

# Stability Region of Opportunistic Scheduling in Wireless Networks

Xuan Wang, *Student Member, IEEE*, and Lin Cai, *Senior Member, IEEE*

**Abstract**—The stability regions of two opportunistic scheduling policies, the utility-based (UB) scheduling and the channel-rate-based (CRB) scheduling, in wireless networks are discussed, respectively. The UB scheduling is a generalized proportional fair scheduling in an unsaturated system, and the CRB scheduling is a variant of the UB scheduling. We give the closed-form expression of the stability region of the CRB scheduling, and a numerical method to obtain the stability region of the UB scheduling. Both of the two scheduling policies are not throughput-optimal, and thus in general their stability regions are less than the ergodic capacity region. With the CRB scheduling, the stability region is a convex hull, while with the UB scheduling, the stability region is generally even non-convex and may exhibit some undesirable properties, such as decreasing the traffic of one flow may lead another flow being unstable, and proportionally decreasing the traffic of all flows may lead a stable system to be unstable. We further show that, as long as the arrival rate lies inside the ergodic capacity region, we can assign a proper weight to each user, and based on the weighted UB/CRB scheduling policies, the system can be stabilized. Detailed numerical examples and simulations are given to illustrate the stability region of the two policies and validate our analysis.

## I. INTRODUCTION

Scheduling and resource allocation is one of the most important tasks in the operation of wireless networks, especially for infrastructure-based wireless networks, since the system performance is mainly determined by the scheduling policy in a multi-user system.

Traditionally, scheduling is a link layer function, which is performed separately from the lower layer functions. For wireless networks, [1] proposed an opportunistic scheduling policy in the scenario that multiple users share the channel, which can enhance the system performance by exploiting the randomness of fading channel. In this work, fading, due to the user mobility and multipath propagation, was firstly treated as a constructive factor to the system. Thereafter, opportunistic scheduling was applied to downlink in [2].

The opportunistic scheduling proposed in [1], [2] has been generalized to the utility-based scheduling in [3] which aims to maximize a pre-defined utility based on the long-term achievable throughput. Based on the stochastic approximation, the convergence of such policy is guaranteed under a mild condition [4], [5]. The work has been further extended to different network scenarios, such as cooperative networks [6], or networks with different wireless techniques, such as the downlink and the uplink of an orthogonal frequency-division multiplexing (OFDM) system [7], [8], [9].

All these works designed the scheduler based on an assumption that each user always has sufficient data to transmit. The assumption simplifies the problem, while as shown in [10], these kinds of schedulers may lead the system to be unstable, while the system in the same circumstance can be stabilized by other scheduling policies, such as max-weight scheduling [11]. The key reason here is because, without considering the stochastic characteristic of incoming traffic, although the arrival rate lies inside the ergodic capacity region, the tie-breaking rule used in the above utility-based scheduling policies is not efficient as these policies schedule some users too frequently and lose the chance to explore the multi-user diversity gain, and thus they are not throughput-optimal.

Little work has been done in quantifying the stability region of the opportunistic scheduling policies. The stability region of an opportunistic scheduling policy in a two-user wireless network with i.i.d. Bernoulli arrival traffic was derived in [12]. Different from the general utility-based scheduling, the scheduler discussed in [12] is a normalized signal-to-noise-ratio (SNR) one, where the user is scheduled based on the normalized instantaneous SNR. The author observed that the stability region is less than the ergodic capacity region, while by varying the normalized factor, the union of the resultant stability region is equal to the ergodic capacity region. Note that, with the identical normalized factor, the scheduler is able to explore the maximal multi-user diversity, but is not easy to explore other features, such as fairness. By changing the normalized factor, the fairness feature can be implicitly explored, while it is unclear how to design the normalized factor for a specific fairness objective. Also, the prior knowledge assumption of the channel in [12] may bring hardness to implement such policy. In [13], the authors discussed the two-user stability region in a static channel configuration with concurrent transmissions. The scheduler discussed is a partial distributed scheduler, combining the user coordination with an Aloha media access control (MAC), which may not be a suitable choice for a centralized wireless network due to the low channel efficiency of the Aloha MAC.

In this work, we quantify the stability region of two opportunistic scheduling policies with a general traffic arrival in a wireless system with  $N$  users. The two scheduling policies include a utility-based (UB) one and a channel-rate-based (CRB) one. The CRB scheduling can be viewed as a variant of the UB scheduling, by treating an intermediate control variable differently. For the UB scheduling, the explicit closed-form stability region generally cannot be obtained, while we develop a theorem to examine the stability of a system given the arrival rate, and a numerical method is provided to obtain the stability

X. Wang and L. Cai are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada, V8W 3P6. e-mail: {xuan\_wang, cai}@ece.uvic.ca.

region in a two-user system. We further study the properties of the stability region of the UB scheduling, and show that it is generally non-convex and may also exhibit some undesirable features. For instance, decreasing the arrival rate of one user may lead the system to be unstable. For the CRB scheduling, we obtain the closed-form expression of the stability region, which is a convex hull. Besides the stability region, we further study the extended stability region by giving a weight to each user. The results show that by varying the weight assigned to each user, the union of the resultant stability region is equal to the ergodic capacity region, for both scheduling policies. This suggests as long as the system can be stabilized, by assigning a proper weight to each user, using a non-throughput-optimal scheduling may also stabilize the system.

It is further noted that, the results of the CRB scheduling is similar to the work in [12], while our work is more general. We use a more general traffic model, consider a general  $N$  user system, and discuss a scheduling algorithm that can be easily designed to achieve certain utility objective.

The rest of the paper is organized as follows. Sec. II introduces the system models, including channel model, queueing model and scheduling policies. In Sec. III and Sec. IV, the stability regions of the CRB scheduling and UB scheduling are presented, respectively. The extended stability region is discussed in Sec. V. Evaluations are shown in Sec. VI, followed by the conclusion and discussion.

In the following, bold face letters represent vectors, and calligraphic letters represent sets.

## II. SYSTEM MODELS

### A. Channel Model

The system has one server who has packets to transmit through a shared wireless channel to  $N$  independent users. The set of users is denoted by  $\mathcal{N} = \{1, 2, \dots, N\}$ . The power set of  $\mathcal{N}$  is denoted by  $\mathbb{S}$ , and the cardinality of  $\mathbb{S}$  is  $|\mathbb{S}| = 2^N$ . We use  $S_i$  to denote the  $i$ -th element in  $\mathbb{S}$ .

We assume that the shared wireless channel is time slotted block fading channel. The set of channel state is finite, which is represented as  $\mathcal{M} = \{1, 2, \dots, M\}$ . Within each time slot, the channel state is constant. Crossing time slots, certain rule is used to govern the transition of the channel state. There is a vector of rates  $\mathbf{u}^m = (u_1^m, u_2^m, \dots, u_N^m)$  associated with each channel state  $m \in \mathcal{M}$ . The element  $u_i^m \in \mathbb{N} \cup \{0\}$  means the number of packets that can be transmitted if the time slot is all allocated to user  $i$  in state  $m$ .

We further assume that the shared wireless channel state process is an irreducible discrete-time Markov chain with the state space  $\mathcal{M}$ . The stationary distribution of this Markov chain is denoted as  $\boldsymbol{\pi} = (\pi^1, \pi^2, \dots, \pi^M)$ .

The capacity region of the system in state  $m$  is denoted as

$$\mathcal{C}_{\mathcal{N}}^m = \bigcup_{\sum_i t_i^m = 1} (u_1^m t_1^m, \dots, u_N^m t_N^m),$$

where  $t_i^m$  is the time portion allocated to user  $i$  in state  $m$ .

The ergodic capacity region of the system is obtained as:

$$\bar{\mathcal{C}}_{\mathcal{N}} = \bigcup_{\sum_i t_i^m = 1} \left( \sum_m u_1^m t_1^m \pi^m, \dots, \sum_m u_N^m t_N^m \pi^m \right). \quad (1)$$

Given the user set  $\mathcal{A}$ , the corresponding capacity region in state  $m$  and the ergodic capacity region can be obtained by assigning  $t_i^m = 0$  for all  $i \notin \mathcal{A}$ , and are denoted by  $\mathcal{C}_{\mathcal{A}}^m$  and  $\bar{\mathcal{C}}_{\mathcal{A}}$  respectively. Denote  $C_{\mathcal{A}}(t)$  as the capacity region of user set  $\mathcal{A}$  in time slot  $t$ . Since state  $m$  and time slot  $t$  are associated, so if the state in  $t$  is  $m$ , we have  $\mathcal{C}_{\mathcal{A}}^m = C_{\mathcal{A}}(t)$ .

### B. Queueing Model

Data packets are arrived randomly and queued up in an infinite buffer reserved for each user. The packet arrival process is considered as a stationary ergodic stochastic process with finite moments. The state of the  $i$ -th buffer is the queue length and denoted by  $q_i(t)$ . All queue states form a vector  $\mathbf{q}(t) \in \mathbb{R}_+^N$ , and are updated by

$$\mathbf{q}(t+1) = [\mathbf{q}(t) - \mathbf{r}(t) + \mathbf{a}(t)]^+, \quad (2)$$

where  $[x]_i^+ = \max\{0, x_i\}, \forall i \in \mathcal{N}$ ,  $\mathbf{r}(t) \in \mathbb{R}_+^N$  is the amount of transmitted data that is determined by the scheduling decision, and  $\mathbf{a}(t) \in \mathbb{R}_+^N$  is the amount of arrived data in time  $t$ , which is a bounded random variable. The average arrival and service rates are  $\boldsymbol{\lambda} = \mathbb{E}_t[\mathbf{a}(t)]$  and  $\boldsymbol{\mu} = \mathbb{E}_t[\mathbf{r}(t)]$ , respectively.

### C. Scheduling Policy

We assume that at the beginning of each time slot, the server can observe the state of the channel and allocate the resource based on the observation.

Under the assumption that each user always has enough data to transmit, a utility-based scheduling policy, which is a generalized proportional fair scheduling [3], [5], allocates the rate to user in time slot  $t$  based on the following problem:

$$\mathbf{r}(t) = \arg \max_{\boldsymbol{\eta} \in \mathcal{C}_{\mathcal{N}}(t)} \sum_{i \in \mathcal{N}} f(R_i(t)) \eta_i, \quad (3)$$

with ties being broken randomly, where function  $f$  is a derivative of a strictly concave smooth utility function  $U$ ,  $R_i(t)$  is the smoothed rate measurement of user  $i$  in time slot  $t$ , which can be updated by an exponentially weighted moving average algorithm [5]

$$\mathbf{R}(t) = \mathbf{R}(t-1) + \epsilon(\mathbf{r}(t) - \mathbf{R}(t-1)),$$

where  $\epsilon$  is the step size.

According to [4], [5], by choosing a proper step size  $\epsilon$ ,  $\mathbf{R}(t)$  weakly converges to the average allocated rate  $\mathbf{R}_{\mathcal{N}}$  which can be obtained based on the following problem

$$\mathbf{R}_{\mathcal{N}} = \arg \max_{\boldsymbol{\eta} \in \mathcal{C}_{\mathcal{N}}} \sum_{i \in \mathcal{N}} U(\eta_i).$$

Note that the online algorithm (3) cannot be directly used in a system without the assumption of enough backlogs, since otherwise it may allocate the resource to users with no packet to transmit. With some modifications to (3), two scheduling policies, the UB and the CRB scheduling, can be obtained for a system with stochastic arrival traffic.

1) *The UB Scheduling*: In time slot  $t$ , a user set  $\mathcal{A}(t)$  is selected satisfying the condition that the queue length of each user in  $\mathcal{A}(t)$  is sufficiently large, for instance  $q_i(t) \geq q_i^{th}$ , where  $q_i^{th}$  is the queue length threshold for user  $i$ . This treatment avoids the wireless resource been wasted that choosing a user without enough data to transmit. The specific value of  $q_i^{th}$  does not affect the stability region, as long as it is sufficiently large. With such treatment, the queue length dynamic in (2) becomes

$$\mathbf{q}(t+1) = \mathbf{q}(t) - \mathbf{r}(t) + \mathbf{a}(t). \quad (4)$$

Then the rate allocated to the user in  $\mathcal{A}(t)$  is

$$\mathbf{r}_{\mathcal{A}(t)}^{\text{UB}}(t) = \arg \max_{\boldsymbol{\eta} \in \mathcal{C}_{\mathcal{A}(t)}(t)} \sum_{i \in \mathcal{A}(t)} f(R_i^{\text{UB}}(t)) \eta_i, \quad (5)$$

with ties being broken randomly, and the rate allocated to the user in  $\mathcal{N} \setminus \mathcal{A}(t)$  is 0. Using  $\mathbf{r}^{\text{UB}}(t)$  to denote the allocated rate in time slot  $t$ , then  $R_i^{\text{UB}}$  is updated based on

$$\mathbf{R}^{\text{UB}}(t) = \mathbf{R}^{\text{UB}}(t-1) + \epsilon(\mathbf{r}^{\text{UB}}(t) - \mathbf{R}^{\text{UB}}(t-1)),$$

which is used to track the average throughput of the system.

2) *The CRB Scheduling*: For the CRB scheduling, in time slot  $t$ , based on the same method as the UB scheduling, we select the candidate user set  $\mathcal{A}(t)$ . The rate allocated to the user in  $\mathcal{A}(t)$  is based on

$$\mathbf{r}_{\mathcal{A}(t)}^{\text{CRB}}(t) = \arg \max_{\boldsymbol{\eta} \in \mathcal{C}_{\mathcal{A}(t)}(t)} \sum_{i \in \mathcal{A}(t)} f(R_i^{\text{CRB}}(t)) \eta_i, \quad (6)$$

with ties being broken randomly, and the rate allocated to user in  $\mathcal{N} \setminus \mathcal{A}(t)$  is 0. We use  $\mathbf{r}^{\text{CRB}}(t)$  to denote the allocated rate in time slot  $t$ .

Different from the UB scheduling, in the CRB scheduling,  $\mathbf{R}^{\text{CRB}}(t)$  is used to track the average channel-rate, and is updated by

$$\mathbf{R}^{\text{CRB}}(t) = \mathbf{R}^{\text{CRB}}(t-1) + \epsilon(\mathbf{r}(t) - \mathbf{R}^{\text{CRB}}(t-1)),$$

where  $\mathbf{r}(t)$  is the solution to (3).

How to update  $\mathbf{R}^{\text{UB}}(t)$  and  $\mathbf{R}^{\text{CRB}}(t)$  is the only difference between the UB and the CRB scheduling policies. For the CRB scheduling, the update is independent of the scheduling decision, while for the UB scheduling, the update depends on the scheduling decision in each time slot.

As shown in [5], under a mild condition,  $\mathbf{R}^{\text{UB}}(t)$  and  $\mathbf{R}^{\text{CRB}}(t)$  are all weakly converge. In the following, we only consider the case that  $\mathbf{R}^{\text{UB}}(t)$  and  $\mathbf{R}^{\text{CRB}}(t)$  converge.

By abusing the notation a bit, we also use  $\mathbf{R}_{\mathcal{A}}$  to denote the rate vector of  $N$  users and satisfies  $\forall j \notin \mathcal{A}, R_j = 0$ , i.e.,  $\mathbf{R}_{\mathcal{A}}^T = [\mathbf{R}_{\mathcal{A}}^T \quad \mathbf{R}_{\mathcal{N} \setminus \mathcal{A}}^T]$ , where  $\mathbf{R}_{\mathcal{N} \setminus \mathcal{A}} = \mathbf{0}$ .

#### D. Stability

In this paper, we apply the stability definition as it is used in [14].

**Definition 1.** A system of queues is said to be strongly stable if

$$\limsup_{t \rightarrow \infty} \mathbb{E}[\|\mathbf{q}(t)\|] < \infty,$$

where  $\|\mathbf{q}(t)\|$  is the norm of vector  $\mathbf{q}(t)$ .

Since we only consider the case that  $\mathbf{R}^{\text{CRB}}(t)$  or  $\mathbf{R}^{\text{UB}}(t)$  converges, and after the convergence of  $\mathbf{R}^{\text{CRB}}(t)$  or  $\mathbf{R}^{\text{UB}}(t)$ , the scheduling decision is only related to the current channel state and the queue state. Therefore, we can simplify the stability condition.

First, for the CRB scheduling we assume that at time slot 0,  $\mathbf{R}^{\text{CRB}}(t)$  has converged. Due to (4), when  $t$  is sufficiently large, we have

$$\begin{aligned} \mathbf{q}(t) &= \mathbf{q}(0) - \sum_{\tau=0}^{t-1} \mathbf{r}(\tau) + \sum_{\tau=0}^{t-1} \mathbf{a}(\tau) \\ &= \mathbf{q}(0) - \boldsymbol{\mu}t + \boldsymbol{\lambda}t \geq 0, \end{aligned}$$

which suggests for all  $i$ ,  $\lambda_i \geq \mu_i$ .

Since the dimension of  $\mathbf{q}(t)$  is finite, here we only consider  $L_1$  norm of  $\mathbf{q}(t)$ , and we have

$$\begin{aligned} \|\mathbf{q}(t)\| &= \|\mathbf{q}(0) - \boldsymbol{\mu}t + \boldsymbol{\lambda}t\| \\ &= t \sum_i (\lambda_i - \mu_i) + \sum_i q_i(0). \end{aligned}$$

Therefore

$$\limsup_{t \rightarrow \infty} \mathbb{E}[\|\mathbf{q}(t)\|] < \infty,$$

requires

$$\begin{aligned} \limsup_{t \rightarrow \infty} \mathbb{E}[t \sum_i (\lambda_i - \mu_i) + \sum_i q_i(0)] \\ = \lim_{t \rightarrow \infty} t \sum_i (\lambda_i - \mu_i) + \sum_i \mathbb{E}[q_i(0)] \leq \infty, \end{aligned}$$

which suggests  $\sum_i (\lambda_i - \mu_i) \leq 0$ .

In summary, the stability of the system requires  $\boldsymbol{\lambda} = \boldsymbol{\mu}$ , i.e., the average arrival rate is the identical to the average throughput. For the UB scheduling, based on the same argument, we can have the same result.

We further define the stability region of the system as follows:

**Definition 2.** The stability region of a system with scheduling policy  $p$  is defined as  $\Lambda^p$ , and we have  $\forall \boldsymbol{\lambda} \in \Lambda^p$ , the system is strongly stable;  $\forall \boldsymbol{\lambda} \notin \Lambda^p$ , the system is not strongly stable;

Without confusing, we also use the stability region of scheduling policy  $p$  to refer to the stability region of a system with scheduling policy  $p$ .

### III. STABILITY REGION OF THE CRB SCHEDULING

We first tackle a simple case, the static channel case ( $M = 1$ ), to obtain the stability region. Thereafter, the general stochastic channel case is discussed. We show that by replacing the capacity region with the ergodic capacity region, all the discussions for the static channel case also hold for the stochastic channel case.

### A. Static Channel Case

Since the channel only has one state, so we have  $\forall \mathcal{A} \in \mathbb{S}$ ,  $\bar{\mathcal{C}}_{\mathcal{A}} = \mathcal{C}_{\mathcal{A}}^1$  and  $\mathcal{C}_{\mathcal{A}}(t) = \mathcal{C}_{\mathcal{A}}^1$ .

**Theorem 1.** *The stability region of the CRB scheduling policy is  $\Lambda^{CRB}$ , and*

$$\Lambda^{CRB} = \bigcup_{\sum_i t_i = 1} \sum_{i=1}^{|\mathbb{S}|} \mathbf{R}_{\mathcal{S}_i}^{CRB} t_i, \quad (7)$$

where  $t_i \in \mathbb{R}_+ \cup \{0\}$ ,

$$\mathbf{R}_{\mathcal{S}_i}^{CRB} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{S}_i}} \sum_{j \in \mathcal{S}_i} f(R_j^{CRB}) r_j, \quad (8)$$

and  $\mathbf{R}^{CRB} = \mathbf{R}_{\mathcal{N}}^{CRB}$ .

*Proof:* Since the scheduler is the CRB one, the update of  $\mathbf{R}^{CRB}(t)$  is independent of the scheduling decision in each time slot, and  $\mathbf{R}^{CRB}(t)$  converges to  $\mathbf{R}_{\mathcal{N}}^{CRB}$ , i.e.,  $\mathbf{R}^{CRB} = \mathbf{R}_{\mathcal{N}}^{CRB}$ .

Comparing (6) with (8), we can conclude that

$$\mathbb{E}_t[\mathbf{r}_{\mathcal{A}(t)}^{CRB}(t)] = \mathbb{E}_i[\mathbf{R}_{\mathcal{S}_i}^{CRB}]$$

if  $\mathbf{R}^{CRB}(t)$  converges. This is because  $\mathbf{R}_{\mathcal{S}_i}^{CRB}$  is the average throughput of user set  $\mathcal{S}_i$  over time, and by taking the expectation over  $i$ ,  $\mathbb{E}_i[\mathbf{R}_{\mathcal{S}_i}^{CRB}]$  is the average throughput of the system. Since  $\mathbf{r}_{\mathcal{A}(t)}^{CRB}(t)$  is the throughput of the system in time slot  $t$ , by taking expectation over time,  $\mathbb{E}_t[\mathbf{r}_{\mathcal{A}(t)}^{CRB}(t)]$  is also the average throughput of the system.

If the system is stable, the average arrival rate should be equal to the average throughput, i.e.,

$$\boldsymbol{\lambda} = \mathbb{E}_t[\mathbf{r}_{\mathcal{A}(t)}^{CRB}(t)] = \mathbb{E}_i[\mathbf{R}_{\mathcal{S}_i}^{CRB}],$$

therefore the necessary condition for the system to be stable is that we can find a  $\mathbf{t}$  ( $\sum_i t_i = 1$ ) such that

$$\boldsymbol{\lambda} = \sum_{i=1}^{|\mathbb{S}|} \mathbf{R}_{\mathcal{S}_i}^{CRB} t_i = \mathbb{E}_i[\mathbf{R}_{\mathcal{S}_i}^{CRB}],$$

which is equivalent to  $\boldsymbol{\lambda} \in \Lambda^{CRB}$ .

The sufficient condition can be proved by contradiction. Suppose  $\boldsymbol{\lambda} \in \Lambda^{CRB}$ , but the system is not stable, and therefore at least one queue is unstable. Suppose that the queues in set  $\mathcal{Q}$  are unstable, and the queues in set  $\mathcal{N} \setminus \mathcal{Q}$  are stable. Since queue  $i \in \mathcal{Q}$  is unstable, we have  $\mathbb{E}[q_i(t)] \rightarrow \infty$  which suggests that user  $i$  is always scheduled. Suppose that user set  $\mathcal{D}$  is the scheduled user set, then we have  $\mathcal{Q} \subseteq \mathcal{D}$ . We further construct a set  $\mathbb{D}$  which is made up of all  $\mathcal{D}$ . Therefore the average throughput of the system is

$$\bar{\mathbf{R}} = \sum_{\mathcal{D} \in \mathbb{D}} \pi_{\mathcal{D}} \mathbf{R}_{\mathcal{D}}^{CRB},$$

and  $\sum_{\mathcal{D} \in \mathbb{D}} \pi_{\mathcal{D}} = 1$ . Because  $\mathcal{D}$  is nonempty, we have

$$\bar{\mathbf{R}} \in \Lambda^{CRB},$$

and for any  $\boldsymbol{\epsilon}$ , with  $\sum_i \epsilon_i > 0$ ,  $\boldsymbol{\epsilon}_i \in \mathbb{R}_+ \cup \{0\}$ ,

$$\bar{\mathbf{R}} + \boldsymbol{\epsilon} \notin \Lambda^{CRB}.$$

Due to the assumption of the stability of the system, we have

$$\begin{cases} \bar{R}_i < \lambda_i, \forall i \in \mathcal{Q}, \\ \bar{R}_i = \lambda_i, \forall i \in \mathcal{N} \setminus \mathcal{Q}, \end{cases}$$

which means there exists an  $\boldsymbol{\epsilon}$  that  $\bar{\mathbf{R}} + \boldsymbol{\epsilon} = \boldsymbol{\lambda} \in \Lambda^{CRB}$ , which is contracted with (9). Thus the assumption cannot hold, and we have proved  $\forall \boldsymbol{\lambda} \in \Lambda^{CRB}$ , the system is stable.

In summary, the stability region of the system is  $\Lambda^{CRB}$ . ■

Here, due to the special property of the capacity region, the stability region equals the capacity region. Note that the capacity region is a Euclidean simplex with  $N + 1$  vertices and each vertex represents a rate vector. Suppose the  $N + 1$  vertices make up a set  $\mathcal{V}$ . Since  $\mathbf{R}_{\mathcal{S}_i}$  is on the boundary of the capacity region  $\mathcal{C}_{\mathcal{S}_i}$  thus lies in the hyperplane determined by the points in  $\mathcal{V}$ . Since the stability region is the convex hull of  $\mathbf{R}_{\mathcal{S}_i}$ , which equals the convex hull of  $\mathcal{V}$ , i.e., the capacity region.

### B. Stochastic Channel Case

For the stochastic channel case, we have a similar result as in the static channel case.

**Theorem 2.** *Theorem 1 holds for the stochastic channel case.*

*Proof:* Similar to the static channel case,  $\mathbf{R}^{CRB}(t)$  converges to

$$\mathbf{R}^{CRB} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{N}}} \sum_{j \in \mathcal{N}} U(r_j),$$

and we have

$$\mathbf{r}_{\mathcal{S}_i}^{CRB}(t) = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{S}_i}(t)} \sum_j f(R_j^{CRB}) r_j.$$

Taking expectation over time, we have

$$\mathbf{R}_{\mathcal{S}_i}^{CRB} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{S}_i}} \sum_j f(R_j^{CRB}) r_j.$$

Also as

$$\mathbf{R}^{CRB} = \mathbf{R}_{\mathcal{N}}^{CRB},$$

and based on the same approach as in the static channel case, we can prove that Theorem 1 holds for the stochastic channel case. ■

Worth to note that, different from the static channel case where the stability region is identical to the capacity region, the stability region in the stochastic channel case is generally less than the capacity region due to the fact that the ergodic capacity region is a convex polytope, but not necessarily a Euclidean simplex.

## IV. STABILITY REGION OF THE UB SCHEDULING

Similar to the discussion of the CRB scheduling, we first discuss the simple case, the static channel case, and then study the complicated stochastic channel case. Furthermore, we show that, the results obtained in the static channel case can be directly used in the stochastic channel case, by replacing the capacity region with the ergodic one. Different from the CRB scheduling, where the stability region can be easily obtained

in closed-form, the stability region of the UB scheduling generally cannot be obtained in closed-form, and therefore we develop a numerical method to tackle the two-user case.

#### A. Static Channel Case

According to [5],  $\mathbf{R}^{\text{UB}}(t)$  converges to the average throughput. Once  $\mathbf{R}^{\text{UB}}(t)$  converges, we have

$$\mathbf{R}_{S_i}^{\text{UB}} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{S_i}} \sum_j f(R_j^{\text{UB}}) r_j,$$

and

$$\mathbf{R}^{\text{UB}} = \mathbb{E}_i[\mathbf{R}_{S_i}^{\text{UB}}].$$

Note that generally

$$\mathbf{R}^{\text{UB}} \neq \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{N}}} \sum_{i \in \mathcal{N}} U(r_i),$$

and  $\mathbf{R}^{\text{UB}}$  may not lie on the boundary of the capacity region.

We have the following theorem to verify whether a system with a specific arrival rate vector is stable or not.

**Theorem 3.** *A system using the UB scheduling policy with average arrival rate  $\lambda$  is stable if and only if  $\lambda \in \tilde{\Lambda}^{\text{UB}}(\lambda)$ , where*

$$\tilde{\Lambda}^{\text{UB}}(\lambda) = \bigcup_{\sum_i t_i = 1} \sum_{i=1}^{|\mathcal{S}|} \mathbf{R}_{S_i}^{\text{UB}} t_i,$$

and

$$\mathbf{R}_{S_i}^{\text{UB}} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{S_i}} \sum_j f(\lambda_j) r_j.$$

*Proof:* Suppose that the system is stable, and then we have  $\lambda = \mathbf{R}^{\text{UB}}$ . Thus, the average rate allocated to user set  $S_i$  is

$$\mathbf{R}_{S_i}^{\text{UB}} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{S_i}} \sum_i f(R_i^{\text{UB}}) r_i = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{S_i}} \sum_i f(\lambda_i) r_i,$$

if  $S_i$  is scheduled. Since  $\mathbf{R}^{\text{UB}} = \mathbb{E}_i[\mathbf{R}_{S_i}^{\text{UB}}]$ , we have  $\lambda = \mathbb{E}_i[\mathbf{R}_{S_i}^{\text{UB}}] \in \tilde{\Lambda}^{\text{UB}}(\lambda)$ .

Based on the same argument as that in Sec. III, we can prove that  $\forall \lambda \in \Lambda^{\text{UB}}$ , the system is stable. Thus, the theorem is proved.  $\blacksquare$

Based on the above theorem, we have the following corollary.

**Corollary 4.** *If  $\lambda = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{N}}} U(r_i)$ , then the system is stable.*

*Proof:* Since

$$\lambda = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{N}}} U(r_i), \quad (9)$$

we have

$$\lambda = \mathbf{R}_{\mathcal{N}}^{\text{UB}} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{N}}} \sum_{i \in \mathcal{N}} f(\lambda_i) r_i,$$

which means  $\lambda \in \tilde{\Lambda}^{\text{UB}}(\lambda)$ , and thus the system is stable.  $\blacksquare$

The corollary states that at least one point on the outer-bound<sup>1</sup> of the capacity region can be stabilized by the UB scheduling.

Based on Theorem 3, we have the following theorem to quantify the stability region of the UB scheduling.

**Theorem 5.** *The stability region of the UB scheduling policy is  $\Lambda^{\text{UB}}$ , and for any  $\lambda \in \Lambda^{\text{UB}}$ , Theorem 3 holds; for any  $\lambda \notin \Lambda^{\text{UB}}$ , Theorem 3 does not hold.*

*Proof:* The theorem can be directly obtained based on the definition of the stability region and Theorem 3.  $\blacksquare$

Similar to the CRB scheduling in the static channel case, the stability region of the UB scheduling also equals to the capacity region in the static channel case.

#### B. Stochastic Channel Case

**Theorem 6.** *Theorem 3 and Theorem 5 hold for the stochastic channel case.*

*Proof:* Similar to the static channel case,  $\mathbf{R}^{\text{UB}}(t)$  converges to the average throughput of the system. Then we have

$$\mathbf{r}_{S_i}^{\text{UB}}(t) = \arg \max_{\mathbf{r} \in \mathcal{C}_{S_i}(t)} \sum_i f(R_i^{\text{UB}}) r_i.$$

Taking expectation over time, we have

$$\mathbf{R}_{S_i}^{\text{UB}} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{S_i}} \sum_i f(R_i^{\text{UB}}) r_i,$$

and  $\mathbf{R}^{\text{UB}} = \mathbb{E}_i[\mathbf{R}_{S_i}^{\text{UB}}]$ . Then we can follow the same approach as in the static channel case. By replacing the static capacity region with the ergodic capacity region, the discussions in the static channel case also hold for the stochastic channel case, therefore proved that Theorem 3 holds for the stochastic channel case. Then based on the definition of the stability region, we can show that Theorem 5 holds for the stochastic channel case.  $\blacksquare$

While different from the static channel case, where the stability region can be obtained in closed-form, the stability region in the stochastic channel case is hard to be derived in closed-form. But we discover two properties as follows.

**Proposition 7.** *The stability region of the UB scheduling policy can be non-convex.*

**Proposition 8.** *With the UB scheduler, even though the system is stable when the arrival rate is  $\lambda$ , the system can be unstable when the arrival rate is reduced to  $x\lambda$ , where  $0 < x < 1$ .*

For these two properties, we only need to show that they hold for some scheduling policies with specific function  $f$ . This will be done in the Sec. VI.

*Remark.* These two properties make the UB scheduling very undesirable if the function  $f$  is selected improperly. The non-convexity means if one user decreases its arrival rate, the system may be unstable which is harmful for the quality of

<sup>1</sup>A point lies on the outer-bound of a set should satisfies two conditions. First is that the point lies on the boundary of a set; second is that the point is no longer belongs to the set if any increment in any dimension is made to the point.

service. The second property means that reducing the traffic intensity may bring a stable system to an unstable system which will also damage the QoS for all on-going traffic.

Although the closed-form expression of the stability region is difficult to obtain, a numerical method can be used to obtain the stability region. Here, we give the method to obtain the stability region of two-user systems, and it can be easily extended to a more general case.

*Numerical method to obtain the stability region of two-user systems:* Since the ergodic capacity region is a compact, convex, coordinate convex polyhedron, it can be represented as

$$\mathcal{C} = \{(r_1, r_2) : a_k r_1 + r_2 \leq b_k, k = 1, 2, \dots, K\},$$

where  $a_k$  is in the increasing order w.r.t.  $k$ , and if  $a_k = \infty$ , then the corresponding equation is  $r_1 = b_k$ .

Let  $\mathbf{r}^k = (r_1^k, r_2^k)$  be the solution of

$$\begin{cases} a_k r_1 + r_2 = b_k, \\ a_{k+1} r_1 + r_2 = b_{k+1}, \end{cases}$$

where  $0 < k < K$ ,  $\mathbf{r}^0$  be the solution of

$$\begin{cases} a_1 r_1 + r_2 = b_1, \\ r_1 = 0, \end{cases}$$

and  $\mathbf{r}^K$  be the solution of

$$\begin{cases} a_K r_1 + r_2 = b_K, \\ r_2 = 0. \end{cases}$$

Geometrically,  $\mathbf{r}^k$  is the vertex on the outer bound of the capacity region.

If  $f(\lambda_1)/f(\lambda_2) \in (a_k, a_{k+1})$ , i.e.,

$$\lambda \in Z_k = \{(\lambda_1, \lambda_2) : f(\lambda_1)/f(\lambda_2) \in (a_k, a_{k+1})\},$$

the stability region is the convex hull of  $\{\mathbf{0}, \mathbf{r}^0, \mathbf{r}^k, \mathbf{r}^K\}$ , which is represented as

$$\Lambda_k = \{(r_1, r_2) = \beta_1 \mathbf{r}^0 + \beta_2 \mathbf{r}^k + \beta_3 \mathbf{r}^K : \forall i, \beta_i > 0, \sum_i \beta_i \leq 1\}.$$

If  $f(\lambda_1)/f(\lambda_2) = a_k$ , then the stability region is

$$\Lambda^k = \{(r_1, r_2) : a_k r_1 + r_2 \leq b_k, f(r_1)/f(r_2) = a_k\}.$$

Overall, the stability region can be represented as

$$\Lambda^{\text{UB}} = \bigcup_k (\Lambda_k \cap Z^k) \bigcup \Lambda^k.$$

*Remark.* The key idea of the numerical method is to partition the capacity region into zones ( $Z_k$ ) and partition curves ( $\Lambda^k$ ). Each partition curve is the curve along the boundary of two neighboring zones. Since the capacity region is a convex polyhedron, the number of zones is finite<sup>2</sup>. For each zone, the allocated rate is identical, and thus the stability region for the arrival rate in each zone can be obtained. Examples are given in Sec. VI to show how to use the proposed method to obtain the stability region.

<sup>2</sup>Note that if the outer bound of the capacity region is strict convex, then the number of zones is infinite, and this method cannot work.

## V. EXTENDED STABILITY REGION

### A. Extended Stability Region of the CRB Scheduling

If we give a weight to each user, then a more general CRB scheduling policy is to allocate the rate based on the following optimization problem if user set  $\mathcal{A}(t)$  is selected:

$$\mathbf{r}_{\mathcal{A}(t)}^{\mathbf{w}, \text{CRB}}(t) = \arg \max_{\boldsymbol{\eta} \in \mathcal{C}_{\mathcal{A}(t)}(t)} \sum_{i \in \mathcal{A}(t)} w_i f(R_i^{\text{CRB}}(t)) \eta_i,$$

where  $w_i \in \mathbb{R}_+ \cup \{0\}$  is the normalized weight, satisfying  $\sum_i w_i = 1$ .

Since with the CRB scheduler,  $\mathbf{R}^{\text{CRB}}(t)$  converges to

$$\mathbf{R}^{\mathbf{w}, \text{CRB}} = \arg \max_{\mathbf{r} \in \mathcal{C}_{\mathcal{N}}} \sum_{i \in \mathcal{N}} w_i U(r_i). \quad (10)$$

Similar to (8), we have

$$\mathbf{R}_{\mathcal{S}_i}^{\mathbf{w}, \text{CRB}} = \arg \max_{\mathbf{r} \in \mathcal{C}_{\mathcal{S}_i}} \sum_{j \in \mathcal{S}_i} w_j f(R_j^{\mathbf{w}, \text{CRB}}) r_j,$$

and  $\mathbf{R}^{\mathbf{w}, \text{CRB}} = \mathbf{R}_{\mathcal{N}}^{\mathbf{w}, \text{CRB}}$ .

Similar to (7), for any given weight  $\mathbf{w}$ , the corresponding stability region is obtained as

$$\Lambda^{\mathbf{w}, \text{CRB}} = \bigcup_{\sum_i t_i = 1} \sum_{i=1}^{|\mathcal{S}|} \mathbf{R}_{\mathcal{S}_i}^{\mathbf{w}, \text{CRB}} t_i,$$

where  $t_i \in \mathbb{R}_+ \cup \{0\}$ , and we have the following theorem.

**Theorem 9.**

$$\bar{\mathcal{C}}_{\mathcal{N}} = \bigcup_{\sum_i w_i = 1} \Lambda^{\mathbf{w}, \text{CRB}},$$

where  $\Lambda^{\mathbf{w}, \text{CRB}}$  is the stability region of the CRB scheduling with weight  $\mathbf{w}$  assigning to users.

*Proof:* In order to prove that the union of the weighted stability region is the ergodic capacity region, essentially we need to show that, all the boundary points of the capacity region are the solutions to (10) by varying the weight  $\mathbf{w}$ .

By mapping  $\bar{\mathcal{C}}_{\mathcal{N}}$  to  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  through  $U(x)$ , i.e.,

$$\bar{\mathcal{C}}_{\mathcal{N}}^U = \{(U(x_1), U(x_2), \dots, U(x_N)) | \mathbf{x} \in \bar{\mathcal{C}}_{\mathcal{N}}\},$$

(10) can be represented as

$$\mathbf{y}^{\mathbf{w}, \text{CRB}} = \arg \max_{\mathbf{y} \in \bar{\mathcal{C}}_{\mathcal{N}}^U} \sum_{i \in \mathcal{N}} w_i y_i. \quad (11)$$

If  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  is a closed convex set, then according to the supporting hyperplane theorem [15], for any point  $\mathbf{y}$  lies on the boundary of  $\bar{\mathcal{C}}_{\mathcal{N}}^U$ , we can find the corresponding  $\mathbf{w}$  such that  $\mathbf{y} = \mathbf{y}^{\mathbf{w}, \text{CRB}}$ . Thus all the boundary points of  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  are the solutions to (11) by varying  $\mathbf{w}$ . Since  $U(x)$  is a monotonic non-decreasing function, the images of the boundary points of  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  in  $\bar{\mathcal{C}}_{\mathcal{N}}$  are still the boundary points. Then, the theorem can be proved if  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  is a closed convex set.

For any  $\mathbf{x}^{\mathbf{m}}, \mathbf{x}^{\mathbf{n}} \in \bar{\mathcal{C}}_{\mathcal{N}}$  whose image in  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  is  $\mathbf{m}, \mathbf{n}$ , respectively, we have

$$\begin{aligned} & \alpha \mathbf{m} + (1 - \alpha) \mathbf{n} \\ &= \alpha (U(x_1^{\mathbf{m}}), \dots, U(x_N^{\mathbf{m}})) + (1 - \alpha) (U(x_1^{\mathbf{n}}), \dots, U(x_N^{\mathbf{n}})) \\ &= (y_1, \dots, y_N), \end{aligned}$$

where

$$y_i = \alpha U(x_i^{\text{m}}) + (1 - \alpha)U(x_i^{\text{n}}).$$

Without loss of generality, we assume  $x_i^{\text{m}} \geq x_i^{\text{n}}$ . Because  $U(x)$  is monotonically non-decreasing,  $y_i \in [U(x_i^{\text{n}}) \ U(x_i^{\text{m}})]$ . i.e.,  $x_i = U^{-1}(y_i) \in [x_i^{\text{n}} \ x_i^{\text{m}}]$ . Consequently,  $\mathbf{x} \in \bar{\mathcal{C}}_{\mathcal{N}}$  due to the convexity of  $\bar{\mathcal{C}}_{\mathcal{N}}$ , and therefore  $\mathbf{y} \in \bar{\mathcal{C}}_{\mathcal{N}}^U$  and the convexity of  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  is proved. Since  $\bar{\mathcal{C}}_{\mathcal{N}}$  is closed,  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  is also closed. Therefore  $\bar{\mathcal{C}}_{\mathcal{N}}^U$  is indeed a closed convex set. ■

### B. Extended Stability Region of the UB Scheduling

Similar to the CRB scheduling, for the UB scheduling, a slight modification to the scheduling policy by giving a weight to each user, the resultant stability region is denoted by  $\Lambda^{\mathbf{w}, \text{UB}}$ , and we have the following theorem.

**Theorem 10.**

$$\bar{\mathcal{C}}_{\mathcal{N}} = \bigcup_{\sum_i w_i = 1} \Lambda^{\mathbf{w}, \text{UB}},$$

where  $\Lambda^{\mathbf{w}, \text{UB}}$  is the stability region of the UB scheduling with weight  $\mathbf{w}$  assigning to users.

*Proof:* First, based on the supporting hyperplane theorem, for any given  $\boldsymbol{\lambda}$  and any boundary point of  $\bar{\mathcal{C}}_{\mathcal{S}_i}$ , we can find a  $\mathbf{w}$  such that the boundary point  $\mathbf{R}_{\mathcal{S}_i}^{\mathbf{w}, \text{UB}}$  is the solution to the following problem,

$$\mathbf{R}_{\mathcal{S}_i}^{\mathbf{w}, \text{UB}} = \arg \max_{\mathbf{r} \in \bar{\mathcal{C}}_{\mathcal{S}_i}} \sum_{j \in \mathcal{S}_i} w_j f(\lambda_j) r_j,$$

where  $\sum_j w_j = 1$ . Also note that the boundary points of  $\bar{\mathcal{C}}_{\mathcal{S}_i}$  also lie on the boundary of  $\bar{\mathcal{C}}_{\mathcal{N}}$ ; therefore, we have

$$\begin{aligned} \bar{\mathcal{C}}_{\mathcal{N}} &= \bigcup_{\sum_i w_i = 1} \text{Co}\{\mathbf{R}_{\mathcal{S}_i}^{\mathbf{w}, \text{UB}} : i \in \mathcal{N}\}, \\ &= \bigcup_{\sum_i w_i = 1} \tilde{\Lambda}^{\mathbf{w}, \text{UB}}(\boldsymbol{\lambda}) \end{aligned}$$

where  $\text{Co}$  means convex hull, and

$$\tilde{\Lambda}^{\mathbf{w}, \text{UB}}(\boldsymbol{\lambda}) = \bigcup_{\sum_i t_i = 1} \sum_{i=1}^{|\mathcal{S}|} \mathbf{R}_{\mathcal{S}_i}^{\text{UB}} t_i.$$

According to Theorem 5, we know that

$$\boldsymbol{\lambda} \in \Lambda^{\mathbf{w}, \text{UB}} \Leftrightarrow \boldsymbol{\lambda} \in \tilde{\Lambda}^{\mathbf{w}, \text{UB}}(\boldsymbol{\lambda}).$$

So we have

$$\boldsymbol{\lambda} \in \bigcup_{\sum_i w_i = 1} \Lambda^{\mathbf{w}, \text{UB}} \Leftrightarrow \boldsymbol{\lambda} \in \bigcup_{\sum_i w_i = 1} \tilde{\Lambda}^{\mathbf{w}, \text{UB}}(\boldsymbol{\lambda}) = \bar{\mathcal{C}}_{\mathcal{N}},$$

which suggests

$$\bigcup_{\sum_i w_i = 1} \Lambda^{\mathbf{w}, \text{UB}} = \bar{\mathcal{C}}_{\mathcal{N}}. \quad \blacksquare$$

### C. Discussion

Although the stability regions of the CRB and the UB scheduling policies are less than the capacity region, respectively, by assigning the weights to users, the resultant scheduling algorithms can stabilize the system. Further note that, by giving the weights to users, the equivalent utility function has changed from a homogeneous one ( $U(\cdot)$ ) to a heterogeneous one ( $w_i U(\cdot)$ ). Therefore, for any given  $\mathbf{w}$ , the discussion in Sec. III and Sec. IV can still be used to analyze the stability of the system.

The advantage of the weighted opportunistic scheduling is that when the arrival rate lies outside the capacity region, the operation point (the throughput) is determined by the utility function  $U$  (in both UB and CRB scheduling policies), which is typically designed based on the fairness concern. Therefore the weighted opportunistic scheduling can provide a better fairness.

Although the approach is promising, it may not be easy. The weight-design is to find the supporting hyperplane (weight) of a closed convex set (capacity region) in a specific boundary point (the intersection of the arrival rate vector and the capacity region). Since the solution highly depends on the shape of the closed convex set, we lack a general analytic method. Further works should be done to obtain a simple method to design the weight.

## VI. EXAMPLES AND SAMPLE VALIDATION

In this section, we give examples about the stability region of the UB scheduling and the CRB scheduling policies. Simulation is conducted to compare the two policies, and validate the analytic results.

### A. Channel Assumption

Considering a two-user four-state channel, the transmission rate vector is

$$\mathbf{u}^m = \begin{cases} [R_1^{\text{ON}}, 0]^T, & m = 1, \\ [0, R_2^{\text{ON}}]^T, & m = 2, \\ [R_1^{\text{ON}}, R_2^{\text{ON}}]^T, & m = 3, \\ [0, 0]^T, & m = 4, \end{cases}$$

and the stationary distribution is  $\boldsymbol{\pi} = (1/4, 1/4, 1/4, 1/4)$ . Note that this is a channel model for a two-user system, where each user has two states (ON and OFF), and the channel states for different users are independent. The achievable throughput of user  $i$  is  $R_i^{\text{ON}}$  when its channel state is ON, and 0 if its channel state is OFF. Without loss of generality, we assume  $R_1^{\text{ON}} \geq R_2^{\text{ON}}$ .

Based on (1), we can obtain the ergodic capacity region as:

$$\begin{aligned} \bar{\mathcal{C}} &= \{(R_1, R_2) : R_1/R_1^{\text{ON}} + R_2/R_2^{\text{ON}} \leq 3/4, \\ & \quad R_1/R_1^{\text{ON}} \leq 1/2, R_2/R_2^{\text{ON}} \leq 1/2\}. \quad \blacksquare \end{aligned}$$

## B. Utility Function

1)  $\alpha$ -Fairness Utility: The utility function chosen to be evaluated is the  $\alpha$ -fairness ones [16]:

$$U(x) = \begin{cases} \log(x), & \alpha = 1, \\ (1 - \alpha)^{-1} x^{1-\alpha}, & \text{otherwise,} \end{cases}$$

where  $x$  is the average throughput, whose unit is bps/Hz in this paper and is omitted in the following. The derivative of  $U(x)$  is

$$f(x) = x^{-\alpha}.$$

By choosing a different  $\alpha$ , the objective is to maximize the fairness measurement based on different principles, and the relative value of the measurement is of more interests. For instance, if  $\alpha = 0$ , then the objective is to maximize the system throughput; if  $\alpha = 1$ , then it is to maximize the proportional fairness; if  $\alpha \rightarrow \infty$ , then it is to maximize the max-min fairness.

2) *Exponential Utility*: Another utility function chosen to be evaluated is exponential utility [17]:

$$U(x) = -\frac{1}{a} e^{-ax},$$

and

$$f(x) = e^{-ax}.$$

For the exponential utility, the marginal utility is exponentially decreasing, and the changing rate of the marginal utility is a constant and independent of  $x$ .

## C. Stability Region of the UB Scheduling

1)  $\alpha$ -Fairness Utility: Based on the numerical method proposed in Sec. IV, we can obtain the stability region, as shown in Fig. 1.

The point P in the figure is the intersection of the boundary of the capacity region and curve  $\frac{f(\lambda_1)}{f(\lambda_2)} = \beta$  where  $\beta = R_2^{\text{ON}}/R_1^{\text{ON}}$ . With the increasing of  $\alpha$ , P moves along the boundary of the capacity region, and results in the shape changing of the stability region. From the figure we also can observe that the stability region is non-convex all the time. When the value of  $\alpha$  is proper, the stability region is the union of a convex set and a line segment. When  $\alpha$  is large or small, P moves to the line  $R_1 = R_1^{\text{ON}}/2$  or  $R_2 = R_2^{\text{ON}}/2$ , then the stability region is a trapezoid minus a triangular. The non-convex property of the stability region makes the system behavior hard to predict and the QoS hard to meet, since decreasing the arrival rate of one flow may lead the system changing from stable to unstable.

2) *Exponential Utility*: We already know that, with the UB scheduler, for a stable system with arrival rate  $\lambda$ , decreasing any element in  $\lambda$  may lead the system to be unstable. Here we give another example to show that, proportionally decreasing all the elements in  $\lambda$  (down-scale  $\lambda$ ) may also lead the system to be unstable.

Based on the same approach as in  $\alpha$ -fairness utility, the stability region can be obtained. Since we change the function  $f$ , so the curve  $\frac{f(\lambda_1)}{f(\lambda_2)} = \beta$  determining P is  $\lambda_2 - \lambda_1 = 1/a \log \beta$ .

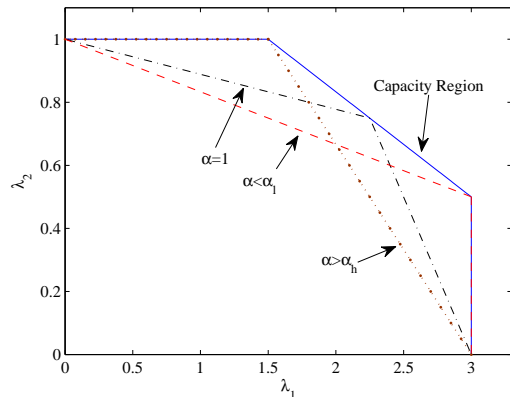


Fig. 3. Stability Region of a system with four-state channel and  $\alpha$ -fairness CRB scheduling.  $R_1^{\text{ON}} = 6$ ,  $R_2^{\text{ON}} = 2$ ,  $\alpha_l = 0.61$ ,  $\alpha_h = 2.71$

If the arrival rate  $\lambda$  is downscaled by  $x$ , the new arrival rate is no longer lies in the partition curve  $\frac{f(\lambda_1)}{f(\lambda_2)} = \beta$  therefore cannot guarantee the stability, and the stability property should be examined by finding which zone the new  $\lambda$  lies in. As illustrated in Fig. 2, when  $a = 1$  or 3, if the system is stable in point P, then ‘down-scale’  $\lambda$  by  $x$ , the system becomes unstable, *i.e.*, suffering the ‘down-scale’ unstable; but when  $a = 0.4$ , the ‘down-scale’ unstable situation does not happen.

## D. Stability Region of the CRB Scheduling

$\alpha$ -Fairness Utility: We enumerate  $\mathbb{S}$  as  $\mathbb{S} = \{(1), (2), (1, 2), \emptyset\}$ . For each  $\mathcal{A} \in \mathbb{S}$ , we have  $\mathbf{R}_{(1)} = [R_1^{\text{ON}}/2, 0]^T$ ,  $\mathbf{R}_{(2)} = [0, R_2^{\text{ON}}/2]^T$ ,  $\mathbf{R}_{\emptyset} = [0, 0]^T$ , and

$$\mathbf{R}_{(1,2)} = \begin{cases} [R_1^{\text{ON}}/2, R_2^{\text{ON}}/4]^T, & \alpha < \alpha_l, \\ \left[ \frac{3R_1^{\text{ON}}/4}{1+\beta^{1/\alpha-1}}, \frac{3R_2^{\text{ON}}/4}{1+\beta^{1-1/\alpha}} \right]^T, & \alpha_l \leq \alpha \leq \alpha_h, \\ [R_1^{\text{ON}}/4, R_2^{\text{ON}}/2]^T, & \alpha > \alpha_h, \end{cases}$$

where  $\alpha_l = \log \beta / \log \frac{\beta}{2}$ ,  $\alpha_h = \log \beta / \log 2\beta$ .

The capacity region and the stability region are illustrated in Fig. 3. By varying  $\alpha$ ,  $\mathbf{R}_{(1,2)}$  is moving on the outer bound of the capacity region, and the stability region is always a convex hull.

Comparing Figs. 1 and 3, under the four-state channel assumption in a two-user system, the CRB scheduling policy can always provide a larger stability region than the UB scheduling policy if using the same utility function.

## E. Scheduling Policy Comparison

We have conducted simulation to compare the UB and the CRB scheduling policies. We choose  $\alpha$ -fairness as the utility function, and  $\alpha = 0.5$ . We use Poisson traffic as the arrival traffic,  $\epsilon$  is chosen as 0.01, and we run 10 times to take the average. We set  $\lambda_1 = R_1^{\text{ON}}/2$  for all 20000 time slots, set  $\lambda_2 = R_2^{\text{ON}}/4$  for the first 10000 time slots and  $\lambda_2 = R_2^{\text{ON}}/10$  for the last 10000 time slots. This is used to simulate the arrival-rate decreasing of one flow.

The throughput comparison is shown in Fig. 4. From the figure, after 10000 time slots, the throughput of Q1 with the UB scheduler (the curve UB Q1) starts to decrease and is



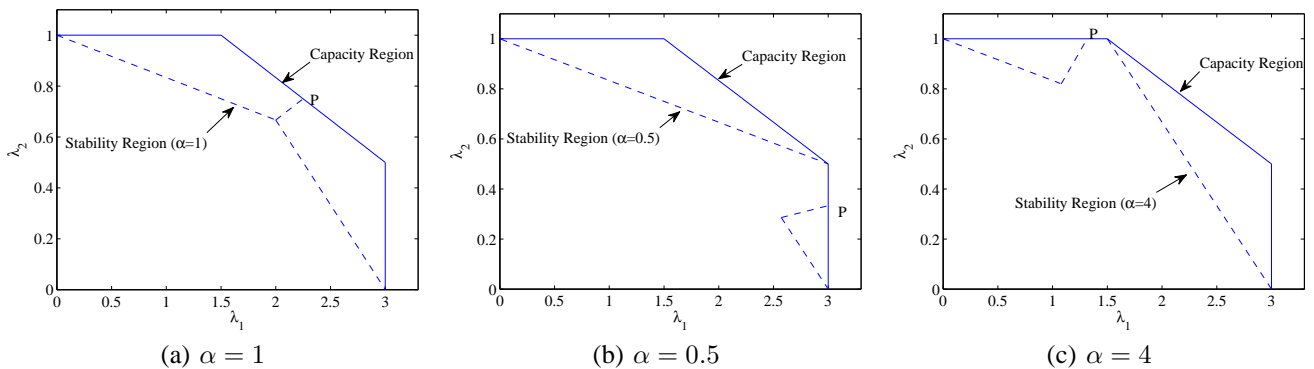


Fig. 1. Stability Region of a system with four-state channel and  $\alpha$ -fairness UB scheduling.  $R_1^{\text{ON}} = 6$ ,  $R_2^{\text{ON}} = 2$ .

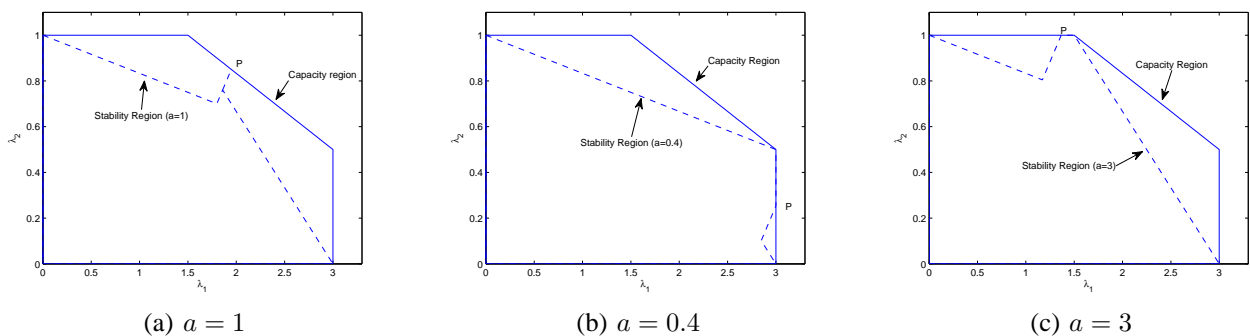


Fig. 2. Stability Region of a system with four-state channel and exponential UB scheduling.  $R_1^{\text{ON}} = 6$ ,  $R_2^{\text{ON}} = 2$ .

less than the throughput of Q1 with the CRB scheduler. For the throughput of Q2, both schedulers can maintain the same throughput, which equals the arrival rate of the second flow. Here we can conclude, by decreasing the arrival rate of one flow, the throughput of another flow can be decreased, if the UB scheduler is used. This phenomenon can be explained by examining the system stability based on Theorem 6 with the new arrival rate. An intuitive explanation is as follows: as the utility function  $U(x)$  is strictly concave, the derivative function  $f(x)$  is a decreasing function. Since  $R_i^{\text{UB}}(t)$  is used to estimate the average throughput of user  $i$  and if the arrival rate of user  $i$  decreases, the estimated average throughput should also decrease, *i.e.*,  $R_i^{\text{UB}}(t)$  decreases. Therefore,  $f(R_i^{\text{UB}}(t))$  will increase. From (5) we can see, this generally results in the increase of  $r_i^{\text{UB}}(t)$ , *i.e.*, the increase of the instantaneous rate of user  $i$ . The probability that the system stays in a state without user  $i$  will increase, as a joint results of the decreasing of the arrival rate and the increasing of the instantaneous rate. As the number of users has decreased, the system will lose certain multi-user diversity, *i.e.*, the achievable rate region thanks to the opportunistic scheduling will shrink. This may lead to a situation that the average throughput of a user except  $i$  is less than its arrival rate, and therefore leads to an unstable flow.

If we give weights to users<sup>3</sup>, we can see the weighted UB scheduler can maintain the throughputs for both users. But further note that, although the specific weighted UB scheduling can stabilize the system with the arrival-rate decreasing of one

<sup>3</sup>Here we assign weight 0.75 to user 1 and 0.25 to user 2, and the corresponding curves are  $\text{UB}^{\text{w}} \text{Q1}$  and  $\text{UB}^{\text{w}} \text{Q2}$ . The weight is specifically designed in order to stabilize the system, and such design is also non-unique.

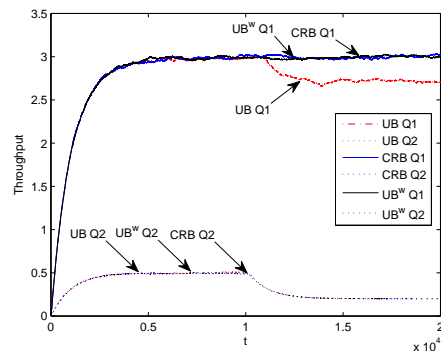


Fig. 4. Throughput Comparison.  $R_1^{\text{ON}} = 6$ ,  $R_2^{\text{ON}} = 2$ ,  $\alpha = 0.5$ .

flow in the given scenario, there will exist some scenarios that the system still cannot be stabilized if one flow decreases its arrival rate, as the stability region of the weighted UB scheduling is still less than the capacity region.

The queue length is compared in Fig. 5. The y-axis is in the logarithm form. After 10000 time slots, while the arrival rate of Q2 is reduced, with the UB scheduler, the queue length of Q1 starts to increase. From the increasing trend we could tell that, the system cannot be stabilized. But with a proper weight assigning to each user, the system can be stabilized by the weighted UB scheduling policy. These results validate our analytical conclusion.

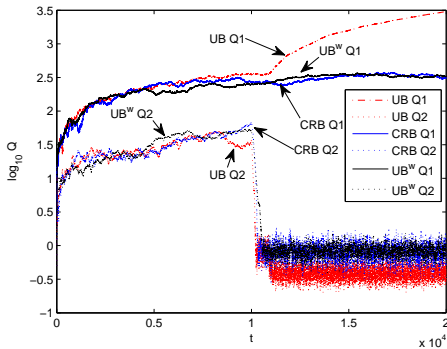


Fig. 5. Queue Length Comparison.  $R_1^{ON} = 6$ ,  $R_2^{ON} = 2$ ,  $\alpha = 0.5$ .

## VII. DISCUSSION AND CONCLUSION

In this paper, the stability regions of two opportunistic scheduling policies have been discussed. One is the UB scheduling policy and the other is a variant of the UB scheduling policy, called the CRB scheduling policy. We have proposed a numerical method to obtain the stability region of the UB scheduling policy, and the results show that the stability region of the UB scheduling policy is generally non-convex and may exhibit some undesirable properties, such as decreasing the arrival rate of one flow may lead the system to be unstable, and proportionally decreasing the arrival rates of flows may lead the system to be unstable. Such properties suggest that in a system using the UB scheduling policy, reducing the traffic intensity may have a negative impact on the QoS for all on-going traffic, which is contradict to the intuition. Different from the UB scheduling policy, the stability region of the CRB scheduling policy is derived in closed-form, and is a convex hull. In addition to the stability region, we have further discussed the extended stability region. The results show that by assigning a proper weight to each user, the weighted scheduling policy can stabilize the system if the arrival process is stationary and the average arrival rate lies inside the capacity region. Simulation and numerical examples have been given to explain the analytical results and validate our analysis.

Although the CRB scheduling policy is better than the UB scheduling policy in terms of the stability region, it needs explicit knowledge of the number of users in the system, which may bring some difficulties to implement, since how frequently to update this information may be hard to design.

There are several open research issues beckon for further research. First, what is the impact of  $\epsilon$  when updating the smoothed rate measurement. Since if  $\epsilon$  is not proper, the smoothed rate measurement may not be able to converge, especially if the flow exhibits a bursty feature. Thus the rate allocation is not stationary, and the resultant impact on the stability region is unclear. Second, given the average arrival rate, how to design the weight to each user to stabilize the system and how the designed weight affects the system performance needs to be investigated further. Under some special assumptions, the system can be modeled as a Markov Chain [18], [19], therefore the performance can be numerically

studied. While a system under general assumptions may not be easy to be modeled as a simple Markov Chain, the performance study is still challenging.

## REFERENCES

- [1] R. Knopp, P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE ICC'95*, June 1995.
- [2] D. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. ISIT'97*, Jun/Jul 1997.
- [3] Xin Wang, G.B. Giannakis, A.G. Marques, "A Unified Approach to QoS-Guaranteed Scheduling for Channel-Adaptive Wireless Networks," *Proceedings of the IEEE*, 95(12):2410-2431, 2007.
- [4] A. L. Stolyar, "On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation," *Operations Research*, 53(1):12-25, 2005.
- [5] H.J. Kushner, P.A. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Transactions on Wireless Communications*, 3(4):1250-1259, 2004.
- [6] J. Kim, J. Lee, K. Son, S. Song, S. Chong, "Two-Hop Opportunistic Scheduling in Cooperative Cellular Networks," *IEEE Transactions on Vehicular Technology*, (61)9:4194-4199, 2012.
- [7] X. Wang and G.B. Giannakis, "Resource Allocation for Wireless Multiuser OFDM Networks," *IEEE Transactions on Information Theory*, 57(7):4359-4372, 2011.
- [8] J. Huang, V.G. Subramanian, R. Agrawal, R.A. Berry, "Downlink scheduling and resource allocation for OFDM systems," *IEEE Transactions on Wireless Communications*, (8)1:288-296, 2009.
- [9] J. Huang, V.G. Subramanian, R. Agrawal, R. Berry, "Joint scheduling and resource allocation in uplink OFDM systems for broadband wireless access networks," *IEEE Journal on Selected Areas in Communications*, (27)2:226-234, 2009.
- [10] M. Andrews, "Instability of the proportional fair scheduling algorithm for HDR," *IEEE Transactions on Wireless Communications*, 3(5):1422-1426, 2004.
- [11] L. Tassiulas, A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inform. Theory*, 39(2):466-478, Mar. 1993.
- [12] Haiyou Guo, Honglin Hu, Yan Zhang, and Hsiao-Hwa Chen, "On stability regions in opportunistic scheduled-packet access networks," *IEEE Transactions on Vehicular Technology*, 59(1):295306, 2010.
- [13] Beiyu Rong and Anthony Ephremides, "Joint mac and rate control for stability and delay in wireless multi-access channels," *Performance Evaluation*, 68(8):658669, 2011.
- [14] E. Leonardi, M. Mellia, F. Neri, M.A. Marsan "Bounds on delays and queue lengths in input-queued cell switches," *J. ACM*, 50(4):520-550, 2003
- [15] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [16] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Networking*, 8(5):556-567, 2000.
- [17] P. C. Fishburn, *Utility Theory for Decision Making*. R. E. Krieger Pub. Co., 1970.
- [18] M. Dianati, R. Tafazolli, X. Shen, S. Naik, "Per-User Service Model for Opportunistic Scheduling Scheme over Fading Channels," *Wireless Communications and Mobile Computing (Wiley)*, 10(1):87-100, 2010.
- [19] L. Lei, C. Lin, J. Cai, X. Shen, "Performance Analysis of Wireless Opportunistic Schedulers using Stochastic Petri Nets," *IEEE Transactions on Wireless Communications*, 8(4):2076-2087, 2009.