Opportunistic Scheduling Revisited Using Restless Bandits: Indexability and Index Policy

Kehao Wang, *Member, IEEE*, Jihong Yu, *Member, IEEE*, Lin Chen, *Member, IEEE*, Pan Zhou, *Member, IEEE*, Xiaohu Ge, *Senior Member, IEEE*, and Moe Z. Win, *Fellow, IEEE*

***Abstract*— We revisit the opportunistic scheduling problem in which a server opportunistically serves multiple classes of users under time-varying multi-state Markovian channels. The aim of the server is to find an optimal policy minimizing the average waiting cost of those users. Mathematically, the problem can be recast to a restless multiarmed bandit one, and a pivot to solve restless bandit by the Whittle index approach is to establish indexability. Despite the theoretical and practical importance of the Whittle index policy, the indexability is still open for opportunistic scheduling in the heterogeneous multi-state channel case. To fill this gap, we mathematically identify a set of sufficient conditions on a channel state transition matrix under which the indexability is guaranteed and consequently, the Whittle index policy is feasible. Furthermore, we obtain the closed-form Whittle index by exploiting the structural property of the channel state transition matrix. For a generic channel state transition matrix, we propose an eigenvalue-arithmetic-mean scheme to obtain the corresponding approximate matrix which satisfies the sufficient conditions, and consequently can get an approximate Whittle index. This paper constitutes a small step toward solving the opportunistic scheduling problem in its generic form involving multi-state Markovian channels and multi-class users.**

***Index Terms*— Restless bandit, indexability, stochastic scheduling, performance evaluation.**

I. Introduction

We revisit the following opportunistic scheduling system involving a base station, also referred to as a server, different classes of users with heterogeneous demands, and time-varying multi-state Markovian channels. Each channel with different states and classes has a different transmission rate, i.e., the evolution of channels is Markovian and class-dependent. For those users connected to (or entering) the system but not served immediately, their waiting costs increase with time. In such an opportunistic scheduling scenario, a central problem is how to exploit the server's capacity to serve the users. This problem can be formalized to the problem of designing an optimum opportunistic scheduling policy to minimize the average waiting cost.

The above opportunistic scheduling problem is fundamental in many classical and emerging wireless communication systems such as mobile cellular systems including 4G LTE and the emerging 5G, heterogeneous networks (HetNet).

A. State of the Art

Due to its fundamental importance, the opportunistic scheduling problem has attracted a large body of research on channel-aware schedulers addressing one or more system performance metrics in terms of throughput, fairness, and stability [1]–[14], [17]–[26].

The seminal work in [2] showed that the system capacity can be improved by opportunistically serving users with maximal transmission rate. Such a scheduler is called $c\mu$-rule or *MaxRate* scheduler. In fact, the *MaxRate* scheduler was myopically throughput-optimal, i.e., maximizing the current slot transmission rate but ignoring the impact of the current scheduling on the future throughput, and consequently, was shown to perform badly in system stability from the long-term viewpoint. For instance, the number of waiting users in the system explodes with the increase of system load. Meanwhile, the *MaxRate* scheduler does not fairly schedule those users with lower transmission rates.

To balance system throughput and fairness, the Proportionally Fair (PF) scheduler was proposed and implemented in CDMA 1xEV-DO system of 3G cellular networks [3]. Technically, the PF scheduler maximizes the logarithmic throughput of the system rather than traditional throughput, and as a result, provides better fairness [4]. In [5], the authors approximated the PF by the relatively best (RB) scheduler, and analyzed the flow-level stability of the PF scheduler. Actually, the RB scheduler gives priority to users according to their ratio of the current transmission rate to the mean transmission rate. Accordingly, it is fair to users by taking future evolution into account. Consequently, it can provide a minimal throughput to the users with low accessible transmission rates, at the price of being not maximally stable at flow-level [6].

Manuscript received January 27, 2019; accepted July 11, 2019. Date of publication August 5, 2019; date of current version October 9, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61672395 and Grant 61911540481. The work of J. Yu was supported in part by the Beijing Institute of Technology Research Fund Program for Young Scholars, in part by the Young Elite Scientist Sponsorship Program of China, and in part by the Chongqing Key Laboratory of Mobile Communications Technology. This article was presented in part at the 2017 IEEE Global Communications Conference (GLOBECOM). *(Corresponding author: Jihong Yu.)*

K. Wang is with the Hubei Key Laboratory of Broadband Wireless Communication and Sensor Networks, Wuhan University of Technology, Wuhan 430070, China (e-mail: kehao.wang@whut.edu.cn).

J. Yu is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: jihong.yu@bit.edu.cn).

L. Chen is with the School of Data Computer Science, Sun Yat-sen University, Guangzhou 510275, China (e-mail: chen@lri.fr).

P. Zhou and X. Ge are with the School of Electrical and Information Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: panzhou@hust.edu.cn; xhge@mail.hust.edu.cn).

M. Z. Win is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: moewin@mit.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2019.2931690

Other schedulers, e.g., *score based (SB)* [7], *proportionally best (PB)*, and *potential improvement (PI)*, belong to the family of the best-condition schedulers. These schedulers give priority to users according to their respective evaluated channel condition and, accordingly, do not have a direct association with transmission rate. They are not myopically throughput-optimal, but rather have a good performance in the long term. They are maximally stable [8], [9], but they do not consider fairness.

The above papers all assume independent and identically distributed (i.i.d.) channels. For the more challenging scenarios, there exist some work on homogeneous channels [10], [11], [18], i.e., i.i.d. in slots, and heterogeneous channels [12]–[16], i.e., discrete-time Markov process in slots. Under the Markovian channel model, the opportunistic scheduling problem can be mathematically recast to a restless multiarmed bandit (RMAB) [17]. The RMAB is of fundamental importance in stochastic decision theory due to its generic nature, and has application in numerous engineering problems. The central problem in investigating and solving an instance of RMAB is to establish its *indexability*. Once the indexability is established, an index policy can be constructed by assigning an index for each state of each arm to measure the reward of activating the arm at that state. The policy thus consists of simply activating those arms with the largest indices.

In the context of opportunistic scheduling, the authors [18] considered a flow-level scheduling problem with time-homogeneous channel state transition where the probability of being in a state is fixed for any time slot, regardless of the evolution of the system. For the same channel model, the authors [10] considered the opportunistic scheduling problem under the assumption of traffic size following the Pascal distribution. In [10], [11], [18], the indexability was first proved and then the similar closed-form Whittle index was obtained [17]. For heterogeneous channels, the authors of [12]–[14] considered a generic flow-level scheduling problem with heterogeneous channel state transition, but carried out their work based on a conjecture that the problem is indexable. As a result, they can only verify the indexability of the proposed policy for some specific scenarios by numerical test before computing the policy index. *The indexability of the opportunistic scheduling for the heterogeneous multi-state Markovian channels, despite its theoretical and practical importance, is still open today.*

B. Main Results and Contributions

To bridge the above theoretical gap, we first carry out a deep investigation into the indexability of the heterogeneous channel case formulated in [12]–[14] and mathematically, identify a set of sufficient conditions on the channel state transition matrix under which the indexability is guaranteed and consequently the Whittle index policy is feasible. Second, by exploiting the structural property of the channel state transition matrix, we obtain the closed-form Whittle index. Third, for a generic channel state transition matrix not satisfying the sufficient conditions, we propose an eigenvalue-arithmetic-mean scheme to approximate this matrix such that the approximate matrix satisfies the sufficient conditions and further the approximate Whittle index is easily obtained. Finally, we present a scheduling algorithm based on the Whittle index, and conduct extensive numerical experiments which demonstrate that the proposed scheduling algorithm can efficiently balance waiting cost and stability.

Our work thus consitutes a small step towards solving the opportunistic scheduling problem in its generic form involving multi-state Markovian channels. As a desirable feature, the indexability conditions established in this work only depend on channel state transition matrix without imposing constraints on those user-dependent parameters such as service rate and waiting cost.

The rest of the paper is organized as follows: Section II presents our system model. Section III formulates the proposed model in Markov decision process language. Section IV studies the index policy, proves its indexability, and gives the closed-form Whittle index. Section V extends the indexability and supplies a simple scheduler. Then a numerical evaluation is presented in Section VI. Finally, Section VII presents the paper's conclusion.

Notation: $\boldsymbol{e}_i$ denotes an N-dimensional column vector with 1 in the i-th element and 0 in others. $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix. $\mathbf{1}_N$ denotes an N-dimensional column vector with 1 in all elements. $\mathbf{0}_N$ denotes an N-dimensional column vector with 0 in all elements. $\mathbf{1}_N^k$ denotes the N-dimensional column vector with 1 in the first k elements and 0 in the remaining $N-k$ elements. $\operatorname{diag}(a_1,\ldots,a_K)$ denotes a diagonal and a block-diagonal matrix with $a_1,\ldots,a_K$. $\operatorname{trace}(\cdot)$ denotes the sum of all elements in a diagonal of a matrix. $(\cdot)^{\mathsf{T}}$ represents the transpose of a matrix or a vector. $(\cdot)^{-1}$ represents the inverse of a matrix.

II. System Model

As mentioned in the introduction, we consider a wireless communication system where a server schedules jobs of heterogeneous users. The system operates in a time-slotted fashion where τ denotes the slot duration and $t \in \mathcal{T} := \{0, 1, \cdots\}$ denotes the slot index.

A. Job, Channel, and User Models

Suppose that there are K classes of users, $k \in \mathcal{K} := \{1, 2, \cdots, K\}$. Each user of class k is uniquely associated with a job of class k which is requested from the server and with a dedicated wireless channel of class k through which the job would be transmitted.

Job sizes: The job (or flow) size b_k of users of class k in bits is geometrically distributed with mean $\mathbb{E}\{\mathsf{b}_k\} < \infty$ for class $k \in \mathcal{K}$.

Channel condition: For each user, the channel condition varies from slot to slot, independently of all other users. For each class k user, the set of discretized channel conditions is denoted by the finite set $\mathcal{N}_k' := \{1, 2, \cdots, N_k\}$.

Channel condition evolution: We assume that at each slot, the channel condition of each user in the system evolves according to a class-dependent Markov chain. Thus, for each user of class $k \in \mathcal{K}$, we can define a Markov chain with state space $\mathcal{N}_k'$. We further define $q_{k,n,m} := \mathbb{P}(Z_k(t+1) = m | Z_k(t) = n)$, where $Z_k(t)$ denotes the channel condition of a class k user at time t. The class k channel condition transition probability matrix is thus defined as:

$$\boldsymbol{Q}^{(k)} := \begin{bmatrix} q_{k,1,1} & q_{k,1,2} & \cdots & q_{k,1,N_k} \\ q_{k,2,1} & q_{k,2,2} & \cdots & q_{k,2,N_k} \\ \vdots & \vdots & \ddots & \vdots \\ q_{k,N_k,1} & q_{k,N_k,2} & \cdots & q_{k,N_k,N_k} \end{bmatrix},$$

where $\sum_{m\in\mathcal{N}'_k} q_{k,n,m} = 1$ for every $n \in \mathcal{N}'_k$.

Transmission rates: When a usr of class k is in channel condition $n \in \mathcal{N}'_k$, he can receive data at transmission rate $s_{k,n}$, i.e., his job is served at rate $s_{k,n}$. We assume that the higher the label of the channel condition, the higher the transmission rate, i.e., $0 \le s_{k,1} < s_{k,2} < \cdots < s_{k,N_k}$.

Waiting costs: For every user of class k, the system operator accrues waiting cost c_k ($c_k > 0$) at the end of every slot if its job is uncompleted.

B. Server Model

The server is assumed to have *full knowledge* of the system parameters. We investigate the case where the server can serve one user each slot. However, our analysis can be straightforwardly generalized to the case where multiple users can be served each slot. At the beginning of every slot, the server observes the actual channel conditions of all users in the system and decides which user to serve during the slot. We assume that the server is preemptive, i.e, at any time it can interrupt the service of a user whose job is not yet completed. For those jobs not completed, they will be saved and served in the future. The server is also allowed to stay idle, and note that it is not work-conserving because of the time-varying transmission rate. We denote by $\mu_{k,n} :\approx \tau s_{k,n}/\mathbb{E}_0\{\mathsf{b}_k\}$ [12] the departure probability that the job is completed within the current time slot when the server serves a user of class k in channel condition $n \in \mathcal{N}'_k$. Note that the departure probabilities are increasing in the channel condition, i.e., $0 \le \mu_{k,1} < \cdots < \mu_{k,N_k} \le 1$, because the transmission rates satisfy $0 \le s_{k,1} < s_{k,2} < \cdots < s_{k,N_k}$.

C. Opportunistic Scheduling Problem

In the above opportunistic scheduling model, a central problem is how to maximally exploit the server's capacity to serve users. This problem can be formalized to the problem of designing an optimum opportunistic scheduling policy to minimize the average waiting cost.

III. Restless Bandit Formulation and Analysis

In this section we analyze the scheduling problem by the approach of RMAB. For the ease of analysis, we investigate the discounted waiting costs by introducing a discount factor $0 \le \beta < 1$. Basically, the time-average case is a special case where $\beta \to 1$.

A. Job-Channel-User Bandit

We denote by $\mathcal{A}_k := \{0,1\}$ the action space of user of class k where action 1 means serving the user and 0 not serving him.

Every job-channel-user couple of class k is characterized by the tuple $\big(\mathcal{N}_k, (\boldsymbol{w}^a_k)_{a\in\mathcal{A}_k}, (\boldsymbol{r}^a_k)_{a\in\mathcal{A}_k}, (\boldsymbol{P}^a_k)_{a\in\mathcal{A}_k}\big)$, where

(1) $\mathcal{N}_k := \{0\} \cup \mathcal{N}'_k$ is the user state space, where state 0 indicates that the job is completed, and state $n \in \mathcal{N}'_k$ indicates that the current channel condition is n and the job is uncompleted;

(2) $\boldsymbol{w}^a_k := (w^a_{k,n})_{n\in\mathcal{N}_k}$, where $w^a_{k,n}$ is the expected one-slot capacity consumption, or work required by a user at state n if action a is chosen. Specifically, for every state $n \in \mathcal{N}_k$, $w^1_{k,n} = 1$ and $w^0_{k,n} = 0$;

(3) $\boldsymbol{r}^a_k := (r^a_{k,n})_{n\in\mathcal{N}_k}$, where $r^a_{k,n}$ is the expected one-slot reward earned by a user at state n if action a is selected. Specifically, for every state $n \in \mathcal{N}'_k$, it is the negative of the expected waiting cost, $r^a_{k,0} = 0$, $r^1_{k,n} = -\bar{\mu}_{k,n} c_k$ where $\bar{\mu}_{k,n} = 1 - \mu_{k,n}$, and $r^0_{k,n} = -c_k$.

(4) $\boldsymbol{P}^a_k := (p^a_{k,n,m})_{n,m\in\mathcal{N}_k}$, where $p^a_{k,n,m}$ is the probability for a user evolving from state n to state m if action a is selected. The one-slot transition probability matrices for action 0 and 1 are as below:

$$\boldsymbol{P}^0_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & q_{k,1,1} & \cdots & q_{k,1,N_k} \\ 0 & q_{k,2,1} & \cdots & q_{k,2,N_k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & q_{k,N_k,1} & \cdots & q_{k,N_k,N_k} \end{bmatrix},$$

$$\boldsymbol{P}^1_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \mu_{k,1} & \bar{\mu}_{k,1} q_{k,1,1} & \cdots & \bar{\mu}_{k,1} q_{k,1,N_k} \\ \mu_{k,2} & \bar{\mu}_{k,2} q_{k,2,1} & \cdots & \bar{\mu}_{k,2} q_{k,2,N_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{k,N_k} & \bar{\mu}_{k,N_k} q_{k,N_k,1} & \cdots & \bar{\mu}_{k,N_k} q_{k,N_k,N_k} \end{bmatrix}.$$

The dynamics of user j of class k is captured by the state process $\mathsf{x}_k(\cdot)$ and the action process $\mathsf{a}_j(\cdot)$, which correspond to state $\mathsf{x}_j(t) \in \mathcal{N}_k$ and action $\mathsf{a}_j(t) \in \mathcal{A}_k$ at any slot t.

B. Restless Bandit Formulation and Opportunistic Scheduling

Let $\Pi^t_{\mathsf{x},\mathsf{a}}$ denote the set of all the policies composed of actions $\mathsf{a}(0), \mathsf{a}(1) \cdots, \mathsf{a}(t)$, where $\mathsf{a}(t)$ is determined by the state history $\mathsf{x}(0), \mathsf{x}(1), \cdots, \mathsf{x}(t)$ and the action history $\mathsf{a}(0), \mathsf{a}(1) \cdots, \mathsf{a}(t-1)$, i.e.,

$$\begin{aligned} \Pi^t_{\mathsf{x},\mathsf{a}} :&= \Big\{\mathsf{a}(i) \Big| \mathsf{a}(i) = \phi\big(\mathsf{x}^{0:i}, \mathsf{a}^{0:i-1}\big), i = 0, 1 \cdots, t\Big\} \\ &\overset{(e)}{=} \Big\{\mathsf{a}(i) \Big| \mathsf{a}(i) = \phi\big(\mathsf{x}(i)\big), i = 0, 1 \cdots, t\Big\}, \end{aligned}$$

where ϕ is a mapping $\phi : (\mathsf{x}^{0:i}, \mathsf{a}^{0:i-1}) \mapsto \mathsf{a}(i)$, $\mathsf{x}^{0:i} := (\mathsf{x}(0), \cdots, \mathsf{x}(i))$ and $\mathsf{a}^{0:i-1} := (\mathsf{a}(0), \cdots, \mathsf{a}(i-1))$, and (e) is due to the Markovian feature.

Let $\Pi^t_{\mathbf{x},\mathbf{a}}$ denote the space of randomized and non-anticipative policies depending on the joint state process $\mathbf{x} := (\mathsf{x}_k(\cdot))_{k\in\mathcal{K}}$ and the joint action process $\mathbf{a} := (\mathsf{a}_k(\cdot))_{k\in\mathcal{K}}$, i.e., $\Pi^t_{\mathbf{x},\mathbf{a}} = \bigcup_{k\in\mathcal{K}} \Pi^t_{\mathsf{x}_k,\mathsf{a}_k}$ is the joint policy space.

Let $\mathbb{E}^\pi_\tau$ denote the expectation over the future states $\mathsf{x}(\cdot)$ and the action process $\mathsf{a}(\cdot)$, conditioned on past states $\mathsf{x}(0)$, $\mathsf{x}(1)$, $\cdots$, $\mathsf{x}(\tau)$ and the policy $\pi \in \Pi^\tau_{\mathsf{x},\mathsf{a}}$.

Consider any expected one-slot quantity $G^{\mathsf{a}(t)}_{\mathsf{x}(t)}$ that depends on state $\mathsf{x}(t)$ and action $\mathsf{a}(t)$ at any time slot t. For any policy

$\pi \in \Pi_{\mathsf{x},\mathsf{a}}^{\infty}$ and any discount factor $0 \le \beta < 1$, we define the infinite horizon β-average quantity as

$$\mathbb{B}_0^{\pi}\left\{G_{\mathsf{x}(\cdot)}^{\mathsf{a}(\cdot)}, \beta, \infty\right\} := \lim_{T\to\infty} \frac{\sum_{t=0}^{T-1} \beta^t \mathbb{E}_t^{\pi}\left\{G_{\mathsf{x}(t)}^{\mathsf{a}(t)}\right\}}{\sum_{t=0}^{T-1} \beta^t}. \quad (1)$$

In the following we consider the discount factor β to be fixed and the horizon to be infinite; therefore we omit them in $\mathbb{B}_0^{\pi}\left\{G_{\mathsf{x}(\cdot)}^{\mathsf{a}(\cdot)}, \beta, \infty\right\}$ and write briefly $\mathbb{B}_0^{\pi}\left\{G_{\mathsf{x}(\cdot)}^{\mathsf{a}(\cdot)}\right\}$. The reason for introducing $\mathbb{B}_0^{\pi}\{\cdot\}$ is that this form can smoothly transit to the average case $\beta = 1$. Henceforth, we always suppose $0 \le \beta < 1$ except when explicitly emphasizing $\beta = 1$.

We are now ready to formulate the opportunistic scheduling problem faced by the server as below.

Problem 1 (Optimum Opportunistic Scheduling): For any discount factor β, the optimum opportunistic scheduling problem is to find a joint policy $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_K) \in \Pi_{\mathbf{x},\mathbf{a}}^{\infty}$ maximizing the total discounted reward (i.e., minimizing the total discounted cost), mathematically defined as below.

$$\text{(P):} \quad \max_{\boldsymbol{\pi}\in\Pi_{\mathbf{x},\mathbf{a}}} \mathbb{B}_0^{\boldsymbol{\pi}}\left\{\sum_{k\in\mathcal{K}} r_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)}\right\} \quad (2)$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{K}} \mathsf{a}_k(t) = 1, \quad t = 0, 1, \cdots. \quad (3)$$

The constraints (3) of problem (P) can be relaxed to the following

$$\mathbb{E}_t^{\boldsymbol{\pi}}\left\{\sum_{k\in\mathcal{K}} \mathsf{a}_k(t)\right\} = 1, \quad t = 0, 1, \cdots$$

$$\Rightarrow \lim_{T\to\infty} \frac{\sum_{t=0}^{T-1} \beta^t \mathbb{E}_t^{\boldsymbol{\pi}}\left\{\sum_{k\in\mathcal{K}} w_{k,\mathsf{x}(t)}^{\mathsf{a}_k(t)}\right\}}{\sum_{t=0}^{T-1} \beta^t} = 1$$

$$\Leftrightarrow \mathbb{B}_0^{\boldsymbol{\pi}}\left\{\sum_{k\in\mathcal{K}} w_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)}\right\} = 1 \quad (4)$$

Using Lagrangian method, we obtain the following by combining (2) and (4),

$$\max_{\boldsymbol{\pi}\in\Pi_{\mathbf{x},\mathbf{a}}} \mathbb{B}_0^{\boldsymbol{\pi}}\left\{\sum_{k\in\mathcal{K}} r_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)}\right\} - \nu\mathbb{B}_0^{\boldsymbol{\pi}}\left\{\sum_{k\in\mathcal{K}} w_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)}\right\}$$
$$= \sum_{k\in\mathcal{K}} \left(\max_{\pi_k\in\Pi_{\mathsf{x}_k,\mathsf{a}_k}} \mathbb{B}_0^{\pi_k}\left\{r_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)} - \nu w_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)}\right\}\right). \quad (5)$$

Thus, we have the subproblem for class $k \in \mathcal{K}$:

$$\text{(SP):} \quad \max_{\pi_k\in\Pi_{\mathsf{x}_k,\mathsf{a}_k}} \mathbb{B}_0^{\pi_k}\left\{r_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)} - \nu w_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)}\right\}. \quad (6)$$

Hence, our goal is to find an optimal policy π_k^* for the subproblem k ($k \in \mathcal{K}$) and then construct a feasible joint policy $\boldsymbol{\pi} = (\pi_1^*, \cdots, \pi_K^*)$ for the problem (P). In the following, we focus on the subproblem (SP) and drop the subscript k.

IV. Indexability Analysis and Index Computation

In this section, we first give a set of conditions on the channel state transition matrix, and, based on which, we obtain the threshold structure of the optimal scheduling strategy for the subproblem. We then establish the indexability under the proposed conditions.

A. Transition Matrices and Threshold Structure

Condition 1: Transition matrix $\boldsymbol{Q}$ can be written as

$$\boldsymbol{Q} = \boldsymbol{O}_0 + \epsilon_1\boldsymbol{O}_1 + \epsilon_2\boldsymbol{O}_2 + \cdots + \epsilon_{2N-2}\boldsymbol{O}_{2N-2},$$

where $\boldsymbol{h} := [h_1, h_2, \cdots, h_N]^{\mathsf{T}}$, $\boldsymbol{O}_j$ is defined in (8) (shown at the top of the next page) and ϵ_j and λ are real numbers satisfying

$$\lambda_j := \lambda - \epsilon_{N-j} - \epsilon_{N-1+j} \le 0, \quad 1 \le j < N. \quad (7)$$

Remark 1: Basically, Condition 1 implies that

- Any two adjacent rows (i.e., $\boldsymbol{Q}_i, \boldsymbol{Q}_{i+1}$) of matrix $\boldsymbol{Q}$ differ in only two adjacent positions (i.e, $i, i+1$). For example, if $N = 3$, $\boldsymbol{Q}$ is written as

$$\boldsymbol{Q} = \begin{bmatrix} h_1 - \epsilon_2 + \lambda & h_2 - \epsilon_1 + \epsilon_2 & h_3 + \epsilon_1 \\ h_1 + \epsilon_3 & h_2 - \epsilon_1 - \epsilon_3 + \lambda & h_3 + \epsilon_1 \\ h_1 + \epsilon_3 & h_2 - \epsilon_3 + \epsilon_4 & h_3 - \epsilon_4 + \lambda \end{bmatrix}.$$

- When $\lambda_j = 0$ for all j ($1 \le j < N$), the $\boldsymbol{Q}$ degenerates into the case of [18].

Now, we give the following lemma on the threshold structure of the optimum scheduling policy for the subproblem.

Lemma 1 (Threshold structure): Under Condition 1, for every real-valued ν, there exists $n \in \mathcal{N} \cup \{-1\}$ such that the optimum scheduling policy only schedules transmission in channel states $\delta_{N-n} := \{m \in \mathcal{N} : m > n\}$.

Proof: Please see Appendix I. □

B. Indexability Analysis

For $\pi_k \in \Pi_{\mathsf{x}_k,\mathsf{a}_k}$, we introduce the concept of serving set, δ ($\delta \subseteq \mathcal{N}_k$), such that the user is served if $n \in \delta$ and not served if $n \notin \delta$. By slightly introducing ambiguity, δ can also be regarded as a policy of serving the set δ.

Thus, the subproblem (6) can be transformed to

$$\max_{\delta\in\mathcal{N}_k} \mathbb{B}_0^{\delta}\left\{r_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)} - \nu w_{k,\mathsf{x}_k(\cdot)}^{\mathsf{a}_k(\cdot)}\right\}. \quad (9)$$

For further analysis, we define

$$\mathbb{R}_n^{\delta} := \frac{\mathbb{B}_0^{\delta}\left\{r_{k,\mathsf{x}_k(\cdot)}^{n,\mathsf{a}_k(\cdot)}\right\}}{1-\beta}, \quad (10)$$

$$\mathbb{W}_n^{\delta} := \frac{\mathbb{B}_0^{\delta}\left\{w_{k,\mathsf{x}_k(\cdot)}^{n,\mathsf{a}_k(\cdot)}\right\}}{1-\beta}, \quad (11)$$

where n refers to the initial state of user of class k.

By Lemma 1, if there exists price ν_n for $n \in \mathcal{N}'$ such that both transmitting and not transmitting are optimal for $\nu = \nu_n$, then there exists a set, δ^*, such that both including state n in δ^* and not including state n lead to the same reward, i.e.,

$$\mathbb{R}_n^{\delta^*\cup\{n\}} - \nu_n\mathbb{W}_n^{\delta^*\cup\{n\}} = \mathbb{R}_n^{\delta^*\setminus\{n\}} - \nu_n\mathbb{W}_n^{\delta^*\setminus\{n\}}. \quad (12)$$

A straightforward consequence is that changing the action only in the initial period must also lead to the same reward, i.e.,

$$\mathbb{R}_n^{\langle 0,\delta^*\rangle} - \nu_n\mathbb{W}_n^{\langle 0,\delta^*\rangle} = \mathbb{R}_n^{\langle 1,\delta^*\rangle} - \nu_n\mathbb{W}_n^{\langle 1,\delta^*\rangle}, \quad (13)$$

$$\boldsymbol{O}_j := \begin{cases} \mathbf{1}_N\big(\boldsymbol{h}\big)^\mathsf{T} + \lambda \boldsymbol{I}_N, & \text{if } j=0,\\ [\underbrace{\mathbf{0}_N,\cdots,\mathbf{0}_N}_{N-j-1}, -\mathbf{1}_N^{N-j}, \mathbf{1}_N^{N-j}, \underbrace{\mathbf{0}_N,\cdots,\mathbf{0}_N}_{j-1}], & \text{if } 1\le j\le N-1,\\ [\underbrace{\mathbf{0}_N,\cdots,\mathbf{0}_N}_{j-N}, \mathbf{1}_N - \mathbf{1}_N^{j-N+1}, \mathbf{1}_N^{j-N+1} - \mathbf{1}_N, \underbrace{\mathbf{0}_N,\cdots,\mathbf{0}_N}_{2N-2-j}], & \text{if } N\le j\le 2N-2. \end{cases} \tag{8}$$

where $\langle a, \delta^* \rangle$ is the policy that employs action a in the initial period and then proceeds according to δ^*.

Then, if $\mathbb{W}_n^{\langle 1,\delta^*\rangle} - \mathbb{W}_n^{\langle 0,\delta^*\rangle} \neq 0$, we have

$$\nu_n = \frac{\mathbb{R}_n^{\langle 1,\delta^*\rangle} - \mathbb{R}_n^{\langle 0,\delta^*\rangle}}{\mathbb{W}_n^{\langle 1,\delta^*\rangle} - \mathbb{W}_n^{\langle 0,\delta^*\rangle}}. \tag{14}$$

We further define

$$\nu_n^\delta := \frac{\mathbb{R}_n^{\langle 1,\delta\rangle} - \mathbb{R}_n^{\langle 0,\delta\rangle}}{\mathbb{W}_n^{\langle 1,\delta\rangle} - \mathbb{W}_n^{\langle 0,\delta\rangle}}. \tag{15}$$

To circumvent the long proof of Whittle indexability, we establish the indexability result by checking the LP-indexability condition [27]. If a problem is LP-indexable, then it is Whittle-indexable. In the following analysis, we show that our problem is LP-indexable; that is, the problem is Whittle-indexable.

Definition 1 ([27]): Problem (6) is LP-indexable with price

$$\nu_n = \nu_n^{\delta_{N-n}} = \frac{\mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle} - \mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle}}{\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle} - \mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}}, \tag{16}$$

if the following conditions hold:

(1) $\forall\, n \in \mathcal{N}$, $\mathbb{W}_n^{\langle 1,\emptyset\rangle} - \mathbb{W}_n^{\langle 0,\emptyset\rangle} > 0$, $\mathbb{W}_n^{\langle 1,\mathcal{N}\rangle} - \mathbb{W}_n^{\langle 0,\mathcal{N}\rangle} > 0$;
(2) $\forall\, n \in \mathcal{N}\setminus\{N\}$, $\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle} - \mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle} > 0$ and $\mathbb{W}_{n+1}^{\langle 1,\delta_{N-n}\rangle} - \mathbb{W}_{n+1}^{\langle 0,\delta_{N-n}\rangle} > 0$;
(3) For each real value ν there exists $n \in \mathcal{N}\cup\{-1\}$ such that the serving set δ_{N-n} is optimal.

To check the LP-indexability, we first characterize the four critical quantities in (16) under δ_{N-n} for any n.

Based on balance equations, when n is not chosen in the initial slot, we have (17) in the matrix language (see the top of the next page) and, further, the following simplified form

$$(\boldsymbol{I}_N - \beta\boldsymbol{M}_0)\cdot \boldsymbol{r}_0 = \boldsymbol{c}_0, \tag{19}$$

where,

$$\begin{aligned} \boldsymbol{M}_0 &= [\boldsymbol{Q}_1^\mathsf{T},\cdots,\boldsymbol{Q}_n^\mathsf{T},\quad \boldsymbol{Q}_{n+1}^\mathsf{T}\bar{\mu}_{n+1},\cdots,\boldsymbol{Q}_N^\mathsf{T}\bar{\mu}_N]^\mathsf{T},\\ \boldsymbol{c}_0 &= [-c,\cdots,-c,\ -c,\quad -c\bar{\mu}_{n+1},\cdots,-c\bar{\mu}_N]^\mathsf{T},\\ \boldsymbol{r}_0 &= [\mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle},\cdots,\mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle},\mathbb{R}_{n+1}^{\langle 1,\delta_{N-n}\rangle},\cdots,\mathbb{R}_N^{\langle 1,\delta_{N-n}\rangle}]^\mathsf{T}. \end{aligned}$$

Similarly, when n is chosen in the initial slot, we have (18) (shown at the top of the next page) and, further, the following

$$(\boldsymbol{I}_N - \beta\boldsymbol{M}_1)\cdot \boldsymbol{r}_1 = \boldsymbol{c}_1, \tag{20}$$

where,

$$\begin{aligned} \boldsymbol{M}_1 &= [\boldsymbol{Q}_1^\mathsf{T},\cdots,\boldsymbol{Q}_{n-1}^\mathsf{T},\boldsymbol{Q}_n^\mathsf{T}\bar{\mu}_n,\cdots,\boldsymbol{Q}_N^\mathsf{T}\bar{\mu}_N]^\mathsf{T},\\ \boldsymbol{c}_1 &= [-c,\cdots,-c,\ -c\bar{\mu}_n,-c\bar{\mu}_{n+1},\cdots,-c\bar{\mu}_N]^\mathsf{T},\\ \boldsymbol{r}_1 &= [\mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle},\cdots,\mathbb{R}_{n-1}^{\langle 0,\delta_{N-n}\rangle},\mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle},\cdots,\mathbb{R}_N^{\langle 1,\delta_{N-n}\rangle}]^\mathsf{T}. \end{aligned}$$

Thus, from (19) and (20), we can obtain

$$\mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle} = \boldsymbol{e}_n^\mathsf{T}\big(\boldsymbol{I}_N - \beta\boldsymbol{M}_0\big)^{-1}\boldsymbol{c}_0, \tag{21}$$

$$\mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle} = \boldsymbol{e}_n^\mathsf{T}\big(\boldsymbol{I}_N - \beta\boldsymbol{M}_1\big)^{-1}\boldsymbol{c}_1. \tag{22}$$

Similarly, replacing $\boldsymbol{c}_0$,Â $\boldsymbol{c}_1$ by $\mathbf{1}_N - \mathbf{1}_N^n$, $\mathbf{1}_N - \mathbf{1}_N^{n-1}$ from (19) and (20), respectively, we have

$$(\boldsymbol{I}_N - \beta\boldsymbol{M}_0)\cdot \boldsymbol{w}_0 = \mathbf{1}_N - \mathbf{1}_N^n, \tag{23}$$

$$(\boldsymbol{I}_N - \beta\boldsymbol{M}_1)\cdot \boldsymbol{w}_1 = \mathbf{1}_N - \mathbf{1}_N^{n-1}, \tag{24}$$

where,

$$\begin{aligned} &\boldsymbol{w}_0\\ &= [\mathbb{W}_1^{\langle 0,\delta_{N-n}\rangle},\cdots,\mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle},\mathbb{W}_{n+1}^{\langle 1,\delta_{N-n}\rangle},\cdots,\mathbb{W}_N^{\langle 1,\delta_{N-n}\rangle}]^\mathsf{T},\\ &\boldsymbol{w}_1\\ &= [\mathbb{W}_1^{\langle 0,\delta_{N-n}\rangle},\cdots,\mathbb{W}_{n-1}^{\langle 0,\delta_{N-n}\rangle},\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle},\cdots,\mathbb{W}_N^{\langle 1,\delta_{N-n}\rangle}]^\mathsf{T}. \end{aligned}$$

Further,

$$\mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle} = \boldsymbol{e}_n^\mathsf{T}\big(\boldsymbol{I}_N - \beta\boldsymbol{M}_0\big)^{-1}(\mathbf{1}_N - \mathbf{1}_N^n), \tag{25}$$

$$\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle} = \boldsymbol{e}_n^\mathsf{T}\big(\boldsymbol{I}_N - \beta\boldsymbol{M}_1\big)^{-1}(\mathbf{1}_N - \mathbf{1}_N^{n-1}). \tag{26}$$

After obtaining the four critical quantities, we now check the LP-indexability condition.

Lemma 2: Under Condition 1, for any $n \in \mathcal{N}\setminus\{N\}$, we have

(1) $\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle} > \mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}$,
(2) $\mathbb{W}_{n+1}^{\langle 1,\delta_{N-n}\rangle} > \mathbb{W}_{n+1}^{\langle 0,\delta_{N-n}\rangle}$.

Proof: Please see Appendix II. □

Lemma 3: Under Condition 1, Problem (6) is LP-indexable with price ν_n in (16).

Proof: According to Definition 1, we prove the indexability by checking three conditions.

(1) Obviously, $\mathbb{W}_n^{\langle 0,\emptyset\rangle} = 0$, $\mathbb{W}_n^{\langle 1,\emptyset\rangle} \ge 1$, and $\mathbb{W}_n^{\langle 1,\mathcal{N}\rangle} = \frac{1}{1-\beta}$. For any δ, $\mathbb{W}_n^{\delta} \le \frac{1}{1-\beta}$, and further $\mathbb{W}_n^{\langle 0,\mathcal{N}\rangle} < \frac{1}{1-\beta}$.
(2) The second condition is proved in Lemma 2.
(3) The third condition is proved in Lemma 1.

Therefore, the LP-indexability is proved. □

Following Lemma 3, the following theorem states our main result on the indexability of Problem (6).

Theorem 1 (indexability): Under Condition 1, we have

(1) if $\nu \le \nu_n$, it is optimal to serve the user in state n;
(2) if $\nu > \nu_n$, it is optimal not to serve the user in state n.

C. Computing Index

In this part, we exploit the structural property of the transition matrix $\boldsymbol{Q}$ to simplify the index computation and further, obtain the closed-form Whittle index.

$$\begin{bmatrix} \mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle} \\ \mathbb{R}_{n+1}^{\langle 1,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_N^{\langle 1,\delta_{N-n}\rangle} \end{bmatrix} = -\beta \begin{bmatrix} q_{1,1} & \cdots & q_{1,n-1} & \cdots & q_{1,N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ q_{n,1} & \cdots & q_{n,n} & \cdots & q_{n,N} \\ \bar{\mu}_{n+1}q_{n+1,1} & \cdots & \bar{\mu}_{n+1}q_{n+1,n} & \cdots & \bar{\mu}_{n+1}q_{n+1,N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \bar{\mu}_N q_{N,1} & \cdots & \bar{\mu}_N q_{N,n-1} & \cdots & \bar{\mu}_N q_{N,N} \end{bmatrix} \begin{bmatrix} \mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle} \\ \mathbb{R}_{n+1}^{\langle 1,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_N^{\langle 1,\delta_{N-n}\rangle} \end{bmatrix} + \begin{bmatrix} -c \\ \vdots \\ -c \\ -c\bar{\mu}_{n+1} \\ \vdots \\ -c\bar{\mu}_N \end{bmatrix} \tag{17}$$

$$\begin{bmatrix} \mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_{n-1}^{\langle 0,\delta_{N-n}\rangle} \\ \mathbb{R}_{n}^{\langle 1,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_N^{\langle 1,\delta_{N-n}\rangle} \end{bmatrix} = -\beta \begin{bmatrix} q_{1,1} & \cdots & q_{1,n-1} & q_{1,n} & \cdots & q_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ q_{n-1,1} & \cdots & q_{n-1,n-1} & q_{n-1,n} & \cdots & q_{n-1,N} \\ \bar{\mu}_n q_{n,1} & \cdots & \bar{\mu}_n q_{n,n-1} & \bar{\mu}_n q_{n,n} & \cdots & \bar{\mu}_n q_{n,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \\ \bar{\mu}_N q_{N,1} & \cdots & \bar{\mu}_N q_{N,n-1} & \bar{\mu}_N q_{N,n} & \cdots & \bar{\mu}_N q_{N,N} \end{bmatrix} \begin{bmatrix} \mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_{n-1}^{\langle 0,\delta_{N-n}\rangle} \\ \mathbb{R}_{n}^{\langle 1,\delta_{N-n}\rangle} \\ \vdots \\ \mathbb{R}_N^{\langle 1,\delta_{N-n}\rangle} \end{bmatrix} + \begin{bmatrix} -c \\ \vdots \\ -c \\ -c\bar{\mu}_n \\ \vdots \\ -c\bar{\mu}_N \end{bmatrix} \tag{18}$$

Proposition 1: Under Condition 1, we have

1) $\mathbb{W}_1^{\langle 0,\delta_{N-n}\rangle} = \cdots = \mathbb{W}_{n-1}^{\langle 0,\delta_{N-n}\rangle} = \mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}$.
2) $\mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle} = \cdots = \mathbb{R}_{n-1}^{\langle 0,\delta_{N-n}\rangle} = \mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle}$.
3) The Whittle index is

$$\nu_n = \frac{-\mu_n \mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle}}{1-\mu_n \mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}}. \tag{27}$$

Proof: Following the proof of Lemma 2, we have

$$\mathbb{W}_1^{\langle 0,\delta_{N-n}\rangle} = \cdots = \mathbb{W}_{n-1}^{\langle 0,\delta_{N-n}\rangle} = \mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}$$

and

$$\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle} - \mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle} = \frac{1-\mu_n \mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}}{1-\mu_n}\left[\boldsymbol{e}_n^{\mathsf{T}}\left(\boldsymbol{I}_N - \beta \boldsymbol{M}_0\right)^{-1}\boldsymbol{e}_n\right]. \tag{28}$$

Similarly,

$$\mathbb{R}_1^{\langle 0,\delta_{N-n}\rangle} = \cdots = \mathbb{R}_{n-1}^{\langle 0,\delta_{N-n}\rangle} = \mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle}$$

and

$$\mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle} - \mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle} = \frac{-\mu_n \mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle}}{1-\mu_n}\left[\boldsymbol{e}_n^{\mathsf{T}}\left(\boldsymbol{I}_N - \beta \boldsymbol{M}_0\right)^{-1}\boldsymbol{e}_n\right]. \tag{29}$$

Therefore,

$$\nu_n = \frac{\mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle} - \mathbb{R}_n^{\langle 0,\delta_{N-n}\rangle}}{\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle} - \mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}} = \frac{-\mu_n \mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle}}{1-\mu_n \mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}}.$$

□

Based on (27), in order to obtain ν_n, we only need to compute $\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}$ and $\mathbb{R}_n^{\langle 1,\delta_{N-n}\rangle}$. Further, by some complex operations, we can obtain the closed-form Whittle index as follows

$$\nu_n = \frac{c\mu_n}{1-\beta+f(n)}, \quad 1 \le n \le N \tag{30}$$

where

$$f(n) = \sum_{i=n+1}^{N}\left(\beta q_{n,i}\left(1-\prod_{j=n+1}^{i}\frac{\frac{1}{\bar{\mu}_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar{\mu}_j}-\beta\lambda_{j-1}}\right)+\mu_n d_{n-1,i}K_i\right)$$

$$d_{n-1,i} = -\beta\left(q_{n,i}+\sum_{k=i+1}^{N} q_{n,k}\prod_{j=i+1}^{k}\frac{\frac{1}{\bar{\mu}_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar{\mu}_j}-\beta\lambda_{j-1}}\right)$$

$$K_i = \frac{\frac{1}{1-\mu_i}-\frac{1}{1-\mu_{i-1}}}{\frac{1}{1-\mu_i}-\beta\lambda_{i-1}}$$

for i $(n+1 \le i \le N)$.

V. Indexability Extension and Scheduling Policy

In this section, we first extend the proposed Condition 1 and obtain the indexability as well as the Whittle index. Next, we propose an eigenvalue-arithmetic-mean scheme to approximate any transition matrix, and further, obtain the corresponding approximate Whittle index. Finally, based on the closed-form Whittle index, we construct an efficient scheduling policy.

A. Indexability Extension

In Section IV-C, the computing process of ν_n shows that the ν_n only depends on the structure of $\boldsymbol{Q}$ rather than the sign of λ_j, i.e., (7). Thus, we release Condition 1 based on the monotonicity of ν_n and obtain the following theorem on the indexability.

Theorem 2: If $\boldsymbol{Q}$ can be written as

$$\boldsymbol{Q} = \boldsymbol{O}_0 + \epsilon_1\boldsymbol{O}_1 + \epsilon_2\boldsymbol{O}_2 + \cdots + \epsilon_{2N-2}\boldsymbol{O}_{2N-2}, \tag{31}$$

then Problem (6) is indexable and the Whittle index for state n $(n = 1, \cdots, N)$ is

$$\nu_n = \begin{cases} \infty, & \text{if } \beta = 1, n = N, \\ \frac{c\mu_n}{1-\beta+f(n)}, & \text{otherwise.} \end{cases} \tag{32}$$

Proof: Please see the Appendix III. □

Remark 2: The critical constraint (7) of Condition 1 is deleted in this theorem compared with Lemma 3.

Corollary 1: If $\boldsymbol{Q}$ can be written as

$$\boldsymbol{Q} = \boldsymbol{O}_0 + \epsilon_1 \boldsymbol{O}_1 + \epsilon_2 \boldsymbol{O}_2 + \cdots + \epsilon_{2N-2} \boldsymbol{O}_{2N-2}$$

and $\lambda_1 = \cdots = \lambda_{N-1} = \lambda$, then Problem (6) is indexable and the Whittle index for state n $(n = 1, \cdots, N)$ is

$$\nu_n = \begin{cases} \infty, & \text{if } \ \beta = 1, n = N, \\ \dfrac{c\mu_n}{1 - \beta + \beta \sum_{i=n+1}^{N} \frac{q_{n,i}(\mu_i - \mu_n)}{1 - \beta\lambda(1-\mu_i)}}, & \text{otherwise.} \end{cases} \tag{33}$$

Remark 3: This corollary shows that the Whittle index degenerates into that of [18] if $\lambda_1 = \cdots = \lambda_{N-1} = \lambda = 0$.

B. Transition Matrix Approximation

Given a generic $\boldsymbol{Q}$, where

$$\boldsymbol{Q} \neq \boldsymbol{O}_0 + \epsilon_1 \boldsymbol{O}_1 + \epsilon_2 \boldsymbol{O}_2 + \cdots + \epsilon_{2N-2} \boldsymbol{O}_{2N-2},$$

thus the result of Theorem 2 cannot be used.

For this case, we approximate $\boldsymbol{Q}$ by the following eigenvalue-arithmetic-mean scheme,

$$\boldsymbol{Q} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}, \tag{34}$$
$$\hat{\boldsymbol{Q}} = \boldsymbol{V}\hat{\boldsymbol{\Lambda}}\boldsymbol{V}^{-1}, \tag{35}$$
$$\hat{\boldsymbol{\Lambda}} = \mathrm{diag}(1, \hat{\lambda}, \cdots, \hat{\lambda}), \tag{36}$$
$$\hat{\lambda} = \frac{\mathrm{trace}(\boldsymbol{Q}) - 1}{N - 1}, \tag{37}$$

where $\hat{\lambda}$ is the arithmetic mean of the $N-1$ eigenvalues of $\boldsymbol{Q}$ (excluding the trivial eigenvalue 1). Thus, the approximate matrix $\hat{\boldsymbol{Q}}$ satisfies the condition of Corollary 1, and furthermore, the Whittle index can be approximated by

$$\nu_n = \begin{cases} \infty, & \text{if } \ \beta = 1, n = N, \\ \dfrac{c\mu_n}{1 - \beta + \beta \sum_{i=n+1}^{N} \frac{\hat{q}_{n,i}(\mu_i - \mu_n)}{1 - \beta\hat{\lambda}(1-\mu_i)}}, & \text{otherwise.} \end{cases} \tag{38}$$

C. Scheduling Policy

In the previous sections, we have obtained the closed-form Whittle index for each subproblem. Now, we construct the joint scheduling policy for the original problem.

In particular, the scheduling policy is to serve the user in $k^*(t)$ with the highest actual price, i.e.,

$$k^*(t) = \mathrm{argmax}_{k \in \mathcal{K}} \left[\nu_{k,x_k(t)}\right], \quad \text{if } \ \nu_{k,x_k(t)} < \infty. \tag{39}$$

Actually, $\nu_{k,x_k(t)} < \infty$ always holds if $0 \leq \beta < 1$. It happens $\nu_{k,x_k(t)} \rightarrow \infty$ only when $\beta = 1$ and $x_k(t) = N_k$, corresponding to the average case.

Therefore, the second item, $c_k\mu_{k,x_k(t)}$, of Laurent expansion of $\nu_{k,x_k(t)}$ would be taken as the secondary index in the case of $\beta = 1$ and $x_k(t) = N_k$ since

$$\lim_{\beta \rightarrow 1} (1-\beta)\nu_{k,N_k} = \frac{(1-\beta)c_k\mu_{k,N_k}}{1-\beta} = c_k\mu_{k,N_k}. \tag{40}$$

Now, we give the marginal productivity index (MPI) scheduler in Algorithm 1. The MPI scheduler always serves the user currently with the best condition, i.e., $\nu_1 \leq \cdots \leq \nu_N$ and is one of the best-condition schedulers, which has the stability property in a Markovian setting [9].

Theorem 3 ([9]): The MPI scheduler with one server is maximally stable under arbitrary arrivals.

Algorithm 1 MPI Scheduler ($\beta = 1$)

```
for t ∈ T
    C ← number of system users in N_k (k ∈ K)
if C ≥ 1 then
    Serve one user in N_k with max{c_k μ_{k,N_k}} (k ∈ K)
    (breaking ties randomly)
else
    if condition (31) is satisfied
        Serve the user k*(t) with highest index value by (33)
    else
        Serve the user k*(t) with highest index value by (38)
    end if
    (breaking ties randomly)
end if
end for
```

VI. Numerical Simulation

In this section, we compare the proposed MPI scheduler with the following policies

- the $c\mu$ rule, $\nu_{k,n}^{c\mu} = c_k\mu_{k,n}$,
- the RB rule, $\nu_{k,n}^{\mathrm{RB}} = \frac{c_k\mu_{k,n}}{\sum_{m=1}^{N_k} q_{k,m}^{\mathrm{SS}}\mu_{k,m}}$,
- the PB rule, $\nu_{k,n}^{\mathrm{PB}} = \frac{c_k\mu_{k,n}}{\mu_{k,N_k}}$,
- the SB rule, $\nu_{k,n}^{\mathrm{SB}} = c_k \sum_{m=1}^{n} q_{k,m}^{\mathrm{SS}}$,
- the PI$^{\mathrm{SS}}$ rule [14], $\nu_{k,n}^{\mathrm{SS}} = \frac{c_k\mu_{k,n}}{\sum_{m>n} q_{k,m}^{\mathrm{SS}}(\mu_{k,m} - \mu_{k,n})}$,

where $q_{k,m}^{\mathrm{SS}}$ is the stationary probability of state m of a user of class k.

Specifically, we only consider the case with at most one user served at each time slot. If there is more than one user having the highest index value, we uniformly choose one of them. In addition, we only consider two classes of users for a clearer performance comparison. Moreover, before evaluating the performance of different schedulers, we first test their similarity for a given scenario by computing the corresponding index and then choose one scheduler as a representative among multiple identical schedulers. In this way, we can decrease the time for numerical simulation and meanwhile, obtain compact figures for performance comparison.

Let $\tau = 1.67$ msec for each slot for practical application [28]. The arrival probability for a new user of class k is characterized by $\xi_k = \rho_k\mu_{k,N_k}$. For comparison, we adopt the transmission rate $s_{k,n}$ in [28], and job size $\mathbb{E}_0\{\mathsf{b}_k\} = 0.5$Mb for HTML, $\mathbb{E}_0\{\mathsf{b}_k\} = 5$Mb for PDF, and $\mathbb{E}_0\{\mathsf{b}_k\} = 50$Mb for MP3. In this case, the departure probability is determined by $\mu_{k,n} = \tau s_{k,n}/\mathbb{E}_0\{\mathsf{b}_k\}$. We assume that $\rho_1 = \rho_2$ and the system load $\rho = \rho_1 + \rho_2$ varies from 0.3 to 1 for a better presentation. The initial channel condition of a new user at the moment of entering the system is assumed to be determined

TABLE I
PARAMETERS ADOPTED IN SIMULATION

No.	$s_{k,n}$ (Mb/s)	(c_1, c_2)	Channel Transition Matrices	Job Size (MB)
1	$\begin{matrix} 8.4 & 50.4 & 53.76 \\ 26.88 & 44.688 & 80.64 \end{matrix}$	$(1, 1)$	$\begin{bmatrix} 0.00 & 0.80 & 0.20 \\ 0.30 & 0.50 & 0.20 \\ 0.30 & 0.60 & 0.10 \end{bmatrix}, \begin{bmatrix} 0.00 & 0.80 & 0.20 \\ 0.30 & 0.50 & 0.20 \\ 0.30 & 0.60 & 0.10 \end{bmatrix}$	$0.5, 0.5$
2	$\begin{matrix} 8.4 & 16.8 & 33.6 \\ 8.4 & 16.8 & 33.6 \end{matrix}$	$(10, 1)$	$\begin{bmatrix} 0.00 & 0.50 & 0.50 \\ 0.10 & 0.40 & 0.50 \\ 0.10 & 0.70 & 0.20 \end{bmatrix}, \begin{bmatrix} 0.25 & 0.60 & 0.15 \\ 0.35 & 0.50 & 0.15 \\ 0.35 & 0.55 & 0.10 \end{bmatrix}$	$5, 0.5$
3	$\begin{matrix} 8.4 & 16.8 & 50.4 & 67.2 \\ 26.88 & 33.6 & 44.688 & 80.64 \end{matrix}$	$(2, 3)$	$\begin{bmatrix} 0.50 & 0.10 & 0.20 & 0.20 \\ 0.15 & 0.45 & 0.20 & 0.20 \\ 0.15 & 0.15 & 0.50 & 0.20 \\ 0.15 & 0.15 & 0.10 & 0.60 \end{bmatrix}, \begin{bmatrix} 0.10 & 0.35 & 0.25 & 0.30 \\ 0.20 & 0.25 & 0.25 & 0.30 \\ 0.20 & 0.30 & 0.20 & 0.30 \\ 0.20 & 0.30 & 0.40 & 0.10 \end{bmatrix}$	$0.5, 0.5$

by the stationary probability vector, i.e., with probability $q^{\text{SS}}_{k,m}$ in state m for a new user of class k. The parameter setting for the following scenarios is stated in Table I.

A. Scenario 1

In this case, the setting is given in Table I. In particular, the users are divided into two different classes. Each user requires a job of expected size of 0.5Mb, and has the same waiting cost $c_1 = c_2 = 1$. The channel state transition matrix is identical. But the second class of users has a better transmission rate than the first class. Our goal is to minimize the number of users waiting for service in the system.

Under this setting, three policies (cμ, RB, and MPI) can be shown to bring about the same scheduling rule, i.e., the scheduling (class, state) order $(2,3) > (1,3) > (1,2) > (2,2) > (2,1) > (1,1)$. Also, PB and PI$^{\text{SS}}$ yield the same result, i.e., $(1,3) = (2,3) > (1,2) > (2,2) > (2,1) > (1,1)$. The SB policy will generate the order $(2,3) = (1,3) > (1,2) = (2,2) > (2,1) = (1,1)$. Thus, Fig. 1 shows that the time-average waiting cost varies with system load ρ for three policies, and the number of users in the system varies with time slots. Obviously, we observe that the behavior of all policies is quite similar. In addition, Fig. 1 clearly shows that cμ, RB, and MPI perform better than PB, SB, and PI$^{\text{SS}}$. This is because cμ, RB, and MPI keep scheduling balance between class 1 and class 2. All those policies perform well with $\rho < 0.9$ but clearly have problems with stability since those policies become unstable close to 1 at which point the time-average waiting cost begins rising very steeply.

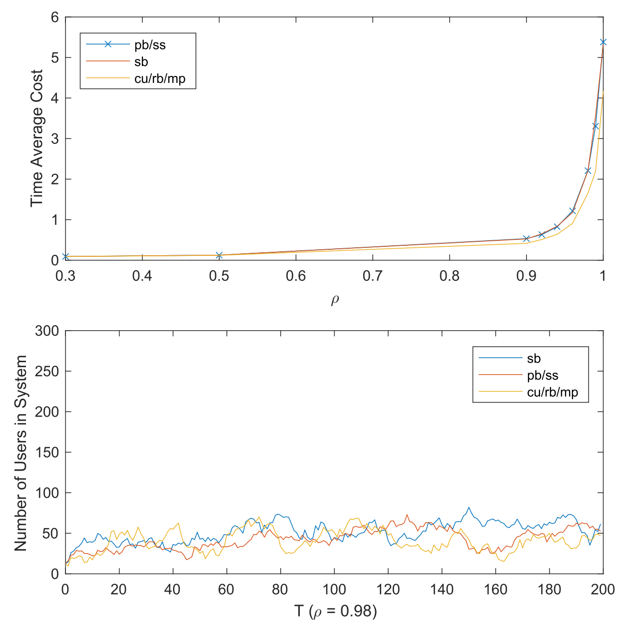


Fig. 1. Scenario 1 [upper]: Time average waiting cost as a function of ρ; [lower]: Number of users in the system as a function of time ($\rho = 0.98$).

B. Scenario 2

In this case, we consider two classes of users with different job sizes: the first one requires a job of expected size 5Mb while the second one requires a job of 0.5Mb. The waiting costs for the two classes are $c_1 = 10$ and $c_2 = 1$, respectively. The two classes have the same transmission rate and different channel state transition matrices.

Thus, we can easily check that PI$^{\text{SS}}$ and MPI generate the same scheduling rule $(1,3) > (2,3) > (1,2) > (2,2) > (1,1) > (2,1)$, and PB and RB have the same rule $(1,3) > (1,2) > (1,1) > (2,3) > (2,2) > (2,1)$. Fig. 2 shows that MPI has comparable performance with $c\mu$ and better performance than other policies in both time-average waiting cost and average number of users in system. From the scheduling order, we observe that PB (or RB) has the worst performance because of the extreme unbalance in user class, i.e., severing class 1 with complete priority than class 2. SB has the worse performance because of partial unbalance in user class from its scheduling order $(1,3) > (1,2) > (2,3) > (1,1) > (2,2) > (2,1)$.

C. Scenario 3

In this case, we assume that every class of users has 4 states, different waiting costs, different transmission rates and different channel transition matrices.

In this case, we can check that PB and RB are same, i.e. $(1,4) > (1,3) > (2,4) > (2,3) > (1,2) > (2,2) > (1,1) > (2,1)$. Fig. 3 shows that MPI policy $(2,4) > (1,4) > (2,3) > (1,3) > (1,2) > (2,2) > (1,1) > (2,1)$ has

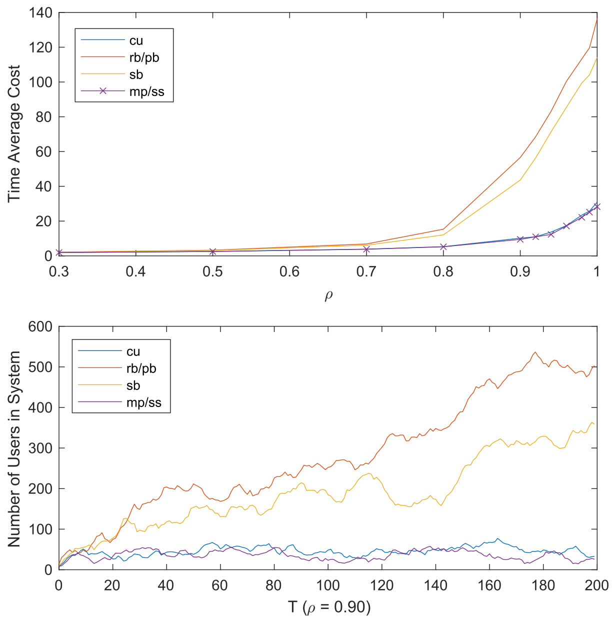


Fig. 2. Scenario 2 [upper]: Time average waiting cost as a function of ρ; [lower]: Number of users in the system as a function of time ($\rho = 0.90$).

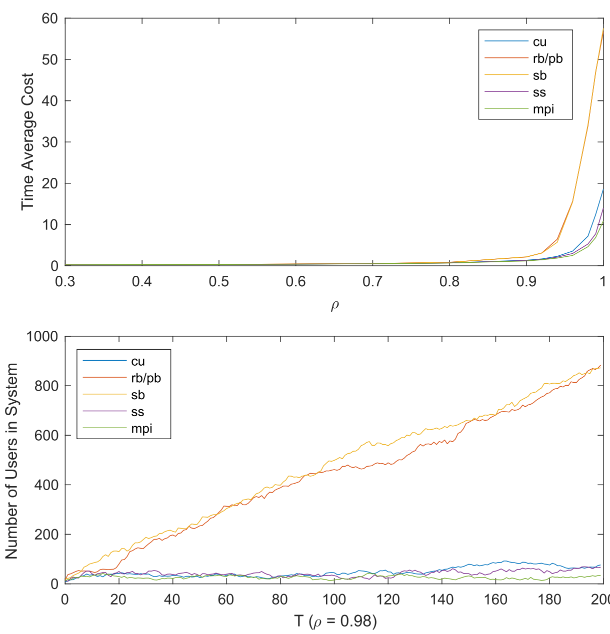


Fig. 3. Scenario 3 [upper]: Time average waiting cost as a function of ρ; [lower]: Number of users in the system as a function of time ($\rho = 0.94$).

comparable performance with $\mathrm{PI}^{\mathrm{SS}}$ policy, $(2,4) = (1,4) > (2,3) > (1,3) > (1,2) > (2,2) > (1,1) > (2,1)$, and better performance than others in both the average cost and the number of waiting users.

VII. Conclusion

In this paper, we have investigated the opportunistic scheduling problem involving multi-class multi-state time-varying Markovian channels. Generally, the problem can be formulated as a restless multiarmed bandit problem. To the best of our knowledge, previous work only established an index policy for a two-state channel process and derived some limited results on multi-state time-varying channels under an assumption of indexability as a prerequisite. To fill this gap, for the class of state transition matrices characterized by our proposed sufficient condition, we prove the indexability of the Whittle index policy and obtain the closed-form Whittle index. Simulation results show that the proposed index scheduler is effective in scheduling multi-class multi-state channels. One future objective is to seek more generic conditions to guarantee the indexability.

Appendix I
Proof of Lemma 1

Let v_n^* denote the optimal value function, and

$$v_n^a := r_n^a - \nu w_n^a + \beta \sum_{m\in\mathcal{N}} p_{n,m}^a v_m^*,$$

$$\begin{aligned} g_n(v_{-n}^*, v_{-(n+1)}^*) &:= \sum_{i=1}^{n-1} \epsilon_{N-1+i} v_i^* - \sum_{i=1}^{n-2} \epsilon_{N-1+i} v_{i+1}^* \\ &\quad + \sum_{i=n+1}^{N-1} \epsilon_{N-i} v_{i+1}^* - \sum_{i=n+2}^{N-1} \epsilon_{N-i} v_i^*, \end{aligned}$$

$$\alpha_n^0 := \begin{cases} -\epsilon_{N-n}, & \text{if } n = 1, \\ -\epsilon_{N-2+n} - \epsilon_{N-n}, & \text{if } 2 \le n \le N-1 \\ -\epsilon_{2N-2}, & \text{if } n = N, \end{cases}$$

$$\alpha_{n+1}^1 := \begin{cases} \epsilon_{N-n} - \epsilon_{N-n-1}, & \text{if } 1 \le n \le N-2, \\ \epsilon_{N-n}, & \text{if } n = N-1, \\ 0, & \text{if } n = N. \end{cases}$$

For state $n \in \mathcal{N}$, the Bellman equation is

$$\begin{aligned} v_n^* &= \max\{v_n^0; v_n^1\} \\ &= \max\Big\{ r_n^0 - \nu w_n^0 + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[(\lambda + \alpha_n^0) v_n^* + \alpha_{n+1}^1 v_{n+1}^*]; \\ &\quad r_n^1 - \nu w_n^1 + \beta \sum_{m\in\mathcal{N}'} (1-\mu_n) h_m v_m^* + \beta \mu_n v_0^* \\ &\quad + \beta(1-\mu_n) g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta(1-\mu_n)[(\lambda + \alpha_n^0) v_n^* + \alpha_{n+1}^1 v_{n+1}^*] \Big\} \\ &= -c + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[(\lambda + \alpha_n^0) v_n^* + \alpha_{n+1}^1 v_{n+1}^*] \\ &\quad + \max\Big\{0; -\nu + \mu_n \Big(c + \beta v_0^* - \beta \sum_{m\in\mathcal{N}'} h_m v_m^* \\ &\quad - \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad - \beta[(\lambda + \alpha_n^0) v_n^* + \alpha_{n+1}^1 v_{n+1}^*] \Big) \Big\}, \end{aligned}$$

where the first term in the curly brackets corresponds to action 0 and the second to action 1.

Obviously, transmitting (i.e., action 1) is optimal in state $n \in \mathcal{N} \setminus \{0\}$ if the first term is less than the second one.

For ease of presentation, let

$$\begin{aligned} Z &:= c + \beta v_0^* - \beta g_n(v_{-n}^*, v_{-(n+1)}^*) - \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^*, \\ Z_n &:= Z - \beta(\lambda + \alpha_n^0 + h_n) v_n^* - \beta(\alpha_{n+1}^1 + h_{n+1}) v_{n+1}^*. \end{aligned}$$

Now, we analyze the Bellman equation by two cases.

Case 1: If $\nu > 0$, we have $v_n^* \leq 0$ for any $n \in \mathcal{N} \setminus \{0\}$. If transmitting is optimal in state $n \in \mathcal{N}\setminus\{0, N\}$, we obtain $-\nu + \mu_n Z_n \geq 0$ indicating $Z_n > 0$, and further, $-\nu + \mu_{n+1} Z_n > -\nu + \mu_n Z_n$ since $\mu_{n+1} > \mu_n$. Thus,

$$\begin{aligned} v_n^* &= -c + \mu_n Z_n - \nu + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* \\ &\quad + \beta[(\lambda + \alpha_n^0)v_n^* + \alpha_{n+1}^1 v_{n+1}^* + g_n(v_{-n}^*, v_{-(n+1)}^*)] \\ &< -c + \mu_{n+1} Z_n - \nu + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* \\ &\quad + \beta[(\lambda + \alpha_n^0)v_n^* + \alpha_{n+1}^1 v_{n+1}^* + g_n(v_{-n}^*, v_{-(n+1)}^*)], \end{aligned}$$

equivalently,

$$\begin{aligned} & v_n^*(1 - \beta(1-\mu_{n+1})(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})) \\ & < -c + \mu_{n+1} Z - \nu + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ & \quad + \beta(1-\mu_{n+1})(h_n + \alpha_n^0 + \epsilon_{N-n} + \epsilon_{N+n-1})v_n^* \\ & \quad + \beta(1-\mu_{n+1})(h_{n+1} + \alpha_{n+1}^1)v_{n+1}^* \\ & = -c + \mu_{n+1} Z - \nu + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ & \quad + \beta(1-\mu_{n+1})(h_n + \gamma_n)v_n^* \\ & \quad + \beta(1-\mu_{n+1})(h_{n+1} + \alpha_{n+1}^1)v_{n+1}^*, \end{aligned} \tag{41}$$

where, $\gamma_n := \alpha_n^0 + \epsilon_{N+n-1} + \epsilon_{N-n}$.

For state $n+1$, if action '1' is adopted, then we have according to Bellman equation

$$\begin{aligned} v_{n+1}^1 &= -c - \nu + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* + \beta g_{n+1}(v_{-(n+1)}^*, v_{-(n+2)}^*) \\ &\quad + \beta[(\lambda + \alpha_{n+1}^0)v_{n+1}^* + \alpha_{n+2}^1 v_{n+2}^*] \\ &\quad + \mu_{n+1}\Big(c + \beta v_0^* - \beta \sum_{m\in\mathcal{N}'} h_m v_m^* \\ &\quad - \beta g_{n+1}(v_{-(n+1)}^*, v_{-(n+2)}^*) \\ &\quad - \beta[(\lambda + \alpha_{n+1}^0)v_{n+1}^* + \alpha_{n+2}^1 v_{n+2}^*]\Big) \\ &\overset{(a)}{=} -c + \mu_{n+1} Z - \nu + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* \\ &\quad + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta(1-\mu_{n+1})(h_n + \alpha_n^0 + \epsilon_{N+n-1} + \epsilon_{N-n})v_n^* \\ &\quad + \beta(1-\mu_{n+1}) \\ &\quad \times(h_{n+1} + \lambda + \alpha_{n+1}^1 - \epsilon_{N-n} - \epsilon_{N+n-1})v_{n+1}^*, \end{aligned}$$

$$\Leftrightarrow$$

$$\begin{aligned} & v_{n+1}^1 - \beta(1-\mu_{n+1})(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})v_{n+1}^* \\ & \quad = -c + \mu_{n+1} Z - \nu + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* \\ & \quad + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) + \beta(1-\mu_{n+1})(h_n + \gamma_n)v_n^* \\ & \quad + \beta(1-\mu_{n+1})(h_{n+1} + \alpha_{n+1}^1)v_{n+1}^*, \end{aligned} \tag{42}$$

where (a) is due to $g_n(v_{-n}^*, v_{-(n+1)}^*) = g_{n+1}(v_{-(n+1)}^*, v_{-(n+2)}^*) + (\epsilon_{N-2+n} - \epsilon_{N-1+n})v_n^* + (\epsilon_{N-n-1} - \epsilon_{N-n-2})v_{n+2}^*$, $\alpha_{n+1}^0 = \alpha_{n+1}^1 - \epsilon_{N-n} - \epsilon_{N+n-1}$, and $\alpha_{n+2}^1 = \epsilon_{N-n-1} - \epsilon_{N-n-2}$.

Thus, combining (41) and (42), we have

$$\begin{aligned} & v_n^*(1 - \beta(1-\mu_{n+1})(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})) \\ & \quad < v_{n+1}^1 - \beta(1-\mu_{n+1})(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})v_{n+1}^* \\ & \quad \leq v_{n+1}^* - \beta(1-\mu_{n+1})(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})v_{n+1}^* \\ & \quad = v_{n+1}^*(1 - \beta(1-\mu_{n+1})(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})), \end{aligned}$$

which indicates $v_n^* < v_{n+1}^*$.

Meanwhile,

$$\begin{aligned} v_n^* &\geq v_n^0 \\ &= -c + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[(\lambda + \alpha_n^0)v_n^* + \alpha_{n+1}^1 v_{n+1}^*] \\ &= -c + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[(h_n + \lambda + \alpha_n^0)v_n^* + (h_{n+1} + \alpha_{n+1}^1)v_{n+1}^*] \end{aligned}$$

$$\Leftrightarrow$$

$$\begin{aligned} & v_n^*(1 - \beta(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})) \\ & \quad \geq -c + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ & \quad + \beta[(h_n + \gamma_n)v_n^* + (h_{n+1} + \alpha_{n+1}^1)v_{n+1}^*]. \end{aligned} \tag{43}$$

On the other hand, we have according to Bellman equation

$$\begin{aligned} v_{n+1}^0 &= -c + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* + \beta g_{n+1}(v_{-(n+1)}^*, v_{-(n+2)}^*) \\ &\quad + \beta[(\lambda + \alpha_{n+1}^0)v_{n+1}^* + \alpha_{n+2}^1 v_{n+2}^*] \\ &\overset{(b)}{=} -c + \beta \sum_{m\in\mathcal{N}'} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[\gamma_n v_n^* + (\lambda + \alpha_{n+1}^1 - \epsilon_{N-n} - \epsilon_{N+n-1})v_{n+1}^*] \\ &= -c + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[(h_n + \gamma_n)v_n^* + (h_{n+1} + \alpha_{n+1}^1)v_{n+1}^*] \\ &\quad + \beta(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})v_{n+1}^* \\ &\overset{(a)}{\leq} -c + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[(h_n + \gamma_n)v_n^* + (h_{n+1} + \alpha_{n+1}^1)v_{n+1}^*] \\ &\quad + \beta(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})v_{n+1}^0 \\ &\Leftrightarrow v_{n+1}^0(1 - \beta(\lambda - \epsilon_{N-n} - \epsilon_{N+n-1})) \\ &\leq -c + \beta \sum_{m\in\mathcal{N}'}^{m\neq n,n+1} h_m v_m^* + \beta g_n(v_{-n}^*, v_{-(n+1)}^*) \\ &\quad + \beta[(h_n + \gamma_n)v_n^* + (h_{n+1} + \alpha_{n+1}^1)v_{n+1}^*] \end{aligned} \tag{44}$$

where (a) is due to $\lambda \leq \epsilon_{N-n} + \epsilon_{N+n-1}$ and $v_{n+1}^0 \leq v_{n+1}^*$, and (b) is due to $g_n(v_{-n}^*, v_{-(n+1)}^*) = g_{n+1}(v_{-(n+1)}^*, v_{-(n+2)}^*) + \gamma_n v_n^* + (\epsilon_{N-n-1} - \epsilon_{N-n-2})v_{n+2}^*$.

Thus, combining (43) and (44), we have $v^0_{n+1} \le v^*_n$. Since $v^*_n < v^*_{n+1}$, we conclude $v^0_{n+1} \le v^*_n < v^*_{n+1} = v^1_{n+1}$, that is, transmitting is optimal in state $n+1$.

Case 2: if $\nu < 0$, then we proceed as follows. First, using the Bellman equation it is easy to obtain that $v^*_0 = -\frac{\nu}{1-\beta}$ because action 1 is optimal in state 0 and thus $-\nu$ is obtained in every period forever. Notice that the one-period net reward, $r^a_n - \nu w^a_n$, is for any state $n \in \mathcal{N}$ and any action $a \in \mathcal{A}$ upperbounded by $-\nu$, i.e., $|r^a_n - \nu r^a_n| \le -\nu$. Hence, $v^*_m \le -\frac{\nu}{1-\beta} = v^*_0$ for any $m \in \mathcal{N}'$, and therefore (using $c > 0$ and $\lambda + \sum_{m\in\mathcal{N}'} h_m = 1$) $Z_n > 0$, and finally, for any state $n \in \mathcal{N} \setminus \{0\}$, $-\nu + \mu_n Z_n > 0$. That is, transmitting is optimal in any state $n \in \mathcal{N}$.

Combining the two cases, we complete the proof.

Appendix II
Proof of Lemma 2

We first show $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n - \mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_n > 0$ by the following four steps.

Step 1: According to the definition of β-average work, we have $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \ge 1$, $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{n+1} \ge 1, \cdots,$ $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_N \ge 1$. To show $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \ge \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{n+1} \ge \cdots \ge \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_N$, we only need to show $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i \ge \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{i+1}$ for any i ($n \le i \le N-1$).

For (24), we perform the following operations sequentially

1) dividing the i-th equation by $1-\mu_i$,
2) dividing the $(i+1)$-th equation by $1-\mu_{i+1}$,
3) subtracting the i-th equation from the $(i+1)$-th one,

then we obtain the i-th equation

$$\left[-\frac{1}{\bar{\mu}_i} + \beta\lambda_{i-1}\right]\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i + \left[\frac{1}{\bar{\mu}_{i+1}} - \beta\lambda_{i-1}\right]\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{i+1} = \frac{1}{\bar{\mu}_{i+1}} - \frac{1}{\bar{\mu}_i}, \quad (45)$$

equivalently,

$$\left[\frac{1}{\bar{\mu}_i} - \beta\lambda_{i-1}\right](\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{i+1} - \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i) = \left[\frac{1}{\bar{\mu}_{i+1}} - \frac{1}{\bar{\mu}_i}\right](1 - \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{i+1}) \le 0,$$

which implies $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{i+1} \le \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i$.

Step 2: To show $\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_1 = \cdots = \mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{n-1}$, we only need to show $\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_i = \mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{i+1}$ for i ($1 \le i \le n-2$).

For (24), we subtract the i-th equation from the $(i+1)$-th one, and come to

$$-\left[1-\beta\lambda_i\right]\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_i + \left[1-\beta\lambda_i\right]\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{i+1} = 0, \quad (46)$$

which indicates $\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_i = \mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{i+1}$.

To show $\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_1 = \cdots = \mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{n-1} \le \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n$, we have by the $(n-1)$-th equation

$$\left[1 - \beta\sum_{i=1}^{n-1} q_{n-1,i}\right]\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{n-1} - \beta\sum_{i=n}^{N} q_{n-1,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i = 0, \quad (47)$$

equivalently,

$$\begin{aligned}\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{n-1} &= \frac{\beta\sum_{i=n}^{N} q_{n-1,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i}{1-\beta\sum_{i=1}^{n-1} q_{n-1,i}} \\ &\overset{(a)}{\le} \frac{\beta\sum_{i=n}^{N} q_{n-1,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n}{1-\beta\sum_{i=1}^{n-1} q_{n-1,i}} \\ &\overset{(b)}{\le} \frac{\sum_{i=n}^{N} q_{n-1,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n}{1-\sum_{i=1}^{n-1} q_{n-1,i}} \\ &= \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n, \end{aligned} \quad (48)$$

where (a) is due to $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \ge \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_{n+1} \ge \cdots \ge \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_N$, and (b) is because $\frac{\beta\sum_{i=n}^{N} q_{n-1,i}}{1-\beta\sum_{i=1}^{n-1} q_{n-1,i}}$ is increasing in β ($0 \le \beta < 1$).

Step 3: Considering the n-th equation of (24), we have

$$\begin{aligned}&-\beta(1-\mu_n)\sum_{i=1}^{n-1} q_{n,i}\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{n-1} \\ &\quad + (1-\beta(1-\mu_n)q_{n,n})\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \\ &\quad - \beta(1-\mu_n)\sum_{i=n+1}^{N} q_{n,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i = 1, \end{aligned} \quad (49)$$

equivalently,

$$\begin{aligned}&(1-\beta(1-\mu_n)q_{n,n})\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \\ &= 1 + \beta(1-\mu_n)\sum_{i=1}^{n-1} q_{n,i}\mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_{n-1} \\ &\quad + \beta(1-\mu_n)\sum_{i=n+1}^{N} q_{n,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i \\ &\overset{(a)}{\le} 1 + \beta(1-\mu_n)\sum_{i=1}^{n-1} q_{n,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \\ &\quad + \beta(1-\mu_n)\sum_{i=n+1}^{N} q_{n,i}\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n, \end{aligned} \quad (50)$$

further,

$$\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \le \frac{1}{1-\beta(1-\mu_n)} < \frac{1}{\mu_n}, \quad (51)$$

due to $0 \le \beta < 1$, and (a) is due to $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \ge \mathbb{W}^{\langle 0,\delta_{N-n}\rangle}_i$ for any i ($1 \le i \le n-1$) and $\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n \ge \mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_i$ for any i ($n+1 \le i \le N$).

Step 4: For the n-th equation of (24) stated as follows

$$(\boldsymbol{e}^{\mathsf{T}}_n - \beta(1-\mu_n)\boldsymbol{e}^{\mathsf{T}}_n\boldsymbol{Q})\boldsymbol{w}_1 = 1, \quad (52)$$

we first subtract $\mu_n\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n$ from both LHS and RHS of (52), and then divide (52) by $1-\mu_n$. As a consequence, (52) can be written as follows

$$(\boldsymbol{e}^{\mathsf{T}}_n - \beta\boldsymbol{e}^{\mathsf{T}}_n\boldsymbol{Q})\boldsymbol{w}_1 = \frac{1-\mu_n\mathbb{W}^{\langle 1,\delta_{N-n}\rangle}_n}{1-\mu_n}. \quad (53)$$

Combined with the other $N-1$ equations of (24), then (24) can be transformed to the following

$$(\boldsymbol{I}_N-\beta\boldsymbol{M}_1)\cdot\boldsymbol{w}_1=\boldsymbol{1}_N-\boldsymbol{1}_N^{n-1}$$
$$\Leftrightarrow(\boldsymbol{I}_N-\beta\boldsymbol{M}_0)\cdot\boldsymbol{w}_1=\boldsymbol{1}_N-\boldsymbol{1}_N^{n}+\frac{1-\mu_n\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}}{1-\mu_n}\boldsymbol{e}_n. \tag{54}$$

Thus,

$$\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}=\boldsymbol{e}_n^{\mathsf{T}}\big(\boldsymbol{I}_N-\beta\boldsymbol{M}_0\big)^{-1}\times(\boldsymbol{1}_N-\boldsymbol{1}_N^{n}+\frac{1-\mu_n\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}}{1-\mu_n}\boldsymbol{e}_n), \tag{55}$$

combined with

$$\mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}=\boldsymbol{e}_n^{\mathsf{T}}(\boldsymbol{I}_N-\beta\boldsymbol{M}_0)^{-1}(\boldsymbol{1}_N-\boldsymbol{1}_N^{n}),$$

then we have

$$\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}-\mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}=\frac{1-\mu_n\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}}{1-\mu_n}\Big[\boldsymbol{e}_n^{\mathsf{T}}\big(\boldsymbol{I}_N-\beta\boldsymbol{M}_0\big)^{-1}\boldsymbol{e}_n\Big]\overset{(a)}{>}0, \tag{56}$$

where (a) is due to $\mu_n\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}<1$ and $\boldsymbol{e}_n^{\mathsf{T}}\big(\boldsymbol{I}_N-\beta\boldsymbol{M}_0\big)^{-1}\boldsymbol{e}_n>0$. Note that $\boldsymbol{e}_n^{\mathsf{T}}\big(\boldsymbol{I}_N-\beta\boldsymbol{M}_0\big)^{-1}\boldsymbol{e}_n>0$ because $\boldsymbol{I}_N-\beta\boldsymbol{M}_0$ is a diagonally dominant matrix and every element in the diagonal line is larger than 0 when $0\le\beta<1$.

To the end, we prove $\mathbb{W}_n^{\langle 1,\delta_{N-n}\rangle}-\mathbb{W}_n^{\langle 0,\delta_{N-n}\rangle}>0$. Following the similar deduction, we can easily prove $\mathbb{W}_{n+1}^{\langle 1,\delta_{N-n}\rangle}-\mathbb{W}_{n+1}^{\langle 0,\delta_{N-n}\rangle}>0$. Therefore, we complete the proof.

Appendix III
Proof of Theorem 2

According to the definition of the Whittle index, we prove the indexability by checking $\nu_1<\nu_2<\cdots<\nu_N$. When $\beta=1$, we have $\nu_N=\frac{c\mu_N}{1-\beta}\to\infty$.

First, we have

$$\begin{aligned}f(n)&=\beta q_{n,n+1}\left(1-\frac{\frac{1}{\bar\mu_n}-\beta\lambda_n}{\frac{1}{\bar\mu_{n+1}}-\beta\lambda_n}\right)+\mu_n d_{n-1,n+1}K_{n+1}\\&\quad+\beta\sum_{i=n+2}^{N}q_{n,i}\left(1-\prod_{j=n+1}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}\right)\\&\quad+\mu_n\sum_{i=n+2}^{N}d_{n-1,i}K_i,\end{aligned}$$

and

$$\begin{aligned}f(n+1)&=\beta\sum_{i=n+2}^{N}q_{n+1,i}\left(1-\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}\right)\\&\quad+\mu_{n+1}\sum_{i=n+2}^{N}d_{n,i}K_i\\&=\beta\sum_{i=n+2}^{N}q_{n,i}\left(1-\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}\right)\\&\quad+\mu_{n+1}\sum_{i=n+2}^{N}d_{n-1,i}K_i.\end{aligned}$$

Further,

$$\begin{aligned}&f(n)-f(n+1)\\&=\beta q_{n,n+1}K_{n+1}+(\mu_n-\mu_{n+1})\sum_{i=n+2}^{N}d_{n-1,i}K_i\\&\quad+\mu_n d_{n-1,n+1}K_{n+1}+\beta\sum_{i=n+2}^{N}q_{n,i}\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}K_{n+1}\\&=\beta q_{n,n+1}K_{n+1}+(\mu_n-\mu_{n+1})\sum_{i=n+2}^{N}d_{n-1,i}K_i\\&\quad-\mu_n\beta\Big(q_{n-1,n+1}+\sum_{i=n+2}^{N}q_{n-1,i}\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}\Big)K_{n+1}\\&\quad+\beta\sum_{i=n+2}^{N}q_{n,i}\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}K_{n+1}\\&=\beta q_{n,n+1}K_{n+1}+(\mu_n-\mu_{n+1})\sum_{i=n+2}^{N}d_{n-1,i}K_i\\&\quad-\mu_n\beta\Big(q_{n,n+1}+\sum_{i=n+2}^{N}q_{n,i}\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}\Big)K_{n+1}\\&\quad+\beta\sum_{i=n+2}^{N}q_{n,i}\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}K_{n+1}\\&=\bar\mu_n\beta q_{n,n+1}K_{n+1}+(\mu_n-\mu_{n+1})\sum_{i=n+2}^{N}d_{n-1,i}K_i\\&\quad+\bar\mu_n\beta\sum_{i=n+2}^{N}q_{n,i}\prod_{j=n+2}^{i}\frac{\frac{1}{\bar\mu_{j-1}}-\beta\lambda_{j-1}}{\frac{1}{\bar\mu_j}-\beta\lambda_{j-1}}K_{n+1}\\&\overset{(a)}{\geq}0,\end{aligned}$$

where (a) is due to $d_{n-1,i}\le 0$, $K_i\ge 0$, and $\mu_n<\mu_{n+1}$.

Hence,

$$\begin{aligned}\nu_n=\frac{c\mu_n}{1-\beta+f(n)}&\le\frac{c\mu_n}{1-\beta+f(n+1)}\\&<\frac{c\mu_{n+1}}{1-\beta+f(n+1)}=\nu_{n+1},\end{aligned}$$

which completes the proof.

References

[1] K. Wang, L. Chen, J. Yu, and M. Win, "Opportunistic scheduling revisited using restless bandits: Indexability and index policy," in *Proc. GlobeCom*, Singapore, Dec. 2017, pp. 1–6.

[2] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE ICC*, Jun. 1995, pp. 331–335.

[3] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and A. Viterbi, "CDMA/HDR: A bandwidth efficient high speed wireless data service for nomadic users," *IEEE Commun. Mag.*, vol. 38, no. 7, pp. 70–77, Jul. 2000.

[4] H. J. Kushner and P. A. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1250–1259, Jul. 2004.

[5] S. Borst, "User-level performance of channel-aware scheduling algorithms in wireless data networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 3, pp. 636–647, Jun. 2005.

[6] S. Aalto and P. E. Lassila, "Flow-level stability and performance of channel-aware priority-based schedulers," in *Proc. 6th EURO-NGI Conf. Next Generation Internet*, Paris, France, Jul. 2010, pp. 1–8.

[7] T. Bonald, "A score-based opportunistic scheduler for fading radio channels," in *Proc. Eur. Wireless*, Jan. 2004, pp. 283–292.

[8] U. Ayesta, M. Erausquin, M. Jonckheere, and I. M. Verloop, "Scheduling in a random environment: Stability and asymptotic optimality," *IEEE/ACM Trans. Netw.*, vol. 21, no. 1, pp. 258–271, Feb. 2013.

[9] J. Kim, B. Kim, J. Kim, and Y. H. Bae, "Stability of flow-level scheduling with Markovian time-varying channels," *Perform. Eval.*, vol. 70, no. 2, pp. 148–159, Feb. 2013.

[10] S. Aalto, P. Lassila, and P. Osti, "Whittle index approach to size-aware scheduling with time-varying channels," in *Proc. ACM Int. Conf. Meas. Modeling Comput. Syst.*, Portland, OR, USA, Jun. 2015, pp. 57–69.

[11] S. Aalto, P. Lassila, and P. Osti, "Whittle index approach to size-aware scheduling for time-varying channels with multiple states," *Queueing Syst.*, vol. 83, nos. 3–4, pp. 195–225, Aug. 2016.

[12] F. Cecchi and P. Jacko, "Scheduling of users with Markovian time-varying transmission rates," in *Proc. Int. Conf. Meas. Modeling Comput. Syst.*, Pittsburgh, PA, USA, Jun. 2013, pp. 129–140.

[13] P. Jacko, "Value of information in optimal flow-level scheduling of users with Markovian time-varying channels," *Perform. Eval.*, vol. 68, no. 11, pp. 1022–1036, Nov. 2011.

[14] F. Cecchi and P. Jacko, "Nearly-optimal scheduling of users with Markovian time-varying transmission rates," *Perform. Eval.*, vols. 99–100, pp. 16–36, May/Jun. 2016.

[15] K. Wang, Q. Liu, Q. Fan, and Q. Ai, "Optimally probing channel in opportunistic spectrum access," *IEEE Commun. Lett.*, vol. 22, no. 7, pp. 1426–1429, Jul. 2018.

[16] K. Wang *et al.*, "On optimality of second-highest policy for opportunistic multichannel access," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12013–12024, Dec. 2018.

[17] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Probab.*, vol. 25, pp. 287–298, Jan. 1988.

[18] U. Ayesta, M. Erausquin, and P. Jacko, "A modeling framework for optimizing the flow-level scheduling with time-varying channels," *Perform. Eval.*, vol. 67, no. 11, pp. 1014–1029, Nov. 2010.

[19] P. Jacko, E. Morozov, L. Potakhina, and I. M. Verloop, "Maximal flow-level stability of best-rate schedulers in heterogeneous wireless systems," *Trans. Emerg. Telecommun. Technol.*, vol. 28, no. 1, Jan. 2017, Art. no. e2930. doi: 10.1002/ett.2930.

[20] I. Taboada, F. Liberal, and P. Jacko, "An opportunistic and non-anticipating size-aware scheduling proposal for mean holding cost minimization in time-varying channels," *Perform. Eval.*, vol. 79, pp. 90–103, Sep. 2014.

[21] I. Taboada, P. Jacko, U. Ayesta, and F. Liberal, "Opportunistic scheduling of flows with general size distribution in wireless time-varying channels," in *Proc. ITC*, Sep. 2014, pp. 1–9.

[22] S. Aalto, P. Lassila, A. Penttinen, and P. Osti, "On the optimal trade-off between SRPT and opportunistic scheduling," in *Proc. ACM Sigmetrics Joint Int. Conf. Meas. Modeling Comput. Syst.*, San Jose, CA, USA, Jun. 2011, pp. 185–196.

[23] T. Bonald, S. Borst, N. Hegde, M. Jonckheere, and A. Proutiere, "Flow-level performance and capacity of wireless networks with user mobility," *Queueing Syst.*, vol. 63, pp. 131–164, Dec. 2009.

[24] S. Borst and M. Jonckheere, "Flow-level stability of channel-aware scheduling algorithms," in *Proc. WiOpt*, Mar. 2006, pp. 1–6.

[25] P. van de Ven, S. Borst, and S. Shneer, "Instability of maxweight scheduling algorithms," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2009, pp. 1701–1709.

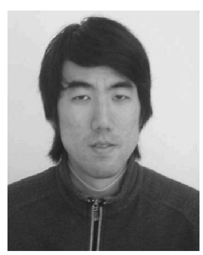
[26] C. Buyukkoc, P. Varaya, and J. Walrand, "The $c\mu$ rule revisited," *Adv. Appl. Probab.*, vol. 17, no. 1, pp. 237–238, 1985.

[27] J. Niño-Mora, "Characterization and computation of restless bandit marginal productivity indices," in *Proc. 2nd Int. Conf. Perform. Eval. Methodol. Tools*, Oct. 2007, Art. no. 74.

[28] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution: From Theory to Practice*. New York, NY, USA: Wiley, 2011.

Kehao Wang received the B.S. degree in electrical engineering, the M.S. degree in communication and information system from the Wuhan University of Technology, Wuhan, China, in 2003 and 2006, respectively, and the Ph.D. degree from the Department of Computer Science, University of Paris-Sud XI, Orsay, France, in 2012. From February 2013 to August 2013, he held a post-doctoral position at The HongKong Polytechnic University. In 2013, he joined the School of Information Engineering, Wuhan University of Technology, where he is currently an Associate Professor. Since December 2015, he has been a Visiting Scholar with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA. His research interests include stochastic optimization, operation research, scheduling, wireless network communications, and embedded operating systems.

Jihong Yu received the B.E. degree in communication engineering and the M.E. degree in communication and information systems from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2010 and 2013, respectively, and the Ph.D. degree in computer science from the University of Paris-Sud, Orsay, France, in 2016. He has been a Research Fellow with the School of Computing Science, Simon Fraser University, Canada, since 2017. He is currently an Associate Professor with the Beijing Institute of Technology. His research interests include RFID technologies, wireless communications, and the Internet of Things.



Lin Chen received the B.E. degree in radio engineering from Southeast University, China, in 2002, the Engineer Diploma and Ph.D. degrees from Telecom ParisTech, Paris, in 2005 and 2008, respectively, and the M.S. degree in networking from the University of Paris 6. He is currently a Professor with the School of Data Computer Science, Sun Yat-sen University, Guangzhou, China. His main research interests include modeling and control for wireless networks, security and cooperation enforcement in wireless networks, and game theory.

Pan Zhou (S'07–M'14) received the B.S. degree from the Advanced Class, Huazhong University of Science and Technology (HUST), Wuhan, China, in 2006, the M.S. degree from the Department of Electronics and Information Engineering, HUST, in 2008, and the Ph.D. degree from the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech.), Atlanta, USA, in 2011. He is currently an Associate Professor with the School of Electronic Information and Communications, HUST, Wuhan, China. He was a Senior Technical Member of Oracle Inc., USA, from 2011 to 2013, where he was involved on Hadoop and distributed storage systems for big data analytics with the Oralce Cloud Platform. His current research interests include big data analytics and machine learning, security and privacy, and information networks. He held an Honorary degree in the bachelor's degree and the Merit Research Award from HUST in the master's degree.

Xiaohu Ge (M'09–SM'11) received the Ph.D. degree in communication and information engineering from Huazhong University of Science and Technology (HUST), China, in 2003. He is currently a Full Professor with the School of Electronic Information and Communications, HUST. He is an Adjunct Professor with the Faculty of Engineering and Information Technology, University of Technology Sydney (UTS), Australia, He has been with HUST since 2005. Prior to that, he was a Researcher with Ajou University, South Korea, and the Politecnico Di Torino, Italy, from January 2004 to 2005. His research interests include the areas of mobile communications, traffic modeling in wireless networks, green communications, and interference modeling in wireless communications. He has published more than 200 papers in refereed journals and conference proceedings. He has been granted about 25 patents in China. He was a recipient of the Best Paper Award from the IEEE GLOBECOM 2010. He has served as the General Chair for the 2015 IEEE International Conference on Green Computing and Communications (IEEE GreenCom 2015). He serves as an Associate Editor for the IEEE WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and IEEE ACCESS.

Moe Z. Win (S'85–M'87–SM'97–F'04) is currently a Professor with the Massachusetts Institute of Technology (MIT) and the Founding Director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and the NASA Jet Propulsion Laboratory. His current research interests include network localization and navigation, network interference exploitation, and quantum information science. He is an Elected Fellow of the AAAS and the IET. He was a recipient of two IEEE technical field awards, including the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (jointly with R. A. Scholtz in 2006). Other recognitions include the U.S. Presidential Early Career Award for Scientists and Engineers in 2004, the Laurea Honoris Causa from the Università degli Studi di Ferrara in 2008, the Copernicus Fellowship in 2011, the International Prize for Communications Cristoforo Colombo in 2013, and the IEEE Communications Society Edwin H. Armstrong Achievement Award in 2016. He has served the IEEE Communications Society as an Elected Member-at-Large on the Board of Governors, the Elected Chair of the Radio Communications Committee, and the IEEE Distinguished Lecturer. He is an ISI Highly Cited Researcher.