



University of Pennsylvania
ScholarlyCommons

Departmental Papers (ESE)

Department of Electrical & Systems Engineering

10-2010

Economy of Spectrum Access in Timy Varying Multichannel Networks

M.H.R. Khouzani

University of Pennsylvania, khouzani@seas.upenn.edu

Saswati Sarkar

University of Pennsylvania, swati@seas.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/ease_papers

 Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

M.H.R. Khouzani and Saswati Sarkar, "Economy of Spectrum Access in Timy Varying Multichannel Networks", . October 2010.

Suggested Citation:

Khouzani, M.H.R. and S. Sarkar. (2010). "Economy of Spectrum Access in Time Varying Multichannel Networks." *IEEE Transactions on Mobile Computing*. Vol. 9(10). pp. 1361-1376.

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/ease_papers/569

For more information, please contact repository@pobox.upenn.edu.

Economy of Spectrum Access in Time Varying Multichannel Networks

Abstract

We consider a wireless network consisting of two classes of potentially mobile users: primary users and secondary users. Primary users license frequency channels and transmit in their respective bands as required. Secondary users resort to unlicensed access of channels that are not used by their primary users. Primaries impose access fees on the secondaries which depend on access durations and may be different for different primary channels and different available communication rates in the channels. The available rates to the secondaries change with time depending on the usage status of the primaries and the random access quality of channels. Secondary users seek to minimize their total access cost subject to stabilizing their queues whenever possible. Our first contribution is to present a dynamic link scheduling policy that attains this objective. The computation time of this policy, however, increases exponentially with the size of the network. We next present an approximate scheduling scheme based on graph partitioning that is distributed and attains arbitrary trade-offs between aggregate access cost and computation times of the schedules, irrespective of the size of the network. Our performance guarantees hold for general arrival and primary usage statistics and multihop networks. Each secondary user is, however, primarily interested in minimizing the cost it incurs, rather than in minimizing the aggregate cost. Thus, it will schedule its transmissions so as to minimize the aggregate cost only if it perceives that the aggregate cost is shared among the users as per a fair cost sharing scheme. Using concepts from cooperative game theory, we develop a rational basis for sharing the aggregate cost among secondary sessions and present a cost sharing mechanism that conforms to the above basis.

Disciplines

Electrical and Computer Engineering | Engineering

Comments

Suggested Citation:

Khouzani, M.H.R. and S. Sarkar. (2010). "Economy of Spectrum Access in Time Varying Multichannel Networks." *IEEE Transactions on Mobile Computing*. Vol. 9(10). pp. 1361-1376.

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Economy of Spectrum Access in Time Varying Multichannel Networks

M.H.R. Khouzani and Saswati Sarkar

Abstract—We consider a wireless network consisting of two classes of potentially mobile users: primary users and secondary users. Primary users license frequency channels and transmit in their respective bands as required. Secondary users resort to unlicensed access of channels that are not used by their primary users. Primaries impose access fees on the secondaries which depend on access durations and may be different for different primary channels and different available communication rates in the channels. The available rates to the secondaries change with time depending on the usage status of the primaries and the random access quality of channels. Secondary users seek to minimize their total access cost subject to stabilizing their queues whenever possible. Our first contribution is to present a dynamic link scheduling policy that attains this objective. The computation time of this policy, however, increases exponentially with the size of the network. We next present an approximate scheduling scheme based on graph partitioning that is distributed and attains arbitrary trade-offs between aggregate access cost and computation times of the schedules, irrespective of the size of the network. Our performance guarantees hold for general arrival and primary usage statistics and multihop networks. Each secondary user is, however, primarily interested in minimizing the cost it incurs, rather than in minimizing the aggregate cost. Thus, it will schedule its transmissions so as to minimize the aggregate cost only if it perceives that the aggregate cost is shared among the users as per a fair cost sharing scheme. Using concepts from cooperative game theory, we develop a rational basis for sharing the aggregate cost among secondary sessions and present a cost sharing mechanism that conforms to the above basis.

Index Terms—Stochastic network optimization, cognitive networks, economy of spectrum access, imperfect scheduling, graph partitioning, cost sharing, Shapley value.

1 INTRODUCTION

IN conventional wireless networks, legacy users license fixed spans of the spectrum. Actual measurements of spectrum usage confirm that since access is fixed to specific frequencies, large portions of potential bandwidth are used only sporadically. This results in poor utilization of the bandwidth, as low as (6 percent) [1]. As the demand for bandwidth grows and industrial, scientific, and medical bands (ISM) become overcrowded, higher utilization of licensed bands becomes imperative. One strategy to overcome the bandwidth underutilization problem, inspired by the major success of unlicensed bands, is spectrum pooling, in which the license owners of the channels (primaries) permit previously specified renters (secondaries) to access their bandwidth during the times or in locations that they themselves do not use it [2].

The temporary underutilized bandwidth at a certain time and location is known as a spectrum hole [3]. The advent of cognitive radio (CR) in the form of software defined radio (SDR) technology enabled wireless devices to continuously scan the spectrum in search of spectrum holes and dynamically adjust their communication parameters at the MAC level to communicate over these frequencies with minimal adverse effect on the primaries [4], [5]. Consequently, the utilization of the bandwidth can improve dramatically and new network capacity of commercial value

can be created out of the already allocated but underutilized spectrum. Commercial network providers can use secondary access to provide wireless service. For example, providers of access networks or mesh networks can allow their users to communicate with access points, base stations, or mesh networks, or directly among each other through secondary access (Fig. 1). On the other hand, the primaries need a financial incentive for this asset sharing. In this regard, each primary may impose access fees on the secondaries which depend on the duration of access and the available communication rate and can be specific to that primary user. Specifically, a single primary user may also offer different service rates at different prices to adjust its own revenue and its own demand. Different primary users may have different (frequency and space) reusability requirements. For instance, radio and TV stations are less dependent on reusability than cellular networks. This can reflect itself in the access costs that they impose on the secondaries: cellular networks may impose higher access fees.

Traffic variation of the primaries, movement of the nodes (secondaries or primaries), and the fluctuations in the quality of the channels lead to variable available service rates with respect to different locations and times for different secondaries. Each secondary session may have a specific traffic demand and needs to transmit data at a certain rate to its respective destination(s). The objective of the network provider(s) is to schedule the access of the users to the time varying available spectrum so as to minimize the aggregate time average cost subject to meeting the traffic demand of the users.

We thus seek to develop scheduling policies that minimize the aggregate time average cost of scheduling subject to meeting the traffic demand of all of the secondaries whenever possible, i.e., supporting the stability region of

• The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, 200 S. 33rd St., Moore 203, Philadelphia, PA 19104. E-mail: {khouzani, swati}@seas.upenn.edu.

Manuscript received 26 Sept. 2008; revised 28 Apr. 2009; accepted 15 Dec. 2009; published online 6 May 2010.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2008-09-0383. Digital Object Identifier no. 10.1109/TMC.2010.90.

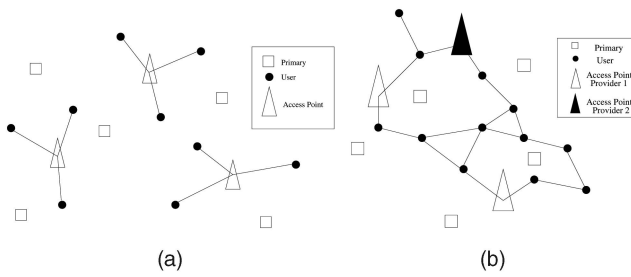


Fig. 1. (a) A sample single-hop cognitive network demonstrating an access network with a single provider. Users communicate with access points using primary channels whenever available (secondary access). Secondary sessions are between a user and an access point. (b) A sample cognitive mesh networks with multiple providers. Users communicate among each other or with the mesh points using secondary access. An MAC layer session is represented by an arc between a user and a mesh point or between two users. An end-to-end session may traverse multiple MAC layer sessions.

the network of the secondaries. Furthermore, the scheduling policy must not require a priori knowledge of the statistics of the arrivals or the channels. Toward this objective, a judicious coordinated decision for each secondary session has to be made about which primary channels to transmit over and with what rate, or whether a session should hold off until a cheaper bandwidth becomes available. This decision should depend on the prices of the available bandwidths and the traffic demand of the sessions.

We illustrate the challenges through a simple example. Suppose that there are two primary channels 1 and 2. Channel 1 is available (idle) in $1/2$ of the time slots and channel 2 in $1/4$. Channels are orthogonal and their availabilities are mutually independent. Channel 1 and channel 2, whenever available, respectively, offer 2 and 1 packet per slot (pps) of communication rates and at \$3/2 and \$1 per pps fees. Suppose that there is a secondary session with arrival rate equal to $1/4$ pps. Now, the question facing the scheduler is whether the secondary should transmit over the expensive channel 1 whenever that is the only available channel, or wait until the cheaper channel 2 becomes available. These decisions should be such that the time average of the incurred cost is minimized, while the backlog of the session does not grow unbounded. Moreover, these decisions should not depend on the knowledge of the arrival and channel availability rates, as a secondary user is not typically aware of them. The scheduling decisions become even more involved in case there is a network of secondaries where interference constraints must also be considered, e.g., when there are multiple interfering sessions with different arrival rates.

Our first contribution is to present a joint channel and rate selection and link scheduling policy that provably attains minimum time average aggregate cost while guaranteeing stability for the secondary arrivals inside their capacity region. Our scheduling policy is dynamic, that is, it does not require any a priori knowledge about the statistics of arrivals, usage of the primaries, quality of the channels, or mobility of the nodes (Section 4). The time required to calculate each schedule, however, grows exponentially with the size of the network. This might become problematic in large networks. Our next contribution is to present a dynamic scheduling policy that provably attains arbitrary trade-offs between 1) time average aggregate cost subject to stability, and

2) complexity and messaging overhead of each schedule in each time slot. The developed trade-offs are independent of the size of the network (Section 5). We discuss in Section 6 how our algorithms can be extended to a multihop network.

The next challenge is to develop a rational basis for sharing the aggregate scheduling cost among the sessions. This goal is motivated by the fact that each session is interested in minimizing its own cost rather than the aggregate cost of the service, and providing a rational basis for cost sharing is, therefore, a prerequisite for motivating the sessions to schedule their transmissions so as to minimize the aggregate cost. A desirable cost sharing mechanism should share the aggregate cost among the sessions in accordance with the amount of stress that each session imposes on the network resources. For instance, a session should naturally pay more if it has a higher traffic demand, or has access to only expensive channels, or is in an area where the congestion is high. We will establish that naive cost sharing solutions, such as splitting the aggregate cost equally among the sessions and charging sessions based on their *direct access cost*, are not desirable. Using tools from cooperative game theory, we present a rational basis for sharing the aggregate cost among sessions and show that this cost sharing mechanism satisfies several intuitively appealing properties (Section 7). The paper is concluded in Section 9 with a discussion of future research directions.

2 RELATED LITERATURE

In [6] and [7], the authors discuss the practical issues of spectrum management and present frameworks that ensure a secondary access that conforms to users' privileges. Kloeck et al. [8] propose a secondary access scheme, based on an auction sequence in a game-theoretical framework, but does not consider the scheduling constraints. Indeed, one of the main objectives of this paper is to bridge the gap between the current research in economical implications of spectrum sharing and effective MAC scheduling policies.

It should be noted that the optimization problem in our paper is different from a utility-based optimization in which the utility is merely a function of the achieved long-term time average service rates when the arrival rates are outside of the stability region [9], [10], [11], etc. Here, the time average scheduling cost depends on specific scheduling decisions in time slots. The latter has been considered in [12] in the context of scheduling rewards and in [13] in the context of energy expenditure. We use the framework of these papers to develop our dynamic optimal scheduling policy. The time required to compute the schedule in each time slot in the policy offered by [13] grows exponentially with the size of the network. In this paper, we proceed to develop a dynamic scheduling policy that: 1) attains a time average cost that is arbitrarily close to the minimum time average cost and 2) has a computation time that is independent of the size of the network. Chaporkar and Sarkar [12] seek a similar goal; however, in that paper, the computation time is reduced at the expense of huge network delays. In contrast, our approximate scheduling policy does not affect the network delay. Also, unlike these two papers which consider iid arrival and channel statistics, we establish our performance guarantees for more general statistical models.

Designing an optimal spectrum assignment for secondaries is considered in [14], [15], [16] among others.

TABLE 1
Table of Important Notations

$\{1 \dots M\}$	set of primary channels
$\{1 \dots N\}$	set of secondary users
$\{1 \dots L\}$	set of links (sessions in a Single-Hop network)
$\{A_l(t)\}$	Arrival to session l during t
Δ	(generic) scheduling policy
Ω	time-average aggregate cost
$c_l(t)$	offered communication rate over link l during t
$i(t)$	scheduling decision during t
\mathcal{I}	set of all feasible scheduling decisions
$\Delta^d(V)$	dynamic scheduling policy in Section-4
$\Delta^{Id}(V)$	Imperfect dynamic scheduling policy in subsection 5.1
$\pi(k, V)$	graph-partitioning scheduling policy in subsection 5.2
$\Delta^*(\lambda)$	stationary scheduling policy in Appendix A

However, in all of these works, spectrum bandwidth is assumed to be free of charge, and thus, other performance metrics including stability are considered. Moreover, Peng and coauthors [14] and [15] assume infinite demand for all secondaries, and Urgaonkar and Neely [16] assume a fixed transmission rate for all of the channels and a single-hop access-point-based network.

Approximate dynamic scheduling policies that achieve a polynomial complexity have been proposed in [17], [18], [19], [20], [21], [22], [23], [24] among others. All of these works focus essentially on approximating the stability region and/or fairness outside of the stability region and consider networks whose topologies do not change with time. Kozat et al. [25] consider the power efficiency problem and present heuristic suboptimal algorithms.

In this paper we develop a dynamic scheduling policy that for a time-varying network attains arbitrary trade-offs between: 1) the time average aggregate cost subject to stability of the secondaries and 2) complexity of the scheduling in each time slot, without affecting the delay.

3 SYSTEM MODEL

3.1 General Framework

A list of major notations is provided in Table 1. Throughout the paper, matrices are represented by bold letters and vectors are specified by vector signs. Also, all of the comparisons between matrices or vectors are considered elementwise. Time is slotted and synchronized. This assumption is justified when clock drifts are negligible at the time scale of control packet transmission; similar assumptions have been made in several papers in this genre (e.g., [10], [11], [12], [13], [14], [16], [19], [24], etc.). Clock synchronization, however, is a challenging problem and an area of active research; addressing the relevant issues is beyond the scope of this paper.

There are M primary channels. Each primary user is the license owner of a unique set of channels and communicates over them whenever needed. There are a total of N wireless secondary users and secondary access points (or base stations) which we will collectively refer to as *nodes*. Let V be the set of the nodes, i.e., $V = \{1, \dots, N\}$. Also, let the set of *links* E be a subset of the ordered triplets (i, j, k) , where $i, j \in V$ and $k \in \{1, \dots, M\}$. The triplet (i, j, m) specifies the link ijm , where i and j are the two *end nodes* and m is the channel of the link ijm . The network of the

secondaries is modeled as the graph $G = (V, E)$. Note that up to M links may exist between two nodes.

Packets may randomly arrive to a secondary node (*source*) to be transmitted to its individual *destination* node. We refer to a source-destination pair as a *session*. Here, we consider single-hop communication among secondaries. Thus, each session can be attributed to a link. The sessions are indexed from 1 to L , where L is the total number of the sessions in the network. Let $A_l(t)$ denote the amount of data arrival during time slot t for session l . The arrivals are according to an aperiodic positive-recurrent Markov process with a countable set of states $\{\vartheta(\tau)\}$, where state $\vartheta(\tau)$ determines the distribution of $\bar{A}(\tau)$.

During each time slot, the network is randomly in one of the finitely many states out of set \mathcal{S} . We assume that the network preserves its state during a time slot and possibly changes its state at the end of a slot. The state of the network embodies information about the scheduling constraints. Specifically, it indicates the availability of a link and the available communication rates over it. This depends on the usage status of the primary users on their respective channels, mobility of the nodes, and fluctuations in the quality of the channels. We assume that the state of the network is also an aperiodic positive-recurrent Markovian process, and thus, has a stationary probability distribution. Let the stationary probability of the state s be denoted by π_s .

We assume that the secondary nodes are capable of sensing the local state of the network at the beginning of each time slot. This assumption follows from the CR capabilities. For each state s , the collective scheduling decision, denoted by i , is confined to be chosen from the finitely many members of the set \mathcal{I}_s . \mathcal{I}_s represents the set of all *valid* scheduling decisions provided that the state of the network is s and depends on the *interference* model as well as the multichannel transmission capability of nodes and the actual deployment of the nodes. We say that two links *interfere* if they cannot be scheduled concurrently. The choice of i along with the state of the network determines the offered transmission rate over each link. Specifically, $c_{lm}(i(t), s(t))$, $i(t) \in \mathcal{I}_{s(t)}$ is the offered transmission rate over link lm during time slot t when the state of the network is $s(t)$ and the scheduling decision is $i(t) \in \mathcal{I}_{s(t)}$. Recall that session l refers to a pair of source-destination nodes say i and j ; thus, link lm is, in fact, a reference to link ijm . The matrix of the offered rates is represented by the $L \times M$ matrix $\mathbf{C}(i, s)$. We assume bounded offered transmission rates. Specifically,

$$c_{max} \triangleq \max_{l,m,s,i \in \mathcal{I}_s} \{c_{lm}(i, s)\} < \infty. \quad (1)$$

Queues evolve according to the following equation:

$$Q_l(t+1) = \left[Q_l(t) - \sum_m c_{lm}(i(t), s(t)) \right]^+ + A_l(t). \quad (2)$$

Let $f_{lm}(t)$ represent the actual flow over link lm during time slot t and let $\mathbf{f}(t)$ be the matrix of the flows over the entire set of links. Mathematically:

$$f_{lm}(t) = \min \left\{ Q_l(t), \sum_m c_{lm}(i(t), s(t)) \right\}.$$

For a network state s , a flow matrix \mathbf{f} is associated with an access cost determined by the scalar nonnegative function

$r(\mathbf{f}, s)$. This is the instantaneous aggregate access cost during time slot t that is incurred by the secondaries by using primary channels. We have allowed the same flow matrix on the network to potentially induce different costs, depending on the state of the network, to preserve its generality. Function r is nondecreasing with respect to each $f_{mn}(t)$ and is bounded by r_{max} . As a specific example, the cost function can be *additive*, i.e.,

$$r(\mathbf{f}, s) = \sum_{l=1}^L \sum_m r_{lm}(f_{lm}, s), \quad (3)$$

where $r_{lm}(f_{lm}, s)$ is the cost incurred by flow f_{lm} of session l on channel m when the network state is s .

A scheduling policy, denoted by Δ , is a rule that determines the scheduling decision $i \in \mathcal{I}_{s(t)}$ in each time slot t .

Definition 1. The aggregate time average cost of a scheduling policy Δ , denoted by Ω^Δ , is defined as follows:

$$\Omega^\Delta = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t r(\mathbf{f}(\tau), s(\tau)).$$

Definition 2. We call a queue stable if

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Q(\tau) < \infty \quad w.p.1.$$

A network is stable if all of the individual queues are stable. Throughout the paper, the term stability is used to refer to this definition of stability. The closure of the set of all arrival rate vectors, $\vec{\lambda}$, for which there exists a scheduling policy that makes the network stable is called the stability region of the network and is denoted by Λ . Also, let $\text{Int}(\Lambda)$ denote the interior of the set Λ .

Definition 3. Δ is ϵ -optimal if and only if for any $\vec{\lambda} \in \text{Int}(\Lambda)$:

1. the network is stable and
2. $\Omega^\Delta(\vec{\lambda}) \leq \Omega_{min}(\vec{\lambda}) + \epsilon \quad w.p.1.$

Let $H(\tau) \triangleq (\vartheta(\tau), s(\tau), \vec{Q}(\tau))$. Also, let $\mathcal{H}(\tau)$ represent the σ -field generated by $H(\tau)$. Since by assumption, the arrivals are due to an aperiodic positive-recurrent (hence, ergodic) Markov chain, for any τ and $\mathcal{H}(\tau - T)$, we have $\lim_{T \rightarrow \infty} E[\vec{A}(\tau) | \mathcal{H}(\tau - T)] = \vec{\lambda}$. Hence,

$$\forall \rho > 0, \quad \exists T_A(\rho) < \infty, \quad \text{s.t.} \quad \forall T > T_A(\rho), \quad \forall \tau: \quad E[\vec{A}(\tau) | \mathcal{H}(\tau - T)] \leq \vec{\lambda} + \rho \vec{1}. \quad (4)$$

We assume that

$$A_{max}^2 \triangleq \max_{l \in \{1, \dots, L\}} EA_l^2(\tau) < \infty. \quad (5)$$

Similar to (4), for the network states, we have

$$\lim_{T \rightarrow \infty} E[1_{s(\tau)=s} | \mathcal{H}(\tau - T)] = \pi_s, \quad (6)$$

where 1 represents an indicator function.

3.2 A Special Case

Here, we elucidate the above framework by constructing \mathcal{I}_s for an important special case. We will focus on this special network setup in Section 5.2.

We assume that all of the nodes are deployed on a 2D plane and in the first quadrant. Note that the latter assumption is made without loss of generality since the choice of the origin and the coordinates is arbitrary. Let two nodes be able to communicate if their distance is less than D , which we assume to be invariant for any channel at any transmission rate. Also, assume that the M primary channels are orthogonal to each other. Recall that each link is represented by a triplet ijm , where i, j designate the end points and m indicates the channel of the link. During a time slot t , link ijm is available in the network graph if first, its two end nodes are within distance D of each other and second, there is no primary user within distance D of nodes i or j which is communicating over channel m . At time slot t , the state of the network $s(t)$ is interpreted as the network graph of the available links in that time slot. The usage status of the primaries is random and may change over time slots. Also, both primary and secondary users may be mobile. Thus, the network graph varies over time.

Now, there can be two different possibilities with respect to the multichannel transmission capability of the nodes:

1. Each node can communicate over at most one channel at a time. In this case, two links interfere whenever a) they are of the same channel, and an end node of one of them is within distance D of an end node of the other, or b) the two links have at least one common end node.
2. Nodes can communicate over different channels simultaneously and with different nodes. In this case, two links interfere only whenever they are of the same channel, and an end node of one of the links is within distance D of an end node of the other.

We have considered this model since similar transmission constraints apply in 802.11 protocols. An *independent set* is a subset of links in which no two links interfere. Let $\mathcal{X}(s)$ be the set of all of independent sets given an instance of the network graph, which is determined through network state s . First, consider the case that each primary channel, whenever available, offers a fixed transmission rate. Now given the state of the network is $s(t)$, the scheduling decision at time slot t is translated into selecting an independent set $X(t)$ from $\mathcal{X}(s(t))$. Thus, we can write $\mathcal{I}_s = \mathcal{X}(s(t))$. More generally, each primary channel, whenever available, may offer different transmission rates per time slot at different prices. Let ς_m represent the set of all available offered transmission rates for channel m . Here, given the state of the network is $s(t)$, the scheduling decision at time slot t is translated into selecting an independent set $X(t)$ from $\mathcal{X}(s(t))$ and also specifying the transmission rate on each of the scheduled links. Thus, a mathematical representation of \mathcal{I}_s is as follows:

$$\mathcal{I}_s = \left\{ (X, \Xi) : X \in \mathcal{X}_s, \Xi : X \rightarrow \prod_{(lm) \in X} \varsigma_m \right\},$$

where Ξ as defined is a vector function from the set of the links in the independent set X to the set of available rates on each of the selected links.

Here, we explain the calculation of the instantaneous cost function in the above framework. Let $\varrho(c, m)$ denote the fee per time slot of using transmission rate c over channel m . Thus, the instantaneous access cost during time slot t is

$$r(\mathbf{C}(i(t), s(t)), s(t)) = \sum_{(lm) \in X(t)} \varrho(f(lm), m),$$

where $f(lm)$ denotes the actual flow on link lm . Note that this is an example of an additive cost model we briefly alluded to in (3).

4 DYNAMIC SCHEDULING FOR SINGLE-HOP NETWORKS AND PERFORMANCE GUARANTEES

In this section, we present an ϵ -optimal dynamic scheduling policy for the general framework presented in Section 3. This scheduling policy is dynamic, in that case, it does not require any a priori knowledge of the statistics of the arrivals or the channels. The presented dynamic policy features a controllable parameter V which provides a trade-off between the time average scheduling cost and network delay.

4.1 Dynamic Scheduling $\Delta^d(V)$

At the beginning of each time slot, given the state is s , the scheduling policy $\Delta^d(V)$ chooses $i(t)$ that is the solution of the following optimization problem:

$$\max_{i \in \mathcal{I}_s} \left[\sum_l \left(Q_l(t) \sum_m c_{lm}(i, s) \right) - V r(\mathbf{C}(i, s), s) \right], \quad (7)$$

where $r(\mathbf{C}(i, s), s)$ is the instantaneous aggregate cost imposed on the secondaries if the state of the network is s and the flows in the network are equal to $\mathbf{C}(i, s)$.

Intuitively, the backlogs function as feedbacks; the scheduling policy $\Delta^d(V)$ assigns the schedule in accordance with the trade-off between stability and the scheduling cost. As some queue backlogs build up, the scheduling decision favors serving them and the effect of the cost becomes less significant. A larger V tunes the scheduler to favor the cost of each schedule at the expense of larger delays. Theorem 1 formalizes the performance guarantees of this scheduler.

Theorem 1. *For any $\vec{\lambda} \in \text{Int}(\Lambda)$ and for all $V \geq 0$, dynamic policy $\Delta^d(V)$ stabilizes the system. Specifically,*

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l Q_l(\tau) \leq \frac{W + V r_{\max}}{\theta_{\max}},$$

where $\theta_{\max} \triangleq \alpha \arg \sup_{\theta > 0} \vec{\lambda} + \theta \vec{1} \in \text{Int}(\Lambda)$ for some $\alpha < 1$ and $W = L(\hat{T} + 1/2)(A_{\max}^2 + c_{\max}^2)$, where \hat{T} is determined through the proof.

Moreover, $\forall \epsilon > 0, \exists \hat{V} > 0$ such that for every $V \geq \hat{V}$, $\Delta^d(V)$ is ϵ -optimal.

Proof. Proof in Appendix A. \square

It becomes clear from the proof that the only reliance on the Markovian assumption for the arrivals and network states processes is through properties in (4) and (6), respectively. Specifically, (4) is used in (23), and (6) is used to develop (28) and (31). Since (4) and (6) are satisfied for *any* stationary and ergodic processes, our performance guarantees hold for any stationary and ergodic arrivals and networks states processes that satisfy (1) and (5).

1. Here, $\alpha < 1$ is used just to avoid the singularities on the border of the stability region.

5 IMPERFECT SCHEDULING AND COMPUTATION SIMPLIFICATION

In the previous section, we presented an ϵ -optimal dynamic scheduling policy for our general framework. Here, we start by applying this dynamic scheduling policy to the important special cases of networks described in Section 3.2. We discuss the issues that arise about the computation time of the algorithm and argue that determining each schedule is an NP-hard problem. In subsequent sections, we take steps to tackle this issue. In Section 5.1, we present a useful lemma for the general framework which enables us to develop approximate scheduling policies that bear less burden of computation for each schedule. In Section 5.2, we again turn our attention to the special network setup of Section 3.2. Inspired by the result of the lemma, we propose a dynamic scheduling policy and establish that for appropriate choice of parameters, our policy is ϵ -optimal. As we next argue, the computation time of each schedule in our algorithm does not depend on the size of the network.

The scheduling policy $\Delta^d(V)$ must solve the optimization problem (7) in every time slot. The computation time of solving this combinatorial optimization can grow exponentially in the size of the network, since for a general cost function and scheduling constraints, there is no other way than to exhaust all of the possible $i \in \mathcal{I}_{s(t)}$. Here, we explain the structure of the above optimization problem for the special case network described in Section 3.2. In the beginning of each time slot, the state of the network is observed which specifies the network of available links. Note that by assumption, altering the transmission rates does not affect the interference constraints of the network. Thus, in the first step, each link lm in the network graph determines its optimum weight $w_{lm}^*(t)$ by individually solving the following optimization and finding the best candidate transmission rate over its channel (c^*):

$$w_{lm}^*(t) = \max_{c \in \mathcal{C}_{lm}} [Q_l(t)c - V \varrho(c, m)], \quad (8)$$

and $c_{lm}^*(t) = \arg \max_{c \in \mathcal{C}_{lm}} [Q_l(t)c - V \varrho(c, m)]$. Note that $c = 0$, i.e., refraining from transmission over a link, can be a valid choice too. In the next step, a maximum-weight-independent set of links using w^* s as weights is found, which is the solution of the following NP-hard [26] combinatorial optimization problem:

$$X^*(t) = \arg \max_{X \in \mathcal{X}_s} \sum_{lm \in X} w_{lm}^*(t). \quad (9)$$

Links in $X^*(t)$ are scheduled at rates equal to their respective c_{lm}^* s. Indeed, as we observe, even for a simple pairwise and symmetric interference model and when the cost function is additive over the scheduled links, the problem is NP-hard, and thus, a solution is impractical to compute in a large network.

5.1 General Framework

In this section, we prove a key lemma that formally specifies the effect of a multiplicative suboptimal decision on the performance guarantees of the scheduling policy. Lemma 1 has an important implication: It asserts that if in each time slot, the optimization problem given in (7) can be approximated arbitrarily closely in the *expected value* sense, then both the stability region and the time average cost of the resulting policy are arbitrarily close to their optimum

values. The extent of generality presented in inequality (10) as the necessary condition of the lemma is quite promising. Next section provides an example in which inequality (10) is satisfied. In fact, as we will see, this inequality is achieved pathwise with respect to $s(t)$ and $\bar{Q}(t)$, and the approximation is introduced by using graph partitioning (Section 5.2). For simplicity of exposition, throughout this section, we assume that the arrivals and network states are iid. Indeed, these results are easily extendable to Markovian model using similar arguments as in the proof of Theorem 1.

Lemma 1. Consider a scheduling policy $\Delta^{Id}(V)$ that in each time slot τ chooses a (suboptimal) scheduling decision $i(\tau) \in \mathcal{I}_{s(\tau)}$ that satisfies the following:

$$E \left[\left(\sum_l Q_l \sum_m c_{lm}(i^{\Delta^{Id}(V)}, s) - Vr(C(i^{\Delta^{Id}(V)}, s), s) \right) \right] \geq (1-\phi) \max_{i \in \mathcal{I}_s} E \left[\left(\sum_l Q_l \sum_m c_{lm}(i, s) - Vr(C(i, s), s) \right) \right], \quad (10)$$

for some constants $0 \leq \phi < 1$ (the τ s are omitted for brevity). Then, for every $V \geq 0$, $\Delta^{Id}(V)$ stabilizes the network for all $\bar{\lambda}$ such that $\bar{\lambda}/(1-\phi) \in \text{Int}(\Lambda)$. Specifically,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l Q_l(\tau) \leq \frac{\hat{W} + V(1-\phi)r_{max}}{\hat{\theta}_{max}}, \quad (11)$$

in which $\hat{W} \triangleq \frac{1}{2} L(A_{max}^2 + c_{max}^2)$ and $\hat{\theta}_{max} \triangleq \alpha \arg \sup_{\theta > 0} (\bar{\lambda} + \theta \bar{\mathbf{I}})/(1-\phi) \in \text{Int}(\Lambda)$ for some $\alpha < 1$.²

Moreover, $\forall \epsilon > 0$ and for every $\bar{\lambda}$ such that $\bar{\lambda}/(1-\phi) \in \text{Int}(\Lambda)$, $\exists \hat{V} > 0$ such that for every $V \geq \hat{V}$ we have

$$\Omega^{\Delta^{Id}(V)}(\bar{\lambda}) \leq (1-\phi)\Omega_{min}(\bar{\lambda}) + \epsilon/2 + \Gamma(\phi),$$

where $\Gamma(\phi) = \phi(1-\phi)\sqrt{A_{max}^2}r_{max}/\hat{\theta}_{max}$, a constant independent of the size of the network that goes to zero as $\phi \rightarrow 0$.

Proof. Proof in Appendix B. \square

Note that Lin and Shroff [23] also consider multiplicative approximate schedulings. However, in that paper, only the stability region and fairness outside of the stability region are addressed for a fixed topology network. Here, we showed that both the stability region and the time average cost of the scheduling subject to stability can indeed be approximated arbitrarily closely by multiplicative suboptimal schedulings in a wireless network with general time-varying topology. In a similar problem, Chaporkar and Sarkar [12] assume a fixed topology network and propose an additive approximate schedulings policy. The proposed algorithm achieves lower complexity by using randomization without discussing the effect on the network delay. However, simulation results [27] reveal that randomized policies impose large delays on the network. On the contrary, the upperbound on the sum of the queue backlogs under our approximate scheduling policy is unaffected provided that we maintain the same distance from the border of the stability region (inequality (11)).

5.2 Graph Partitioning

Here, we design a scheduling policy that: 1) for any given $\epsilon > 0$, stabilizes the network for every $\bar{\lambda}$ such that

2. Here, $\alpha < 1$ is used just to avoid the singularities on the border of the stability region. Also, refer to footnote 1 on page 1365.

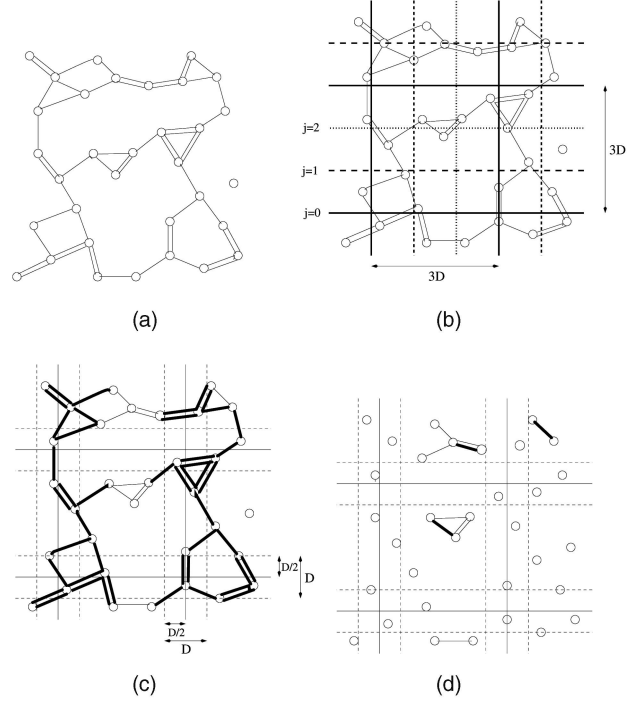


Fig. 2. Step by step illustration of the $\pi(k, V)$ scheduling policy. Here, $k = 3$ and there are two channels (i.e., $m = 2$). (a) Geometric graph of the available links during time slot t . (b) Depiction of the three grids associated with $j = 0, 1, 2$. Links choose each grid with probability $1/3$. (c) Here, the grid for $j = 0$ is selected. Bold links construct the set L^0 and the rest of the links comprise $G^{(0)}$. Each link in $G^{(0)}$ calculates its optimal weight and finds its best transmission rate over its channel (12), (13). (d) Then, independently in each component of $G^{(0)}$, a maximum-weight-independent set of links is scheduled at each link's individual calculated best transmission rate.

$\bar{\lambda}/(1-\epsilon) \in \text{Int}(\Lambda)$, 2) the cost it incurs is at most ϵ more than the minimum time average scheduling cost, and 3) has a computational complexity that is independent of the size of the network.

Consider the same network setup described as the special case in Section 3.2. We further assume that each node is aware of its own coordinates. We now present the scheduling policy $\pi(k, V)$, where k and V are control parameters. An illustration of the algorithm is provided in Fig. 2.

Scheduling policy $\pi(k, V)$. Consider k different grids: each grid consists of a series of horizontal and vertical lines, respectively, parallel to x - and y -axes. The distance between any two consecutive vertical or horizontal lines is kD . Each grid is specified by its first horizontal and vertical lines. The first horizontal and vertical lines of the j th grid are $x = jD$ and $y = jD$, respectively, for $j = 0, \dots, k-1$. Define $L^{(j)}$ to be the set of links for which at least one end point is within a distance $D/2$ of a vertical or a horizontal line of the j th grid. Let $G^{(j)}$ represent the remainder of the graph after removing all of the links in $L^{(j)}$. Note that $G^{(j)}$ comprises a series of decoupled subsets of links such that the links in a component do not interfere with links in any other component. At the beginning of each time slot, scheduling policy $\pi(k, V)$ performs the following:

1. Every link selects a $j \in 0, \dots, k-1$ with probability $1/k$. Links share the same random seed in their pseudorandom number generators, and hence, will

select the same number. Upon selection of j , links $L^{(j)}$ are removed.

2. Each link lm in the graph $G^{(j)}$ finds its optimum weight $w_{lm}^*(t)$ by individually solving the following optimization and finding the best candidate transmission rate over its channel:

$$w_{lm}^*(t) = \max_{c \in \zeta_m} [Q_l(t)c - V\varrho(c, m)], \quad (12)$$

$$c_{lm}^*(t) = \arg \max_{c \in \zeta_m} [Q_l(t)c - V\varrho(c, m)]. \quad (13)$$

3. Within each component of $G^{(j)}$, the maximum-weight-independent set is scheduled using the calculated optimal weights. Links in each selected independent set are scheduled at rates equal to their respective c^* s.

Here, we explain the intuition how this algorithm yields a controllable trade-off between computational complexity and optimality. At each time slot, a set of the links in the network of the secondaries is deactivated such that the remainder is a set of decoupled components where the scheduling decision in each component is performed independently of (and thus, in parallel with) other components. The complexity and messaging overhead of the scheduling decisions depend only on the size of the largest component. However, in order to construct smaller components, more links need to be deactivated, which implies that the scheduling decision deviates more from the optimal. We formally present these trade-offs and establish bounds for suboptimality versus complexity.

5.2.1 Performance Guarantees

Theorem 2. For every $V \geq 0$, scheduling policies $\pi(k, V)$ stabilize the network for all $\vec{\lambda}$ such that $\vec{\lambda}/(1 - \beta/k) \in \text{Int}(\Lambda)$, where constant $\beta = 120$ for case 1 and $\beta = 96$ for case 2, both described in Section 3.2. Specifically,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l Q_l(\tau) \leq \frac{\hat{W} + V(1 - \beta/k)r_{\max}}{\hat{\theta}_{\max}}, \quad (14)$$

in which $\hat{W} \triangleq \frac{1}{2}L(A_{\max}^2 + c_{\max}^2)$ and $\hat{\theta}_{\max} \triangleq \alpha \arg \sup_{\theta > 0} (\vec{\lambda} + \theta \vec{1})/(1 - \beta/k) \in \text{Int}(\Lambda)$ for some $\alpha < 1$.

Moreover, $\forall \epsilon > 0$ and for every $\vec{\lambda}$ such that $\vec{\lambda}/(1 - \beta/k) \in \text{Int}(\Lambda)$, $\exists \hat{V} > 0$ such that for every $V \geq \hat{V}$, we have

$$\Omega^{\pi(k, V)}(\vec{\lambda}) \leq (1 - \beta/k)\Omega_{\min}(\vec{\lambda}) + \epsilon/2 + \hat{\Gamma}(k),$$

where $\hat{\Gamma}(k)$ is achieved constant independent of the size of the network and that goes to zero as k goes to infinity.

Proof. Proof in Appendix C. \square

Theorem 2 and Lemma 1 imply that for any given $\epsilon > 0$, $\pi(k, V)$ is ϵ -optimal for a large enough V and k . Our analysis in Section 5.2.2 shows that the complexity of $\pi(k, V)$ depends on k and the maximum degree of a node, but is indeed independent of the size of the network.

5.2.2 Discussion of Complexity

Finding the maximum-weight-independent set in different components can be performed in parallel. Hence, the computation time depends only on the size of the

components (and not on the size of the entire network). Each component in $G^{(j)}$ has $O(D_G^2 k^2)$ links of the same channel [24], where D_G is the maximum degree of that channel in the network graph. In case 2, the maximum-weight-independent set problem can be broken into M decoupled problems for each channel, and thus, can potentially be performed in parallel. Thus, the complexity is the same as when there were only a single channel, which is $O(D_G^2 k^2)^{O(k^2)}$ [24] if each node has M processors, and is $MO(D_G^2 k^2)^{O(k^2)}$ if each node has only one processor. For case 1, the maximum-weight-independent set problem must be considered for all frequencies jointly, and thus, the computation time is $O(MD_G^2 k^2)^{O(k^2)}$.

6 MULTIHOP NETWORKS

Assume that there are N secondary nodes and there are M different channels. As before, we represent each link by a triplet xyz which shows a link from node x to node y over channel z . Assume now that data can be routed over multihop paths. We refer to all of the data that originates from a particular node and is destined for another particular node as a commodity (or a class). Hence, each commodity pertains to a certain source-destination pair. We use (κ) to represent commodity κ . Assume that there are K different commodities in the network. We define $A_n^{(\kappa)}$ to be the amount of data of commodity κ exogenously entering node n during time slot t . Let $Q_n^{(\kappa)}(t)$ be the amount of data of commodity κ buffered at node n at time t . $f_{xyz}^{(\kappa)}(t)$ is the flow of commodity κ over link xyz during time slot t , which is the amount of data served from $Q_n^{(\kappa)}(t)$ over link xyz during time slot t . $\mathbf{f}(t)$ is the matrix of such flows during time slot t . For each network state s , the scheduler can select j from a set of scheduling decisions \mathcal{J}_s , each of which corresponds to a specific offered transmission rate. j also indicates that on each link, which commodity is the bandwidth allocated to.

Let $c_{xyz}^{(\kappa)}(j(t), s(t))$ represent the offered transmission rate over link xyz to the data of commodity κ at node x during time slot t . $\mathbf{C}(j(t), s(t))$ is the matrix of the offered transmission rates during time slot t . As before we assume that the cost function also depends on the state of the network and we represent the cost function by $r(\mathbf{f}(t), s(t))$.

The network is called stable if and only if each of queues $Q_n^{(\kappa)}$ is stable. The closure of the set of all $\vec{\lambda}$ for which there exists a scheduling policy that stabilizes the network is called the (multihop) stability region and is denoted by Λ . For simplicity of the analysis, we assume that the arrivals and network states are all iid. The definition of the throughput and the ϵ -optimal scheduling is the same as was defined for the single-hop network. We now present a dynamic joint routing and scheduling policy which we prove to be ϵ -optimal.

Multihop dynamic scheduling Θ^d .

1. At the beginning of each time slot and for each link ijk , find the commodity $(\kappa) \in \{1 \dots K\}$ that maximizes $Q_i^{(\kappa)}(t) - Q_j^{(\kappa)}(t)$. Call it $(\kappa_{ijk}^*(t))$ and define

$$W_{ijk}^*(t) \triangleq Q_i^{(\kappa_{ijk}^*(t))}(t) - Q_j^{(\kappa_{ijk}^*(t))}(t).$$

Note that both $(\kappa_{ijk}^*(t))$ and $W_{ijk}^*(t)$ do not change for a fixed i, j and different ks .

2. Observe the state of the channel $s(t)$. Out of the set \mathcal{I}_s , select $i(t) = i$ which solves the following optimization:

$$\max_{i \in \mathcal{I}_s} \left[\sum_{a,b,m} W_{abm}^* c_{abm}(i, s) - Vr(\mathbf{C}(i, s), s) \right].$$

ϵ -optimality of the above scheduling policy can be obtained as a special case of Theorem (3) by taking $\phi = 0$.

We now consider the issue of imperfect scheduling. As before, we first consider a general framework and state a useful lemma, and then, focus on our special network setup and present a scheduling that, given $\epsilon > 0$, stabilizes the network for every $\bar{\lambda}$ such that $\bar{\lambda}/(1 - \epsilon) \in \text{Int}(\Lambda)$ incurs a cost that is at most ϵ more than the minimum time average scheduling cost for any given $\epsilon > 0$ and its complexity is independent of the size of the network.

6.1 General Framework

Lemma 2. Consider a scheduling policy Θ^{Id} that in each time slot τ chooses a (suboptimal) scheduling decision $i(\tau) \in \mathcal{I}_{s(\tau)}$ that satisfies the following:

$$E \left[\sum_{a,b,m} \sum_{(\kappa)} c_{abm}^{(\kappa)} [Q_a^{(\kappa)} - Q_b^{(\kappa)}] - Vr(\mathbf{C}(j, s), s) \right] \geq (1 - \phi) \max E \left[\sum_{a,b,m} \sum_{(\kappa)} c_{abm}^{(\kappa)} [Q_a^{(\kappa)} - Q_b^{(\kappa)}] - Vr(\mathbf{C}(j, s), s) \right],$$

for some constants $0 \leq \phi < 1$ (τ s are omitted for brevity). Then, for every $V \geq 0$, Θ^{Id} stabilizes the network for all $\bar{\lambda}$ such that $\bar{\lambda}/(1 - \phi) \in \text{Int}(\Lambda)$. Specifically,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_{n,(\kappa)} Q_n^{(\kappa)}(\tau) \leq \frac{\hat{W} + V(1 - \phi)r_{max}}{\hat{\theta}_{max}},$$

in which $\hat{W} \triangleq \frac{1}{2} NK[(A_{max} + c_{max}^{in})^2 + (c_{max}^{out})^2]$, where $A_{max}^2 \triangleq \max_{n,(\kappa)} E(A_n^\kappa)^2$ and

$$c_{max}^{out} \triangleq \max_{\{n,s,j \in \mathcal{J}_s\}} \sum_{b,m} c_{nbm}(j, s)$$

and $c_{max}^{in} \triangleq \max_{n,s,j \in \mathcal{J}_s} \sum_{a,m} c_{anm}(j, s)$. Also, $\hat{\theta}_{max} \triangleq \alpha \arg \sup_{\theta > 0} (\bar{\lambda} + \theta \bar{\mathbf{1}})/(1 - \phi) \in \text{Int}(\Lambda)$ for some $\alpha < 1$.³

Moreover, $\forall \epsilon > 0$ and for every $\bar{\lambda}$ such that $\bar{\lambda}/(1 - \phi) \in \text{Int}(\Lambda)$, $\exists \hat{V} > 0$ such that for every $V \geq \hat{V}$, we have

$$\Omega^{\Theta^{Id}}(\bar{\lambda}) \leq (1 - \phi)\Omega_{min}(\bar{\lambda}) + \epsilon/2 + \Gamma(\phi),$$

where $\Gamma(\phi)$ is a constant independent of the size of the network and that goes to zero as $\phi \rightarrow 0$.

6.2 Graph Partitioning

The setup is the same as in Section 5. The scheduling is quite similar. The only difference is in the last two steps:

Scheduling Policy $\pi(k, V)$ and Multihop version.

1. Every link selects a $j \in 0, \dots, k - 1$ with probability $1/k$. Links share the same random *seed* in their

3. Here, $\alpha < 1$ is used just to avoid the singularities on the border of the stability region. Also, refer to footnote 1 on page 1365.

pseudorandom number generators, and hence, will select the same number. Upon selection of j , links $L^{(j)}$ are removed.

2.

- a. First, each link xyz specifies its candidate commodity by individually solving for $\kappa_{xyz}^*(t)$ and its respective best *backpressure* B_{xyz}^* :

$$B_{xyz}^*(t) = \max_{(\kappa)} [Q_x^{(\kappa)}(t) - Q_y^{(\kappa)}(t)],$$

$$c_{xyz}^*(t) = \arg \max_{(\kappa)} [Q_x^{(\kappa)}(t) - Q_y^{(\kappa)}(t)].$$

- b. Then, each link lm in the graph $G^{(j)}$ finds its optimum weight $w_{lm}^*(t)$ by individually solving the following optimization and finding the best candidate transmission rate over its channel:

$$w_{xyz}^*(t) = \max_{c \in \mathcal{C}_{lm}} [B_{xyz}^*(t)c - V\rho(c, m)],$$

$$\mu_{xyz}^* = \arg \max_{c \in \mathcal{C}_{lm}} [B_{xyz}^*(t)c - V\rho(c, m)].$$

3. Within each component of $G^{(j)}$, the maximum-weight-independent set is found using the calculated optimal weights. A link xyz in each selected independent set is scheduled for data of commodity $\kappa_{xyz}^*(t)$ at rate $\mu_{xyz}^*(t)$.

Theorem 3. For every $V \geq 0$, scheduling policies $\pi(k, V)$ stabilize the network for all $\bar{\lambda}$ such that $\bar{\lambda}/(1 - \beta/k) \in \text{Int}(\Lambda)$, where constant β is the same constant as in the single-hop case. Specifically,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l Q_l(\tau) \leq \frac{W + V(1 - \beta/k)r_{max}}{\hat{\theta}_{max}},$$

in which $\hat{W} \triangleq NK(A_{max} + c_{max}^{in})^2 + c_{max}^{out}$. Also, $\hat{\theta}_{max} \triangleq \alpha \arg \sup_{\theta > 0} (\bar{\lambda} + \theta \bar{\mathbf{1}})/(1 - \beta/k) \in \text{Int}(\Lambda)$ for some $\alpha < 1$.

Moreover, $\forall \epsilon > 0$ and for every $\bar{\lambda}$ such that $\bar{\lambda}/(1 - \beta/k) \in \text{Int}(\Lambda)$, $\exists \hat{V} > 0$ such that for every $V \geq \hat{V}$, we have

$$\Omega^{\pi(k,V)}(\bar{\lambda}) \leq (1 - \beta/k)\Omega_{min}(\bar{\lambda}) + \epsilon/2 + \hat{\Gamma}(k),$$

where $\hat{\Gamma}(k) = \phi \sqrt{A_{max}^2 r_{max}} / \hat{\theta}_{max}$, a constant independent of the size of the network and that goes to zero as k goes to infinity.

The proofs of Lemma 2 and Theorem 3 follow by using similar techniques used for the single-hop network model, and thus, are included in our technical report [28].

7 COST SHARING AND SHAPLEY VALUE

We have developed ϵ -optimal dynamic scheduling policies that arbitrarily closely approximate the aggregate time average cost of the scheduling subject to supporting the stability region of the network. We now focus on developing a rational basis for sharing the aggregate cost between: 1) the different sessions, or 2) the different providers in case there are more than one providing service. Our investigation is motivated by the fact that each user (provider or session) is interested in minimizing its own cost rather than the aggregate cost of the service, and providing a rational basis for cost sharing is, therefore, a prerequisite for motivating

them to schedule their transmissions so as to minimize the aggregate cost. We develop this rational basis using principles from cooperative game theory, more specifically, the notion of *Shapley value*, which satisfies several intuitively appealing properties for a rational cost sharing solution. The framework we present can easily consider both the above cases, but has been stated for only case 1): The aggregate time average cost is shared among the sessions.

We first point out the deficiencies of two naive cost sharing mechanisms: 1) *equal splitting* and 2) *direct cost*. As the name suggests, *equal splitting* splits the cost equally among all sessions. The second, *direct cost*, assumes that the cost structure is additive (3) and charges session i the time average of the additive cost (i.e., its *direct cost*) it imposes on the system. Specifically, if session l uses flow f_{lm} over channel m for $\gamma_{lm,s}$ fraction of time when the system is in state s , then it incurs a cost of $\sum_s \pi_s \sum_m \gamma_{lm,s} r_{lm}(f_{lm}, s)$. Intuitively, the cost sharing mechanism should be such that the cost incurred by a session depends on 1) its traffic demand, 2) the access fees imposed by the primary channels available to that session, and 3) the traffic demand of the sessions which interfere with it. We demonstrate, using simple examples, that the above naive cost sharing mechanisms do not consider the above.

Assume that there are only two interfering sessions with arrival rates equal to 1 and 2 pps, respectively. The aggregate time average scheduling cost is \$3 per slot (ps) where session 1 is responsible for \$1 ps and session 2 for \$2 ps. But equal splitting will charge each session 1.5 ps. Next, consider a network consisting of two noninterfering sessions with identical arrival rates equal to 1 pps. Assume that there are two available channels. Session 1 has access to the *expensive* channel 1 which offers 1 pps transmission rate and imposes \$1, i.e., \$1 pps. Session 2 has access to *cheap* channel 2 which offers 4 pps and imposes \$2 per instance of access, i.e., \$0.5 pps. Thus, the aggregate time average cost is \$1.5 ps, where sessions are responsible for 1 and \$0.5 ps, respectively. However, again, equal splitting charges each session 0.75 ps. These two examples show that equally splitting the aggregate cost among sessions is not desirable.

We now present an example to demonstrate the deficiencies of the direct mechanism. Suppose that there are two interfering secondary sessions each with an arrival rate equal to 0.75 pps. Also, suppose that there are two primary channels that both offer 1 pps transmission rates but charge 2 and \$4 per access, respectively. Suppose that the first session has access to both channels and the second session has access only to the cheaper channel. The scheduling cost of each of the sessions if they were the only session to be scheduled is \$1.5 ps. However, when they are both present, the direct access costs of the first and second sessions are \$2.5 and \$1.5 ps, respectively. Hence, the first session might argue that it is charged excessively in favor of the second session, and thus, feels unfairly treated. This example shows that charging the sessions based on their direct cost is not desirable as well, and a careful design should also take these subtle mutual effects into account and share the aggregate cost accordingly.

We now provide a rational basis for splitting the aggregate time average scheduling cost between different sessions. First, let us introduce the notations that are used in this section. Let Z denote the set of all secondary sessions, i.e., $Z = \{1, \dots, L\}$, and let Y be an arbitrary subset of Z , with arrival rate vector equal to $\vec{\lambda}_Y$. Note that $\vec{\lambda}_Z$ is simply $\vec{\lambda}$. Let

$v(Y)$ represent the aggregate time average cost of scheduling the sessions in Y subject to their stability, when only sessions in Y are present in the network. The value of v clearly depends on the scheduling policy. For our ϵ -optimal scheduling policies, we have $v(Y) = U(\vec{\lambda}_Y)$, where $U(\vec{\lambda}_Y)$ is given by the optimization problem defined in (17), (18), (19), and (20). This follows from Lemma 5 and Theorems 1 and 2. Now, the problem is to devise a *rule* for splitting $v(Z)$ among sessions of Z . Assume that the assignment of the time average costs to the sessions of the set Z is represented by the vector function $\vec{\phi}(Z) = (\phi_1(Z), \dots, \phi_L(Z))$ and is defined for any Z and any $\|Z\| = L$. We assume that $\vec{\phi}(\emptyset) = 0$. Also, we assume that the total cost is shared among the sessions, i.e., $\sum_i \phi_i(Z) = v(Z)$. We refer to this condition as *feasibility*. Trivially, $\vec{\phi}(\{i\}) = v(\{i\})$, where i is a single session. Now, let $L \geq 2$ and consider two distinct sessions i and j . Suppose that $\phi_j(Z) - \phi_j(Z \setminus \{i\}) > 0$, then session j would pay less if session i were not present in the network. Thus, session j might object that $\vec{\phi}$ is an *unfair* assignment, unless session i can counterobject that it is at least as much disadvantaged due to the presence of session j , and vice versa.⁴ This results in the following condition for the assignment function:

$$\forall L > 0: \quad \phi_i(Z) - \phi_i(Z \setminus \{j\}) = \phi_j(Z) - \phi_j(Z \setminus \{i\}), \quad (15)$$

which is known as the *balanced contributions property*.

It is known [30] that there is only one assignment function $\vec{\phi}$ that satisfies both feasibility and balanced contributions properties and is called the *Shapley Value*, which is defined as follows:

$$\phi_i(Z) = \sum_{Y \subseteq Z \setminus \{i\}} \frac{\|Y\|!(L - \|Y\| - 1)!}{L!} [v(Y \cup \{i\}) - v(Y)]. \quad (16)$$

The Shapley value has an interesting intuitive interpretation: suppose that all of the sessions are arranged in a random order. Whenever a session with nonzero demand is added to the network, the aggregate time average cost of the scheduling increases. The incremental cost of scheduling a session i depends on the specific order, and thus, is a random variable. Assume that all permutations are assigned equal probabilities. Now, the Shapley value of a session i is the expectation of these cost increments due to addition of that session under the above probability distribution for the random ordering.

Now, as an example, let us calculate the Shapley value of each of the sessions in our three simple examples. In the first example, we have $v(\{1, 2\}) = \$3$, $v(\{1\}) = \$1$, and $v(\{2\}) = \$2$ ps. Thus, the Shapley values are as follows: $\phi_1(\vec{\lambda}) = 1$ and $\phi_2(\vec{\lambda}) = \2 ps. Note that other things being identical, the session that has higher arrival rate incurs a higher cost. In the second example, $v(\{1, 2\}) = 1.5$, $v(\{1\}) = 1$, and $v(\{2\}) = \$0.5$ ps. Thus, $\phi_1(\vec{\lambda}) = 1$ and $\phi_2(\vec{\lambda}) = \0.5 ps. Thus, other things being identical, the session that has access to the cheaper channel incurs lower cost. Finally, in the third example, we have $v(\{1, 2\}) = 4$ and $v(\{1\}) = 1.5$ and $v(\{2\}) = \$1.5$ ps. Thus, $\phi_i(\vec{\lambda}) = \2.5 and $\phi_2(\vec{\lambda}) = \2.5 ps. We see that now none of the sessions can *unilaterally* argue that it is disadvantaged due to the presence of the other session.

4. A similar argument can be constructed when $\phi_j(Z) - \phi_j(Z \setminus \{i\}) < 0$ [29, pp. 170-171].

In our first approach, we started from feasibility and balanced contributions (15) properties and reached the Shapley value function. Alternatively, it can be shown [30] that the Shapley value $\varphi(Z)$ is the unique assignment function that satisfies feasibility along with the following three axioms:

1. **Symmetry:** If for every subset Y that includes i but not j , we have $v((Y \setminus \{i\}) \cup \{j\}) = v(Y)$, then $\phi_i(Z) = \phi_j(Z)$.
2. **Dummy player:** If for every Y , we have $v(Y \cup \{i\}) - v(Y) = v(\{i\})$, then $\phi_i(Z) = v(\{i\})$.
3. **Additivity:** If for every Y , a new cost function u is defined as the sum of the two separate cost functions v and w , i.e., if for every Y , $u(Y) = w(Y) + v(Y)$, then for all $i \in Z$, we have $\phi_i^u(Z) = \phi_i^w(Z) + \phi_i^v(Z)$, where $\phi_i^v(Z)$ is the assignment to the player i when the cost function is v , etc.

These axioms have interesting interpretations in the context of our problem. *Symmetry* guarantees that if two sessions impose identical incremental costs irrespective of the set of sessions, they are scheduled with, then they are charged the same. *Dummy player* says that a session i pays exactly $v(\{i\})$ when the increment in cost of scheduling is $v(\{i\})$ irrespective of the set of sessions that i is scheduled with. For example, a session is a dummy player if it does not interfere with other sessions or when its arrival rate is zero. In the latter case, Shapley value guarantees that sessions with zero arrival rate are not charged. Finally, *additivity* ensures that the cost incurred by the session for two different types of service is the sum of the costs incurred for each service.

Note that in order to be able to calculate the Shapley values, we now require knowledge of stationary statistics of the arrivals and the channels; however, this should not impose a significant problem. It is because billing the sessions, unlike scheduling the sessions, is not delay-sensitive, and thus, can be performed after sufficient data about the statistics of the network are collected.

8 SIMULATION RESULTS

In this section, we apply both the dynamic scheduling policy $\Delta^d(V)$ (Section 4) and the graph-partitioning approximation algorithm $\pi(k, V)$ (Section 5) to a sample network. Particularly, we are interested to see how sensitive the algorithms are to the choice of parameters V and k when they take values much smaller than what the theorems prescribe.

First, we consider 25 nodes located on a square grid of 5×5 nodes, where the distance between every pair of adjacent nodes on the grid is D . The communication is single-hop and there is a session between every two adjacent nodes on the grid; thus, there are a total of 40 sessions to be scheduled. The topology of the network does not change over the duration of simulation, which is $T = 10,000$ slots. There are two orthogonal channels and the nodes have multiradio capability, i.e., they can communicate over different channels simultaneously. We use the same interference model described in case 2 in Section 3.2.

The arrivals to each session are according to an ON/OFF Markovian source and are independent of other sessions. We consider a symmetric arrival vector, i.e., the same arrival rate for all sessions. The transition from ON to ON state in the arrival generating Markov chain is set to 0.7 to model the bursty nature of the data. The ON state represents arrival of a packet. The transition between the OFF to OFF state is

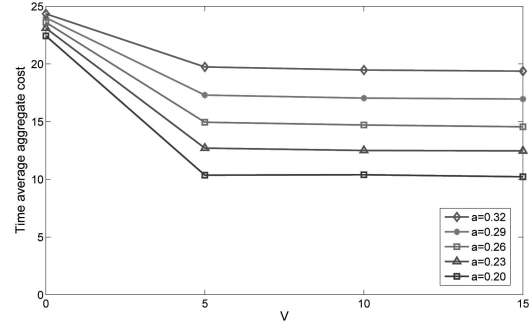


Fig. 3. Time average aggregate cost of $\Delta^d(V)$ scheduling policy versus V for different symmetric arrival rates.

chosen so as to yield the desired arrival rate (denoted here by a) as an input parameter. The first channel has capacity 1 and the second channel has capacity 2 packets per slot. The access fee for channel one is 1 unit per slot and for channel two is 4 units per slot. Both channels are always connected.

We first apply the $\Delta^d(V)$ algorithm developed in Section 4. We observe that the scheduling policy $\Delta^d(V)$ stabilizes the network for all the symmetric arrival rates less than 0.35 packets per slot. The time average aggregate cost of the scheduling is depicted in Fig. 3. For $V = 0$, $\Delta^d(V)$ treats the two channels as if there was one channel of sum of their capacities, and is completely blind to the access fees. We observe that even for a small V , i.e., $V = 5$, the algorithm yields a substantial improvement over fee-blind access to both channels, and the time average aggregate cost improves only slightly by further increase in V . Recall from Theorem 1 (Section 4) that a larger V translates into a higher bound on the sum of backlogs. Thus, our experiments show that, in practice, we can choose a small V and thereby attain the minimum time average aggregate cost without suffering from increased backlogs/network delays.

In Fig. 4, we have gradually increased the symmetric arrival rate (a) and depict the proportion of frequency of access to channel 2 (the more expensive channel) to channel 1. It clearly shows that as the arrival rates increase, the relative frequency of accessing the expensive channel increases. This is because the algorithm is required to stabilize the queues, and thus, has to resort to the expensive channel to prevent growing backlogs.

We now demonstrate the benefits of the approximate scheduling policy by considering a large network. With the same channel and interference settings, we consider a grid of 100×100 nodes. The implementation of the $\Delta^d(V)$ scheduling policy is thus clearly impractical as there are 19,800 links (sessions) in the network, and an exponentially large number of independent sets. However, graph-partitioning technique allows us to perform an approximate scheduling over the whole network. We specifically choose $k = 7$, which renders the largest size of the subgraphs 5×5 . As before, we run the simulation for $T = 10,000$ slots. As is clear from Fig. 5, even for V as small as 5, we witness a significant improvement compared with the case of $V = 0$, and minor sensitivity to further increasing of V .

9 CONCLUSION AND FUTURE DIRECTIONS

In this paper, we studied the problem of spectrum access for a network of secondaries where the primaries impose access

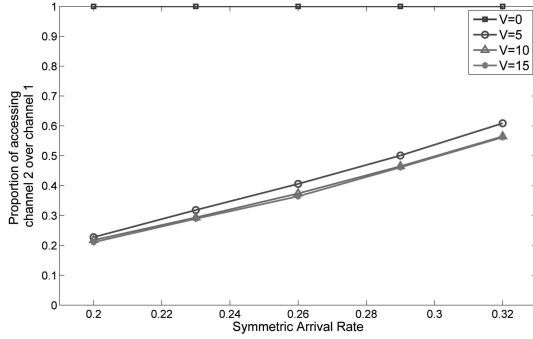


Fig. 4. Time average fraction of access to channel 2 over channel 1 in $\Delta^d(V)$ scheduling policy for different symmetric arrival rates and different values of V .

fees. We developed a dynamic scheduling policy that, without knowledge of arrivals' and channels' statistics, 1) supports the stability region of the network of the secondaries and 2) attains a time average cost that is arbitrarily close to its minimum. We used a Lyapunov drift technique to establish these performance guarantees for a Markovian arrival and network statistics. Next, we considered the issue of the computation time required for calculation of each schedule and proceeded to develop an approximate scheme that attains arbitrary trade-offs between: 1) the complexity of the schedules and 2) the time average access cost subject to stability. The design of the approximate scheme relied on a combination of Lyapunov drift and graph partitioning. We also showed that our results are extendable to multihop networks. Next, we proposed a cost sharing policy based on the concept of the Shapley value that attains a set of desirable properties. Our paper provides a formal framework to the problem of dynamic wireless scheduling with economical consideration. The results in our paper are rigorously proved based on a general Markov chain model of the states of the network and arrivals, from which the i.i.d. case follows as a special case.

The problem of effectively sharing the scheduling cost can open doors to new investigations of this kind. For instance, by introducing new objectives such as variations of fairness and/or collusion prevention from cooperative game theory, other interesting cost sharing policies can be developed. Another interesting direction of research is the issue of pricing: we started with a given set of access fees by primaries. The choice of these costs by each primary will determine the rate of their revenue. In a more advanced model, the announced access fees may also affect the demands of the secondaries. Now, since more than one primary can be present in the network, possibility of competition, co-operation, and collusion can be investigated. Inspired by works on pricing the Internet (e.g., [31]), new pricing mechanisms may be developed to achieve a set of desirable properties and potentially engineer the congestion in a multichannel wireless network in the presence of such dynamics.

APPENDIX A

PROOF OF THEOREM 1

Here, we provide a brief overview of the proof: We establish the existence of a stationary randomized scheduler, $\Delta^*(\vec{\lambda})$,

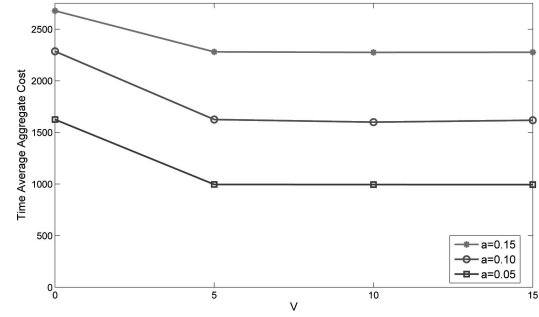


Fig. 5. Time average aggregate cost of partitioning algorithm $\pi(k, V)$ for $k = 7$ versus V for different symmetric arrival rates in the graph of 100×100 nodes.

which provably attains minimum time average cost subject to stability of the network for any arrival rate in the stability region of the network (Lemmas 3, 4, and 5). Next, we compare the performance of our dynamic scheduling policy ($\Delta^d(V)$) against the stationary scheduler, and hence, deduce our results.

Before we get to the proof of Theorem 1, we state four lemmas which we later use. Consider the following optimization problem which we will refer to as $MC(\vec{\lambda})$, where the variables are ω_i^s :

$$U(\vec{\lambda}) = \min_{\{\omega\}} \sum_s \pi_s \sum_{i \in \mathcal{I}_s} r(\mathbf{C}(i, s), s) \omega_i^s, \quad (17)$$

$$\text{s.t. } \sum_s \pi_s \sum_m \sum_{i \in \mathcal{I}_s} c_{lm}(i, s) \omega_i^s = \lambda_l \quad \forall l \in \{1, \dots, L\}, \quad (18)$$

$$\sum_i \omega_i^s = 1 \quad \forall s \in \mathcal{S}, \quad (19)$$

$$\omega_i^s \geq 0 \quad \forall s, i \in \mathcal{I}_s. \quad (20)$$

Lemma 3. $MC(\vec{\lambda})$ is feasible for all $\vec{\lambda} \in \text{Int}(\Lambda)$.

Lemma 4. For any $\vec{\lambda} \in \text{Int}(\Lambda)$, $\Omega_{\min}(\vec{\lambda}) \geq U(\vec{\lambda})$.

In words, any scheduling policy that stabilizes $\vec{\lambda}$ has an aggregate time average cost at least $U(\vec{\lambda})$.

Minimum cost stationary single-hop scheduling policy. Let ω^* be the minimizer of the $MC(\vec{\lambda})$. During each time slot, observe the state of the network $S(t) = s$ and choose scheduling decision $i \in \mathcal{I}_s$ with probability ω_i^{s*} . We refer to this scheduling policy as $\Delta^*(\vec{\lambda})$.

Lemma 5. For any $\vec{\lambda} \in \text{Int}(\Lambda)$ and any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\Omega^{\Delta^*(\vec{\lambda} + \delta \vec{1})} \leq U(\vec{\lambda} + \delta \vec{1}) \leq \Omega_{\min}(\vec{\lambda}) + \epsilon \quad \text{w.p.1.}$$

Also, $\delta \rightarrow 0$ as $\epsilon \rightarrow 0$.

Lemma 6. If for nonnegative real variables X, Y, Z, W , we have $X \leq [Y - Z]^+ + W$, then the following inequality holds: $X^2 \leq Y^2 + Z^2 + W^2 - 2Y(Z - W)$.

Lemmas 3, 4, and 5 follow by similar arguments used in [13] and are relegated to our technical report [28]. The only significant difference in the assumptions is that in [13], energy consumption is associated with allocation of transmission rates. In contrast, we attribute cost only to the actual transmitted flows. Lemma 6 is a simple algebraic relation which can be found in [32, p. 54].

Proof of Theorem 1. Throughout the proof, $\vec{Q}(\tau)$ is the vector of the secondary queue backlogs at time τ under the $\Delta^d(V)$ scheduling policy. We also assume that the network starts at $\tau = 1$ with finite queue backlogs.

Define $L(\tau) \triangleq \frac{1}{2} \sum_l Q_l^2(\tau)$. Also, define $\Delta(L(\tau), \tau) \triangleq E[L(\tau+1) - L(\tau)]$ in which the expectation is taken with respect to the joint distribution of $H(t), t = 1, \dots, \tau+1$ and possibly any randomization used in the scheduling policy. For brevity, we abuse the notation and use $c_{lm}^\Delta(\tau)$ to represent the offered transmission rate assigned to link lm during time slot τ by scheduling policy Δ , which is, in fact, equal to $c_{lm}(i^\Delta(\tau), s(\tau))$. Also, whenever $c_{lm}(t)$ is used without superscript, it pertains to the $\Delta^d(V)$ scheduling policy. Applying Lemma 6 to the network dynamics equation (2) yields:

$$\begin{aligned} \Delta(L(\tau), \tau) &\leq \frac{1}{2} \sum_l E[Q_l^2(\tau)] + \frac{1}{2} \sum_l E[A_l^2(\tau)] \\ &\quad + \frac{1}{2} \sum_l E \left[\left(\sum_m c_{lm}(\tau) \right)^2 \right] \\ &\quad - \sum_l E \left[Q_l \left(\sum_m c_{lm}(\tau) - A_l(\tau) \right) \right] - \frac{1}{2} \sum_l E[Q_l^2(\tau)] \\ &\leq \hat{W} - \sum_l E \left[Q_l \left(\sum_m c_{lm}(\tau) - A_l(\tau) \right) \right], \end{aligned}$$

where $\hat{W} \triangleq \frac{1}{2} [A_{\max}^2 L + (c_{\max})^2 L]$. The last inequality follows from (1) and (5). Adding $VE[r(\mathbf{C}(\tau), s(\tau))]$ to both sides, we obtain

$$\begin{aligned} \Delta(L(\tau), \tau) + VE[r(\mathbf{C}(\tau), s(\tau))] &\leq \hat{W} - E \left[\left(\sum_l Q_l(\tau) \sum_m c_{lm}(\tau) \right) - Vr(\mathbf{C}(\tau), s(\tau)) \right] \\ &\quad + \sum_l E[Q_l(\tau) A_l(\tau)]. \end{aligned} \quad (21)$$

Now, note that by the law of iterated expectations, we have

$$E[Q_l(\tau) A_l(\tau)] = E[E[Q_l(\tau) A_l(\tau) \mid \mathcal{H}(\tau-T)]].$$

Clearly, $Q_l(\tau) \leq Q_l(\tau-T) + \sum_{\nu=\tau-T}^{\tau-1} A_l(\nu)$. Thus,

$$\begin{aligned} E[Q_l(\tau) A_l(\tau)] &\leq E[E[Q_l(\tau-T) A_l(\tau) \mid \mathcal{H}(\tau-T)]] \\ &\quad + TA_{\max}^2. \end{aligned} \quad (22)$$

Since $Q_l(\tau-T) \in \mathcal{H}(\tau-T)$, we have

$$E[Q_l(\tau-T) A_l(\tau) \mid \mathcal{H}(\tau-T)] \quad (23)$$

$$= Q_l(\tau-T) E[A_l(\tau) \mid \mathcal{H}(\tau-T)]. \quad (24)$$

Hence, by using inequality (4) in (22), for $T > T_A(\rho)$, we get

$$E[Q_l(\tau) A_l(\tau)] \leq (\lambda_l + \rho) E[Q_l(\tau-T)] + TA_{\max}^2.$$

Applying the above inequality in (21) yields

$$\begin{aligned} \Delta(L(\tau), \tau) + VE[r(\mathbf{C}(\tau), s(\tau))] &\leq \hat{W} - E \left[\left(\sum_l Q_l(\tau) \sum_m c_{lm}(\tau) \right) - Vr(\mathbf{C}(\tau), s(\tau)) \right] \\ &\quad + (\lambda_l + \rho) \sum_l E[Q_l(\tau-T)] + TLA_{\max}^2. \end{aligned} \quad (25)$$

Now, note that the following inequality holds *pathwise* in $\vec{Q}(\tau)$ and $s(\tau)$:

$$\begin{aligned} \sum_l \left(Q_l(\tau) \sum_m c_{lm}(\tau) \right) - Vr(\mathbf{C}(\tau), s(\tau)) \\ \geq \sum_l \left(Q_l(\tau) \sum_m c_{lm}^{\Delta^*(\cdot)}(\tau) \right) - Vr(\mathbf{C}^{\Delta^*(\cdot)}(\tau), s(\tau)). \end{aligned} \quad (26)$$

The inequality follows because referring to the definition of the scheduling policy $\Delta^d(V)$, at each time slot τ , it observes $\vec{Q}(\tau)$ and $s(\tau)$ and maximizes the left-hand side over *any* possible scheduling decisions, including those made by *any* $\Delta^*(\cdot)$. Taking the expectation of both sides of (26) and applying the result in inequality (25), we obtain

$$\begin{aligned} \Delta(L(\tau), \tau) + VE[r(\mathbf{C}(\tau), s(\tau))] &\leq \hat{W} + TLA_{\max}^2 \\ &\quad - E \left[\sum_l Q_l(\tau) \sum_m c_{lm}^{\Delta^*(\cdot)}(\tau) \right] + VE[r(\mathbf{C}^{\Delta^*(\cdot)}(\tau), s(\tau))] \\ &\quad + (\lambda_l + \rho) \sum_l E[Q_l(\tau-T)]. \end{aligned} \quad (27)$$

Consider the scheduling policy $\Delta^*(\vec{\lambda} + \theta \vec{\mathbf{1}})$, where $\vec{\lambda} + \theta \vec{\mathbf{1}} \in \text{Int}(\Lambda)$. According to Lemma 3, $MC(\vec{\lambda} + \theta \vec{\mathbf{1}})$ is feasible. The stationary probability of the network state s is π_s . Hence, from (6) and following constraint (18) of $MC(\vec{\lambda} + \theta \vec{\mathbf{1}})$ and the definition of $\Delta^*(\vec{\lambda} + \theta \vec{\mathbf{1}})$, we conclude

$$\begin{aligned} \forall \rho > 0, \quad \exists T_C(\rho) < \infty, \text{ s.t. } \forall T > T_C(\rho), \forall \tau : \\ E \left[\sum_m \tilde{c}_m^{\Delta^*(\vec{\lambda} + \theta \vec{\mathbf{1}})}(\tau) \mid \mathcal{H}(\tau-T) \right] &\geq \vec{\lambda} + \theta \vec{\mathbf{1}} - \rho \vec{\mathbf{1}}, \end{aligned} \quad (28)$$

where $\tilde{c}_l \triangleq (c_{l1}, \dots, c_{lm})$. Clearly, $Q_l(\tau) \geq Q_l(\tau-T) - c_{\max} T$. Therefore,

$$\begin{aligned} E \left[Q_l(\tau) \sum_m c_{lm}^{\Delta^*(\vec{\lambda} + \theta \vec{\mathbf{1}})}(\tau) \right] \\ = E[E[Q_l(\tau) \sum_m c_{lm}^{\Delta^*(\vec{\lambda} + \theta \vec{\mathbf{1}})}(\tau) \mid \mathcal{H}(\tau-T)]] \\ \geq E[Q_l(\tau-T) E[c_{lm}^{\Delta^*(\vec{\lambda} + \theta \vec{\mathbf{1}})}(\tau-T) \mid \mathcal{H}(\tau-T)]] - c_{\max}^2 T. \end{aligned} \quad (29)$$

From (27), (28), and (29), for $T > \max\{T_A(\rho), T_C(\rho)\}$, we have

$$\begin{aligned} \Delta(L(\tau), \tau) + VE[r(\mathbf{C}(\tau), s(\tau))] &\leq W \\ &\quad - (\theta - 2\rho) \sum_l E[Q_l(\tau-T)] + VE[r(\mathbf{C}^{\Delta^*}(\tau), s(\tau))], \end{aligned} \quad (30)$$

where

$$W \triangleq \hat{W} + TLA_{\max}^2 + TLe_{\max}^2 = L(T+1/2)(A_{\max}^2 + c_{\max}^2).$$

Stability. We can ignore the second term in the left-hand side of the inequality (30), i.e., $VE[r(\mathbf{C}(\tau), s(\tau))]$, for it is nonnegative for all τ . Also, note that $\vec{\lambda} + \theta \vec{\mathbf{1}} \in \text{Int}(\Lambda)$; thus, (30) holds for θ_{\max} :

$$\begin{aligned} \frac{1}{2} \sum_l E[Q_l^2(\tau+1)] - \frac{1}{2} \sum_l E[Q_l^2(\tau)] \\ \leq W - (\theta_{\max} - 2\rho) \sum_l E[Q_l(\tau-T)] + Vr_{\max}. \end{aligned}$$

This inequality holds for every $\tau > 0$. Summation of these inequalities for $\tau = T + 1, \dots, t + T$ and simplifying the telescopic sum and reordering yield:

$$\begin{aligned} & \frac{1}{2} \sum_l EQ_l^2(t + T + 1) - \frac{1}{2} \sum_l EQ_l^2(T + 1) \\ & + (\theta_{\max} - 2\rho) \sum_{\nu=1}^t \sum_l EQ_l(\nu) \leq tW + tVr_{\max}. \end{aligned}$$

Note that $1/2 \sum_l EQ_l^2(t + T + 1) \geq 0$; hence, we can remove it from the inequality. Now, we take the $\limsup_{t \rightarrow \infty} \frac{1}{t}$ of both sides. Note the fact that $1/2 \sum_l EQ_l^2(T + 1) < \infty$, therefore, as long as $\rho < \theta_{\max}/2$, we have

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\nu=1}^t \sum_l EQ_l(\nu) \leq \frac{(W + Vr_{\max})}{\theta_{\max} - 2\rho}.$$

Hence, for instance, for $\rho = \theta_{\max}/4$, we obtain

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\nu=1}^t \sum_l EQ_l(\nu) \leq \frac{2(W + Vr_{\max})}{\theta_{\max}},$$

where the value of \hat{T} in calculation of W is thus equal to $\max\{T_A(\theta_{\max}/4), T_C(\theta_{\max}/4)\}$. We notice that under the $\Delta^d(V)$ scheduling policy, the random process $H(\tau)$ is a countably infinite-state discrete time Markov chain. We conclude from the above \limsup inequality that the Markov chain of $H(\tau)$ is positive-recurrent. Hence, the inequality is valid in almost sure sense as well:

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l Q_l(\tau) \leq \frac{1}{\theta_{\max}} (W + Vr_{\max}) \quad w.p.1,$$

concluding the result for stability.

Time average cost. Note that by definition of $U(\vec{\lambda} + \theta\vec{1})$, we have

$$U(\vec{\lambda} + \theta\vec{1}) = \sum_s \pi_s \sum_{i \in \mathcal{I}_s} r(\mathbf{C}(i, s), s) \omega_s^{\Delta^*(\vec{\lambda} + \theta\vec{1})}.$$

Thus, from (6) and the definition of $\Delta^*(\vec{\lambda} + \theta\vec{1})$, we have

$$\forall \rho > 0, \quad \exists T_U(\rho) < \infty, \text{ s.t. } \forall T > T_U(\rho), \forall \tau : \quad E[r(\mathbf{C}^{\Delta^*(\vec{\lambda} + \theta\vec{1})}(\tau), s(\tau)) \mid \mathcal{H}(\tau - T)] \leq U(\vec{\lambda} + \theta\vec{1}) + \rho. \quad (31)$$

Using (31) in inequality (30), for $T > \max\{T_U(\rho), T_A(\rho), T_C(\rho)\}$, we obtain

$$\begin{aligned} & \Delta(L(\tau), \tau) + VE[r(\mathbf{C}(\tau), s(\tau))] \leq W \\ & - (\theta - 2\rho) E \left[\sum_l Q_l(\tau - T) \right] + VU(\vec{\lambda} + \theta\vec{1}) + V\rho. \end{aligned}$$

As long as $\rho \leq \theta/2$, we can ignore the term $E[\sum_l Q_l(\tau - T)]$ in the right-hand side of the inequality. Hence,

$$\begin{aligned} & \frac{1}{2} \sum_l EQ_l^2(\tau + 1) - \frac{1}{2} \sum_l EQ_l^2(\tau) + VE[r(\mathbf{C}(\tau), s(\tau))] \\ & \leq W + VU(\vec{\lambda} + \theta\vec{1}) + V\rho. \end{aligned}$$

This inequality holds for every $\tau > 0$. Summing up the inequalities for $\tau = 1 \dots t$ and simplifying the telescopic sum yield:

$$\begin{aligned} & \frac{1}{2} \sum_l EQ_l^2(t + 1) - \frac{1}{2} \sum_l EQ_l^2(1) + V \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \\ & \leq tW + tVU(\vec{\lambda} + \theta\vec{1}) + tV\rho. \end{aligned}$$

We can ignore the term $1/2 \sum_l EQ_l^2(t + 1)$ on the left-hand side as it is nonnegative. Dividing both sides by t and taking the \limsup as t go to infinity, noting the fact that $1/2 \sum_l EQ_l^2(1) < \infty$, we obtain

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \leq \frac{W}{V} + U(\vec{\lambda} + \theta\vec{1}) + \rho.$$

Since $\vec{\lambda}$ is strictly interior to Λ , following Lemma 4, for any given $\epsilon > 0$, we can find a $\underline{\theta} > 0$ such that $U(\vec{\lambda} + \underline{\theta}\vec{1}) \leq \Omega_{\min}(\vec{\lambda}) + \epsilon/4$. Thus, for such $\underline{\theta}$ and $\rho \leq \min\{\underline{\theta}/2, \epsilon/4\}$, we get

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \leq \frac{W}{V} + \Omega_{\min}(\vec{\lambda}) + \epsilon/2.$$

Thus, for every choice of $V \geq \hat{V} \geq 2W/\epsilon$, we have

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \leq \Omega_{\min}(\vec{\lambda}) + \epsilon. \quad (32)$$

Now, to relate the left-hand side of the inequality to $\Omega^{\Delta^d(V)}(\vec{\lambda})$, we first note that due to nondecreasing property of the r function, the following holds:

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t r(\mathbf{f}(\tau), s(\tau)) \leq \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t r(\mathbf{C}(\tau), s(\tau)). \quad (33)$$

Also, note that under $\Delta^d(V)$, $H(t)$ is a discrete-time Markov chain process with countably infinite states. By establishing the stability of the queues, we indeed showed that this Markov chain is positive-recurrent. Hence, any bounded function defined on the $H(t)$ almost surely converges to its mean:

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) = E_0 r(\mathbf{C}(t), s(t)) \quad w.p.1, \quad (34)$$

where the last expectation is taken with respect to the stationary distribution of $H(t)$. Following the same argument, we conclude

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t r(\mathbf{C}(\tau), s(\tau)) = E_0 r(\mathbf{C}(t), s(t)) \quad w.p.1. \quad (35)$$

The result follows from (32), (33), (34), and (35). \square

APPENDIX B

PROOF OF LEMMA 1

Proof. Throughout the proof, $\vec{Q}(\tau)$ is the queue backlogs under the $\Delta^{Id}(V)$ scheduling policy. Also, $\mathbf{C}^\Delta(\tau)$ is used to refer to $\mathbf{C}(i^\Delta(\tau), s(\tau))$ and those without superscript pertain to $\Delta^{Id}(V)$. Following the same steps as in the proof of Theorem 1 until inequality (21), we obtain

$$\begin{aligned} & \Delta(L(\tau), \tau) + VEr(\mathbf{C}(\tau), s(\tau)) \\ & \leq \hat{W} - E \left[\sum_l Q_l(\tau) \sum_m c_{lm}(\tau) - Vr(\mathbf{C}(\tau), s(\tau)) \right] \\ & \quad + \sum_l E[Q_l(\tau) A_l(\tau)]. \end{aligned} \quad (36)$$

From the iid assumption of the arrivals, we get

$$E[Q_l(\tau) A_l(\tau)] = EQ_l(\tau) \lambda_l. \quad (37)$$

Applying (37) in (36) yields:

$$\begin{aligned} & \Delta(L(\tau), \tau) + VEr(\mathbf{C}(\tau), s(\tau)) \\ & \leq \hat{W} - E \left[\sum_l Q_l(\tau) \sum_m c_{lm}(\tau) - Vr(\mathbf{C}(\tau), s(\tau)) \right] \\ & \quad + \lambda_l \sum_l EQ_l(\tau) \leq \hat{W} - (1-\phi) \left\{ E \left[\sum_l Q_l(\tau) \sum_m c_{lm}^{\Delta^*(\cdot)}(\tau) \right] \right. \\ & \quad \left. - VEr(\mathbf{C}^{\Delta^*(\cdot)}(\tau), s(\tau)) \right\} + \lambda_l \sum_l EQ_l(\tau). \end{aligned} \quad (38)$$

Inequality (38) is obtained from inequality (10) which defines the scheduling policy $\Delta^{Id}(V)$. Now, note that scheduling policies $\Delta^*(\cdot)$ make their scheduling decisions independent of the queue lengths. Moreover, due to the assumption of iid network states, $\vec{Q}(\tau)$ is independent of $s(\tau)$. Therefore,

$$E \left[Q_l(\tau) \sum_m c_{lm}^{\Delta^*(\cdot)}(\tau) \right] = EQ_l(\tau) E \left[\sum_m c_{lm}^{\Delta^*(\cdot)}(\tau) \right]. \quad (39)$$

Stability. Consider the policy $\Delta^*((\vec{\lambda} + \hat{\theta}_{max} \vec{1})/(1-\phi))$. Since $(\vec{\lambda} + \hat{\theta}_{max} \vec{1})/(1-\phi) \in \text{Int}(\Lambda)$, according to Lemma 3, $MC((\vec{\lambda} + \hat{\theta}_{max} \vec{1})/(1-\phi))$ is feasible. Note that constraint (18) of $MC((\vec{\lambda} + \hat{\theta}_{max} \vec{1})/(1-\phi))$ and the iid assumption of the states guarantee that

$$E \left[\sum_m c_{lm}^{\Delta^*((\vec{\lambda} + \hat{\theta}_{max} \vec{1})/(1-\phi))}(\tau) \right] = \lambda_l / (1-\phi) + \hat{\theta}_{max} / (1-\phi).$$

Hence, by referring to (39) and (38) and canceling the common terms, we obtain

$$\begin{aligned} & \Delta(L(\tau), \tau) + VEr(\mathbf{C}(\tau), s(\tau)) \leq \hat{W} - \hat{\theta}_{max} \sum_l EQ_l(\tau) \\ & \quad + (1-\phi) VEr(\mathbf{C}^{\Delta^*((\vec{\lambda} + \hat{\theta}_{max} \vec{1})/(1-\phi))}(\tau), s(\tau)) \\ & \leq \hat{W} - \hat{\theta}_{max} \sum_l Q_l(\tau) + (1-\phi) Vr_{max}. \end{aligned}$$

Following similar steps as in the proof of Theorem 1 after (30), we reach the following relation:

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l EQ_l(\tau) \leq \frac{(\hat{W} + (1-\phi) Vr_{max})}{\hat{\theta}_{max}}. \quad (40)$$

Under $\Delta^{Id}(V)$, $\vec{Q}(t)$ is a Discrete-time Markov chain process with countably infinite states. We conclude from the above limsup inequality that the Markov chain of $\vec{Q}(t)$ is positive-recurrent. Hence, the inequality in almost sure sense is also implied:

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l Q_l(\tau) \leq \frac{(\hat{W} + (1-\phi) Vr_{max})}{\hat{\theta}_{max}} \quad w.p.1.$$

Time average cost. Since $\vec{\lambda} \in \text{Int}(\Lambda)$, Lemma 3 guarantees that $MC(\vec{\lambda})$ is feasible. Also, note that

$$Er(\mathbf{C}^{\Delta^*(\vec{\lambda})}(\tau), s(\tau)) = \sum_s \pi_s \sum_{i \in \mathcal{I}_s} \omega_i^{s*} r(\mathbf{C}(i, s), s) = U(\vec{\lambda}).$$

This follows by the iid assumption of the network states and the definition of $U(\vec{\lambda})$. Also, referring to constraint (18) of $MC(\vec{\lambda})$ and the iid assumption of the states, we have: $E[\sum_m c_{lm}^{\Delta^*(\vec{\lambda})}(\tau)] = \lambda_l$. Hence, from (38) and (39):

$$\begin{aligned} & \Delta(L(\tau), \tau) + VEr(\mathbf{C}(\tau), s(\tau)) \\ & \leq \hat{W} + \phi \sum_l \lambda_l EQ_l(\tau) + (1-\phi) VU(\vec{\lambda}). \end{aligned}$$

Taking similar steps as in the proof of Theorem 1 after (31), we achieve the following:

$$\begin{aligned} & \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \leq \frac{\hat{W}}{V} \\ & \quad + \frac{\phi}{V} \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \sum_l \lambda_l EQ_l(\tau) + (1-\phi) U(\vec{\lambda}). \end{aligned}$$

Note that from Cauchy-Schwartz inequality,

$$\lambda_l \leq \sqrt{EA_l^2(t)}.$$

Thus, by applying inequality (40), we obtain

$$\begin{aligned} & \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \leq \frac{\hat{W}}{V} \\ & \quad + \frac{\phi \sqrt{A_{max}^2} (\hat{W} + (1-\phi) Vr_{max})}{V \hat{\theta}_{max}} + (1-\phi) U(\vec{\lambda}). \end{aligned}$$

Since $\vec{\lambda} \in \text{Int}(\Lambda)$, following Lemma 4, we have $U(\vec{\lambda}) \leq \Omega_{min}(\vec{\lambda})$. Therefore,

$$\begin{aligned} & \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \leq \frac{\hat{W}}{V} \\ & \quad + \frac{\phi \sqrt{A_{max}^2} (\hat{W} + (1-\phi) Vr_{max})}{V \hat{\theta}_{max}} + (1-\phi) \Omega_{min}(\vec{\lambda}). \end{aligned} \quad (41)$$

Thus, by choosing a large enough V , we can have

$$\begin{aligned} & \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t Er(\mathbf{C}(\tau), s(\tau)) \leq \phi(1-\phi) \frac{\sqrt{A_{max}^2} r_{max}}{\hat{\theta}_{max}} \\ & \quad + (1-\phi) \Omega_{min}(\vec{\lambda}) + \epsilon/2. \end{aligned}$$

Following a similar argument as in the proof of Theorem 1, we can conclude that the inequality for limsup holds in almost sure sense as well. Hence, the result follows with $\Gamma(\phi) = \phi(1-\phi) \sqrt{A_{max}^2} r_{max} / \hat{\theta}_{max}$. \square

APPENDIX C

PROOF OF THEOREM 2

Proof. Note that by assumption, altering the transmission rates does not affect the interference constraints of the network. Therefore, at time slot t , $w_{lm}^*(t)$ is indeed the weight that any solution of the optimization problem in (9) chooses for link lm if that link is a part of $X^*(t)$.

Let $\iota(t)$ be the integer selected by links at the beginning of time slot t . Define

$$B(t) \triangleq \arg \max_{\substack{X \in \mathcal{X}(s(t)) \\ X \subseteq L(\iota(t))}} \sum_{lm \in X} w_{lm}^*(t).$$

For any given $\vec{Q}(t)$ and $s(t)$, the following identity is obvious: $\sum_{lm \in X^*} w_{lm}^* = \sum_{lm \in X^* \cap G^{(j)}} w_{lm}^* + \sum_{lm \in X^* \cap L^{(j)}} w_{lm}^*$. Thus, from the definition of $B(t)$ and $\pi(k, V)$, we have

$$\sum_{lm \in X^{\pi(k, V)}(t)} w_{lm}^*(t) \geq \sum_{lm \in X^*(t)} w_{lm}^*(t) - \sum_{lm \in B(t)} w_{lm}^*(t), \quad (42)$$

where $X^{\pi(k, V)}(t)$ is the independent set selected by the scheduling policy $\pi(k, V)$ at time slot t . Now,

$$\begin{aligned} & E \left[\sum_{lm \in B(t)} w_{lm}^*(t) \mid \vec{Q}(t), s(t) \right] \\ &= \sum_{j=0}^{k-1} P(\iota(t) = j \mid \vec{Q}(t), s(t)) \\ & \quad \left\{ E \left[\sum_{lm \in B(t)} w_{lm}^*(t) \mid \vec{Q}(t), s(t), \iota(t) = j \right] \right\} \\ &= (1/k) \sum_{j=0}^{k-1} \max_{\substack{X \in \mathcal{X}(s(t)) \\ X \subseteq L^{(j)}}} \sum_{lm \in X} w_{lm}^*(t). \end{aligned} \quad (43)$$

The above inequality holds for any instance of network graph and queue backlogs, i.e., holds *pathwise* in $s(t)$ and $\vec{Q}(t)$.

We now bound the right-hand side of the inequality (43). Let $\vec{\omega}$ be an arbitrary vector of nonnegative real weights for the links in an instance of the network graph. Let S_{lm} be the set of links that if scheduled will interfere with link lm . For any link lm , $\omega_{lm} \leq \sum_{i \in X^* \cap S_{lm}} \omega_i$, since, otherwise, X^* could be improved by instead selecting lm and deselecting the other links in $S_{lm} \cap X^*$. Let X_0, \dots, X_{k-1} be k -arbitrary-independent sets such that $X_j \subseteq L^{(j)}$, for $j = 0, \dots, k-1$. Let $\eta_{lm}^{(j)} \triangleq \|X_j \cap S_{lm}\|$. Thus,

$$\sum_{lm \in X_j} \omega_{lm} \leq \sum_{lm \in X_j} \sum_{i \in X^* \cap S_{lm}} \omega_i = \sum_{i \in X^*} \sum_{lm \in X_j \cap S_i} \omega_i \quad (44)$$

$$= \sum_{lm \in X^*} \eta_{lm}^{(j)} \omega_{lm}, \quad (45)$$

where the equality in (44) follows from pairwise and symmetric property of the interference model, and in equality (45), we have used the definition of $\eta_{lm}^{(j)}$ along with a change of indexing. Thus,

$$\sum_{j=0}^{k-1} \sum_{lm \in X_j} \omega_{lm} \leq \sum_{lm \in X^*} \left(\sum_{j=0}^{k-1} \eta_{lm}^{(j)} \right) \omega_{lm}. \quad (46)$$

Let the *supergrid* be the set of all lines of all grids. Then, the *supergrid* is a grid where the distance between any two consecutive horizontal (vertical) lines is D .

Let $\psi_{lm} \triangleq \{j : lm \in L^{(j)}\}$. Then, $\|\psi_{lm}\| \leq 4$, since an end node of link lm can be within a distance $D/2$ from at most two horizontal and two vertical lines. Now, let $\hat{\psi}_{lm} \triangleq \{j : lm \notin L^{(j)} \text{ \& } S_{lm} \cap L^{(j)} \neq \emptyset\}$. Then, $\|\hat{\psi}_{lm}\| \leq 8$. This

is because, lm can interfere with a link in $L^{(j)}$ but not be member of it, only if one of its end nodes is within a distance of $5D/2$ from a horizontal or a vertical line of grid j and none of its end nodes are within $D/2$ distance of any line of grid j . This can occur at most four times for vertical lines and four times for horizontal lines of *supergrid*. Therefore, $\eta_{lm}^{(j)} > 0$ for at most $4 + 8 = 12$ different js in $\{0, \dots, k-1\}$.

Now, for each cases 1 and 2, we upperbound the value of $\eta_{lm}^{(j)}$. Since each X_j is an independent set, for all $j \in \psi_{lm}$, we have $\eta_{lm}^{(j)} = 1$. Now, let us focus on the js in $\hat{\psi}_{lm}$. In case 2, as channels are assumed orthogonal, only links of the same channel can interfere with each other. Hence, for any $lm \in X^*$, the maximum number of the links that interfere with lm but do not interfere with each other is 8 [18]. In case 1, similarly, we can have up to eight links of the same channel that interfere with lm but not with each other. In addition, up to two extra links of dissimilar channels can have a common end node with lm , and thus, by description of case 1, interfere with lm . (Note that for the special case of $M = 2$, only one such extra link is possible.)

We can now upperbound $\sum_{j=0}^{k-1} \eta_{lm}^{(j)}$. Following the above observations, for any $lm \in X^*$, $\sum_{j=0}^{k-1} \eta_{lm}^{(j)} < 4 + 8 \times 10 = 84$ in case 1 and $\sum_{j=0}^{k-1} \eta_{lm}^{(j)} < 4 + 8 \times 8 = 68$ in case 2. Applying these inequalities in (46), we obtain

$$\sum_{j=0}^{k-1} \sum_{lm \in X_j} \omega_{lm} \leq \beta \sum_{lm \in X^*} \omega_{lm}, \quad (47)$$

where $\beta = 84$ in case 1 and $\beta = 68$ in case 2. Considering the relations (42), (43), and (47), we obtain

$$\begin{aligned} & E \left[\sum_{lm \in X^{\pi(k, V)}} w_{lm}^*(t) \mid \vec{Q}(t), s(t) \right] \\ & \leq (1 - \beta) E \left[\sum_{lm \in X^*} w_{lm}^*(t) \mid \vec{Q}(t), s(t) \right]. \end{aligned}$$

Now, since the inequality holds *pathwise* in $s(t)$ and $\vec{Q}(t)$, we can take the expectation of both sides w.r.t. $s(t)$, $\vec{Q}(t)$ to obtain:

$$E \left[\sum_{lm \in X^{\pi(k, V)}} w_{lm}^*(t) \right] \leq (1 - \beta/k) E \left[\sum_{lm \in X^*} w_{lm}^*(t) \right].$$

Comparing the above inequality with (8) and (9) implies that scheduling policies $\pi(k, V)$ satisfy inequality (10), as the necessary condition of Lemma 1 for $\phi = \beta/k$, where $\beta = 84$ for case 1 and $\beta = 68$ for case 2. Thus, the performance guarantees of Lemma 1 hold with the respective ϕ s, and where

$$\hat{\Gamma}(k) = \beta/k(1 - \beta/k) \sqrt{A_{max}^2 r_{max} / \hat{\theta}_{max}}.$$

□

Remarks on graph partitioning. Note that condition $0 \leq \phi < 1$ requires $k > \beta$. If in the scheduling policy $\pi(k, V)$, the grids are $\mu k D$ distanced, where $\mu \geq 1$ is an integer constant, then β in inequality (47) decreases as μ is increased from 1. The best upperbound is realized for $\mu = 6$ where $\beta = 20$ and 16 for cases 1 and 2, respectively. This can be established by

showing that now $\|\psi_{lm}\| + \|\hat{\psi}_{lm}\| \leq 2$. The details are straightforward and are omitted for brevity. Also, note that similar performance guarantees can be obtained for a 3D network, where grids are replaced with 3D lattices and each lattice is specified by its first three planes. The calculations of the constants are quite identical to the 2D case and are omitted for brevity. Finally, note that our graph-partitioning analysis is not specific to the interference model assumed in this paper and is readily extendable to any other pairwise symmetric interference model.

REFERENCES

- [1] M. McHenry, "NSF Spectrum Occupancy Measurements Project Summary," technical report, Shared Spectrum Company, 2005.
- [2] T. Weiss and F. Jondral, "Spectrum Pooling: An Innovative Strategy for the Enhancement of Spectrum Efficiency," *IEEE Comm. Magazine*, vol. 42, no. 3, pp. S8-14, Mar. 2004.
- [3] R. Tandra, S. Mishra, and A. Sahai, "What Is a Spectrum Hole and What Does It Take to Recognize One?" *Proc. IEEE*, vol. 97, no. 5, pp. 824-848, May 2008.
- [4] I.F. Akyildiz, W.-Y. Lee, M.C. Vuran, and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," *Computer Networks*, vol. 50, no. 13, pp. 2127-2159, 2006.
- [5] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79-89, 2007.
- [6] W. Xu, P. Kamat, and W. Trappe, "TRIESTE: A Trusted Radio Infrastructure for Enforcing Spectrum Etiquettes," *Proc. First IEEE Workshop Networking Technologies for Software Defined Radio Networks (SDR '06)*, 2006.
- [7] W. Lehr and J. Crowcroft, "Managing Shared Access to a Spectrum Commons," *Proc. First IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, 2005.
- [8] C. Kloeck, H. Jaekel, and F. Jondral, "Dynamic and Local Combined Pricing, Allocation and Billing System with Cognitive Radios," *Proc. First IEEE Int'l Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, 2005.
- [9] M. Andrews, "Maximizing Profit in Overloaded Networks," *Proc. IEEE INFOCOM*, 2005.
- [10] M. Neely, E. Modiano, and C. Li, "Fairness and Optimal Stochastic Control for Heterogeneous Networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 2, pp. 396-409, Apr. 2008.
- [11] A. Eryilmaz and R. Srikant, "Joint Congestion Control, Routing, and MAC for Stability and Fairness in Wireless Networks," *IEEE J. Selected Areas in Comm.*, vol. 24, no. 8, pp. 1514-1524, Aug. 2006.
- [12] P. Chaporkar and S. Sarkar, "Stable Scheduling Policies for Maximizing Throughput in Generalized Constrained Queueing Systems," *IEEE Trans. Automatic Control*, vol. 53, no. 8, pp. 1913-1931, Sept. 2008.
- [13] M. Neely, "Energy Optimal Control for Time Varying Wireless Networks," *IEEE Trans. Information Theory*, vol. 52, no. 7, pp. 2915-2934, July 2006.
- [14] C. Peng, H. Zheng, and B. Zhao, "Utilization and Fairness in Spectrum Assignment for Opportunistic Spectrum Access," *Mobile Networks and Applications*, Springer, 2006.
- [15] L. Cao and H. Zheng, "Distributed Spectrum Allocation via Local Bargaining," *Proc. IEEE CS Conf. Sensor and Ad Hoc Comm. and Networks*, 2005.
- [16] R. Urgaonkar and M. Neely, "Opportunistic Scheduling with Reliability Guarantees in Cognitive Radio Networks," *IEEE Trans. Mobile Computing*, vol. 8, no. 6, pp. 766-777, June 2008.
- [17] L. Tassiulas, "Linear Complexity Algorithms for Maximum Throughput in Radionetworks and Input Queued Switches," *Proc. IEEE INFOCOM*, 1998.
- [18] P. Chaporkar, K. Kar, and S. Sarkar, "Throughput Guarantees through Maximal Scheduling in Wireless Networks," *Proc. 43rd Ann. Allerton Conf. Comm., Control and Computing*, 2005.
- [19] A. Eryilmaz, A. Ozdaglar, and E. Modiano, "Polynomial Complexity Algorithms for Full Utilization of Multi-Hop Wireless Networks," *Proc. IEEE INFOCOM*, 2007.
- [20] A. Gupta, X. Lin, and R. Srikant, "Low-Complexity Distributed Scheduling Algorithms for Wireless Networks," *IEEE/ACM Trans. Networking*, vol. 17, no. 6, pp. 1846-1859, Dec. 2009.
- [21] Y. Yi and M. Chiang, "Wireless Scheduling Algorithms with $o(1)$ Overhead for M-Hop Interference Model," *Proc. IEEE Int'l Conf. Comm. (ICC '08)*, 2008.
- [22] X. Lin and S. Rasool, "Constant-Time Distributed Scheduling Policies for Ad Hoc Wireless Networks," *IEEE Trans. Automatic Control*, vol. 54, no. 2, pp. 231-242, Feb. 2009.
- [23] X. Lin and N. Shroff, "The Impact of Imperfect Scheduling on Cross-Layer Congestion Control in Wireless Networks," *IEEE/ACM Trans. Networking*, vol. 14, no. 2, pp. 302-315, Nov. 2006.
- [24] S. Sarkar and S. Ray, "Arbitrary Throughput versus Complexity Tradeoffs in Wireless Networks Using Graph Partitioning," *IEEE Trans. Automatic Control*, vol. 53, no. 10, pp. 2307-2323, 2008.
- [25] U. Kozat, I. Koutsopoulos, and L. Tassiulas, "Cross-Layer Design for Power Efficiency and QoS Provisioning in Multi-Hop Wireless Networks," *IEEE Trans. Wireless Comm.*, vol. 5, no. 11, pp. 3306-3315, 2006.
- [26] G. Sharma, R. Mazumdar, and N. Shroff, "On the Complexity of Scheduling in Wireless Networks," *Proc. 12th Ann. Int'l Conf. Mobile Computing and Networking*, pp. 227-238, 2006.
- [27] D. Shah, P. Giaccone, and B. Prabhakar, "Efficient Randomized Algorithms for Input-Queued Switch Scheduling," *IEEE Micro*, vol. 22, no. 1, pp. 10-18, Jan. 2002.
- [28] M. Khouzani and S. Sarkar, "Cost Efficient Secondary Access in Time Varying Cognitive Networks," technical report, Univ. of Pennsylvania, 2008.
- [29] C. Courcoubetis and R. Weber, *Pricing Communication Networks*. Wiley Hoboken, 2003.
- [30] M. Osborne and A. Rubinstein, *A Course in Game Theory*. MIT Press, 1994.
- [31] J. MacKie-Mason and H. Varian, "Pricing Congestible Network Resources," *IEEE J. Selected Areas in Comm.*, vol. 13, no. 7, pp. 1141-1149, Sept. 1995.
- [32] L. Georgiadis, L. Tassiulas, and M. Neely, *Resource Allocation and Cross Layer Control in Wireless Networks*. Now Publishers, Inc., 2006.



M.H.R. Khouzani received the BSc degree from the Sharif University of Technology, Tehran, Iran, in 2006, and the MSE degree in electrical and systems engineering in 2008 from the University of Pennsylvania, Philadelphia, where he is currently working toward the PhD degree in the Multimedia and Networking Laboratory. His research interests are in stochastic optimization, resource allocation, and dynamic games in wireless networks.



Saswati Sarkar received the BS degree from Jadavpur University in 1994, the MS degree from the Indian Institute of Science Bengaluru in 1996, and the PhD degree from the Electrical and Computer Engineering Department, University of Maryland, College Park, in 2000. She is currently an associate professor in the Electrical and Systems Engineering Department, University of Pennsylvania, where she joined as an assistant professor in 2000. She received the Motorola Gold Medal for being adjudged the best masters student in the division of electrical sciences at the Indian Institute of Science, Bengaluru, and also received the US National Science Foundation Career Award in 2003. Her research interests are in resource allocations, security, and economics of wireless networks. She has served as an associate editor of the *IEEE Transactions in Wireless Networks* from 2001 to 2006 and is currently in the editorial board of the *IEEE/ACM Transactions on Networking*.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.