. We show that a policy where each node makes decisions based only on the age of the packets present at it, and a prior computable price of transmission, oblivious to all else in the network, is optimal. This vastly simplifies the network operation. If the nodes instead have peak-power constraints, then the decentralized policy can be simply truncated to yield a policy that is near-optimal in the same quantifiable sense as Whittle's relaxation for multiarmed bandits [5], [6].

R. Singh is at 32-D716, LIDS, MIT, Cambridge, MA 02139; P. R. Kumar is at Dept. of ECE, Texas A&M Univ., 3259 TAMU, College Station, TX 77843-3259. rsingh12@mit.edu, prk@tamu.edu.

Preferred address for correspondence: P. R. Kumar, Dept. of ECE, Texas A&M Univ., 3259 TAMU, College Station, TX 77843-3259.

This material is based upon work partially supported by NSF under Contract Nos. CNS-1302182 and NSF Science & Technology Center Grant CCF-0939370, and USARO under Contract W911NF-15-1-0279.

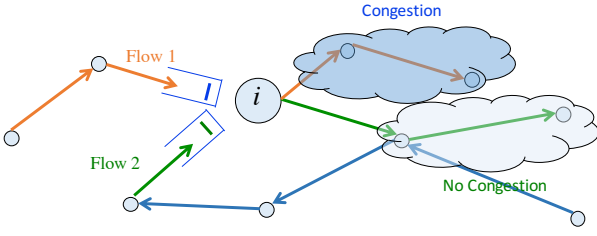


Fig. 1: The challenge of distributed scheduling. Suppose that Node i wishes to transmit one packet, either a packet of Flow 1 or a packet of Flow 2. Suppose packets of Flow 1 are experiencing downstream congestion, but packets of Flow 2 face no downstream congestion. Then serving a packet of Flow 1 is not useful since it will just get stuck in downstream congestion and not make it to its destination in time. Hence Node i should serve a packet of Flow 2. Therefore network state is useful information. This creates a chicken and egg situation since obtaining instantaneous network state information requires zero-delay communication of information over the network, while the very purpose of obtaining network state information is to provide low delay communication. We show that this conundrum can be completely resolved when nodes have average-power constraints.

In this paper we address the case where links are unreliable. This is of interest in networks with directional antennas, or networks of microwave repeaters, or even networks composed of unreliable wireline links. In a companion paper [7] we address the case where the links face interference.

Delivering packets on time is of great interest in emerging applications such as cyber-physical systems, where control-loops are closed over networks, are sensitive to delays. Similarly, Quality of Service (QoS) requirements for real-time applications such as video streaming, VoIP, surveillance, sensor networks, mobile ad-hoc networks (MANETS), and in-vehicular networks, all entail that packets should be delivered on time [8].

The rest of this paper is organized as follows. In Section II we describe the problem considered and the main results. In Section III we summarize previous work and set this work in context. In Section IV we describe the system model. In Section V we describe the constrained network-wide MDP. In Section VI we show the packet-level decomposition of the network-wide Lagrangian. In Section VII we describe the Single-Packet Transportation MDP that arises. In Section VIII we consider the Dual Problem of the constrained network-wide MDP and establish strong duality. In Section IX we show that the implementation of the packet-level optimal transportation policies yields an overall optimal policy for the entire system. In Section X we show how the overall optimal policy is obtained through a tractable linear program. In Section XI we show that there is a further simplifying threshold structure for the optimal policy. In Section XII

we show how the optimal prices can be precomputed. In Section XIII we address the problem with link capacities or, equivalently, peak-power constraints. In Section XIV we establish the asymptotic optimality of the truncated policy. In Section XV we address the problem when the channel condition changes with time. In Section XVI we address the problem where there are both real-time flows as well as non-real-time flows. In Section XVII we provide examples showing how the theory can be used to determine optimal distributed policies, and also present a comparative simulation study of the performance of the truncated policy for link-capacity constraints. We conclude in Section XVIII.

II. PROBLEM CONSIDERED AND MAIN RESULTS

We consider multi-hop, multi-flow networks in which packets of all flows have a hard end-to-end relative deadlines. (The relative deadline is the remaining time-till-deadline when a packet arrives). Nodes can transmit packets at varying power levels. Since the wireless channel is unreliable, the outcome of packet transmissions is modeled as a random process. Nodes can transmit and receive packets simultaneously. The throughput of packets of a flow that meet the end-to-end relative deadline constraint is called the timely-throughput. We consider the following two types of nodal constraints in this paper: a) An average-power constraint on each node in the network, or b) A link-capacity constraint on each network link which bounds the number of concurrent packets that can be transmitted on it at any given time t , or, equivalently, a peak-power constraint at each node.

Our goal is to design a decentralized, joint scheduling, transmission and routing policy, abbreviated as a “scheduling policy,” that maximizes the weighted sum of the timely-throughputs of the flows, for any given nonnegative choice of weights. That is, it can attain any point on the Pareto frontier of the timely-throughput vector. To be optimal, the policy should be dynamic enough and take into account in an online fashion the following factors:

- *Routing*: The policy will need to dynamically route packets so as to avoid paths that have a higher delay or nodes with lower power budgets.
- *Scheduling*: The policy will need to prioritize packet transmissions based on their age, the channel conditions, and the congestion at the nodes lying on the paths to their destinations.
- *Energy Efficiency and Channel Reliability*: It will need to choose the power levels of packet transmissions to balance between reliability and energy consumption. If channel states are time-varying, the policy will have to carry out packet transmissions opportunistically when the channel states are “good” so that the maximum throughput under deadline constraints is attained in an energy-efficient manner [4], [9], [10]. This also involves a trade-off between packets missing their deadlines on account of bad channel conditions, and between spending more energy to

transmit it in a bad channel state so that it reaches the destination within its deadline.

The main result is the determination that a Markov Decision Process for a certain “Single-Packet Transportation Problem” governs the behavior of each packet, oblivious to all other traffic or network state. In this standalone problem, a single packet optimizes its progress through the network, paying prices to nodes every time it requests transmission, but is compensated with a reward if it reaches its destination prior to the hard deadline. The only manner in which this optimal single-packet transportation problem is coupled to the overall network, nodal power constraints, other flows and other packets, is through predetermined prices for nodal transmissions. The optimal prices can be tractably computed off-line and stored. Determining the optimal policies for all packets is also of tractable complexity, involving a linear program with the number of variables equal to the product of the square of the number of nodes, the number of flows and the maximum relative deadline, rather than the network state that is exponential in the above quantity. Thereby, we obtain optimal distributed policies for maximizing the network’s timely-throughput of packets meeting hard per-packet deadlines under average-power constraints on the nodes.

The key to these results is to pursue a fundamentally stochastic approach that considers the Lagrangian of the constrained network-wide Markov Decision Process (MDP) governing the entire network, and showing how it decomposes into packet-by-packet decisions. This decomposition approach allows treatment of the intrinsically variability related aspects such as delay, in sharp contrast to the backpressure approach that considers the decomposition of the Lagrangian of the fluid model. Thus the approach of this paper is able to address delay rather than just throughput. Through this novel decomposition, we can address timely-throughput optimality of packets that meet hard per-packet delay deadlines, rather than just throughput optimality. Moreover, it does so producing a completely distributed and tractable policy, for systems with average-power constraints at nodes and unreliable links.

When nodes have peak-power constraints in addition to, or in place of, average-power constraints, one can simply truncate the above optimal policy dynamically to respect the peak-power constraint, and obtain a policy that is quantifiably near-optimal. Specifically, it is asymptotically optimal in the same manner as Whittle’s relaxation for multi-armed bandits [5], [6].

The exposition proceeds as follows. We begin with the overall problem with average nodal power constraints and invoke the scalarization principle to pose the problem of maximizing the network’s weighted timely-throughput subject to nodal average-power constraints as a constrained network-wide Markov Decision Process (MDP) [11]. We then solve this problem by considering the Lagrangian dual of this constrained network-wide MDP. The Lagrange multipliers associated with the

average-power constraints are interpreted as prices paid by a packet for utilizing energy every time its transmission is attempted by a node. As recompense, a packet collects a reward equal to the weight of its flow when it reaches its destination within its specified hard deadline. This results in a very convenient packet-level decomposition into an Optimal Single-Packet Transportation Problem. The Markov Decision Process for this problem has a small-sized state-space: (Number of nodes)*(Bound on relative deadline), much smaller than the exponentially large number of states in the network-wide problem. The prices can be calculated off-line just by price tatonnement that drives the “excess power consumption” at nodes to zero. The optimal policies for all packets of all flows can be determined by solving a linear program with a small number of states. Importantly, the overall approach yields a completely distributed solution, where a node only needs to know the remaining times-till-deadlines of packets present at that node, that is timely-throughput optimal under average-power constraints at nodes.

When the constraints are on link-capacity and not its average utilized capacity, or equivalently on peak-power and not average-power, one obtains a near-optimal policy by simply truncating the average power-optimal policy. The result is asymptotically optimal as the network capacity is increased, in the same sense as Whittle’s indexability [5] approach is asymptotically optimal as the population of bandits increases in proportion [6].

III. PREVIOUS WORKS

Over the past twenty-five years there have been several notable advances in the theory of networking.

In pioneering work, Tassiulas and Ephremides [1], Lin and Shroff [2], Lin, Shroff and Srikant [3], and Neely, Modiano and Rohrs [4] have shown that scheduling networks based on Max-Weight and backpressure are throughput optimal. The backpressure policy emerges naturally from a decomposition of the Lagrangian for the fluid network.

In another breakthrough, Jiang and Walrand [12] have designed a novel Adaptive Carrier Sensing Multiple Access algorithm for a general interference model that achieves maximal throughput through completely distributed scheduling under slow adaptation, without slot synchrony, if packet collisions are ignored. Combined with end-to-end control it also achieves fairness among the multiple flows.

In another seminal contribution, Kelly, Maulloo and Tan [13] have shown that the problem of congestion control of the Internet can be formulated as convex programming problem and have provided a quantitative framework for design based on primal or dual approaches.

Eryilmaz, Ozdaglar and Modiano [14] have developed a throughput optimal randomized algorithm for routing and scheduling of the common two-hop interference model that can be implemented in a distributed way with polynomial complexity. They have also developed such

a policy for inelastic flows that takes flow control into account and results in a fair allocation of the network's capacity.

Any effort at developing a delay-optimal scheduling policy needs to take into account the time-till-deadline of packets in the network. The CSMA algorithm does not do so. The backpressure policy schedules packets only on the basis of rate-weighted queue lengths of nodes and provides no delay guarantees. In fact, fluid-based policies such as the backpressure policy, should be expected to and have been shown to be throughput optimal, but they should not be expected to provide delay optimality. They can perform poorly with regard to delay performance [15]–[18]. For optimal delay performance, one needs to start with a fundamentally stochastic framework that takes fluctuations into account. This is akin to the difference between the law of large numbers and the central limit theorem. Such a stochastic framework is the path that is pursued in this paper.

There has been considerable progress on the problem of scheduling an access point, in which multiple one-hop flows with hard relative deadlines share a wireless channel. The Pareto optimal frontier of timely-throughput vectors has been characterized, and simple optimal policies have been determined [19]–[32].

Li and Eryilmaz [33] consider the problem of scheduling deadline-constrained packets over a multi-hop network. However, the proposed policies are not shown to have any provable guarantees on the resulting timely-throughput. To the best of the authors' knowledge, Mao, Koksall and Shroff [34] is the only work which provides a provable sub-optimal policy for deadline-constrained networks, though the policies proposed therein guarantee only a fraction, $(\text{Length of the longest route in the network})^{-1}$, of the timely-throughput capacity region.

IV. THE SYSTEM MODEL

We consider networks in which the data-packets have a hard deadline constraint on the time within which they should be delivered to their destination nodes if they are to be counted in the throughput.

The communication network of interest is described by a directed graph $G = (V, \mathcal{E})$ as shown in Figure 2, where $V = \{1, 2, \dots, |V|\}$ is the set of nodes that are connected via communication links. A directed edge $(i, j) \in \mathcal{E}$ signifies that node i can transmit data packets to node j . We will call this link $\ell = (i, j)$.

We assume that time is discrete, and evolves over slots numbered $1, 2, \dots$. One time-slot is the time taken to attempt a packet transmission over any link in the network.

There are a finite number of transmit power levels at which a packet can be transmitted. For convenience, we normalize each time slot to 1 second, so that power and energy of a transmission are interchangeable.

The outcome of a transmission over a link between any two nodes is allowed to be random, which enables us to

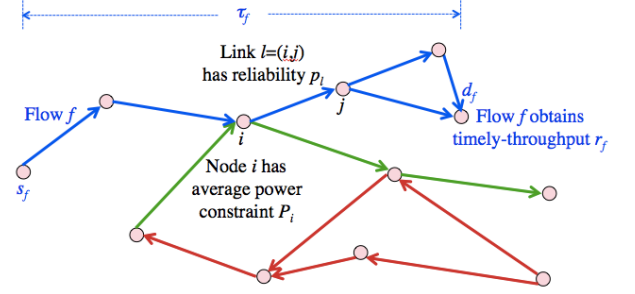


Fig. 2: Multi-hop network serving F flows. Flow f , with source s_f and destination d_f , has several feasible routes. Its end-to-end relative deadline is τ_f . Node i has an average-power constraint P_i . A packet transmitted on link $\ell = (i, j)$ has a probability p_ℓ of being successfully received by node j . Though not so indicated, the probability p_i may depend on the power level of the packet's transmission by node i .

model unreliable channels. If a packet transmission occurs on the link ℓ at a certain power level that consumes energy E , then the transmission is successful with probability $p_\ell(E)$, which is monotone decreasing in E . We can model the phenomena of wireless fading by allowing the success probability $p_\ell(t, E)$ to also be a function of time that can be assumed to be governed by a finite-state Markov process, whose state is known at the transmitting node. However, for simplicity of exposition, we consider time-invariant $p_\ell(E)$'s only. In this paper we do not consider contention for the transmission medium; the case of interference is considered in the companion paper [7].

The network is shared by F flows. Packets of flow f have source node s_f and destination node d_f . They may traverse one of several alternative routes.

The numbers of packet arrivals of a flow at its source node are i.i.d. across time-slots, though the distribution can vary from flow to flow. For simplicity of exposition we suppose that these distributions have bounded support, i.e., the number of arrivals is bounded, though we can relax this to merely assuming they are finite valued. Packets across flows are independent. The analysis below carries over to the case when the arrivals and relative deadlines (detailed below) are governed by a finite-state Markov process. We will denote the average arrival rate of flow f in packets/time slot by A_f .

Each packet of flow f has a “relative-deadline,” or “allowable delay” τ_f . If a packet of flow f arrives to the network at time t , then it needs to be either delivered to its destination node by time-slot $t + \tau_f$, or else it is discarded from the network at time $t + \tau_f$ if it has not yet reached its destination d_f . We suppose that all relative deadlines of packets are bounded by a quantity Δ . We can allow packets of a flow to have random independent and identically distributed (i.i.d.) deadlines, independent across flows, however, for simplicity of exposition we will suppose that all packets of flow f have the same relative

deadline τ_f .

The “timely-throughput” r_f attained by a flow f under a policy is the average number of packets delivered prior to deadline expiry per unit time, i.e.,

$$r_f := \liminf_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \delta_f(t), \quad (1)$$

where the random variable $\delta_f(t)$ is equal to the number of packets of a packet of flow f that are delivered in time to their destination at time t , with the expectation taken under the policy being applied.

The vector $\mathbf{r} := (r_1, r_2, \dots, r_F)$ is called the “timely-throughput vector.” A timely-throughput vector \mathbf{r} that can be achieved via some scheduling policy will be called an “achievable timely-throughput vector”. The set of all achievable timely-throughput vectors constitutes the “rate-region,” denoted by Λ .

Given weights $\beta_i \geq 0$ for the timely-throughput of each flow f_i , we will define the “weighted timely-throughput” as $\beta^T \mathbf{r}$, where $\beta = (\beta_1, \beta_2, \dots, \beta_F)$.

In Sections V–XII we consider an average-power constraint on each node $i \in V$. If the total energy consumed by all the concurrent packet transmissions on link ℓ at time t is $E_\ell(t)$ units of energy, then the nodal average-power constraints are given by,

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell: \ell=(i, \cdot)} E_\ell(t) \leq P_i, \forall i \in \{1, 2, \dots, V\}. \quad (2)$$

The second summation above is taken over all links ℓ , where $\ell = (i, j)$ for some node j . For simplicity of exposition, we suppose that the number of power levels available to choose from for any transmission is finite. We note that the above constraint on the average-power allows a node to transmit packets simultaneously over several outgoing links, which can be achieved via employing various techniques such as TDMA, OFDMA, CDMA etc., [35]–[37]. We suppose that nodes can simultaneously receive any number of packets while they are transmitting.

We will derive completely distributed scheduling policies that maximize the weighted timely-throughput for any given weight vector β for the unreliable multi-hop network under the average-power constraint (2) on nodes.

As an alternative to (2), or in addition to it, in Section XIII we will consider peak-power constraints on each link:

$$E_\ell(t) \leq C_\ell, \forall \ell \in \mathcal{E}, \text{ and } t = 1, 2, \dots \quad (3)$$

Alternatively, we can constraint the number of concurrent packets that can be transmitted on a link ℓ at each time t . For either of these situations, we will obtain quantifiably near-optimal distributed scheduling policies.

V. CHARACTERIZING THE RATE REGION: THE CONSTRAINED NETWORK-WIDE MDP

In order to characterize the network’s rate-region Λ , it is sufficient to characterize the set of Pareto-optimal

vectors $\{\mathbf{r} : \mathbf{r} \text{ is a timely-throughput vector and } \exists \beta \in \mathbb{R}_+^N \text{ such that } \mathbf{r} \in \arg \max_{\mathbf{y} \in \Lambda} \sum_f \beta_f y_f\}$, since Λ is simply its closed convex hull. The problem of obtaining Λ therefore reduces to that of finding scheduling policies which maximize weighted timely-throughputs of the form $\beta^T \mathbf{r}$.

The latter problem can be posed as a Constrained Network-wide Markov Decision Process (CMDP) [38]. The state of an individual packet present in the network at time t is described by the flow f to which it belongs, and the two tuple (i, s) , where i is the node at which it is present, and s is the time-to-go till its deadline. The state of the network at time t , $X(t)$, is described by specifying the state of each packet present in the network at time t . Since the time spent by a packet in the network is bounded by Δ , and since the number of arrivals in any time slot is also bounded due to the bounded support assumption, the system state $X(t)$ takes on only finitely many values, though it will be exponentially large.

A scheduling policy π has to choose, at each time t , possibly in a randomized way, which packets to transmit at each node from the set of available packets, and over which links and at what powers. The link choice allows routing to be optimized. The choice made at time t will be denoted $U(t)$.

Since the probability distribution of the system state $X(t+1)$ at time $t+1$ depends only $X(t)$ and $U(t)$, the problem of maximizing the timely-throughput subject to node-capacity constraints (2) is a CMDP, where a reward of β_f is received when a packet of flow f is delivered to its destination before its deadline expires:

$$\text{Maximize}_\pi \liminf_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_f \sum_{t=1}^T \beta_f \delta_f(t),$$

Subject to (2).

Above, $\delta_f(t)$ is the number of packets of flow f delivered in time to d_f at time t .

The above CMDP, parameterized by the vector $\beta := (\beta_1, \dots, \beta_F)$, is optimized by a stationary randomized policy [38]. Since the numbers of states and actions, are both finite, there is a finite set $\{\pi_1, \pi_2, \dots, \pi_M\}$ of stationary randomized policies such that for each value of β , there is a policy that belongs to this set and solves the CMDP [38]. Let $\gamma_1, \gamma_2, \dots, \gamma_M$ be the vectors of timely-throughputs associated with the policies $\pi_1, \pi_2, \dots, \pi_M$. We then have the following characterization of Λ :

Lemma 1:

$$\Lambda = \left\{ \mathbf{r} : \mathbf{r} = \sum_{i=1}^M \gamma_i c_i, c_i \geq 0, \sum_i c_i \leq 1 \right\}.$$

Note that, though finite, using the above lemma is intractable for computation since the number of stationary Markov policies is exponentially large in the following parameter: Maximum possible number of packets in the network \times Maximum path length of the flows \times Maximum possible relative deadline. We therefore seek to design significantly lower complexity decentralized scheduling policies that achieve the region Λ .

Since we can restrict ourselves to stationary randomized policies, we can replace \limsup and \liminf by \lim :

Lemma 2: Maximizing the timely-throughput (1) subject to nodal average power constraints (2) can equivalently be posed as the following CMDP:

$$\begin{aligned} & \text{Max}_{\pi} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_f \sum_{t=1}^T \beta_f d_f(t), \text{ subject to:} \\ & \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell: \ell=(i, \cdot)} E_{\ell}(t) \leq P_i, \forall i \in \{1, 2, \dots, V\}. \end{aligned} \quad (4)$$

VI. THE PACKET-LEVEL DECOMPOSITION OF THE LAGRANGIAN OF THE CONSTRAINED NETWORK-WIDE MDP

Defining λ_i as the Lagrange multiplier associated with the power constraint on node i , and $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_{|V|})$, we can write the Lagrangian for (4) as,

$$\begin{aligned} \mathcal{L}(\pi, \lambda) = & \lim_{T \rightarrow \infty} \frac{1}{T} \left\{ \mathbb{E} \sum_f \sum_{t=1}^T \beta_f \delta_f(t) \right. \\ & \left. - \sum_i \lambda_i \left(\mathbb{E} \sum_{t=1}^T \sum_{\ell: \ell=(i, \cdot)} E_{\ell}(t) \right) \right\} + \sum_i \lambda_i P_i, \end{aligned} \quad (5)$$

where the expectation is w.r.t. the policy π that is being used, the random packet transmission outcomes, and the randomness of the packet arrivals and relative deadlines, if the latter are random.

Denoting by $E_{\ell, f, n}(t)$ the amount of energy spent on transmitting the n -th packet of flow f at time t on link ℓ , we have,

$$E_{\ell}(t) = \sum_{f, n} E_{\ell, f, n}(t).$$

The Lagrangian (5) reduces to,

$$\mathcal{L}(\pi, \lambda) = \sum_i \lambda_i P_i \quad (6)$$

$$+ \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{f, n} \mathbb{E} \sum_{t=1}^T \left(\beta_f d_f(t) - \sum_i \lambda_i \sum_{\ell: \ell=(i, \cdot)} E_{\ell, f, n}(t) \right).$$

This can be decoupled completely on a packet-by-packet basis for any fixed value of the vector λ , as follows.

Let $\text{Packets}(f) := \text{set of all packets of flow } f$. We will denote a packet by σ . Let $\text{Packets}(f, t) := \text{subset of packets of flow } f \text{ that arrive before time } t$. Let $e(\sigma, i) := \text{total energy consumed by packet } \sigma \text{ at node } i$, $\phi(\sigma) := \text{flow that } \sigma \text{ belongs to}$, and $\delta(\sigma)$ be the random variable that assumes value one if packet σ reaches its destination before its deadline and zero otherwise.

Since the relative deadlines of packets are bounded, $\mathcal{L}(\pi, \lambda)$ can be rewritten as a sum over packets:

$$\mathcal{L}(\pi, \lambda) = \sum_i \lambda_i P_i \quad (7)$$

$$+ \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_f \sum_{\sigma \in \text{Packets}(f, T)} \left(\beta_f \delta(\sigma) - \sum_i \lambda_i e(\sigma, i) \right).$$

The term corresponding to packet σ of flow f ,

$$\mathbb{E} \left(\beta_f \delta(\sigma) - \sum_i \lambda_i e(\sigma, i) \right), \quad (8)$$

can be interpreted as follows. The packet incurs a payment of λ_i for using unit energy at node i , and accrues a reward of β_f if it reaches its destination before its deadline expires. This *Optimal Single-Packet Transportation Problem* is of very low complexity and is addressed further in Section VII. Let $R(f)$ denote its optimal expected cost.

Due to the decomposition of (7) over packets, we can optimize packet-by-packet. Hence we obtain,

$$\text{Max}_{\pi} \mathcal{L}(\pi, \lambda) = \sum_f A_f R_f(\lambda) + \sum_i \lambda_i P_i,$$

since A_f is the arrival rate of packets of flow f .

Therefore, for designing the policy π for maximizing the Lagrangian, we can simply solve the Optimal Single-Packet Transportation problem and let each packet make its own decision at each node on whether it wants to be transmitted, and if so, at what power level. No network state knowledge is needed by a packet to determine its optimal decision. Importantly, each packet's actions are independent of the actions chosen for all other packets in the network. It is very noteworthy that this results in a completely scheduling decentralized policy.

VII. THE SINGLE-PACKET TRANSPORTATION PROBLEM

The Optimal Single-Packet Transportation Problem is described as follows. A single packet of flow f is generated at time $t = 0$ at source node s_f , with state (s_f, τ_f) , where τ_f is the time-to-deadline. At each time step thereafter, the time-to-deadline is decremented by one. If it is not delivered to the destination node d_f by time $t = \tau_f$, then it is discarded from the network. A reward of β_f units is accrued if the packet reaches the destination node d_f by time τ_f . A price λ_i per unit energy has to be paid by the packet for transmission over an outgoing link at node i .

With the state of the packet described by the two tuple (i, s) , where i is the node at which it is present, and s is the time-till-deadline, we can use Dynamic Programming to solve for the value function V^f ,

$$\begin{aligned} V^f(i, s) = & \max\{V^f(i, s-1), \\ & \max_{j: (i, j) \in \mathcal{E}, E} \{-\lambda_i + p_{(i, j)}(E) V^f(j, s-1) \\ & + (1 - p_{(i, j)}(E)) V^f(i, (s-1)^+)\}\}, \end{aligned} \quad (9)$$

$$V^f(d_f, s) = \beta_f \text{ if } s \geq 0.$$

Solving for the maximizer on the RHS yields the optimal action, i.e., whether to transmit or not, and if so, at what level, for the packet of flow f in the state (i, s) .

It is important to note that this is a low complexity problem. Each packet's state only consists of the two tuple, (Node it is currently at, Time remaining to its deadline). This is simply a dynamic programming problem over a time horizon of Δ , with $|V|\Delta$ states, in contrast to the exponentially large size, $(V\Delta)^{F\Delta}$, of the original network's state-space.

VIII. THE DUAL OF THE CONSTRAINED NETWORK-WIDE PROBLEM AND STRONG DUALITY

The dual function is

$$D(\lambda) = \max_{\pi} \mathcal{L}(\pi, \lambda). \quad (10)$$

Importantly, it can be obtained in a decentralized fashion for any price vector λ , due to the decomposition into a collection of Optimal Single-Packet Transportation problems coupled only through the node prices λ_i .

The Dual Problem is:

$$\max_{\lambda \geq 0} D(\lambda). \quad (11)$$

Lemma 3: There is no duality gap.

Proof: The CMDP (4) can equivalently be posed as a linear program, in which the variables to be optimized are the *occupation measures* [11], [38]–[40] induced by the policy π on the joint state-action space. Being a linear program, the duality gap is zero. ■

IX. SYNTHESIZING OPTIMAL SINGLE-PACKET TRANSPORTATION PROBLEMS TO OBTAIN THE NETWORK-WIDE OPTIMAL POLICY

We now elaborate on how the optimal single-packet transport problem yields the overall network-wide optimal joint routing, scheduling and transmission policy.

The key is to use randomization when packets are indifferent to being transmitted at two power levels. This arises when two different choices both attain the maximum of the RHS of the dynamic programming equation (9). In such cases, the action taken can be chosen randomly from one of the maximizers. Such randomization allows satisfaction of the power constraints with equality, or, to put it another way, it allows us to fully use up all the power that is available at a node if beneficial.

Theorem 1: Let $\lambda^* \geq 0$ be a price vector. Denote by $\pi_f(\lambda^*)$ an optimal randomized policy for packets of flow f , and by $\pi(\lambda^*)$ the policy that implements $\pi_f(\lambda^*)$ for each packet belonging to flow f . Suppose that, at every node i , either the average-power constraint (2) is satisfied with equality by $\pi(\lambda^*)$, or $\lambda_i^* = 0$. Then $\pi(\lambda^*)$ is optimal for CMDP, and λ^* is optimal for the Dual Problem.

Proof: The dual function (10) is

$$D(\lambda^*) = \mathcal{L}(\pi(\lambda^*), \lambda^*). \quad (12)$$

The result simply follows from Complementary Slackness [41] since the primal problem can be written as a linear program over variables that are occupation measures. ■

Theorem 2 (Optimality of Decentralized Policy): The optimal policy for CMDP (4) is fully decentralized. In order for any node i to make a decision regarding a packet σ present with it at any time t , the node only needs to know the flow f that the packet belongs to, and the time-to-deadline of the packet. The optimal policy for node i simply consists of implementing the policy $\pi_f(\lambda^*)$ for packets of flow f , where λ^* is the optimal price.

We may observe the following key features of the solution. In order to solve the Primal Problem (4) in its original form, the network is required to make decisions based on the knowledge of the network state $X(t)$. The size of the state-space in which the network state $X(t)$ resides is exponential, $(V\Delta)^{F\Delta}$, since there can be $F\Delta$ packets in the network, with each being in one of $V\Delta$ states. Moreover, the optimal policy requires the entire network state information to be instantaneously known at each node at each slot. Indeed, one of the key reasons why optimal policies for communication networks (and other distributed systems) are generally intractable is that every decision requires instantaneous knowledge of the complete network state, which is something that cannot be obtained since the entire purpose of determining the optimal policy is to communicate information with deadlines. Thus, an approach based on directly solving the Constrained Network-Wide MDP (4) would have been computationally and implementationally futile.

These serious limitations have led us to instead formulate the Optimal Single-Packet Transportation Problem with the nodal transmission energy prices λ_i^* . This reduces the computational complexity from *exponential* to *linear*. Moreover the resulting solution can be implemented locally at each node. It is highly decentralized with no coupling between flows or nodes or even packets.

X. DIRECT OPTIMAL SOLUTION OF THE NETWORK-WIDE PROBLEM

After establishing the above completely decentralized structure in Theorem 2, we can directly obtain the network-wide optimal policy by embedding the very low-dimensional single-packet transportation problems of each flow into one flow-level problem, and then optimally allocating the power at each node among all the flows. We do this by considering the linear program involving “state-action probabilities” [42]. In this approach we do not need to first explicitly solve for the optimal prices.

Let us consider a single packet of flow f . From Theorem 2 we can restrict attention to randomized Markov policies where the packet is transmitted with a certain probability over a certain outgoing link, or not transmitted at all, with the probabilities depending only on the state (i, s) of the packet. Note also that links are unreliable; a packet transmitted over link (i, j) reaches j with probability p_{ij} . Under a randomized Markov policy, the packet moves stochastically through the network. We can delete a packet and remove it from the network as soon as s hits 0. Let $\xi^f(i, j, s)$ denote the probability that the packet is transmitted over link (i, j) when its time-till-deadline is s , where we use the convention that $\xi^f(i, i, s)$ is the probability that it is not transmitted, and also define $p_{ii} = 1$ correspondingly. (Note that $s = 0$ means that this is the last allowed transmission of the packet, and the packet will be deleted after this transmission). The “state-

action probabilities" $\{\xi^f(i, j, s)\}$ satisfy the constraints:

$$\begin{aligned} & \sum_{j \in V, j \neq d_f} \xi^f(j, i, s) p_{j,i} + \sum_{m \in V} \xi^f(i, m, s) (1 - p_{i,m}) \\ &= \sum_{k \in V} \xi^f(i, k, s-1) \quad \forall i \neq d_f, 1 \leq s \leq \tau_f, \end{aligned} \quad (13)$$

with the initial starting state s_f captured by the equation $\sum_{j \in V} \xi^f(s_f, j, \tau_f) = 1$. The probability that the packet reaches its destination d_f before deadline expiry is $\sum_{s=0}^{\tau_f} \sum_{i \in V} \xi^f(i, d_f, s) p_{i,d_f}$. If this policy is applied to all packets of flow f , then the average reward per unit-time obtained by this policy is obtained by simply multiplying by the arrival rate A_f , $\sum_{s=0}^{\tau_f} \sum_{i \in V} A_f \xi^f(i, d_f, s) p_{i,d_f}$. The energy consumed by a single packet at node i is $\sum_{s=0}^{\tau_f} \sum_{j \neq i} \xi^f(i, j, s) E$. The power consumption at node i due to the packets of flow f is $\sum_{s=0}^{\tau_f} \sum_{j \neq i} A_f \xi^f(i, j, s) E$.

Combining all the flows, we obtain the following *Direct Linear Program* to determine the optimal reward and the optimal Markov randomized policy:

$$\text{Max} \sum_{f \in F} \sum_{s=0}^{\tau_f} \sum_{i \in V} A_f \xi^f(i, d_f, s) p_{i,d_f}$$

Subject to:

$$\sum_{f \in F} \sum_{s=0}^{\tau_f} \sum_{j \neq i} A_f \xi^f(i, j, s) E \leq P_i \quad \forall i \in V,$$

$$\sum_{j \in V} \xi^f(s_f, j, 0) = 1 \quad \forall f,$$

$$\xi^f(i, j, s) \geq 0 \quad \forall i \neq d_f, \text{ with } i, j \in V, f \in F, 0 \leq s \leq \tau_f,$$

and equality constraints (13) for all f .

This linear program directly determines the optimal power allocation over flows at each node, and the optimal transportation policy for each packet. It also randomizes the actions of all packets of a particular identically so that the power available is utilized optimally.

This is a low complexity linear program with only $|V|^2 F \Delta$ variables and $|V| + |V| F \Delta + F + |V|^2 F \Delta$ constraints. This is a dramatic reduction of the complexity, and is eminently tractable being a linear program.

XI. OPTIMALITY OF A THRESHOLD POLICY

In fact, there is further structure that further reduces the complexity. For simplicity we illustrate this when there is only one transmit power level that corresponds to a fixed energy usage E for any transmission. We show that each packet's decision is simply governed by a threshold on time-to-deadline.

Theorem 3 (Threshold Policy): For each flow f , and node i , there is a threshold $\tau_f(i)$, such that the optimal decision for a packet of flow f at node i with a time-to-deadline s is to be transmitted/not transmitted according to whether the time-to-deadline s is strictly greater than/equal to or lesser than the threshold $\tau_f(i)$.

Proof: In a state where the decisions to transmit/not transmit are both optimal (i.e., the minimizer on the RHS

of the Dynamic Programming equation (9) is not unique), we choose "not to transmit," so that we thereby obtain an optimal policy that uniquely assigns an optimal action to each state. We will prove the following property (P) of this optimal policy, from which the theorem readily follows: (P) If the optimal decision is to not transmit a packet at a node, then it is optimal to never again transmit that packet at that node.

The reason is that one can then simply define $\tau_f(i)$ as the maximum value of s at which the decision to not transmit is the optimal action. Now we prove property (P) by using stochastic coupling. Suppose that for a packet of flow f at a node i it is optimal to not transmit it at time-to-deadlines equal to $\sigma, \sigma-1, \dots, \sigma-k$, but it is optimal to transmit it when its time-to-deadline is $\sigma-k-1$. Consider a packet, called Packet-1, that follows this optimal policy. It waits for k slots at node i , and then gets transmitted when its time-to-deadline is $\sigma-k$. Now consider another packet, called Packet-2, that waits no time at node i , and is transmitted when its time-to-deadline is σ . We will couple the subsequent experiences of Packet-1 and Packet-2, i.e., whether a transmission at a link is failure or success, after that transmission. Then, if Packet-1 reaches the destination d in time, then so does Packet-2. Hence the reward accrued by Packet-2 is no less than the reward accrued by Packet-1, while its costs are the same. Hence the decision of Packet-2 to immediately get transmitted at time-to-deadline σ is optimal. ■

However, one should not search for an optimal policy by trying to find the optimal thresholds. It is far better to search for optimal *prices* since they are the *same* for all packets of all flows, but the optimal threshold is only for a specific flow. That is, the prices result in the right trade-off between packets of different flows. However, if one obtains a set of thresholds that is "Person-by-Person" (or Nash) optimal for the flows, in the sense that each threshold is optimal for a particular flow when the thresholds of the other flows are fixed, then the entire set of thresholds may not be "Team" Optimal.

XII. DETERMINING THE OPTIMAL PRICES

Now we consider the problem of determining the correct prices $\lambda^* := \{\lambda_i^* : i \in V\}$ to be charged by the nodes. The key point to note is that price-determination can be done offline; prices can be pre-computed and stored.

One method is to just obtain them simply as the sensitivities of the power constraints in the Direct Linear Program. This requires a good model of the network and its reliabilities and the demands. However, one can also obtain them by "tatonnement" over a running system. Such price discovery is based on the Dual Function (10). First we discuss a hybrid of optimization and simulation, and subsequently a purely learning approach.

A. Repeated simulations

Price discovery can be performed offline by repeated simulation. Since the Dual Problem is convex, each node

can use subgradient descent to converge to the optimal price vector λ^* . The sub-gradient iteration is simply Walrasian tatonnement [43], which drives the “excess power consumption” at each node towards zero. Specifically, at the n -th iterate of the price vector, node i chooses

$$\lambda_i^{n+1} = \lambda_i^n + \epsilon [\text{Power consumed by node } i - P_i].$$

which simply amounts to a standard subgradient iteration [44] with a step-size $\epsilon > 0$. For any fixed price vector λ one can solve the dynamic programming equations for the optimal packet policy $\pi(\lambda)$.

B. Employing on-line learning

Instead of using simulation-based optimization to determine the optimal prices, one can determine both the optimal policy as well as the optimal policy contingent on that price, by using two time-scale stochastic approximation [45]–[47]. The faster time-scale stochastic approximation for the policy is

$$V_{n+1}(i, s) = 1_n(i, s) \{V_n(i, s) (1 - a_n) + a_n \max\{V_n(i, (s-1)^+), X\}\} + (1 - 1_n(i, s)) V_n(i, s),$$

where X is the maximum, over all links $\ell = (i, j)$, of

$$\lambda_{i,n} + p_\ell(E) V_n(j, (s-1)^+) + (1 - p_\ell(E)) V_n(i, (s-1)^+).$$

In the above, $1_n(i, s)$ assumes the value 1 if the packet-state at iteration n is (i, s) . $\{a_n\}$ is a positive sequence that satisfies $\sum_n a_n = \infty$, $\sum_n a_n^2 < \infty$.

We can use a slower time-scale stochastic approximation for the prices,

$$\lambda_{n+1} = \lambda_n (1 - b_n) + b_n (P - \bar{P}(\pi(\lambda_n))), \quad (14)$$

where P is the vector consisting of nodal power bounds, $\bar{P}_i(\pi(\lambda))$ is the average-power utilization at node i under $\pi(\lambda)$, and the sequence b_n satisfies $\sum_n b_n = \infty$, $\sum_n b_n^2 < \infty$, as well as $\sum_n (\frac{b(n)}{a(n)})^\gamma < \infty$, where $\gamma \geq 1$ [46]. The iterations converge to the optimal prices λ^* [48].

When network parameters are not known, one can both solve for the optimal policy $\pi(\lambda^*)$ as well as the optimal nodal prices λ^* in a decentralized manner. One way to achieve this task is to perform the Value Iterations using reinforcement learning for each price vector λ until convergence, and then to update the price λ using a gradient descent method.

XIII. LINK-CAPACITY CONSTRAINTS: THE TRUNCATION POLICY FOR NEAR-OPTIMAL ROUTING AND SCHEDULING

The analysis so far has addressed the nodal *average*-power constraints (2). In this section, we impose more stringent peak link-capacity constraints on the number of concurrent packets that can be scheduled at any given time slot t . $E_\ell(t)$, now re-defined as the number of packets transmitted on link ℓ at time t , has to satisfy

$$E_\ell(t) \leq C_\ell, \forall \text{ links } \ell \in \mathcal{E}, \text{ and } t = 1, 2, \dots \quad (15)$$

The more stringent problem that results is,

$$\begin{aligned} & \text{Maximize}_\pi \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_f \sum_{t=1}^T \beta_f d_f(t), \\ & \text{Subject to (15)}. \end{aligned} \quad (16)$$

We now proceed to construct a distributed policy with a provably close approximation to optimality. We begin with the policy that is optimal for the *average* power constraint

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T E_\ell(t) \leq C_\ell.$$

This is a distributed policy, as we have shown in the preceding sections, and is moreover tractable to compute. However it only ensures that the constraint (15) is met on average, and *not* at each time t . On the occasions that the number of packets that it prescribes for concurrent transmission does not exceed the constraint (15), we just transmit all the packets specified by that policy. However, on the occasions that it specifies an excessive number of transmissions exceeding the RHS of (15), we simply truncate the list of packets in the manner to be specified below, and only transmit a total of C_ℓ of those packets. Clearly, this leads to a policy that does satisfy the constraint (15) at each time instant. What we will show is that this policy is nearly optimal in a certain precise sense to be quantified below.

We first note the connection of our approach to that advocated by Whittle [5] for multiarmed bandits. Since there is no simple index policy [49] that is optimal when one is allowed to pull n arms concurrently, if $n > 1$, Whittle has suggested relaxing this constraint for *each* time t to a constraint that the *average* number of arms concurrently pulled is n . This relaxed problem has a tractable solution under an “indexability” condition [50]. Importantly, it is near-optimal when the number of arms available goes to infinity, with the proportion of arms of each type held constant [6]. Our approach can be regarded as an extension to multi-hop networks.

We first take care of one detail. The average-power constraints in (4) are nodal, while the link-capacity constraints in (16) are link-dependent. To reconcile this, we consider the following version of the problem (4) which involves average *link-wise* power constraints,

$$\begin{aligned} & \max_\pi \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_f \sum_{t=1}^T \beta_f d_f(t), \text{ subject to:} \\ & \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T E_{(i,j)}(t) \leq C_{(i,j)}, \forall \text{ links } (i,j) \in \mathcal{E}. \end{aligned} \quad (17)$$

The optimal policy for the above problem can be obtained in exactly the same fashion as for the problem (4), except that now there are link-based prices $\lambda_{(i,j)}$, instead of nodal prices λ_i .

Now we define the truncation policy for the problem (16) which involves link-capacity constraints $\{C_{(i,j)}\}$: If the policy π^* specifies that node i transmit more

than $C_{(i,j)}$ packets at some time t , then the truncation policy can pick any $C_{(i,j)}$ of these packets and transmit them. Moreover, we eject from the network those packets which π^* dictated to be scheduled, but were not picked for transmission. (Discarding the packets is not strictly required, but it simplifies the discussion). Let us denote this modified policy by $\tilde{\pi}^*$. It may be noted that under this policy, the evolution of the network is not independent across different packets, as was the case with π^* .

For the policy π^* , let us denote by $p_f(\tau, \ell)$ the probability that under π^* a packet (of flow f) with age τ time-slots would be attempted on link ℓ . Since the arrival rate of flow f packets is A_f , then, on account of the fact that the policy π^* satisfies the average-power-constraints C_ℓ imposed by the network, we have

Lemma 4:

$$\sum_f \sum_{\tau=1}^{\Delta} A_f p_f(\tau, \ell) \leq C_\ell, \forall \ell \in \mathcal{E}. \quad (18)$$

Next, we determine the level of sub-optimality of $\tilde{\pi}^*$.

A. A Lower Bound on Performance of $\tilde{\pi}^*$

We will now obtain lower bounds on the performance of the policy $\tilde{\pi}^*$. The following arguments are based on analysis of the evolutions of policies on an appropriately constructed probability space. Let us denote by r_0 the (average) reward earned by policy π^* . First note that the reward collected by $\tilde{\pi}^*$ (denoted by r_1) does not increase if it were to, instead of dropping a packet because of capacity constraint violation, schedule it as dictated by π^* , but no reward is given to it if this packet is delivered to its destination node (denoted by r_2). However r_2 is more than the reward if now a penalty of β_f units per-packet was imposed for scheduling a packet via utilizing “excess capacity” at some link $l \in \mathcal{E}$, but it were given a reward in case this packet reaches the destination node (denoted by r_3). r_3 is certainly more than the reward which π^* earns if it is penalized an amount equal to the sum of the excess bandwidths that its links utilize (denoted by r_4) multiplied by β_f , since any individual packet may be scheduled multiple times by utilizing excess bandwidth. Thus, the difference $r_0 - r_4$ is less than the sum of the excess bandwidths utilized by the links operating under the policy π^* , scaled by $\text{Max}_f \beta_f$'s. Let $E_{f,\ell,\tau}(t)$ denote the number of packets of flow f that have an age of τ time-slots, and are served on link ℓ at time t under the policy π^* . We thus have

Lemma 5:

$$r_0 - r_4 \leq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell \in \mathcal{E}} \left[\sum_g (\text{Max}_g \beta_g) (E_{f,\ell,\tau}(t) - C_\ell) \right]^+. \quad (19)$$

B. Asymptotic Optimality of $\tilde{\pi}^*$ in the limit

Next, we will scale the network parameters in two equivalent ways, and show that the policy $\tilde{\pi}^*$ is asymptotically optimal. In the discussion below, N is a scaling

parameter. In the first approach to scaling, the link capacities C_ℓ and the mean arrival rates A_f will be kept fixed, while the size of an individual packet will be scaled as $\frac{1}{N}$, with the arrivals being i.i.d. with binomial parameters $(N, A_f/N)$. The second equivalent formulation is to keep the size of packets fixed, while the link capacities for the N -th system are scaled as NC_ℓ , with the arrivals being i.i.d. with binomial parameters (N, A_f) . In the rest of the discussion, we will confine ourselves to the former formulation; however a similar analysis can be performed for the latter case.

Theorem 4: Consider the sequence of systems described in problem (17) parameterized by N , in which the arrivals for the N -th system are i.i.d. with binomial parameters $(N, A_f/N)$. The deviation from optimality of the N -th system in the sequence operating under the policy $\tilde{\pi}^*$ is $O(\frac{1}{\sqrt{N}})$, and hence the policy $\tilde{\pi}^*$ is asymptotically optimal for the joint routing-scheduling problem (16) under hard link-capacity constraints.

Proof: Below, $MAD(X) := \mathbb{E}|X - \bar{X}|$ is the mean absolute deviation of X with respect to its mean \bar{X} .

The deviation from optimality satisfies,

$$\begin{aligned} r_0 - r_4 &\leq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell \in \mathcal{E}} \left(\sum_{f,\tau} \beta_f (E_{f,\ell,\tau}(t) - C_\ell) \right)^+ \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell \in \mathcal{E}} \left\{ \sum_{f,\tau} \beta_f (E_{f,\ell,\tau}(t) - \bar{E}_{f,\ell,\tau}(t)) \right. \\ &\quad \left. + \sum_f \beta_f (\bar{E}_{f,\ell}(t) - C_\ell) \right\}^+ \\ &\leq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell \in \mathcal{E}} \left\{ \sum_{f,\tau} \beta_f (E_{f,\ell,\tau}(t) - \bar{E}_{f,\ell,\tau}(t)) \right\}^+ \\ &\leq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell \in \mathcal{E}} \sum_{f,\tau} \beta_f \{E_{f,\ell,\tau}(t) - \bar{E}_{f,\ell,\tau}(t)\}^+ \\ &\leq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell \in \mathcal{E}} \sum_{f,\tau} \beta_f (|E_{f,\ell,\tau}(t) - \bar{E}_{f,\ell,\tau}(t)|) \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{\ell \in \mathcal{E}} \sum_{f,\tau} \beta_f MAD(E_{f,\ell,\tau}(t)) \\ &= O\left(\frac{1}{\sqrt{N}}\right), \end{aligned}$$

where the last equality follows from [51]. \blacksquare

XIV. NEAR-OPTIMAL SCHEDULING WITH PEAK-POWER CONSTRAINTS

We can similarly consider the problem of maximizing the network's timely-throughput (1) subject to additional bounds on the *peak* power that can be utilized by each node $i \in V$. The treatment is similar to that in Section XIII, in which we designed optimal policies under

hard constraints on the *number* of packets that can be transmitted at any given time t .

The problem is formally stated as,

$$\text{Maximize}_{\pi} \liminf_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_f \sum_{t=1}^T \beta_f d_f(t), \text{ subject to} \quad (20)$$

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^T \sum_{\ell: \ell=(i,j)} E_{\ell}(t) \leq P_i, \forall i \in \{1, \dots, V\}, \quad (21)$$

$$\text{and} \quad \sum_{\ell: \ell=(i,j)} E_{\ell}(t) \leq P_i^{\max}, \forall i \in \{1, \dots, V\} \text{ and } t,$$

where P_i^{\max} (with value greater than P_i) is the peak-power constraint on node i .

Denote the policy which maximizes the objective function (20) under the average-power constraints (21) by π^* . Now we modify it to a policy $\tilde{\pi}^*$ described as follows: at each time t , each node $i \in V$ looks up the decision rule π^* and obtains the optimal power levels at which π^* would have carried out transmissions of the packets available with it. For this purpose, each node i only needs to have knowledge of the age of packets present with it. Node i then chooses a maximal subset of the packets present with it, such that the transmission power levels assigned to them by π^* sum to less than the bound P_i^{\max} . One way to choose such a set of packets and associated power levels is as follows: arrange the packets in decreasing order of the transmission power assigned by π^* , and label them. Then $\tilde{\pi}^*$ schedules the largest index packet such that the energy of all packets upto that index sum to less than P_i^{\max} . The asymptotic optimality of $\tilde{\pi}^*$ follows as in Theorem 4.

Theorem 5: Consider the sequence of networks operating under the policy $\tilde{\pi}^*$, in which the arrivals for the N -th system are i.i.d. with binomial parameters $(N, A_f/N)$. The deviation from optimality of the N -th system in the sequence operating under the policy $\tilde{\pi}^*$ is $O(\frac{1}{\sqrt{N}})$, and hence the policy $\tilde{\pi}$ is asymptotically optimal for the peak-power problem (20).

XV. WIRELESS FADING

Our model allows us to incorporate wireless fading. We model the channel state as a finite-state Markov process $Y(t)$, with the link transmission success probabilities $p_{\ell}(Y(t), E)$ a function of the channel state $Y(t)$ and the transmit power level. As before, we assume that the probabilities are monotone decreasing in E .

The network state is then described by a) the state of each packet, and b) the channel state $Y(t)$. The optimal policy can be determined along similar lines as before, by augmenting the system state with the channel state $Y(t)$. The optimal policy will be of the following form: the decision to be taken by a node i at time t will depend on the state of the packet and the channel state $Y(t)$.

The above assumes that the channel condition is known to each transmitter. A simplification is possible if we assume that the process $Y(t)$ is i.i.d., which would eliminate the need for communicating $Y(t)$. Alternately it could be

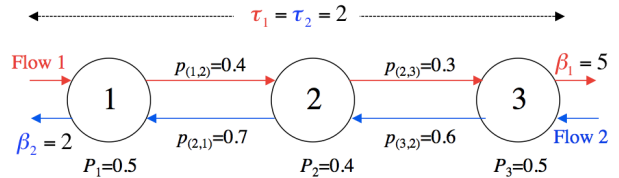


Fig. 3: System considered in Example 1.

deterministically time-varying. A common model which can be approximated is block fading [52], [53], under which the channel state need only be communicated periodically.

XVI. JOINTLY SERVING REAL-TIME AND NON-REAL-TIME FLOWS

In the previous sections we have considered networks exclusively serving real-time flows for which the utility of a packet arriving after its deadline is zero. Often one is interested in networks that serve a mix of real and non-real-time flows [54]. The system model can be easily extended. To incorporate this case, we simply set the relative deadlines of the packets belonging to the non-real-time flows as $+\infty$, so that they are never dropped.

XVII. ILLUSTRATIVE EXAMPLES AND SIMULATIONS

We first illustrate the amenability of the theory by explicitly hand-computing the optimal distributed policy in two examples. In the second example, the deadlines are slightly more relaxed than in the first example, and we show both how the prices change as a consequence, and how the optimal policy reacts to this.

Subsequently, we consider a more complex example and provide a comparative simulation illustrating the performance of the asymptotically optimal policy for the case of link-capacity constraints, comparing it with well known routing/scheduling policies such as the Backpressure, Shortest Path, and Earliest Deadline First (EDF) policies.

A. Two Illustrative Examples

Example 1: Consider the system shown in Figure 3. It consists of two flows traversing the Nodes 1, 2, and 3, in opposite directions. Flow 1, with source node $s_1 = 1$ and destination node $d_1 = 3$, has an end-to-end deadline τ_1 of 2 slots. Flow 2, with source node $s_2 = 3$ and destination node $d_2 = 1$, also has an end-to-end deadline τ_2 of 2 slots. Packets cannot afford even one failure on any transmission if they are to reach their destinations in time, since the relative deadlines for the flows are exactly equal to the total number of hops to be traversed. One packet of each flow arrives in every time slot, so $A_1 = A_2 = 1$. Each packet transmission at any node is at 1 watt, so $E = 1$ since all time slots are one second. Nodes 1, 2 and 3, have average-power constraints

$P_1 = 0.5, P_2 = 0.4$ and $P_3 = 0.5$ watts, respectively. Links (1,2), (2,3), (2,1) and (3,2) have reliabilities of $p_{(1,2)} = 0.4, p_{(2,3)} = 0.3, p_{(2,1)} = 0.7$ and $p_{(3,2)} = 0.6$, respectively. Denoting by r_1 and r_2 the timely-throughputs of Flows 1 and 2, we wish to maximize $5r_1 + 2r_2$, i.e., packets of Flow 1 are 2.5 times more valuable than packets of Flow 2. So $\beta_1 = 5$ and $\beta_2 = 2$.

The dynamic programming equations for the optimal single-packet transportation problem for Flow 1 yield:

$$\begin{aligned} V^1(1, 2) &= \text{Max}\{0, -\lambda_1 + 0.4V^1(2, 1)\}, \\ V^1(2, 1) &= \text{Max}\{0, -\lambda_2 + (0.3) \cdot 5\}. \end{aligned}$$

So $V^1(2, 1) = (1.5 - \lambda_2)^+$ and $V^1(1, 2) = [(0.6 - 0.4\lambda_2)^+ - \lambda_1]^+$. Similarly, for Flow 2, $V^2(3, 2) = [(0.84 - 0.6\lambda_2)^+ - \lambda_3]^+$ and $V^2(2, 1) = (1.4 - \lambda_2)^+$.

Packets of Flow 1 at Node 2 are more valuable than packets of Flow 2 at Node 2, since packets of Flow 1 have expected reward of $(0.3)5 = 1.5$, while packets of Flow 2 have expected reward of $(0.7)2 = 1.4$. So we will push as many packets of Flow 1 as possible to Node 2.

In order for a packet of Flow 1 to choose to be transmitted at Node 2, however, the price λ_2 that it pays needs to be less than the expected reward $(0.3)5$ that it can obtain in the future. Hence

$$\lambda_2 \leq 1.5.$$

Similarly, in order for a packet of Flow 1 to choose to be transmitted at Node 1, the total expected price it expects to pay, $\lambda_1 + 0.4\lambda_2$ (since λ_1 is the price it pays at Node 1, and if it succeeds to reach Node 2, which happens with probability 0.4, it then pays a price λ_2) must be less than the expected reward, which is $(0.4)(0.3)5$. Hence,

$$\lambda_1 + 0.4\lambda_2 \leq 0.6, \quad (22)$$

But Flow 1 can only push $(0.5)(0.4) = 0.2$ of its packets to Node 2. So there is spare capacity at Node 2 that Flow 2 can use. For Flow 2 to use that we need $\lambda_2 \leq (0.7)2 = 1.4$. Now, Flow 2 needs to utilize the spare capacity of 0.2 left at Node 2. So it needs to ensure a flow of 0.2 reaches Node 2. To do that it needs to transmit $\frac{1}{3}$ of the packets that arrive since $\frac{1}{3}(0.6) = 0.2$. So it needs to randomize at Node 3. By Complementary Slackness, this can only happen if packets at Node 3 are indifferent to being transmitted or not. So,

$$\lambda_3 + 0.6\lambda_2 = (0.6)(0.7)(2) = 0.84.$$

Since we want to maximize $D(\lambda)$ we choose $\lambda_2 = 1.4, \lambda_3 = 0$, and, from (22), $\lambda_1 = 0.04$.

Therefore, we arrive at the following solution, where we denote by $\pi^f(i, s)$ the probability with which a packet of flow f is transmitted at Node i when the time-to-deadline is s :

$$\begin{aligned} \lambda^* &= (0.04, 1.4, 0). \\ \pi^1(1, 2) &= 0.5, \pi^1(2, 1) = 1, \\ \pi^2(3, 2) &= \frac{1}{3}, \pi^2(2, 1) = 1. \end{aligned}$$

Now we verify that this policy is optimal. $\lambda_2 = 1.4$ implies $\pi^1(2, 1) = 1$ since $1.4 \leq (0.3)5$. Now $\lambda_1 + 0.4\lambda_2 = 0.6$ implies $\pi^1(1, 2) = 1$ and $\pi^1(2, 1) = 0$ are both optimal, i.e., a packet is indifferent to them, and so one may randomize between them to satisfy the average-power constraint. Similarly, $\lambda_2 = 1.4$ implies that both decisions $\pi^2(2, 1) = 1$ and $\pi^2(2, 1) = 0$ are optimal. Also, $\lambda_3 + 0.6\lambda_2 = 0.84$ implies both $\pi^2(3, 2) = 1$ and $\pi^2(3, 2) = 0$ are both optimal. So we can randomize the transmission of packets of Flow 2 in state (3, 2). The average-power usages are 0.5 watts at Node 1, 0.4 watts at Node 2, and $\frac{1}{3}$ watt at Node 3. The average-power constraints of $P_1 = 0.5$ and $P_2 = 0.4$ at Nodes 1 and 2, respectively, are met with equality. The average-power constraint at Node 3 is slack but $\lambda_3 = 0$. So complementary slackness holds. Hence the policy is optimal.

Example 2: We now consider the same system as in Example 1, except that we relax the relative deadlines to $\tau^1 = \tau^2 = 3$, so that every packet can afford to have one hop that is retransmitted and still make it to its destination in time.

Consider a packet that has just arrived at Node 1. It can either make it to its destination in two hops if both transmissions are successful the first time they are attempted, or it can fail once at Node 1 and then be successful on subsequent transmissions at Nodes 1 and 2, or it can succeed the first time at Node 1, fail once at Node 2, and then succeed at Node 2 on the second attempt. If it does so reach its destination, it obtains a reward of 5. Hence taking these possibilities into account, if a packet of Flow 1 gets transmitted at every available opportunity, then the Expected reward for a packet of Flow 1 at its first visit to Node 1 = $[(0.4)(0.3) + (0.6)(0.4)(0.3) + (0.4)(0.7)(0.3)]5 = 1.38$. Similarly, Expected reward for a packet of Flow 2 at its first visit to Node 3 = 1.428, Expected reward for a packet of Flow 1 at its second visit to Node 1 = 0.6, Expected reward for a packet of Flow 2 at its second visit to Node 3 = 0.84, Expected reward for a packet of Flow 1 at its first visit to Node 2 = 2.55, Expected reward for a packet of Flow 2 at its first visit to Node 2 = 1.82, Expected reward for a packet of Flow 1 at its second visit to Node 2 = 1.5. Expected reward for a packet of Flow 2 at its second visit to Node 2 = 1.4.

Packets of Flow 1 are more valuable at Node 2 than Flow 2. So we want to maximize the throughput of packets of 1 to Node 2. If we transmit with probability 0.5 on the first attempt at Node 1 then all power is used up. The maximum power that can be consumed by packets of Flow 1 at Node 2 = $(0.5)(0.4) + (0.5)(0.4)(0.7) = 0.34$ watts. So there is still 0.06 watts left at Node 2 that can be used by packets of Flow 2. After arriving at Node 2 for the first time, a packet of Flow 2 can use a maximum power 1.3 watts. So Flow 2 at Node 3 needs to make $\frac{0.06}{(1.3)(0.6)}$ attempts which amounts to randomization with probability $1/13$. In order to transmit a packet of Flow 2 on its second visit to Node 2, the price λ_2 cannot be any more than the expected reward $(0.7)2 = 1.4$. So we could attempt some of the packets of Flow 2 that arrive at

Node 3, and transmit some packets of Flow 2 that arrive at Node 2.

With $\lambda_2 = 1.4$, λ_1 needs to satisfy $\lambda_1 + (0.4)\lambda_2 + (0.4)(0.7)\lambda_2 = [(0.4)(0.7) + (0.4)(0.7)(0.3)]5$, so $\lambda_1 = 0.068$. Similarly, λ_3 needs to satisfy, $\lambda_3 + (0.6)\lambda_2 + (0.6)(0.3)\lambda_2 = [(0.6)(0.7) + (0.6)(0.7)(0.3)]2$, which yields $\lambda_3 = 0$. The power constraint at Node 3 is slack, but $\lambda_3 = 0$. So the price vector is $\lambda = (0.068, 1.4, 0)$. The corresponding probabilities of transmission are

$$\begin{aligned}\pi^1(1, 3) &= 0.5, \pi^1(1, 2) = 0, \\ \pi^1(2, 2) &= 1, \pi^1(2, 1) = 1, \\ \pi^2(3, 3) &= 1/13, \pi^2(3, 2) = 0, \\ \pi^2(2, 2) &= 1, \pi^2(2, 1) = 1.\end{aligned}$$

The optimal single-packet transportation dynamic programming equations yield:

$$\begin{aligned}V^1(1, 3) &= \text{Max} \{0, -0.068 + (0.6)V^1(1, 2) + 0.4V^1(2, 2)\} \\ &= 0, \text{ so both choices are optimal,} \\ &\text{permitting randomization,} \\ V^1(1, 2) &= \text{Max} \{0, -0.68 + (0.6)(0) + 0.4V^1(2, 1)\} = 0, \\ &\text{again, both choices are optimal,} \\ &\text{permitting randomization,} \\ V^1(2, 2) &= \text{Max} \{0, -1.4 + (0.7)V^1(2, 1) + (0.3)5\} = 0.17, \\ V^1(2, 1) &= \text{Max} \{0, -1.4 + (0.3)5\} = 0.1.\end{aligned}$$

In all of the below, both choices are again optimal,

$$\begin{aligned}V^2(3, 3) &= \text{Max} \{0, 0 + (0.6)V^2(2, 2) + 0.4V^2(3, 2)\} = 0, \\ V^2(3, 2) &= \text{Max} \{0, 0 + 0.6V^2(2, 1)\} = 0, \\ V^2(2, 2) &= \text{Max} \{0, -1.4 + (0.3)V^2(2, 1) + (0.7)2\} = 0, \\ V^2(2, 1) &= \text{Max} \{0, -1.4 + (0.7)2\} = 0.\end{aligned}$$

Note that the power consumptions are

$$\begin{aligned}P_1 &= (1)(0.5) = 0.5, \text{ so it is tight,} \\ P_2 &= (0.5)(0.4)[1 + (0.7)1] + \frac{1}{13}(0.6)[1 + (0.3)1] = 0.4, \text{ tight,} \\ P_3 &= \frac{1}{13}.\end{aligned}$$

The last constraint is loose, but then $\lambda_3 = 0$, and we still have complementary slackness. So the solution is optimal.

B. Simulations

Now we consider the case of link-capacity constraints (or equivalently peak-power constraints). We present a comparative simulation study of the asymptotically optimal policy with respect to the following two policies: a) Earliest Deadline First scheduling combined with Backpressure routing (EDF-BP), and b) Earliest Deadline First scheduling combined with Shortest Path routing (EDF-SP) that routes packets along the shortest path from source to destination with ties broken randomly. We consider the systems shown in Figures 4 and 5. All link capacities are just 1, so the asymptotically optimal policy is noteworthy

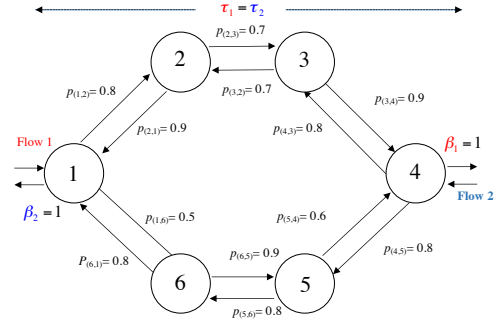


Fig. 4: Network with two source-destination pairs ($s_1 = 1, d_1 = 4$) and ($s_2 = 4, d_2 = 1$). Arrivals are deterministic with rates $A_1 = A_2 = 1$ per time-slot. Link capacities are $C_{(i,j)} \equiv 1$ packet/time-slot for all links (i, j) shown.

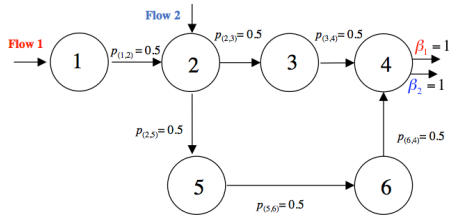


Fig. 5: Network with two source-destination pairs ($s_1 = 1, d_1 = 4$), ($s_2 = 2, d_2 = 4$). The arrivals are deterministic with rates $A_1 = A_2 = 1$ per time-slot. Link capacities are $C_{(i,j)} \equiv 1$ packet/time-slot for all links (i, j) shown.

for its excellent performance seen below even in the very much non-asymptotic regime.

We compare the performance of the asymptotically optimal policy $\tilde{\pi}^*$ of Theorem 4, with the following EDF-SP policy:

- 1) The link $\ell = (i, j)$ chosen for scheduling packet transmissions for flow f lies on the shortest path that connects the source and destination nodes of flow f .
- 2) Thereafter, on each link (i, j) , it gives higher priority to packets having earlier deadlines. It then serves a maximum of $C_{(i,j)}$ packets in decreasing order of priority.

We also compare the performance with the EDF-BP policy. Under the EDF-BP policy, each node i maintains queues for each flow f and possible age s . Denoting by $Q_{i,f}(t, s)$ the queue length at node i at time t , and by $Q_{f,i}(t) = \sum_s Q_{i,f}(t, s)$ the total number of packets of flow f at node i at time t , the policy functions as follows:

- 1) For each outgoing link $\ell = (i, j)$, EDF-BP calculates the backlogs $Q_{f,i}(t) - Q_{f,j}(t)$ of flow f .
- 2) On each link $\ell = (i, j)$ it prioritizes packets on the basis of the backlogs associated with their flows. For packets belonging to the same flow, higher priority is given to packets having earlier deadlines.
- 3) It then serves a maximum of $C_{(i,j)}$ highest priority packets from amongst the packets whose flows have

(τ_1, τ_2)	Links (i, j)					
	(1,2)	(2,3)	(3,4)	(2,5)	(5,6)	(6,4)
(4,3)	0.0	24.078	0.0	-0.5	22.7674	0.0
(5,4)	0.0107	48.3409	0.0518	0.0	48.7170	0.0
(6,5)	0.0037	40.3655	0.0276	0.0	39.7392	0.0

Fig. 6: Prices of links for the network in Fig. 5.

a positive backlog $Q_{f,i}(t) - Q_{f,j}(t)$.

Both EDF-SP and EDF-BP eject packets that have crossed their deadlines.

Plots 7 and 9 show the comparative performances of the policies for the networks in Fig. 4 and 5 as the relative deadlines of the flows are varied. The performance of the asymptotically optimal policy is superior even in the non-asymptotic regime. Plots 8 and 10 show the comparative performance as network capacities are increased.

Observe that for the network in Figure 5, a shortcoming of EDF-SP is that it is unable to utilize the path $1 \rightarrow 2 \rightarrow 5 \rightarrow 6 \rightarrow 4$, and therefore performs worse than EDF-BP. Though it seems that in a general network the EDF-BP should be able to utilize all source-destination paths, it will neither be able to efficiently prioritize packets based on their age, nor discover which paths are more efficient at delivering packets within their deadlines.

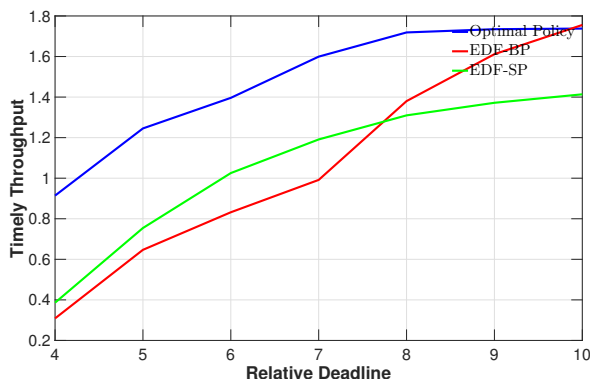


Fig. 7: Timely throughputs of the asymptotically optimal (labeled “Optimal”), EDF-BP and EDF-SP policies for the network in Fig. 4 as the relative deadlines of both flows are increased.

XVIII. CONCLUSIONS AND FUTURE WORK

We have addressed the problem of designing optimal distributed policies that maximize the timely-throughput of multi-hop wireless networks with average nodal power constraints and unreliable links, in which data packets are useful only when they are delivered by their deadline. The key to our results is the observation that if the nodes are subject to average-power constraints, then the optimal solution is decoupled not only along nodes and flows, but also along packets within the same flow at a node. Each packet can be treated exclusively in terms of its time-to-deadline at a node. The decision to transmit a packet is

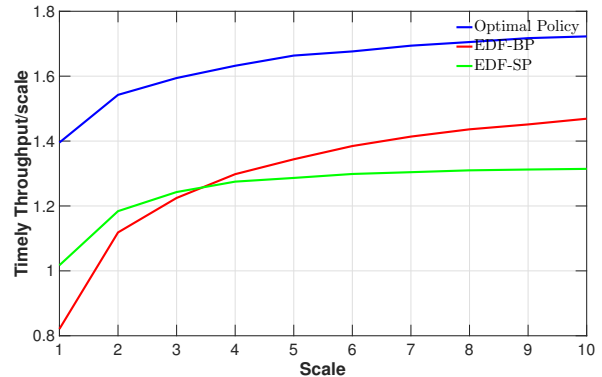


Fig. 8: Timely throughputs of the policies for the network in Fig. 4 as link capacities and arrival rate are scaled. The relative deadlines for both flows are set at 6 time-slots.

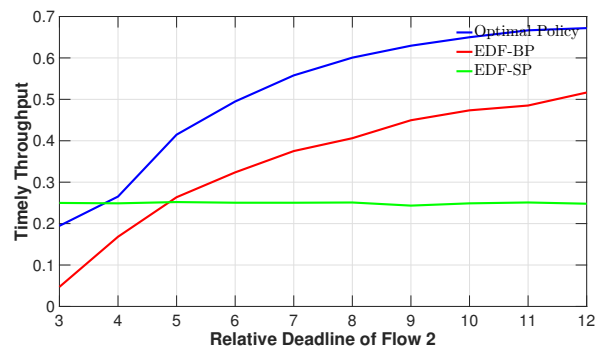


Fig. 9: Timely-throughputs for the network in Fig. 5 as the relative deadlines of flows are increased. Relative deadline of Flow 1 is one more than that of Flow 2.

governed by a “transmission price” that the packet pays at each node, weighed against the reward that it collects at the destination if it reaches it before the deadline expires.

The above policies are highly decentralized; a node’s decisions regarding a packet can be taken solely on the basis of its age and flow classification. The nodes need not share any information such as queue lengths, etc., amongst themselves in order to schedule packets. This approach is notable since obtaining optimal distributed policies for networks has long been considered an intractable problem. Thus, our work fills two important gaps in the existing literature of policies for multi-hop networks a) hard per-packet end-to-end delay guarantees, b) optimal distributed policies.

The traditional approach to scheduling has been to consider the Lagrangian of the fluid model, and interpret the queue lengths as prices. This addresses throughput optimality, but not delay, as one would expect from any fluid model-based analysis. The key to our analysis consists of posing the problem of joint routing- and scheduling packets under deadline constraints over a multi-hop network as an intrinsically stochastic problem involving unreliabilities, and consider its Lagrangian and the Dual. This intrinsically captures variabilities in packet

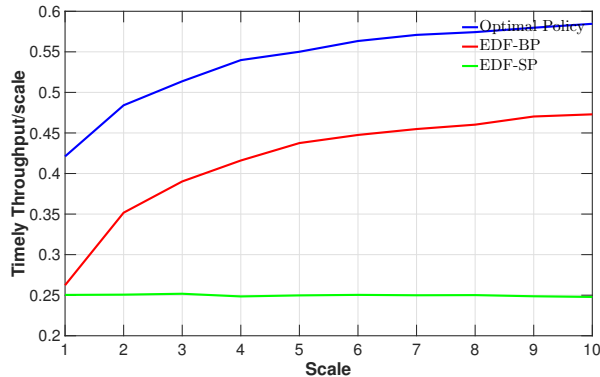


Fig. 10: Timely-throughputs for the network in Figure 5 as link capacities and arrival rates are scaled. Relative deadlines for Flows 1 and 2 are 6 and 5 respectively.

movement which critically affect delays, and allows us to address the timely-throughput optimality of packets that meet hard end-to-end deadlines. The Lagrange multipliers associated with the average power or rate constraints are then the prices paid by a packet to a node for transmitting its packet, rather than queue lengths. This yields a completely decentralized policy, in which decisions are taken by a packet based solely on its age and location in the network, for which the accompanying dynamic programming equations are very tractable. The overall solution is eminently tractable, being completely determined by a linear program with the number of variables equal to the product of the square of the number of nodes, the number of flows and the maximum relative deadline, rather than exponential in problem size.

We also consider the case of peak-power constraints at each node, which may be present in addition to, or as a replacement of, average-power constraints. It is interesting that a minor modification of the optimal policy for the case of average-power constraints is asymptotically optimal as the network capacity is scaled.

This approach of dualizing the stochastic problem has broad ramifications, as has been explored in [55], [56] for problems such as video transmission and energy storage.

This paper has considered only the case of unreliable links, which is of interest in networks consisting of microwave repeaters, networks with directed antennas, or even unreliable wireline links. The case of networks with contention for the medium is addressed in a companion paper.

REFERENCES

- [1] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, Dec 1992.
- [2] X. Lin and N. Shroff, "Joint rate control and scheduling in multihop wireless networks," in *Proceedings of 43rd IEEE Conference on Decision and Control (CDC)*, Dec. 14–17 2004, pp. 1484–1489.
- [3] X. Lin, N. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1452–1463, Aug 2006.
- [4] M. Neely, E. Modiano, and C. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," in *Proceedings of INFOCOM*, vol. 1, March 2003, pp. 745–755.
- [5] P. Whittle, "Restless bandits: Activity allocation in a changing world," *Journal of Applied Probability*, vol. 25, pp. 287–298, 1988.
- [6] R. R. Weber and G. Weiss, "On an Index Policy for Restless Bandits," *Journal of Applied Probability*, vol. 27, no. 3, pp. 637–648, 1990.
- [7] Rahul Singh and P. R. Kumar, "Throughput optimal decentralized scheduling of multi-hop wireless networks with end-to-end deadline constraints: Interfering links," *To be submitted*, 2016.
- [8] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener, "Rethinking information theory for mobile ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 94–101, December 2008.
- [9] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, pp. 1135–1149, 2002.
- [10] Nitin Salodkar, Abhay Karandikar and V. S. Borkar, "A stable on-line algorithm for energy efficient multi-user scheduling," *IEEE Transactions on Mobile Computing*, vol. 9, no. 1, pp. 1391–1406, 2010.
- [11] E. Altman, *Constrained Markov Decision Processes*. Chapman and Hall/CRC, March 1999.
- [12] L. Jiang and J. Walrand, "A Distributed CSMA Algorithm for Throughput and Utility Maximization in Wireless Networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 3, pp. 960–972, June 2010.
- [13] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *The Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.
- [14] A. Eryilmaz, A. Ozdaglar, and E. Modiano, "Polynomial complexity algorithms for full utilization of multi-hop wireless networks," in *Proc. of IEEE INFOCOM*, 2007, pp. 499–507.
- [15] L. Jiang, M. Leconte, J. Ni, R. Srikant, and J. Walrand, "Fast Mixing of Parallel Glauber Dynamics and Low-Delay CSMA Scheduling," *IEEE Transactions on Information Theory*, vol. 58, no. 10, pp. 6541–6555, Dec 2012.
- [16] L. Bui, R. Srikant, and A. Stolyar, "A novel architecture for reduction of delay and queueing structure complexity in the back-pressure algorithm," *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1597–1609, Dec 2011.
- [17] G. Gupta and N. Shroff, "Delay analysis for multi-hop wireless networks," in *Proceedings of Twenty-Eighth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, April 2009, pp. 2356–2364.
- [18] L. Ying, S. Shakkottai, A. Reddy, and S. Liu, "On combining shortest-path and back-pressure routing over multihop wireless networks," *IEEE/ACM Trans. Networking*, vol. 19, no. 3, pp. 841–854, Jun. 2011.
- [19] I-Hong Hou, V. Borkar, and P. R. Kumar, "A theory of QoS for wireless," in *Proceedings of Twenty-Eighth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, April 2009, pp. 486–494.
- [20] I-Hong Hou and P. R. Kumar, "Admission control and scheduling for QoS guarantees for variable-bit-rate applications on wireless channels," in *Proceedings of the Tenth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, May 2009.
- [21] —, "Scheduling heterogeneous real-time traffic over fading wireless channels," in *Proceedings of Twenty-Ninth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, San Diego, CA, March 2010.
- [22] —, "Utility maximization for delay constrained QoS in wireless," in *Proceedings of Twenty-Ninth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, March 2010, pp. 1–9.
- [23] —, "Utility-optimal scheduling in time-varying wireless networks with delay constraints," in *Proceedings of the Eleventh ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2010, pp. 31–40.
- [24] —, "Broadcasting delay-constrained traffic over unreliable wireless links with network coding," in *Proceedings of the Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2011, pp. 33–42.

- [25] I-Hong Hou, Anh Truong, Santanu Chakraborty and P. R. Kumar, "Optimality of periodwise static priority policies in real-time communications," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, 2011, pp. 5047–5051.
- [26] I-Hong Hou and P. R. Kumar, "Scheduling periodic real-time tasks with heterogeneous reward requirements," in *Proceedings of the 32nd IEEE Real-Time Systems Symposium (RTSS)*, 2011, pp. 282–291.
- [27] —, "Real-time communication over unreliable wireless links: a theory and its applications," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 48–59, 2012.
- [28] —, "A survey of recent results on real-time wireless networking," *Proceedings of Real-Time Wireless for Industrial Applications*, 2011.
- [29] —, "Queueing systems with hard delay constraints: A framework and solutions for real-time communication over unreliable wireless channels," *Queueing Systems: Theory and Applications*, vol. 71, pp. 151–177, 2012.
- [30] —, *Packets with Deadlines: A Framework for Real-Time Wireless Networks*, ser. Synthesis Lectures on Communication Networks. Morgan & Claypool Publishers, 2013.
- [31] Rahul Singh, I-Hong Hou and P. R. Kumar, "Pathwise performance of debt based policies for wireless networks with hard delay constraints," in *Proceedings of the 52nd IEEE Conference on Decision and Control (CDC)*, Dec. 10-13 2013, pp. 7838–7843.
- [32] —, "Fluctuation analysis of debt based policies for wireless networks with hard delay constraints," in *Proceedings of Thirty-Third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, April 2014, pp. 2400–2408.
- [33] R. Li and A. Eryilmaz, "Scheduling for end-to-end deadline-constrained traffic with reliability requirements in multi-hop networks," in *Proceedings of Thirtieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, April 2011, pp. 3065–3073.
- [34] Z. Mao, C. Koksai, and N. Shroff, "Optimal online scheduling with arbitrary hard deadlines in multihop communication networks," *Networking, IEEE/ACM Transactions on*, vol. PP, no. 99, pp. 1–1, 2014.
- [35] S. Zhang and V. K. N. Lau, "Resource allocation for OFDMA system with orthogonal relay using rateless code," *IEEE Transactions on Wireless Communications*, vol. 7, Nov. 2008.
- [36] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proceedings of ASILOMAR*, 2010, pp. 1558–1562.
- [37] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *ACM MobiCom*, 2010.
- [38] E. Altman, "Constrained Markov Decision Processes with Total Cost Criteria: Occupation Measures and Primal LP," *Mathematical Methods of Operations Research*, vol. 43, no. 1, pp. 45–72, 1996.
- [39] Alan S. Manne, "Linear programming and sequential decisions," *Management Science*, vol. 6, pp. 259–267, 1960.
- [40] V. S. Borkar, "Control of Markov Chains with Long-Run Average Cost Criterion," in *The IMA Volumes in Mathematics and Its Applications*, W. Fleming and P. Lions, Eds. Springer, 1988, pp. 57–77.
- [41] D. P. Bertsekas, A. E. Ozdaglar, and A. Nedic, *Convex analysis and optimization*, ser. Athena scientific optimization and computation series. Belmont (Mass.): Athena Scientific, 2003.
- [42] S. M. Ross, *Applied Probability Models with Optimization Applications*. San Francisco: Holden-Day, 1970.
- [43] M. Morishima, *Walras' Economics: A Pure Theory of Capital and Money*. Cambridge University Press, 1977.
- [44] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [45] H. Robbins and S. Monro, "A stochastic approximation method," *The Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, Sept. 1951.
- [46] V. S. Borkar, *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge: Cambridge University Press New Delhi, 2008.
- [47] H. J. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*. New York: Springer Verlag, 1997.
- [48] V. S. Borkar, "Stochastic approximation with two time scales," *Systems & Control Letters*, vol. 29, no. 5, pp. 291–294, 1997.
- [49] J. Gittins and D. Jones, "A dynamic allocation index for the sequential design of experiments," in *Progress in Statistics*, J. Gani, Ed. Amsterdam, NL: North-Holland, 1974, pp. 241–266.
- [50] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed Bandit Allocation Indices*. John Wiley & Sons, 2011.
- [51] Persi Diaconis and Y. Zabeil, "Closed form summation for classical distributions: Variations on theme," *Statistical Science*, vol. 6, no. 3, pp. 284–302, 1991.
- [52] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005.
- [53] T. Rappaport, *Wireless Communications: Principles and practice*. New York: Prentice Hall, 2002.
- [54] S. Shakkottai and A. L. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR," in *Teletraffic Engineering in the Internet Era. Proceedings of the International Teletraffic Congress - ITC-17*, ser. Teletraffic Science and Engineering, N. L. d. F. Jorge Moreira de Souza and E. A. de Souza e Silva, Eds. Elsevier, 2001, vol. 4, pp. 793–804.
- [55] Rahul Singh and P. R. Kumar, "Optimizing quality of experience of dynamic video streaming over fading wireless networks," in *Proceedings of 54th IEEE Conference on Decision and Control (CDC)*, Dec. 15-18 2015, pp. 7195–7200.
- [56] —, "On storage and renewables: A theory of sizing and uncertainty," in *2015 IEEE Power & Energy Society General Meeting*, July 26-30 2015, pp. 7195–7200.

Rahul Singh received the B.E. degree in electrical engineering from Indian Institute of Technology, Kanpur, India, in 2009, the M.Sc. degree in Electrical Engineering from University of Notre Dame, South Bend, IN, in 2011, and the Ph.D. degree in electrical and computer engineering from the Department of Electrical and Computer Engineering Texas A&M University, College Station, TX, in 2015.



He is currently a Postdoctoral Associate at the Laboratory for Information Decision Systems (LIDS), Massachusetts Institute of Technology. His research interests include decentralized control of large-scale complex cyberphysical systems, operation of electricity markets with renewable energy, and scheduling of networks serving real time traffic.



P. R. Kumar B. Tech. (IIT Madras, '73), D.Sc. (Washington University, St. Louis, '77), was a faculty member at UMBC (1977-84) and Univ. of Illinois, Urbana-Champaign (1985-2011). He is currently at Texas A&M University. His current research is focused on stochastic systems, energy systems, wireless networks, security, automated transportation, and cyberphysical systems.

He is a member of the US National Academy of Engineering and The World Academy of Sciences. He was awarded a Doctor Honoris Causa by ETH, Zurich. He has received the IEEE Field Award for Control Systems, the Donald P. Eckman Award of the AACC, Fred W. Ellersick Prize of the IEEE Communications Society, the Outstanding Contribution Award of ACM SIGMOBILE, the Infocom Achievement Award, and the SIGMOBILE Test-of-Time Paper Award. He is a Fellow of IEEE and ACM Fellow. He was Leader of the Guest Chair Professor Group on Wireless Communication and Networking at Tsinghua University, is a D. J. Gandhi Distinguished Visiting Professor at IIT Bombay, and an Honorary Professor at IIT Hyderabad. He was awarded the Distinguished Alumnus Award from IIT Madras, the Alumni Achievement Award from Washington Univ., and the Daniel Drucker Eminent Faculty Award from the College of Engineering at the Univ. of Illinois.