Scheduling in Time-correlated Wireless Networks with Imperfect CSI and Stringent Constraint

Wenzhuo Ouyang, Atilla Eryilmaz, and Ness B. Shroff

Abstract— **In a wireless network, the efficiency of scheduling algorithms over time-varying channels depends heavily on the accuracy of the Channel State Information (CSI), which is usually quite "costly" in terms of consuming network resources. Scheduling in such systems is also subject to stringent constraints such as power and bandwidth, which limit the maximum number of simultaneous transmissions. In the meanwhile, communication channels in wireless systems typically fluctuate in a time-correlated manner. We hence design schedulers to exploit the temporal-correlation inherent in channels with memory and Automatic Repeat reQuest (ARQ) feedback from the users for better channel state knowledge, under the assumption of Markovian channels and the stringent constraint on the maximum number of simultaneously active users. We model this problem under the framework of a Partially Observable Markov Decision Processes.**

In recent work, a low-complexity optimal solution was developed for this problem under a long-term time-average resource constraint. However, in real systems with instantaneous resource constraints, how to optimally exploit the temporal correlation and satisfy realistic stringent constraint on the instantaneous service remains elusive. In this work, we incorporate a stringent constraint on the simultaneously scheduled users and propose a low-complexity scheduling algorithm that dynamically implements user scheduling and dummy packet broadcasting. We show that the throughput region of the optimal policy under the long-term average resource constraint can be asymptotically achieved in the stringent constrained scenario by the proposed algorithm, in the many users limiting regime.

I. INTRODUCTION

In wireless networks, the states of the wireless channels fluctuate in time. This characteristic calls for designing resource allocation algorithms that dynamically adapt to the random variation of the wireless channels. Scheduling algorithms are essential components of resource allocation. A scheduling algorithm is designed to control a subset of users to consume the scarce network resources (e.g., bandwidth, power), so that the overall network utility is maximized subject to link interference and queue stability constraints. Under the assumption that accurate instantaneous Channel State Information (CSI) is available at the scheduler, maximum-weight-type scheduling algorithms (e.g., [1]-[3]) are known to be throughput-optimal, i.e., they can maintain system stability for arrival rates that are supportable by any other scheduler.

The performance of efficient scheduling algorithm relies heavily on the accurate instantaneous CSI at the scheduler. In practice, however, accurate instantaneous CSI is difficult to obtain at the scheduler, i.e., a significant amount of system resources must be spent to accurately estimate the instantaneous CSI. Therefore, acquiring CSI continuously from all users is resource-consuming and impractical as the size of network increase. Hence, in this work we consider the important scenario where the instantaneous CSI is not directly accessible to the scheduler, but is instead learned at the user and fed back to the scheduler via ARQ feedback after a certain delay. Many scheduling algorithms have been designed that consider imperfect CSI, where the channel state is considered as independent and identically distributed (i.i.d.) processes over time (e.g., [4]-[6]). However, although the *i.i.d.* channel models facilitate trackable analysis, it does not capture the time-correlation of the fading channels.

Because perfect instantaneous CSI is costly to acquire, the time-correlation or channel memory inherent in the fading channels is an important resource that can be exploited by the scheduler to make more informed decisions, and hence to obtain significant throughput/utility gains (e.g., [7]-[17]). Under imperfect CSI, channel memory, and resources constraint, the scheduler needs to intelligently balance the intricate 'exploitation-exploration tradeoff', i.e., to decide at each slot whether to exploit the channels with more up-to-date CSI, or to explore the channels with outdated CSI.

We consider the downlink of a single cell, where the packets for each user are stored in a data queue for transmission. Under the complicated channel memory evolution and queue evolution, traditional Dynamic Programming based approaches can be used for designing scheduling schemes, but are intractable due to the well-known 'curse of dimensionality'. Recently, a low-complexity algorithm was proposed in [7] that considers throughput-optimal downlink scheduling with imperfect CSI over time-correlated fading channels, under a constraint on the *long-term average* number of transmissions.

Scheduling in wireless systems is typically subject to *stringent instantaneous* constraints, such as instantaneous resource limitations from bandwidth, power, interference, etc. In this work, we study scheduling with imperfect CSI over time-correlated channels and under stringent resource constraint where the instantaneous scheduling decision is subject to constraint on the maximum number of scheduled

Wenzhuo Ouyang is with Department of ECE, Rice University (e-mail: wenzhuo.ouyang@rice.edu). Atilla Eryilmaz is with the Department of ECE, The Ohio State University (e-mail: eryilmaz.2@osu.edu). Ness B. Shroff holds a joint appointment in both the Department of ECE and the Department of CSE at The Ohio State University (e-mail: shroff.11@osu.edu).

This work is supported by the QNRF grants NPRP 7-923-2-344, NPRP 09-1168-2-455, DTRA grant # BRBAA11-PerC-4-2-0034, HDTRA1-11-16-BRCWMD, ARO MURI grant W911NF-14-1-0368, and NSF grants CAREER-CNS-0953515, CCSS-EARS-1444026, CNS-1421576, and CNS-1065136.

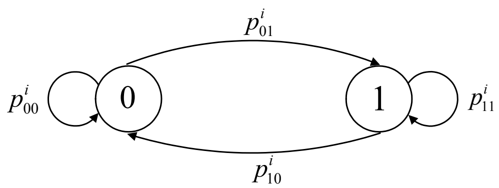

Fig. 1: Two state Markov Chain model.

users. The stringent constraint brings with it significant challenges, and to the best of our knowledge under the setting of imperfect CSI, no low-complexity algorithm exists that is optimal for general scenarios. Under the restrictive regime where users have identical ON/OFF Markovian channel statistics, round-robin based scheduling policies are shown to be throughput optimal in [12][13]. Further, under these settings, it has been shown in [14][15] that greedy scheduling algorithms are also optimal. In [16][17], throughput-optimal frame-based policies are proposed. These policies rely on solving a Linear Programming in each frame, which is hindered by the curse of dimensionality where the computational complexity grows exponentially with the network size.

In this paper, we propose a low-complexity algorithm under stringent constraint and heterogeneous Markovian transition statistics across users. We prove that the proposed algorithm has asymptotical optimal properties in the regime of a large number of users. Our contributions are as follows:

- Under stringent constraint on the instantaneous number of transmissions, we propose a novel low-complexity *joint scheduling and broadcasting* algorithm. At each slot, the scheduler dynamically decides whether to schedule a subset of users and learn their channel state via ARQ feedback, or to broadcast a dummy packet to a larger set of users to learn their channel states from ARQ feedback but with no throughput gain.
- We conduct our analysis in the framework of Partially Observable Markov Decision Process, where we utilize Whittle's index analysis of Restless Multi-armed Bandit Problem (RMBP) [18]. We then use a *Large-Deviation-based Lyapunov* technique over time frames to prove the throughput performance of the proposed algorithm.
- We prove that the throughput region in [7], which is achieved by an optimal policy under a relaxed constraint on the long-term average number of transmissions, can be asymptotically achieved in the stringent constrained scenario by the proposed algorithm, in the regime of a large number of users.

II. System Model

A. Downlink Scheduling Problem

We study a time-slotted wireless downlink network with one Base Station (BS) and N users. Each user i occupies a dedicated wireless channel. The channel state of user i, denoted by $C_i[t]$, evolves as an ON/OFF Markov chain across time slots with state space $\mathcal{S} = \{0, 1\}$, independently of other channels. State '1' represents high channel gain where one packet can be transmitted successfully, whereas state '0' represents deep fading state where no packet can be delivered. The Markovian channel state evolution is depicted in Fig. 1, represented by the transition probabilities

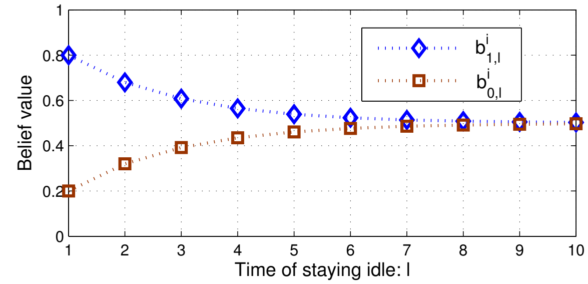

Fig. 2: Belief value evolution, $p^i_{11} = 0.8$, $p^i_{01} = 0.2$.

$$p^i_{jk} := \Pr\big(C_i[t]{=}k \big| C_i[t{-}1]{=}j\big), j, k \in \mathcal{S}.$$

We assume that $p^i_{11} > p^i_{01}$ for $i{=}1, 2, \cdots, N$. This assumption is called positive correlation and is commonly made in this field (e.g., [9][10][12][16][19]). We also assume the existence of a constant $\delta > 0$ so that $p^i_{01}{>}\delta$ and $p^i_{10}{>}\delta$ for all i to allow a minimum cross-state transition probability, which captures the random varying nature of the wireless channels.

Data packets for different users are first stored in separate queues at the BS before successful transmission. The queue length for user i at slot t is denoted by $q_i[t]$. The number of data packet that arrives at queue i is denoted as $A_i[t]$, which forms an *i.i.d.* process with mean λ_i and a bounded second moment. At the beginning of every time slot, the scheduler at the BS selects users for data transmission. We let $a_i[t] \in \{0, 1\}$ indicate whether user i is scheduled at slot t. The i-th data queue evolves as $q_i[t{+}1]{=}\max\{0, q_i[t]{-}a_i[t]{\cdot}C_i[t]\}{+}A_i[t]$.

The scheduling decisions are made without the exact knowledge of the channel state in the current slot. In our model, the scheduler at the BS instead obtains the accurate CSI via ACK/NACK feedback, only from the scheduled users *at the end of each slot* following data transmission, i.e., an ACK from scheduled user i implies $C_i[t] = 1$, while an NACK implies $C_i[t] = 0$.

We consider the class Φ of (possibly non-stationary) scheduling policies that make decisions based on the history of observed channel states, arrival processes, and scheduling decisions. Under the aforementioned instantaneous constraint, the scheduling schemes are subject to the constraint on the number of scheduled transmissions at each slot, i.e.,

$$\sum_{i=1}^{N} a^\phi_i[t] \leq M, \tag{1}$$

where $M{\leq}N$, and $a^\phi_i[t]{\in}\{0, 1\}$ indicates if the i-th user is scheduled at slot t under policy $\phi{\in}\Phi$.

B. Belief Value Evolution

The scheduler maintains a belief value $\pi_i[t]$ for each channel i, which is the probability of channel i being in state 1 at the beginning of t-th slot conditioned on the past channel state observations. The belief values are hence updated according to the scheduling decisions and channel state feedbacks,

$$\pi_i[t+1] = \begin{cases} p_{11}^i & \text{if } a_i[t]=1 \text{ and } C_i[t]=1, \\ p_{01}^i & \text{if } a_i[t]=1 \text{ and } C_i[t]=0, \\ Q_i(\pi_i[t]) & \text{if } a_i[t]=0, \end{cases} \quad (2)$$

where $Q_i(x){=}xp_{11}^i{+}(1{-}x)p_{01}^i$. For our downlink scheduling problem, the belief values are known to be sufficient statistics to represent the past channel state feedback [20]. In the meanwhile, the value $\pi_i[t]$ is also the expected throughput for user i if it is scheduled in slot t.

We use $b_{c,l}^i$ to denote the state of i-th channel's belief value when the most recent channel state was observed l time slots ago and was in state $c \in \{0,1\}$, given by

$$b_{0,l}^i{=}\frac{p_{01}^i-(p_{11}^i{-}p_{01}^i)^l p_{01}^i}{1+p_{01}^i-p_{11}^i}, b_{1,l}^i{=}\frac{p_{01}^i{+}(1{-}p_{11}^i)(p_{11}^i{-}p_{01}^i)^l}{1+p_{01}^i-p_{11}^i}.$$

As depicted in Fig. 2, when the scheduler is never informed of the i-th user's channel state, $\pi_i[t]$ monotonically converges to the stationary probability $b_s^i{:=}p_{01}^i/(1+p_{01}^i-p_{11}^i)$ of the channel being in state 1. We assume that all belief values are initially set to their stationary values. Hence, based on (2), each belief value $\pi_i[t]$ evolves over a countable state space, denoted by $\mathcal{B}_i{=}\{b_s^i, b_{c,l}^i \; : \; c{\in}\{0,1\}, l{\in}\mathbb{Z}^+\}$. We also consider the τ-truncated state apace $\mathcal{B}_i^\tau{=}\{b_s^i, b_{c,l}^i : c{\in}\{0,1\}, l{=}1,2,\cdots,\tau\}$.

C. Network Stability Regions

Queue i is said to be stable if there exists a limiting stationary distribution F_i such that $\lim_{t\to\infty} P(q_i[t] \le q) = F_i(q)$ [1]. When there are N total downlink users and at most M users can be simultaneously scheduled, the *network stability region* $\mathbf{\Lambda}_{str}^{N,M}$ is defined as the closure of the set of arrival rate vectors stably supported by all policies in class Φ while abiding by the stringent constraint (1).

For comparison, we introduce another region $\mathbf{\Lambda}_{rel}^{N,M}$ as the closure of the set of arrival rate vectors supported by all policies in class Φ that maintains queue stability and satisfies the following *relaxed constraint* that only requires an *average number* of M users to be activated in the long run,

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N} a_i^\phi[t]\Big] \le M. \quad (3)$$

The region $\mathbf{\Lambda}_{rel}^{N,M}$ provides a benchmark for our analysis on stringent constrained scenarios. Contrary to the stringent constraint (1), the relaxed constraint (3) allows the activation of more than M users at a time if the long term average number of transmissions does not exceed M. Hence the corresponding region $\mathbf{\Lambda}_{rel}^{N,M}$ provides an upper bound to $\mathbf{\Lambda}_{str}^{N,M}$.

III. Optimal Policy for Weighted Sum-throughput Maximization under a Relaxed Constraint

We begin our analysis by introducing the weighted sum-throughput maximization problem under the relaxed constraint, which serves as a benchmark for our main result.

Specifically, consider the following weighted sum-throughput maximization problem $\Psi_{rel}(\mathbf{r}, N, M)$ for a given non-negative vector $\mathbf{r} = (r_i)_{i=1}^N$,

$$\max_{\phi\in\Phi} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N} r_i{\cdot}\pi_i[t]{\cdot}a_i^\phi[t]\Big] \quad \text{s.t. constraint (3).}$$

The above problem is a constrained Partially Observable Markov Decision Process (CPOMDP) and can be tackled in the framework of the Restless Multiarmed Bandit Problem (RMBP) [18] by making use of the associated Whittle's index value. Specifically, for problem $\Psi_{rel}(\mathbf{1}, N, M)$, a closed-form *Whittle's index value* $W_i^{\mathbf{1}}(\pi)$ is assigned to each belief state $\pi \in \mathcal{B}_i$ of channel i. These indices intelligently capture the exploitation-exploration value to be gained from scheduling the user at the corresponding belief state [9]. The $\mathbf{r}$-weighted index value is defined as $W_i^{\mathbf{r}}(\pi) = r_i \cdot W_i^{\mathbf{1}}(\pi)$ for all i.

For details of Whittle's indexability analysis, please refer to [11][7]. The following policy, denoted as $\phi_{rel}^\tau(\mathbf{r}, N, M)$, was proposed in [7] to tackle problem $\Psi_{rel}(\mathbf{r}, N, M)$.

Algorithm $\phi_{rel}^\tau(\mathbf{r}, N, M)$**:** $\mathbf{r}$-weighted Index Policy

1: **Initialization phase:** The parameters ω_τ and ρ_τ are calculated by algorithm $G^\tau(\mathbf{r}, N, M)$.

2: **At slot** t**:** user i is scheduled if the $\mathbf{r}$-weighted index value $W_i^{\mathbf{r}}(\pi_i[t]) > \omega_\tau$, and stays passive if $W_i^{\mathbf{r}}(\pi_i[t]) < \omega_\tau$. If $W_i^{\mathbf{r}}(\pi_i[t]){=}\omega_\tau$, user i is scheduled with probability ρ_τ.

The algorithm $G^\tau(\mathbf{r}, N, M)$ in the initialization phase is given in the next page, where the closed-form expression of function α_i^τ and β_i can be found in [7].

We henceforth use $V^*(\mathbf{r}, N, M)$ to denote the optimal weighted sum-throughput of problem $\Psi_{rel}(\mathbf{r}, N, M)$. We let $V_\tau(\mathbf{r}, N, M)$ be the throughput under policy $\phi_{rel}^\tau(\mathbf{r}, N, M)$,

$$V_{rel}^\tau(\mathbf{r},N,M){=}\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N} r_i\pi_i[t]a_i^{\phi_{rel}^\tau(\mathbf{r},N,M)}[t]\Big]. \quad (4)$$

It was shown in [7] that $V_\tau(\mathbf{r}, N, M)$ is arbitrarily close to $V^*(\mathbf{r}, N, M)$ for large values of τ.

IV. Weighted Sum-throughput Maximization Problem under Stringent Constraint

Next we also consider the $\mathbf{r}$-weighted sum throughput optimization problem as in the last section, but under the stringent constraint, i.e., no more than M users are scheduled for data transmission at each time slot. We propose a joint scheduling and broadcasting algorithm that abides by the stringent constrained. This algorithm has novelty of incorporating the possibility of *broadcasting* a dummy packet at a slot, and can provide performance asymptotically close to the optimal algorithm for the relaxed problem.

A. Policy with Joint Scheduling and Broadcasting

The proposed policy, denoted by $\phi_{str}^\tau\big(\mathbf{r}, N, M, K\big)$ with $K{\le}M$, builds on the policy $\phi_{rel}^\tau\big(\mathbf{r}, N, M\big)$ for the relaxed problem. However, it fundamentally differs from

Algorithm $G^\tau(\mathbf{r}, N, M)$**:** Calculation of ω_τ and ρ_τ

TxTime$[i] = 1$ for all $i \in \{1, \cdots, N\}$
TotalTime $= N$
struct Index
{ float value
  int user
} $\mathbf{I}[(2\tau+1)N], \mathbf{w}[(2\tau+1)N]$
$j = 0$
for $i = 1$ to N **do**
    for each $\pi_i \in \mathcal{B}_i^\tau$ **do**
        $W_i^{\mathbf{r}}(\pi_i) = r_i \cdot W_i^{\mathbf{1}}(\pi_i)$
        $\mathbf{I}[j]$.value= $W_i^{\mathbf{r}}(\pi_i)$
        $\mathbf{I}[j]$.user= i
        $j \leftarrow j+1$
    end for
end for
$\mathbf{w}$ =sort($\mathbf{I}$)
for $k = 1$ to size($\mathbf{w}$) **do**
    NewTime$[\mathbf{w}[k]$.user$] = \alpha^\tau_{\mathbf{w}[k].\mathrm{user}}(\mathbf{w}[k].\mathrm{value}, 1)$
    TimeDiff = TxTime$[\mathbf{w}[k]$.user$]-$NewTime$[\mathbf{w}[k]$.user$]$
    TotalTime = TotalTime $-$ TimeDiff
    if TotalTime $< M$ **then**
        $\omega_\tau = \mathbf{w}[k-1]$.value
        TxTime$[\mathbf{w}[k-1]$.user$] = M - \sum_{i \neq \mathbf{w}[k-1].\mathrm{user}}$ TxTime$[i]$
        $\rho_\tau = \beta_{\mathbf{w}[k-1].\mathrm{user}}(\omega_\tau, \mathrm{TxTime}[\mathbf{w}[k-1].\mathrm{user}])$
        Break
    end if
    TxTime$[\mathbf{w}[k]$.user$]$=NewTime$[\mathbf{w}[k]$.user$]$
end for
return ω_τ, ρ_τ

$\phi^\tau_{rel}\big(\mathbf{r}, N, M\big)$ in the following way. At the beginning of each slot, algorithm $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$ carefully makes one of two choices: 1) transmit data packets to no more than M users and receive ARQ-type feedback from them, or 2) *broadcast* a short dummy packet to more than M users, and learn their channel states from their ARQ-type feedback. Note that, the dummy packet is known to the users and contains no new information and hence does not bring throughput gains if it is broadcasted. However, the scheduler still receive ARQ feedback from the candidates, and hence obtain CSI update from possibly more than M users. The parameter K controls how aggressively the dummy packets are broadcasted. Algorithm $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$ is proposed in the right-hand column.

Algorithm $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$ **under stringent constraint**

1. **Initialization phase:** The parameters ω_τ and ρ_τ are calculated by algorithm $G^\tau(\mathbf{r}, N, K)$.
2. **At slot t, candidate selection:** user i is called a '*candidate*', represented by $\theta_i[t]$=1, if the $\mathbf{r}$-weighted index value $W_i^{\mathbf{r}}(\pi_i[t]) > \omega_\tau$, and is not a candidate, i.e., $\theta_i[t]$=0, if $W_i^{\mathbf{r}}(\pi_i[t]) < \omega_\tau$. If $W_i^{\mathbf{r}}(\pi_i[t]) = \omega_\tau$, user i becomes a 'candidate' with probability ρ_τ.
3. **At slot t, transmission:** If the total number of candidates is under M, i.e., $\sum_{i=1}^N \theta_i[t] \le M$, then all the candidates are scheduled for data transmission, i.e., $a_i^{\phi^\tau_{str}\left(\mathbf{r},N,M,K\right)}[t] = \theta_i[t]$ for all i. If there are more than M candidates, then $a_i^{\phi^\tau_{str}\left(\mathbf{r},N,M,K\right)}[t] = 0$ for all i, and dummy packet is *broadcasted*.
4. **At slot t, feedback:** At the end of each slot, if data packets are transmitted, the scheduled users send ARQ feedback to the BS; if the dummy packet is broadcasted, the *candidates* send ARQ feedback to the BS. The belief values are updated based on the feedback.

Remarks:

(1) Steps 1-2 of algorithm $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$ is exactly algorithm $\phi^\tau_{rel}\big(\mathbf{r}, N, K\big)$, where the scheduled users in algorithm $\phi^\tau_{rel}\big(\mathbf{r}, N, K\big)$ becomes the candidates in $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$.
(2) Step 3 ensures that the stringent interference constraint is met so that data packets are transmitted to no more than M users. If the number of candidates exceeds M, a dummy packet is broadcasted for the scheduler to learn the channel states of the candidates and no throughput is accrued.
(3) Because of step 4, the scheduler receives channel state feedback from all the candidates, although data packets may not be transmitted. By taking this approach, the channel memory evolution in the relaxed constrained algorithm $\phi^\tau_{rel}\big(\mathbf{r}, N, K\big)$ is maintained in the stringent constrained algorithm $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$, which facilities much more trackable performance analysis.
(4) In step 4, only the candidates (instead of all users) send feedback to the BS if dummy packet is broadcasted. By allowing only the candidates to feedback[1], the algorithm not only helps maintain the tractability of channel memory evolution, it also fits with the realistic scenario where it is costly (in terms of time, power, bandwidth, etc.) to obtain feedback from a large number of users.

We henceforth let $V^\tau_{str}\big(\mathbf{r}, N, M, K\big)$ be the weighted sum-throughput under policy $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$, i.e.,

$$\begin{aligned} &V^\tau_{str}\big(\mathbf{r}, N, M, K\big) \\ &= \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}\Big[\sum_{t=0}^{T-1} \sum_{i=1}^{N} r_i \cdot \pi_i[t] \cdot a_i^{\phi^\tau_{str}\left(\mathbf{r},N,M,K\right)}[t]\Big]. \end{aligned} \tag{5}$$

B. Performance of the algorithm under stringent constraint

From the algorithm and Remark (1) thereafter, in each slot, if the number of scheduled users exceeds M under algorithm $\phi^\tau_{rel}\big(\mathbf{r}, N, K\big)$ for the relaxed problem, the number of candidates under algorithm $\phi^\tau_{str}\big(\mathbf{r}, N, M, K\big)$ exceeds M and a dummy packet is broadcasted, otherwise all candidates are scheduled for data transmission. Hence in the regime when K is close to M, the larger the K, the more aggressively are dummy packets broadcasted, which bring more updated system-level channel state information, but with a tradeoff that no throughput is obtained in these broadcasting slots. On

[1]This can be achieved by marking the corresponding bits in the dummy packet.

the other hand, in the regime when K is away from M, the smaller the K, on average there are less candidates and hence scheduled users, which also brings down the throughput.

The next lemma bounds the difference between the throughput performance $V^{\tau}_{str}\big(\mathbf{r}, N, M, K\big)$ of algorithm $\phi^{\tau}_{str}\big(\mathbf{r}, N, M, K\big)$ for the stringent constrained problem, and the throughput $V^{\tau}_{rel}\big(\mathbf{r}, N, M\big)$ of $\phi^{\tau}_{rel}\big(\mathbf{r}, N, M\big)$ for the problem under relaxed constraint. Recall that $V^{\tau}_{rel}\big(\mathbf{r}, N, M\big)$ and $V^{\tau}_{str}\big(\mathbf{r}, N, M, K\big)$ were defined in (4) and (5), and δ was defined in the introduction so that $p^i_{01}{>}\delta$ and $p^i_{10}{>}\delta$ for all i.

Lemma 1: The following bounds hold for the values of $V^{\tau}_{str}(\mathbf{r}, N, M, K)$ and $V^{\tau}_{rel}\big(\mathbf{r}, N, M\big)$,

$$\mu(M, K) \leq \frac{V^{\tau}_{str}(\mathbf{r}, N, M, K)}{V^{\tau}_{rel}\big(\mathbf{r}, N, M\big)} \leq 1, \tag{6}$$

where

$$\mu(M, K){=}\Big[1{-}\exp({-}\frac{(M{-}K)^2}{3K})\Big]\cdot\Big[1{-}\frac{M{-}K}{\delta(K{-}1)}\Big]^+, \tag{7}$$

and $[\cdot]^+$ represents $\max\{0, \cdot\}$.

Proof Outline: In the proof, we first bound the steady state probability that dummy packets are transmitted using Large Deviation techniques, from which we obtain the first multiplicand in (7). We next bound the effect of K in the throughput different between $V^{\tau}_{str}(\mathbf{r}, N, M, K)$ and $V^{\tau}_{rel}\big(\mathbf{r}, N, M\big)$, which brings us the second multiplicand in (7). Details of the proof can be found in [23]. ■

The previous lemma is important to derive the asymptotic throughput performance of the stringent constrained policy, captured in the next proposition. The proposition shows that as both N and M become large, if the parameter K is kept an *appropriate distance* $g(N, M)$ from M, then the throughput performance of policy $\phi^{\tau}_{str}\big(\mathbf{r}, N, M, K\big)$ becomes asymptotically close to $\phi^{\tau}_{rel}\big(\mathbf{r}, N, M\big)$ of the relaxed policy.

Proposition 1: Suppose $K = M - g(M)$ when M users can be simultaneously scheduled, where $g(M) \geq 0$ is a function of M.
If function $g(M)$ satisfies $\lim_{M\to\infty} g(M)/M{=}0$ with $\lim_{M\to\infty} g^2(M)/M = \infty$, the throughput performance of policy $\phi^{\tau}_{str}\big(\mathbf{r}, N, M, M - g(M)\big)$ is asymptotically close to that of $\phi^{\tau}_{rel}\big(\mathbf{r}, N, M\big)$, i.e.,

$$\lim_{M\to\infty} \frac{V^{\tau}_{str}\big(\mathbf{r}, N, M, M{-}g(M)\big)}{V^{\tau}_{rel}\big(\mathbf{r}, N, M\big)} = 1. \tag{8}$$

Proof Outline: The proposition is proven by substituting $K{=}M{-}g(M)$ to (6)-(7). For details, please refer to [23]. ■

Remark: Proposition 1 states that, if the distance between K and M grows at an order larger than $O(\sqrt{M})$ but lower than $O(M)$, the performance of the proposed algorithm $\phi^{\tau}_{str}(\mathbf{r}, N, M, M - g(N))$ is asymptotically close to $\phi^{\tau}_{rel}(\mathbf{r}, N, M)$, which is optimal for the relaxed problem. This is an interesting finding, as it quantities the trade-off between scheduling data packets and broadcasting of dummy packets. When K is less than $O(\sqrt{M})$ to M, excessive training leaves insufficient slots for data transmission. If K is more than $O(M)$ from M, the scheduler is over-conservative on data transmission, which in turn reduces the throughput.

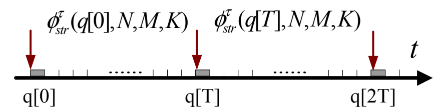

Fig. 3: Illustration of algorithm Frame$_\tau(T, N, M, K)$.

V. Queue-based Joint Scheduling and Broadcasting Policy over Time Frames

Note that, in the two last sections, we considered weighted sum-throughput. In this section, we consider the system model with data queues where *queue stability* is taking into account. Next, we propose a joint scheduling and broadcasting algorithm based on the algorithm $V^{\tau}_{str}(\mathbf{r}, N, M, K)$ in the last section. The policy is implemented over separate time-frames and has low-complexity.

We divide the time slots $\{0, 1, 2, \cdots\}$ into separate *time frames* of length T, i.e., the k-th frame, $k \in \{0, 1, 2, \cdots\}$, includes time slots $kT, \cdots, (k+1)T{-}1$. The scheduling decisions in the k-th frame are made based on the queue length information $\mathbf{q}[kT]$ at the beginning of that frame. During the k-th frame, the policy $\phi^{\tau}_{str}(\mathbf{q}[kT], N, M, K)$, developed in the last section, is implemented. This algorithm is illustrated in Fig. 3. Formally, with N users in the network and under stringent M constraint, the T-frame queue-based policy Frame$_\tau$(T,N,M,K) is introduced next.

Algorithm Frame$_\tau(T, N, M, K)$**:** T-Frame Queue-based Policy

---

1: The time slots are divided into frames of length T. Slot t is in the k-th frame if $kT \leq t < (k+1)T$, $k \in \{0, 1, \cdots\}$.
2: **At the beginning of the k-th frame:** At the beginning of slot kT, implement the algorithm $G^{\tau}(\mathbf{q}[kT], N, K)$ that outputs ω_τ and ρ_τ for the frame.
3: **At slot t, candidate selection:** Each user i becomes a *candidate* if the $\mathbf{q}[kT]$-weighted index value $W^{\mathbf{q}[kT]}_i(\pi_i[t]){>}\omega_\tau$, and is not a candidate if $W^{\mathbf{q}[kT]}_i(\pi_i[t]){<}\omega_\tau$. If $W^{\mathbf{q}[kT]}_i(\pi_i[t]){=}\omega_\tau$, user i becomes a 'candidate' with probability ρ_τ.
4: **At slot t, transmission:** If there are no more than M total candidates, then all the candidates are scheduled for data transmission. If there are more than M candidates, then a dummy packet is *broadcasted*.
5: **At slot t, feedback:** At the end of each slot, if data packets are transmitted, the scheduled users send ARQ feedback to the BS; if the dummy packet is broadcasted, the *candidates* send ARQ feedback to the BS. The belief values are updated correspondingly.

The next proposition establishes that the throughput region $\mathbf{\Lambda}^{N,M}_{rel}$, which is achieved by the optimal policy under a relaxed constraint on the long-term average number of transmissions, can be asymptotically achieved in the stringent constrained scenario by the frame-based algorithm, in the regime of a large number of users. In the proposition, $\mathbf{1}$

is an all 1 vector, and $g(N,M)$, $\mu(N,M,K)$ are given in Proposition 1.

Proposition 2: We let $l(M,K) = 1 - \mu(M,K)$. If $\tau \geq \tau_0 := \left\lceil 4\max\left\{\frac{1}{-\log(2\delta)}, \frac{1}{\log^2(2\delta)}\right\}\right\rceil$, we have
(i) if $K > M/2$, for all arrival rate $\boldsymbol{\lambda}$ with $\boldsymbol{\lambda} + \big(f(\tau) + 2l(M, M-g(M))\big)\mathbf{1} \in \boldsymbol{\Lambda}_{rel}^{N,M}$, there exists T_0 such that, if $T > T_0$, all queues are stable under the T-frame queue-based policy $Frame_\tau(T,N,M,M-g(M))$. The function $f(\tau)$ satisfies $\lim_{\tau\to\infty} f(\tau) = 0$.
(ii) if $\lim_{M\to\infty}\frac{g(M)}{M} = 0$ and $\lim_{M\to\infty}\frac{g^2(M)}{M} = \infty$, then the function $l(M, M-g(M))$ satisfies

$$\lim_{N\to\infty} l(M, M-g(M)) = 0. \tag{9}$$

Proof Outline: We prove the proposition using a Large-Deviation-based Lyapunov technique over time frames. Specifically, we combine the Large Deviation result in Lemma 1 with uniform convergence of the finite horizon throughput to the infinite horizon throughput performance. We then prove that the average Lyapunov drift of the queue lengths in each time frames is negative, which leads to the stability of the queues. Details are included in [23]. ■

Remark:
(1) Note that, in Proposition 2, the parameter K is kept a distance $g(M)$ from M. This mechanism is optimally controls the trade-off between transmitting data packets and broadcasting dummy packets so that we can apply Proposition 1 to guarantee the supportable stability region is asymptotic close to the relaxed constrained region $\Lambda_{rel}^{N,M}$, if $g(M)$ scales up at an appropriate rate.
(2) In the proposed algorithm, a user is selected based on its $\mathbf{q}[kT]$-weighted Whittle's index value in step 3. Since the Whittle's index value measures the importance of scheduling a user under the time-correlated channel, this multiplication captures the importance of scheduling a user under both queue evolution and the time correlation.
(3) In each frame of algorithm $\text{Frame}_\tau(T,N,M,M-g(N,M))$, implementation of $G^\tau(\mathbf{q}[kT],N,M,M-g(N,M))$ in step 2 has computational complexity $O((2\tau+1)N\log(2\tau+1)N)$ [7], while implementing step 3 over the frame has complexity $O(TN)$. Hence the *per-frame* complexity is $O((2\tau+1)N\log(2\tau+1)N + TN)$. As the frame length T scales up, the *per-slot* complexity decreases toward $O\big(N\big)$.

VI. Conclusion

In this work, we study downlink scheduling algorithm design over Markovian ON/OFF channels, where the scheduler does not possess accurate instantaneous channel state information. The scheduler instead exploits the Markovian channel memory and channel state feedback from users to make scheduling decisions. We proposed a low-complexity frame-based algorithm in downlink queuing networks with stringent constraint on the number of simultaneously scheduled users. The proposed algorithm dynamically determines whether to schedule data transmission or broadcast a dummy packet in a slot. By carefully choosing its parameter, the proposed algorithm stably supports arrival rates in a region asymptotically close to that under a relaxed constraint, when the number of users is large. Our on-going work involves designing scheduler for the scenario with both stringent and average constraints, as well as designing throughput optimal scheduler under stringent constraint for finite number of users.

References

[1] L. Tassiulas, A. Ephremides,"Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Transactions on Information Theory,* vol. 39, pp. 466-478, 1993.
[2] X. Lin, N. B. Shroff, "Joint rate control and scheduling in multihop wireless networks," *IEEE CDC*, Atlantis, Bahamas, Dec. 2004.
[3] A. Eryilmaz, R. Srikant, *"Fair resource allocation in wireless networks using queue-length based scheduling and congestion control," IEEE/ACM transactions on networking*, vol. 15, no. 6, pp. 1333-1344, Dec. 2007.
[4] W. Ouyang, S. Murugesan, A. Eryilmaz, N. B. Shroff, "Scheduling with rate adaptation under incomplete knowledge of channel/estimator statistics," in *Allerton Conference,* 2010.
[5] J. Huang, R. A. Berry, and M. L. Honig, "Wireless scheduling with hybrid ARQ", IEEE transactins on wireless communicatins, vol. 4, no. 6, 2005.
[6] W. Ouyang, N. Prasad, S. Rangarajan, "Exploiting hybrid channel information for downlink multi-user mimo scheduling," *IEEE WiOpt 2013*.
[7] W. Ouyang, A. Eryilmaz, N. B. Shroff, "Low-complexity Optimal Scheduling Over Correlated Fading Channels with ARQ Feedback," *IEEE WiOpt 2012*, Paderborn, Germany. ArXiv:1206.1634.
[8] W. Ouyang, S. Murugesan, A. Eryilmaz, N. Shroff, "Exploiting channel memory for joint estimation and scheduling in downlink networks," *IEEE INFOCOM*, Shanghai, China, Apr. 2011.
[9] W. Ouyang, A. Erilmaz, N. B. Shroff, "Asymptotically optimal downlink scheduling over markovian fading channels," *IEEE INFOCOM 2012*.
[10] S.H. Ahmad, M. Liu, T. Javidi, Q. Zhao and B. Krishnamachari, "Optimality of myopic sensing in multi-Channel opportunistic access," *IEEE Trans. on Info. Theory,* 2009.
[11] K. Liu, Q. Zhao, "Indexability of restless bandit problems and optimality of whittle's index for dynamic multichannel access," *IEEE Transactions on Information Theory,* vol. 56, pp. 5547-5567, 2008.
[12] C. Li, M. J. Neely, "Exploiting channel memory for multiuser wireless scheduling without channel measurement: capacity regions and algorithms," *Elsevier Performance Evaluation,* 2011.
[13] C. Li, M. J. Neely, "Network utility maximization over partially observable markovian channels," *Elsevier Performance Evaluation*, 2013.
[14] S. H. Ahmad, M. Liu, T. Javidi, Q. Zhao, B. Krishnamachari, "Optimality of myopic sensing in multi-channel opportunistic access," *IEEE Transactions on Information Theory,* vol. 55, pp. 4040–4050, 2009.
[15] S. Murugesan, P. Schniter, N. B. Shroff, "Multiuser scheduling in a Markov-modeled downlink using randomly delayed ARQ feedback," *IEEE Transactions on Information Theory,* vol. 58, no. 2, 2012.
[16] K. Jagannathan, S. Mannor, I. Menache, E. Modiano, "A state action frequency approach to throughput maximization over uncertain wireless channels," *IEEE INFOCOM*, Shanghai, China, Apr. 2011.
[17] G. Celik, E. Modiano, "Scheduling in networks with time-varying channels and reconfiguration delay," *IEEE INFOCOM*, 2012.
[18] P. Whittle, "Restless bandits: activity allocation in a changing world," *Journal of Applied Probability,* vol. 25, pp. 287-298, 1988.
[19] S. Murugesan, P. Schniter, N. B. Shroff, "Opportunistic Scheduling using ARQ feedback in Multi-Cell Downlink,", Asilomar, 2010.
[20] E. J. Sondik, *"The optimal control of partially observable Markov Decision Processes,"* PhD thesis, Stanford University, 1971.
[21] Eitan Altman, *"Constrained Markov Decision Processes"*, 1999.
[22] J. D. Isom, S. Meyn, R. D. Braatz, "Piecewise linear dynamic programming for constrained POMDPs," *National Conference on Artificial Intelligence*, pp. 291-296, 2008.
[23] W. Ouyang, A. Erilmaz, N. B. Shroff, "Scheduling in Time-correlated Wireless Networks with Imperfect CSI and Stringent Constraint," *Technical Report*, arXiv:1403.7773.