Throughput-optimal Scheduling in Multichannel Access Point Networks under Infrequent Channel Measurements

Koushik Kar, Xiang Luo and Saswati Sarkar

***Abstract*—We consider the problem of uplink/downlink scheduling in a multichannel wireless access point network where channel states differ across channels as well as users, vary with time, and can be measured only infrequently. We demonstrate that, unlike infrequent measurement of queue lengths, infrequent measurement of channel states reduce the maximum attainable throughput. We then prove that in frequency division multiplexed systems, a dynamic scheduling policy that depends on both the channel rates (averaged over the measurement interval) and the queue lengths, is throughput optimal. We also generalize the scheduling policy to solve the joint power allocation and scheduling problem. In addition, we provide simulation studies that demonstrate the impact of the frequency of channel and queue state measurements on the average delay and attained throughput.**

***Index Terms*—Infrequent channel measurements, multi-channel access point networks, throughput-optimal scheduling.**

I. Introduction

Future wireless networks are likely to provide each user access to multiple channels. The dynamic scheduling problem at any given time in such networks is to determine (i) the set of users that can transmit/receive, and (ii) the set of channels that a user can use. Our goal is to optimally determine the above so as to maximize the system throughput using on-line adaptive policies. The availability of multiple channels gives rise to several unique challenges in attaining the above goal. Channel characteristics at any given time will typically be different for different channels, and these characteristics will also vary with time. In a system with a large number of users and channels, an individual user could use only a small number of channels at any time. Therefore, measuring the channel quality perceived by each user for each channel would require additional probe packets, which introduces a significant measurement overhead. Thus unlike single-channel networks, scheduling in multichannel networks must be done under inaccurate channel state information, resulting from infrequent channel measurements. Moreover, in a multichannel wireless system, the scheduling questions depend strongly on the transmission mechanisms. Specifically, the scheduling constraints differ significantly based on whether the transmission by each user is single-channel or multi-channel, and the manner in which power is allocated across channels. Our contribution in this paper is to develop optimal scheduling policies that address the above challenges.

Our first contribution is to demonstrate that infrequent channel state measurements affect the system throughput in a fundamentally different way than infrequent measurements of other state variables. Specifically, it is well-known that infrequent measurements of queue lengths of users do not alter the maximum attainable throughput region, as long as the measurement intervals are upper bounded by a constant. We however show that infrequent measurement of channel states does reduce the maximum attainable throughput region. We further prove that a weighted queue-length based scheduling policy attains the maximum attainable throughput region under partial information about channel states. The weights must be chosen based on the average channel rates till the next measurement instant. We also investigate the structure of the optimal scheduling policy under specific scheduling constraints. We show that for single-channel transmission by users, the throughput-optimal scheduling policy is a maximum weighted matching between the users and the channels, and for multi-channel transmission by users, on the other hand, the scheduling policy corresponds to a maximum weighted *poly*-matching. We then show how our results can be extended to jointly optimize the scheduling and power allocation under multi-channel transmission. From a practical perspective, the algorithms that we present in this paper can be used for uplink/downlink scheduling and power assignment for multichannel wireless systems like 802.16 access point networks.

II. Related Work

There is a rich body of literature on the subject of throughput-optimal scheduling in a wide variety of queueing networks [1], [2], [6], [7], [15], [19], [20], [21], [22], [23], [24], [25], [26], [29], [31], [30]. These papers either assume that the service rates of the queues do not vary with time, or if the service rates vary, the schedulers know the service rates of the queues before each scheduling decision. The equivalent assumption in our context is that the schedulers know the instantaneous channel states. Recently, Neely *at al.* have addressed the problem of jointly selecting the queues

K. Kar and X. Luo are with the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA (email: {kark,luox3}@rpi.edu).

S. Sarkar is with the Department of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA (email: swati@seas.upenn.edu).

This work was supported by the National Science Foundation under grants NCR-0238340, CNS-0435306, NCR-0448316 and CNS-0435141.

This is an extended version of a work presented in *IEEE INFOCOM 2007*, Anchorage, AK, USA, May 2007.

to serve and determining the service rates of the selected queues by appropriately regulating the transmission power levels [18]. They also assume that the scheduler always knows the instantaneous states of the channels. Our main contribution is to develop throughput optimal scheduling policies when the scheduler knows the channel states only infrequently. We also demonstrate that the impact of infrequent knowledge of channel states is substantially different from that of infrequent knowledge of queue lengths. While infrequent knowledge of queue lengths does not alter the maximum achievable throughput region (as shown by several previous results in different settings [1], [2], [20], [21], [22], [23], [24], [25], [26], [27]), we show in this paper that infrequent knowledge of service rates substantially reduces the maximum achievable throughput region.

Several interesting medium access control protocols, e.g., [9], [14], [8], [16], [28], [32], have been proposed for selecting channels in context of specific wireless technologies, e.g., IEEE 802.11, which do not however guarantee throughput optimality. Our contribution lies in the development of scheduling algorithms that provably maximize throughput in presence of time variations, asymmetry in the rates of different channels, and infrequent measurements.

For the case where each user can transmit over multiple channels simultaneously, there have been several recent papers that address a problem that is closely related to ours [5], [33], [10], [12]. The authors in [5], [33] have addressed the question of how resources (like bandwidth and power) should be allocated to users in an multi-channel transmission system to maximize system throughput. However, in these works, the resource allocation problem is not considered in a stochastic setting, and therefore the problem addressed in [5], [33] is quite different from the stochastic dynamic optimization problem that we consider here. In [10], [12], the authors address the multi-channel transmission case of our problem for two-state (on-off) channel models. In contrast, we consider channel models that are much more general (can have any number of states) and address both the cases of single-channel and multi-channel transmission by users. More importantly, unlike our work, the results in [10], [12] assume that the instantaneous channel states are always known, and do not jointly optimize the channel and power allocations.

III. Formulation

A. System Model and Assumptions

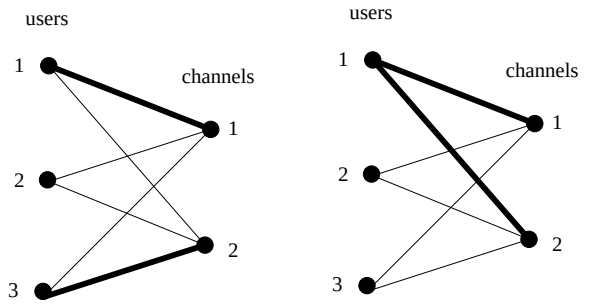
Our system consists of a set of users sharing a set of channels to communicate with an access point (AP). Let M denote the set of channels, and N denote the set of users. The access point network that we consider is a centralized network, where the scheduling decisions (both uplink and downlink) are taken by the AP. In the following, we focus most of our discussion on uplink scheduling, where the users are transmitting data to the AP; the formulation and approach presented here can easily be extended to the downlink case. We assume that the AP is equipped with a separate transceiver for each channel, and is thus capable of receiving data simultaneously from multiple users provided they receive on different channels. However, the AP cannot successfully receive data from multiple users over a single channel. In this scenario, whether a user can simultaneously transmit on multiple channels or not, depends on the specific system considered, and is discussed in Section III-B.

We allow channel conditions to vary across channels as well as users. Channel conditions depend on various factors like fading and interference (from neighboring access point networks), which typically depend on the channel frequency, as well as the user location. Therefore, the attainable rate on a channel may be different for different channels; moreover, the attainable rate may also depend on the user using the channel. Let α_{ij} ($0 \leq \alpha_{ij} \leq 1$) denote the packet success probability when user i transmits a packet on channel j. In the rest of the paper, we will therefore refer to α_{ij} as the *channel rate* of user i on channel j. Note that the channel rates are typically functions of time, since fading and interference levels at any location can vary with time. These variations will be more pronounced when the users are mobile.

We assume that time is slotted, and the slots are denoted by $t = 1, 2, ...$ All packets have the same length, and the transmission time of a packet equals a slot length. We assume that packet arrivals occur at the beginning of any time slot, and packet departures occur at the end of the time slot. At any given time slot, the number of packet arrivals for different users can be arbitrarily correlated. For user i, the number of arrivals in any slot follows an i.i.d. process, with mean λ_i. Let $\vec{\lambda} = (\lambda_i, i \in N)$ denote the vector of average arrival rates. Note that while our results assume i.i.d. traffic arrival patterns, they can be extended to more general arrival patterns using fluid flow techniques [4]. We assume that each channel rate, α_{ij}, evolves in time according to a finite-state Markov chain. At any given time, the different α_{ij}s can be arbitrarily correlated. Finally, we state our assumptions on the sampling of channel and queue states. Let the time slots be grouped into *intervals* of time T. Thus the $(k+1)$th interval consists of slots $kT, ..., (k+1)T-1$. Although the channel conditions and queue lengths can change in each slot, these are measured only at the beginning of each interval, i.e., at the beginning of slot kT, for $k = 0, 1, ...$ Thus the interval length T denotes the duration between successive sampling instances of the channel conditions and queue lengths.

B. Scheduling Constraints

Next, we describe the constraints on our scheduling policy. At the beginning of each interval, for each channel, a user is selected to transmit on that channel during the interval. Note that a channel cannot be assigned to multiple users in the same interval. Under single-channel transmission, a user can transmit on only one channel at any given time. Therefore, in this case, the scheduling policy across channels corresponds to a *matching* [3] in a bipartite graph, where the users and the channels represent the two sets of vertices that need to be matched. Under multi-channel transmission, however, a user can transmit on multiple channels at the same time. Thus in this case, a user can be matched to multiple channels, but not vice versa. In this paper, we refer to such a one-to-many



(a) Matching (b) Poly−matching

Fig. 1. *Matching vs. poly-matching:* The figure shows one possible matching and one possible poly-matching for 3 users and 2 channels. (Note that the matching/poly-matching is represented by the bold edges.)

matching between the users and channels as a *poly-matching*. Figure 1 explains the difference between matchings and poly-matchings.

Note that there can be multiple matchings or poly-matchings in the bipartite graph of users and channels (the total number of matchings or poly-matchings is in fact, exponential in the size of the user-channel graph), and different matchings and poly-matchings will provide significantly different throughputs. A good choice of matching or poly-matching is critical to attaining high system throughput. Therefore, the key challenge in the dynamic scheduling question considered here is to select the right matching or poly-matching at any time slot, so as to maximize the long-term system throughput.

C. Stability Region and Throughput-optimal Scheduling

The notion of throughput-optimal scheduling is based on the notion of a "stability region"; so we define the latter first. A system is said to be *stable* for an arrival rate vector $\vec{\lambda}$ under a scheduling policy Ψ, if the expected lengths of all queues in the system remain bounded over all time, when the packet arrival rate vector is $\vec{\lambda}$ and Ψ is used as the scheduling policy. In such a case, scheduling policy Ψ is said to *stabilize* the system for arrival rate vector $\vec{\lambda}$. The *stability region* of the system is the set of all arrival rate vectors for which the system can be stabilized by *some* scheduling policy. Intuitively, the arrival rate vector belonging to the stability region is "attainable", since there exists a scheduling policy under which the system is stable for that arrival rate vector. Moreover, a rate vector outside the stability region is not attainable, since all scheduling policies would lead to unbounded queues in the system for that arrival rate vector. As we argue later in the paper, the stability region in our system depends on the measurement interval T. Let Λ_T denote the stability region of the system for interval length T. An analytical characterization of the stability region of the system that we consider can be found in the appendix (refer to (8)-(10)).

A scheduling policy is said to be *throughput-optimal* if it stabilizes the system for all arrival rate vectors that are *strictly within* the stability region. In other words, a throughput-optimal scheduling policy can "attain" all arrival rate vectors that belong to the *interior* of the stability region Λ_T. In the next few sections, we present throughput-optimal scheduling policies for the multichannel wireless system described above.

IV. Throughput-Optimal Scheduling

Before we present our scheduling policy and argue about its throughput-optimality, we discuss some properties of the stability region Λ_T.

A. Characterization of the Stability Region

In the following lemma, we prove that the stability region reduces with increase in T. Let $\mathrm{Int}(\Lambda_T)$ represent the interior of the stability region, Λ_T.

Lemma 1: For any $T \geq 1$, $\Lambda_{lT} \subseteq \Lambda_T \ \ \forall$ positive integers l. If $l > 1$, there exists systems where $\mathrm{Int}(\Lambda_{lT}) \subset \mathrm{Int}(\Lambda_T)$.

Lemma 1 is proved in the appendix. Intuitively, Lemma 1 states that the stability region "shrinks" as the measurement interval increases.

Note that in practice, some inference on the channel states can be drawn from the success or failure of packets transmitted during an interval. However, in our definition of Λ_T, we assume that such information is not used by the scheduling policy.

Let us now consider a scenario where the queue states are measured only at the beginning of each interval (of T time slots), but the channel states are measured at the beginning of *every* time slot. Let $\hat{\Lambda}_T$ denote the stability region in this case. The following result can be easily shown, and has been observed in the existing literature in different contexts [21], [23], [24], [25], [26], [27]:

Observation 2: For any $T \geq 1$, $\mathrm{Int}(\hat{\Lambda}_T) = \mathrm{Int}(\Lambda)$.

The above observation (proof outline in appendix) states that the stability region remains the same if the queue measurement interval is increased, as long as the channel states are measured every time slot.

From the lemma and observation stated above, we can conclude that the shrinking of the stability region Λ_T with increasing T, is a result of the reduction in the channel rate measurement frequency, and not due to the reduction in the frequency of queue-length measurements. Increasing the queue measurement interval (while keeping the channel measurement interval fixed) does not affect the maximum achievable throughput; it usually results only in an increase in the average packet delay. Increasing the channel measurement interval, however, not only increases the average delay, but also leads to a reduction in the maximum achievable throughput. Thus the reduction in the frequency of measurement in the channel rates affects the system in a fundamentally different way than that of the queue-lengths. The optimal scheduling policy which we state in the next section provides more intuition behind these results. We also substantiate these observations through simulation results in Section V.

B. Scheduling Policy

We now describe our scheduling policy Ψ_T, which is parameterized by the length T of the measurement interval.

The scheduling policy consists of two components: (i) packet queueing policy and (ii) packet service policy. Both of these can be executed in parallel. We will first describe the packet queuing policy which assigns the service channel to each packet of each user. Each user maintains a queue for each of the channels (see Figure 2). A queue for channel j at user i contains packets of i that will be scheduled on channel j. A packet, on arrival, is stored in the queue with the *smallest queue-length*, amongst all queues for that user. Thus the channel on which a packet will be scheduled is assigned on packet arrival. Let $Q_{ij}(t)$ denote the length of the queue for channel j at user i at time slot t. In computing $Q_{ij}(t)$, the packets that enter the corresponding queue at the beginning of time slot t are also taken into account. In our packet queueing policy, the arriving packets are routed to the corresponding queue (i.e., are considered eligible for scheduling) only at the beginning of each interval. Thus, a packet of user i arriving at a time slot t, where t satisfies $(k-1)T \leq t < kT$, will enter a queue only at the beginning of time slot kT, i.e., at the beginning of the $(k+1)$th interval. Moreover, the packet will enter the queue for channel j at user i, where j satisfies

$$j = \arg \min_{j' \in M} Q_{ij'}(kT). \tag{1}$$

We now describe the packet service policy. Our packet service policy selects the matching (poly-matching) at the beginning of the $(k+1)$th interval, and uses it for the rest of that interval. Recall that the channel rates α_{ij} are functions of time, and let $\alpha_{ij}(t)$ denote the corresponding values in time slot t. Now, for any user i and channel j, define $\tilde{\alpha}_{ij}(kT)$ as follows

$$\tilde{\alpha}_{ij}(kT) = \frac{1}{T} E\Big(\sum_{t=kT}^{(k+1)T-1} \alpha_{ij}(t) | \alpha_{ij}(kT) \Big), \tag{2}$$

where $E(\cdot)$ denotes the expectation of a random variable. In other words, $\tilde{\alpha}_{ij}(kT)$ denotes the average channel rate until the next channel measurement instant, given the current (observed) channel state, $\alpha_{ij}(kT)$. Note that since the channel rate α_{ij} evolves according to a finite-state Markov chain, $\tilde{\alpha}_{ij}(kT)$ can be computed from (2) using the multi-step transition probabilities of the Markov chain.

Now associate a weight of $\tilde{\alpha}_{ij}(kT)Q_{ij}(kT)$ with each "edge" (i,j) in the user-channel bipartite graph (note that an edge corresponds to a user-channel pair). Note that a matching (poly-matching) can be viewed as a collection of edges. The weight of a matching (poly-matching) is the sum of the weights of the edges belonging to the matching (poly-matching). Thus, the weight of a matching (poly-matching) Φ, computed at time kT, is given by

$$\sum_{(i,j) \in \Phi} \tilde{\alpha}_{ij}(kT) Q_{ij}(kT). \tag{3}$$

Then the packet service policy is to assign channels to users so that (3) is maximized. Thus the user-channel assignment for the single-channel transmission (multi-channel transmission) case corresponds to the *maximum weighted matching (poly-matching)* in the user-channel bipartite graph. Figure 2 provides a schematic diagram that explains our queueing and service policies for the two cases.

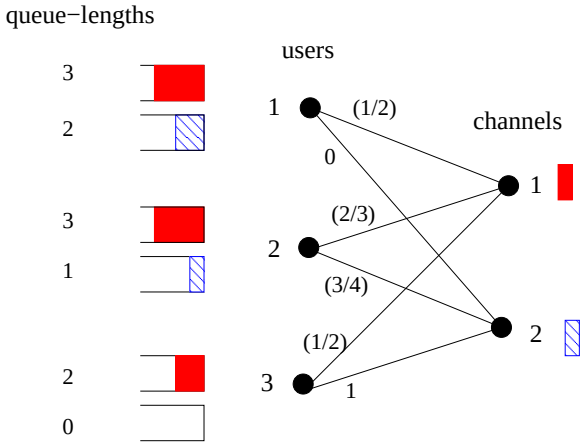


Fig. 2. *Example: Scheduling for single-channel and multi-channel transmission:* In this case the optimal matching is {(1,1),(2,2)}, with a total weight of $\frac{9}{4}$, and the optimal poly-matching is {(2,1),(2,2)} with a total weight of $\frac{11}{4}$. The number shown across each edge represents, $\tilde{\alpha}_{ij}(kT)$, the average channel rate until the next measurement instant, given the current (observed) channel state. Note that the queue-length for channel 2 is smaller than that for channel 1, at all users; therefore, any packets arriving at any user before the next measurement instant will be stored in the queue corresponding to channel 2 at that user.

We now describe the intuition behind the design. For higher system throughput, we would like to schedule user i to transmit on channel j if the expected rate (in the current interval) of i on j, expressed by $\tilde{\alpha}_{ij}$, is high. In other words, in selecting the user-channel pairs (edges) for scheduling, preference should be given to those with higher expected channel rates in the current interval. Moreover, for stability of the system, we would prefer to choose user-channel pairs for which the corresponding queue-lengths are large. This intuitively justifies the term Q_{ij} in the weight of edge (i,j). A user i should transmit more packets on a channel j in which it has higher channel rate. The queue length Q_{ij} in such a channel will be low due to frequent service of packets. This justifies the selection of the least loaded queue for each new packet in the packet queueing policy.

C. Optimality Result

Theorem 2: The scheduling policy Ψ_T stabilizes the system for all arrival rate vectors $\vec{\lambda} \in \text{Int}(\Lambda_T)$, for any $T \geq 1$.

The above result (see the appendix for proof) states that our policy stabilizes the system for all arrival rate vectors that are strictly within the stability region. In other words, Theorem 2 states that our scheduling policy, Ψ_T, is throughput-optimal. Note that the throughput-optimality of the maximum queue-length matching based scheduling in input-queued switches, as shown in [15], follows as a special case of the above result, by considering the case $T = 1$ and $\alpha_{ij}(t) = 1 \ \forall i,j,t$.

D. Discussion

The design of Ψ_T also explains why the impact of infrequent channel measurements is fundamentally different from that of infrequent queue-length measurements. Note that the

packet service policy for Ψ_T depends on the products of the queue lengths and corresponding average channel rates. If the queue lengths known by the scheduler differ from the actual queue lengths by an amount that is upper bounded by a finite constant that independent of the queue-lengths (which is the case as long as the measurement intervals are finite), then the difference between the weights are also bounded since the channel rates are upper bounded by 1. This difference constitutes a negligible fraction of the actual weights for large queue lengths. However, when the channel rates are inaccurate, then the difference in the weights become arbitrarily large as queue lengths increase. This holds even when the inaccuracy in the channel weights is small. Thus the performance of the optimal strategy deteriorates primarily due to infrequent channel measurements.

The design of Ψ_T also demonstrates that using the current queue-lengths in the edge-weight computation is enough to ensure throughput-optimality. However, it can be shown that the use of the current channel rate in the edge-weight computation need not guarantee throughput-optimality; some measure of the average channel rate till the next measurement instant can be used instead, as in (2).

Finally, the assumption that channel and queue state measurements are made at the same time is mainly for the ease of exposition and analysis. Our results can be extended to the case where the channel and queue state measurements occur at different instants (and even different frequencies), as long as the measurement intervals remain bounded. In that case, the matching/poly-matching computation is done whenever the channel states (channel rates) are measured. In the edge-weight calculations, the last observed queue-lengths can be used. The calculation procedure of the average channel rate (see (2)) remains the same.

Wireless systems often transmit the same data (or interleave coded data) across multiple frequency channels to achieve frequency diversity. This variant of multi-channel transmission can be incorporated into our throughput-optimality framework as well, as discussed next. In this case, a valid channel assignment corresponds to a poly-matching, and it can be shown that the optimal channel assignment corresponds to finding a poly-matching Φ that maximizes an expression similar to that given by (2)-(3). However, the only (although very signicant) difference is that the channel rate α_{ij} in this case is a function of the poly-matching (Φ) itself (not just kT). Since $\alpha_{ij}(\Phi, k, T)$ (and therefore, $\tilde{\alpha}_{ij}(\Phi, k, T)$) will typically be a complex, non-linear function of Φ, optimizing the expression in (3) may be a computationally difcult problem.

E. Computational Aspects

The maximum weighted bipartite matching problem, also popularly known as the *assignment problem*, can be solved efficiently using the well-known *Hungarian Method* [13]. Let $m = |M|$ and $n = |N|$. Then the maximum weighted bipartite matching problem can be solved in $O(mn^2)$ time if $m \leq n$, and in $O(m^2 n)$ time if $m > n$.

The maximum weighted poly-matching can be computed as follows: each channel greedily selects the "best" user on that channel, irrespective of whether the user was selected by other channels or not. Thus during the $(k+1)$th interval, a channel j will select user i satisfying $i = \arg\max_{i' \in N} \tilde{\alpha}_{i'j}(kT) Q_{i'j}(kT)$ for receiving traffic on channel j. Note that in this assignment, a user can be assigned to multiple channels, but a channel can be assigned to at most one user. The algorithm requires $O(mn)$ time under sequential computation. However, note that the user selections across different channels are independent of each other, and can be executed in parallel; in that case, the algorithm can be completed in $O(n)$ time.

Our scheduling policy can be somewhat generalized, without affecting throughput optimality. More specifically, in the weight computation procedure, the queue length $Q_{ij}(kT)$ could be replaced by $f_{ij}(Q_{ij}(kT))$, where f_{ij} is some function of the queue-length. As long as the functions f_{ij} are strictly increasing, and satisfy some additional (fairly general) criteria, throughput optimality is achieved by our scheduling policy. The choice of the function, however, affects the average packet delays of different users. This fact can be exploited to provide delay differentiation to users. For example, if we use linear functions $f_{ij}(Q_{ij}(kT)) = w_i Q_{ij}(kT)$, we can attain delay differentiation by associating larger weights w_i with higher priority users. We explore this issue further through simulations in Section V.

F. Joint Scheduling and Power Allocation under Multi-channel Transmission

In a multi-channel transmission system, as mentioned earlier, data of a single user can be transmitted on multiple channels simultaneously. In the uplink case, the user might have fixed power budget per slot, which can be split across the different channels used by the user. The channel rates depend on the power allocation in these different channels. Also note that the optimal power allocation across different channels depends on the poly-matching chosen. In this scenario, therefore, the optimal scheduling and power allocation questions are closely coupled, and both scheduling and power allocation need to be jointly optimized for maximizing system throughput. We next show how our scheduling policy described earlier can be extended so as to solve this joint optimization question.

Let P_i denote the maximum power at which user i can transmit (over all channels). Let p_{ij} denote the transmission power used by user i on channel j in any time slot. Thus $\sum_{j \in M} p_{ij} \leq P_i$. We assume that a user i can transmit on any channel using only a finite number of power levels; let Ω_i denote the set of these power levels. Thus $p_{ij} \in \Omega_i \ \forall j = 1, 2, ..., M$. The constraints on the power allocations, p_{ij}, are

$$p_{ij} \in \Omega_i, \ \forall j \in M, \quad \forall i \in N, \tag{4}$$

$$\sum_{j \in M} p_{ij} \leq P_i, \quad \forall i \in N. \tag{5}$$

We assume that the channel rates are functions of the power allocation of user i on channel j. Thus $\alpha_{ij}(t)$, the channel rate of user i on channel j at time t, can be written as $\alpha_{ij}(t) = \hat{\alpha}_{ij}(p_{ij}, t)$. Typically, $\hat{\alpha}_{ij}$ is a concave function of p_{ij}.

The packet queuing policy remains the same as the one described in Section IV-B. Let $\bar{\alpha}_{ij}(kT)$ denote the average channel quality in the $(k+1)$th interval, derived using (2), while replacing $\alpha_{ij}(t)$ by $\hat{\alpha}_{ij}(p_{ij},t)$. Then our scheduling and power allocation policy for the $(k+1)$th interval involves finding the power allocations p_{ij}s and the poly-matching Υ so as to maximize

$$\sum_{(i,j)\in\Upsilon} \bar{\alpha}_{ij}(p_{ij},kT)Q_{ij}(kT), \tag{6}$$

where p_{ij}s must satisfy (4)-(5). We can show that the stability result (Theorem 2) holds in this case as well; a proof outline is provided in the appendix. It is worth noting, however, that computing the optimal power allocations and poly-matching that maximizes (6) is in general a difficult problem. Efficient computation of the optimal power allocations and the poly-matching for cases where the user-channel graph is large, remains an open question.

V. Simulation Studies

In this section, we evaluate the performance of our scheduling policy through simulations. More specifically, we demonstrate that our maximum weighted matching (poly-matching) based scheduling policy attains maximum achievable throughput when the channel/queue state measurement interval (T) is set to unity. We also study the reduction of the attained throughput, and the increase in the average delays, as the measurement interval T increases. Finally, we demonstrate that the maximum achievable throughput remains unaltered when the queue state measurement interval increases, provided the channel states are measured every time slot.

We consider downlink data transmission in an access point network with 6 users and 4 channels. We consider two channel models. In the first model, each channel has two states – "good" and "bad", and the channel rates associated with the two states are 1 and 0, respectively. In the second model, each channel has three states – "good", "intermediate" and "bad", and the rates associated with the three states are 0.9, 0.5 and 0.1, respectively. The state of each channel for each user varies in time according to a Markov chain, with a symmetric transition probability matrix. At any time slot, channel states (rates) for different channels or different users are independent of each other. The packet arrival process for each user is Bernoulli; packet arrival processes for different users are independent of each other. The nature of the simulation results for both the single-channel and multi-channel transmission cases are similar; therefore, we only present results for the single-channel transmission case.

Figure 3 shows how the average packet delay varies with the arrival rate, for different values of the measurement interval T, for the 2-state channel model. Note that we assume that the channel and queue state measurements, as well as the scheduling decisions, are made once every interval (of T time slots). The packet arrival rate for users $1,2,3$ is λ per user, and that for users $4,5,6$ is $\lambda/2$ per user. The figure shows that for a given arrival rate parameter λ, the average delay increases with an increase in T. Moreover, the maximum λ that can be supported (and therefore, the maximum throughput

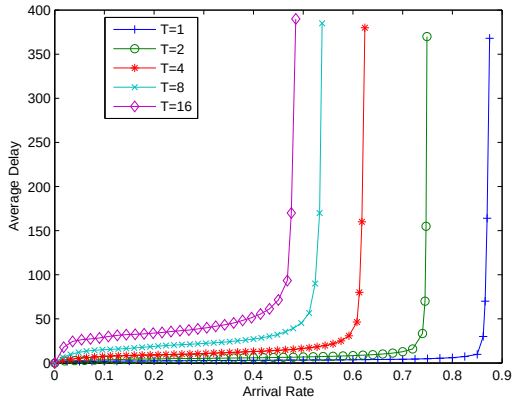


Fig. 3. Average packet delay vs. λ (2-state channel)

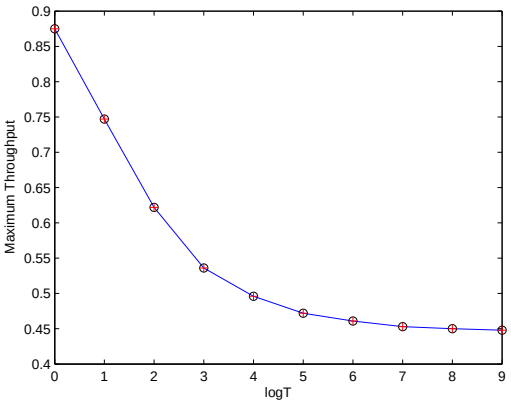


Fig. 4. Maximum attainable λ vs. Measurement interval T (2-state channel)

attained per user) decreases as T increases, as expected from Lemma 1. This is also evident from Figure 4, which plots the maximum supportable λ versus T, in a semi-log scale. Note that for $T=1$, the maximum attainable λ in this case can easily be calculated as $(8/9)\approx 0.889$. Figures 5-6 are similar to Figures 3-4, but for the 3-state channel model. The trends observed in this case are also similar to the ones discussed above.

Let us next explore how delay differentiation can be attained by associating different weights with different queues. In Figure 7 we plot the average delay vs. λ curves in the case where user 1 is associated with a higher weight than the rest of the users. More specifically, in computing the matching/poly-matching, user 1's queue-length in multiplied by a factor of 4, while the weight calculations for the other users remain unaltered. The average delay vs. λ curve for the undifferentiated case (where all users are associated with equal weight, and therefore treated uniformly) is also shown in the figure. Figure 7 shows that with this weight-based differentiation, the average delay of user 1 decreases, while that of the other users increases, compared to the undifferentiated case. However, note that the maximum attainable λ remains the same.

Now, let us consider the case where the channel state measurement and scheduling decisions are made every time

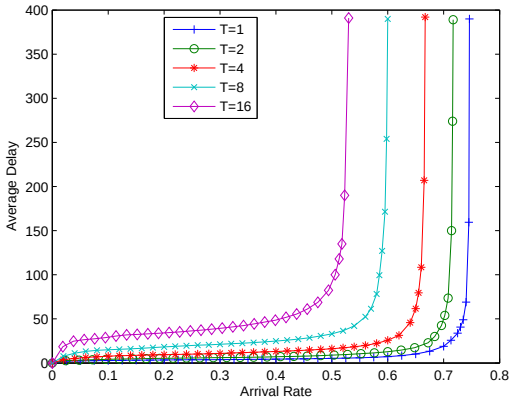


Fig. 5. Average packet delay vs. λ (3-state channel)

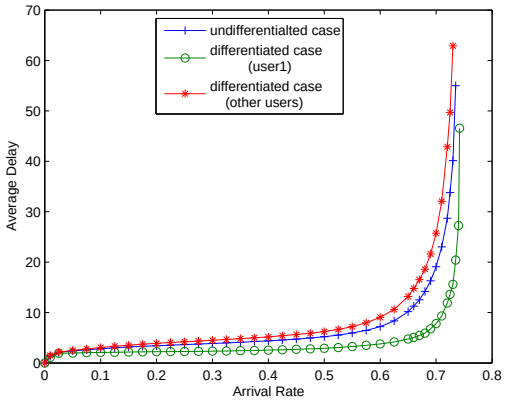


Fig. 7. Delay differentiation: Average packet delay vs. λ (3-state channel)

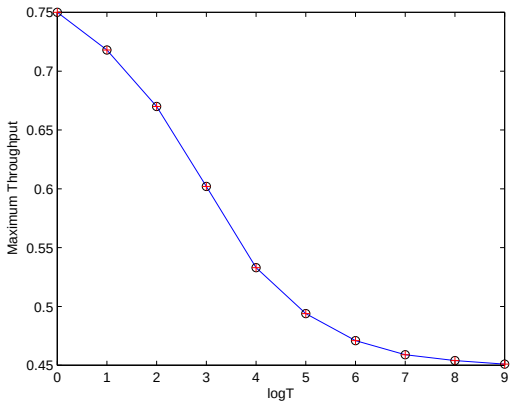


Fig. 6. Maximum attainable λ vs. Measurement interval T (3-state channel)

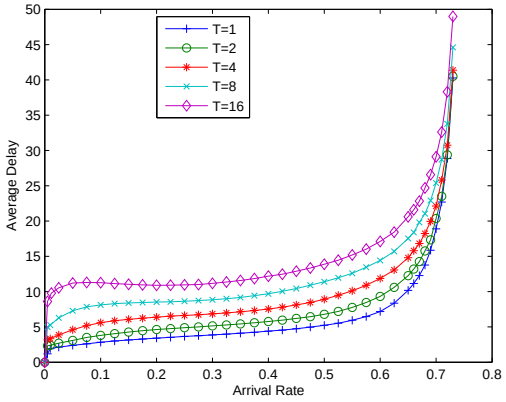


Fig. 8. Average packet delay vs. λ, when only queue state measurements are made infrequently (3-state channel)

slot, but the queue measurement decisions are made only once every T-slot interval. Figure 8 plots the average packet delay versus λ in this scenario, for different values of T. The plots demonstrate that in this scenario, the maximum achievable throughput does not change as T increases, as discussed in Section IV-A. The plots also demonstrate that our weighted matching based scheduling algorithm attains the maximum achievable throughput, for every value of T considered.

Finally we consider the joint scheduling and power allocation problem; figures 9-10 show the average delay and maximum attainable λ in this case, for 3 users and 2 channels. Here, the channel rate α_{ij} is expressed as $\alpha_{ij} = B\log(1+\kappa\frac{p_{ij}}{n_{ij}})$, where B and κ are constants, p_{ij} represents the transmission power allocated on channel j by user i, and n_{ij} is the noise power on channel j for user i. We assume that there are three noise power level values $-$ 0.1, 0.5, 0.9, and the noise power levels vary according to a Markov chain with a symmetric transition probability matrix. The maximum power P_i is unity for each user i, and the transmission power p_{ij} can be chosen from three different levels $-$ 0, 0.5 and 1. We consider non-uniform traffic, where the packet arrival rate for users 1 and 2 is λ per user, and that for user 3 is $\lambda/2$. The optimal power and channel assignments (which maximize (6) subject to (4)-(5)) are computed by complete enumeration over all possible power levels and channel allocations. Figures 9-10 show a trend similar to the cases without power allocation discussed earlier.

VI. Concluding Remarks

We have presented a throughput-optimal uplink/downlink scheduling policy in a multichannel wireless access point network where the time-varying channel rates can be measured only infrequently. We identified a fundamental disparity between the roles played by the queue and channel state measurements: less frequent queue-length measurements do not affect the maximum throughput achieved, but a reduction in the channel rate measurement frequency reduces the maximum achievable throughput. Finally, we have also shown how our approach can be used for joint optimization of power allocation and scheduling in a multi-channel transmission system. Computationally efficient approaches of computing the optimal power allocations and schedules in this case remains an open question.

Appendix

Proof of Lemma 1

Proof: Consider any positive integer l. Choose any $\vec{\lambda} \in \Lambda_{lT}$. Then there exists a scheduling policy, say $\hat{\Psi}$, that achieves

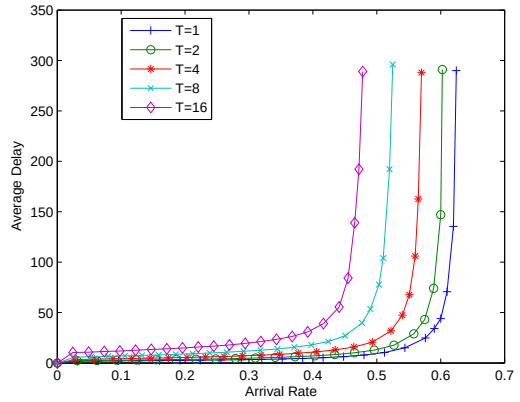


Fig. 9. Average packet delay vs. λ (joint power allocation and scheduling)

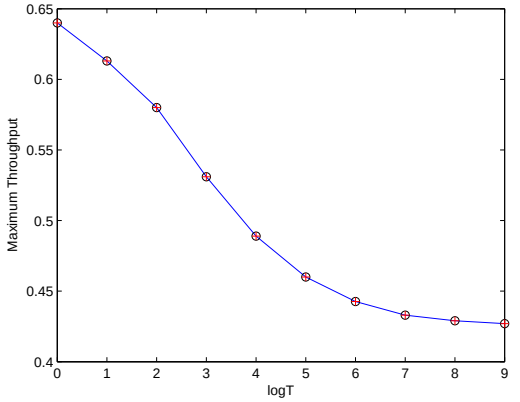


Fig. 10. Maximum attainable λ vs. Measurement interval T (joint power allocation and scheduling)

stability for the arrival rate vector $\vec{\lambda}$, in a system where the measurement interval has length lT, i.e., measurements are made only at time slots kT for which k is a multiple of l. Now consider using $\hat{\Psi}$ in the system where the measurement interval has length T, i.e. measurements are made only at time slots kT, for $k = 0, 1, \ldots$. In this case, our policy simply ignores the measurements made at time slot kT (and keeps using the previously computed matching/poly-matching), unless k is a multiple of l. Clearly, this policy will also achieve stability for the arrival rate vector $\vec{\lambda}$, in the system where the measurement interval has length T. Therefore $\vec{\lambda} \in \Lambda_T$. Since $\vec{\lambda}$ was chosen arbitrarily from Λ_{lT}, we conclude $\Lambda_{lT} \subseteq \Lambda_T$.

We now provide an example scenario where $\text{Int}(\Lambda_{lT}) \subset \text{Int}(\Lambda_T)$. Consider a single-channel transmission system with $T = 1, l = 2$, 1 user and 2 channels. In any slot, for both channels j, α_{1j} is a Bernoulli random variable with probability p of failure, i.e., $\alpha_{1j} = 1$ with probability $1 - p$, and 0 otherwise. Let $0 < p < 1$. Consider measurement intervals of size T, i.e., when the channel rates and queue lengths are measured every slot. Since the channels are statistically identical and there is only one user, it can be shown that in each slot the optimal policy is to transmit a packet in any channel that has rate 1 provided the user has a packet to transmit. Thus, the user does not transmit only when it does not have a packet to transmit or both channels have rate 0. This policy can stabilize the system as long as the arrival rate λ of the user is less than $1 - p^2$. Thus, the interior of the stability region Λ_T is given by $0 < \lambda < 1 - p^2$. Now, consider measurement intervals of size lT, i.e., channel measurements are done in alternate slots. Again, since the channels are statistically identical and there is only one user, the optimal policy is to select a channel that has rate 1 in the slot in which the channel is measured, and transmit packets in the same channel during the interval while the user has a packet to transmit. This policy can stabilize the system as long as the arrival rate λ of the user is less than $1-(p^2+p)/2$. The interior of the stability region Λ_{lT} is given by $0 < \lambda < 1-(p^2+p)/2$. Clearly, this region is a proper subset of $0 < \lambda < 1-p^2$. Thus, $\text{Int}(\Lambda_{lT}) \subset \text{Int}\Lambda_T$. ∎

Proof Outline of Theorem 2

Preliminaries: First we introduce some notation which will be used in the proof. Let $\vec{\alpha}(t) = (\alpha_{ij}(t), i \in N, j \in M)$ denote the vector of channel rates at time t. Let $p_{\vec{\theta}} = \Pr\left(\vec{\alpha}(t) = \vec{\theta}\right)$, denote the stationary probabilities of the Markov chain of the channel rate vector $\vec{\alpha}(t)$. Let $\Theta = \{\vec{\theta} : p_{\vec{\theta}} > 0\}$, and $|\Theta|$ be finite.

In the following, we use a vector representation of matching (poly-matching) Φ, where Φ is represented as a NM-dimensional vector $\vec{\phi}$ with components ϕ_{ij}, where

$$\phi_{ij} = \begin{cases} 1 & \text{if channel } j \text{ is used by user } i, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that $\vec{\alpha}(lT), l = 0, 1, \ldots$ constitutes a positive recurrent Markov chain with stationary probabilities $p_{\vec{\theta}}, \theta \in \Theta$. Let $\vec{Q}(t) = (Q_{ij}(t), i \in N, j \in M)$ denote the vector of queue-lengths at time t.

Let us consider the $(l+1)$th interval, i.e. the interval $[lT, ..., (l+1)T - 1]$, for any non-negative integer l. Define $\vec{\theta} = \vec{\alpha}(lT)$, i.e., $\vec{\theta}$ is the vector of the channel rates at the beginning of the interval. Let $\gamma^{\vec{\phi},\vec{\theta}} = (\gamma_{ij}^{\vec{\phi},\vec{\theta}}, i \in N, j \in M)$, denote the vector of expected throughputs in that interval, if matching (poly-matching) $\vec{\phi}$ is chosen, and if all queues are continuously backlogged during that interval. Define $H_{\vec{\theta}}$ as the set of $\vec{\gamma}^{\vec{\phi},\vec{\theta}}$ for all possible $\vec{\phi}$ when $\vec{\theta}$ is the vector of channel rates at the beginning of the chosen interval.

Let $\vec{D}^{l+1}$ be a NM-dimensional vector representing the number of packet departures from the different queues in the $(l+1)$th interval. Also, let $\vec{A}^{l+1}$ be a NM-dimensional vector representing the number of arrivals entering the different queues at the beginning of the $(l+2)$th interval. (Recall that the packets entering the queues at the beginning of the $(l+2)$th interval are those that arrive during the $(l+1)$th interval.) Let a denote an upper bound on the number of arrivals, and the number of departures, in any interval. For simplicity of exposition, we prove Theorem 2 under an additional restriction (R) on the scheduling policy: if $Q_{ij}(lT) < aT$, then no packets of user i is scheduled on channel j during the $(l+1)$th interval. It should however be noted that the proof presented here can be extended to work even in the absence of restriction R.

Let $\vec{\phi}(lT)$ denote the matching (poly-matching) selected at the beginning of the $(l+1)$th interval, i.e., the matching (poly-matching) selected at time lT. Let $J(lT) = \{\vec{\phi} : \vec{\phi}_{ij} = 0 \text{ if } Q_{ij}(lT) < aT\}$. Then, for our scheduling policy Ψ_T under restriction R, we have

$$\vec{\phi}(lT) = \arg \max_{\vec{\phi} \in J(lT)} \left(\vec{Q}(lT)\right)^T \vec{\gamma}^{\vec{\phi},\vec{\alpha}(lT)}, \tag{7}$$

where we use $(\cdot)^T$ to denote the transpose of a vector, with slight abuse of notation.

We now proceed with the proof of Theorem 2.

Proof: First we characterize the interior of the stability region Λ_T. A rate vector $\vec{\lambda} \in \text{Int}(\Lambda_T)$ if there exist non-negative real numbers μ_{ij}s and $\beta_{\vec{\gamma}\vec{\theta}}$s such that

$$\lambda_i = \sum_{j \in M} \mu_{ij} \ \forall i, \tag{8}$$

$$\vec{\mu} = (1/T) \sum_{\vec{\theta} \in \Theta} p_{\vec{\theta}} \sum_{\vec{\gamma} \in H_{\vec{\theta}}} \beta_{\vec{\gamma}\vec{\theta}} \vec{\gamma}, \tag{9}$$

$$\sum_{\vec{\gamma} \in H_{\vec{\theta}}} \beta_{\vec{\gamma}\vec{\theta}} < 1 \ \ \forall \ \vec{\theta} \in \Theta. \tag{10}$$

Note that in (9), $\vec{\mu} = (\mu_{ij}, i \in N, j \in M)$ denotes the vector of the μ_{ij}s.

We define $V(t)$ as $V(t) = \left(\vec{Q}(t)\right)^T \vec{Q}(t)$.

Let j be a non-negative integer. We will show that there exist a positive integer k and a negative real number K such that $E\left(V\left((j+k)T\right) - V(jT) / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu}\right) < K$ for all $\vec{\nu}$, whenever $\vec{\lambda} \in \text{Int}(\Lambda_T)$, and $||\vec{Q}(jT)||$ is sufficiently large.

Let $q_{\vec{\theta}}^{\vec{\nu},i}$ be the probability that $\vec{\alpha}\left((j+i)T\right) = \vec{\theta}$ given that $\vec{\alpha}(jT) = \vec{\nu}$. Let

$$\epsilon = \max_{\vec{\theta},\vec{\nu} \in \Theta} \left| \frac{\sum_{i=1}^{k} q_{\vec{\theta}}^{\vec{\nu},i}}{k p_{\vec{\theta}}} - 1 \right|. \tag{11}$$

Let k be a large enough integer such that

$$\epsilon < 1 - \max_{\vec{\theta} \in \Theta} \sum_{\vec{\gamma} \in H_{\vec{\theta}}} \beta_{\vec{\gamma}\vec{\theta}}. \tag{12}$$

Clearly, there exists one such k since $\sum_{\vec{\gamma} \in H_{\vec{\theta}}} \beta_{\vec{\gamma}\vec{\theta}} < 1 \ \forall \ \vec{\theta} \in \Theta$ (from (10)), $|\Theta|$ is finite, and $p_{\vec{\theta}}, \vec{\theta} \in \Theta$, is the stationary distribution of the positive recurrent Markov chain $\vec{\alpha}(lT), l = 0, 1, \ldots$ Clearly,

$$\vec{Q}\left((j+k)T\right) = \vec{Q}(jT) + \sum_{i=1}^{k} \vec{A}^{j+i} - \sum_{i=1}^{k} \vec{D}^{j+i}.$$

Then,

$$\begin{aligned} & E\left(V\left((j+k)T\right) - V(jT) / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu}\right) \\ = \ & 2 \sum_{i=1}^{k} E\left\{ \left(\vec{Q}(jT)\right)^T \left(\vec{A}^{j+i} - \vec{D}^{j+i}\right) / \right. \\ & \left. \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\} \\ & + \sum_{i=1}^{k} E\left\{ \left(\vec{A}^{j+i} - \vec{D}^{j+i}\right)^T \left(\vec{A}^{j+i} - \vec{D}^{j+i}\right) / \right. \\ & \left. \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\}. \end{aligned}$$

Since the number of arrivals and departures in any interval is bounded (by a), there exists a constant K_1 such that for any $\vec{\nu}$,

$$\sum_{i=1}^{k} E\left\{ \left(\vec{A}^{j+i} - \vec{D}^{j+i}\right)^T \left(\vec{A}^{j+i} - \vec{D}^{j+i}\right) / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\} < kTK_1.$$

Thus,

$$\begin{aligned} & E\left(V\left((j+k)T\right) - V(jT) / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu}\right) \\ < \ & 2\left(\vec{Q}(jT)\right)^T E\left\{ \sum_{i=1}^{k} \left(\vec{A}^{j+i} - \vec{D}^{j+i}\right) / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\} \\ & + kTK_1 \\ = \ & 2\left(\vec{Q}(jT)\right)^T E\left\{ \sum_{i=1}^{k} \vec{A}^{j+i} / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\} \\ & - 2\left(\vec{Q}(jT)\right)^T E\left\{ \sum_{i=1}^{k} \vec{D}^{j+i} / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\} \\ & + kTK_1. \end{aligned} \tag{13}$$

We can derive the following inequalities:

$$\begin{aligned} & \left(\vec{Q}(jT)\right)^T E\left\{ \sum_{i=1}^{k} \vec{A}^{j+i} / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\} \\ & \leq \ kT \left(\vec{Q}(jT)\right)^T \vec{\mu} + f_1(k), \end{aligned} \tag{14}$$

$$\begin{aligned} & \left(\vec{Q}(jT)\right)^T E\left\{ \sum_{i=1}^{k} \vec{D}^{(j+i)} / \vec{Q}(jT), \vec{\alpha}(jT) = \vec{\nu} \right\} \\ & \geq \ \sum_{i=1}^{k} \sum_{\vec{\theta} \in \Theta} q_{\vec{\theta}}^{\vec{\nu},i} \max_{\vec{\gamma} \in H_{\vec{\theta}}} \left(\vec{Q}(jT)\right)^T \vec{\gamma} - f_2(k), \end{aligned} \tag{15}$$

where $f_1(k) = aT^2k^2(aMN + \sum_{l=1}^{N} \lambda_l)$ and $f_2(k) = aTMNk^2 + 2a^2T^2kMN$. Note that $f_1(k)$ and $f_2(k)$ are both positive terms. The derivations of (14) and (15) are rather tedious, and can be found in [11]; these are omitted here due to space constraints. Both (14) and (15) derived using the fact that the arrivals/departures in any slot is upper bounded by a. Furthermore, (14) is derived using the fact that arrivals are routed to the queue with the smallest length, as in (1). On the other hand, (15) is derived using the fact that our scheduling policy corresponds to maximum weight matching/poly-matching, as in (3). Intuitively, the term $f_1(k)$

can be viewed as a "correction factor" needed due to the deviation of the expected arrival rate in the k intervals considered, from the arrival rate vector $\vec{\lambda}$. Similarly, the term $f_2(k)$ can be intuitively viewed as a "correction factor" needed due to the inaccuracy in the scheduling policy, resulting from infrequent measurements/scheduling, and restriction R.

From (14) and (9),

$$\begin{aligned}
&\left(\vec{Q}(jT)\right)^T E\left\{\sum_{i=1}^{k}\vec{A}^{t+i}/\vec{Q}(jT),\vec{\alpha}(jT)=\vec{\nu}\right\}\\
&\leq\ k\left(\vec{Q}(jT)\right)^T\sum_{\vec{\theta}\in\Theta}p_{\vec{\theta}}\sum_{\vec{\gamma}\in H_{\vec{\theta}}}\beta_{\vec{\gamma}\vec{\theta}}\vec{\gamma}+f_1(k)\\
&=\ k\sum_{\vec{\theta}\in\Theta}p_{\vec{\theta}}\sum_{\vec{\gamma}\in H_{\vec{\theta}}}\beta_{\vec{\gamma}\vec{\theta}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}+f_1(k)\\
&\leq\ k\sum_{\vec{\theta}\in\Theta}p_{\vec{\theta}}\sum_{\vec{\gamma}\in H_{\vec{\theta}}}\beta_{\vec{\gamma}\vec{\theta}}\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}+f_1(k)\\
&\leq\ k\left(\max_{\vec{\theta}\in\Theta}\sum_{\vec{\gamma}\in H_{\vec{\theta}}}\beta_{\vec{\gamma}\vec{\theta}}\right)\sum_{\vec{\theta}\in\Theta}p_{\vec{\theta}}\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}\\
&\quad+f_1(k). \qquad (16)
\end{aligned}$$

From (15) and (11),

$$\begin{aligned}
&\left(\vec{Q}(jT)\right)^T E\left\{\sum_{i=1}^{k}\vec{D}^{(j+i)}/\vec{Q}(jT),\vec{\alpha}(jT)=\vec{\nu}\right\}\\
&\geq\ \sum_{i=1}^{k}\sum_{\vec{\theta}\in\Theta}q_{\vec{\theta}}^{\vec{\nu},i}\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}-f_2(k)\\
&=\ \sum_{\vec{\theta}\in\Theta}\sum_{i=1}^{k}q_{\vec{\theta}}^{\vec{\nu},i}\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}-f_2(k)\\
&=\ \sum_{\vec{\theta}\in\Theta}\left(\sum_{i=1}^{k}q_{\vec{\theta}}^{\vec{\nu},i}\right)\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}-f_2(k)\\
&\geq\ k(1-\epsilon)\sum_{\vec{\theta}\in\Theta}p_{\vec{\theta}}\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}-f_2(k). \qquad (17)
\end{aligned}$$

From (13), (16) and (17), for any $\vec{\nu}$ and sufficiently large $||\vec{Q}(jT)||$,

$$\begin{aligned}
&E\left(V\left((t+k)T\right)-V(jT)/\vec{Q}(jT),\vec{\alpha}(jT)=\vec{\nu}\right)\\
&<\ -2k(1-\epsilon-\max_{\vec{\theta}\in\Theta}\sum_{\vec{\gamma}\in H_{\vec{\theta}}}\beta_{\vec{\gamma}\vec{\theta}})\times\sum_{\vec{\theta}\in\Theta}p_{\vec{\theta}}\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(jT)\right)^T\vec{\gamma}\\
&\quad+kTK_1+f_1(k)+f_2(k).
\end{aligned}$$

From the last inequality and (12), we see that $E\left(V\left((j+k)T\right)-V(jT)/\vec{Q}(jT),\vec{\alpha}(jT)=\vec{\nu}\right) < K$ for some negative number K, when $||\vec{Q}(jT)||$ is sufficiently large. From standard results in stochastic stability (see pages 330-331 of [17]), it now follows that the expected queue-lengths are bounded. ∎

Proof Outline of Observation 2

Clearly, $\Lambda_1 = \Lambda$. Thus, we only need to show that $\mathrm{Int}(\hat{\Lambda}_T) = \mathrm{Int}(\Lambda_1)$ for all $T \geq 1$. Consider the policy $\hat{\Psi}_T$ which is the same as Ψ_1 except that it measures queue lengths at the beginning of the first slot of each interval of T slots, and uses this measurement to compute the channel assignment to be used over the entire interval. Consider the proof for Theorem 2 with $T = 1$. All the arguments in this proof, except possible equation (15), hold for $\hat{\Psi}_T$ irrespective of the value of T. Since $\hat{\Psi}_T$ does not select the schedule $\vec{\gamma}$ that attains $\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(j)\right)^T\vec{\gamma}$ in each slot j, it is not clear that (15) holds. Nevertheless, as we argue next, (15) still holds in this case with a different value for $f_2(k)$, which does not depend on the queue lengths. Note that the proof holds as long as $f_2(k)$ does not depend on the queue lengths, irrespective of the exact value of $f_2(k)$. Thus, the proof holds in this case as well.

Note that $\hat{\Psi}_T$ selects the schedule $\vec{\gamma}$ that attains $\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(\hat{j})^T\vec{\gamma}\right)$ in each slot j, where $\hat{j} = \lfloor j/T \rfloor T$ is the first slot of the queue length measurement interval containing j). Now, $|Q_{xy}(j) - Q_{xy}(j-1)| \leq a$ for any user x and channel y, since the number of arrivals and departures in any slot is bounded by a. Since $j - \hat{j} \leq T$, $|Q_{xy}(j) - Q_{xy}(\hat{j})| \leq (j-\hat{j})a \leq Ta$, for each x, y. Thus, $\max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(\hat{j})^T\vec{\gamma}\right) \geq \max_{\vec{\gamma}\in H_{\vec{\theta}}}\left(\vec{Q}(j)^T\vec{\gamma}\right) - MNTa^2$. Thus, (15) now holds with $f_2(k) = aMNk^2 + 2a^2kMN + MNTa^2$. We obtain this expression for $f_2(k)$ by setting the "channel measurement interval T" to 1 in the proof for Theorem 2, and augmenting the resulting expression with $MNTa^2$ as per the above discussions, where T is the queue length measurement interval in this case. ∎

Proof Outline of Theorem 2 for Joint Scheduling and Power Allocation

The proof is similar to that for Theorem 2. The only difference is that now, ϕ_{ij} is used to denote the transmission power used by user i on channel j. If the pair (i, j) is not selected in the poly-matching, i.e., $(i, j) \notin \Upsilon$, then $\phi_{ij} = 0$. Thus, ϕ_{ij} also specifies the poly-matching. The constraints on the transmission power determines the set of $\vec{\phi}$ that can be used in any given slot. Now, as before, $\gamma^{\vec{\phi},\vec{\theta}} = (\gamma_{ij}^{\vec{\phi},\vec{\theta}}, i \in N, j \in M)$, denotes the vector of expected throughputs in that interval, if $\vec{\phi}$ is chosen, and if all queues are continuously backlogged during that interval; similarly, $H_{\vec{\theta}}$ denotes the set of $\vec{\gamma}^{\vec{\phi},\vec{\theta}}$ for all possible $\vec{\phi}$ when $\vec{\theta}$ is the vector of channel rates at the beginning of the chosen interval. The rest of the proof for Theorem 2 can now be used as is. ∎

References

[1] M. Armony, N. Bambos, Queueing Networks with Interacting Service Resources, *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing,* pp. 42-51, Allerton Park, Monticello, IL, September, 1999.

[2] M. Armony, N. Bambos, Queueing Dynamics and Maximal Throughput Scheduling in Switched Processing Systems, Queueing Systems, 44, 209-252, 2003.

[3] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein, *Introduction to Algorithms*, 2nd Ed., MIT Press, 2001.

[4] JG Dai, B. Prabhakar, "The throughput of data switches with and without speedup", *Proceedings of the IEEE INFOCOM 2000*, Tel Aviv, Israel, March 2000.

[5] Mustafa Ergen, Sinem Coleri, and Pravin Varaiya, "QoS Aware Adaptive Resource Allocation Techniques for Fair Scheduling in OFDMA Based Broadband Wireless Access Systems", *IEEE Trans on Broadcasting* 94(4): 362-370, Dec. 2003.

[6] A. Eryilmaz and R. Srikant, "Fair Resource Allocation in Wireless Networks using Queue-length based scheduling and Congestion Control", *Proceedings of INFOCOM 2005*, Miami, March 2005.

[7] B. Hajek and G. Sasaki, "Link Scheduling in Polynomial Time", *IEEE Transactions on Information Theory*, vol. 34, no. 5, 1988, pp. 910-917.

[8] W.-C. Hung and K.L. Eddie Law, A. Leon-Garcia, "A Dynamic Multi-Channel MAC for Ad-Hoc LAN", *21st Biennial Symposium on Communications*, Kingston, Canada, Jun 2002.

[9] N. Jain and S. R. Das and A. Nasipuri, "A Multichannel MAC Protocol with Receiver-Based Channel Selection for Multihop Wireless Networks", *9th Int. Conf. on Computer Communications and Networks (IC3N)*, Phoenix, Oct 2001.

[10] T. Javidi, "Rate Stable Resource Allocation in OFDM Systems: From Waterfilling to Queue-Balancing," *Allerton Conference on communication, Control and Computing*, September 2004.

[11] K. Kar, X. Luo and S. Sarkar, "Throughput-optimal Scheduling in Multichannel Access Point Networks under Time-Varying Channel Rates", *Technical Report*, Rensselaer Polytechnic Institute, March 2005. http://www.ecse.rpi.edu/∼koushik/thruput-opt.pdf.

[12] S. Kittipiyakul and T. Javidi, "Subcarrier allocation in OFDMA, systems: beyond water-filling", *Asilomar Conference on Signals, Systems and Computers*, November 2004.

[13] H. W. Kuhn, "The Hungarian Method for the assignment problem", *Naval Research Logistic Quarterly*, 2 (1955), pp. 83-97.

[14] P. Kyasanur and N. H. Vaidya, "Routing and interface assignment in multi-channel multi-interface wireless network", *Proceedings of IEEE WCNC 2005*.

[15] N. McKeown, V. Anantharam and J. Walrand, "Achieving 100% Throughput in an Input-Queued Switch", *Proceedings of IEEE Infocom '96*, San Francisco, Mar 1996.

[16] A. Muir and J.J. Garcia-Luna-Aceves, "Channel-Hopping Multiple Access", *International Conf. on Communications (ICC)*, Jun 2000.

[17] S. Meyn and R. Tweedie. *Markov Chains and Stochastic Stability.* Springer Verlag, 1994.

[18] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic Power Allocation and Routing for Time Varying Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Special Issue on Wireless Ad-Hoc Networks, vol. 23, no. 1, pp. 89-103, Jan. 2005.

[19] M. Neely and E. Modiano and C. Li, "Fairness and Optimal Stochastic Control for Heterogeneous Networks", *Proceedings of IEEE INFOCOM 2005*, Miami, March 2005.

[20] K. Ross and N. Bambos, "Dynamic Quality of Service Control in Packet Switch Scheduling", *Proceedings of IEEE International Conference on Communications 2005*, Seoul, Korea, May 2005.

[21] K. Ross and N. Bambos, "Packet Scheduling Across Networks of Switches", *Proceedings of IEEE International Conference on Communications 2005*, Seoul, Korea, May 2005.

[22] K. Ross and N. Bambos, "Optimizing Quality of Service in Prioritized Packet Switch Scheduling", *Proceedings of IEEE International Conference on Communications 2004*, Paris, France, June 2004.

[23] K. Ross and N. Bambos, "Local Search Scheduling Algorithms for Maximal Throughput in Packet Switches Optimizing", *Proceedings of IEEE INFOCOM 2004*, Hong Kong, China, March 2004.

[24] K. Ross, N. Bambos, K. Kumaran, I. Saniee, I. Wadjaja, "Scheduling Bursts in Time-Domain Wavelength Interleaved Networks", *IEEE Journal on Selected Areas in Communications; Optical Communications and Networking Series*, November 2003.

[25] K. Ross and N. Bambos, "Dynamic Scheduling of Optical Data Bursts in Time-Domain Wavelength Interleaved Networks", *Proceedings of IEEE Symposium on High Performance Interconnects 2003*, August, 2003.

[26] K. Ross and N. Bambos, "Projective Cone Schedules in Queueing Structures; Geometry of Packet Scheduling in Communication Network Switches", *Allerton Conference on Communication, Control and Computing*, Monticello, IL, October, 2002.

[27] S. Sarkar and L. Tassiulas, "A Framework for Routing and Congestion Control for Multicast Information Flows", *IEEE Transactions on Information Theory*, vol. 48, no. 10, 2002, pp. 2690-2708.

[28] J. So and N. H. Vaidya, "Multi-channel mac for ad hoc networks: Handling multi-channel hidden terminals using a single tranceiver", *Proceedings of ACM MOBIHOC 2004*.

[29] A. Stolyar and S. Shakkottai and R. Srikant, "Pathwise optimality of the exponential scheduling rule for wireless channels", *Advances in Applied Probability*, vol. 36, no. 4, 2004, pp. 1021-1045.

[30] L. Tassiulas, "Adaptive back-pressure congestion control based on local information", *IEEE Transactions on Automatic Control*, vol. 40, no. 2, 1995.

[31] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling for maximum throughput in multihop radio networks", *IEEE Transactions on Automatic Control*, vol. 37, no. 12, 1992, pp. 1936-1949.

[32] A. Tzamaloukas and J. J. Garcia-Luna-Aceves, "A Receiver-Initiated Collision-Avoidance Protocol for Multi-channel Networks", *Proceedings of IEEE INFOCOM 2001*.

[33] Ian C. Wong, Zukang Shen, Brian L. Evans, and Jeffrey G. Andrews, "A Low Complexity Algorithm for Proportional Resource Allocation in OFDMA Systems", *IEEE Workshop on Signal Processing Systems (SIPS04)*, October 13-15, 2004.

Koushik Kar (M'02) received the Ph.D. and M.S. degrees in Electrical and Computer Engineering from the University of Maryland, College Park, in 2002 and 1999, respectively. He received the B.Tech. degree in Electrical Engineering from Indian Institute of Technology, Kanpur, in 1997. Since 2002, he has been an assistant professor in the Electrical, Computer and Systems Engineering department at Rensselaer Polytechnic Institute, Troy, NY. His research interests include performance optimization questions in ad-hoc and sensor networks, traffic engineering, congestion control and multicasting. Dr. Kar received the CAREER Award from the National Science Foundation in 2005.

Xiang Luo (S'08) Xiang received his B.E. and M.S. degrees, both in electrical engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2001 and 2004, respectively. He is currently a Ph.D. student in the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute (RPI), Troy, NY. His research interests include medium access control and link-layer scheduling in wireless networks.

Saswati Sarkar (S'98, M'00) received Master of Engineering in Electrical Communication Engineering from the Indian Institute of Science, Bangalore in 1996 and Ph.D. in Electrical and Computer Engineering from University of Maryland, College Park in 2000. She is currently an Associate Professor in the department of Electrical and Systems Engineering in University of Pennsylvania. Her research interests are in resource allocation and performance analysis in communication networks. She received the Motorola gold medal for the best masters student in the division of electrical sciences at the Indian Institute of Science and a National Science Foundation (NSF) Faculty Early Career Development Award in 2003. She was an associate editor of IEEE Transaction on Wireless Communications from 2001 to 2006.