HAL
open science

Low-complexity Optimal Scheduling over Correlated Fading Channels with ARQ Feedback

Wenzhuo Ouyang, Atilla Eryilmaz, Ness B. Shroff

To cite this version:

Wenzhuo Ouyang, Atilla Eryilmaz, Ness B. Shroff. Low-complexity Optimal Scheduling over Correlated Fading Channels with ARQ Feedback. WiOpt'12: Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, May 2012, Paderborn, Germany. pp.270-277. hal-00764151

HAL Id: hal-00764151

https://inria.hal.science/hal-00764151v1

Submitted on 12 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Low-complexity Optimal Scheduling over Correlated Fading Channels with ARQ Feedback

Wenzhuo Ouyang, Atilla Eryilmaz, and Ness B. Shroff

***Abstract*—We investigate the downlink scheduling problem under Markovian ON/OFF fading channels, where the instantaneous channel state information is not directly accessible, but is revealed via ARQ-type feedback. The scheduler can exploit the temporal correlation/channel memory inherent in the Markovian channels to improve network performance. However, designing low-complexity and throughput-optimal algorithms under temporal correlation is a challenging problem. In this paper, we find that under an average number of transmissions constraint, a low-complexity index policy is throughput-optimal. The policy uses Whittle's index value, which was previously used to capture opportunistic scheduling under temporally correlated channels. Our results build on the interesting finding that, under the intricate queue length and channel memory evolutions, the importance of scheduling a user is captured by a simple multiplication of its queue length and Whittle's index value. The proposed queue-weighted index policy has provably low complexity which is significantly lower than existing optimal solutions.**

I. Introduction

In wireless networks with randomly fluctuating channels, intelligently scheduling users is critical for achieving high network efficiency. Under the assumption that the scheduler possesses accurate instantaneous Channel State Information (CSI), maximum-weight scheduling algorithms (e.g., [1]-[3]) are known to be throughput-optimal, i.e., no scheduling policy can ensure system stability for arrival rates that are not supportable by a max-weight scheduler.

In practice, accurate instantaneous CSI is difficult to obtain at the scheduler. Hence, in this work we consider the important scenario where the instantaneous CSI is not directly accessible to the scheduler, but is instead revealed through ARQ-type feedback only *after* each scheduled data transmission.

The time-correlation or channel memory inherent in the fading channels can be exploited by the scheduler for more informed decisions, and hence to obtain large throughput gains (e.g., [4][5]). In this paper, we incorporate the temporal correlation by modeling the fading channels as Markov-modulated ON/OFF processes.

Under imperfect CSI, channel memory, and limited network resources, designing throughput-optimal scheduling schemes is highly challenging. This is because the scheduler needs to optimally balance the intricate 'exploitation-exploration tradeoff', i.e., to decide whether to exploit the channels with more up-to-date CSI, or to explore the channels with outdated CSI. The packets destined to each user are stored in a corresponding data queue before transmission. Due to this temporal correlation and imperfect ARQ-based CSI, to develop throughput-optimal scheduler requires a complex characterization of the interplay between user scheduling, channel memory evolution and queue evolution. Therefore, traditional Lyapunov drift minimization technique do not apply in this context.

Under the aforementioned complications, traditional Dynamic Programming based approaches can be used, but are intractable due to the well-known 'curse of dimensionality'. In related works [5][6], a simple round-robin based scheduling policy is shown to possess the throughput-optimality property. However, such a scheme is only optimal in the regime of a large number of users with *identical* Markovian channel statistics. In [7][8], a throughput-optimal frame-based policy is proposed. This policy relies 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 work, we study throughput-optimal downlink scheduling under imperfect CSI over heterogeneous Markovian fading channels. We assume that each user occupies a dedicated channel, i.e., all users can transmit simultaneously, but the *long-term average* number of transmissions is limited. Such a constraint can be used to limit the energy consumption or interference effect depending on the context. An example to limit the energy consumption is the *green cellular networks* (e.g., [9]-[11]). It is estimated that the cellular base stations consume 4.5 GW of power globally, which corresponds to more than 40 million metric tons of CO_2 emission and over $\$10$ billion electricity bill annually [9][10]. With energy expenditure rising by 15-20% each year, an important objective in green cellular networks design is to reduce the long-run average number of data transmissions to decrease energy consumption [10]. Therefore, it is of great interest to understand the relationship between the achievable throughput region and the constraint on the long-term average number of transmissions. In the meanwhile, restricting the average number of transmissions also helps to reduce interference between concurrent transmissions in the network. Specifically, our contributions are as follows:

- Under the constraint on the average number of transmissions, we propose a low-complexity *throughput-optimal* pol-

Wenzhuo Ouyang and Atilla Eryilmaz are with the Department of ECE, The Ohio State University (e-mails: ouyangw@ece.osu.edu, eryilmaz@ece.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@ece.osu.edu).

This work is partly supported by NSF grants CNS-0721434, CNS-0831919, CNS-0953515, CCF-0916664, DTRA grant HDTRA 1-08-1-0016, Army Research Office MURI Awards W911NF-08-1-0238 and W911NF-07-1-0376.

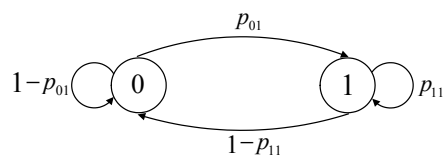

Fig. 1. Two state Markov Chain model.

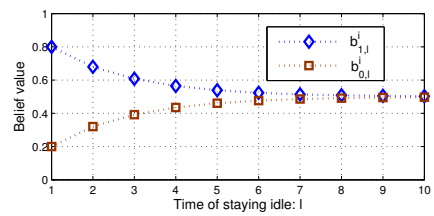

Fig. 2. Belief value evolution.

icy. The policy operates over separate time frames, where the per-frame computational complexity is at most $O(N \log N)$ with the number of users N. Therefore, the policy does not suffer from the curse of dimensionality.

• The proposed policy builds on Whittle's index analysis of Restless Multi-armed Bandit Problem [12], where Whittle's index value is used to measure the importance of scheduling a user under the time-correlated channel [13]. We find that, interestingly, under the coupled queue length and channel memory evolution, the importance of scheduling a user is measured by a simple *multiplication* of the queue length and Whittle's index value.

II. System Model

A. Downlink Scheduling Problem

We consider a time-slotted wireless downlink network with one base station and N users, where each user i occupies a dedicated wireless channel. The channel state of user i, denoted by $C_i[t]$ at slot t, evolves according to an ON/OFF Markov chain across time slots within the state space $\mathcal{S} = \{0, 1\}$, independently across channels. When the channel is in state '1', one packet can be successfully transmitted, otherwise no packet can be delivered [1]. As shown in Fig. 1, the channel state evolution is represented by the transition probabilities

$$p_{11} := \Pr\left(C_i[t]{=}1 \middle| C_i[t{-}1]{=}1\right),$$
$$p_{01} := \Pr\left(C_i[t]{=}1 \middle| C_i[t{-}1]{=}0\right).$$

We assume that the Markovian channels are positively correlated, i.e., $p_{11} > p_{01}$ for $i{=}1, 2, \cdots, N$, which has been commonly used to model the wireless channels in slow fading environment (e.g., [5][13]).

At the beginning of each time slot, the scheduler chooses users for data transmission. The scheduling decisions are made without the exact knowledge of the channel state in the current slot. Instead, the accurate ON/OFF channel state of a scheduled user is revealed via ACK/NACK feedback from the receiver, only at the end of each slot following data transmission.

We consider the class Φ of (possibly non-stationary) scheduling policies that make scheduling decisions based on the history of observed channel states, arrival processes, and scheduling decisions. Under the aforementioned restrictions on average energy consumption, the scheduling schemes are subject to the constraint that the long-term average number of scheduled transmissions is under M,

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N} a_i^{\phi}[t]\Big] \le M, \tag{1}$$

where $a_i^{\phi}[t]$ indicates whether user i is scheduled at slot t under policy $\phi \in \Phi$.

Data packets destined for different users are stored in separate queues before transmission. The queue length for user i is denoted by $q_i[t]$ at slot t. We assume that the packet arrivals for the i-th user form an *i.i.d.* process $A_i[t]$ with mean λ_i and a bounded second moment. Hence, the i-th data queue evolves as $q_i[t{+}1]{=}\max\{0, q_i[t]{-}a_i[t]{\cdot}C_i[t]\}{+}A_i[t]$.

B. Belief Value Evolution

The scheduler maintains a belief value $\pi_i[t]$ for each channel i, defined as 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 accurate channel state feedbacks as follows,

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

where $Q_i(x){=}xp_{11} + (1{-}x)p_{01}$ is the belief evolution operator when user i is not scheduled in the current slot. In our setup, the belief values are known to be sufficient statistics to represent the past scheduling decisions and feedback [14]. In the meanwhile, the belief value $\pi_i[t]$ is the expected throughput for user i if it is scheduled in slot t.

For the i-th user, we use $b^i_{c,l}$ to denote the state of its belief value when the most recent channel state was observed l time slots ago to be $c \in \{0, 1\}$. The closed form expression of $b^i_{c,l}$ can be calculated from (2) and is given as follows,

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

As depicted in Fig. 2, if the scheduler is never informed of the i-th user's channel state, the belief value monotonically converges to the stationary probability $b^i_s{:=}p_{01}/(1+p_{01}-p_{11})$ of the channel being in state 1. We assume that the belief values of all channels are initially set

[1]Our results easily generalize to the scenario where multiple packets, different across channels, can be transmitted in state '1'.

to their stationary values. It is then clear that, based on (2), each belief value $\pi_i[t]$ evolves over a countable state space, denoted by $\mathcal{B}_i=\{b_s^i, b_{c,l}^i : c\in\{0,1\}, l\in\mathbb{Z}^+\}$.

C. Network Stability Region and Achievable Rate Region

We adopt the following definition of queue stability [16]: queue i is stable if there exists a limiting stationary distribution F_i such that $\lim_{t\to\infty} P(q_i[t]\le q)=F_i(q)$. The *network stability region* $\mathbf{\Lambda}$ is defined as the closure of the set of arrival rate vectors supported by all policies in class Φ that does not lead to system instability while abiding by the constraint (1).

In the meanwhile, we define the *achievable rate region* $\mathbf{\Gamma}$ as the closure of the set of service rate vectors $\boldsymbol{\lambda}$ that can be achieved by all policies, i.e.,

$$\mathbf{\Gamma}=Cl\big\{\boldsymbol{\lambda} : \exists\phi\in\Phi \text{ with } \lambda_i=\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\big[\sum_{t=0}^{T-1}\pi_i[t]\cdot a_i^{\phi}[t]\big], \quad i=1,\cdots,N, \text{ subject to constraint (1)}\big\}, \tag{3}$$

where $Cl\{\cdot\}$ denotes the closure of the set. The rate region is hence a convex set, since, by appropriately randomizing between any two policies, all the rate vectors between the corresponding two rate points can be achieved.

The rate region $\mathbf{\Gamma}$ corresponds to the expected throughput that can be achieved in the system with infinitely backlogged queues. Therefore it provides an upper bound on the stability region $\mathbf{\Lambda}$. As we shall see in the following sections, the two regions $\mathbf{\Gamma}$ and $\mathbf{\Lambda}$ turn out to share the same interior and are, therefore, "equal".

III. Optimal Policy for Weighted Sum-throughput Maximization

In this section, we consider a *weighted sum-throughput maximization* problem. The policy introduced here, which is based on scaling the Whittle's index values, not only achieves the transmission rate at the *boundary* of the achievable rate region $\mathbf{\Gamma}$, it also plays an important part in the throughput-optimal policy in the next section that stabilizes all arrival rates within the system stability region $\mathbf{\Lambda}$– the main result of the paper.

A. Weighted Sum-throughput Maximization Problem

Consider the following weighted sum-throughput maximization problem $\Psi(\mathbf{r},M)$ for a given $\mathbf{r}=(r_i)_{i=1}^N$, where the expected service rate for each user i is scaled by a positive factor r_i,

$$V(\mathbf{r},M)=\max_{\phi\in\Phi}\ \liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N} r_i\cdot\pi_i[t]\cdot a_i^{\phi}[t]\Big] \tag{4}$$

$$\text{s.t.}\ \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N} a_i^{\phi}[t]\Big]\le M. \tag{5}$$

The above problem $\Psi(\mathbf{r},M)$ is a constrained Partially Observable Markov Decision Process (CPOMDP) [15]. Consider the optimal policy $\phi^*(\mathbf{r},M)$ (if it exists) for the problem

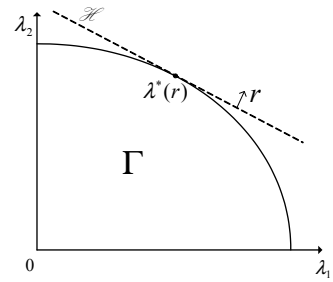

Fig. 3. Illustration of the achieved rate vector $\boldsymbol{\lambda}^*(\mathbf{r})$ under policy $\phi^*(\mathbf{r},M)$ with weight vector $\mathbf{r}$.

$\Psi(\mathbf{r},M)$ and let $\lambda_i^*(\mathbf{r})=\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\big[\sum_{t=0}^{T-1}\pi_i[t]\cdot a_i^{\phi^*(\mathbf{r},M)}[t]\big]$. Then, as illustrated in Fig. 3, the achieved rate vector $\boldsymbol{\lambda}^*(\mathbf{r})$ is at the intersection of the achievable rate region $\mathbf{\Gamma}$ and the supporting hyper-plane $\mathscr{H}$ with normal vector $\mathbf{r}$. We proceed to characterize the optimal policy $\phi^*(\mathbf{r},M)$.

Under uniform weights $\mathbf{r}=\mathbf{1}$, an optimal policy for problem $\Psi(\mathbf{1},M)$ is proposed in [13] based on Whittle's indexability analysis of Restless Multi-armed Bandit Problem [12]. Specifically, for channel i, a closed form *Whittle's index value* $W_i(\pi)$ is assigned to each belief state $\pi\in\mathcal{B}_i$. The index value intelligently captures the exploitation-exploration value to be gained from scheduling the user at the corresponding belief state [13]. Details of Whittle's indexability analysis can be found in [12][13][17]. The closed form expression of the Whittle's index value $W_i(\pi), \pi\in\mathcal{B}_i$ is given as follows [13],

$$W_i(\pi)=\begin{cases}\frac{(b_{0,l}^i-b_{0,l+1}^i)(l+1)+b_{0,l+1}^i}{1-p_{11}+(b_{0,l}^i-b_{0,l+1}^i)l+b_{0,l+1}^i} & \text{if } p_{11}\le\pi=b_{0,l}^i<b_s^i\\ \frac{p_{01}}{(1-p_{11})(1+p_{01}-p_{11})+p_{11}} & \text{if } b_s^i\le\pi\le p_{11}\end{cases} \tag{6}$$

It can be observed from (6) that $W_i(\pi)$ monotonically increases with π and satisfies $W_i(\pi)\in[0,1]$. In the following key lemma, we relate the optimal algorithm developed for problem $\Psi(\mathbf{1},M)$ to the problem $\Psi(\mathbf{r},M)$ with arbitrary weight vector $\mathbf{r}$. The proof of the lemma can be found in Appendix A.

Lemma 1. *There exists an optimal stationary policy $\phi^*(\mathbf{r},M)$ for problem $\Psi(\mathbf{r},M)$ (cf. (4)-(5)), parameterized by a threshold ω^* and a randomization factor $\rho^*\in(0,1]$, such that*

(i) The scheduler maintains an $\mathbf{r}$-weighted index value $W_i^{\mathbf{r}}(\pi_i[t])=r_i\cdot W_i(\pi_i[t])$ for user i.

(ii) User i is scheduled if $W_i^{\mathbf{r}}(\pi_i[t])>\omega^*$, and stays idle if $W_i^{\mathbf{r}}(\pi_i[t])<\omega^*$. If $W_i^{\mathbf{r}}(\pi_i[t])=\omega^*$, it is scheduled with probability ρ^*.

(iii) The parameters ω^* and ρ^* are such that the long-term average number of transmissions equals M.

Remark: Interestingly, by multiplying the Whittle's index values $W_i(\pi_i[t])$ with r_i, the optimal policy $\phi^*(\mathbf{1},M)$ proposed in [13] for problem $\Psi(\mathbf{1},M)$ is extended to solve the

more general problem $\Psi(\mathbf{r}, M)$. This property is important for designing the low-complexity and throughput-optimal policy in Section IV.

B. State Space Truncation

Recall that the belief value evolves over a countable state space $\mathcal{B}_i$ for user i and approaches the stationary value if the channel is not active for a long time. This motivates us to consider a truncated version of the belief value evolution whereby the belief value of a user is set to its steady state (i.e., its channel state history is entirely forgotten) if the corresponding channel has not been scheduled for a long time, say τ slots. The finite space truncation not only facilitates more trackable analysis, it also provides a close approximation to the countable state space. We let $\mathcal{B}_i^\tau$ denote the truncated state space for the i-th user, i.e., $\mathcal{B}_i^\tau = \{b_s^i, b_{c,l}^i : c \in \{0,1\}, l = 1, 2, \cdots, \tau\}$ and let $\mathbf{B}^\tau = [\mathcal{B}_1^\tau, \cdots, \mathcal{B}_N^\tau]$.

In what follows, we introduce the $\mathbf{r}$-weighted index policy $\phi_\tau(\mathbf{r}, M)$ that operates over the truncated state space. We shall prove (in Lemma 2) that, under sufficiently large truncation size, the throughput performance of $\phi_\tau(\mathbf{r}, M)$ is very close to that of $\phi^*(\mathbf{r}, M)$.

$\mathbf{r}$-weighted Index Policy $\phi_\tau(\mathbf{r}, M)$

(1). At time 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 $W^{\mathbf{r}}(\pi_i[t]) < \omega_\tau^*$. If $W^{\mathbf{r}}(\pi_i[t]) = \omega_\tau^*$, user i is scheduled with probability ρ_τ^*.

(2). The parameters ω_τ^* and ρ_τ^*, calculated in the *initialization phase*, are such that the long-term average number of transmissions equals M.

Note that, to implement the policy $\phi_\tau(\mathbf{r}, M)$, the parameters ω_τ^* and ρ_τ^* need to be calculated at the *initialization phase*. We next design the initialization phase based on the observation that the average number of transmissions decreases when either the threshold ω increases or the randomization factor ρ decreases [13]. Hence, during initialization, we first identify the parameter ω_τ^* by increasing the threshold ω until the constraint (1) is satisfied. Then we select the randomization factor ρ_τ^* so that the constraint (1) is strictly satisfied with equality. We let the parameter $\alpha_i(\omega, \rho)$ denote the expected transmission time to user i under a policy with threshold ω and randomization factor ρ. The closed form expression of $\alpha_i(\omega, \rho)$ is derived as follows,

$$\alpha_i(\omega, \rho) = \begin{cases} \frac{\rho(b_{0,h}^i - b_{0,h+1}^i) + 1 - p_{11} + b_{0,h+1}^i}{\rho b_{0,h}^i + (1-\rho) b_{0,h+1}^i + (1-p_{11})(h+1-\rho)} & \text{if } \omega = W_i^{\mathbf{r}}(b_{0,h}^i),\ h < \tau; \\ \frac{\rho(b_{0,h}^i - b_s^i) + 1 - p_{11} + b_s^i}{\rho b_{0,\tau}^i + (1-\rho) b_s^i + (1-p_{11})(\tau+1-\rho)} & \text{if } \omega = W_i^{\mathbf{r}}(b_{0,\tau}^i); \\ \frac{\rho(1 - p_{11} + b_s^i)}{(1+\tau p_{11})(1 - p_{11} + \rho b_s^i)} & \text{if } \omega = W_i^{\mathbf{r}}(b_s^i); \\ 0 & \text{if } \omega > W_i^{\mathbf{r}}(b_s^i). \end{cases} \tag{7}$$

We formally introduce the initialization phase next.

Initialization phase: calculation of ω_τ^* and ρ_τ^*

1. Calculate the $\mathbf{r}$-weighted index value $W_i^{\mathbf{r}}(\pi_i) = r_i \cdot W_i(\pi_i)$ for all $\pi_i \in \mathcal{B}_i^\tau$, $i = 1, \cdots, N$;

2. Sort the $\mathbf{r}$-weighted index values of each belief states of all users to a $(2\tau+1)N$-dimensional vector $\mathbf{w}$ in increasing order. Let $\gamma(k)$ be the user index corresponding to the k-th element w_k of vector $\mathbf{w}$.

3. Let $k=1$ and $\hat{\alpha}_i = 1, i = 1, \cdots, N$.

4. Calculate the activation time $\alpha_{\gamma(k)}(w_k, 1)$ of user $\gamma(k)$ from (7), and update $\hat{\alpha}_{\gamma(k)} = \alpha_{\gamma(k)}(w_k, 1)$.

5. If $\sum_{i=1}^N \hat{\alpha}_i < M$, then $\omega_\tau^* = w_{k-1}$; calculate the randomization factor ρ_τ^* from (7) such that $\sum_{j \neq \gamma(k)} \hat{\alpha}_j + \alpha_{\gamma(k)}(\omega_\tau^*, \rho_\tau^*) = M$; output ω_τ^* and ρ_τ^*. Otherwise, let $k = k+1$, and go to Step 4.

Remark: The computational complexity of the initialization phase is dominated by sorting the index values in the second step, which has complexity $O\big((2\tau+1)N \cdot \log\big((2\tau+1)N\big)\big)$. After initialization, the $\mathbf{r}$-weighted Index Policy $\phi_\tau(\mathbf{r}, M)$ takes a very simple form: in each slot, schedule a user (possibly with randomization) if its $\mathbf{r}$-weighted index value is above a threshold. Therefore, the per-slot computational complexity is $O(N)$.

We let the value τ_0 be

$$\tau_0 = 4 \max\Big\{\frac{1}{-\log(p_{11} - p_{01})}, \frac{1}{\log^2(p_{11} - p_{01})}, i = 1, \cdots, N\Big\}. \tag{8}$$

Let $V_\tau(\mathbf{r}, M)$ be the weighted sum-throughput under policy $\phi_\tau(\mathbf{r}, M)$, i.e.,

$$V_\tau(\mathbf{r}, M) = \liminf_{T \to \infty} \frac{1}{T} \mathbb{E}\Big[\sum_{t=0}^{T-1} \sum_{i=1}^N r_i \cdot \pi_i[t] \cdot a_i^{\phi_\tau(\mathbf{r}, M)}[t]\Big]. \tag{9}$$

The next lemma bounds the throughput performance difference between policies $\phi^*(\mathbf{r}, M)$ and $\phi_\tau(\mathbf{r}, M)$.

Lemma 2. *For $\tau > \tau_0$, the throughput performance difference between the policy $\phi^*(\mathbf{r}, M)$ and $\phi_\tau(\mathbf{r}, M)$ is upper bounded as follows,*

$$|V(\mathbf{r}, M) - V_\tau(\mathbf{r}, M)| \leq f(\tau) \sum_{i=1}^N r_i, \tag{10}$$

where $f(\tau) = \sum_{i=1}^N \alpha_i\big(W_i(b_{0,\tau}^i), 1\big)$, which satisfies $f(\tau) \to 0$ as $\tau \to \infty$.

Proof: We prove this lemma by carefully studying the relationship between the truncation size τ and the achieved transmission rate. For details, please refer to [20]. ■

IV. Frame-based queue-weighted index policy

In this section, we propose a throughput-optimal scheduling policy that operates over the truncated state space. The policy is based on the $\mathbf{r}$-weighted index policy proposed in the last section and has low-complexity.

A. Throughput Optimal Algorithm

We divide the time slots into separate *time frames* of length T, where the k-th frame includes time slots $kT, \ldots, (k+1)T-1$. 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 $\phi_\tau(\mathbf{q}[kT], M)$, developed in the last section, is implemented with the *queue-weighted index values*. Formally, the T-frame queue-weighted index policy $QWI_\tau(T, M)$ is introduced next.

> **T-Frame Queue-Weighted Index Policy** $QWI_\tau(T, M)$
>
> The time slots are divided into frames of length T. Within the k-th frame, the $\mathbf{q}[kT]$-weighted index policy $\phi_\tau(\mathbf{q}[kT], M)$ is implemented for T consecutive slots, over the truncated state space $\mathbf{B}^\tau$.

The next proposition establishes throughput-optimality of the frame-based queue-weighted index policy.

Proposition 1. *For any $\epsilon > 0$, there exist T' and τ' such that, if $T \geq T'$ and $\tau \geq \tau'$, then for any arrival rate $\boldsymbol{\lambda}$ within the achievable rate region $\mathbf{\Gamma} - \epsilon\mathbf{1}$, under the T-frame queue-weighted index policy $QWI_\tau(T, M - \epsilon/2)$: (i) all queues are stable, (ii) the constraint (1) on the average number of transmissions is satisfied.*

Proof: We prove the proposition by first establishing the uniform convergence of the finite horizon throughput performance in a frame to the infinite horizon throughput. We then apply Lemma 1 to show that the average Lypunov drift in each frame is negative, which establishes the throughput-optimality. Details of the proof are given in Appendix B. ■

Remarks: (1) Note that, in Proposition 1, the parameter M in the queue-weighted index policy is scaled down by $\epsilon/2$. This mechanism is needed to guarantee the constraint on the long-term average number of transmission. The details are given in the proof.
(2) In the queue-weighted index policy, a user is scheduled based on its queue-weighted Whittle's index value. This is especially interesting because of the following: a simple multiplication of queue length and Whittle's index value captures the importance of scheduling a user under two sophisticated system features – the queue evolution and the fundamental exploration-exploitation tradeoff.
(3) Calculation of queue-weighted index value is very simple, which only requires scaling the *pre-calculated* Whittle's index value. Under the queue-weighted index policy, in each frame, the initialization phase of $\phi_\tau(\mathbf{q}[kT], M-\epsilon/2)$ has computational complexity $O(N \log N)$, while implementing $\phi_\tau(\mathbf{q}[kT], M-\epsilon/2)$ over the frame has complexity $O(TN)$ (see the remark in Section III-B). Accordingly, the *per-frame* complexity is $O(N \log N + TN)$. Therefore, as the frame length T scales up, the *per-slot* complexity decreases toward $O(N)$.
(4) The scheduling decisions are made by comparing each user's own index value to a threshold, independently with other users. Hence our policy is also applicable for *distributed implementation* in uplink scenarios.

Corollary 1. *The achievable rate region $\mathbf{\Gamma}$, expressed in (3), is equal to the stability region $\mathbf{\Lambda}$.*

Proof: Recall that the achievable rate region $\mathbf{\Gamma}$ provides an upper bound to the stability region $\mathbf{\Lambda}$. Since the previous proposition states that the queue-weighted index policy stabilizes arrival rates arbitrarily close to the boundary of the achievable rate region $\mathbf{\Gamma}$, hence the achievable rate region $\mathbf{\Gamma}$ and the stability region $\mathbf{\Lambda}$ share the same interior. Because both regions $\mathbf{\Gamma}$ and $\mathbf{\Lambda}$ are defined over closure of sets, we have $\mathbf{\Gamma} = \mathbf{\Lambda}$. ■

Proposition 1 requires the state-space truncation size τ to be large enough for throughput-optimality in the region $\mathbf{\Gamma} - \epsilon\mathbf{1}$. We next characterize the relationship between the truncation size τ and the size of the corresponding supportable region, where, recall that, the expression of τ_0 is given in (8).

Proposition 2. *(a). If $\tau \geq \tau_0$, there exist T_0 and function $g(\tau)$ such that, if $T > T_0$, for all arrival rates within the stability region $\mathbf{\Lambda} - g(\tau)\mathbf{1}$, under the T-frame queue-weighted index policy $QWI_\tau(T, M - g(\tau)/2)$, all queues are stable, and constraint (1) on the average number of transmissions is satisfied.*
(b). The function $g(\tau) = 3\sum_{i=1}^{N} \alpha_i \big(W_i(b^i_{0,\tau}), 1\big)$ and satisfies $\lim_{\tau\to\infty} g(\tau) = 0$.

Proof: In the proof, we used Lemma 2 to bound the throughput performance difference between the truncated scenario and the non-truncated case. For details, please refer to Appendix C. ■

Remark: Proposition 2 allows one to upper bound the state-space truncation size τ that ensures the throughput-optimality in any region $\mathbf{\Lambda} - \epsilon\mathbf{1}$, when the frame length T is sufficiently large. We believe that, by implementing the policy with *expanding frame duration*, the dependence on T_0 in Proposition 2 can be removed while preserving the low-complexity.

V. Conclusion

In this work, we have studied downlink scheduling problem over Markovian evolving ON/OFF fading channels and imperfect instantaneous channel state information. The scheduling decisions are made based on the single-bit ARQ-type feedback and the channel memory inherent in the

Markovian channels. We propose a throughput-optimal policy that operates over time frames and appropriately truncated belief state space. In the proposed policy, the importance of scheduling a user is measured by a simple multiplication of the queue length and Whittle's index value. Based on this key observation, we develop an index-based policy that is not only throughput-optimal, but also has low-complexity per frame in the network size and the truncation level of the belief state space. Most notably, our policy does not suffer from the curse of dimensionality that is observed in earlier works in this context. We further identified a closed form relationship between the size of the state space truncation and the achievable throughput region, which is important in the practical implementation of our low-complexity solution.

REFERENCES

[1] L Tassiulas, A Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936-1948, Dec. 1992.

[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 Transaction on Networking*, vol. 15, no. 6, pp. 1333-1344, Dec. 2007.

[4] 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.

[5] C. Li, M. J. Neely, "Exploiting channel memory for multiuser wireless scheduling without channel measurement: capacity regions and algorithms," *Elsevier Performance Evaluation,* vol 68, no. 8, pp. 631-657, Aug. 2011.

[6] C. Li, M. J. Neely, " Network utility maximization over partially observable markovian channels," *IEEE WiOpt,* May 2011.

[7] 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.

[8] G. Celik, E. Modiano, "Scheduling in networks with time-varying channels and reconfiguration delay," *IEEE INFOCOM*, Orlando, FL, Mar. 2012.

[9] H. Bogucka, A. Conti, "Degrees of freedom for energy savings in practical adaptive wireless systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 38-45, 2011.

[10] E. Oh, B. Krishnamachari, X. Liu, Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 56 -61, 2011.

[11] K. Son, H. Kim, Y. Yi, B.Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1525 - 1536, 2011.

[12] P. Whittle, "Restless bandits: activity allocation in a changing world," *Journal of Applied Probability,* vol. 25, pp. 287-298, 1988.

[13] W. Ouyang, A. Erilmaz, N. B. Shroff, "Asymptotically optimal downlink scheduling over markovian fading channels," *IEEE INFOCOM 2012*, Orlando, Frorida (ArXiv Preprint: 1108.3768).

[14] E. J. Sondik, *"The optimal control of partially observable Markov Decision Processes,"* PhD thesis, Stanford University, 1971.

[15] J. D. Isom, S. Meyn, R. D. Braatz, "Piecewise linear dynamic programming for constrained POMDPs," *National Conference on Artificial Intelligence*, pp. 291-296, 2008.

[16] 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.

[17] 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.

[18] L. Georgiadis, M. Neely, L. Tassiulas, *"Resource allocation and cross-Layer control in wireless networks,"* NOW Publishers Inc., 2006

[19] S. Meyn, *"Control Techniques for Complex Networks,"* Cambridge University Press, 2007.

[20] W. Ouyang, A. Eryilmaz, N. B. Shroff, "Low-complexity optimal scheduling over correlated fading channels with ARQ feedback," *Technical Report*, (www.ece.osu.edu/∼ouyangw/wiopt12.pdf).

APPENDIX A
PROOF OF LEMMA 1

The proof of the lemma is an extension of the proof of Proposition 1 in [13]. Consider the problem $\Psi(\mathbf{r}, M)$ with weight vector $\mathbf{r}$. The constraint (1) can be written in an equivalent form that requires at least $N - M$ channels to be *passive* on average, i.e.,

$$\liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N}(1-a_i^{\phi}[t])\Big] \geq N - M. \qquad (11)$$

Associating a Lagrange multiplier ω to the constraint (11), we have the following Lagrangian function $L(\phi, \omega)$ for problem $\Psi(\mathbf{r}, M)$,

$$L(\phi,\omega)=\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N} r_i\cdot\pi_i[t]\cdot a_i^{\phi}[t]\Big]$$
$$+\omega\cdot\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\sum_{i=1}^{N}(1-a_i^{\phi}[t])\Big]-\omega\cdot(N-M). \quad (12)$$

The dual function $D(\omega)$ is defined as $D(\omega) = \max_{\phi\in\Phi} L(\phi, \omega)$. Following the lines of proof in [13] we have

$$D(\omega) = \sum_{i=1}^{N} U_i^{r_i}(\omega) + \omega(N - M).$$

in which $U_i^{r_i}(\omega)$ is a ω-subsidy problem under weight r_i,

$$U_i^{r_i}(\omega) = \max_{\phi\in\Phi_i}\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\big[r_i\cdot\pi_i[t]\cdot a_i^{\phi}[t] + \omega\cdot(1-a_i^{\phi}[t])\big]\Big], \quad (13)$$

where Φ_i denotes the set of scheduling policies that activate and idle the user i according to the observed channel history. In the above problem (13), for each channel i at belief state π_i, it will receive a reward $r_i\pi_i$ when it activates, otherwise it will receive a subsidy ω for passivity. We let $\mathcal{I}_i^{r_i}(\omega) \subseteq \mathcal{B}_i$ be the set of belief states for which it is optimal to stay idle.

Under the unit weight $r_i = 1$, it was shown in [17] that the problem is Whittle indexable, i.e., $\mathcal{I}_i^1(\omega)$ monotonically increases from $\emptyset$ to $\mathcal{B}_i$ as ω increase from 0 to ∞ for each user i. The Whittle's index value $W_i(\pi)$ is defined as the infimum subsidy value for which the belief state π is at the

boundary of $\mathcal{I}_i^1(\omega)$, i.e.,

$$W_i(\pi) = \inf\{\omega : \pi \in \mathcal{I}_i^1(\omega)\}.$$

It follows from [13] that, for the ω-subsidy problem under unit weight $r_i = 1$, the optimal policy is to activate the user at time slot t if $W_i(\pi) > \omega$, and to stay idle if $W_i(\pi) < \omega$, with tie breaking arbitrarily if $W_i(\pi) = \omega$.

We next extend the optimal algorithm for the ω-subsidy problem under unit weight to the general case with arbitrary weight r_i. An equivalent form of $U_i^{r_i}(\omega)$ is as follows,

$$U_i^{r_i}(\omega) = r_i \max_{\phi\in\Phi_i} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\big[\pi_i[t]a_i^{\phi}[t]+\frac{\omega}{r_i}(1-a_i^{\phi}[t])\big]\Big]. \quad (14)$$

Therefore, the optimal solution for the ω-subsidy problem (13) with weight r_i takes the same form as the optimal solution for the ω/r_i-subsidy problem with weight 1. Accordingly, the optimal solution takes the following form: a user i is scheduled at slot t if $W_i(\pi_i[t]) > \omega/r_i$, and stay idle if $W_i(\pi) < \omega/r_i$, with tie breaking arbitrarily if $W_i(\pi) = \omega/r_i$.

We define the $\mathbf{r}$-weighted index value as $W_i^{\mathbf{r}}(\pi) = r_i \cdot W_i(\pi), \pi \in \mathcal{B}_i, i \in \{1, \cdots, N\}$. The optimal policy for the reward maximization problem in (14) is then to activate the user i at time slot t if $W_i^{\mathbf{r}}(\pi) > \omega$, and to stay idle if $W_i^{\mathbf{r}}(\pi) < \omega$, with tie breaking arbitrarily if $W_i^{\mathbf{r}}(\pi) = \omega$. Therefore the dual function value $D(\omega)$ can be achieved by a threshold-based policy implemented over the $\mathbf{r}$-weighted index values $W_i^{\mathbf{r}}(\pi)$. We shall denote the policy as $\phi(\omega, \rho)$ where each user i at the threshold $W_i^{\mathbf{r}}(\pi) = \omega$ is activated with probability ρ.

Following the similar proof techniques of Lemma 11 in [13], by appropriately choosing ω^* and the randomization parameter ρ^* such that the constraint (1) on the average number of transmissions is strictly satisfied with equality, the corresponding policy $\phi(\omega^*, \rho^*)$ is optimal for the problem $\Psi(\mathbf{r}, M)$.

Appendix B
Proof of Proposition 1

Define Lyapunov function $L(\mathbf{q}) = \frac{1}{2}\sum_{i=1}^N q_i^2$. We consider *the T-frame average Lyapunov drift $\Delta L(\mathbf{q}[kT])$ over the k-th frame*, expressed as,

$$\begin{aligned}
&\Delta L(\mathbf{q}[kT])/T \\
=&\frac{1}{T}\mathbb{E}\Big[L(\mathbf{q}[(k+1)T]) - L(\mathbf{q}[kT])\Big|\mathbf{q}[kT], \boldsymbol{\pi}[kT]\Big] \\
\le& B + \sum_{i=1}^N q_i[kT]\cdot\lambda_i - \sum_{i=1}^N q_i[kT]\cdot\frac{1}{T} \\
&\cdot\mathbb{E}\Big[\sum_{t=0}^{T-1}\pi_i[kT+t]\cdot a_i^{\phi_\tau(\mathbf{q}[kT], M-\epsilon/2)}[kT+t]\Big|\boldsymbol{\pi}[kT]\Big], \quad (15)
\end{aligned}$$

where B is a constant whose value is determined by the second moment of the arrival process [18]. Because $\boldsymbol{\lambda}$ lies within the stability region $\boldsymbol{\Gamma} - \epsilon\mathbf{1}$, we have $\boldsymbol{\lambda} + \epsilon\mathbf{1} \in \boldsymbol{\Gamma}$. Therefore, for any vector $\mathbf{q}$,

$$\sum_{i=1}^N q_i \cdot (\lambda_i + \epsilon) \le V(\mathbf{q}, M)$$

where $V(\mathbf{q}, M)$ is defined in (4)-(5). The Lyapunov drift (15) now becomes,

$$\begin{aligned}
\Delta L(\mathbf{q}[kT])/T \le& B - \epsilon\sum_{i=1}^N q_i[kT] + \\
& V(\mathbf{q}[kT], M) - V_\tau^T(\mathbf{q}[kT], M-\epsilon/2), \quad (16)
\end{aligned}$$

where $V_\tau^T(\mathbf{q}[kT], M)$ is the T-horizon expected transmission rate achieved under the policy $\phi_\tau(\mathbf{q}[kT], M)$, i.e.,

$$\begin{aligned}
&V_\tau^T(\mathbf{q}[kT], M) \\
=&\sum_{i=1}^N q_i[kT]\frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1}\pi_i[kT+t]\cdot a_i^{\phi_\tau(\mathbf{q}[kT], M)}[kT+t]\Big|\boldsymbol{\pi}[kT]\Big].
\end{aligned}$$

We denote $Z_\tau^T(\mathbf{q}, M)$ as the finite T-horizon expected number of transmissions, under the policy $\phi_\tau(\mathbf{q}[kT], M)$, i.e.,

$$Z_\tau^T(\mathbf{q}, M) = \sum_{i=1}^N \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1} a_i^{\phi_\tau(\mathbf{q}, M)}[t]\Big].$$

The next lemma states that, as the length of the time horizon tends to infinity, the expected achieved rate in finite horizon asymptotically converges to infinite horizon achievable rate, and the expected number of transmissions converges to the value M.

Lemma 3. *For any M and $\kappa > 0$, we have, uniformly over $\mathbf{q}$, M, and the initial state $\boldsymbol{\pi}[kT]$,*

(a) there exists positive constants c_1 and c_2 such that

$$\Big|V_\tau(\mathbf{q}, M) - V_\tau^T(\mathbf{q}, M)\Big| < \big(\kappa + c_1\exp(-c_2T)\big)\sum_{i=1}^N q_i.$$

(b) there exists positive constants d_1 and d_2 such that

$$\Big|Z_\tau^T(\mathbf{q}, M) - M\Big| < \big(\kappa + d_1\exp(-d_2T)\big).$$

Proof: The proof of this lemma is based on the observation that the belief state of each user evolves as a finite state space, aperiodic Markov Chain within one communicating class. By bounding the mixing time of the belief state evolution, we can obtain the exponential decay rate as in the lemma. Details of the proof are moved to our technical report [20]. ■

The next lemma bounds the difference between the reward function $V_\tau(\mathbf{q}, M-\epsilon)$ and $V_\tau(\mathbf{q}, M)$.

Lemma 4. *When $\tau > \tau_0$, the difference between the expected transmission rate achieved under policy $\phi_\tau(\mathbf{q}, M)$*

and $\phi_\tau(\mathbf{q}, M-\epsilon)$ satisfies the following bound,

$$\left|V_\tau(\mathbf{q}, M) - V_\tau(\mathbf{q}, M-\epsilon)\right| \le \epsilon \sum_{i=1}^{N} q_i.$$

Proof: We use $\nu_i(\omega, \rho)$ to denote the expected throughput contributed by user i under a policy with threshold ω and randomization factor ρ. The closed form expression of $\nu_i(\omega, \rho)$ can be found in [20]. We proceed with the following lemma.

Lemma 5. *For a user i, if $\tau \ge \tau_0$, we have*
(a) fixing ω, both $\alpha_i(\omega, \rho)$ and $\nu_i(\omega, \rho)$ increase with ρ;
(b) for any two parameter pairs (ω_1, ρ_1) and (ω_2, ρ_2),

$$\Big|\, \nu_i(\omega_1, \rho_1) - \nu_i(\omega_2, \rho_2)\Big| \le q_i \cdot \Big|\, \alpha_i(\omega_1, \rho_1) - \alpha_i(\omega_2, \rho_2)\Big|.$$

Proof: The lemma is proven via studying the closed form relationship between $\alpha_i(\omega, \rho)$ and $\nu_i(\omega, \rho)$. Details of the proof are available in [20]. ■

Suppose, under the weight $\boldsymbol{q}$, the policies $\phi_\tau(\boldsymbol{q}, M)$ and $\phi_\tau(\boldsymbol{q}, M-\epsilon)$ correspond to parameter pairs $(\omega_M^\tau, \rho_M^\tau)$ and $(\omega_{M-\epsilon}^\tau, \rho_{M-\epsilon}^\tau)$, respectively.

For user i, we let $y_i(\epsilon)$ denote be the difference between activation time under policy $\phi_\tau(\boldsymbol{q}, M-\epsilon)$ and $\phi_\tau(\boldsymbol{q}, M)$, i.e., $y_i(\epsilon) = \alpha_i(\omega_M^\tau, \rho_M^\tau) - \alpha_i(\omega_{M-\epsilon}^\tau, \rho_{M-\epsilon}^\tau)$. From Lemma 5(a), we have $y_i(\epsilon) \ge 0, \forall i$. Since the difference of the total expected number of transmissions between the two policies is ϵ, we have $\sum_{i=1}^N y_i(\epsilon) = \epsilon$. From Lemma 5(b), we have,

$$\begin{aligned}
&\left|V_\tau(\boldsymbol{q}, M) - V_\tau(\boldsymbol{q}, M-\epsilon)\right| \\
\le& \sum_{i=1}^{N} \left|\nu_i(\omega_M^\tau, \rho_M^\tau) - \nu_i(\omega_{M-\epsilon}^\tau, \rho_{M-\epsilon}^\tau)\right| \\
\le& \sum_{i=1}^{N} q_i \cdot \left|\alpha_i(\omega_M^\tau, \rho_M^\tau) - \alpha_i(\omega_{M-\epsilon}^\tau, \rho_{M-\epsilon}^\tau)\right| \\
=& \sum_{i=1}^{N} q_i \cdot y_i(\epsilon) \\
\le& \sum_{i=1}^{N} q_i \sum_{i=1}^{N} y_i(\epsilon) \\
=& \epsilon \sum_{i=1}^{N} q_i.
\end{aligned}$$

We hence have proved Lemma 4. ■

From Lemma 2-4, the Lyapunov drift (16) can be further bounded as follows,

$$\begin{aligned}
&\Delta L(\mathbf{q}[kT])/T \\
\le& B - \epsilon \sum_{i=1}^{N} q_i[kT] + V(\mathbf{q}[kT], M) - V_\tau(\mathbf{q}[kT], M) \\
&+ V_\tau(\mathbf{q}[kT], M) - V_\tau(\mathbf{q}[kT], M-\epsilon/2) \\
&+ V_\tau(\mathbf{q}[kT], M-\epsilon/2) - V_\tau^T(\mathbf{q}[kT], M-\epsilon/2) \\
\le& B + \Big[-\epsilon + f(\tau) + \epsilon/2 + \big(\kappa + c_1 \exp(-c_2 T)\big)\Big] \cdot \sum_{i=1}^{N} q_i[kT] \\
=& B + \Big[-\epsilon/2 + f(\tau) + \big(\kappa + c_1 \exp(-c_2 T)\big)\Big] \sum_{i=1}^{N} q_i[kT]. \quad (17)
\end{aligned}$$

Since $f(\tau) = \sum_{i=1}^N \alpha_i(b_{0,\tau}^i, 1)$ can get arbitrarily small as τ becomes large, and $c_1 \exp(-c_2 T)$ approaches zero as T scales, and also noting that κ can be arbitrarily small, the Lyapunov drift becomes negative as both $\tau > \tau' > \tau_0$, and T is large enough, e.g., $T > T_1$. From Foster-Lyapunov stability criterion [19], all the queues in the system are hence stable.

Note that, under the queue weighted policy $QWI_\tau(T, M-\epsilon/2)$, the expected number of transmissions in the k-th frame, $Z_\tau^T(\mathbf{q}[kT], M-\epsilon/2)$, is bounded by Lemma 3 as,

$$\Big|Z_\tau^T(\mathbf{q}[kT], M-\epsilon/2) - (M-\epsilon/2)\Big| < \big(\kappa + d_1 \exp(-d_2 T)\big).$$

Therefore, there exists T_2 such that $Z_\tau^T(\mathbf{q}[kT], M-\epsilon/2) < M$ regardless of $\mathbf{q}[kT]$ and $\boldsymbol{\pi}[kT]$. Therefore, the long term constraint on the average number of transmissions is satisfied. Letting $T' = \max\{T_1, T_2\}$, the proposition is established.

Appendix C
Proof of Proposition 2

The proof of the proposition 2 follows the similar lines of the proof for proposition 2. For all arrival rates within the stability region $\boldsymbol{\Lambda} - g(\tau)\mathbf{1}$, under the T-frame queue-weighted index policy $QWI_\tau(T, M-g(\tau)/2)$ with $\tau \ge \tau_0$, we have the following upper bound on the average Lyapunov drift over the k-th frame similar to (17),

$$\begin{aligned}
&\Delta L(\mathbf{q}[kT])/T \\
\le& B + \Big[-g(\tau)/2 + f(\tau) + \big[\kappa + c_1 \exp(-c_2 T)\big]\Big] \sum_{i=1}^{N} q_i[kT] \\
=& B + \Big[-f(\tau)/2 + \big[\kappa + c_1 \exp(-c_2 T)\big]\Big] \sum_{i=1}^{N} q_i[kT],
\end{aligned}$$

where the last equality holds because $g(\tau) = 3f(\tau)$. For fixed τ, by choosing κ sufficiently small and T sufficiently large, the Lyapunov drift is negative. Therefore, the queues are stable according to the Foster-Lyapunov criterion. Also similar to the proof of proposition 2, the long-term constraint on the average number of transmissions is satisfied for sufficiently large T. Letting T_0 be the value of frame length that guarantees the negative Lyapunov drift and also satisfies the constraint, part (a) of the proposition will hold.

Note that we have $g(\tau) = 3f(\tau)$. From Lemma 2, we have $\lim_{\tau\to\infty} g(\tau) = 0$. Therefore, part (b) of the proposition also holds.