On Scheduling in Multi-Channel Wireless Downlink Networks with Limited Feedback

Ming Ouyang and Lei Ying
Department of Electrical and Computer Engineering
Iowa State University
Email: {mouyang, leiying}@iastate.edu

***Abstract*—We consider a wireless downlink network with a single base-station, N mobile users and L shared on-off channels. Each mobile user receives a downlink traffic flow from the base-station where a separate queue is maintained for each flow. In this multi-channel downlink network, throughput-optimal scheduling algorithms such as the MaxWeight scheduling require the complete channel state information (i.e., NL channel states) for scheduling. This could be a significant overhead when the number of mobile users is large. This paper considers wireless downlink networks with limited feedback bandwidth so that at most F of the NL channel states can be reported at each time slot. We propose dynamic feedback allocation schemes, named as Longest-Queue-First Feedback-Allocation (LQF-FA) and Modified-Longest-Queue-First Feedback-Allocation (MLQF-FA), which dynamically and adaptively allocate the feedback resource according to queue-lengths and channel statistics. We prove that given a fixed feedback resource F, the LQF-FA+MaxWeight is throughput-optimal under a mean approximation; and the throughput difference between the MLQF-FA+MaxWeight and the MaxWeight with the complete channel state information decreases exponentially as a function of F/L when $F = O(L^2)$.**

I. Introduction

Scheduling is one of the key challenges in wireless network design. In a seminal work [1], the authors have shown that with the complete queue and channel state information, the MaxWeight scheduling can stabilize the network given any traffic load that is within the network throughput region. For a cellular downlink network consisting of a base station and multiple mobile users, the queue lengths are readily available at the base-station, but the channel states, in many cases, need to be measured by mobile users, and reported to the base-station. With the increase of network capacity (i.e., the increase of supportable mobile users), reporting the complete channel state information could consume a significant amount of communication resource. Because of that, in current and next generation cellular standards, such as 802.16e [2], 802.11m [3], and LTE [4], the feedback bandwidth is limited and each mobile can only report a few channel states. A mobile user measures all the channels and selects a subset of them to report. In most existing systems, the feedback bandwidth is evenly distributed to mobile users. While this feedback allocation provides a simple solution under feedback resource constraint, the efficiency of this scheme is questionable. As we have learnt from the MaxWeight scheduling [1], [5], [6], [7], [8], a high-performance scheduling algorithm should dynamically allocate resource based on channel conditions and traffic demands, which motivates us to consider dynamical feedback resource allocation schemes under limited feedback bandwidth.

Scheduling with limited network state information has been studied in [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19]. This paper studies a different model, which is motivated by a downlink network using multi-carrier techniques such as Orthogonal Frequency Division Multiplexing (OFDM). We consider a wireless downlink network with a base station, N mobile users, and L shared fading channels (frequency bands). During each time slot, a channel can be used by at most one mobile user. The channel is assumed to be on-off channels, which have identical and independent distributions across channels, time and users. We assume the transmission on each channel is with the same modulation and coding scheme, which implies that the transmission rate is the same when a channel is on. We also assume the probability that a channel is on is known by the base station. A mobile user can measure the states of all channels, but can only report a subset of them to the base-station. The number of channels a mobile user can report (i.e., the feedback bandwidth of the mobile user) is decided by the base-station. For this model, we address the following two basic questions:

(i) *how should we allocate the limited feedback bandwidth to maximize the network throughput;* and
(ii) *how much feedback bandwidth is necessary for a near-optimal throughput?*

The main contributions of this paper include:

- We characterize the network throughput region under limited feedback bandwidth.
- Under a mean approximation, i.e., assuming that the number of on channels is pL for each of the mobile users where p is the probability a channel being on, we develop the Longest-Queue-First Feedback-Allocation (LQF-FA) scheme. Combined with the MaxWeight scheduling, LQF-FA+MW is throughput optimal under the mean approximation.
- We then propose a Modified-Longest-Queue-First Feedback-Allocation (MLQF-FA) scheme. We prove that the throughput difference between the MLQF-FA+MaxWeight and the MaxWeight with the complete channel state information decreases exponentially as a

function of F/L when $F = O(L^2)$.

II. System Model

We consider a cellular downlink network with a single base station and N mobile users. Each user associates with a downlink data flow, and the base station maintains a separate queue for each flow. The N flows are transmitted over L shared on-off channels. Denote by $X_{ij}(t)$ the state of channel j to user i at time slot t. We assume that $X_{ij}(t)$ are independently and identically (i.i.d.) distributed over users, channels, and time, and

$$X_{ij}(t) = \begin{cases} R, & \text{with probability } p; \\ 0, & \text{with probability } 1-p. \end{cases}$$

At time t, we define $Y_{ij}(t)$ to be the state of channel j available at the base station, i.e.,

$$Y_{ij}(t) = \begin{cases} X_{ij}(t), & \text{if channel } j \text{ is reported by user } i \text{ ;} \\ 0, & \text{otherwise.} \end{cases}$$

We further define $Z_{ij}(t)$ to be the scheduling decision of the base station such that

$$Z_{ij}(t) = \begin{cases} 1, & \text{if user } i \text{ is served over channel } j \text{ at time } t; \\ 0, & \text{otherwise.} \end{cases}$$

We assume that a channel can be used to served at most one user at a time so that

$$\sum_{j=1}^{L} Z_{ij}(t) \leq 1$$

for all i and t.

We assume that the transmission to user i over channel j cannot succeed if the channel state is not reported. Denoting by $D_i(t)$ the service rate allocated to user i, we have that

$$D_i(t) = \sum_{j=1}^{L} Y_{ij}(t) Z_{ij}(t).$$

Next let $A_i(t)$ denote the number of packets injected to the queue maintained for user i at time t. Assume that $A_i(t)$ are bounded random variables, which are independent across users and time, and are independent of $X_{ij}(t)$. We further assume that the packets are injected at the beginning of each time slot, and are served at the end of each time slot. To this end, the evolution of queue i can be written as:

$$Q_i(t+1) = (Q_i(t) + A_i(t) - D_i(t))^+ \qquad (1)$$

where $Q_i(t)$ is the queue length for user i at the beginning of time slot t, and $(x)^+ = \max\{x, 0\}$. Furthermore, define $U_i(t)$ to be the unused service rate due to the lack of packets, i.e.,

$$U_i(t) = \begin{cases} 0, & \text{if } D_i(t) \leq Q_i(t) + A_i(t); \\ D_i(t) - Q_i(t) - A_i(t), & \text{otherwise.} \end{cases}$$

The queue evolution can be written as:

$$Q_i(t+1) = Q_i(t) + A_i(t) - D_i(t) + U_i(t). \qquad (2)$$

Assuming that the base-station has the complete channel state information, we can treat each channel as a single-server and apply the MaxWeight scheduling to select a user for each channel.

MaxWeight Scheduling: Channel j is allocated to to user $i^*(t)$ such that

$$i^*(t) \in \arg\max_i Q_i(t) X_{ij}(t).$$

□

Given the complete complete channel state, the MaxWeight scheduling solves the following optimization problem:

$$\begin{aligned} &\max_{\{Z_{ij}\}} \textstyle\sum_{i,j} Q_i(t) X_{ij}(t) Z_{ij} \\ \text{Subject to:} \quad &\textstyle\sum_j Z_{ij} \leq 1, Z_{ij} \in \{0,1\}. \end{aligned}$$

While the Max-Weight scheduling is throughput-optimal [1], the bases-station needs to collect the complete channel state information ($X_{ij}(t)$). For the system considered in this paper, the base-station needs to obtain the $N \times L$ channel states for scheduling, which means the feedback bandwidth required is $\Theta(NL)$. With the increase of mobile users, obtaining the complete channel state information will require a significant amount of uplink bandwidth. We assume that the wireless downlink network only allocate limited bandwidth for feedback so that at most F channel states can be reported to the base station at a time. In this system, the base-station needs to allocate the feedback bandwidth intelligently to maximize the network throughput. In this paper, we assume that the channel feedback is collected with two steps:

Two-step channel state feedback:

- At the beginning of each time slot, the base station decides a feedback allocation vector $\vec{m}(t)$, which is of length N, $m_i(t) \geq 0$ for all i and t, and $\sum_{i=1}^{F} m_i(t) \leq L$. This feedback allocation vector is then broadcast to all mobile users.
- At each time slot, mobile user i can measure all L channel states $\{X_{ij}\}_{j=1,\ldots,L}$, and can report at most $m_i(t)$ of them to the base station. Denote by $O_i(t)$ the number of on channels related to user i at time t. If $O_i(t) \leq m_i(t)$, user i reports all on channels to the base-station; otherwise, user i randomly and uniformly selects $m_i(t)$ of the $O_i(t)$ users to report.

□

Note that in most existing cellular systems, the feedback bandwidth is evenly distributed to all users, i.e., $m_i(t) = \frac{F}{N}$ for all i and t. The efficiency of this feedback resource allocation, however, is questionable. We note that the network throughput is constrained by the feedback resource because the traffic flows can only be sent over those reported channels. Therefore we are interested in feedback allocation scheme that can maximize the network throughput.

To evaluate the efficiency of a feedback allocation $\vec{m}(t)$, we study the expected value of the following optimization problem for a given $\vec{m}(t)$:

$$\begin{aligned} &\max_{\{Z_{ij}\}} \sum_{i,j} Q_i(t) Y_{ij}(t) Z_{ij} \\ \text{Subject to:} \quad &\textstyle\sum_j Z_{ij} \leq 1, Z_{ij} \in \{0,1\}, \end{aligned}$$

where the distribution of $Y_{ij}(t)$ is determined by the feedback allocation vector $\vec{m}(t)$ and the channel statistics under the two-step channel state feedback mechanism. The reason we choose $\mathbf{E}\left[\max_{\{Z_{ij}\}}\sum_{i,j}Q_i(t)Y_{ij}(t)Z_{ij}\right]$ as the performance metrics is that the ratio of $\mathbf{E}\left[\max_{\{Z_{ij}\}}\sum_{i,j}Q_i(t)Y_{ij}(t)Z_{ij}\right]$ and $\mathbf{E}\left[\max_{\{Z_{ij}\}}\sum_{i,j}Q_i(t)X_{ij}(t)Z_{ij}\right]$ is closely related to the throughput loss due to limited feedback bandwidth.

More specifically, we define

$$g(full,\vec{Q}(t))=\mathbf{E}\left[\max_{\sum_j Z_{ij}\leq 1}\sum_{i,j}Q_i(t)X_{ij}(t)Z_{ij}(t)\right].$$

Given $\vec{m}(t)$, we define

$$g(\vec{m}(t),\vec{Q}(t))=\mathbf{E}\left[\max_{\sum_j Z_{ij}\leq 1}\sum_{i,j}Q_i(t)Y_{ij}(t)Z_{ij}(t)|\vec{m}(t)\right],$$

where the distribution of $Y_{ij}(t)$ is determined by $\vec{m}(t)$.

It is easy to see that given $Y_{ij}(t)$, the MaxWeight scheduling, which schedules user

$$i^*(t)\in\arg\max_i Q_i(t)Y_{ij}(t)$$

for channel j, always maximizes $\sum_{i,j}Q_i(t)Y_{ij}(t)Z_{ij}(t)$. Therefore, denoting $Z_{ij}^{MX}(t)$ to be the scheduling decision under the MaxWeight scheduling, $g(full,\vec{Q}(t))$ and $g(\vec{m}(t)$ can be re-written as:

$$\begin{aligned} g(full,\vec{Q}(t)) &= \mathbf{E}\left[\sum_{i,j}Q_i(t)Y_{ij}(t)Z_{ij}^{MW}(t)\right] \\ g(\vec{m}(t),\vec{Q}(t)) &= \mathbf{E}\left[\sum_{i,j}Q_i(t)Y_{ij}(t)Z_{ij}^{MW}(t)|\vec{m}(t)\right]. \end{aligned}$$

Next we say a traffic load $\vec{A}(t)$ is supportable if there exists a scheduling algorithm and feedback allocation scheme that guarantee that the means of the queues are bounded. Based on the notations and definition above, we have the following theorems:

Theorem 1: Assume that the network can support F feedback at a time. Given a feedback allocation scheme, which guarantees that

$$g(\vec{m}(t),\vec{Q}(t))=\max_{\vec{m}:\sum_i m_i\leq F}g(\vec{m},\vec{Q}(t)) \tag{3}$$

holds for all t, then the feedback allocation scheme, combined with the MaxWeight scheduling, can support any traffic load that is supportable under feedback constraint F.

□

Theorem 2: Given a feedback allocation scheme, which guarantees that $(1+\delta)g(\vec{m}(t),\vec{Q}(t))\geq g(full,\vec{Q}(t))$ holds for all t, then the feedback allocation scheme, combined with the MaxWeight scheduling, can support any traffic load $\vec{A}(t)$ such that $(1+\delta)\vec{A}(t)$ supportable under the MaxWeight scheduling with the complete channel state information.

□

Remark 1: These two theorems can be proved using the standard Lyapunov drift analysis, and are ignored in this paper.

Remark 2: Theorem 1 provides a feedback-resource allocation scheme, which guarantees throughput optimality when combined with the MaxWeight scheduling. However, solving the optimization problem $\max_{\vec{m}:\sum_i m_i\leq F}g(\vec{m}(t),\vec{Q}(t))$ could be computationally expensive. We therefore derive simple feedback-resource allocation schemes based on a mean approximation.

III. Optimal feedback scheme under a mean approximation

In this section, we derive a feedback scheme that is optimal under the *mean approximation,* i.e., we assume that the number of on channels of each user is exactly pL. Under the mean approximation, it is easy to see that there is no need to ask a user to report more than pL channels, so we consider $\vec{m}(t)$ such that $m_i(t)\leq pL$ for all i. Next we first derive the closed form expression of $Q_i(t)\mathbf{E}\left[D_i(t)\middle|\vec{Q}(t),\vec{m}(t)\right]$ under the mean approximation.

Lemma 1: Given a feedback-resource allocation vector $\vec{m}(t)$ where $m_i(t)\leq pL$ for all i, we have

$$\begin{aligned} & Q_i(t)\mathbf{E}\left[D_i(t)|\,\vec{Q}(t),\vec{m}(t)\right] \\ = {} & Q_i(t)Rm_i(t)\prod_{k=1}^{i-1}\left(1-\frac{m_k(t)}{L}\right). \end{aligned} \tag{4}$$

Proof: Since the channels are i.i.d., we can obtain that

$$\begin{aligned} & Q_i(t)\mathbf{E}[D_i(t)|\vec{Q}(t),\vec{m}(t)] \\ = {} & Q_i(t)\mathbf{E}\left[\sum_{j=1}^{L}Y_{ij}(t)Z_{ij}^{MW}(t)\middle|\vec{Q}(t),\vec{m}(t)\right] \\ = {} & LQ_i(t)\mathbf{E}[Y_{ij}(t)Z_{ij}^{MW}(t)|\vec{Q}(t),\vec{m}(t)]. \end{aligned} \tag{5}$$

Without loss of generality, we now assume $Q_i(t)$ is sorted in a descent order such that $Q_i\leq Q_j$ if $i>j$. Therefore, channel j is allocated to user i if and only if the following two conditions hold:

(i) channel j is on and is reported by user i; and
(ii) the base station does not receive any report of channel j from users with an index smaller than i.

Recall that we only need to consider $\vec{m}(t)$ such that $m_i(t)\leq pL$, so we first obtain that

$$\begin{aligned} & \Pr(\text{channel 1 to user } i \text{ is on and reported}|\vec{m}(t)) \\ = {} & p\times\frac{m_i(t)}{pL}=\frac{m_i(t)}{L}, \end{aligned}$$

which implies that

$$\begin{aligned} & Q_i(t)\mathbf{E}\left[Y_{ij}(t)Z_{ij}^{MX}(t)\middle|\,\vec{Q}(t),\vec{m}(t)\right] \\ = {} & Q_i(t)R\frac{m_i(t)}{L}\prod_{k=1}^{i-1}\left(1-\frac{m_k(t)}{L}\right), \end{aligned} \tag{6}$$

holds for all channel j. Therefore,

$$\begin{aligned} & Q_i(t)\mathbf{E}\left[D_i(t)|\,\vec{Q}(t),\vec{m}(t)\right] \\ = \;& Q_i(t)Rm_i(t)\prod_{k=1}^{i-1}\left(1-\frac{m_k(t)}{L}\right), \end{aligned} \tag{7}$$

and the lemma holds. ∎

Now based on the expression (4), we will derive two important lemmas that will lead to the optimal feedback allocation policy. Assume that the queues are sorted in a descent order. Lemma 2 shows that if user i receives less feedback resource than user $i+1$, then we can swap their feedback resource to improve $g(\cdot,\vec{m}(t))$. Lemma 3 shows that we can move the feedback resource from user i, where $i=\max\{k:m_k(t)>0\}$ to user $i-1$ to improve $g(\cdot,\vec{m}(t))$ if $m_i(t)\leq m_{i-1}(t)$. The optimal policy is then obtained by repeatedly re-arrange the feedback allocation according to Lemma 2 and Lemma 3.

Lemma 2: Given a feedback allocation $\vec{m}(t)$ such that $m_i(t)<m_{i+1}(t)$, we can construct an alternative feedback allocation $\vec{m}'(t)$ such that $m'_i(t)=m_{i+1}(t)$, $m'_{i+1}(t)=m_i(t)$, and $m'_j(t)=m_j(t)$ for $j\neq i,i+1$, and

$$g(\vec{m}(t),\vec{Q}(t))\geq g(\vec{m}'(t),\vec{Q}(t)).$$

Proof: First, it is easy to verify that for any user k such that $k<i$,

$$\mathbf{E}\left[D_k(t)\left|\vec{Q}(t),\vec{m}(t)\right.\right]=\mathbf{E}\left[D_k(t)\left|\vec{Q}(t),\vec{m}'(t)\right.\right] \tag{8}$$

because $D_k(t)$ is independent of the channel states of user h such that $h>k$ (Lemma 1).

Further, according to Lemma 1, it is easy to see that exchanging the feedback resource of user i and user $i+1$ will not change the conditional expectation of $D_k(t)$ for $k>i+1$, i.e.,

$$\mathbf{E}\left[D_k(t)\left|\vec{Q}(t),\vec{m}(t)\right.\right]=\mathbf{E}\left[D_k(t)\left|\vec{Q}(t),\vec{m}'(t)\right.\right]$$

for any $k>i+1$.

From the observations above and Lemma 1, we can finally obtain that

$$\begin{aligned} & g(\vec{m}'(t),\vec{Q}(t))-g(\vec{m}(t),\vec{Q}(t)) \\ = \;& \mathbf{E}\left[Q_i(t)D_i(t)+Q_{i+1}(t)D_{i+1}(t)\left|\vec{Q}(t),\vec{m}(t)\right.\right] \\ & -\mathbf{E}\left[Q_i(t)D_i(t)+Q_{i+1}(t)D_{i+1}(t)\left|\vec{Q}(t),\vec{m}'(t)\right.\right] \\ = \;& R\prod_{k=1}^{i-1}\left(1-\frac{m_i(t)}{L}\right)\times \\ & \left(m_{i+1}(t)Q_i(t)+\left(1-\frac{m_{i+1}(t)}{L}\right)m_i(t)Q_{i+1}(t)\right. \\ & \left.-m_i(t)Q_i(t)-\left(1-\frac{m_i(t)}{L}\right)m_{i+1}(t)Q_{i+1}(t)\right) \\ = \;& R\prod_{k=1}^{i-1}\left(1-\frac{m_i(t)}{L}\right)\times \\ & (m_{i+1}(t)Q_i(t)+m_i(t)Q_{i+1}(t) \\ & -m_i(t)Q_i(t)-m_{i+1}(t)Q_{i+1}(t)) \\ = \;& R\prod_{k=1}^{i-1}\left(1-\frac{m_i(t)}{L}\right)\times \\ & (m_{i+1}(t)-m_i(t))(Q_i(t)-Q_{i+1}(t)) \\ \geq \;& 0, \end{aligned}$$

where the last inequality holds because $m_{i+1}(t)>m_i(t)$ and $Q_i(t)\geq Q_{i+1}(t)$ according to the assumptions. ∎

Lemma 3: Given a feedback allocation $\vec{m}(t)$ such that $pL>m_i(t)\geq m_{i+1}(t)>0$ and $m_k(t)=0$ for all $k>i+1$, we can construct an alternative feedback allocation $\vec{m}'(t)$ such that $m'_i(t)=m_i(t)+1$, $m'_{i+1}(t)=m_{i+1}(t)-1$, and $m'_j(t)=m_j(t)$ for $j\neq i,i+1$, and

$$g(\vec{m}(t),\vec{Q}(t))\geq g(\vec{m}'(t),\vec{Q}(t)).$$

Proof: We compare $g(\cdot)$ under the two different feedback allocations. Since $m_k(t)=m'_i(t)=0$ for all $k>i+1$, and $D_k(t)$ is independent of the channel states of user h for any

k and any $h > k$, we can obtain that

$$
\begin{aligned}
& g(\vec{m}'(t), \vec{Q}(t)) - g(\vec{m}(t), \vec{Q}(t)) \\
= & \mathbf{E}\left[Q_i(t)D_i(t) + Q_{i+1}(t)D_{i+1}(t) \middle| \vec{Q}(t), \vec{m}(t)\right] \\
& -\mathbf{E}\left[Q_i(t)D_i(t) + Q_{i+1}(t)D_{i+1}(t) \middle| \vec{Q}(t), \vec{m}'(t)\right] \\
= & RQ_i(t)\left(\prod_{k=1}^{i-1}\left(1 - \frac{m_k(t)}{L}\right)\right)((m_i(t)+1) - m_i(t)) \\
+ & RQ_{i+1}(t)\left(\prod_{k=1}^{i-1}\left(1 - \frac{m_k(t)}{L}\right)\right) \times \\
& \left(\left(1 - \frac{m_i(t)+1}{L}\right)(m_{i+1}(t) - 1)\right. \\
& \left. - \left(1 - \frac{m_i(t)}{L}\right) m_{i+1}(t)\right) \\
= & R\left(\prod_{k=1}^{i-1}(1 - \frac{m_k(t)}{L})\right) \times \\
& \left(Q_i(t) - \frac{L - m_i(t) + m_{i+1}(t) - 1}{L} Q_{i+1}(t)\right) \\
\geq & R\prod_{k=1}^{i-1}\left(1 - \frac{m_k(t)}{L}\right)(Q_i(t) - Q_{i+1}(t)) \\
\geq & 0
\end{aligned}
$$

■

Next, we consider a simple resource allocation algorithm based on the mean approximation, which we name as Longest-Queue-First Feedback-Allocation (LQF+FA):

Longest-Queue-First Feedback-Allocation (LQF+FA): Assuming that at most F channel states can be reported at each time slot and the users are sorted in a descent order according to their queue lengths, the feedback resource allocation vector $\vec{m}(t)$ is as follows:

$$
m_i(t) = \begin{cases} pL, & i \leq \left\lfloor \frac{F}{pL} \right\rfloor; \\ F \pmod{pL}, & i = \left\lfloor \frac{F}{pL} \right\rfloor; \\ 0, & \text{otherwise.} \end{cases}
$$

□

Theorem 3: Given the feedback resource constraint F, the LQF-FA+MaxWeight maximizes $g(\vec{m}(t), \vec{Q}(t))$. Therefore the LQF-FA+MaxWeight is throughput optimal under the mean approximation.

Proof: Denote by $\vec{m}^{LQF-FA}(t)$ to be the feedback allocation vector under LQF-FA. In the following analysis, we prove that given the feedback budget F,

$$
g(\vec{m}^{LQF-FA}(t), \vec{Q}(t)) \geq g(\vec{m}(t), \vec{Q}(t))
$$

for any $\vec{m}(t)$ such that $\sum_i m_i(t) \leq F$. The idea is to transform any feedback allocation $\vec{m}(t)$ to $\vec{m}^{LQF+FA}(t)$ in a way that the value of $g(\cdot, \vec{Q}(t))(t)$ is not reduced.

Given an allocation scheme $\vec{m}(t)$, we transform it to $\vec{m}^{LQF+FA}(t)$ by repeating the following two steps:

(1) **Step 1 - Reordering:** If $m_i(t) < m_{i+1}(t)$, we exchange the feedback resource allocated to user i and user $i+1$. Repeat this until we get a feedback allocation $\vec{m}(t)$ such that $m_i(t) \geq m_j(t)$ if $i < j$. According to Lemma 2, this reordering will not reduce the value of $g(\cdot, \vec{Q}(t))$. After the reordering, the process stops if $\vec{m}(t) = \vec{m}^{LQF+FA}(t)$; otherwise goes to step 2.

(2) **Step 2 - Reallocation:** After the reordering, we find user i such that

$$
i = \max\{k : m_k(i) \neq 0\}.
$$

We then reallocate the feedback resource of user i to user $i-1$ until $m_{i-1}(t) = pL$. According to Lemma 3, this reallocation will not reduce the value of $g(\cdot, \vec{Q}(t))$. After the reallocation, the process stops if $\vec{m}^{LQF+FA}(t)$ is obtained; otherwise, goes to step 1.

It is easy to see after a complete operation of step 1 — step 2 — step 1, we can obtain a feedback allocation scheme $\vec{m}(t)$ such that $m_1(t) = \min\{pL, F\}$. Therefore the LQF-FA allocation can be obtained in a finite number of iterations. Since $g(\cdot, \vec{Q}(t))$ is not reduced in either step 1 or step 2, so

$$
g(\vec{m}^{LQF-FA}(t), \vec{Q}(t)) \geq g(\vec{m}(t), \vec{Q}(t))
$$

for any $\vec{m}(t)$. The theorem holds due to Theorem 1. ■

IV. Modified LQF-FA Scheme

Motivated by the LQF-FA, we propose the following Modified LQF-FA and analyze its performance without the mean approximation.

Modified-Longest-Queue-First Feedback-Allocation (MLQF-FA): Assuming that at most F channel states can be reported at each time slot and the users are sorted in a descent order according to their queue lengths, the feedback resource allocation vector $\vec{m}(t)$ is as follows:

$$
m_i(t) = \begin{cases} (1+\delta)pL, & i \leq \left\lfloor \frac{F}{(1+\delta)pL} \right\rfloor; \\ F \pmod{(1+\delta)pL}, & i = \left\lfloor \frac{F}{(1+\delta)pL} \right\rfloor; \\ 0, & \text{otherwise.} \end{cases},
$$

where $\delta > 0$.

□

We will prove that compared to the MaxWeight scheduling with the complete channel state information, the throughput loss of the MLQF-FA+MaxWeight decreases exponentially as a function of F/L. Before the detail analysis, we first introduce several new notations. We denote by $D_i^{full}(t)$ the service rate allocated to user i under the MaxWeight with the complete network state information, and $D_{ij}^{full}(t)$ the rate contribution from channel j to user i under the MaxWeight with the complete network state information.

Theorem 4: Given a feedback resource constraint F, the MLQF-FA+MaxWeight can support any traffic load $\vec{A}(t)$ such that $(1+\delta)\vec{A}(t)$ is supportable under the MaxWeight with the complete channel state information, where

$$
\delta = \frac{1}{p^2}\exp\left(-\frac{\delta^2 pL}{3}\right) + \frac{1}{p}\exp\left(-\left\lfloor \frac{F}{(1+\delta)pL} \right\rfloor \log\frac{1}{1-p}\right).
$$

Proof: We first consider the MaxWeight scheduling with the complete channel state information:

$$\begin{aligned}
& Q_i(t)\mathbf{E}\left[D_i^{full}(t)|\vec{Q}(t)\right] \\
= & \sum_{j=1}^{L} Q_i(t)\mathbf{E}\left[D_{ij}^{full}(t)|\vec{Q}(t)\right] \\
= & LQ_i(t)\mathbf{E}\left[D_{i1}^{full}(t)|\vec{Q}(t)\right] \\
= & LQ_i(t)pR \times \\
& \Pr(\text{channel 1 is not on for users } \{1, \ldots, i-1\}) \\
= & LRp(1-p)^{i-1}Q_i(t).
\end{aligned}$$

As a result, we can obtain that

$$\begin{aligned}
\sum_{i=1}^{N} Q_i(t)\mathbf{E}[D_i^{full}(t)|\vec{Q}(t)] &= \sum_{i=1}^{N} LRpQ_i(t)(1-p)^{i-1} \\
&\leq LRpQ_1(t)\sum_{i=1}^{N}(1-p)^{i-1} \\
&< LRQ_1(t). \qquad (9)
\end{aligned}$$

Now we consider MLQF-FA+MaxWeight. For any user i such that $i \geq n+2$ where $n = \left\lfloor \frac{F}{(1+\delta)pL} \right\rfloor$, we have

$$Q_i(t)\mathbf{E}[D_i(t)|\vec{Q}(t), \vec{m}^{MLQF-FA}(t)] = 0. \qquad (10)$$

Now we consider user i such that $i \leq n$. Recall that $O_i(t)$ is the number of on channels to user i. Since the channels are i.i.d. on-off channels, according to the Chernoff's bound, we have

$$\Pr\left(O_i(t) \geq (1+\delta)pL\right) \leq \exp\left(-\frac{\delta^2 pL}{3}\right).$$

Therefore, when a user is assigned with $(1+\delta)pL$ feedback budgets, the probability that a channel is on and reported is at least

$$p - \exp\left(-\frac{\delta^2 pL}{3}\right),$$

which implies for any $i \leq n$,

$$\begin{aligned}
& Q_i(t)\mathbf{E}[D_i(t)|\vec{Q}(t), \vec{m}^{MLQF-FA}(t)] \\
\geq\ & LR(1-p)^{i-1}\left(p - \exp\left(-\frac{\delta^2 pL}{3}\right)\right)Q_i(t). \qquad (11)
\end{aligned}$$

According inequalities (9) and (11), we can conclude that

$$\begin{aligned}
& g(full, \vec{Q}(t)) - g(\vec{m}^{MLQF-FA}(t), \vec{Q}(t)) \\
\leq & \sum_{i=1}^{n}(1-p)^{i-1}LRQ_i(t)\exp\left(-\frac{\delta^2 pL}{3}\right) \\
& + \sum_{i=n+1}^{N}(1-p)^{i-1}LRpQ_i(t) \\
\leq & \frac{LRQ_1(t)}{p}\exp\left(-\frac{\delta^2 pL}{3}\right) + LRQ_1(t)(1-p)^n.
\end{aligned}$$

Note that $g(full, \vec{Q}(t)) \geq LRpQ_1(t)$, so we have that

$$\begin{aligned}
& \frac{g(full, \vec{Q}(t)) - g(\vec{m}^{MLQF+FA}(t), \vec{Q}(t))}{g(full, \vec{Q}(t))} \\
\leq & \frac{1}{p^2}\exp\left(-\frac{\delta^2 pL}{3}\right) + \frac{(1-p)^n}{p} \\
\leq & \frac{1}{p^2}\exp\left(-\frac{\delta^2 pL}{3}\right) + \frac{1}{p}\exp\left(-n\log\frac{1}{1-p}\right) \\
= & \frac{1}{p^2}\exp\left(-\frac{\delta^2 pL}{3}\right) + \frac{1}{p}\exp\left(-\left\lfloor\frac{F}{(1+\delta)pL}\right\rfloor\log\frac{1}{1-p}\right),
\end{aligned}$$

and the theorem holds due to Theorem 2. ■

Remark 3: In typical multichannel wireless networks, we expect that $F = O(L^2)$. In this case, the performance loss decreases exponentially as a function of F/L.

V. Simulation

In this section, we present simulation results to evaluate the performance of the proposed MLQF-FA-MaxWeight algorithm. The parameters of the simulation are summarized in Table I.

TABLE I
Simulation Parameters

Parameter	Value
Number of users N	50
Number of channels L	50
On probability p	0.4
Channel rate R	1 packet/channel/timeslot
Overall feedback number M	100–1000
Arrival Process	Poisson
Simulation duration	30000 iterations

There are five classes of traffic flows in the network: the mean arrival rate of user 1 is 10α, the mean arrival rate of users 2 is 6α, the mean arrival rate of users 3 and 4 is 4α, and the mean arrival rate of users 5 and 6 is 2α, and the mean arrival rate of users 7, 8, 9, 10 is α, and the mean arrival rate of users 11 to 50 is 0.2α. In our simulations,α varies from 0.2 to 1.2.

We consider three different scenario: (i) the complete channel state information is available at the base-station, (ii) the available feedback bandwidth is limited and is allocated according to the MLQF-FA scheme, (iii) the available feedback bandwidth is limited and is evenly allocated to all mobile users. When the base-station receives the reported feedback, it schedules the mobile users using the MaxWeight scheduling. Figure 1 illustrates the average total queue lengths under different schemes with different traffic loads. We can see that the MLQF-FA can support a much higher throughput compared to the even allocation scheme. For example, when 100 channel states can be reported at a time, the MLQF-FA+MaxWeight stabilizes the network when the mean of the sum arrival rates is less than 32, while the evenly-allocating+MaxWeight cannot stabilize the network when the sum of the arrival rates is more than 4. Further, when the number of reported channel states is 200, the performance of the MLQF-FA+MaxWeight is almost the same as the MaxWeight with the complete channel state information (note that the complete channel state information requires *2,500* channel states).

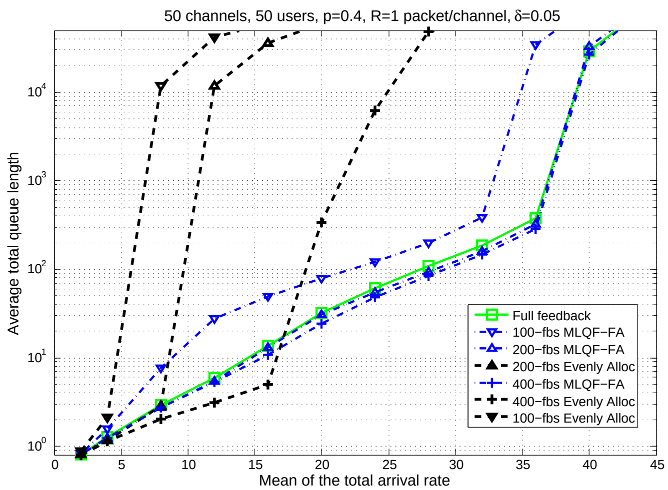

Fig. 1. Average total queue lengths under various feedback allocation schemes

VI. Conclusion

In this paper, we studied multi-channel downlink networks with limited feedback bandwidth F. We proposed the MLQF-FA scheme, which dynamically allocates the feedback resource according to the queue-lengths and channel statistics. We proved that the throughput difference between the MLQF-FA+MaxWeight and the MaxWeight with the complete channel state information decreases exponentially as a function of F/L when $F = O(L^2)$, where the L is the number of shared fading channels.

Acknowledgment: This work was partially supported by NSF Grant CNS-0831756 and the DTRA grant HDTRA1-09-1-0055.

References

[1] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inform. Theory*, vol. 39, pp. 466–478, March 1993.

[2] IEEE, "Ieee std 802.16-2009," *IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems*, 2009.

[3] R. Vannithamby, G. Li, H. Yin, and S. Ahmadi, "Proposal for IEEE 802.16m CQI feedback framework," *IEEE C802.16m-08/391*, 2008.

[4] 3GPP, "3GPP TS 36.300-870," *3rd Generation Partnership Project*.

[5] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Automat. Contr.*, vol. 4, pp. 1936–1948, December 1992.

[6] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting, "Scheduling in a queueing system with aynchronously varying service rates," *Probability in the Engineering and Informational Sciences*, vol. 18, pp. 191–217, 2004.

[7] A. Eryilmaz, R. Srikant, and J. Perkins, "Stable scheduling policies for fading wireless channels," *IEEE/ACM Trans. Network.*, vol. 13, pp. 411–424, 2005.

[8] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," vol. 23, no. 1, pp. 89–103, Janurary 2005.

[9] X. Qin and R. Berry, "Opportunistic splitting algorithms for wireless networks," in *Proc. IEEE Infocom.*, March 2004.

[10] Z. Ji, Y. Yang, J. Zhou, M. Takai, and R. Bagrodia, "Exploiting medium access diversity in rate adaptive wireless LANs," in *Proc. ACM MOBICOM*, 2004.

[11] T. Tang and R. Heath, "Opportunistic feedback for downlink multiuser diversity," *IEEE Communications Letters*, vol. 9, no. 10, pp. 948–950, 2005.

[12] S. Guha, K. Munagala, and S. Sarkar, "Performance guarantees through partial information based control in multichannel wireless networks," University of Pennsylvania, Tech. Rep., 2006, http://www.seas.upenn.edu/∼swati/report.pdf.

[13] A. Sabharwal, A. Khoshnevis, and E. Knightly, "Opportunistic spectral usage: Bounds and a multi-band CSMA/CA protocol," *IEEE/ACM Transactions on Networking*, 2006.

[14] S. Patil and G. Veciana, "Managing resources and quality of service in heterogeneous wireless systems exploiting opportunism," *IEEE/ACM Trans. Netw.*, vol. 15, no. 5, pp. 1046–1058, 2007.

[15] N. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," in *ACM Int. Conf. on Mobile Computing and Networking (MobiCom)*, September 2007.

[16] A. Gopalan, C. Caramanis, and S. Shakkottai, "On wireless scheduling with partial channel-state information," in *Proc. Ann. Allerton Conf. Communication, Control and Computing*, Urbana, IL, September 2007.

[17] D. Zheng, W. Ge, and J. Zhang, "Distributed opportunistic scheduling for ad-hoc communications: An optimal stopping approach," in *ACM Mobihoc*, 2007.

[18] L. Ying and S. Shakkottai, "On throughput optimality with delayed network-state information," in *Technical Report*, 2008.

[19] ——, "Scheduling in mobile wireless networks with topology and channel-state uncertainty," in *Proc. IEEE Infocom.*, Rio de Janeiro, Brazil, 2009.