Scheduling with Limited Information in Wireless Systems

Prasanna Chaporkar
Dept. of Electrical Engineering
Indian Inst. of Technology, Mumbai
chaporkar@ee.iitb.ac.in

Alexandre Proutiere
Microsoft Research
Cambridge, UK
alexandre.proutiere@microsoft.com

Himanshu Asnani
Dept. of Electrical Engineering
Indian Inst. of Technology, Mumbai
himanshu_asnani@iitb.ac.in

Abhay Karandikar
Dept. of Electrical Engineering
Indian Inst. of Technology, Mumbai
karandi@ee.iitb.ac.in

ABSTRACT

Opportunistic scheduling is a key mechanism for improving the performance of wireless systems. However, this mechanism requires that transmitters are aware of channel conditions (or CSI, Channel State Information) to the various possible receivers. CSI is not automatically available at the transmitters, rather it has to be acquired. Acquiring CSI consumes resources, and only the remaining resources can be used for actual data transmissions. We explore the resulting trade-off between acquiring CSI and exploiting channel diversity to the various receivers. Specifically, we consider a system consisting of a transmitter and a fixed number of receivers/users. An infinite buffer is associated to each receiver, and packets arrive in this buffer according to some stochastic process with fixed intensity. We study the impact of limited channel information on the stability of the system. We characterize its stability region, and show that an adaptive queue length-based policy can achieve stability whenever doing so is possible. We formulate a Markov Decision Process problem to characterize this queue length-based policy. In certain specific and yet relevant cases, we explicitly compute the optimal policy. In general case, we provide a scheduling policy that achieves a fixed fraction of the system's stability region. Scheduling with limited information is a problem that naturally arises in cognitive radio systems, and our results can be used in these systems.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless communication

General Terms

Algorithms, Design, Performance

MobiHoc'09, May 18–21, 2009, New Orleans, Louisiana, USA.

Keywords

Throughput optimality, queue stability, optimal stopping time, cognitive radio

1. INTRODUCTION

Wireless systems have limited resources, and they must be exploited optimally. Recently, opportunistic scheduling has emerged as an attractive solution for improving the efficiency of these systems. The basic principle behind opportunistic scheduling is to exploit the independent variations of fading experienced by different interfering links/users [16, 3]. The throughput improvement is achieved by always scheduling users with relatively favorable radio conditions. When the number of users grows large, the throughput gain may become significant since each time a scheduling decision is made, there is, with high probability, a user with better radio conditions than average.

The design and the performance of opportunistic scheduling policies have been extensively explored under various network and traffic scenarios. In their seminal papers [25, 24], Tassiulas et al. have proposed the opportunistic *max-weight* scheduling strategy and have shown that it could ensure stability of the user buffers (i.e., the expected backlog is finite) whenever this is at all possible. This schemes has been generalized to adapt to various network scenarios, see e.g. [26]. Opportunistic scheduling strategies providing delay guarantees have also been proposed and analyzed, see e.g. [2, 23, 21, 18]. Finally, in wireless systems providing data *elastic* services, e.g. CDMA 1EvDo [3] and UMTS/HSDPA [1], opportunistic and fair schedulers have been proposed [3, 13, 5, 17, 27, 22] and have been actually implemented in 3G cellular systems. There, fairness is formalized through the notion of user utility.

To achieve the promised substantial performance gains, the transmitter (an access point or a base station) implementing opportunistic scheduling has to be aware of the radio conditions or CSI (Channel state Information) of the various users so as to take the right scheduling decision. The channel states are by default not known and they have to be acquired. In 3G cellular systems, users report their CSI to the base station using dedicated uplink channels (one uplink channel per user), which can be quite costly in terms of radio resources (spectrum and power). The amount of resources used to acquire CSI is proportional to the number of users whose channels have to be known, and hence one has to carefully evaluate the trade-off between these wasted resources and the performance gain achieved by employing opportunism. Here, we typically have a problem of characterizing the optimal *exploration vs. exploitation* trade-off: acquiring the CSI of many of users (exploration) has a cost, but at the same time, it allows to increase the transmission rate (exploitation). The problem of designing joint opportunistic scheduling and CSI acquisition strategies also naturally arises in cognitive radio networks. For example, consider a cellular network shared by primary and secondary users. In a given cell or area, secondary users may access the spectrum only when primary users are not present in the system. Now since the secondary users may have very sporadic access to the spectrum, it makes no sense to maintain a dedicated channel between each of them and the base station. Instead, the base station could acquire the CSI only when primary users are not present in the system. Again when this occurs, one has to optimally design joint scheduling and CSI acquisition strategies, i.e., to define from how many secondary users the channel state should be acquired.

In all the work mentioned above, complete knowledge of CSI has been assumed and the cost of CSI acquisition has not considered. Recently, however, due to the practical importance of the issue, a significant amount of research efforts has been devoted towards designing schemes that optimize system performance when CSI is not automatically available, and when it rather has to be acquired (we present the detailed related work at the end of this section). In this paper, we consider a system consisting of a transmitter and a fixed number of receivers / users. To each user, an infinite buffer is associated and it stores packets arriving at a fixed intensity according to some stochastic process. Our aim is to design a joint CSI acquisition and scheduling strategy that stabilizes the buffers whenever this is at all possible.

Our main contributions are as follows:

- First, we characterize the set of arrival rates for which buffers can be stabilized (stability region). Next, we show that analogous to the max-weight policies proposed for the wireless systems with complete CSI information, queue length-based policies stabilizing the buffers whenever doing so is possible can be designed for wireless systems where the CSIs have to be acquire.
- Even though we could identify optimal queue length-based policies, implementing them turns out to be computationally expensive. Indeed, to compute an optimal policy, we need to solve generalized versions of optimal stopping time problems, which is known to be difficult [4]. To circumvent this difficulty, we first obtain key structural properties of the optimal policies, allowing us to reduce their computational complexity. In certain special cases of practical interest, we propose simple (polynomial time) algorithms to compute optimal policies.
- When an optimal policy cannot be obtained in polynomial time, we design computationally efficient approximate policies that provide a guaranteed fraction of the stability region.
- We evaluate the performance of the proposed policies using numerical experiments, and also quantify the cost in terms of stability due to the lack of CSI.

Related work. Identifying an optimal channel state acquisition and scheduling strategy has been addressed in the literature only recently [14, 20, 11, 12, 7, 8]. These papers study the exploration vs. exploitation trade-off in the saturated case, i.e., all users always have packets in their corresponding buffers. They characterize strategies that maximize the throughput in such systems. To our knowledge, the present work provides the first analysis of the joint CSI acquisition and scheduling problem with system stability as a performance objective. As we demonstrate later, accounting for the stochastic queueing behavior of the system greatly increases the problem complexity, and our results are non-trivial extensions of those presented in [8].

Most of the papers mentioned above consider that the cost of channel state acquisition is linear in the number of probed users and is independent of the transmission rate of the user finally scheduled. In other words, when, after acquiring the CSI of p users, a user is scheduled and the corresponding transmission rate is R, then the *reward* or throughput is equal to $R - p\beta$. This cost structure simplifies the problem as explained in [7, 8], but proves difficult to justify. Here we consider a more natural cost structure: time is slotted and acquiring the CSI of one user takes a fraction β of the slot, so that the throughput (in a slot) is equal to $R \times (1-p\beta)$. This cost model is referred to as a *logarithmic* cost in [8] (since the log of the throughput is just $\log(R) + log(1-p\beta)$).

It is worth mentioning here that the stability of wireless systems has been investigated when infrequent channel information is available, see e.g. [15]. Here, the CSI is available infrequently but acquiring this information has no cost, and controlling this acquisition is not considered.

As we will show later on in the paper, designing joint CSI acquisition and scheduling strategies to ensure system stability can be formalized as a stochastic control problem [4], however as already noticed in [11], even for saturated systems, this problem is different from all the classical control problems studied in the literature (bandit problems, secretary problems, stopping time problems, see e.g. [19, 28, 9]).

2. SYSTEM MODEL

Consider a broadcast channel with N receivers. Time is slotted, and the duration of a slot corresponds to the coherence time the channel of the various users. Let $C_i(t)$ denote a random variable representing the channel gain of user i in slot t. We assume that $\{C_i(t)\}_{t\geq 1}$ is an i.i.d. sequence, and that $C_i(t)$ can take a finite number of values. Without loss of generality, let $C_i(t) < c_{\max} < \infty$ for all i. Moreover, we assume that $\boldsymbol{C}(t) = [C_1(t) \ \cdots \ C_N(t)]$ is a random vector with independent components. Clearly, $\boldsymbol{C}(t)$ can take only finite number of values, say M, i.e., $\boldsymbol{C}(t) \in \{\boldsymbol{C}_1, \ldots, \boldsymbol{C}_M\}$ for every t. Let $\tilde{F}_i(\cdot)$ denote the distribution for $C_i(t)$. We assume that the distributions are known at the transmitter.

Notation: Capital letters will indicate random variables, while the corresponding small letters will indicate the observed value of the random variable. E.g. $C_i(t)$ is a random variable, while $c_i(t)$ is the observed value of $C_i(t)$. Moreover, bold letters will indicate vectors.

Let $A_i(t)$ denote the number of bits arriving in the buffer of receiver $i \in \{1, \ldots, N\}$ in slot t. We assume that the queue for each receiver can potentially store infinite number of bits, i.e., the arriving bits are not lost due to buffer overflow. Assume that $\{A_i(t)\}_{t\geq 1}$ be an i.i.d. sequence for each i, and $\boldsymbol{A}(t) = [A_1(t) \ \cdots \ A_N(t)]$ is a random vector with independent components with $A_i(t) < a_{\max} < \infty$ for all i. The expected arrival rate for the receiver i is denoted by $a_i = \mathbb{E}[A_i(t)]$.

We assume that the transmitter transmits at the fixed average power, but it can adapt transmission rate based on the channel state (CSI) of the receiver to which it transmits. Unlike previous work, we assume that the CSI for each of the receivers are not known at the beginning of the slot, rather the transmitter has to acquire this information. Acquiring the CSI consumes resources like time, bandwidth and power. In this paper, we assume that the transmitter can acquire the CSI of a receiver by probing, which consumes a fraction β of a slot. Thus, if the transmitter probes k receivers and then decides to transmit to a probed receiver (say i), then it can transmit only $(1 - k\beta)R(c_i(t))$ bits, where $c_i(t)$ is the observed value of the random variable $C_i(t)$. Since the transmitter can transmit to a receiver only after knowing its CSI, probing more receivers provide more options at the transmitter, but on the other hand, it consumes a larger fraction of resources and hence leaves a smaller fraction of resources for actual data transmission.

Let us assume that the transmitter probes k receivers. The set of probed receivers is denoted by $\mathcal{P}_k$ and the vector of observed CSI values is denoted by $\boldsymbol{c}^{\mathcal{P}_k}$. Similarly, the CSI of the unprobed receivers is denoted by $\boldsymbol{C}^{\overline{\mathcal{P}}_k}$, where $\overline{\mathcal{P}}_k = \{1, \ldots, N\} \setminus \mathcal{P}_k$.

Definition 1. *A joint probing and transmission strategy π is an algorithm that, given $(\mathcal{P}_k, \boldsymbol{c}^{\mathcal{P}_k})$, decides either to probe a receiver $j \in \overline{\mathcal{P}}_k$ or to transmit to a receiver $i \in \mathcal{P}_k$. When π decides to probe the channel of user j, the state changes to $(\mathcal{P}_{k+1}, \boldsymbol{c}^{\mathcal{P}_{k+1}})$, where $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{j\}$. When π decides to transmit, it transmits $(1 - \beta k)R(c_i)$ bits.*

Let $\mathcal{S}$ denote the set of all joint probing and transmission strategies. Note that the above definition does not prohibit a strategy π to choose i or j randomly from sets $\mathcal{P}_k$ and $\overline{\mathcal{P}}_k$, respectively, according to some distributions that may depend on the state $(\mathcal{P}_k, c^{\mathcal{P}_k})$. Thus, the class of all the joint probing and transmission strategies is uncountable. However, we show that it is sufficient to consider only a finite number of joint probing and transmission strategies in order to quantify the stability region.

Definition 2. *A deterministic joint probing and transmission strategy is generated by a set of functions $g_k^\pi : (\mathcal{P}_k, \boldsymbol{c}^{\mathcal{P}_k}) \rightarrow \{1, \ldots, N\}$ for every k. Let $g_k^\pi(\mathcal{P}_k, \boldsymbol{c}^{\mathcal{P}_k}) = i$. If $i \in \mathcal{P}_k$, then π transmits $(1 - k\beta)R(c_i)$ bits to user i. Moreover, if $i \in \overline{\mathcal{P}}_k$, then π probes user i when $\overline{\mathcal{P}}_k \neq \emptyset$; otherwise transmitter remains idle.*

Let $\widehat{\mathcal{S}}$ denote the set of all deterministic joint probing and transmission strategies. Note that each $\pi \in \widehat{\mathcal{S}}$ is defined recursively using functions $g_k^\pi(\cdot)$. In a given slot t, the initial state is $(\emptyset, \boldsymbol{c}^\emptyset)$, and $\mathcal{P}_0 = \emptyset$. If $g_0^\pi(\emptyset, \boldsymbol{c}^\emptyset) = i_1$, π probes i_1, and the state becomes $(\mathcal{P}_1, \boldsymbol{c}^{\mathcal{P}_1})$ with $\mathcal{P}_1 = \{i_1\}$. The process continues until for some k, $g_k^\pi(\mathcal{P}_k, \boldsymbol{c}^{\mathcal{P}_k}) \in \mathcal{P}_k$. Note that there are at most $O(M \times N!)$ deterministic strategies. Moreover, any policy in $\mathcal{S}$ can be obtained as a convex combination of policies in $\widehat{\mathcal{S}}$.

Definition 3. *A scheduling policy Δ is an algorithm that selects a strategy $\pi \in \widehat{\mathcal{S}}$ in every slot t.*

Denote by $\pi^\Delta(t)$ the joint probing and transmission strategy chosen by Δ in slot t. In general, we use the term "policy" for a scheduling policy and the term "strategy" for the a joint probing and transmission strategy. Also, to simplify the notation, we will drop t whenever the time slot of interest can be identified unambiguously.

Let $\boldsymbol{B}^\pi(t) = [B_1^\pi(t) \ \cdots \ B_N^\pi(t)]$ denote the vector indicating the number of bits served under strategy π in slot t. Since at most one receiver is served in each slot, at most one component of $\boldsymbol{B}^\pi(t)$ is greater than zero. The strategy π, however, may transmit to different receivers in different slots depending on the channel states observed. Now, we define the throughput of a scheduling policy.

Definition 4. *Throughput of a scheduling policy Δ is*

$$\boldsymbol{T}^\Delta \stackrel{\text{def}}{=} \liminf_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} \boldsymbol{B}^{\pi^\Delta}(s).$$

Let $\boldsymbol{Q}^\Delta(t)$ denote the queue length vector under policy Δ in slot t. The queue length dynamics is then given by

$$\boldsymbol{Q}^\Delta(t+1) = \max\{\boldsymbol{Q}^\Delta(t) + \boldsymbol{A}(t) - \boldsymbol{B}^{\pi^\Delta}(t), \boldsymbol{0}\}, \qquad (1)$$

where the maximum is component-wise.

Definition 5. *We say that the system is stable if the mean queue length for all the receivers is finite. A scheduling policy that stabilizes the system is a stable scheduling policy.*

Definition 6. *A policy is said to be throughput optimal if it can stabilize the system whenever there exists a scheduling policy that can stabilize the system.*

In the following section, we characterize the stability region of the system defined by the set of arrival rates $\boldsymbol{a} = (a_1, \ldots, a_N)$ such that there exists a policy stabilizing the system. We also present a queue length-based throughput optimal scheduling policy.

3. THROUGHPUT OPTIMAL SCHEDULING

First, we characterize the stability region. Let us define $\boldsymbol{T}^\pi = \mathbb{E}[\boldsymbol{B}^\pi(t)]$, where the expectation is with respect to the channel state distribution. Note that $\boldsymbol{T}^\pi$ denotes the throughput of a scheduling policy Δ that chooses $\pi^\Delta(t) = \pi$ in every slot t. Let $\Lambda = co(\{\boldsymbol{T}^\pi : \pi \in \widehat{\mathcal{S}}\})$, where $co(\mathcal{A})$ denotes the convex hull of the set $\mathcal{A}$. Also, let Λ° denote the set of all $\boldsymbol{\mu} \in \Lambda$ such that there exists $\boldsymbol{\nu} \in \Lambda$ satisfying $\boldsymbol{\nu} > \boldsymbol{\mu}$, where the inequality is component-wise. Now, the stability region is characterized by:

Theorem 1. *If $\boldsymbol{a} \in \Lambda^\circ$, then there exists Δ such that the system is stable under Δ. Moreover, if $\boldsymbol{a} \notin \Lambda$, then no scheduling policy can stabilize the system.*

To prove the above theorem, we use the following randomized policy $\Delta_R(\delta)$, where $\delta > 0$ is a parameter that will be chosen appropriately depending on $\boldsymbol{a}$. In each slot, $\Delta_R(\delta)$ chooses a strategy $\pi \in \widehat{\mathcal{S}}$ w.p. p_π independent of the choices in the previous slots. The distribution $\boldsymbol{p}$ is a feasible solution of the following equation:

$$\sum_{\pi \in \widehat{\mathcal{S}}} p_\pi \mathbb{E}[\boldsymbol{B}^\pi(t)] = \boldsymbol{a} + \boldsymbol{\delta}, \qquad (2)$$

where $\boldsymbol{\delta}$ is N-dimensional vector with all components equal to δ. $\Delta_R(\delta)$ is well defined only if (2) admits a solution. It might happen that in a given slot, say t, the service

$B_i^{\pi^{\Delta_R(\delta)}}(t)$ obtained by some user i under $\pi^{\Delta_R(\delta)}$ exceeds the buffer content $Q_i^{\Delta_R(\delta)}(t)$, in which case we assume that buffer i just empties in slot t. Let $r_{\max} = R(c_{\max})$. Whenever $Q_i(t) \geq r_{\max}$, $B_i^{\pi^{\Delta_R(\delta)}}(t)$ bits depart from queue i under $\Delta_R(\delta)$ in slot t.

Lemma 1. *If $\boldsymbol{a} \in \Lambda^\circ$, then there exists $\delta_{\boldsymbol{a}} > 0$ such that $\Delta_R(\delta)$ is well defined for every $\delta \in (0, \delta_{\boldsymbol{a}})$. Moreover, for every $\delta \in (0, \delta_{\boldsymbol{a}})$, $\Delta_R(\delta)$ is stable.*

Proof. When $\boldsymbol{a} \in \Lambda^\circ$, there exists $\widehat{\delta} > 0$ such that for every $\delta > \widehat{\delta}$, $\boldsymbol{a}+\boldsymbol{\delta} \notin \Lambda$; while for every $\delta \in (0, \widehat{\delta})$, $\boldsymbol{a}+\boldsymbol{\delta} \in \Lambda^\circ$. The above statement follows directly from the definition of Λ°. Choose $\widehat{\delta} = \delta_{\boldsymbol{a}}$, and fix $\delta \in (0, \delta_{\boldsymbol{a}})$. Now, since $\boldsymbol{a} + \boldsymbol{\delta} \in \Lambda^\circ$, i.e., $\boldsymbol{a} + \boldsymbol{\delta}$ is in the convex hull of $\{\mathbb{E}[\boldsymbol{B}^\pi(t)]\}_{\pi \in \widehat{\mathcal{S}}}$, then clearly (2) admits a solution.

Now, we show that $\Delta_R(\delta)$ is stable by proving that $\{\boldsymbol{Q}^{\Delta_R(\delta)}(t)\}_{t\geq 1}$ is a positive recurrent Markov chain. We use Foster's theorem [10] to establish the positive recurrence. Define Lyapunov function $l(\boldsymbol{Q}) = \boldsymbol{Q} \cdot \boldsymbol{Q}$, and consider the drift $D^{\Delta_R(\delta)}(\boldsymbol{Q})$ of $l(\cdot)$ under $\Delta_R(\delta)$ in state $\boldsymbol{Q}$.

$$\begin{aligned} &D^{\Delta_R(\delta)}(\boldsymbol{Q}) \\ &= \mathbb{E}\left[l\left(\boldsymbol{Q}^{\Delta_R(\delta)}(t+1)\right) - l\left(\boldsymbol{Q}^{\Delta_R(\delta)}(t)\right) | \boldsymbol{Q}^{\Delta_R(\delta)}(t) = \boldsymbol{Q}\right] \\ &\leq N[a_{\max}]^2 + N[r_{\max}]^2 + 2\boldsymbol{Q}\cdot\boldsymbol{a} - 2\mathbb{E}[\boldsymbol{Q}\cdot\boldsymbol{B}^{\pi^{\Delta_R(\delta)}}(t)|\boldsymbol{Q}]. \end{aligned}$$

Let $\mathcal{M} = \{i : Q_i(t) \geq r_{\max}\}$. Note that $\mathcal{M}$ is a set of receivers for which the number of bits that depart are exactly equal to $B_i^{\pi^{\Delta_R(\delta)}}(t)$. Moreover, for every $i \in \mathcal{M}$, $B_i^{\pi^{\Delta_R(\delta)}}(t)$ is independent of the queue length. Thus, we have the following.

$$\begin{aligned} &D^{\Delta_R(\delta)}(\boldsymbol{Q}) \\ &\leq N[a_{\max}]^2 + N[r_{\max}]^2 + 2r_{\max}\sum_{i\in\overline{\mathcal{M}}} a_i \\ &\quad +2\sum_{i\in\mathcal{M}} Q_i a_i - 2\sum_{i\in\mathcal{M}} Q_i \mathbb{E}[B_i^{\pi^{\Delta_R(\delta)}}(t)] \\ &\leq N[a_{\max}]^2 + N[r_{\max}]^2 + 2r_{\max}\sum_{i\in\overline{\mathcal{M}}} a_i - 2\delta\sum_{i\in\mathcal{M}} Q_i \\ &\leq N(a_{\max}+r_{\max})^2 - 2\delta\sum_{i\in\mathcal{M}} Q_i. \end{aligned}$$

Thus, $D^{\Delta_R(\delta)}(\boldsymbol{Q}) < 0$ for every $\boldsymbol{Q}$ such that there exists i such that $Q_i > \frac{N(a_{\max}+r_{\max})^2}{2\delta}$. This shows that the Lyapunov drift is negative in all except finite number of states. Thus, $\Delta_R(\delta)$ stabilizes $\boldsymbol{a}$. ☐

Now, we are ready to prove Theorem 1.

Proof Proof of Theorem 1. In view of Lemma 1, it suffices to show that when $\boldsymbol{a} \notin \Lambda$, then no scheduling policy can stabilize the system. We prove this using contradiction. Fix $\boldsymbol{a} \notin \Lambda$, and let there exist a scheduling policy Δ that can stabilize the system. Let $\zeta_\pi^\Delta(t)$ denote the number of slots until t in which Δ has chosen π. We assume that $\lim_{t\to\infty} \frac{\zeta_\pi^\Delta(t)}{t}$ exists w.p. 1. However, we do not assume that the limit is same on all the sample paths. Fix a sample path and let $\tilde{p}_\pi^\Delta = \lim_{t\to\infty} \frac{\zeta_\pi^\Delta(t)}{t}$. Note that $\{\tilde{p}_\pi^\Delta : \pi \in \widehat{\mathcal{S}}\}$ is a distribution on $\widehat{\mathcal{S}}$. Moreover, since Δ is stable, the departure rate from each queue is equal to the arrival rate in the queue. Thus, $\boldsymbol{a} \in \Lambda$, a contradiction. ☐

Now, we design an adaptive queue length-based scheduling policy $\Delta^\star$ as follows. The policy $\Delta^\star$ that we propose chooses $\pi^{\Delta^\star}(t)$ such that

$$\pi^{\Delta^\star}(t) \in \arg\max_{\pi\in\widehat{\mathcal{S}}}\{\boldsymbol{Q}(t)\cdot\mathbb{E}[\boldsymbol{B}^\pi(t)]\}.$$

In the following theorem, we show that $\Delta^\star$ is throughput optimal. First, we note that in every t, for every Δ

$$\mathbb{E}[\boldsymbol{Q}(t)\cdot\boldsymbol{B}^{\pi^{\Delta^\star}}(t)|\boldsymbol{Q}(t)] \geq \mathbb{E}[\boldsymbol{Q}(t)\cdot\boldsymbol{B}^{\pi^{\Delta}}(t)|\boldsymbol{Q}(t)].$$

The above inequality is satisfied for all values of $\boldsymbol{Q}(t)$.

Theorem 2. *If $\boldsymbol{a} \in \Lambda^\circ$, then $\Delta^\star$ stabilizes $\boldsymbol{a}$. Thus, $\Delta^\star$ is a throughput optimal scheduling policy.*

Proof. We show that the Markov chain $\{\boldsymbol{Q}^{\Delta^\star}(t)\}_{t\geq1}$ is positive recurrent, and as in the proof of Theorem 1, we use Foster's theorem to establish stability. Again we use the Lyapunov function $l(\cdot)$: $l(\boldsymbol{Q}) = \boldsymbol{Q}\cdot\boldsymbol{Q}$. The Lyapunov drift under policy $\Delta^\star$ becomes

$$\begin{aligned} &\mathbb{E}[l(\boldsymbol{Q}(t+1)) - l(\boldsymbol{Q}(t))|\boldsymbol{Q}(t)] \\ &\leq N[a_{\max}]^2 + N[r_{\max}]^2 \\ &\quad +2\boldsymbol{Q}(t)\cdot\boldsymbol{a} - 2\mathbb{E}[\boldsymbol{Q}(t)\cdot\boldsymbol{B}^{\pi^{\Delta^\star}}(t)|\boldsymbol{Q}(t)] \\ &\leq N[a_{\max}]^2 + N[r_{\max}]^2 \\ &\quad +2\boldsymbol{Q}(t)\cdot\boldsymbol{a} - 2\mathbb{E}[\boldsymbol{Q}(t)\cdot\boldsymbol{B}^{\pi^{\Delta_R(\delta)}}(t)|\boldsymbol{Q}(t)]. \end{aligned}$$

Note that the last expression above provides an upper bound on the drift under policy $\pi^{\Delta_R(\delta)}$. Now, using the same arguments as that in the proof of Lemma 1, it can be shown that the drift is negative outside the compact set $\{\boldsymbol{Q} : \boldsymbol{Q} \leq N(r_{\max}+a_{\max})^2/2\delta$ element-wise$\}$. ☐

Though Theorem 2 provides a throughput optimal policy, it does not specify how $\pi^{\Delta^\star}$ should be computed in each slot. We address this in the next section.

4. OPTIMAL JOINT-PROBING AND TRANSMISSION STRATEGY $\pi^{\Delta^\star}$

The problem of maximizing $\boldsymbol{Q}(t)\cdot\mathbb{E}[\boldsymbol{B}^\pi(t)]$ given the queue lengths $\boldsymbol{Q}(t)$ over all joint probing and transmission strategies π can be addressed using Markov Decision Process (MDP) formulation. Indeed, the problem is a generalization of optimal stopping time problem as can be seen from Definition 1. We say that the problem is a generalization because in this case, in addition to determining when to stop, the strategy has to determine which user probe next; while in the classical optimal stopping time problem, the probing sequence is given a priori. Next, we formally formulate the problem of determining $\pi^{\Delta^\star}$.

4.1 Generalized Stopping Time Problem

Consider any slot t. Let $\boldsymbol{Q}(t) = [Q_1(t) \ \cdots \ Q_N(t)]$ denote the queue lengths in slot t. For conciseness, we omit t. Let π be a joint probing and transmission strategy, and assume that in state $(\mathcal{P}_k, \boldsymbol{c}^{\mathcal{P}_k})$ it decides to transmit. To maximize $\boldsymbol{Q}\cdot\mathbb{E}[\boldsymbol{B}^\pi]$, then π should transmit to user $i = \arg\max_{j\in\mathcal{P}_k} Q_j R(c_j)$. Thus, to define an optimal strategy, it suffices to record the *modified* state $(\mathcal{P}_k, w)$, where $\mathcal{P}_k$ is the set of k probed users, and $w = \max_{j\in\mathcal{P}_k} Q_j R(c_j)$ is the maximum weighted rate in the set of users in $\mathcal{P}_k$. We refer to $\boldsymbol{Q}\cdot\mathbb{E}[\boldsymbol{B}^\pi]$ as the expected weighted throughput.

Consider the state $(\mathcal{P}_k, w)$, and define $T_{\rm tr}(\mathcal{P}_k, w)$ as the weighted throughput if the joint probing and transmission strategy decides to transmit in the state $(\mathcal{P}_k, w)$. Thus, $T_{\rm tr}(\mathcal{P}_k, w) = (1 - k\beta)w$. Furthermore, let $T^\star(\mathcal{P}_k, w)$ denote the maximum expected value of the weighted throughput that can be achieved from state $(\mathcal{P}_k, w)$ under any strategy. Then, $T^\star(\mathcal{P}_k, w)$ can be recursively computed as follows:

$$\begin{aligned} &T^\star(\mathcal{P}_k, w) \\ &= \max\left\{T_{\rm tr}(\mathcal{P}_k, w), \max_{j\in\overline{\mathcal{P}}_k} \mathbb{E}[T^\star(\mathcal{P}_k \cup \{j\}, w \vee W_j)]\right\}, \end{aligned} \quad (3)$$

where $W_j \stackrel{\rm def}{=} Q_j R(C_j)$, and $(a \vee b) = \max\{a, b\}$. Note that the first term in the right hand side of (3) corresponds to the weighted throughput given that the decision is to transmit, while the second term corresponds to the maximum expected weighted throughput obtained by a system given that it probes some user $j \in \overline{\mathcal{P}}_k$. After probing j, the state becomes $(\mathcal{P}_{k+1}, w \vee W_j)$, where $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{j\}$. In each slot, the system starts from the state $(\emptyset, 0)$, and hence the maximum expected value of the weighted throughput is $T^\star(\emptyset, 0)$. Let us denote the distribution of W_i by $F_i(\cdot)$. Note that given $\boldsymbol{Q}$, $F_i(\cdot)$ can be easily obtained from the distribution $\tilde{F}_i(\cdot)$ of C_i.

At this point, it is worth mentioning that $T^\star(\emptyset, 0)$ can be obtained using dynamic programming. But the computational complexity of this procedure is of the order of the cardinality of the state space. Note that the cardinality of the state space in our case is at least $N! \times M$. Thus, it is easy to see that when the number of users N is large or/and the number of channel state values M is large, then using dynamic programming is computationally infeasible. In the following subsection, our aim is to find structural properties of $\pi^{\Delta^\star}$ so as to reduce the computational complexity.

4.2 Structural Properties of $\pi^{\Delta^\star}$

Before obtaining some key structural properties, let us define some important terms. Let $T_{\rm tr,pr(\it i)}(\mathcal{P}_k, w)$ for $i \in \overline{\mathcal{P}}_k$ and $T_{\rm tr,pr}(\mathcal{P}_k, w)$ be defined as follows.

$$\begin{aligned} T_{\rm tr,pr(\it i)}(\mathcal{P}_k, w) &\stackrel{\rm def}{=} (1 - (k+1)\beta)\mathbb{E}[w \vee W_i], \\ T_{\rm tr,pr}(\mathcal{P}_k, w) &\stackrel{\rm def}{=} \max_{i\in\overline{\mathcal{P}}_k} \left\{T_{\rm tr,pr(\it i)}(\mathcal{P}_k, w)\right\}. \end{aligned}$$

We note that $T_{\rm tr,pr(\it i)}(\mathcal{P}_k, w)$ denotes the expected weighted throughput of the system, if in state $(\mathcal{P}_k, w)$, the probing and transmission strategy probes user $i \in \overline{\mathcal{P}}_k$ and then transmits to the user that provides the largest weighted throughput. Thus, $T_{\rm tr,pr}(\mathcal{P}_k, w)$ denotes the maximum expected weighted throughput given that from state $(\mathcal{P}_k, w)$ exactly one additional user is probed. We refer to $T_{\rm tr,pr}(\mathcal{P}_k, w)$ as a one-step-look-ahead throughput from state $(\mathcal{P}_k, w)$.

Next, fix the set $\mathcal{P}_k$ and let us define $\mathcal{D}_k$ as follows:

$$\mathcal{D}_k \stackrel{\rm def}{=} \{w \;:\; T_{\rm tr}(\mathcal{P}_k, w) \geq T_{\rm tr,pr}(\mathcal{P}_k, w)\}. \quad (4)$$

Using this notation, we prove the following key result.

THEOREM 3. *In state $(\mathcal{P}_k, w)$, if $w \in \mathcal{D}_k$, then the optimal decision is to transmit to the user $i \in \mathcal{P}_k$ such that $Q_i R(c_i) = w$, otherwise the optimal decision is to probe a user in $\overline{\mathcal{P}}_k$.*

PROOF. We fix any arbitrary $\mathcal{P}_{N-1} \supset \mathcal{P}_k$, and let us assume that the users in $\mathcal{P}_{N-1}$ are probed. Then, the resulting system state is $(\mathcal{P}_{N-1}, w_{N-1})$. Note that $w_{N-1} \geq w$, therefore $w_{N-1} \geq w_{max}(\mathcal{P}_k)$, or by Lemma 3, $w_{N-1} \in \mathcal{D}_k$ and by Lemma 4, $w_{N-1} \in \mathcal{D}_{N-1}$. Lemmas 3-4 and the definition of $w_{max}(\mathcal{P}_k)$ are provided in Appendix. Thus,

$$\begin{aligned} &T_{tr}(\mathcal{P}_{N-1}, w_{N-1}) \\ &\quad\geq\; T_{pr,tr}(\mathcal{P}_k, w_{N-1}) \\ &\quad=\; \max_{i\in\overline{\mathcal{P}}_{N-1}} \{\mathbb{E}[T_{tr}(\mathcal{P}_{N-1} \cup \{i\}, w_{N-1} \vee W_i)]\} \\ &\quad=\; \max_{i\in\overline{\mathcal{P}}_{N-1}} \{\mathbb{E}[T^\star(\mathcal{P}_{N-1} \cup \{i\}, w_{N-1} \vee W_i)]\}. \end{aligned}$$

The last relation follows because after probing the last user, transmit is the only optimal decision. Hence if $N - 1 > k$, we have

$$T^\star(\mathcal{P}_{N-1}, w_{N-1}) = T_{tr}(\mathcal{P}_{N-1}, w_{N-1}). \quad (5)$$

We next consider state $(\mathcal{P}_{N-2}, w_{N-2})$ and prove the result similarly. In fact we show by induction down to $k + 1$, that $T^\star(\mathcal{P}_{k+1}, u_{k+1}) = T_{tr}(\mathcal{P}_{k+1}, w_{k+1})$ which completes the proof. □

Theorem 3 shows that the decision of transmitting in the current state or to probe a new user can be obtained by considering the one-step-look-ahead weighted throughput. However, if the optimal decision is to probe a user in $\overline{\mathcal{P}}_k$, then the above theorem does not elaborate on which user to probe. Indeed, the intuitive choice of probing a user that maximizes the one-step-look-ahead weighted throughput is not optimal. We demonstrate this using an example.

EXAMPLE 1. *Assume that slots are of unit duration. Let us consider two receivers U_1 and U_2, and let their respective queue lengths be Q_1 and Q_2. In each slot, let the maximum rate to U_1 be 2 w.p. $(k-1)/k$ and k w.p. $1/k$, and for receiver U_2 let it be 1 w.p. $(2k-1)/2k$ and $2k$ w.p. $1/2k$. Hence the one-step-look-ahead weighted throughputs for U_1 and U_2 at the initial state $(\emptyset, 0)$ are respectively*

$$\begin{aligned} T_{tr,pr(U_1)}(\emptyset, 0) &= \frac{Q_1(3k-2)}{k}, \text{ and} \\ T_{tr,pr(U2)}(\emptyset, 0) &= \frac{Q_2(4k-1)}{2k}. \end{aligned}$$

Let us fix $k > 3$ and $Q_1 = 2Q_2 = 2Q$. In this case, $T_{tr,pr(U_1)}(\emptyset, 0) > T_{tr,pr(U_2)}(\emptyset, 0)$. Hence intuitively one would expect to probe U_1. However, we show that for

$$\beta < \min\left\{\frac{k^2 - 3k + 2}{7k^2 - 10k + 4}, \frac{2k^2 - 3k + 1}{5k^2 - 6k + 2}\right\}, \quad (6)$$

probing U_2 first would provide a higher expected weighted throughput. We can show that the optimal policy $\pi^\star$ is to probe U_2 first in every slot, and if the achievable rate is $2k$, then transmit to U_2; otherwise probe U_1 and transmit to it at appropriate rate. The expected weighted throughput of $\pi^\star$ is $T_{\pi^\star} = Q(1-\beta) + 2Q(1-2\beta)\frac{6k^2-7k+2}{2k^2}$. To show that this policy is optimal we need to compare it with π^1 that probes U_1 and transmits at appropriate rate, and with π^2 which probes U_1 first. If the achievable rate is k, then π^2 transmits to U_1 otherwise it probes U_2 and transmits to it if the achievable rate is $2k$, else it transmits to U_1 at rate 2. The expected weighted throughput of π^1 is $T_{\pi^1} = 2Q(1-\beta)(\frac{2(k-1)}{k} + 1)$ while that of π^2 is $T_{\pi^2} = 2Q(1-\beta) + Q(1-$

$2\beta)\frac{3k^2-4k+1}{k^2}$. $T_{\pi^1} < T_{\pi^\star}$ *if* $\beta < \frac{k^2-3k+2}{7k^2-10k+4}$ *while* $T_{\pi^2} < T_{\pi^\star}$ *if* $\beta < \frac{2k^2-3k+1}{5k^2-6k+2}$. *Hence if (6) holds, then* $\pi^\star$ *is optimal.*

In the following theorem, we provide guidelines to determine which user to probe in a special case. First, let us denote $W_i \leq_{st} W_j$ whenever the random variable W_i is stochastically smaller than the random variable W_j, i.e., $\mathbb{E}[f(W_i)] \leq \mathbb{E}[f(W_j)]$ for every increasing function $f(\cdot)$.

THEOREM 4. *If there exists a user* $i \in \overline{\mathcal{P}}_k$ *such that* $W_j \leq_{st} W_i$ *for every* $j \in \overline{\mathcal{P}}_k$*, then in state* $(\mathcal{P}_k, w)$ *such that* $w \notin \mathcal{D}_k$*, the optimal decision is to probe user* i.

PROOF. We prove the theorem by induction on the number of unprobed users. When this number is 1, the result holds since we can only probe this user. Now assume that the result holds when the number of unprobed users is strictly less than $N-k$. Denote $(\mathcal{P}_k, w)$ as the system state. Since $T_{pr,tr}(\mathcal{P}_k, w) \geq T_{tr}(\mathcal{P}_k, w)$, $w \leq w_{max}(\mathcal{P}_k)$. Define for all i, j $\alpha_i = w_{max}(\mathcal{P}_k \cup \{i\})$, $\alpha_j = w_{max}(\mathcal{P}_k \cup \{j\})$ and $\alpha = w_{max}(\mathcal{P}_k \cup \{i\} \cup \{j\})$. Note the $\alpha_i \geq \alpha_j$ and let $i \in \overline{\mathcal{P}_k}$ with $i \neq j$. If $w \geq \alpha_i$, then after probing i or j we should transmit. It is then optimal to probe j. Hence we will assume $w \leq \alpha_i$
We compare the expected throughput obtained starting from the state $(\mathcal{P}_k, w)$, (a) when first probing user i and the user j, (b) when first probing user j and then user i.

- In scenario (a), probing i (with weight w_i) results in channel state x_i. By induction we know the next user to probe is j. Then if $W_i = Q_i R(x_i) \geq \alpha_i$, we should not probe j and transmit. If $Q_i R(x_i) < \alpha_i$, we should probe j. If $w \vee W_i \vee W_j \geq \alpha$, we should transmit otherwise we should probe further.
- In scenario (b), we first probe j. If $W_j \geq \alpha_j$, we should transmit. Otherwise we probe i. Then if $w \vee W_j \vee W_i \geq \alpha$, we should transmit otherwise we should probe further.

We need to compare expected throughput in scenarios (a) and (b) in cases where we transmit after probing i and/or j. We will hence consider cases when $w \leq \alpha_i$. This is simply due to the fact that if we have to probe further i and j, the system (a) and (b) are identical. We denote by $T^{(a)}(w)$ and $T^{(b)}(w)$ the expected throughputs in scenarios (a) and (b) when we do not probe more users than i and j. Also w_i and w_j are weights are associated with user i and j and F_i and F_j are pdf's associated with W_i and W_j:

$$\begin{aligned} T^{(a)}(w) = \quad & a_{k+1}\int_{\infty}^{\alpha_i} dF_i(w_i)w_i + a_{k+2}\int_0^{\alpha_i} dF_i(w_i) \\ & \times \int_0^{\infty} dF_j(w_j)\mathbf{1}_{w\vee w_i\vee w_j\geq\alpha}\max\{w,w_i,w_j\} \\ T^{(b)}(w) = \quad & a_{k+1}\int_{\infty}^{\alpha_j} dF_j(w_j)w_j + a_{k+2}\int_0^{\alpha_j} dF_j(w_j) \\ & \times \int_0^{\infty} dF_i(w_i)\mathbf{1}_{w\vee w_i\vee w_j\geq\alpha}\max\{w,w_i,w_j\} \end{aligned}$$

where $a_k = (1-k\beta)$. We want to prove $G(w) = T^{(b)}(w) - T^{(a)}(w) \geq 0$, which follows from Lemmas 5 and 6 in Appendix. □

Though Theorem 4 provides useful guidelines for deciding which user to probe, we note that in general the random variables W_j and W_i may not be stochastically ordered even when $R(C_i) \leq_{st} R(C_j)$ for certain values of Q_i and Q_j. Moreover, since the weight Q_i corresponds to the queue length of user i, we can not assume any specific structure on the values of Q_i's. Thus, the complete characterization of the optimal policy through Theorems 3 and 4 is available only in limited cases. We note that the problem of determining which user to probe is difficult in general (see [11]) even when all the weights are equal; the solutions are known only in special cases [8]. In the following section, we consider certain special cases that are relevant in practice, and obtain scheduling policies with provable stability guarantees.

5. PROPORTIONAL FADING

We consider the special case of *proportional fading* where for every i, $C_i = x_i Y_i$, where x_i is a constant and Y_i is a random variable with distribution $F(\cdot)$ independent of i. The Y_i's are assumed to be independent across users. This choice of modelling is relevant in practical scenarios, where the channel state distribution of each user is of the same form, but the means are different. In wireless scenario, such cases will appear when the fast fading is i.i.d. but the slow fading for each user is different on account of the different distances between the transmitters and their respective receivers. Such models are widely used in the literature, see e.g. [6]. We also assume that $R(c) = \log(1 + c/N_0)$, where N_0 is the power spectral density of noise.

With these additional assumptions, we can completely characterize the optimal probing sequence using Theorem 4 in certain special cases that are described in Corollaries 1 and 2.

COROLLARY 1. *If* $x_i = x_j$ *for all* i *and* j*, i.e. the users are i.i.d., then the optimal strategy is defined recursively, starting from state* $(\emptyset, 0)$ *as follows: In any state* $(\mathcal{P}_k, w)$*, if* $w \in \mathcal{D}_k$*, then transmit to the user that maximizes the weighted throughput in the set* $\mathcal{P}_k$*; otherwise probe user*

$$i = \arg\max_{j\in\overline{\mathcal{P}}_k}\{Q_j\}.$$

PROOF. Let $x_i = x$ for every i. Now, note that

$$Q_i \log\left(1 + \frac{xY}{N_0}\right) \leq_{st} Q_j \log\left(1 + \frac{xY}{N_0}\right),$$

whenever $Q_i \leq Q_j$. Thus, the result follows from Theorems 3 and 4. □

COROLLARY 2. *Suppose that, for every* i*, the rates* $R(C_i)$ *can be approximated by* $\rho_i R_i$*, where* ρ_i *is a constant and* R_i *is a random variable whose distribution does not depend on* i*. Then, the optimal strategy is defined recursively, starting from state* $(\emptyset, 0)$ *as follows: In any state* $(\mathcal{P}_k, w)$*, if* $w \in \mathcal{D}_k$*, then transmit to the user that maximizes the weighted throughput in the set* $\mathcal{P}_k$*; otherwise probe user*

$$i = \arg\max_{j\in\overline{\mathcal{P}}_k}\{Q_j \times \rho_j\}.$$

PROOF. Note that the weighted throughput of any user i is $W_i = \frac{Q_i \rho_i R}{N_0}$. Thus, clearly, $W_i \leq_{st} W_j$ whenever $Q_i\rho_i \leq Q_j\rho_j$. Thus, the result follows from Theorem 4. □

Note that the condition in Corollary 2 is satisfied for the low SNR case because $R(C_i) \approx \frac{x_i}{N_0}Y$ for low SNR. Now, choose $\rho_i = x_i/N_0$ and $R = Y$.

Even in this simplified channel model, apart from the special cases discussed in Corollaries 1 and 2, it is difficult to determine the optimal probing sequence. In the following subsection, we obtain a scheduling policy that achieves a guaranteed fraction of the stability region.

5.1 Approximate Probing and Transmission Strategy

Fix slot t and the queue lengths in this slot. Let $[a,b]$ denote the support for the r.v. Y_i's, and let $x_{\min} = \min_i x_i$ and $x_{\max} = \max_i x_i$. Also, define

$$\gamma = \frac{\log\left(\frac{x_{\max}}{x_{\min}}\right)}{\log\left(\frac{x_{\min} a}{N_0}\right)} + 1.$$

We assume that $\frac{a x_{\min}}{N_0} \gg 1$, i.e., we consider the high SNR regime. Hence, for every i, $R(x_i c) \approx \log\left(\frac{x_i c}{N_0}\right)$ for every $c \in [a,b]$. In these settings, we define a probing and transmission strategy π_A as follows:

Initial step: (system state = $(\emptyset, 0)$)

1. Find set $\mathcal{A} = \arg\max_{i=1,\ldots,N} Q_i R(x_i a)$. Now, probe user $j_1^{\pi_A} = \min\{\arg\max_{i\in\mathcal{A}} x_i\}$. Let the weighted rate for the probed user be w_1, i.e., $w_1 = Q_{j_1^{\pi_A}} R(c_{j_1^{\pi_A}})$.

2. If $\gamma w_1 \in \mathcal{D}_1$, where $\mathcal{D}_1$ is defined with respect to set $\mathcal{P}_1 = \{j_1^{\pi_A}\}$, then transmit to the user; otherwise update the state to $(\mathcal{P}_1, \gamma w_1)$ and continue as described below.

Subsequent steps: (system state = $(\mathcal{P}_{k-1}, \gamma w)$)

1. Find set $\mathcal{A} = \arg\max_{i\in\overline{\mathcal{P}}_{k-1}} Q_i R(x_i a)$. Now, probe user $j_k^{\pi_A} = \min\{\arg\max_{i\in\mathcal{A}} x_i\}$. Let the weighted throughput for the probed user be w_k.

2. If $\gamma(w \vee w_k) \in \mathcal{D}_k$, where $\mathcal{D}_k$ is defined with respect to set $\mathcal{P}_k = \mathcal{P}_{k-1} \cup \{j_1^{\pi_A}\}$, then transmit to the user that provides the maximum weighted throughput; otherwise update the state to $(\mathcal{P}_k, \gamma(w \vee w_k))$ and go to 1).

We note that under π_A, the sequence in which the users are probed is deterministic given the queue lengths. That is for a given $\boldsymbol{Q}$, the probing sequence under π_A can be completely determined independently of the observed channel states of the probed users. Only the decision to transmit now or probe further is determined by the channel state observations.

Let $T^{\pi_A}(\mathcal{P}_k, \gamma w)$ be the expected weighted throughput obtained under policy π_A from state $(\mathcal{P}_k, \gamma w)$. Note that if π_A decides to transmit in the state, then $T^{\pi_A}(\mathcal{P}_k, \gamma w) = (1-k\beta)w$; otherwise $T^{\pi_A}(\mathcal{P}_k, \gamma w) = \mathbb{E}[T^{\pi_A}(\mathcal{P}_{k+1}, \gamma(w \vee W_{j_{k+1}^{\pi_A}}))]$, where $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{j_{k+1}^{\pi_A}\}$. Thus,

$$\begin{aligned}
&\mathbb{E}[T^{\pi_A}(\mathcal{P}_{k+1}, \gamma(w \vee W_{j_{k+1}^{\pi_A}}))] \\
&= (1-(k+1)\beta)\int_{\gamma(w\vee v)\in\mathcal{D}_{k+1}} (u \vee v)\mathbf{d}F_{j_{k+1}^{\pi_A}}(v) \\
&\quad + \int_{\gamma(w\vee v)\notin\mathcal{D}_{k+1}} \mathbb{E}[T^{\pi_A}(\mathcal{P}_{k+2}, \gamma(w \vee v \vee W_{j_{k+2}^{\pi_A}}))]\mathbf{d}F_{j_{k+1}^{\pi_A}}(v).
\end{aligned}$$

With this notation, our key result is the following.

Theorem 5. *For any queue length values $\{Q_i\}_{i=1,\ldots,N}$,*

$$\frac{1}{\gamma}T^\star(\emptyset, 0) \le T^{\pi_A}(\emptyset, 0).$$

To prove Theorem 5, we use the following lemma.

Lemma 2. *Fix any system state $(\mathcal{P}_k, w)$. Then, $W_i \le_{st} \gamma W_{j_k^{\pi_A}}$ for every $i \in \overline{\mathcal{P}}$.*

Proof. For brevity, let $j_k^{\pi_A} = j$. Now, to establish $W_i \le_{st} \gamma W_j$, it suffices to show that for every $u \in [a,b]$, $Q_i R(x_i u) \le \gamma Q_j R(x_j u)$. We prove the afore-mentioned by considering the various possible cases. Consider any $i \in \overline{\mathcal{P}}_k$. First, let us consider the case when $x_i \ge x_j$. Here, $Q_i \le Q_j$. Otherwise, π_A will choose i instead of j. Now, let us define the function $g_{ij}(u)$ as follows: $g_{ij}(u) = Q_j R(x_j u) - Q_i R(x_i u)$. Moreover,

$$\frac{\partial g_{ij}(u)}{\partial u} = \frac{Q_j - Q_i}{u}.$$

Now, since $Q_j \ge Q_i$, $g_{ij}(\cdot)$ is an increasing function. Furthermore, by the choice of j, $g_{ij}(a) \ge 0$. Thus, $g_{ij}(u) \ge 0$ for every $u \in [a,b]$. Since $\gamma > 1$, the required result follows.

Now, let us consider the case $x_i \le x_j$. If $Q_i \le Q_j$, then clearly $g_{ij}(u) \ge 0$ for every $u \in [a,b]$. So, we consider $w_i \ge w_j$. Let $m = Q_i/Q_j$. It suffices to show that $\gamma \ge m$. First, note that $g_{ij}(u)$ is a monotone decreasing function with $g_{ij}(a) \ge 0$. Moreover, if $g_{ij}(b) \ge 0$, then the required follows immediately. Hence, we need to consider $g_{ij}(b) < 0$. Since $g_{ij}(\cdot)$ is a continuous decreasing function on a compact set $[a,b]$ with boundary conditions $g_{ij}(a) \ge 0$ and $g_{ij}(b) < 0$, there exists $u' \in [a,b]$ such that $g_{ij}(u') = 0$. Thus,

$$\begin{aligned}
&Q_i \log\left(\frac{x_i u'}{N_0}\right) = Q_j \log\left(\frac{x_j u'}{N_0}\right) \\
&\Rightarrow m = \frac{\log\left(\frac{x_j}{x_i}\right)}{\log\left(\frac{x_i u'}{N_0}\right)} + 1 \Rightarrow m \le \frac{\log\left(\frac{x_{\max}}{x_{\min}}\right)}{\log\left(\frac{x_{\min} a}{N_0}\right)} + 1.
\end{aligned}$$

Thus, $m \le \gamma$. This concludes the proof. ☐

Now, we prove Theorem 5.

Proof of Theorem 5. First, using induction k, we show that for every $(\mathcal{P}_k, \gamma w)$, $T^\star(\mathcal{P}_k, \gamma w) \le \gamma T^{\pi_A}(\mathcal{P}_k, \gamma w)$. Consider any state $(\mathcal{P}_N, \gamma w)$. Since there are no more receivers to probe, the optimal decision is to transmit. Now, note that $T^\star(\mathcal{P}_N, \gamma w) = \gamma(1-N\beta)w$. Also, note that $T^{\pi_A}(\mathcal{P}_N, \gamma w) = (1-N\beta)w$. Thus, the result holds for $k = N$. Now, by induction hypothesis, we assume that the claim is true for every $u \ge k+1$.

Now, consider state $(\mathcal{P}_k, \gamma w)$. As before, if $\gamma w \in \mathcal{D}_k$, then $T^\star(\mathcal{P}_k, \gamma w) = (1-k\beta)\gamma w$. Under π_A also, the decision is to transmit if $\gamma w \in \mathcal{D}_k$. Hence, $T^{\pi_A}(\mathcal{P}_k, \gamma w) = (1-k\beta)w$. Thus, the result follows whenever $\gamma w \in \mathcal{D}_k$. Now, consider $\gamma w \notin \mathcal{D}_k$. Then,

$$\begin{aligned}
&T^\star(\mathcal{P}_k, \gamma w) \\
&= \max_{i\notin\overline{\mathcal{P}}_k} \mathbb{E}[T^\star(\mathcal{P}_k \cup \{i\}, \gamma(w \vee W_i))] \\
&\le \max\Bigg\{ \max_{\substack{i\in\overline{\mathcal{P}}_k \\ i\ne j_{k+1}^{\pi_A}}} \mathbb{E}[T^\star(\mathcal{P}_k \cup \{i\}, \gamma(w \vee W_i))], \\
&\qquad \mathbb{E}[T^\star(\mathcal{P}_k \cup \{j_{k+1}^{\pi_A}\}, \gamma(w \vee W_{j_{k+1}^{\pi_A}}))]\Bigg\} \\
&= \mathbb{E}[T^\star(\mathcal{P}_k \cup \{j_{k+1}^{\pi_A}\}, \gamma(w \vee W_{j_{k+1}^{\pi_A}}))]. \qquad (7)
\end{aligned}$$

The last relation follows from Lemma 2 and Theorem 4. Note that Lemma 2 shows that for every i, $W_i \le_{st} \gamma W_{j_{k+1}^{\pi_A}}$. Now, (7) follows using Theorem 4. Next, note that

$$T^{\pi_A}(\mathcal{P}_k, \gamma w) = \mathbb{E}[T^{\pi_A}(\mathcal{P}_k \cup \{j_{k+1}^{\pi_A}\}, \gamma(w \vee W_{j_{k+1}^{\pi_A}}))]. \qquad (8)$$

Now, the required result follows using (7), (8) and the induction hypothesis. Since, we have shown that $T^\star(\mathcal{P}_k, \gamma w) \le \gamma T^{\pi_A}(\mathcal{P}_k, \gamma w)$ for every $\mathcal{P}_k$ and w. Thus, the result holds for $(\emptyset, 0)$. This concludes the proof. $\square$

5.2 Approximate Scheduling Policy

Here, we define a scheduling policy using the approximate probing and transmission strategy developed in the last subsection. Let us define a scheduling policy Δ_A as follows: $\pi^{\Delta_A}(t) = \pi_A(t)$ in every slot t. The policy Δ_A is called approximate as it satisfies for every queue length vector $\boldsymbol{Q}$

$$\mathbb{E}[\boldsymbol{Q}(t)\cdot \boldsymbol{B}^{\pi^{\Delta^\star}}(t)|\boldsymbol{Q}(t)=\boldsymbol{Q}] \le \gamma \mathbb{E}[\boldsymbol{Q}(t)\cdot \boldsymbol{B}^{\pi^{\Delta_A}}(t)|\boldsymbol{Q}(t)=\boldsymbol{Q}].$$

Thus, Δ_A approximates $\Delta^\star$ to the fraction $1/\gamma$. Now, using similar arguments to those used in the proof of Theorem 2, the following theorem can be shown.

THEOREM 6. *The policy Δ_A stabilizes every $\boldsymbol{a} \in \frac{1}{\gamma}\Lambda^\circ$.*

The above result shows that the policy Δ_A achieves $1/\gamma$ fraction of the system's stability region. Note that γ depends on $x_{\min}$, $x_{\max}$ and $R(x_{min}a)$. In high SNR regime, $R(x_{min}a)$ is large. In this scenario, given that $x_{\max}/x_{\min}$ is not too large, $\gamma \approx 1$. Then, by Theorem 6, Δ_A is stable for most of the stabilizable arrival rate vectors.

6. NUMERICAL EXPERIMENTS

In this section, we provide numerical experiments to illustrate the theoretical findings of the previous sections. Our objectives are here (i) to show the probing cost in terms of throughput (stability region) one has to pay when the CSIs have to be acquired, and (ii) to assess the performance of the approximate probing and transmission strategy proposed in Section 5. We consider systems with proportional fading as considered in Section 5: first we investigate the case of homogeneous fading and then that of heterogeneous fading.

6.1 Homogeneous Rayleigh fading

For this scenario, we assume that the channel condition of user i during slot t is $C_i(t)$ and that $C_i(t)$ are i.i.d. across users and slots. We consider Rayleigh fading, i.e., $C_i(t)$ is exponentially distributed (actually we consider a discretized version of the exponential distribution with 600 possible values). We further consider the case where the service rate of a user with channel condition c is given by $R(c) = \log(1 + c/N_0)$. The transmitted power is 40 dBm and the noise power is -100dBm. With such a system, an optimal probing strategy is to probe user channels in the decreasing order of their queue lengths as shown in Corollary 1. When no probing is required, i.e., when the CSIs are known in advance, a throughput optimal policyy is to transmit at each time t to the user with the highest product $R(C_i(t)) \times Q_i(t)$.

Assume that the expected arrival rate at the various buffers are identical and equal to a. we investigate the maximum total arrival rate $N \times a$ such that there exists a scheduling policy stabilizing the network. With perfect knowledge of the channel states, an alternative throughput optimal policy is obtained when, each slot t, one transmits to the user with the best channel. Indeed, since all users are equivalent, when one approaches the stability limit (i.e. a is close to the maximum arrival rate compatible with stability), then all queues tend to saturate simultaneously. Hence all users have packets in their buffer. As a result, the optimal scheduling decision is to maximize the system service rate and hence to transmit to user with the best channel. Under this policy, the stability condition can be written as:

$$N \times a < \mathbb{E}[\max_{i=1,\dots,N} R(C_i(t))].$$

Note that $\mathbb{E}[\max_{i=1,\dots,N} R(C_i(t))]$ scales as $\log\log(N)$ when N grows large. This contrasts with the case where the channel states have to be acquired, since in this case with a fixed probing time β, the total throughput of the system remains bounded as N grows large. In fact, one can easily observed that when N is large enough, the optimal probing and transmission strategy remains the same as N varies. In Figure 1, we present the maximum total throughput with or without the knowledge of channel states. In the latter case, we consider different values of β, 0.01, 0.05, 0.1, and 0.2. Figure 2 presents the optimal probing policy when N is large enough (greater than $\lfloor \beta/N \rfloor$). More specifically, it provides the thresholds (in terms of rate) below which it is optimal to probe the channel of another channel as a function of the number of users already probed. when the threshold reaches 0, it is always optimal to transmit.

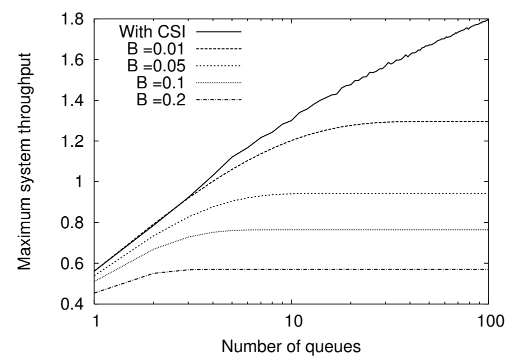

Figure 1: Maximum system throughput compatible with stability in the case of homogeneous Rayleigh fading. Results are in bit/s/Hz.

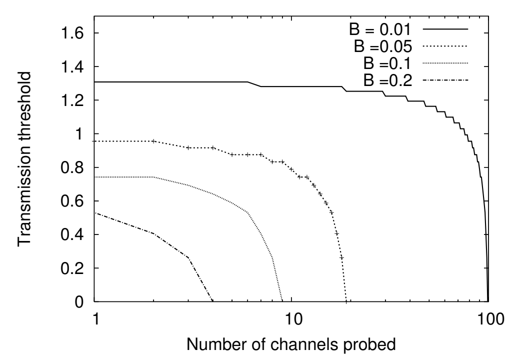

Figure 2: Successive thresholds of the optimal probing and transmission strategy.

Observe in Figure 1 that the price of acquiring the channel states is not negligible in general, unless the number of queues is limited and the time to probe a channel remains small. In Figure 2, note that as β decreases, the optimal policy consists in probing more and more user channels.

6.2 Heterogeneous fading

We now consider heterogeneous fading, and to simplify the exhaustive search of optimal probing and transmission strategy, we consider 2 users / queues only, each with 3 possible channel states. In Figure 3, we plot the Pareto-boundary of the stability region for different strategies: (1) a throughput-optimal strategy when the channel states are perfectly known, (2) a throughput-optimal policy when they have to be acquired (we use different values of β), and (3) the approximate policy (for the case $\beta = 0.1$).

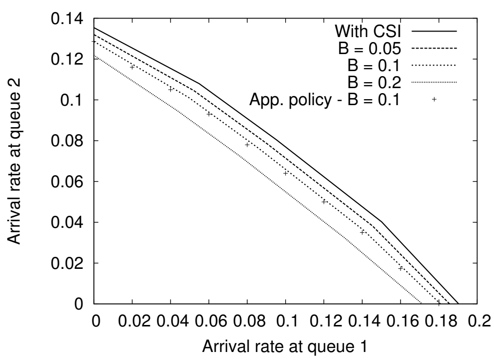

Figure 3: Stability region for a 2-user system with heterogeneous fading.

Note that the throughput region are all convex and that their Pareto-boundary is strictly concave. This strict concavity is due to the fact that the policies considered are really able to take advantage of multi-user diversity. Observe also that the approximate policy proves to be almost throughput optimal, even if in this case we are not in an high SNR regime. We know that in general the latter policy cannot be throughput optimal, however we believe that it performs quite well in most system scenarios.

7. CONCLUSION

We explored the problem of designing throughput optimal policies in wireless systems with limited information. As a first step, we showed that the policies, analogous to the adaptive queue length-based policies that are known to be throughput optimal when the CSIs are completely known, are also throughput optimal when the CSIs have to be acquired. But, to completely characterize these queue length-based policies when only a limited channel state information is available, we need to design joint probing and transmission strategies that maximize the weighted throughput of the system. We showed that this problem is a generalized version of the optimal stopping time problem, and that the solutions are, in general, difficult to obtain. Hence, we first derived certain structural properties of the optimal solution to reduce the computational complexity of designing these optimal strategies. Using these structural properties, we completely characterized optimal strategies in certain practically relevant cases. Moreover, when a complete characterization was not possible, we proposed an approximate scheduling policy that achieves a guaranteed fraction of the stability region. Our results are relevant to practical systems like broadcast with limited information (e.g. downlink of wireless LANs or cellular systems) and cognitive radio systems.

Acknowledgement

Work of P. Chaporkar, H. Asnani and A. Karandikar is supported by TTSL-IIT BOMBAY Center for Excellence in Telecom (TICET), Mumbai.

8. REFERENCES

[1] 3GPP TS 25.308. Utra high speed downlink packet access (hsdpa); overall description. In *Release 5*, 2003.

[2] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar. Providing quality of service over a shared wireless link. *IEEE Communications Magazine*, 39(2):150–154, Feb 2001.

[3] P. Bender, . Black, M. Grob, Padovani R, N. Sindhushyana, and S. Viterbi. Cdma/hdr: a bandwidth efficient high speed wireless data servicefor nomadic users. *IEEE Communications Magazine*, 38(7):70–77, Jul 2000.

[4] D. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 3rd edition, 2007.

[5] S. Borst and P. Whiting. Dynamic rate control algorithms for hdr throughput optimization. In *IEEE INFOCOM*, Anchorage, Alaska, Apr 2001.

[6] G. Caire, R. Müller, and R. Knopp. Multiuser diversity in wireless systems: Delay-limited vs. scheduling. *IEEE Transactions on Information Theory*, 2007.

[7] N. Chang and M. Liu. Optimal channel probing and transmission scheduling for opportunistic spectrum access. In *ACM MOBICOM'07*, pages 27–38, Montrï£¡al, Quï£¡bec, Canada, 2007.

[8] P. Chaporkar and A. Proutiere. Optimal joint probing and transmission strategy for maximizing throughput in wireless systems. *IEEE Journal on Selected Areas in Communications*, 26(18):1546–1556, Oct 2008.

[9] Yuan-Shih Chow, H. Robbins, and D. Siegmund. *Great Expectations: Theory of Optimal Stopping*. Houghton Mifflin Co, Jun 1972.

[10] G. Fayolle, V. A. Malyshev, and M. V. Menshikov. *Topics in the Constructive Theory of Countable Markov Chains*. Cambridge University Press, 1995.

[11] S. Guha, K. Mungala, and S. Sarkar. Jointly optimal transmission and probing strategies for multichannel wireless systems. In *CISS'06*, Mar 2006. (Invited Paper).

[12] S. Guha, K. Mungala, and S. Sarkar. Optimizing transmission rate in wireless channels using adaptive probes. In *ACM SIGMETRICS Performance Evaluation Review*, pages 381–382, New York, NY, 2006. (Poster Paper).

[13] J. Holtzman. Asymptotic analysis of proportional fair algorithm. In *IEEE PIMRC*, 2001.

[14] Z. Ji, Y. Yang, J. Zhou, M. Takai, and R. Bagrodia. Exploiting medium access diversity in rate adaptive wireless lans. In *ACM MOBICOM'04*, pages 345–359, Philadelphia, PA, 2004.

[15] K. Kar, X. Luo, , and S. Sarkar. Throughput-optimal scheduling in multichannel access point networks

under infrequent channel measurements. In *IEEE INFOCOM*, Anchorage, USA, May 2007.

[16] R. Knopp and P. Humblet. Information capacity and power control in single cell multiuser communications. In *IEEE ICC'95*, pages 331–335, Seattle, WA, Jun. 1995.

[17] H.J. Kushner and P.A. Whiting. Asymptotic properties of proportional fair sharing algorithms. In *40th Annual Allerton Conf. Commun. Control Comp.*, 2002.

[18] M. Neely and E. Modiano. Capacity and delay tradeoffs for ad-hoc mobile networks. *IEEE Trans. on Inform. Theory*, 51(6):1917–1937, Jun 2005.

[19] H. Robbins. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.*, 1952.

[20] A. Sabhrawal, A. Khoshnevis, and E. Knightly. Opportunistic spectral usage: bounds and a multi-band csma/ca protocol. *IEEE/ACM Trans. on Networking (TON)*, 5(3):533–545, Jun 2007.

[21] S. Shakkottai and A. Stolyar. Scheduling for multiple flows sharing a time-varying channel: The exponential rule. *American Mathematical Society Translations, Series 2, A volume in memory of F. Karpelevich*, 2002.

[22] A. Stolyar. Greedy primal-dual algorithm for dynamic resource allocation in complex networks. *Queueing Systems*, 54(3):203–220, 2006.

[23] A. Stolyar and K. Ramanan. Largest weighted delay first scheduling: Large deviations andoptimality. *Annals of Applied Probability*, 11(1):1–48, 2001.

[24] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. on Automatic Control*, 37(12):1936–1948, Dec 1992.

[25] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Trans. Infor. Theory*, 39(2):466–478, Mar 1993.

[26] V. Tsibonis, L. Georgiadis, and L. Tassiulas. Exploiting wireless channel state information for throughput maximization. In *IEEE Infocom*, 2003.

[27] P. Viswanath, D. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory*, 48(6), June 2002.

[28] P. Whittle. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25:287–298, 1988. Special Issue: A Celebration of Applied Probability.

APPENDIX

Supporting Lemmas for Theorem 3:

LEMMA 3. *There exists $w_{\max}(\mathcal{P}_k)$ such that $\mathcal{D}_k = \{w : w \geq w_{\max}(\mathcal{P}_k)\}$.*

PROOF. Consider a user with $w \in \mathcal{D}_k$. It suffices to prove that for all $w' > w$, $w' \in \mathcal{D}_k$. Since $w \in \mathcal{D}_k$,

$$(1-k\beta)w \geq (1-(k+1)\beta)\max_{i\in\overline{\mathcal{P}}_k}\{\mathbb{E}[w \vee W_i]\}$$

$$(1-k\beta)\max_{i\in\mathcal{P}_k}\{\mathbb{E}_i[w-(w\vee W_i)]\} \geq -\beta\max_{i\in\overline{\mathcal{P}}_k}\{\mathbb{E}_i[(w'\vee W_i)]\}$$

as $\max\{w', W_i\} \geq \max\{w, W_i\}$. It can be seen that $w' - (w' \vee W_i) \geq w - (w \vee W_i)$. Hence we obtain,

$$(1-k\beta)\max_{i\in\overline{\mathcal{P}}_k}\{\mathbb{E}_i[w'-\max\{w,W_i\}]\} \geq -\beta\max_{i\in\overline{\mathcal{P}}_k}\{\mathbb{E}_i[\max\{w',W_i\}]\}$$

which implies $w' \in \mathcal{D}_k$. □

LEMMA 4. *For any sequence of sets of probed users such that $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{i\}$ for some $i \in \overline{\mathcal{P}}_k$ for $k \in \{0,\ldots,N-1\}$, we have: for all k, $\mathcal{D}_k \subseteq \mathcal{D}_{k+1}$.*

PROOF. The proof is by contradiction. Assume that there is some w, such that $w \in \mathcal{D}_k$, but $w \notin \mathcal{D}_{k+1}$. Thus,

$$(1-(k+1)\beta)w < (1-(k+2)\beta)\max_{i\in\overline{\mathcal{P}}_{k+1}}\{\mathbb{E}_i[\max\{w,W_i\}]\}$$

Thus,

$$\begin{aligned} T_{tr}(\mathcal{P}_k,u) - T_{pr,tr}(\mathcal{P}_k,u) &< \beta\max_{i\in\overline{\mathcal{P}}_{k+1}}\mathbb{E}_i[(w-\max\{w,W_i\})] \\ &\leq 0, \end{aligned}$$

which is a contradiction as assumed $w \in \mathcal{D}_k$. □

Supporting Lemmas for Theorem 4:

LEMMA 5. *For all $w \leq \alpha_j$, we have $G(w) = G(\alpha_j)$.*

PROOF. We note that if $w \leq \alpha$, then $T^{(b)}(w)$ and $T^{(a)}(w)$ are independent of w and hence is $G(w)$. Now let us assume $\alpha \leq w \leq \alpha_j$, The first terms in both $T^{(b)}(w)$ and $T^{(a)}(w)$ do not depend upon w, while the second terms are respectively,

$$a_{k+2}\int_0^{\alpha_i} dF_i(w_i)\int_0^{\infty}\mathbf{1}_{w_i\vee w_j\geq\alpha}\max\{w_i,w_j\}$$
$$-\int\int_{\Gamma(\alpha,w)} dF_i(w_i)dF_j(w_j)(\max\{w_i,w_j\}-w).$$

and

$$a_{k+2}\int_0^{\alpha_j} dF_j(w_j)\int_0^{\infty}\mathbf{1}_{w_j\vee w_i\geq_W\alpha}\max\{w_j,w_i)\}$$
$$-\int\int_{\Gamma(\alpha,w)} dF_i(w_i)dF_j(w_j)(\max\{w_i,w_j\}-w)$$

where $\Gamma(\alpha,w) = \{(w_i,w_j) : \alpha \leq w_i, w_j \leq w\}$. Hence difference $G(w)$ is still independent of w. □

LEMMA 6. *For all w such that $\alpha_j \leq w \leq \alpha_i$, $G(w) \geq 0$.*

PROOF. We prove the result in discrete setting, proof for continuous setting is similar. We employ the perturbation approach.Without loss of generality, let $\mathbf{N}$ be the channel state space. Denote by $p_i(l)$ the probability of the user i to be in state l. We observe that when $F_i = F_j$ result holds. Now assuming that result is true for F_j and show that stochastically increasing F_j does not change this conclusion. We use F_j^+ defined by: for $\epsilon > 0$, for a particular $l_0 \in \mathbf{N}$, $p_j^+(l_0) = p_j(l_0) - \epsilon$, $p_j^+(l_0+1) = p_j(l_0+1) + \epsilon$ and for all $l \neq l_0, l_0+1, p_j^+(l) = p_j(l)$. It can be shown that using these kind of perturbations we can start from F_i and modify it to obtain F_j.

$$G^+(w) \geq G(w) + o(\epsilon) + \epsilon \times \mathbf{1}_{l_0\geq u}(a_{k+1} - a_{k+2}F_i(\alpha_i \vee l_0)). \quad (9)$$

From (9) we conclude $G^+(w) \geq 0$. From this for $\alpha \leq w \leq \alpha_j$, we have $G(w) = G(\alpha_j) \geq 0$. □