This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Aalto, Samuli; Lassila, Pasi; Taboada, Ianire

Whittle index approach to opportunistic scheduling with partial channel information

Published in:
Performance Evaluation

DOI:
[10.1016/j.peva.2019.102052](https://doi.org/10.1016/j.peva.2019.102052)

Published: 01/12/2019

Document Version
Peer reviewed version

Published under the following license:
CC BY-NC-ND

Please cite the original version:
Aalto, S., Lassila, P., & Taboada, I. (2019). Whittle index approach to opportunistic scheduling with partial channel information. *Performance Evaluation*, 136, [102052]. <https://doi.org/10.1016/j.peva.2019.102052>

Whittle index approach to opportunistic scheduling with partial channel information

Samuli Aalto, Pasi Lassila

Department of Communications and Networking, Aalto University, Finland

Ianire Taboada

University of the Basque Country (UPV/EHU), Spain

Abstract

Opportunistic scheduling in wireless cellular systems utilizes random channel quality variations in time by favoring the users with good channel conditions. However, the success of such schedulers is heavily depending on the accuracy of the available information on the channel states of users. In this paper, we consider the opportunistic scheduling problem of downlink data traffic with partial channel information, where the target is to minimize the flow-level holding costs. In earlier works, the Whittle index approach has successfully been utilized to develop near-optimal scheduling policies for the corresponding problems, however, typically with exact channel information. Using the same approach, we complement and extend the results found thus far. More specifically said, our novel contributions are (i) proving that the flow-level opportunistic scheduling problem with partial channel information is indexable at least in certain parts of the parameter space; (ii) deriving an explicit formula for the corresponding Whittle index; (iii) proving that, in these parts of the parameter space, the optimal policy for the relaxed optimization problem is of threshold type; and (iv) demonstrating that, in the remaining parts of the parameter space, it is possible that the optimal policy for the relaxed problem is not of threshold type. In addition, we evaluate the performance of the derived Whittle index policies and compare them with some greedy policies by numerical simulations.

Keywords: Whittle index, wireless cellular networks, opportunistic scheduling, partial channel information, POMDP

1. Introduction

The channel quality in wireless cellular systems is inherently randomly varying in time. Opportunistic schedulers utilize such variations by favoring the users with good channel conditions [19, 10, 17, 33, 7]. However, their success is heavily depending on the accuracy of the available information on the channel states of users.

In this paper, we consider the opportunistic scheduling problem of down-link data traffic with *partial channel information*, where the target is to minimize the flow-level holding costs. In line with many related papers [25, 22, 5, 24, 16, 27], we assume that the scheduler estimates the channel state based on ARQ-type feedback from the *scheduled* users. Thus, from the users that are not scheduled in the time slot, the scheduler does not get any new channel state information. The wireless channel dynamics are described by discrete-time Markov processes, and the traffic consists of elastic flows corresponding to downloading of files using TCP.

Such a flow-level opportunistic scheduling problem with IID or Markovian channel dynamics has earlier been studied under the assumption of *exact channel information* [8, 14, 12, 9, 18, 30, 32, 1, 2, 13, 34, 15, 3, 6]. Near-optimal scheduling policies have successfully been developed by applying the *Whittle index approach*, which was originally developed for the restless bandits problems [36]. In this approach, the original (mathematically intractable) problem is first relaxed by allowing the scheduler to choose multiple users in the same time slot and only requiring that the time slot is given on average for at most one user, which makes the problem more tractable by decomposing it to separate user-related subproblems. While the problem becomes separable, it is, however, not clear, a priori, whether the problem is indexable, or not. If the problem is indexable, the resulting Whittle index policy is known to be asymptotically optimal, at least under certain technical conditions [35].

Under the more realistic assumption of *partial channel information*, the main emphasis has been on the opportunistic scheduling problem where the target is to maximize the average or discounted packet-level throughput of a fixed user population [20, 37, 5, 25, 22, 24, 23, 28, 26, 27, 21]. With partial information, the problem can be formulated in the framework of Partially Observable Markov Decision Processes. Even here, the Whittle index approach has been used in many papers to develop near-optimal scheduling policies [25, 22, 24, 23, 28, 26, 27, 21].

The main difference between the maximization of packet-level throughput and the minimization of flow-level holding costs arises from the fact that in the previous problem a user stays in the system forever, while in the latter one, a user leaves the system as soon as the service of the corresponding flow has been completed. So the cost structure is different and, therefore, the solutions, as well. In particular, the optimality and near-optimality results for packet-level models found in [20, 37, 5, 25, 22, 24, 23, 28, 26, 27, 21] are not applicable in our flow-level problem.

The flow-level opportunistic scheduling problem with partial channel information has, thus far, only been considered in [16, 31]. In [16], Jacko and Villar assume ARQ-type feedback and two-state positively autocorrelated Markovian channels. By applying the Whittle index approach, the scheduling problem is decomposed into separate user-related subproblems. However, they are not able to prove but just *conjecture* that the subproblems are indexable. Under this conjecture and some additional assumptions concerning the model parameters, they establish the optimality of threshold type policies for the subproblems and also present a formula for the corresponding Whittle index. However, the formula for the Whittle index given in [16] is not fully explicit. In [31], Taboada et al. study a slightly different model related to periodic feedback relying on the Whittle index approach. They allow multi-state Markovian channels but present only an approximate analysis, where the indexability property is again conjectured.

In this paper, we consider the model introduced in [16] complementing and extending the results given there. More precisely said, our contributions in the present paper are as follows:

- (i) We prove that the flow-level opportunistic scheduling problem with partial channel information is indexable at least in certain parts of the parameter space.
- (ii) We derive an explicit formula for the corresponding Whittle index.
- (iii) We prove that, in these parts of the parameter space, the optimal policy for the relaxed optimization problem is of threshold type.
- (iv) We characterize the optimal policy even in the remaining parts of the parameter space, and demonstrate that it is, indeed, possible that the optimal policy for the relaxed problem is not of threshold type, which significantly complicates the task of formally proving the indexability.

Moreover, we evaluate the performance of the derived Whittle index policies and compare them with greedy policies by numerical simulations. The results show that the Whittle index policies perform well in all our scenarios.

This paper is an extended version of our earlier conference paper [4], where we focused on a certain part of the parameter space, while the present paper covers the whole parameter space.

The rest of the paper is organized as follows. The flow-level opportunistic scheduling problem with partial channel information is described in Section 2 and the mathematically more tractable relaxed version of the problem in Section 3. In Section 4, we introduce a class of threshold policies and characterize the optimal policy for the relaxed problem. It turns out that the optimal policy is not necessarily of the threshold type. However, in those cases where the optimal policy is provably of threshold type, the relaxed problem is shown to be indexable and the corresponding Whittle index is derived, which is done in Section 5. On the other hand, for those cases where the optimal policy is not necessarily of threshold type, we provide, in Section 6, concrete examples of the optimal policies, which conflict with the idea of threshold policies. In Section 7, we focus on the part of parameter space where the relaxed problem is indexable and define the Whittle index policy for the original opportunistic scheduling problem. Its performance is compared to some greedy policies by numerical simulations in Section 8. Section 9 concludes the paper.

2. Scheduling problem with partial channel information

We consider the following opportunistic scheduling problem with partial channel information in discrete time. The same problem was earlier introduced and partially solved in [16].

At time 0 (i.e., in the beginning of the first time slot), there are K users in the area of a wireless base station, each with a file (a.k.a. flow) to be downloaded. For any time slot $t \in \{0, 1, \dots\}$, the scheduler in the base station chooses (at most) M of the flows for service until all the flows have been downloaded. Denote $A_k^\pi(t) = 1$ if flow k is chosen in time slot t , where π refers to the scheduling policy used; otherwise $A_k^\pi(t) = 0$. Thus, for any policy π and time slot t , we have the constraint

$$\sum_{k=1}^K A_k^\pi(t) \leq M, \quad (1)$$

where $M \in \{1, 2, \dots, K-1\}$.

We assume that, for any user k , the wireless channel is described by a two-state discrete-time Markov chain $N_k(t) \in \{1, 2\}$. Let $q_{k,n,m} > 0$ denote the transition probability from channel state n to channel state m for user k . We assume that these Markov chains are independent of each other and, furthermore, *positively autocorrelated*, i.e.,

$$\rho_k = q_{k,2,2} - q_{k,1,2} > 0. \quad (2)$$

In addition, let $q_{k,n}$ denote the steady-state probability to be in channel state n ,

$$q_{k,1} = \frac{q_{k,2,1}}{q_{k,2,1} + q_{k,1,2}}, \quad q_{k,2} = \frac{q_{k,1,2}}{q_{k,2,1} + q_{k,1,2}}. \quad (3)$$

Finally, let $\mu_{k,n}$ denote the probability that flow k is completed in a time slot where it is scheduled and its channel state is *known* to be n . Without loss of generality, we may assume that

$$\mu_{k,2} \geq \mu_{k,1} > 0, \quad (4)$$

i.e., channel state 2 is better than channel state 1.

The scheduling decisions are, however, not based on full information but just partial information on the channel states of the users. We assume that any user scheduled in time slot t sends back its true channel state at the end of this time slot. This assumption leads to a Partially Observable Markov Decision Process (POMDP).

In the POMDP formulation, the state of user k is given by the *belief state* $X_k^\pi(t)$, which denotes the current probability to be in the good channel state 2, unless the flow has already been completed, which is denoted by a special symbol '*'. The set of possible belief states for each user k is countably infinite:

$$\mathcal{S}_k = \{x_{k,2,j} : j \in \{0, 1, \dots\}\} \cup \{q_{k,2}\} \cup \{x_{k,1,j} : j \in \{0, 1, \dots\}\} \cup \{*\},$$

where $x_{k,2,0} = q_{k,2,2}$, $x_{k,1,0} = q_{k,1,2}$, and

$$x_{k,n,j+1} = x_{k,n,j}q_{k,2,2} + (1 - x_{k,n,j})q_{k,1,2}. \quad (5)$$

Note that user k with an uncompleted flow is in belief state $x_{k,n,j}$ if and only if this user was last time scheduled for $j+1$ slots ago and its channel state

was then known to be n . When the flow of user k is completed, it will enter the absorbing belief state $*$.¹

It is easy to show by induction that

$$x_{k,2,j} = q_{k,2}(1 - \rho_k^{j+1}) + \rho_k^{j+1}, \quad x_{k,1,j} = q_{k,2}(1 - \rho_k^{j+1}). \quad (6)$$

Now it follows from our assumption (2) that

$$x_{k,2,0} > x_{k,2,1} > \dots > q_{k,2} > \dots > x_{k,1,1} > x_{k,1,0}. \quad (7)$$

Recall also that belief state $q_{k,2}$ refers to the steady-state probability of channel state 2.

Let $\mu_k(x)$ denote the probability that flow k is completed in a time slot where it is scheduled and its belief state is $x \in \mathcal{S}_k \setminus \{*\}$. We adopt the same *rate adaptation* scheme as in [16]: If flow k is scheduled in belief state x for which $x > \mu_{k,1}/\mu_{k,2}$, the base station transmits with the best rate. In such a case, the flow is completed with probability $x\mu_{k,2}$. But if $x \leq \mu_{k,1}/\mu_{k,2}$, the base station transmits with the worst rate so that the flow is completed with probability $\mu_{k,1}$. In other words, the flow completion probability if scheduled in belief state x is given by

$$\mu_k(x) = \max\{x\mu_{k,2}, \mu_{k,1}\}. \quad (8)$$

In addition, let $\mu_k(x, n)$ denote the probability that flow k is completed in a time slot where it is scheduled, its belief state is $x \in \mathcal{S}_k \setminus \{*\}$, and the channel state is n . According to our rate adaptation scheme, we have

$$\begin{aligned} \mu_k(x, 2) &= \begin{cases} \mu_{k,2}, & \text{if } x > \mu_{k,1}/\mu_{k,2}, \\ \mu_{k,1}, & \text{otherwise;} \end{cases} \\ \mu_k(x, 1) &= \begin{cases} 0, & \text{if } x > \mu_{k,1}/\mu_{k,2}, \\ \mu_{k,1}, & \text{otherwise.} \end{cases} \end{aligned}$$

Note that $\mu_k(x) = x\mu_k(x, 2) + (1 - x)\mu_k(x, 1)$.

For any user k , holding costs are accumulated at rate $h_k > 0$ until the whole file is downloaded and the flow is completed. Thus, the objective function in our scheduling problem is given by

$$E \left[\sum_{t=0}^{\infty} \sum_{k=1}^K h_k 1_{\{X_k^\pi(t) \neq *\}} \right]. \quad (9)$$

¹This is a characteristic property of flow-level opportunistic scheduling models, which the corresponding packet-level models lack.

The aim is to find the optimal scheduling policy π that minimizes the expected total holding costs (9) subject to the strict capacity constraint (1) for all t and assuming that the scheduling decisions at each time slot t are based on the belief states $X_k^\pi(t)$ of users.

3. Relaxed optimization problem

The original optimization problem described in the previous section belongs to the class of *restless bandit* problems, which are known to be mathematically intractable. Jacko and Villar [16] followed Whittle's approach [36] to develop near-optimal heuristic solutions for the problem: They modified the original problem by replacing the strict capacity constraint (1) by an averaged one and approached the relaxed problem by the Lagrangian methods. In this paper, we use the same approach, which results in the following (separate) subproblems for each user k : Find the optimal policy π that minimizes the objective function

$$f_k^\pi + \nu g_k^\pi, \quad (10)$$

where ν can be interpreted as the unit price of work, f_k^π as the expected total holding costs of user k , and g_k^π as the expected total amount of work needed for user k ,

$$f_k^\pi = E \left[\sum_{t=0}^{\infty} h_k 1_{\{X_k^\pi(t) \neq *\}} \right], \quad g_k^\pi = E \left[\sum_{t=0}^{\infty} A_k^\pi(t) \right].$$

The separable subproblems of the Lagrangian version of the relaxed scheduling problem are now considered in the context of partially observed Markov decision processes. The possible actions $a \in \mathcal{A} = \{0, 1\}$ are “to schedule” ($a = 1$) and “not to schedule” ($a = 0$). For any belief state $x \in \mathcal{S}_k$, both actions $a \in \{0, 1\}$ are allowed.

Let $p_k(y|x, a) \geq 0$ denote the transition probability from belief state $x \in \mathcal{S}_k$ to belief state $y \in \mathcal{S}_k$ after action $a \in \mathcal{A}$. It follows from the previous discussion that the non-zero transition probabilities are as follows:

$$\begin{aligned} p_k(x_{k,n,j+1}|x_{k,n,j}, 0) &= 1, & n \in \{1, 2\}, j \in \{0, 1, \dots\}, \\ p_k(q_{k,2}|q_{k,2}, 0) &= 1, \\ p_k(x_{k,2,0}|x, 1) &= x(1 - \mu_k(x, 2)), & x \in \mathcal{S}_k \setminus \{*\}, \\ p_k(x_{k,1,0}|x, 1) &= (1 - x)(1 - \mu_k(x, 1)), & x \in \mathcal{S}_k \setminus \{*\}, \\ p_k(*|x, 1) &= \mu_k(x), & x \in \mathcal{S}_k \setminus \{*\}, \\ p_k(*|*, 0) &= p_k(*|*, 1) = 1. \end{aligned} \quad (11)$$

Note that $*$ is an absorbing belief state for any policy.

Finally, let $c_k(x, a)$ denote the immediate cost in belief state $x \in \mathcal{S}_k$ after action $a \in \mathcal{A}$. In our model,

$$\begin{aligned} c_k(x, 0) &= h_k, & x \in \mathcal{S}_k \setminus \{*\}, \\ c_k(x, 1) &= h_k + \nu, & x \in \mathcal{S}_k \setminus \{*\}, \\ c_k(*, 0) &= 0, \\ c_k(*, 1) &= \nu. \end{aligned} \tag{12}$$

Thus, the state space $\mathcal{S}_k$ is discrete, the action space $\mathcal{A}$ is finite, and the immediate costs are bounded.

Note that if $\nu > 0$, then the expected total cost (10) remains finite (instead of increasing without limits) only for those policies π that finally end up in state $*$ and choose action 0 there. In addition, if $\nu = 0$, then the expected total cost (10) remains again finite only for those policies π that finally end up in state $*$ but which can choose either action there. Finally, if $\nu < 0$, the expected total costs (10) can be taken to $-\infty$ by finally ending up in state $*$ and choosing action 1 there.

Now, to get a well-defined optimization problem for any ν , we modify the objective function from the expected total cost (10) to the *expected long-run average cost*: Find the optimal policy π that minimizes the objective function

$$\tilde{f}_k^\pi + \nu \tilde{g}_k^\pi, \tag{13}$$

where ν can still be interpreted as the unit price of work but $\tilde{f}_k^\pi$ refers to the expected long-run average holding costs per time slot of user k and $\tilde{g}_k^\pi$ to the expected long-run average amount of work per time slot needed for user k ,

$$\tilde{f}_k^\pi = \lim_{T \rightarrow \infty} E \left[\frac{1}{T} \sum_{t=0}^T h_k 1_{\{X_k^\pi(t) \neq *\}} \right], \quad \tilde{g}_k^\pi = \lim_{T \rightarrow \infty} E \left[\frac{1}{T} \sum_{t=0}^T A_k^\pi(t) \right].$$

Let $V_k(x; \nu)$ denote the *value function* for belief state $x \in \mathcal{S}_k$ related to the minimization of the expected long-run average cost (13) with Lagrangian parameter ν , and $\bar{c}_k$ the corresponding optimal average cost (per time slot).

The *optimality equations* [29, Thm. 6.7] in this model read as follows:

$$\begin{aligned}
\bar{c}_k + V_k(*; \nu) &= \min\{0, \nu\} + V_k(*; \nu), \\
\bar{c}_k + V_k(q_{k,2}; \nu) &= h_k + \min\{V_k(q_{k,2}; \nu), \\
&\quad \nu + q_{k,2}(1 - \mu_k(q_{k,2}, 2))V_k(x_{k,2,0}; \nu) + \\
&\quad (1 - q_{k,2})(1 - \mu_k(q_{k,2}, 1))V_k(x_{k,1,0}; \nu)\}, \\
\bar{c}_k + V_k(x_{k,n,j}; \nu) &= h_k + \min\{V_k(x_{k,n,j+1}; \nu), \\
&\quad \nu + x_{k,n,j}(1 - \mu_k(x_{k,n,j}, 2))V_k(x_{k,2,0}; \nu) + \\
&\quad (1 - x_{k,n,j})(1 - \mu_k(x_{k,n,j}, 1))V_k(x_{k,1,0}; \nu)\}.
\end{aligned} \tag{14}$$

In addition, let $V_k^\pi(x; \nu)$ denote the corresponding value function for policy π . The policy π_k^* for which

$$V_k^{\pi_k^*}(x; \nu) = V_k(x; \nu)$$

for all $x \in \mathcal{S}_k$ is said to be ν -*optimal* for user k .

From optimality equations (14), we see that if $\nu \geq 0$, then the minimum expected long-run average cost (13) equals clearly to $\bar{c}_k = 0$ and is achieved by those policies π that finally end up in state $*$ and choose action 0 there (or even action 1 if $\nu = 0$). On the other hand, if $\nu < 0$, the minimum expected long-run average cost (13) equals to $\bar{c}_k = \nu$ and is achieved by those policies π that finally end up in state $*$ and choose action 1 there. However, from optimality equations (14), we also see that, in this case with $\nu < 0$, the only ν -optimal policy is such that user k is scheduled in each time slot. In other words, if $\nu < 0$, there is no need to postpone the scheduling decision but user k should be scheduled in each time slot until its completion.

Thus, from this on, we assume that $\nu \geq 0$. In this case with $\nu \geq 0$, the optimality equations (14) read as follows:

$$\begin{aligned}
V_k(*; \nu) &= 0, \\
V_k(q_{k,2}; \nu) &= h_k + \min\{V_k(q_{k,2}; \nu), \\
&\quad \nu + q_{k,2}(1 - \mu_k(q_{k,2}, 2))V_k(x_{k,2,0}; \nu) + \\
&\quad (1 - q_{k,2})(1 - \mu_k(q_{k,2}, 1))V_k(x_{k,1,0}; \nu)\}, \\
V_k(x_{k,n,j}; \nu) &= h_k + \min\{V_k(x_{k,n,j+1}; \nu), \\
&\quad \nu + x_{k,n,j}(1 - \mu_k(x_{k,n,j}, 2))V_k(x_{k,2,0}; \nu) + \\
&\quad (1 - x_{k,n,j})(1 - \mu_k(x_{k,n,j}, 1))V_k(x_{k,1,0}; \nu)\}.
\end{aligned} \tag{15}$$

Clearly, the last equation can also be written as follows:

$$\begin{aligned}
V(x_{k,n,j}; \nu) = \min\{ & h_k + \nu + x_{k,n,j}(1 - \mu_k(x_{k,n,j}, 2))V_k(x_{k,2,0}; \nu) + \\
& (1 - x_{k,n,j})(1 - \mu_k(x_{k,n,j}, 1))V_k(x_{k,1,0}; \nu), \\
& 2h_k + \nu + x_{k,n+1,j}(1 - \mu_k(x_{k,n+1,j}, 2))V_k(x_{k,2,0}; \nu) + \\
& (1 - x_{k,n+1,j})(1 - \mu_k(x_{k,n+1,j}, 1))V_k(x_{k,1,0}; \nu), \quad (16) \\
& 3h_k + \nu + x_{k,n+2,j}(1 - \mu_k(x_{k,n+2,j}, 2))V_k(x_{k,2,0}; \nu) + \\
& (1 - x_{k,n+2,j})(1 - \mu_k(x_{k,n+2,j}, 1))V_k(x_{k,1,0}; \nu), \\
& \dots\}.
\end{aligned}$$

As a final remark of this section, we note that the optimality equations (15) related to the optimal expected long-run average cost (13) are exactly the same as we would have if we would use the expected total cost (10) as our objective function under assumption $\nu \geq 0$.

4. Characterization of the optimal policy for the relaxed problem

In this section, we characterize the ν -optimal policy π_k^* for the relaxed problem (13) related to user k assuming that $\nu \geq 0$. We first define a family of scheduling policies, called threshold policies. Then we give seven auxiliary results (Lemma 1 and Propositions 1-6) needed for the characterization given in Theorem 1, which is the main result of this section.

For the definition of threshold policies, we recall that, due to our assumption (2), we have the following ordering among the belief states:

$$0 < x_{k,1,0} < x_{k,1,1} < \dots < q_{k,2} < \dots < x_{k,2,1} < x_{k,2,0} < 1.$$

Definition 1. Let $\pi(\theta)$, where $\theta \in [0, 1]$, refer to the policy π such that

$$a^\pi(x) = \begin{cases} 0, & \text{if } x < \theta \text{ or } x = *, \\ 1, & \text{otherwise.} \end{cases} \quad (17)$$

In addition, let Π^T denote the whole family of such policies

$$\Pi^T = \{\pi(\theta) \mid \theta \in [0, 1]\}, \quad (18)$$

which are called *threshold policies*.

Note that, for policy $\pi(x_{k,1,0})$, user k is scheduled in any belief state $x \in \mathcal{S}_k \setminus \{*\}$, while for policy $\pi(1)$, user k is not scheduled in any belief state $x \in \mathcal{S}_k$. Note also that, in the rest of this section, we leave out the user index k to simplify notations. For example, the value function $V_k(x; \nu)$ is briefly referred to by $V(x; \nu)$ and so on.

Lemma 1. *If $a \leq b$ and $0 < q \leq p < 1$, then*

$$pa + (1 - p)b \leq qa + (1 - q)b.$$

The proof of Lemma 1 is elementary and, thus, omitted here.

Proposition 1. *If $\nu \geq 0$, $V(x; \nu) < \infty$ for all $x \in \mathcal{S}$ and $V(x_{2,0}; \nu) \leq V(x_{1,0}; \nu)$.*

Proof. 1° Assume first that $x_{2,0} \leq \mu_1/\mu_2$. It follows from (8) and (7) that under this assumption the base station transmits with the worst rate μ_1 in any belief state $x \in \mathcal{S} \setminus \{*\}$ so that

$$\mu(x) = \mu(x, 2) = \mu(x, 1) = \mu_1.$$

Thus, there is no reason to postpone the decisions to schedule but it is optimal to schedule the user in each time slot. In addition, since the completion trials in each time slot are independent, the value function for any $x \in \mathcal{S} \setminus \{*\}$ is equal to

$$V(x; \nu) = \frac{h + \nu}{\mu_1} < \infty, \quad (19)$$

which is also easy to verify from the optimality equations (15) in this case. In particular, we have $V(x_{2,0}; \nu) - V(x_{1,0}; \nu) = 0$.

2° Assume now that $x_{2,0} > \mu_1/\mu_2$. It follows from (8) and (7) that under this assumption the flow completion probability $\mu(x)$ in any belief state x is upper [lower] bounded by $x_{2,0}\mu_2$ [μ_1]. In addition, since the completion trials in each time slot are independent and $h + \nu \geq 0$, the value function for any $x \in \mathcal{S} \setminus \{*\}$ is lower and upper bounded by

$$\frac{h + \nu}{x_{2,0}\mu_2} \leq V(x; \nu) \leq \frac{h + \nu}{\mu_1} < \infty. \quad (20)$$

Since $x_{2,0} > \mu_1/\mu_2$, we also have

$$\mu(x_{2,0}) = x_{2,0}\mu_2, \quad \mu(x_{2,0}, 2) = \mu_2, \quad \mu(x_{2,0}, 1) = 0.$$

Thus, from the optimality equation (15), we get

$$V(x_{2,0}; \nu) \leq h + \nu + x_{2,0}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,0})V(x_{1,0}; \nu)$$

implying, together with the lower bound of (20), that

$$(1 - x_{2,0})(V(x_{2,0}; \nu) - V(x_{1,0}; \nu)) \leq h + \nu - x_{2,0}\mu_2 V(x_{2,0}; \nu) \leq 0.$$

Thus, $V(x_{2,0}; \nu) - V(x_{1,0}; \nu) \leq 0$, which completes the proof. $\square$

Proposition 2. *Let $\nu \geq 0$. If π is ν -optimal and $x_{2,j+1} > \mu_1/\mu_2$, then*

$$a^\pi(x_{2,j+1}) = 1 \quad \Rightarrow \quad a^\pi(x_{2,j}) = 1.$$

Proof. Assume that π is ν -optimal. Let $j \in \{0, 1, \dots\}$, and assume that $x_{2,j+1} > \mu_1/\mu_2$ and $a^\pi(x_{2,j+1}) = 1$. By (7), we deduce that $x_{2,j} > \mu_1/\mu_2$. Now if $a^\pi(x_{2,j}) = 0$, then it follows from the optimality equations (15), Proposition 1, and Lemma 1 that

$$\begin{aligned} V(x_{2,j+1}; \nu) &\leq \nu + x_{2,j}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,j})V(x_{1,0}; \nu) \\ &\leq \nu + x_{2,j+1}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,j+1})V(x_{1,0}; \nu) \\ &< h + \nu + x_{2,j+1}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,j+1})V(x_{1,0}; \nu) \\ &= V(x_{2,j+1}; \nu), \end{aligned}$$

which is a contradiction. Thus, we have $a^\pi(x_{2,j}) = 1$. $\square$

Proposition 3. *Let $\nu \geq 0$. If π is ν -optimal and $x_{2,j} \leq \mu_1/\mu_2$, then*

$$a^\pi(x_{2,j+1}) = 1 \quad \Rightarrow \quad a^\pi(x_{2,j}) = 1.$$

Proof. Assume that π is ν -optimal. Let $j \in \{0, 1, \dots\}$, and assume that $x_{2,j} \leq \mu_1/\mu_2$ and $a^\pi(x_{2,j+1}) = 1$. By (7), we deduce that $x_{2,j+1} \leq \mu_1/\mu_2$. Now if $a^\pi(x_{2,j}) = 0$, then it follows from the optimality equations (15), Proposition 1, and Lemma 1 that

$$\begin{aligned} V(x_{2,j+1}; \nu) &\leq \nu + x_{2,j}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{2,j})(1 - \mu_1)V(x_{1,0}; \nu) \\ &\leq \nu + x_{2,j+1}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{2,j+1})(1 - \mu_1)V(x_{1,0}; \nu) \\ &< h + \nu + x_{2,j+1}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{2,j+1})(1 - \mu_1)V(x_{1,0}; \nu) \\ &= V(x_{2,j+1}; \nu), \end{aligned}$$

which is a contradiction. Thus, we have $a^\pi(x_{2,j}) = 1$. $\square$

Proposition 4. *Let $\nu \geq 0$. If π is ν -optimal and $x_{1,j} > \mu_1/\mu_2$, then*

$$a^\pi(x_{1,j}) = 1 \quad \Rightarrow \quad a^\pi(x_{1,j+1}) = 1.$$

Proof. Assume that π is ν -optimal. Let $j \in \{0, 1, \dots\}$, and assume that $x_{1,j} > \mu_1/\mu_2$ and $a^\pi(x_{1,j}) = 1$. By (7), we deduce that $x_{1,k} > \mu_1/\mu_2$ for any $k > j$. Now it follows from the optimality equations (15) and (16) that

$$V(x_{1,j}; \nu) = h + \nu + x_{1,j}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,j})V(x_{1,0}; \nu)$$

and, for any $k > j$,

$$\begin{aligned} & h + \nu + x_{1,j}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,j})V(x_{1,0}; \nu) \\ & \leq (k - j + 1)h + \nu + x_{1,k}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,k})V(x_{1,0}; \nu). \end{aligned}$$

Thus,

$$(k - j)h + (x_{1,k} - x_{1,j})((1 - \mu_2)V(x_{2,0}; \nu) - V(x_{1,0}; \nu)) \geq 0.$$

In addition, from (6), it is easy to verify that, for any $k > j$,

$$x_{1,k} - x_{1,j} > x_{1,k+1} - x_{1,j+1} > 0.$$

By combining this with Proposition 1, we deduce that

$$\begin{aligned} & (k - j + 1)h + \nu + x_{1,k+1}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,k+1})V(x_{1,0}; \nu) \\ & - h - \nu - x_{1,j+1}(1 - \mu_2)V(x_{2,0}; \nu) - (1 - x_{1,j+1})V(x_{1,0}; \nu) \\ & = (k - j)h + (x_{1,k+1} - x_{1,j+1})((1 - \mu_2)V(x_{2,0}; \nu) - V(x_{1,0}; \nu)) \\ & > (k - j)h + (x_{1,k} - x_{1,j})((1 - \mu_2)V(x_{2,0}; \nu) - V(x_{1,0}; \nu)) \geq 0. \end{aligned}$$

Since this is true for any $k > j$, it follows from the optimality equation (16) that

$$V(x_{1,j+1}; \nu) = h + \nu + x_{1,j+1}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,j+1})V(x_{1,0}; \nu)$$

so that $a^\pi(x_{1,j+1}) = 1$. □

Proposition 5. *Let $\nu \geq 0$. If π is ν -optimal and $q_2 \leq \mu_1/\mu_2$, then*

$$a^\pi(x_{1,j}) = 1 \quad \Rightarrow \quad a^\pi(x_{1,j+1}) = 1.$$

Proof. Assume that π is ν -optimal and $q_2 \leq \mu_1/\mu_2$. By (7), we deduce that $x_{1,j} \leq \mu_1/\mu_2$ for any $j \geq 0$. Let then $j \in \{0, 1, \dots\}$, and assume that $a^\pi(x_{1,j}) = 1$. Now it follows from the optimality equations (15) and (16) that

$$V(x_{1,j}; \nu) = h + \nu + x_{1,j}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,j})(1 - \mu_1)V(x_{1,0}; \nu)$$

and, for any $k > j$,

$$\begin{aligned} & h + \nu + x_{1,j}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,j})(1 - \mu_1)V(x_{1,0}; \nu) \\ & \leq (k - j + 1)h + \nu + x_{1,k}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,k})(1 - \mu_1)V(x_{1,0}; \nu). \end{aligned}$$

Thus,

$$(k - j)h + (x_{1,k} - x_{1,j})(1 - \mu_1)(V(x_{2,0}; \nu) - V(x_{1,0}; \nu)) \geq 0.$$

In addition, from (6), it is easy to verify that, for any $k > j$,

$$x_{1,k} - x_{1,j} > x_{1,k+1} - x_{1,j+1} > 0.$$

By combining this with Proposition 1, we deduce that

$$\begin{aligned} & (k - j + 1)h + \nu + x_{1,k+1}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,k+1})(1 - \mu_1)V(x_{1,0}; \nu) \\ & \quad - h - \nu - x_{1,j+1}(1 - \mu_1)V(x_{2,0}; \nu) - (1 - x_{1,j+1})(1 - \mu_1)V(x_{1,0}; \nu) \\ & = (k - j)h + (x_{1,k+1} - x_{1,j+1})(1 - \mu_1)(V(x_{2,0}; \nu) - V(x_{1,0}; \nu)) \\ & > (k - j)h + (x_{1,k} - x_{1,j})(1 - \mu_1)(V(x_{2,0}; \nu) - V(x_{1,0}; \nu)) \geq 0. \end{aligned}$$

Since this is true for any $k > j$, it follows from the optimality equation (16) that

$$V(x_{1,j+1}; \nu) = h + \nu + x_{1,j+1}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,j+1})(1 - \mu_1)V(x_{1,0}; \nu)$$

so that $a^\pi(x_{1,j+1}) = 1$. □

Proposition 6. *Let $\nu \geq 0$. If policy π is ν -optimal, then*

- (i) $a^\pi(x_{2,j}) = 1$ for infinitely many j ,
- (ii) $a^\pi(q_2) = 1$, and
- (iii) $a^\pi(x_{1,j}) = 1$ for infinitely many j .

Proof. It is easy to see from the optimality equations (15) that if any of the three condition is violated, $V^\pi(x; \nu) = \infty$ for some belief state $x \in \mathcal{S} \setminus \{*\}$, which contradicts Proposition 1. $\square$

Now we are ready to characterize the ν -optimal policy for the relaxed problem (13) related to user k . More precisely said, we reveal the cases when threshold policies are optimal. As seen from the previous propositions, the ratio μ_1/μ_2 has an important role here.

Theorem 1. (On the optimality of threshold policies)

Let $\nu \geq 0$.

- (i) If $\mu_1/\mu_2 < x_{1,0}$, then there is $j \geq 0$ such that the threshold policy $\pi(x_{1,j})$ is ν -optimal.
- (ii) If $x_{1,0} \leq \mu_1/\mu_2 < q_2$, then there is $j_1 > k_1$ and $\mathcal{S}_{k_1} \subset \{x_{1,0}, \dots, x_{1,k_1}\}$, where

$$k_1 = \max \{j \in \{0, 1, \dots\} : x_{1,j} \leq \mu_1/\mu_2\}, \quad (21)$$

such that the policy π defined by

$$a^\pi(x) = \begin{cases} 0, & \text{for } x \in \mathcal{S}_{k_1} \cup \{x_{1,k_1+1}, \dots, x_{1,j_1-1}\} \cup \{*\}, \\ 1, & \text{otherwise.} \end{cases} \quad (22)$$

is ν -optimal.

- (iii) If $\mu_1/\mu_2 = q_2$, then there is $j \geq 0$ such that the threshold policy $\pi(x_{1,j})$ is ν -optimal.
- (iv) If $q_2 < \mu_1/\mu_2 < x_{2,0}$, then there is $j_1 \geq 0$ and $j_2 \leq k_2$, where

$$k_2 = \min \{j \in \{1, 2, \dots\} : x_{2,j} \leq \mu_1/\mu_2\}, \quad (23)$$

such that the policy π defined by

$$a^\pi(x) = \begin{cases} 0, & \text{for } x \in \{x_{1,0}, \dots, x_{1,j_1-1}\} \cup \\ & \{x_{2,j_2}, \dots, x_{2,k_2-1}\} \cup \{*\}, \\ 1, & \text{otherwise.} \end{cases} \quad (24)$$

is ν -optimal.

- (v) If $\mu_1/\mu_2 \geq x_{2,0}$, then the threshold policy $\pi(x_{1,0})$ is ν -optimal.

Proof. (i) Assume first that $\mu_1/\mu_2 < x_{1,0}$, and let π denote the ν -optimal policy in this case.

By (7), we deduce that $x_{2,j} > \mu_1/\mu_2$ for any $j \in \{0, 1, \dots\}$. Now it follows from Propositions 2 and 6(i) that $a^\pi(x_{2,j}) = 1$ for all $j \in \{0, 1, \dots\}$.

By (7), we also deduce that $x_{1,j} > \mu_1/\mu_2$ for any $j \in \{0, 1, \dots\}$. Now it follows from Propositions 4 and 6(iii) that there is $j_1 \in \{0, 1, \dots\}$ such that $a^\pi(x_{1,j}) = 0$ for all $j \in \{0, \dots, j_1 - 1\}$ and $a^\pi(x_{1,j}) = 1$ for all $j \in \{j_1, j_1 + 1, \dots\}$. In addition, since $a^\pi(q_2) = 1$ by Proposition 6(ii), the threshold policy $\pi(x_{1,j_1})$ is ν -optimal.

(ii) Assume now that $x_{1,0} \leq \mu_1/\mu_2 < q_2$, and let π denote the ν -optimal policy in this case.

By (7), we deduce that $x_{2,j} > \mu_1/\mu_2$ for any $j \in \{0, 1, \dots\}$. Now it follows from Propositions 2 and 6(i) that $a^\pi(x_{2,j}) = 1$ for all $j \in \{0, 1, \dots\}$.

By (7), we also deduce that $x_{1,j} > \mu_1/\mu_2$ for any $j \in \{k_1 + 1, k_1 + 2, \dots\}$, where k_1 is defined in (21). Thus, by Propositions 4 and 6(iii), there is $j_1 > k_1$ such that $a^\pi(x_{1,j}) = 0$ for all $j \in \{k_1 + 1, \dots, j_1 - 1\}$ and $a^\pi(x_{1,j}) = 1$ for all $j \in \{j_1, j_1 + 1, \dots\}$. In addition, since $a^\pi(q_2) = 1$ by Proposition 6(ii), we deduce that there is subset $\mathcal{S}_{k_1} \subset \{x_{1,0}, \dots, x_{1,k_1}\}$ such that the policy π defined in (22) is ν -optimal.

(iii) Assume then that $\mu_1/\mu_2 = q_2$, and let π denote the ν -optimal policy in this case.

By (7), we deduce that $x_{2,j} > \mu_1/\mu_2$ for any $j \in \{0, 1, \dots\}$. Now it follows from Propositions 2 and 6(i) that $a^\pi(x_{2,j}) = 1$ for all $j \in \{0, 1, \dots\}$.

On the other hand, it follows from Propositions 5 and 6(iii) that there is $j_1 \in \{0, 1, \dots\}$ such that $a^\pi(x_{1,j}) = 0$ for all $j \in \{0, \dots, j_1 - 1\}$ and $a^\pi(x_{1,j}) = 1$ for all $j \in \{j_1, j_1 + 1, \dots\}$. In addition, since $a^\pi(q_2) = 1$ by Proposition 6(ii), the threshold policy $\pi(x_{1,j_1})$ is ν -optimal.

(iv) Assume then that $q_2 < \mu_1/\mu_2 < x_{2,0}$, and let π denote the ν -optimal policy in this case.

By (7), we deduce that $x_{2,j} \leq \mu_1/\mu_2$ for any $j \in \{k_2, k_2 + 1, \dots\}$, where k_2 is defined in (23). Now it follows from Propositions 3 and 6(i) that $a^\pi(x_{2,j}) = 1$ for all $j \in \{k_2, k_2 + 1, \dots\}$. On the other hand, by (7), we also deduce that $x_{2,j} > \mu_1/\mu_2$ for any $j \in \{0, 1, \dots, k_2 - 1\}$. Now it follows from Proposition 2 that there is $j_2 \in \{0, \dots, k_2\}$ such that $a^\pi(x_{2,j}) = 1$ for all $j \in \{0, \dots, j_2 - 1\}$ and $a^\pi(x_{2,j}) = 0$ for all $j \in \{j_2, \dots, k_2 - 1\}$.

As above in (iii), it follows from Propositions 5 and 6(iii) that there is $j_1 \in \{0, 1, \dots\}$ such that $a^\pi(x_{1,j}) = 0$ for all $j \in \{0, \dots, j_1 - 1\}$ and

$a^\pi(x_{1,j}) = 1$ for all $j \in \{j_1, j_1 + 1, \dots\}$. In addition, since $a^\pi(q_2) = 1$ by Proposition 6(ii), we deduce that the policy π defined in (24) is ν -optimal.

(v) Assume finally that $\mu_1/\mu_2 \geq x_{2,0}$. It follows from (8) and (7) that in this case the base station transmits with the worst rate in any belief state $x \in \mathcal{S} \setminus \{*\}$ so that

$$\mu(x) = \mu(x, 2) = \mu(x, 1) = \mu_1.$$

The optimality equations (15) read now as follows:

$$\begin{aligned} V(q_2; \nu) &= h + \min\{V(q_2; \nu), \\ &\quad \nu + q_2(1 - \mu_1)V(x_{2,0}; \nu) \\ &\quad + (1 - q_2)(1 - \mu_1)V(x_{1,0}; \nu)\}, \\ V(x_{n,j}; \nu) &= h + \min\{V(x_{n,j+1}; \nu), \\ &\quad \nu + x_{n,j}(1 - \mu_1)V(x_{2,0}; \nu) \\ &\quad + (1 - x_{n,j})(1 - \mu_1)V(x_{1,0}; \nu)\}. \end{aligned} \tag{25}$$

Let π denote the policy that schedules flow k in any belief state (but $*$), i.e., $\pi = \pi(x_{1,0})$. Clearly, we have, for any $x \in \mathcal{S} \setminus \{*\}$,

$$V^\pi(x; \nu) = \frac{h + \nu}{\mu_1}.$$

Now it is easy to check that the optimality equations (25) are satisfied with this value function $V(x; \nu) = V^\pi(x; \nu)$. In other words, the threshold policy $\pi(x_{1,0})$ is ν -optimal. $\square$

Note that in cases (i), (iii), and (v) of Theorem 1, the optimal policy belongs, for sure, to the threshold policies Π^T .² In cases (ii) and (iv), it is also possible but not necessary. Case (v) is trivial, since then it is always optimal to schedule the user, but the other cases are non-trivial.

In the following section, we continue exploring the non-trivial cases (i) and (iii). More precisely said, we prove that the optimization problem (13) is indexable in these cases, and we also derive the corresponding Whittle index values for all belief states. Thereafter, in Section 6, we give some insight into the remaining non-trivial cases (ii) and (iv) by presenting examples where the optimal policy does *not* belong to the class of threshold policies Π^T , which makes the analysis even more challenging.

²Note also that, in [16, Proposition 4], which explicitly considered only case (i), the optimality of threshold policies was based on a conjecture that the problem itself is indexable, while Theorem 1 above does not require any such conjecture.

5. Whittle index with partial channel information

In this section, we continue the characterization of the optimal solution of the relaxed problem (13) related to user k . More precisely said, we consider cases (i), (iii), and (v) of Theorem 1 and show that the relaxed optimization problem is indexable in these cases. In addition, we derive explicit expressions for the corresponding Whittle index values; see Theorems 2 and 3, together with Corollary 1 below.

We start by defining the indexability property, and then consider the two cases separately in their own subsections. After the definition, we again leave out the user index k to simplify notations in the three subsections.

Definition 2. *The relaxed optimization problem (13) related to user k is indexable if, for any belief state $x \in \mathcal{S}_k \setminus \{*\}$, there exists $W_k(x) \in [-\infty, \infty]$ such that*

- (i) *decision $a = 1$ (to schedule user k) is optimal in belief state x if and only if $\nu \leq W_k(x)$;*
- (ii) *decision $a = 0$ (not to schedule user k) is optimal in belief state x if and only if $\nu \geq W_k(x)$.*

Note that, according to this definition, the two actions are equally good (and, thus, optimal) when $\nu = W_k(x)$.

5.1. Case (i) of Theorem 1

In this subsection, we assume that

$$\mu_1/\mu_2 < x_{1,0}, \quad (26)$$

which is the same case as explicitly considered in [16]. It follows from (7) and (26) that in this case the base station transmits with the best rate in any belief state $x \in \mathcal{S} \setminus \{*\}$:

$$\mu(x) = x\mu_2, \quad \mu(x, 2) = \mu_2, \quad \mu(x, 1) = 0.$$

The optimality equations (15) read now as follows:

$$\begin{aligned} V(*; \nu) &= 0, \\ V(q_2; \nu) &= h + \min\{V(q_2; \nu), \\ &\quad \nu + q_2(1 - \mu_2)V(x_{2,0}; \nu) + (1 - q_2)V(x_{1,0}; \nu)\}, \\ V(x_{n,j}; \nu) &= h + \min\{V(x_{n,j+1}; \nu), \\ &\quad \nu + x_{n,j}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{n,j})V(x_{1,0}; \nu)\}. \end{aligned} \quad (27)$$

Recall also that in this case the optimal policy belongs to the threshold policies Π^T for sure by Theorem 1(i).

For any $j \in \{0, 1, \dots\}$, let us define

$$\nu_{1,j} = h \left(\frac{x_{1,j}}{x_{1,j+1} - x_{1,j}} - j - 1 \right). \quad (28)$$

It follows from (6) that

$$\nu_{1,j} = h \left(\frac{1 - \rho^{j+1}}{(1 - \rho)\rho^{j+1}} - j - 1 \right). \quad (29)$$

Below we show that these values are positive and increasing unboundedly with respect to j .

Proposition 7. *For any $j \in \{0, 1, \dots\}$,*

$$0 < \nu_{1,j} < \nu_{1,j+1}. \quad (30)$$

In addition, $\lim_{j \rightarrow \infty} \nu_{1,j} = \infty$.

Proof. By (29),

$$\nu_{1,0} = h \left(\frac{1}{\rho} - 1 \right) > 0,$$

since $0 < \rho < 1$. Let then $j \in \{0, 1, \dots\}$. Now, again by (29), we have

$$\nu_{1,j+1} - \nu_{1,j} = h \left(\frac{1}{\rho^{j+2}} - 1 \right),$$

which is strictly positive and an increasing function of j , since $0 < \rho < 1$. The claims follow from these facts. $\square$

Now we are ready to show that the relaxed optimization problem is indexable under assumption (26), and the corresponding Whittle index is given by (28). The proof relies on Propositions 8 and 9 given below. The main result is thereafter presented in Theorem 2.

Proposition 8. *Assume (26). If policy $\pi(x_{1,0})$ is ν -optimal, then $\nu \leq \nu_{1,0}$.*

Proof. Assume that policy $\pi(x_{1,0})$ is ν -optimal, and denote $\pi = \pi(x_{1,0})$. Now it follows from the optimality equations (27) that

$$\begin{aligned} V(x_{1,0}; \nu) &= h + \nu + x_{1,0}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,0})V(x_{1,0}; \nu), \\ V(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,0})V(x_{1,0}; \nu), \end{aligned}$$

which gives

$$\begin{aligned} V(x_{1,0}; \nu) &= (h + \nu) \frac{1 - (x_{2,0} - x_{1,0})(1 - \mu_2)}{x_{1,0}\mu_2}, \\ V(x_{2,0}; \nu) &= (h + \nu) \frac{1 - x_{2,0} + x_{1,0}}{x_{1,0}\mu_2}. \end{aligned} \tag{31}$$

Similarly, by the optimality equations (27),

$$V(x_{1,1}; \nu) = h + \nu + x_{1,1}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,1})V(x_{1,0}; \nu).$$

On the other hand, since $a^\pi(x_{1,0}) = 1$, we have, again by the optimality equations (27), that

$$V(x_{1,1}; \nu) \geq \nu + x_{1,0}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,0})V(x_{1,0}; \nu).$$

Thus,

$$V(x_{1,0}; \nu) \leq V(x_{2,0}; \nu)(1 - \mu_2) + \frac{h}{x_{1,1} - x_{1,0}}.$$

By combining this with (31), we finally deduce that

$$\nu \leq h \left(\frac{x_{1,0}}{x_{1,1} - x_{1,0}} - 1 \right) = \nu_{1,0},$$

which completes the proof. $\square$

Proposition 9. Assume (26), and let $j \in \{1, 2, \dots\}$. If policy $\pi(x_{1,j})$ is ν -optimal, then

$$\nu_{1,j-1} \leq \nu \leq \nu_{1,j}.$$

Proof. Assume that policy $\pi(x_{1,j})$ is ν -optimal, and denote $\pi = \pi(x_{1,j})$. First, for any belief state $\{x_{1,0}, \dots, x_{1,j-1}\}$, we have, by the optimality equations (27),

$$V(x_{1,i}; \nu) = h + V(x_{1,i+1}; \nu)$$

implying that

$$V(x_{1,0}; \nu) = jh + V(x_{1,j}; \nu)$$

In addition, it follows from the optimality equations (27) that

$$\begin{aligned} V(x_{1,j}; \nu) &= h + \nu + x_{1,j}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,j})V(x_{1,0}; \nu), \\ V(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,0})V(x_{1,0}; \nu). \end{aligned} \quad (32)$$

Thus,

$$\begin{aligned} V(x_{1,0}; \nu) &= \frac{(h + \nu)(1 - (x_{2,0} - x_{1,j})(1 - \mu_2))}{x_{1,j}\mu_2} \\ &\quad + \frac{jh(1 - x_{2,0}(1 - \mu_2))}{x_{1,j}\mu_2}, \\ V(x_{1,j}; \nu) &= \frac{(h + \nu)(1 - (x_{2,0} - x_{1,j})(1 - \mu_2))}{x_{1,j}\mu_2} \\ &\quad + \frac{jh(1 - x_{2,0}(1 - \mu_2) - x_{1,j}\mu_2)}{x_{1,j}\mu_2}, \\ V(x_{2,0}; \nu) &= \frac{(h + \nu)(1 - x_{2,0} + x_{1,0}) + jh(1 - x_{2,0})}{x_{1,j}\mu_2}. \end{aligned} \quad (33)$$

Now, since $a^\pi(x_{1,j-1}) = 0$, we have, by the optimality equations (27),

$$V(x_{1,j}; \nu) \leq \nu + x_{1,j-1}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,j-1})V(x_{1,0}; \nu),$$

which implies, by (32), that

$$V(x_{2,0}; \nu)(1 - \mu_2) + \frac{h}{x_{1,j} - x_{1,j-1}} \leq V(x_{1,0}; \nu).$$

By combining this with (33), we deduce that

$$\nu \geq h \left(\frac{x_{1,j-1}}{x_{1,j} - x_{1,j-1}} - j \right) = \nu_{1,j-1}.$$

On the other hand, since $a^\pi(x_{1,j}) = 1$, we have, by the optimality equations (27) and, that

$$V(x_{1,j+1}; \nu) \geq \nu + x_{1,j}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{1,j})V(x_{1,0}; \nu).$$

In addition, again by the optimality equations (27),

$$V(x_{1,j+1}; \nu) = h + \nu + x_{j+1,1}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{j+1,1})V(x_{1,0}; \nu).$$

Thus,

$$V(x_{1,0}; \nu) \leq V(x_{2,0}; \nu)(1 - \mu_2) + \frac{h}{x_{1,j+1} - x_{1,j}}.$$

By combining this with (33), we finally deduce that

$$\nu \leq h \left(\frac{x_{1,j}}{x_{1,j+1} - x_{1,j}} - j - 1 \right) = \nu_{1,j},$$

which completes the proof. $\square$

Theorem 2. *Under assumption (26), the relaxed optimization problem (13) is indexable, and the corresponding Whittle index $W(x)$ is given by*

$$\begin{aligned} W(x_{2,j}) &= \infty, & j \in \{0, 1, \dots\}, \\ W(q_2) &= \infty, \\ W(x_{1,j}) &= \nu_{1,j}, & j \in \{0, 1, \dots\}, \end{aligned} \tag{34}$$

where $\nu_{1,j}$ is given in (29).

Proof. The results follow straightforwardly from Theorem 1(i) and Propositions 7-9. $\square$

Observe that, under assumption (26), the Whittle index values do *not* depend on the “rate” parameters μ_1 and μ_2 at all but only on the “channel dynamics” parameters $q_{1,2}$ and $q_{2,2}$ together with the holding cost weight h and the time $j + 1$ from the previous channel state update, as seen from (34) and (29), since $\rho = q_{2,2} - q_{1,2}$ by (2). However, this fact is not clearly visible from the implicit Whittle index formula given in [16, Theorem 4.1] under the same assumption (26).

Note also that, since $W(x_{2,j}) = \infty$ for any j and $W(q_2) = \infty$, it is always optimal (for the relaxed optimization problem) to schedule flow k when the last observation from the channel state is the good one or when there is not yet any observations available (so that the belief state is equal to the stationary probability of the good channel state).

5.2. Case (iii) of Theorem 1

Assume now that

$$\mu_1/\mu_2 = q_2. \quad (35)$$

In this case, for any $x \in \{x_{2,j} : j \in \{0, 1, \dots\}\}$,

$$\mu(x) = x\mu_2, \quad \mu(x, 2) = \mu_2, \quad \mu(x, 1) = 0,$$

and, for any $x \in \{x_{1,j} : j \in \{0, 1, \dots\}\} \cup \{q_2\}$,

$$\mu(x) = \mu(x, 2) = \mu(x, 1) = \mu_1.$$

The optimality equations (15) read now as follows:

$$\begin{aligned} V(*; \nu) &= 0, \\ V(x_{2,j}; \nu) &= h + \min\{V(x_{2,j+1}; \nu), \\ &\quad \nu + x_{2,j}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,j})V(x_{1,0}; \nu)\}, \\ V(q_2; \nu) &= h + \min\{V(q_2; \nu), \\ &\quad \nu + q_2(1 - \mu_1)V(x_{2,0}; \nu) + (1 - q_2)(1 - \mu_1)V(x_{1,0}; \nu)\}, \\ V(x_{1,j}; \nu) &= h + \min\{V(x_{1,j+1}; \nu), \\ &\quad \nu + x_{1,j}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,j})(1 - \mu_1)V(x_{1,0}; \nu)\}. \end{aligned} \quad (36)$$

Recall also that in this special case the optimal policy belongs to the threshold policies Π^T for sure by Theorem 1(iii).

For any $j \in \{0, 1, \dots\}$, let us define

$$\begin{aligned} \tilde{\nu}_{1,j} &= \frac{h x_{2,0}\mu_2}{x_{2,0}\mu_2 - \mu_1} \left(\frac{x_{1,j}}{x_{1,j+1} - x_{1,j}} - j - 1 \right) \\ &\quad + \frac{h \mu_1}{x_{2,0}\mu_2 - \mu_1} \left(\frac{1 - x_{2,0}(1 - \mu_2)}{(x_{1,j+1} - x_{1,j})(1 - \mu_1)} + 1 \right). \end{aligned} \quad (37)$$

It follows from (6) that

$$\begin{aligned} \tilde{\nu}_{1,j} &= \frac{h q_{2,2}\mu_2}{q_{2,2}\mu_2 - \mu_1} \left(\frac{1 - \rho^{j+1}}{(1 - \rho)\rho^{j+1}} - j - 1 \right) \\ &\quad + \frac{h \mu_1}{q_{2,2}\mu_2 - \mu_1} \left(\frac{1 - q_{2,2}(1 - \mu_2)}{q_{1,2}(1 - \mu_1)\rho^{j+1}} + 1 \right). \end{aligned} \quad (38)$$

Below we first show that these values $\tilde{\nu}_{1,j}$ are positive and increasing unboundedly with respect to j , and then compare them to values $\nu_{1,j}$ defined in Section 5.1.

Proposition 10. For any $j \in \{0, 1, \dots\}$,

$$0 < \tilde{\nu}_{1,j} < \tilde{\nu}_{1,j+1}. \quad (39)$$

In addition, $\lim_{j \rightarrow \infty} \tilde{\nu}_{1,j} = \infty$.

Proof. By (38),

$$\begin{aligned} \tilde{\nu}_{1,0} &= \frac{h x_{2,0} \mu_2}{x_{2,0} \mu_2 - \mu_1} \left(\frac{1}{\rho} - 1 \right) \\ &\quad + \frac{h \mu_1}{x_{2,0} \mu_2 - \mu_1} \left(\frac{1 - x_{2,0}(1 - \mu_2)}{x_{1,0}(1 - \mu_1)\rho} + 1 \right) > 0, \end{aligned}$$

since $x_{2,0} \mu_2 > \mu_1$ by (35), and $0 < \rho < 1$. Let then $j \in \{0, 1, \dots\}$. Now, again by (38), we have

$$\begin{aligned} \tilde{\nu}_{1,j+1} - \tilde{\nu}_{1,j} &= \frac{h x_{2,0} \mu_2}{x_{2,0} \mu_2 - \mu_1} \left(\frac{1}{\rho^{j+2}} - 1 \right) \\ &\quad + \frac{h \mu_1}{x_{2,0} \mu_2 - \mu_1} \left(\frac{(1 - x_{2,0}(1 - \mu_2))(1 - \rho)}{x_{1,0}(1 - \mu_1)\rho^{j+2}} \right), \end{aligned}$$

which is strictly positive and an increasing function of j , since $0 < \rho < 1$. The claims follow from these facts. $\square$

Proposition 11. For any $j \in \{0, 1, \dots\}$,

$$\nu_{1,j} < \tilde{\nu}_{1,j}. \quad (40)$$

Proof. This follows from (29) and (38). In particular, note that the latter part of the first term for $\tilde{\nu}_{1,j}$ in (38) is equal to $\nu_{1,j}$ given in (29). $\square$

Now we are ready to show that the relaxed optimization problem is indexable under assumption (35), and the corresponding Whittle index is given by (37). The proof relies on Propositions 12 and 13 given below. The main result is thereafter presented in Theorem 3.

Proposition 12. Assume (35). If policy $\pi(x_{1,0})$ is ν -optimal, then $\nu \leq \tilde{\nu}_{1,0}$.

Proof. Assume that policy $\pi(x_{1,0})$ is ν -optimal, and denote $\pi = \pi(x_{1,0})$. Now it follows from the optimality equations (36) that

$$\begin{aligned} V(x_{1,0}; \nu) &= h + \nu + x_{1,0}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,0})(1 - \mu_1)V(x_{1,0}; \nu), \\ V(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,0})V(x_{1,0}; \nu), \end{aligned}$$

which gives

$$\begin{aligned} V(x_{1,0}; \nu) &= \frac{(h + \nu)(1 - x_{2,0} + x_{1,0}(1 - \mu_1) + x_{2,0}\mu_2)}{x_{1,0}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1}, \\ V(x_{2,0}; \nu) &= \frac{(h + \nu)(1 - x_{2,0} + x_{1,0}(1 - \mu_1) + \mu_1)}{x_{1,0}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1}. \end{aligned} \quad (41)$$

Similarly, by the optimality equations (36),

$$V(x_{1,1}; \nu) = h + \nu + x_{1,1}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,1})(1 - \mu_1)V(x_{1,0}; \nu).$$

On the other hand, since $a^\pi(x_{1,0}) = 1$, we have, again by the optimality equations (36),

$$V(x_{1,1}; \nu) \geq \nu + x_{1,0}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,0})(1 - \mu_1)V(x_{1,0}; \nu).$$

Thus,

$$V(x_{1,0}; \nu) \leq V(x_{2,0}; \nu) + \frac{h}{(x_{1,1} - x_{1,0})(1 - \mu_1)}.$$

By combining this with (41), we finally deduce that

$$\begin{aligned} \nu &\leq h \left(\frac{x_{1,0}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1}{(x_{2,0}\mu_2 - \mu_1)(x_{1,1} - x_{1,0})(1 - \mu_1)} - 1 \right) \\ &= \frac{h x_{2,0}\mu_2}{x_{2,0}\mu_2 - \mu_1} \left(\frac{x_{1,0}}{x_{1,1} - x_{1,0}} - 1 \right) \\ &\quad + \frac{h \mu_1}{x_{2,0}\mu_2 - \mu_1} \left(\frac{1 - x_{2,0}(1 - \mu_2)}{(x_{1,1} - x_{1,0})(1 - \mu_1)} + 1 \right) = \tilde{\nu}_{1,0}, \end{aligned}$$

which completes the proof. $\square$

Proposition 13. Assume (35), and let $j \in \{1, 2, \dots\}$. If policy $\pi(x_{1,j})$ is ν -optimal, then

$$\tilde{\nu}_{1,j-1} \leq \nu \leq \tilde{\nu}_{1,j}.$$

Proof. Assume that policy $\pi(x_{1,j})$ is ν -optimal, and denote $\pi = \pi(x_{1,j})$. First, for any belief state $\{x_{1,0}, \dots, x_{1,j-1}\}$, we have, by the optimality equations (36),

$$V(x_{1,i}; \nu) = h + V(x_{1,i+1}; \nu)$$

implying that

$$V(x_{1,0}; \nu) = jh + V(x_{1,j}; \nu)$$

In addition, it follows from the optimality equations (36) that

$$\begin{aligned} V(x_{1,j}; \nu) &= h + \nu + x_{1,j}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,j})(1 - \mu_1)V(x_{1,0}; \nu), \\ V(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V(x_{2,0}; \nu) + (1 - x_{2,0})V(x_{1,0}; \nu). \end{aligned} \quad (42)$$

Thus,

$$\begin{aligned} V(x_{1,0}; \nu) &= \frac{(h + \nu)(1 - x_{2,0} + x_{1,j}(1 - \mu_1) + x_{2,0}\mu_2)}{x_{1,j}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1} \\ &\quad + \frac{jh(1 - x_{2,0}(1 - \mu_2))}{x_{1,j}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1}, \\ V(x_{1,j}; \nu) &= \frac{(h + \nu)(1 - x_{2,0} + x_{1,j}(1 - \mu_1) + x_{2,0}\mu_2)}{x_{1,j}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1} \\ &\quad + \frac{jh(1 - \mu_1)(1 - x_{2,0}(1 - \mu_2(1 - x_{1,j})))}{x_{1,j}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1}, \\ V(x_{2,0}; \nu) &= \frac{(h + \nu)(1 - x_{2,0} + x_{1,j}(1 - \mu_1) + \mu_1)}{x_{1,j}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1} \\ &\quad + \frac{jh(1 - x_{2,0})}{x_{1,j}(1 - \mu_1)x_{2,0}\mu_2 + (1 - x_{2,0}(1 - \mu_2))\mu_1}. \end{aligned} \quad (43)$$

Now, since $a^\pi(x_{1,j-1}) = 0$, we have, by the optimality equations (36),

$$V(x_{1,j}; \nu) \leq \nu + x_{1,j-1}(1 - \mu_1)V(x_{2,0}; \nu) + (1 - x_{1,j-1})(1 - \mu_1)V(x_{1,0}; \nu),$$

which implies, by (42), that

$$V(x_{2,0}; \nu) + \frac{h}{(x_{1,j} - x_{1,j-1})(1 - \mu_1)} \leq V(x_{1,0}; \nu).$$

By combining this with (43), we deduce that

$$\begin{aligned}
\nu &\geq h \left(\frac{x_{1,j}(1-\mu_1)x_{2,0}\mu_2 + (1-x_{2,0}(1-\mu_2))\mu_1}{(x_{2,0}\mu_2 - \mu_1)(x_{1,j} - x_{1,j-1})(1-\mu_1)} - \frac{jx_{2,0}\mu_2}{x_{2,0}\mu_2 - \mu_1} - 1 \right) \\
&= \frac{h x_{2,0}\mu_2}{x_{2,0}\mu_2 - \mu_1} \left(\frac{x_{1,j-1}}{x_{1,j} - x_{1,j-1}} - j \right) \\
&\quad + \frac{h \mu_1}{x_{2,0}\mu_2 - \mu_1} \left(\frac{1 - x_{2,0}(1-\mu_2)}{(x_{1,j} - x_{1,j-1})(1-\mu_1)} + 1 \right) = \tilde{\nu}_{1,j-1}.
\end{aligned}$$

On the other hand, since $a^\pi(x_{1,j}) = 1$, we have, by the optimality equations (36),

$$V(x_{1,j+1}; \nu) \geq \nu + x_{1,j}(1-\mu_1)V(x_{2,0}; \nu) + (1-x_{1,j})(1-\mu_1)V(x_{1,0}; \nu).$$

In addition, again by the optimality equations (36),

$$V(x_{1,j+1}; \nu) = h + \nu + x_{j+1,1}(1-\mu_1)V(x_{2,0}; \nu) + (1-x_{j+1,1})(1-\mu_1)V(x_{1,0}; \nu).$$

Thus,

$$V(x_{1,0}; \nu) \leq \frac{h}{(x_{1,j+1} - x_{1,j})(1-\mu_1)} + V(x_{2,0}; \nu).$$

By combining this with (43), we finally deduce that

$$\begin{aligned}
\nu &\leq h \left(\frac{x_{1,j}(1-\mu_1)x_{2,0}\mu_2 + (1-x_{2,0}(1-\mu_2))\mu_1}{(x_{2,0}\mu_2 - \mu_1)(x_{1,j+1} - x_{1,j})(1-\mu_1)} - \frac{jx_{2,0}\mu_2}{x_{2,0}\mu_2 - \mu_1} - 1 \right) \\
&= \frac{h x_{2,0}\mu_2}{x_{2,0}\mu_2 - \mu_1} \left(\frac{x_{1,j}}{x_{1,j+1} - x_{1,j}} - j - 1 \right) \\
&\quad + \frac{h \mu_1}{x_{2,0}\mu_2 - \mu_1} \left(\frac{1 - x_{2,0}(1-\mu_2)}{(x_{1,j+1} - x_{1,j})(1-\mu_1)} + 1 \right) = \tilde{\nu}_{1,j},
\end{aligned}$$

which completes the proof. $\square$

Theorem 3. *Under assumption (35), the relaxed optimization problem (13) is indexable, and the corresponding Whittle index $W(x)$ is given by*

$$\begin{aligned}
W(x_{2,j}) &= \infty, \quad j \in \{0, 1, \dots\}, \\
W(q_2) &= \infty, \\
W(x_{1,j}) &= \tilde{\nu}_{1,j}, \quad j \in \{0, 1, \dots\},
\end{aligned} \tag{44}$$

where $\tilde{\nu}_{1,j}$ is given in (38).

Proof. The results follow straightforwardly from Theorem 1(iii) and Propositions 10-13. $\square$

Under assumption (35), the Whittle index values depend on the “rate” parameters μ_1 and μ_2 as well as the “channel dynamics” parameters $q_{1,2}$ and $q_{2,2}$, the holding cost weight h , and the time $j + 1$ from the previous channel state update. However, as already mentioned earlier, the latter part of the first term for $\tilde{\nu}_{1,j}$ in (38) is equal to $\nu_{1,j}$ given in (29), which does not depend on μ_1 and μ_2 .

Note also that $W(x_{2,j}) = \infty$ for any j and $W(q_2) = \infty$, just as in case (i) above. Thus, it is again always optimal (for the relaxed optimization problem) to schedule flow k when the last observation from the channel state is the good one or when there is not yet any observations available.

5.3. Case (v) of Theorem 1

Assume finally that

$$\mu_1/\mu_2 \geq x_{2,0}. \quad (45)$$

By Theorem 1(v), the optimal policy in this special case is, for any ν , the threshold policy $\pi(x_{1,0}) \in \Pi^T$, in which user k is scheduled in any belief state $x \in \mathcal{S}_k \setminus \{*\}$. As an immediate consequence, we see that the problem is (trivially) indexable and the corresponding index is infinite.

Corollary 1. *Under assumption (45), the relaxed optimization problem (13) is indexable, and the corresponding Whittle index $W(x)$ is given by*

$$W(x) = \infty \quad (46)$$

for any belief state $x \in \mathcal{S}_k \setminus \{*\}$.

6. Examples of optimal policies which are not of threshold type

In this section, we give more insight into the remaining cases (ii) and (iv) of Theorem 1 by presenting examples where the optimal policy of the relaxed problem (13) related to user k does not belong to the class of threshold policies Π^T defined in (17). As in Sections 4 and 5, we leave out here the user index k to simplify notations.

6.1. Case (ii) of Theorem 1

In this subsection, we consider the case (ii) of Theorem 1, for which it is required that

$$x_{1,0} \leq \mu_1/\mu_2 < q_2. \quad (47)$$

Assume that

$$\begin{aligned} h &= 1, \quad \nu = 1.25, \\ q_{1,2} &= 0.1, \quad \rho = 0.8, \quad \mu_1 = 0.0179, \quad \mu_2 = 0.1. \end{aligned} \quad (48)$$

It follows from (3) and (6) that

$$x_{1,0} = 0.1, \quad \mu_1/\mu_2 = 0.179, \quad x_{1,1} = 0.18 \quad q_2 = 0.5$$

so that

$$x_{1,0} < \mu_1/\mu_2 < x_{1,1} < q_2, \quad (49)$$

which satisfies condition (47).

Let us now define policy π by requiring that

$$a^\pi(x) = \begin{cases} 0, & \text{for } x \in \{x_{1,1}\} \cup \{*\}, \\ 1, & \text{otherwise.} \end{cases} \quad (50)$$

Note that π is *not* a threshold policy as defined in (17), since $a^\pi(x_{1,0}) = 1$, but $a^\pi(x_{1,1}) = 0$. The value function for policy π at belief state $x_{1,1}$ equals

$$V^\pi(x_{1,1}; \nu) = 52.0270, \quad (51)$$

which can be verified by solving the following system of linear equations based on condition (49) and applying then the parameters given in (48):

$$\begin{aligned} V^\pi(x_{1,0}; \nu) &= h + \nu + x_{1,0}(1 - \mu_1)V^\pi(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,0})(1 - \mu_1)V^\pi(x_{1,0}; \nu), \\ V^\pi(x_{1,1}; \nu) &= h + V^\pi(x_{1,2}; \nu), \\ V^\pi(x_{1,2}; \nu) &= h + \nu + x_{1,2}(1 - \mu_2)V^\pi(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,2})V^\pi(x_{1,0}; \nu), \\ V^\pi(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V^\pi(x_{2,0}; \nu) \\ &\quad + (1 - x_{2,0})V^\pi(x_{1,0}; \nu). \end{aligned} \quad (52)$$

Let us then consider the threshold policy $\pi(x_{1,0})$. The value function for this policy at belief state $x_{1,1}$ equals

$$V^{\pi(x_{1,0})}(x_{1,1}; \nu) = 52.1295, \quad (53)$$

which can be verified by solving the following system of linear equations based on condition (49) and applying then the parameters given in (48):

$$\begin{aligned} V^{\pi(x_{1,0})}(x_{1,0}; \nu) &= h + \nu + x_{1,0}(1 - \mu_1)V^{\pi(x_{1,0})}(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,0})(1 - \mu_1)V^{\pi(x_{1,0})}(x_{1,0}; \nu), \\ V^{\pi(x_{1,0})}(x_{1,1}; \nu) &= h + \nu + x_{1,1}(1 - \mu_2)V^{\pi(x_{1,0})}(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,1})V^{\pi(x_{1,0})}(x_{1,0}; \nu), \\ V^{\pi(x_{1,0})}(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V^{\pi(x_{1,0})}(x_{2,0}; \nu) \\ &\quad + (1 - x_{2,0})V^{\pi(x_{1,0})}(x_{1,0}; \nu). \end{aligned} \quad (54)$$

Let us finally consider a generic threshold policy $\pi(x_{1,j})$ with $j \geq 1$. The value function for this policy at belief state $x_{1,1}$ equals

$$V^{\pi(x_{1,j})}(x_{1,1}; \nu) = 19.25 + \frac{10.6875 + 4.75j}{1.25 - (0.8)^j}, \quad (55)$$

which can be verified by solving the following system of linear equations based on condition (49) and applying then the parameters given in (48):

$$\begin{aligned} V^{\pi(x_{1,j})}(x_{1,0}; \nu) &= h + V^{\pi(x_{1,j})}(x_{1,1}; \nu), \\ V^{\pi(x_{1,j})}(x_{1,1}; \nu) &= (j - 1)h + V^{\pi(x_{1,j})}(x_{1,j}; \nu), \\ V^{\pi(x_{1,j})}(x_{1,j}; \nu) &= h + \nu + x_{1,j}(1 - \mu_2)V^{\pi(x_{1,j})}(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,j})V^{\pi(x_{1,j})}(x_{1,0}; \nu), \\ V^{\pi(x_{1,j})}(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V^{\pi(x_{1,j})}(x_{2,0}; \nu) \\ &\quad + (1 - x_{2,0})V^{\pi(x_{1,j})}(x_{1,0}; \nu). \end{aligned} \quad (56)$$

From (55), we see that $V^{\pi(x_{1,j})}(x_{1,1}; \nu)$ is an increasing function of j for any $j \geq 2$, for which

$$\begin{aligned} V^{\pi(x_{1,1})}(x_{1,1}; \nu) &= 53.5556, & V^{\pi(x_{1,2})}(x_{1,1}; \nu) &= 52.3443, \\ V^{\pi(x_{1,3})}(x_{1,1}; \nu) &= 53.0407, & V^{\pi(x_{1,4})}(x_{1,1}; \nu) &= 54.5754, \\ V^{\pi(x_{1,5})}(x_{1,1}; \nu) &= 56.5879, & \dots & \end{aligned} \quad (57)$$

Based on (51), (53), and (57), we deduce that the policy π defined in (50) results in a smaller cost at belief state $x_{1,1}$ than any of the threshold policies $\pi(x_{1,j})$. In addition, from Proposition 6, we can deduce that the threshold policies $\pi(\theta)$ where $\theta \in [q_2, 1]$ are never optimal. Thus, the optimal policy is not of threshold type in this example.

6.2. Case (iv) of Theorem 1

In this subsection, we consider the case (iv) of Theorem 1, for which it is required that

$$q_2 < \mu_1/\mu_2 < x_{2,0}. \quad (58)$$

Assume now that

$$\begin{aligned} h &= 1.0, & \nu &= 200.0, \\ q_{1,2} &= 0.1, & \rho &= 0.05, & \mu_1 &= 0.05374, & \mu_2 &= 0.5. \end{aligned} \quad (59)$$

It follows from (3) and (6) that

$$\begin{aligned} q_2 &= 0.105263, & x_{2,2} &= 0.105375, & \mu_1/\mu_2 &= 0.10748, \\ x_{2,1} &= 0.1075 & x_{2,0} &= 0.15 \end{aligned}$$

so that

$$q_2 < x_{2,2} < \mu_1/\mu_2 < x_{2,1} < x_{2,0}, \quad (60)$$

which satisfies condition (58).

Let us now define policy π by requiring that

$$a^\pi(x) = \begin{cases} 0, & \text{for } x \in \{x_{2,1}\} \cup \{*\}, \\ 1, & \text{otherwise.} \end{cases} \quad (61)$$

Note that π is *not* a threshold policy as defined in (17), since $a^\pi(x_{2,1}) = 0$ while $a^\pi(x_{2,2}) = 1$. The value function for policy π at belief state $x_{2,1}$ equals

$$V^\pi(x_{2,1}; \nu) = 3608.39, \quad (62)$$

which can be verified by solving the following system of linear equations

based on condition (60) and applying then the parameters given in (59):

$$\begin{aligned}
V^\pi(x_{1,0}; \nu) &= h + \nu + x_{1,0}(1 - \mu_1)V^\pi(x_{2,0}; \nu) \\
&\quad + (1 - x_{1,0})(1 - \mu_1)V^\pi(x_{1,0}; \nu), \\
V^\pi(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V^\pi(x_{2,0}; \nu) \\
&\quad + (1 - x_{2,0})V^\pi(x_{1,0}; \nu), \\
V^\pi(x_{2,1}; \nu) &= h + V^\pi(x_{2,2}; \nu), \\
V^\pi(x_{2,2}; \nu) &= h + \nu + x_{2,2}(1 - \mu_1)V^\pi(x_{2,0}; \nu) \\
&\quad + (1 - x_{2,2})(1 - \mu_1)V^\pi(x_{1,0}; \nu).
\end{aligned} \tag{63}$$

Let us then consider a generic threshold policy $\pi(x_{1,j})$ with $j \geq 0$. The value function for this policy at belief state $x_{2,1}$ equals

$$V^{\pi(x_{1,j})}(x_{2,1}; \nu) = 2736.95 + \frac{132899.302 + 2332.516j}{153.083 - (0.05)^j}, \tag{64}$$

which can be verified by solving the following system of linear equations based on condition (60) and applying then the parameters given in (59):

$$\begin{aligned}
V^{\pi(x_{1,j})}(x_{1,0}; \nu) &= jh + V^{\pi(x_{1,j})}(x_{1,j}; \nu), \\
V^{\pi(x_{1,j})}(x_{1,j}; \nu) &= h + \nu + x_{1,j}(1 - \mu_1)V^{\pi(x_{1,j})}(x_{2,0}; \nu) \\
&\quad + (1 - x_{1,j})(1 - \mu_1)V^{\pi(x_{1,j})}(x_{1,0}; \nu), \\
V^{\pi(x_{1,j})}(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V^{\pi(x_{1,j})}(x_{2,0}; \nu) \\
&\quad + (1 - x_{2,0})V^{\pi(x_{1,j})}(x_{1,0}; \nu) \\
V^{\pi(x_{1,j})}(x_{2,1}; \nu) &= h + \nu + x_{2,1}(1 - \mu_2)V^{\pi(x_{1,j})}(x_{2,0}; \nu) \\
&\quad + (1 - x_{2,1})V^{\pi(x_{1,j})}(x_{1,0}; \nu).
\end{aligned} \tag{65}$$

From (64), we see that $V^{\pi(x_{1,j})}(x_{2,1}; \nu)$ is an increasing function of j for any $j \geq 1$, for which

$$\begin{aligned}
V^{\pi(x_{1,0})}(x_{2,1}; \nu) &= 3610.81, & V^{\pi(x_{1,1})}(x_{2,1}; \nu) &= 3620.63, \\
V^{\pi(x_{1,2})}(x_{2,1}; \nu) &= 3635.59, & V^{\pi(x_{1,3})}(x_{2,1}; \nu) &= 3650.82, \\
V^{\pi(x_{1,4})}(x_{2,1}; \nu) &= 3666.05, & \dots &
\end{aligned} \tag{66}$$

Based on (62) and (66), we deduce that the policy π defined in (61) results in a smaller cost at belief state $x_{2,1}$ than any of the threshold policies $\pi(x_{1,j})$. In addition, from Proposition 6, we can deduce that the threshold policies $\pi(\theta)$ where $\theta \in [q_2, 1]$ are never optimal. Thus, the optimal policy is not of threshold type in this example.

7. Whittle index policy

Now we are ready to define the Whittle index policy for the original opportunistic scheduling problem described in Section 2, where it is strictly required that (at most) M users are scheduled at each timeslot t , see (1).

Assume that there are K users (with $K > M$) whose channel dynamics satisfy either condition (26) or (35). Let $x_k(t)$ denote the belief state of user k in the beginning of timeslot t . The initial belief state $x_k(0)$ is assumed to be the equilibrium state of the corresponding two-state Markov process, $x_k(0) = q_{k,2}$, given by (3). At each timeslot t , the *Whittle index policy* schedules the M users with the highest Whittle index values $W_k(x_k(t))$, where index $W_k(x_k(t))$ is given by (34) for those users that satisfy condition (26) and by (44) for those users that satisfy condition (35). All ties are here broken randomly. Let us use acronym *WhR* for this policy.

We also define a slightly modified version of the Whittle index policy, where the difference comes from the tie-breaking rule. Let us use acronym *WhM* for the Whittle index policy that breaks the ties related to an infinite Whittle index value by applying the rule that is based on a myopic index $M_k(x_k(t))$ defined by

$$M_k(x) = h_k \max\{x\mu_{k,2}, \mu_{k,1}\}. \quad (67)$$

All the other ties are broken randomly.

For comparison, we define two other index policies (cf. [16]). The *myopic index policy* (*Mpc*) schedules greedily the M users with the highest myopic index values $M_k(x_k(t))$ given in (67), thus maximizing the short-term utilization. The other one, called the *belief state index policy* (*Blf*) schedules the M users with the highest belief state values $x_k(t)$. For both of these policies, all ties are broken randomly.

If all users are homogeneous, *Mpc* is throughput-optimal for the corresponding packet-level scheduling problem with partial channel information [20, 37, 5]. In this case, both *Blf* and *WhM* make, in fact, identical decisions with *Mpc* being, thus, also optimal. For our flow-level problem with heterogeneous users, *Mpc* becomes optimal in the limit $\rho_k \rightarrow 0$ for all k (as already noticed in [16]), since in this limit the channels become IID in time, and the optimal scheduling policy is known to be the $c\mu$ -rule [11], which is equivalent with the *Mpc* policy in this case.

Note that, due to mathematical tractability, the development of the Whittle index policies in our flow-level opportunistic scheduling problem (as in all

similar flow-level problems [8, 14, 16, 12, 30, 32, 1, 2, 13, 34, 15, 3, 6]) is based on a transient setting with a finite number of jobs in the beginning and no new arrivals later on. Still, as also done in many of the papers mentioned above, the developed Whittle index policies can be applied in the dynamic setting with new flows arriving randomly. Next we evaluate numerically the performance of our Whittle index policies in such a dynamic setting.

8. Numerical results

In this section, we illustrate the properties of the Whittle index policies, WhR and WhM, and compare them against greedy policies Mpc and Blf, which were introduced in the previous section. In addition, as a reference scheduler providing insight on the value of *exact* channel state information we use the well-known *PF* scheduler [10], which applies an index rule, where the index is the instantaneous channel rate divided by the current throughput. Another extreme is provided by the *Rnd* policy, which does not use any channel information but the scheduled flow is selected purely randomly. The rate adaptation is still greedily optimized as in other policies. The performance of the policies using partial channel information (WhR, WhM, Mpc, Blf) should lie in between Rnd and PF.

We implemented a discrete event simulator to simulate a dynamic setting with two classes of flows indexed by $k \in \{1, 2\}$. In the simulations, flows with a random size arrive according to a Poisson process with rate λ_k in each class k , and each class has its own channel dynamics. The simulation proceeds in discrete time from one time slot to the next one, and the scheduling is decided at the beginning of the time slot, as in actual cellular systems. In each time slot, exactly one flow is scheduled (i.e., $M = 1$). During the time slot, the channel rate of each flow is constant and changes according to the considered two-state Markov chain between the successive time slots. Let $r_{k,n}$ denote the channel rate (Mbit/s) of class- k flows in channel state $n \in \{1, 2\}$. In LTE the time slot length is 1 ms, but in our simulations, the time slot length was $\Delta = 10$ ms in order to keep simulation times manageable. The results of each simulation with a given load consist of at least 10^6 flow arrivals. The original flow size is exponentially distributed with a given mean $E[B]$, which is the same for both classes k . Thus, the corresponding completion probabilities are $\mu_{k,n} = r_{k,n}\Delta/E[B]$, $n = 1, 2$. This kind of mapping of the parameters between the model and a flow-level discrete event simulation has been used in many similar studies, see, e.g. [8, 12, 16].

The load in the system is defined as $\gamma = \lambda_1 E[B]/r_{1,2} + \lambda_2 E[B]/r_{2,2}$, and the maximal stability condition (with exact channel information) is $\gamma < 1$ [18]. In our simulation scenarios, we study the performance as a function of the load γ fixing $\lambda = \lambda_1 = \lambda_2$. Thus, for a given load γ , we have

$$\lambda = \frac{\gamma}{E[B](1/r_{1,2} + 1/r_{2,2})}.$$

Below we first consider scenarios, where the parameters, to be defined below, satisfy condition (26), i.e., $q_{k,1,2} > \mu_{k,1}/\mu_{k,2}$ for both classes k , which guarantees indexability by Theorem 2. We do this separately for the cases, where the flows are long and short. After that we also consider scenarios, where indexability is no longer guaranteed.

8.1. Scenarios with long flows

First we consider the case, where the flows are rather long with mean size $E[B] = 5$ Mbit. We consider three scenarios with varying rates and degrees of positive correlation under the sufficient indexability condition (26) for both classes k . The results are presented in Figure 1, where each panel shows the mean total number of flows in the system for different policies relative to the corresponding performance of our WhR policy as a function of the load γ .

In Scenario 1, we consider a setting where the rates in both classes are the same ($r_{1,1} = r_{2,1} = 0.01$ Mbit/s, $r_{1,2} = r_{2,2} = 50$ Mbit/s), but the channel dynamics are asymmetric ($q_{1,1,2} = 0.06$, $q_{1,2,2} = 0.96$, $q_{2,1,2} = 0.06$, $q_{2,2,2} = 0.86$), and thus have relatively high positive correlation with $\rho_1 = 0.9$ and $\rho_2 = 0.8$. The results are shown in Figure 1 (upper left panel). Rnd performs very poorly, which highlights the importance of the belief state information with highly positively correlated channels. On the other hand, the performance of PF gives a very loose lower bound, showing that exact information gives by far superior results. All policies utilizing belief state information perform quite well with WhM being somewhat better than the rest. Note that since the rates are symmetric in both classes, Mpc and Blf are identical policies in this scenario.

In Scenario 2, we change the maximum rate parameters to be highly asymmetric by reducing the maximum rate of class-2 flows to $r_{2,2} = 3$ Mbit/s. In this way, under the Mpc policy, the system behaves, in fact, as a strict priority system so that class-1 flows always have absolute priority over class-2 flows independent of the belief state. The results are depicted in Figure 1 (middle panel). The Rnd policy performs even worse and the PF slightly

Figure 1: Mean total number of flows in the system for different policies relative to WhR policy as a function of the load for Scenario 1 (upper left panel), Scenario 2 (upper right panel) and Scenario 3 (lower panel).

better than in Scenario 1. However, Mpc is indeed now showing clearly worse performance: because of the $c\mu$ -rule it is favoring too heavily the class-1 flows with very high $r_{1,2}$. The other policies using belief state information (WhM, WhR, Blf) all perform roughly equally well.

Finally, in Scenario 3, we reduce the degree of positive correlation significantly by choosing $\rho_1 = 0.4$ ($q_{1,1,2} = 0.3$, $q_{1,2,2} = 0.7$) and $\rho_2 = 0.1$ ($q_{2,1,2} = 0.6$, $q_{2,2,2} = 0.7$). Also, the maximum rates are asymmetric between the classes, $r_{1,2} = 33.6$ Mbit/s and $r_{2,2} = 16.8$ Mbit/s, but the minimum rates are the same, $r_{1,1} = r_{2,1} = 8.4$ Mbit/s. This scenario should favor the Mpc policy as it is known to be optimal if the channel statistics are i.i.d. between the time slots (i.e., there is no correlation). Also, the Rnd policy should benefit from this as the belief state information used by the other policies is not in this context as useful as in the case with higher positive correlation. The results are shown in Figure 1 (lower panel). Indeed, as predicted, Rnd is performing better than in the previous scenarios. Also, Mpc gives the best performance among the policies utilizing the belief state.

Both Whittle index policies (WhM, WhR) are also performing well, but Blf is clearly worse than all these.

Overall, the Whittle index policies (WhM, WhR) perform well in all these scenarios, independent of the degree of positive correlation in channel dynamics of the flows. This is in contrast with the properties of the more heuristic Mpc and Blf policies, which may change with the correlations.

8.2. Scenarios with short flows

Here we study the case, where flows are significantly shorter than previously with mean size $E[B] = 0.5$ Mbit. We consider two scenarios with asymmetric rates and high/low positive correlation. The parameters are still selected such that the sufficient indexability condition (26) holds for both classes k . The results are presented in Figure 2, where each panel shows the mean total number of flows in the system for different policies relative to the corresponding performance of our WhR policy as a function of the load γ . Here each simulation consists of $2 \cdot 10^6$ arrivals.

Figure 2: Mean total number of flows in the system for different policies relative to WhR policy as a function of the load for Scenario 4 (left panel) and Scenario 5 (right panel).

In Scenario 4, we have high positive correlation in the channel dynamics, and these parameters are the same as we used in Scenarios 1 and 2: $\rho_1 = 0.9$ ($q_{1,1,2} = 0.06$, $q_{1,2,2} = 0.96$) and $\rho_2 = 0.8$ ($q_{2,1,2} = 0.06$, $q_{2,2,2} = 0.86$). The rates are somewhat asymmetric between the classes such that for class-1 flows we have $r_{1,1} = 1$ Mbit/s and $r_{1,2} = 20$ Mbit/s, and for class-2 flows $r_{1,1} = 1$ Mbit/s and $r_{1,2} = 18$ Mbit/s. As the rate asymmetry is not that extreme, this scenario is, in fact, close to Scenario 1. The results are depicted in Figure 2 (left panel). The Rnd policy performs poorly, as expected. Comparing to these (see Figure 1 upper panels), now at higher load both WhR

and WhM are performing better than the heuristic policies Blf and Mpc with the WhM policy being overall the best. At lower loads, the WhR policy is slightly worse than the Blf and Mpc policies.

Next, in Scenario 5, the rates are the same as in Scenario 4 but the channel dynamics are the same as in Scenario 3 with low positive correlation, i.e., we have $\rho_1 = 0.4$ ($q_{1,1,2} = 0.3$, $q_{1,2,2} = 0.7$) and $\rho_2 = 0.1$ ($q_{2,1,2} = 0.6$, $q_{2,2,2} = 0.7$). This favors the Mpc policy, as mentioned earlier. The results are given in Figure 2 (right panel). As in Scenario 3 earlier, Rnd policy is indeed performing better but it is of course worse than all others. Notably, now the WhM policy still remains the best performing policy and not Mpc as in Scenario 3, see Figure 1 (lower panel). In both Scenarios 4 and 5, the PF policy continues naturally to perform best.

In summary, it appears that when the flows sizes are smaller, which effectively increases the completion probabilities $\mu_{k,i}$ in our model, our Whittle-based policies, in particular the WhM policy with the myopic tie-breaking rule, start performing better relative to the heuristic Blf and Mpc policies.

8.3. Scenarios where indexability not guaranteed

Finally, we experiment with scenarios, where the sufficient indexability condition (26) no longer holds, but we still apply, in our WhR and WhM policies, the Whittle index values derived under this assumption. More precisely said, we consider the case (47), i.e., $q_{k,1,2} < \mu_{k,1}/\mu_{k,2} < q_{k,2}$ for both classes k , where indexability is not guaranteed. This is achieved by using previous scenarios as a baseline and increasing significantly the rate of the channel state 1, $r_{k,1}$, which increases $\mu_{k,1}$. The results are presented in Figure 3, where each panel shows the mean total number of flows in the system for different policies relative to the corresponding performance of our WhR policy as a function of the load γ .

In Scenario 6, the channel dynamics have high positive correlation with the same parameters as earlier in such scenarios and the flows are long with $E[B] = 5$ Mbit, as in Section 8.1. The rates are asymmetric such that for class-1 flows we have $r_{1,1} = 10$ Mbit/s and $r_{1,2} = 50$ Mbit/s, and for class-2 flows $r_{1,1} = 10$ Mbit/s and $r_{1,2} = 25$ Mbit/s. As the rate asymmetry is not that high, this scenario is close to Scenario 1 (see Figure 1, left panel). The results of Scenario 6 with 10^6 arrivals for each simulation are shown in Figure 3 (left panel). Different to all earlier scenarios, now the Blf policy performs overall best at higher load. The performance of the Mpc and WhM

Figure 3: Mean total number of flows in the system for different policies relative to WhR policy as a function of the load for Scenario 6 (left panel) and Scenario 7 (right panel).

policies are very close to each other and slightly worse than for Blf. Rnd and PF policies continue to behave as earlier.

In Scenario 7, we continue having high positive correlation, but the flows are short with $E[B] = 0.5$ Mbit. The rates are for class-1 flows $r_{1,1} = 5$ Mbit/s and $r_{1,2} = 20$ Mbit/s, and for class-2 flows $r_{1,1} = 2$ Mbit/s and $r_{1,2} = 18$ Mbit/s. Thus, this scenario is similar to Scenario 4, but the rates of channel state 1 have been increased. The results with $2 \cdot 10^6$ arrivals in each simulation are given in Figure 3 (right panel). Comparing these results to Scenario 4, see Figure 2 (left panel), it can be seen that the results are very similar with the WhM policy performing best. Rnd and PF policies continue to behave as earlier.

Overall, the results here demonstrate that the Whittle index based policies can still continue to perform well even when the indexability condition is not necessarily valid anymore.

9. Conclusions

We considered flow-level opportunistic scheduling problem of downlink data traffic with partial channel information, where the channel dynamics are described by two-state positively autocorrelated Markov processes and the scheduler estimates the channel state based on ARQ-type feedback from the scheduled users.

Since the original problem belongs to the class of restless bandit problems, which are known to be mathematically intractable, the Whittle index approach has been utilized in earlier studies to develop near-optimal scheduling policies for the corresponding problems. We used the same approach and

managed to extend the results found thus far: We were able to find sufficient conditions under which the problem with partial channel information is provably indexable, derived an explicit formula for the corresponding Whittle index, and proved that the optimal policy for the relaxed optimization problem is of threshold type in these cases. In addition, we demonstrated that, for a certain range of parameter values, it is possible that the optimal policy for the relaxed problem is not of threshold type, which significantly complicates the task of formally proving the indexability property.

In our numerical simulations with dynamic flow-level traffic, the derived Whittle index policies performed well, independent of the degree of positive correlation in the channel dynamics and the size of the flows, which affects the difference in the time scales between the time slots and the flow arrivals and departures. Even in cases where the indexability property is not guaranteed, the Whittle index based policies demonstrated good performance.

As further work, it would be interesting to find out whether this flow-level opportunistic scheduling problem with partial channel information, where the channel dynamics are described by two-state positively autocorrelated Markov processes, is indexable for any parameter values. Furthermore, the case with negatively autocorrelated channels remains yet unexplored. Under this condition the ordering of the belief states is reversed, but similar sufficient conditions for indexability may be possible to establish. Technically even harder would still be to consider the corresponding problems with more than two channel states.

Acknowledgements

We thank the anonymous reviewers for their insightful comments that helped us in improving the paper.

References

- [1] S. Aalto, P. Lassila, and P. Osti. Whittle index approach to size-aware scheduling with time-varying channels. In *Proc. of ACM Sigmetrics*, pages 57–69, June 2015.
- [2] S. Aalto, P. Lassila, and P. Osti. Whittle index approach to size-aware scheduling for time-varying channels with multiple states. *Queueing Systems*, 83:195–225, 2016.

- [3] S. Aalto, P. Lassila, and P. Osti. Opportunistic scheduling with flow size information for Markovian time-varying channels. *Performance Evaluation*, 112:27–52, 2017.
- [4] S. Aalto, P. Lassila, and I. Taboada. Indexability of an opportunistic scheduling problem with partial channel information. In *Proc. of VALUETOOLS*, pages 95–102, March 2019.
- [5] S.H.A. Ahmad, M. Liu, T. Javadi, Q. Zhao, and B. Krishnamachari. Optimality of myopic sensing in multichannel opportunistic access. *IEEE Transactions on Information Theory*, 55:4040–4050, 2009.
- [6] A. Anand and G. de Veciana. A Whittle’s index based approach for QoE optimization in wireless networks. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2(1), 2018.
- [7] A. Asadi and V. Mancuso. A survey on opportunistic scheduling in wireless communications. *IEEE Communications Surveys & Tutorials*, 15:1671–1688, 2013.
- [8] U. Ayesta, M. Erausquin, and P. Jacko. A modeling framework for optimizing the flow-level scheduling with time-varying channels. *Performance Evaluation*, 67:1014–1029, 2010.
- [9] U. Ayesta, M. Erausquin, M. Jonckheere, and I.M. Verloop. Scheduling in a random environment: Stability and asymptotic optimality. *IEEE/ACM Transactions on Networking*, 21:258–271, 2013.
- [10] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi. CDMA/HDR: A bandwidth efficient high speed wireless data service for nomadic users. *IEEE Communications Magazine*, 38(7):70–77, 2000.
- [11] C. Buyukkoc, P. Varaiya, and J. Walrand. The $c\mu$ rule revisited. *Advances in Applied Probability*, 17:237–238, 1985.
- [12] F. Cecchi and P. Jacko. Scheduling of users with Markovian time-varying transmission rates. In *Proc. of ACM Sigmetrics*, pages 129–140, June 2013.

- [13] F. Cecchi and P. Jacko. Nearly-optimal scheduling of users with Markovian time-varying transmission rates. *Performance Evaluation*, 99-100:16–36, 2016.
- [14] P. Jacko. Value of information in optimal flow-level scheduling of users with Markovian time-varying channels. *Performance Evaluation*, 68:1022–1036, 2011.
- [15] P. Jacko, E. Morozov, L. Potakhina, and I.M. Verloop. Maximal flow-level stability of best-rate schedulers in heterogeneous wireless systems. *Transactions on Emerging Telecommunications Technologies*, 28(1), 2017.
- [16] P. Jacko and S.S. Villar. Opportunistic schedulers for optimal scheduling of flows in wireless systems with ARQ feedback. In *Proc. of ITC 24*, September 2012.
- [17] A. Jalali, R. Padovani, and R. Pankaj. Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system. In *Proc. of IEEE VTC 2000 Spring*, pages 1854–1858, May 2000.
- [18] J. Kim, B. Kim, J. Kim, and Y.H. Bae. Stability of flow-level scheduling with Markovian time-varying channels. *Performance Evaluation*, 70:148–159, 2013.
- [19] R. Knopp and P. Humblet. Information capacity and power control in single-cell multiuser communications. In *Proc. of IEEE ICC*, pages 331–335, June 1995.
- [20] G. Koole, Z. Liu, and R. Righter. Optimal transmission policies for noisy channels. *Operations Research*, 49:892–899, 2001.
- [21] M. Larrañaga, M. Assaad, A. Destounis, and G.S. Paschos. Asymptotically optimal pilot allocation over Markovian fading channels. *IEEE Transactions on Information Theory*, 64:5395–5418, 2018.
- [22] J. Le Ny, M. Dahleh, and E. Feron. Multi-UAV dynamic routing with partial observations using restless bandit allocation indices. In *Proc. of ACC*, pages 4220–4225, June 2008.

- [23] K. Liu, R. Weber, and Q. Zhao. Indexability and Whittle index for restless bandit problems involving reset processes. In *Proc. of IEEE CDC-ECC*, pages 7690–7696, December 2011.
- [24] K. Liu and Q. Zhao. Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access. *IEEE Transactions on Information Theory*, 56:5547–5567, 2010.
- [25] J. Niño-Mora. An index policy for dynamic fading-channel allocation to heterogeneous mobile users with partial observations. In *Proc. of NGI*, April 2008.
- [26] W. Ouyang, A. Eryilmaz, and N. Shroff. Asymptotically optimal downlink scheduling over Markovian fading channels. In *Proc. of IEEE Infocom*, pages 1224–1232, March 2012.
- [27] W. Ouyang, A. Eryilmaz, and N. Shroff. Low-complexity optimal scheduling over time-correlated fading channels with ARQ feedback. *IEEE Transactions on Mobile Computing*, 15:2275–2289, 2016.
- [28] W. Ouyang, S. Murugesan, A. Eryilmaz, and N. Shroff. Exploiting channel memory for joint estimation and scheduling in downlink networks. In *Proc. of IEEE Infocom*, pages 3056–3064, April 2011.
- [29] S.M. Ross. *Applied Probability Models with Optimization Applications*. Holden-Day, 1970.
- [30] I. Taboada, P. Jacko, U. Ayesta, and F. Liberal. Opportunistic scheduling of flows with general size distribution in wireless time-varying channels. In *Proc. of ITC 26*, September 2014.
- [31] I. Taboada, F. Liberal, J.O. Fajardo, and B. Blanco. An index rule proposal for scheduling in mobile broadband networks with limited channel feedback. *Performance Evaluation*, 117:130–142, 2017.
- [32] 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. *Performance Evaluation*, 79:90–103, 2014.

- [33] P. Viswanath, D.N.C. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory*, 48:1277–1294, 2002.
- [34] K. Wang, J. Yu, L. Chen, and M. Win. Opportunistic scheduling revisited using restless bandits: Indexability and index policy. In *Proc. of IEEE Globecom*, December 2017.
- [35] R. Weber and G. Weiss. On an index policy for restless bandits. *Journal of Applied Probability*, 27:637–648, 1990.
- [36] P. Whittle. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25A:287–298, 1988.
- [37] Q. Zhao, B. Krishnamachari, and K. Liu. On myopic sensing for multichannel opportunistic access: Structure, optimality, and performance. *IEEE Transactions on Wireless Communications*, 7:5431–5440, 2008.