Minimizing Age-of-Information in Heterogeneous Multi-Channel Systems: A New Partial-Index Approach

Yihan Zou, Kwang Taik Kim, Xiaojun Lin, and Mung Chiang

Electrical and Computer Engineering, Purdue University

Emails: {zou59, kimkt, linx, chiang}@purdue.edu

Abstract

We study how to schedule data sources in a wireless time-sensitive information system with multiple heterogeneous and unreliable channels to minimize the total expected Age-of-Information (AoI). Although one could formulate this problem as a discrete-time Markov Decision Process (MDP), such an approach suffers from the curse of dimensionality and lack of insights. For single-channel systems, prior studies have developed lower-complexity solutions based on the Whittle index. However, Whittle index has not been studied for systems with multiple heterogeneous channels, mainly because indexability is not well defined when there are multiple dual cost values, one for each channel. To overcome this difficulty, we introduce new notions of partial indexability and partial index, which are defined with respect to one channel's cost, given all other channels' costs. We then combine the ideas of partial indices and max-weight matching to develop a Sum Weighted Index Matching (SWIM) policy, which iteratively updates the dual costs and partial indices. The proposed policy is shown to be asymptotically optimal in minimizing the total expected AoI, under a technical condition on a global attractor property. Extensive performance simulations demonstrate that the proposed policy offers significant gains over conventional approaches by achieving a near-optimal AoI. Further, the notion of partial index is of independent interest and could be useful for other problems with multiple heterogeneous resources.

Index Terms

Age-of-Information, Whittle Index, restless bandits, Markov decision processes, heterogeneous channels.

1. Introduction

Many emerging wireless applications (e.g., real-time control in robotics systems, and data collection for IoT applications) rely on timely status updates from information sources [1], [2]. In these applications, oftentimes only the information with the latest timestamp is valuable to the receiver, while out-dated packets have little value. These applications have motivated a growing body of literature in optimizing the Age-of-Information (AoI), which is defined as the elapsed time of the last-received information packet since it was generated (at the source). Intuitively, AoI captures the freshness of information from the data source's perspective, and is considered a more useful metric for time-sensitive information systems than packet-level delays [3].

In this paper, we are interested in minimizing AoI for a wireless system with multiple heterogeneous sources and channels. This is a difficult setting that still lacks effective solutions in the literature. Many existing work on minimum-AoI scheduling policies study only a *single-source* system [3], [4]. For multiple sources, most of the existing work assumes that data sources are transmitting in a *single shared channel* [5], [6], [7], [8], [9], [10]. Further, most of these studies assume the channel to be reliable, with only a few extension to the case of a single unreliable channel [6], [9], [10]. For studies that do involve multiple channels, a recent article [11] assumes a *homogeneous* channel model, where each user-channel pair has equal ON/OFF probability. Thus, the solutions in these studies cannot be used in wireless systems that exhibit heterogeneous channel condition (e.g., transmission success probability) for each source-channel pair, which is common due to antenna beamforming, frequency selectivity and location-dependent fading [12], [13], [14]. [15] and [16] are the closest work to ours, as they also study heterogeneous multi-channel systems. [15] proposes a scheduling policy for ON/OFF multi-channel systems based on max-age matching. Under a similar setting, [16] proposes a policy that is asymptotically 8-optimal in minimizing the total weighted age. However, [15] and [16] assume that the ON/OFF states of all channels are known before the scheduling decisions are made. This assumption, combined with the setting that the number of channels are large, ensures that with high probability each source sees at least one ON channel. In this way, the impact of unreliable channels can be absorbed by an event with negligible probability in their analysis. In contrast, we are interested in a model where the channel states are unknown when scheduling decisions are made. Therefore, the question remains open on *how to design a provably optimal scheduling policy to minimize AoI in time-sensitive information systems with multiple heterogeneous and unreliable channels.*

One of the key obstacles in deriving the optimal scheduling policy under multiple heterogeneous sources and channels is the complexity of the associated Markov decision problem. Note that such AoI optimization problems (regardless of the channel conditions) are often formulated as Markov Decision Processes (MDP) or Restless Multi-armed Bandits (RMAB), which in theory can be optimally solved by value iteration [17], [18]. However, this approach suffers from *the curse of dimensionality* and lack of insights. Therefore, it is highly desirable to develop low-complexity and near-optimal solutions. For single-channel systems, policies based on Whittle index [19], whose complexity does not grow with the number of sources, have been found to exhibit good performance. Further, they are known to be asymptotically optimal when the number of sources and the channel capacity both grow to infinity [20], [8], [7], [6], [9]. However, to the best of our knowledge, there have been no such Whittle index policies for systems with multiple heterogeneous channels/resources. Part of the difficulty is that Whittle's notion of "indexability" [19] is not well-defined when there are multiple heterogeneous channels. Specifically, in [19], a project is indexable if there is a single threshold for the channel cost, above which the optimal action of the project will be passive (i.e., not to consume the channel resource). Thus, Whittle indexability critically relies on the assumption that there is only one dual cost for either a single channel or a single group of homogeneous channels. For heterogeneous multi-channel systems, each channel naturally has a different dual cost. The optimal action of the project will also depend on all channel costs. As a result, one cannot even define such a threshold or index.

In this paper, we propose a new Whittle-like scheduling policy for heterogeneous and unreliable multi-channel systems. Similar to [19], we first formulate the MDP for the system, and decompose the problem into per-source sub-problems using Lagrange relaxation [18] (Section 2). However, to overcome the difficulty of Whittle indexability as mentioned above, we introduce the new notions of *partial indexability* and partial index, which are defined with respect to the cost of one channel, given the costs of all other channels (see Section 3 for detailed definitions). Then, we propose a low-complexity Whittle-like scheduling policy, which we call the Sum Weighted Index Matching (SWIM) policy, by computing a maximum-weighted matching (MWM) between the sources and channels, where the weight between each source-channel pair is the above-defined partial index. Our key contribution in Section 3 is to identify a precise-division condition, under which the SWIM policy is asymptotically optimal, under a technical assumption on a global attractor property (which has also been used in the literature [21], [22]). To the best of our knowledge, our work is the first in the literature to extend the concept of indexability to heterogeneous multi-channel settings. We note that both the notion of partial indexability and the SWIM policy are very general, and can be applied to various large-scale MDP problems with multiple heterogeneous channels. We then verify in Section 4 that our AoI problem indeed satisfies the partial indexability and precise-division property. Our simulation results in Section 5 shows that applying the SWIM policy to our AoI problem produces significant performance gains over conventional approaches, and achieves a near-optimal average AoI.

We note that in the RMAB literature there is also a line of work on multi-action bandits [21], [22]. However, we emphasize that "actions" and "channels" are very different, because the multiple actions in [21], [22] are still applied to a *single resource*. This is the reason why [21] can still define a Whittle index based on the (single) dual cost associated with the resource. In contrast, multiple heterogeneous channels correspond to *multiple resources* and multiple dual costs. Therefore, the techniques in [21] cannot be directly applied to our setting with heterogeneous channels.

2. Model and Problem Formulation

We consider a wireless system where a base station (BS) is scheduling N data sources or sensors on multiple channels for timely status updates in the uplink (Fig. 1). Each source corresponds to one sender node on the left in Fig. 1. Note that each source may experience different channel conditions due to their locations. As a result, sources may have different preferences on the set of communication channels. To model such heterogeneity, we assume that the sources are divided into G groups. Let $\mathcal{N}_g$ be the set of sources in group g. Then, the set of all sources is $\mathcal{N} = \bigcup_{g=1}^{G} \mathcal{N}_g$. The sources $n \in \mathcal{N}_g$ in the same group g experience the same condition on each channel. We consider a discrete-time system where time is indexed by $t \in \mathcal{T}$. We assume that the transmission from the source to the BS takes one time slot.

Heterogeneous and Unreliable Channels: As shown in Fig. 1, the BS is capable of communicating in multiple channels at each time. Depending on the frequency, modulation, and beam-forming schemes used, the channels may have similar or different qualities. To model such heterogeneity, we divide the channels also into $M > 1$ types. We assume that each type $m \in \mathcal{M} = \{1, \ldots, M\}$ of channels has $\overline{C}$ identical instances (which we refer to in the future as "channels"). As we explain below, all sources in a given group g sees the same channel quality in the $\overline{C}$ channel instances of a given type m. We adopt the standard collision channel model [9] as follows. At each time, the BS can schedule at most one source to transmit update packets on each channel. The channel is potentially *unreliable*, due to wireless channel fading. In contrast to most existing work in the literature, we consider *heterogeneous source-channel conditions*. Specifically, we assume that each transmission from source $n \in \mathcal{N}_g$ on a channel of type $m \in \mathcal{M}$ succeeds with probability $p_{gm} \in (0, 1]$, independently from all other transmissions. We denote the *channel quality vector* for group g as $\vec{p}_g = [p_{g1}, \ldots, p_{gM}]^T$. With slight abuse of notation, we denote $p_{nm} = p_{g(n),m}$ where $g(n)$ is the group containing n.

Packet Generation: To focus our discussion on the effect of multiple heterogeneous channels, we adopt the generate-at-will model as [5], [6]. Specifically, whenever a source is scheduled for transmission, it can generate a fresh update. In this work, we use age-of-information (or simply age) to measure the information freshness, which is defined as the elapsed time of the last-received information packet since it was generated (at the source). Denote $h_n(t)$ as the age of source n at time t. If the transmission is successful, the age of this source reduces to 1. If the source is not scheduled for transmission, or if the transmission fails, the age increases by 1. Then, the AoI evolution of source n can be written as

$$h_n(t+1) = \begin{cases} 1, & \text{successful update from } n; \\ h_n(t) + 1, & \text{otherwise.} \end{cases} \tag{1}$$

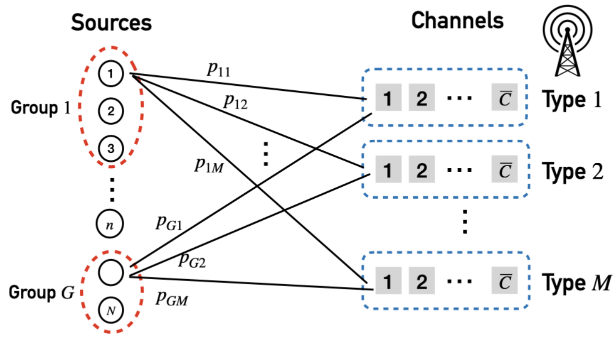


Fig. 1. Uplink of a heterogeneous, unreliable multi-channel system with a base station and N data sources.

Intuitively, to avoid wasting channel resources, the BS should only schedule each source on at most one channel instance. Let $u_n(t)$ be the decision variable at time t such that $u_n(t) = m$ if source n is scheduled to transmit on channel of type m, and $u_n(t) = 0$ if the source is not scheduled for transmission. In summary, we have the constraints that $\sum_{n=1}^{N} \mathbb{1}\{u_n(t) = m\} \leq \overline{C}$ for all channel type m, and $\sum_{m=1}^{M} \mathbb{1}\{u_n(t) = m\} \leq 1$ for all source n.

A. MDP-based Formulation

Now, we can formulate the average AoI minimization problem for the above heterogeneous and unreliable multi-channel system as an MDP. Let $\overline{\mathcal{S}}(t) \triangleq \{h_1(t), h_2(t), \ldots, h_N(t)\} \in \mathbb{N}_+^N$ be the system state at time t. Denote the action space of the entire system as $\overline{\mathcal{U}} \triangleq \{0, 1, \ldots, M\}^N$. (Recall that action 0 denotes no scheduled transmission and action $m \in \mathcal{M}$ denotes the scheduled channel type). A policy $\boldsymbol{\pi}$ maps from the system state $\overline{\mathcal{S}}(t)$ to the action in $\overline{\mathcal{U}}$. The state transition probability of source n when it is passive is

$$\mathbb{P}\{h_n(t+1) = d+1 | h_n(t) = d, u_n(t) = 0\} = 1. \tag{2}$$

The state transition probabilities when source n is scheduled on a channel of type m are

$$\begin{aligned}
&\mathbb{P}\{h_n(t+1) = d+1 | h_n(t) = d, u_n(t) = m\} = 1 - p_{nm}, \\
&\mathbb{P}\{h_n(t+1) = 1 | h_n(t) = d, u_n(t) = m\} = p_{nm}.
\end{aligned} \tag{3}$$

We can define the T-horizon average AoI and the long-term average AoI of the system under policy $\boldsymbol{\pi}$ as

$$H_{\boldsymbol{\pi}}^{(T)} = \frac{1}{TN} \sum_{t=1}^{T} \sum_{n=1}^{N} \mathbb{E}[h_n^{\boldsymbol{\pi}}(t)], \text{ and } \overline{H}_{\boldsymbol{\pi}} \triangleq \limsup_{T \to \infty} H_{\boldsymbol{\pi}}^{(T)}, \tag{4}$$

respectively, where T is the length of time horizon, and $h_n^{\boldsymbol{\pi}}(t)$ is the AoI of source n at time t under policy $\boldsymbol{\pi}$. The objective of the MDP is to minimize the long-term average system AoI in (4), i.e.,

$$\min_{\boldsymbol{\pi}\in\overline{\mathcal{U}}^T} \limsup_{T\to\infty} \frac{1}{TN} \sum\nolimits_{t=1}^{T} \sum\nolimits_{n=1}^{N} \mathbb{E}\left[h_n^{\boldsymbol{\pi}}(t)\right]. \tag{5}$$

In theory, the above MDP can be solved optimally as an infinite-horizon average cost per stage problem using relative value iteration [17]. However, this approach suffers from the curse of dimensionality and lack of insights for the solution structure. Hence, many efforts have been focusing on developing low-complexity solutions.

B. Decomposition Using Lagrange Relaxation

For lower-complexity solutions, two representative approaches in the literature are based on the relaxed problem and index policies. In this section, we will discuss how they are related to a Lagrange relaxation of the MDP, and the challenges of applying these existing approaches to our setting with multiple heterogeneous channels.

We first introduce the relaxed problem. Denote $u_{nm}^{\boldsymbol{\pi}}(t) \triangleq \mathbb{1}\{u_n(t)=m\}$, i.e., the indicator variable that source n is scheduled on channel type m at time t under policy $\boldsymbol{\pi}$. The MDP formulated in Section 2-A can be equivalently written in the following optimization form:

$$\begin{aligned}
\underset{\boldsymbol{\pi}}{\text{minimize}} \quad & \limsup_{T\to\infty} \frac{1}{TN} \sum\nolimits_{t=1}^{T} \sum\nolimits_{n=1}^{N} \mathbb{E}\left[h_n^{\boldsymbol{\pi}}(t)\right] \\
\text{subject to} \quad & \sum\nolimits_{n=1}^{N} u_{nm}^{\boldsymbol{\pi}}(t) \le \overline{C}, \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, && \text{(6a)} \\
& \sum\nolimits_{m=1}^{M} u_{nm}^{\boldsymbol{\pi}}(t) \le 1, \quad \forall n \in \mathcal{N}, t \in \mathcal{T}, && \text{(6b)} \\
& u_{nm}^{\boldsymbol{\pi}}(t) \in \{0,1\}, \qquad \forall t \in \mathcal{T}. && \text{(6c)}
\end{aligned}$$

Following Whittle's approach [19], we relax the instantaneous constraint (6a) to an average constraint, and obtain the *relaxed problem*

$$\begin{aligned}
\underset{\boldsymbol{\pi}}{\text{minimize}} \quad & \limsup_{T\to\infty} \frac{1}{TN} \sum\nolimits_{t=1}^{T} \sum\nolimits_{n=1}^{N} \mathbb{E}\left[h_n^{\boldsymbol{\pi}}(t)\right] \\
\text{subject to} \quad & \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}[\sum\nolimits_{t=1}^{T} \sum\nolimits_{n=1}^{N} u_{nm}^{\boldsymbol{\pi}}(t)] \le \overline{C}, \quad \forall m \in \mathcal{M}, && \text{(7a)} \\
& \text{(6b)}, \text{(6c)} && \text{(7b)}
\end{aligned}$$

Next, we use Lagrange relaxation in [18, Chapter 6]. Specifically, we introduce a dual cost λ_m to each of (7a), and decouple the relaxed problem of (7) into N *sub-problems*, i.e., $\forall n \in \mathcal{N}$,

$$\underset{\boldsymbol{\pi}}{\text{minimize}} \quad \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\big[h_n^{\boldsymbol{\pi}}(t) + \sum_{m\in\mathcal{M}} \lambda_m u_{nm}^{\boldsymbol{\pi}}(t)\big] \tag{8a}$$

$$\text{subject to} \quad \sum\nolimits_{m=1}^{M} u_{nm}^{\boldsymbol{\pi}}(t) \leq 1, \qquad \forall t \in \mathcal{T}, \tag{8b}$$

$$u_{nm}^{\boldsymbol{\pi}}(t) \in \{0,1\}, \qquad \forall t \in \mathcal{T}, \forall m \in \mathcal{M}. \tag{8c}$$

It is easy to see that, given channel costs $\vec{\lambda} = [\lambda_m, \forall m \in \mathcal{M}]$, each sub-problem (8) is an average cost per stage problem [17] for optimizing the long-term average AoI of source n plus the costs of its channel use, which by itself is a decoupled MDP with state space $\mathcal{S} = \mathbb{N}_+$, action space $\mathcal{U} = \{0, 1, \ldots, M\}$, and transition probabilities in (2) and (3).

With this dual decomposition, the relaxed problem can be solved by iteratively solving all independent sub-problems (8) given the current $\vec{\lambda}^{(k)}$ (denote the resulting decisions by $u_{nm}^{\boldsymbol{\pi}(k)}(t)$) and updating the dual costs $\vec{\lambda}^{(k)}$ by dual gradient ascent in the k-th iteration [23], i.e., for all $m \in \mathcal{M}$

$$\lambda_m^{(k+1)} = \left[\lambda_m^{(k)} + \gamma \cdot \left(\mathbb{E}\left[\sum\nolimits_{n \in \mathcal{N}} u_{nm}^{\boldsymbol{\pi}(k)}(t)\right] - \overline{C}\right)\right]^+ \tag{9}$$

where $\gamma > 0$ is the step size, and $[x]^+ = \max\{x, 0\}$. When $\vec{\lambda}^{(k)}$ converges, the corresponding solution of the relaxed problem is known to provide us with a lower bound for the objective of the original MDP (6) [20]. However, this solution does not always provide a feasible scheduling decision, because, in the real system, resource constraints (6a) must be met at all time, instead of just in the average sense as in (7a). Moreover, before the primal-dual iteration converges, the average constraint may be violated severely, resulting in poor policy performance.

For single-channel systems ($M = 1$), Whittle's index policy in the literature overcomes these drawbacks by producing a scheduling decision that is always feasible, and is near-optimal [19], [20], [18], [9]. However, Whittle index or indexability have not been defined for multiple heterogeneous channels. Note that in the RMAB literature, the indexability is defined based on the following: for each state of a project, there exists a *scalar* price threshold such that, when the price is above (or below) that threshold, the resource is not used (or used). However, as we have shown above, the sub-problem (8) in our model is parameterized by multiple channel costs with distinct values. Obviously, the decision to use each channel type m depends on not only the cost λ_m of this type, but also the costs of other channel types. As a result, there is no longer a single threshold that can divide the spaces of cost vectors into one where the resource is used, and the opposite one where the resource is not used. Next, we overcome this difficulty by introducing the new notions of *partial indexability* and *partial index*.

3. Partial Indexability and Asymptotically optimal Policies

In this section, we will propose a powerful framework to design asymptotically optimal scheduling policies, which generalizes the notion of indexability to heterogeneous multi-channel settings. Specifically, we introduce a new notion of *partial indexability*, which are defined with respect to the cost of one channel, given the costs of the others. Partial indexability and the corresponding partial index then allow us to develop a near-optimal policy for heterogeneous multi-channel systems, which is a key contribution of our work.

Our proposed solution framework in this section is based on only the relaxed-problem formulation in Section 2-B. Note that the formulation of the relaxed problem in Section 2-B can be applied to any MDP with the cost function given by $h(\cdot)$. Thus, our methodology not only applies to the AoI minimization problem in this paper, but also other large MDP problems with multiple heterogeneous channels (or resources). In that sense, the applicability of our proposed framework in this section is beyond the current problem. Thus, although we still use the notions of "sources/channels" in this section, they could be easily applied to more general notions of "projects/resources" as in the typical Whittle-index literature [20].

A. Partial Indexability

We first focus on the sub-problem (8) with a given vector $\vec{\lambda} = [\lambda_1, \ldots, \lambda_M]^T$ of costs for all channels. As we mentioned in Sec. 2-B, the MDP of each sub-problem is an infinite-horizon average cost per stage problem with countably infinite state space [17]. Since the sub-problems of all sources n in the same group g are identical and independent, next we can write the Bellman Equation of the sub-problem (8) for each group g as

$$f^g(s) + J^* = \min_{u \in \{0,\ldots,M\}} \left[g_u^g(s, \vec{\lambda}) + \sum\nolimits_{d \in \mathcal{S}} p_{sd}^g(u) f^g(d) \right], \tag{10}$$

where $f^g(\cdot)$ is the optimal relative value function, J^* is the optimal average cost. Here, to keep our notations general, we have used $g_u(s, \vec{\lambda}) = C_{gs}^u + \lambda_u$ to denote the stage cost in (8a) at state $s \in \mathcal{S}$ under action $u \in \mathcal{U} = \{0\} \bigcup \mathcal{M}$, and $p_{sd}^g(u)$ to denote the transition probability from state s to state d by taking action u. For the AoI minimization problem, $C_{gs}^u = h(s)$ and $p_{sd}^g(u)$ specializes to the transition probabilities in (2) and (3).

Next, we define the partial indexability and partial index that generalize Whittle's index [19]. Given the cost vector $\vec{\lambda}$, let

$$\mu_u^g(s, \vec{\lambda}) \triangleq g_u^g(s, \vec{\lambda}) + \sum\nolimits_{d \in \mathcal{S}} p_{sd}^g(u) f^g(d)$$

denote the expected cost-to-go from state s under action u, assuming that the optimal policy is used in the future. We first define the following concepts that are analogous to Whittle's notations [19].

Definition 3.1 (Passive Set)**.** *Given the cost vector $\vec{\lambda}$, the set of passive states for channel-type m is*

$$\mathcal{P}_m^g(\vec{\lambda}) \triangleq \{s \in \mathcal{S} | \mu_m^g(s, \vec{\lambda}) > \min_{u \neq m, u \geq 0} \mu_u^g(s, \vec{\lambda})\}. \tag{11}$$

In other words, if the current state of a source $n \in \mathcal{N}_g$ is $s \in \mathcal{P}_m^g(\vec{\lambda})$, the solution to the relaxed problem under $\vec{\lambda}$ will not schedule source n on channel-type m. Let $\vec{\lambda}_{-m}$ denote the cost vector of all channels except for channel type m. We now fix all channel costs $\vec{\lambda}_{-m}$ except that of type m, but vary the channel cost of type m to λ'_m. Let the new cost vector be $\vec{\lambda}' = [\lambda'_m, \vec{\lambda}_{-m}]$. We define the partial indexability as follows.

Definition 3.2 (Partial Indexability)**.** *Given the cost vector $\vec{\lambda}$, the sub-problem* (8) *is partially indexable (or indexable as abbr.) if, for all $m \in \mathcal{M}$, the size of the passive set $|\mathcal{P}_m^g(\vec{\lambda}')|$ increases monotonically to the entire state space as λ'_m increases from 0 to ∞ (while fixing other channels' costs $\vec{\lambda}_{-m}$).*

If the sub-problem (8) is partially indexable, then for each state s, there is a largest value of λ'_m such that the passive set $\mathcal{P}_m^g(\vec{\lambda}')$ no longer includes the state s. We refer to this value of λ'_m as the partial index, as defined below.

Definition 3.3 (Partial Index)**.** *Given channel vector $\vec{p}$ and cost vector $\vec{\lambda}$, the partial index (or index, as abbr.) $I_m^g(s, \vec{\lambda}_{-m})$ of state $s \in \mathcal{S}$ for channel type $m \in \mathcal{M}$ is defined as the supremum of cost λ'_m such that the expected cost-to-go from state s for using channel type m is no larger than that under any other actions, i.e.,*

$$I_m^g(s, \vec{\lambda}_{-m}) \triangleq \left[\sup\{\lambda'_m \big| \mu_m^g(s, \vec{\lambda}') \le \mu_k^g(s, \vec{\lambda}'), \forall k \ge 0\}\right]^+. \tag{12}$$

In addition, we define the index for passive action ($m = 0$) as

$$I_0^g(s, \vec{\lambda}) \triangleq \left[\sup\{\lambda' \big| \mu_0^g(s, \vec{\lambda}) + \lambda' \le \mu_k^g(s, \vec{\lambda}), \forall k \in \mathcal{M}\}\right]^-, \tag{13}$$

where $[x]^- \triangleq \min\{x, 0\}$.

Similar to Whittle policy, partial indexability allows us to characterize the urgency of each state by its indices, based on which an efficient solution for the original problem can be derived. However, in contrast to standard Whittle indexability, partial indexability is defined given *all the channel costs other than channel type m.* Like Whittle indexability, verifying such partial indexability is non-trivial, and often requires significant work. We will show how to verify partial indexability for the AoI minimization problem in Section 4.

Next, we are interested in designing a Whittle-like policy that can utilize partial indices. For single-channel system, the Whittle index policy simply picks the project with the highest index. However, such a simple decision will not work for multi-channel systems anymore, because each source is also restricted to transmit on one channel at a time. Intuitively, to respect the capacity constraints (6a) and (6b) for each channel and each source, the decision should involve some matching between sources and channels. The goal of the next section is to establish this matching formally.

B. Max-Weight Matching of Partial Indices

Motivated by Whittle's index policy, we aim to schedule a group of users with higher partial indices, while satisfying the resource constraints on each channels. The problem can be naturally formulated as a Maximum Weighted Matching problem based on partial indices (MWM-PI). Define the graph $\mathcal{R} \triangleq (\mathcal{N} \bigcup \mathcal{M}, \mathcal{E})$, where $\mathcal{E}$ is the set of all source-channel-type pairs. We define the problem of MWM-PI as follows,

$$\underset{y_{nm} \in \{0,1\}}{\text{maximize}} \quad \sum\nolimits_{n \in \mathcal{N}} \sum\nolimits_{m=0}^{M} w_{nm} y_{nm} \tag{14a}$$

$$\text{subject to} \quad \sum\nolimits_{n \in \mathcal{N}} y_{nm} \leq \overline{C}, \quad \forall m \in \mathcal{M}, \tag{14b}$$

$$\sum\nolimits_{m \in \mathcal{M}} y_{nm} \leq 1, \quad \forall n \in \mathcal{N} \tag{14c}$$

where y_{nm} is the binary decision to schedule source n on channel type m, $w_{nm} \triangleq I_m^{g(n)}(s_n, \vec{\lambda}_{-m})$ is the edge weight given by the partial index in (12) and (13), and $g(n)$ is the group index for user n. We then schedule the sources according to $u_{nm}(t) = y_{nm}$.

Note that MWM-PI is based on the current set of prices $\vec{\lambda}$. As we present next, the outcome of the MWM-PI will also guide us in updating the prices $\vec{\lambda}$. This idea leads to the proposed Sum Weighted Index Matching (SWIM) policy in Algorithm 1. Specifically, Line 1 initializes the system parameters. Lines 3-5 compute the scheduling decision for time t by solving the MWM-PI problem. Lines 6-7 correspond to the transmission phase of the update packets. Line 8 updates each channel type's cost for the next time $t+1$ as a weighted average (by the parameter β) of the previous channel cost and the optimal dual cost associated with (14b) at time t.

Remark. Clearly, Algorithm 1 is a generalization of Whittle's index policy. In fact, in the single-channel case, the MWM-PI reduces to Whittle's policy. The critical difference is that, in heterogeneous multi-channel systems, source n's index for channel type m depends on other channels' costs $\vec{\lambda}_{-m}$, whose optimal value also needs to be found. To address this difficulty, Algorithm 1 uses adaptive updates to approach the optimal channel costs in Line 8.

C. Fluid Analysis and Asymptotic Optimality

In the literature, the optimality of Whittle index policies is often shown using a fluid limit argument, by considering the regime of a large-scale system. Specifically, [20] shows that the difference between

Algorithm 1: Sum Weighted Index Matching (SWIM)

1 At $t = 0$: Initialize parameters $N, M, \overline{C}, \beta$, and $\vec{\lambda}(1)$;

2 At time $t \geq 1$:

3 Compute partial indices $\vec{I}^{g(n)}(t) = [I_m^{g(n)}(s_n(t), \vec{\lambda}_{-m}(t))]$ for every source $n \in \mathcal{N}$, given current cost $\vec{\lambda}(t)$;

4 Solve MWM-PI in (14) with $w_{nm} \leftarrow I_m^{g(n)}(s_n(t), \vec{\lambda}_{-m}(t))$, and obtains the scheduling decision $\vec{y}(t)$;

5 Schedule sources according to $\vec{u}(t) \stackrel{\Delta}{=} [u_{nm}(t)] = \vec{y}(t)$;

6 Wait for updates from scheduled sources on all channels;

7 Broadcast an ACK message to indicate all successful updates;

8 Update channel cost as $\lambda_m(t+1) \leftarrow (1-\beta)\lambda_m(t) + \beta\nu_m(t)$, where $\nu_m(t)$ is the optimal dual variable associated with (14b) for channel type m in the MWM-PI problem at time t.

the state distribution under the Whittle index policy and the steady-state distribution under the optimal policy for the relaxed problem (7) diminishes to zero, when $N, \overline{C} \to \infty$ and $\alpha = \overline{C}/N$ is kept constant. Similarly, in this section, we will focus on such a fluid limit. We will show that the fixed point of the MWM-PI problem is equivalent to that of the relaxed problem (i.e., when dual gradient descent on $\vec{\lambda}$ converges and when the steady-state distribution is reached). Since the optimal solution for the relaxed problem at the fixed point is a lower bound for the original MDP (6), the above-mentioned equivalence relationship is essential for establishing the asymptotic optimality of our proposed SWIM policy later.

We first define the fluid limit model of the relaxed problem and its fixed point as follows. For any group g, let z_{gs} be the fraction of sources of group g that is in state s, with $\sum_{s\in\mathcal{S}} z_{gs} = 1$. Thus, $\vec{z}_g \stackrel{\Delta}{=} [z_{gs}, s \in \mathcal{S}]$ denotes the state distribution vector of group g. Given the current cost vector $\vec{\lambda}$, we assume that the distribution under the relaxed policy has reached the steady state. Let $x_{gs}^u \in [0, 1]$ be the fraction of sources of state s in group g that is assigned to channel u by the relaxed policy $\boldsymbol{\pi}_{\text{rel}}$. We use $(\vec{x}, \vec{z}^*, \vec{\lambda}^*)$ to denote a fluid fixed point of the relaxed problem at steady state (i.e., when the dual gradient ascent on $\vec{\lambda}$ converges). Similar to the fluid analysis in [22], we can verify that, at the fixed-point channel cost $\vec{\lambda}^*$, $(\vec{x}, \vec{z}^*)$ also solves the following fluid problem (which is a linear program (LP))

$$\underset{\vec{x}, \vec{z}}{\text{minimize}} \quad \sum\nolimits_{g\in\mathcal{G}} \sum\nolimits_{s\in\mathcal{S}} \sum\nolimits_{u=0}^{M} z_{gs} C_{gs}^u x_{gs}^u \tag{15a}$$

subject to

$$\sum\nolimits_{g\in\mathcal{G}} \sum\nolimits_{s\in\mathcal{S}} z_{gs} x_{gs}^u \leq \overline{C}_u, \quad \forall u \in \mathcal{M}, \tag{15b}$$

$$\sum\nolimits_{u\in\mathcal{M}} x_{gs}^u \le 1. \quad \forall g\in\mathcal{G}, \forall s\in\mathcal{S}, \tag{15c}$$

$$\sum_{u\ge 0}\sum_{d\in\mathcal{S}} z_{gs}x_{gs}^u p_{sd}^g(u) = \sum_{u'\ge 0}\sum_{d\in\mathcal{S}} z_{gd}x_{gd}^{u'} p_{ds}^g(u'), \forall s\in\mathcal{S}, \forall g\in\mathcal{G}, \tag{15d}$$

where C_{gs}^u and $p_{sd}^g(u)$ are defined in (10) (recall that $u=0$ corresponds to passive). Thus, the primal and dual variables $(\vec{x}, \vec{z}^*, \vec{\lambda}^*)$ will satisfy the KKT conditions of (15). Similar to [22, Lemma 4.3], it can be shown that the optimal solution of the fluid problem (15), denoted as $V^*(\vec{x}, \vec{z}^*, \vec{\lambda}^*)$, is a lower bound for the original MDP.

For Algorithm 1, we can similarly define its fluid limit and fixed point as follows. Suppose that the steady state is reached. Denote the corresponding state distribution, channel cost vector and decision vector as $\vec{z}', \vec{\lambda}'$ and $\vec{y}$, respectively. Recall that MWM-PI is based on a set of dual costs $\vec{\lambda}$, and the edge weight is computed by $w_{gs}^m = I_m^g(s, \vec{\lambda}_{-m})$ for $m\in\mathcal{M}$, and $w_{gs}^0 = I_0^g(s,\vec{\lambda})$. At steady state, $(\vec{y}, \vec{z}', \vec{\lambda}')$ must solve the following fluid problem

$$\underset{\vec{y}}{\text{maximize}} \quad \sum\nolimits_{g\in\mathcal{G}}\sum\nolimits_{s\in\mathcal{S}}\sum\nolimits_{u=0}^{M} z_{gs}w_{gs}^u y_{gs}^u \tag{16a}$$

$$\text{subject to} \quad \sum\nolimits_{g\in\mathcal{G}}\sum\nolimits_{s\in\mathcal{S}} z_{gs}y_{gs}^u \le \overline{C}_u, \quad \forall u\in\mathcal{M}, \tag{16b}$$

$$\sum\nolimits_{u=0}^{M} y_{gs}^u = 1, \quad \forall g\in\mathcal{G}, \forall s\in\mathcal{S} \qquad , \tag{16c}$$

$$\text{replace } \vec{x} \text{ by } \vec{y} \text{ in (15d).} \tag{16d}$$

Denote the Lagrange multiplier associated with (16b) as ν_u (define $\nu_0 = 0$). At the fixed point, $\vec{\lambda}' = \vec{\nu}$ must hold. Thus, we denote such $(\vec{y}, \vec{z}', \vec{\lambda}')$ as the fixed point of Algorithm 1, which should also satisfy the KKT conditions of (16).

Ideally, our goal is to show that the fixed point of the relaxed problem is identical to that of Algorithm 1, and thus they produce the same near-optimal objective value. However, we need a slightly stronger condition than partial indexability, as follow.

Definition 3.4 (Precise Division)**.** *Given state space $\mathcal{S}$ and channel costs $\vec{\lambda}$, suppose that the sub-problem* (10) *is partially indexable with index $I_m^g(s, \vec{\lambda}_{-m})$ for $s\in\mathcal{S}$. We say that the preference for channel-type m is precisely divisible by its partial-index $I_m^g(s,\vec{\lambda}_{-m})$, if the following holds: for all $s\in\mathcal{S}$ and $m\ge 1$,*

(i) If $I_m^g(s,\vec{\lambda}_{-m}) = \lambda_m$, then $\mu_m^g(s,\vec{\lambda}) \le \mu_u^g(s,\vec{\lambda}), \forall u\ne m, u\ge 0$.

(ii) If $I_m^g(s,\vec{\lambda}_{-m}) > \lambda_m$, then $\mu_m^g(s,\vec{\lambda}) < \mu_u^g(s,\vec{\lambda}), \forall u\ne m, u\ge 0$.

(iii) Otherwise, there exists $u\ne m, u\ge 0$ s.t. $\mu_m^g(s,\vec{\lambda}) > \mu_u^g(s,\vec{\lambda})$.

Note that Definition 3.4 implies partial indexability in Definition 3.2. To see this, note that given $\vec{\lambda}_{-m}$, for any state $s\in\mathcal{S}$, its partial index $I_m^g(s,\vec{\lambda}_{-m})$ is independent of λ_m. Thus, as λ_m increases, we transition from $I_m^g(s, \vec{\lambda}_{-m}) > \lambda_m$ (i.e., using channel type m per Definition 3.4-(ii)) to $I_m^g(s, \vec{\lambda}_{-m}) < \lambda_m$ (i.e., not using channel type m per Definition 3.4-(iii)). In other words, as λ_m increases, $\mathcal{P}_m^g(\vec{\lambda})$ increases monotonically to the entire state space $\mathcal{S}$. On the other hand, Definition 3.4 is stronger than partial indexability because it states that this transition occurs precisely at $I_m^g(s, \vec{\lambda}_{-m}) = \lambda_m$.

Condition 3.5. *The sub-problem* (10) *is partially indexable (Definition 3.2) and satisfies the precise division property in Definition 3.4*

The next theorem, which is one of our main contributions in this work, establishes the connection between the fixed point of the relaxed problem (7) and the fixed point of Algorithm 1.

Theorem 3.6. *Suppose that Condition 3.5 holds. Then, any fixed-point solutions $\{\vec{x}, \vec{z}^*, \vec{\lambda}^*\}$ of the relaxed problem are equivalent to the fixed-point solutions $\{\vec{y}, \vec{z}', \vec{\lambda}'\}$ of Algorithm 1 at the fluid limit.*

Proof. Here, we provide the complete proof for Theorem 3.6 (including the missing details in our conference paper). To prove the equivalence in the theorem, we need to show the statement in both direction. We first consider the fixed point of the relaxed problem. For any group g, let z_{gs} be the fraction of users of group g that is in state s, i.e., $\sum_{s\in\mathcal{S}} z_{gs} = 1$. Thus, $\vec{z_g} \triangleq [z_{gs}]_{s\in\mathcal{S}}$ denotes the state distribution vector of group g. We consider the case where the distribution of the relaxed problem has reached the steady state. Let $x_{gs}^m \in [0, 1]$ be the fraction of users of state s in group g that is assigned to channel m by the relaxed policy $\boldsymbol{\pi}_{\text{rel}}$. Therefore, the corresponding decisions of the relaxed problem at the fixed point $(\vec{x}, \vec{z}^*, \vec{\lambda}^*)$ of the relaxed policy have to satisfy the following KKT conditions of the relaxed problem (15).

i) (Dual feasible) The dual cost vector satisfies $\vec{\lambda}^* \geq \vec{0}$.

ii) The steady state user distribution $[\vec{z}_1^*, \ldots, \vec{z}_G^*]$ satisfies (15d) i.e., $\forall s \in \mathcal{S}, \forall g \in \mathcal{G}$,

$$\sum_{m\geq 0}\sum_{d\in\mathcal{S}} z_{gs}^* x_{gs}^m p_{sd}^g(m) = \sum_{m\geq 0}\sum_{d\in\mathcal{S}} z_{gd}^* x_{gd}^m p_{ds}^g(m),$$

where $p_{sd}^g(m)$ is the transition probability of group-g user from state s to d using channel m ($m = 0$ means passive) according to the Markov model. In other words, for each g and s, the amount of fluid going into each state (g, s) must equal to the amount of fluid going out of that state.

iii) For each state $s \in \mathcal{S}$ in group $g \in \mathcal{G}$, $[x_{gs}^m]$ must satisfy

$$\sum\nolimits_{(g,s)} z_{gs}^* x_{gs}^m \leq C_m, \quad \forall m \in \mathcal{M}, \qquad \text{(Primal feasible)}$$

$$\sum\nolimits_{(g,s)} z_{gs}^* x_{gs}^m = C_m, \quad \forall m \in \mathcal{M}, \quad \text{if } \lambda_m^* > 0. \qquad \text{(Complementary slackness)}$$

$$\sum\nolimits_{m\in\mathcal{M}} x_{gs}^m \leq 1, \quad \forall g \in \mathcal{G}, s \in \mathcal{S}. \qquad \text{(Primal feasible)}$$

iv) (Optimizing the Lagrangian) For each g and s, $[x_{gs}^m]$ denotes the optimal decisions for the sub-problem (8), given $\vec{\lambda}^*$, i.e., if $x_{gs}^m > 0$ for some m, it must be true that

$$\mu_m^g(s, \vec{\lambda}^*) \leq \mu_{u'}^g(s, \vec{\lambda}^*), \quad \forall u' \neq m, 0 \leq j' \leq M, \tag{17}$$

where $\mu_m^g(s, \vec{\lambda}^*)$ is the payoff of channel m in the Bellman equation (10) under $\vec{\lambda}^*$.

On the other hand, suppose that the SWIM policy (Algorithm 1) has also reached its steady state solution. Denote the steady state distribution of Algorithm 1 as $[\vec{z}'_1, \ldots, \vec{z}'_G]$. Recall that the Algorithm 1 is based on a set of dual costs $\vec{\lambda}$, and the edge weight is computed by $w_{gs}^u = I_u^g(s, \vec{\lambda}_{-u})$ and $w_{gs}^0 = I_0^g(s, \vec{\lambda})$. Let $[y_{gs}^u]$ be the optimal solution for MWM at the fixed point. At the fixed point, the optimal policy for MWM-PI solves (16). When Algorithm 1 reaches the fixed point $(\vec{y}, \vec{z}', \vec{\lambda}')$, the variables must satisfy the following conditions, where most (except (A)) comes from the KKT conditions of of (16) as follows:

A) At the fixed point, $\lambda'_u = \nu_u$ must hold for all $u \in \mathcal{M}$.

B) (Dual feasibility) The Lagrange multipliers satisfy $\nu_u \geq 0$ for all $u \in \mathcal{M}$.

C) The steady state user distribution $[z'_{gs}]$ have been reached, i.e., $\forall s \in \mathcal{S}, \forall g \in \mathcal{G}$,

$$\sum_{u \geq 0} \sum_{d \in \mathcal{S}} z'_{gs} y_{gs}^u p_{sd}^g(u) = \sum_{u \geq 0} \sum_{d \in \mathcal{S}} z'_{gd} y_{gd}^u p_{ds}^g(u).$$

D) For each state $s \in \mathcal{S}$ in group $g \in \mathcal{G}$, $[y_{gs}^u]$'s must satisfy

$$\sum\nolimits_{(g,s)} z'_{gs} \cdot y_{gs}^u \leq C_u, \quad \forall j \in \mathcal{M}, \qquad \text{(Primal feasibility)}$$

$$\sum\nolimits_{(g,s)} z'_{gs} \cdot y_{gs}^u = C_u, \quad \text{if } \nu_u > 0. \qquad \text{(Complementary slackness)}$$

$$\sum\nolimits_{u=0}^{M} y_{gs}^u = 1. \qquad \text{(Primal feasibility)}$$

E) (Optimizing the Lagrangian) For each (g, s), the solution $[y_{gs}^u]$ should be optimal for the following

$$\underset{y_{gs}^u}{\text{maximize}} \quad \sum\nolimits_{u=0}^{M} y_{gs}^u (w_{gs}^u - \nu_u) \tag{18a}$$

$$\text{subject to} \quad \sum\nolimits_{u=0}^{M} y_{gs}^u = 1. \tag{18b}$$

Denote $J^{\max} \triangleq \{u | w_{gs}^u - \nu_u \geq w_{gs}^{u'} - \nu_{u'}, \forall u' \neq u, u' \geq 0\}$. Then, $\sum_{u \in J^{\max}} y_{gs}^u = 1$, and $y_{gs}^{u'} = 0$ for $u' \notin J^{\max}$.

Before we proceed, we first state a corollary and a lemma as follows.

Corollary 3.7. *Suppose that Condition 3.5 holds. For any state $d \in \mathcal{S}$, suppose that there exists one channel-type m such that $I_m^g(d, \vec{\lambda}_{-m}) > \lambda_m$. Then, the other channels $u \neq m$ must have $I_u^g(d, \vec{\lambda}_{-u}) < \lambda_u$.*

Proof. By Definition 3.4-(ii), since $I^g_m(d, \vec{\lambda}_{-m}) > \lambda_m$, we must have $\mu^g_u(d, \vec{\lambda}) > \mu^g_m(d, \vec{\lambda})$ for all channels $u{\neq}m$. Suppose in contrary that $I^g_u(d, \vec{\lambda}_{-u}) \geq \lambda_u$ for $u{\neq}m$. Then, we would have $\mu^g_m(d, \vec{\lambda}) \geq \mu^g_u(d, \vec{\lambda})$ by Definition 3.4-(i),(ii), which is a contradiction. ■

Lemma 3.8. *In Condition (E), at least one action* $u \geq 0$ *should satisfy* $w^u_{gs} \geq \nu_u$ *at the fixed point.*

Proof. For the purpose of contradiction, we assume $w^u_{gs} < \nu_u, \forall u \geq 0$. Recall that $\vec{\lambda}' = \vec{\nu}$. Thus, we have $I^g_u(s, \vec{\nu_{-u}}) < \nu_u, \forall u \geq 1$. By the definition (3) of the partial index for $u \geq 1$, at cost $\vec{\nu}$, the following must be true for all $u \geq 1$: there must exist some $u' \neq u, u' \geq 0$ such that $\mu^g_u(s, \vec{\nu}) > \mu^g_{u'}(s, \vec{\nu})$. Since this is true for all $j \geq 1$, the only possibility is that $\mu^g_0(s, \vec{\nu}) < \mu^g_u(s, \vec{\nu}), \quad \forall u \geq 1$. This implies that $I^g_0(s, \vec{\nu}) = \nu^*_0 \triangleq 0$ by the definition of index for channel-0 in (13). However, this contradicts with the assumption of $I^g_0(s, \vec{\nu}) < 0$. Hence, at least one $u \geq 0$ should satisfy $w^u_{gs} \geq \nu_u$ at the fixed point. ■

Now, we proceed to show that the fixed point for the relaxed problem is equivalent to the fixed point for the MWM-PI. We first show the "$\Longrightarrow$" direction. Suppose that we are given the fixed point $[x^u_{gs}]$'s, $[\vec{z}^*_1, \ldots, \vec{z}^*_G]$ and $\vec{\lambda}^*$ of the relaxed problem, which satisfy conditions (i)-(iv). Our goal is to show that conditions (A)-(E) are satisfied by letting $\vec{\lambda}' = \vec{\lambda}^*$, $[\vec{z}'_g] = [\vec{z}^*_g]$, $y^m_{gs} = x^m_{gs}, \forall m \geq 1$ and $y^0_{gs} = 1 - \sum_{m \geq 1} y^m_{gs}$. Clearly, (A) holds by our construction. (B) $\nu_m = \lambda'_m = \lambda^*_m \geq 0$ is true since $\vec{\lambda}^* \geq 0$ at the fixed point. (C) follows from (ii) and our construction of $[\vec{z}'_g] = [\vec{z}^*_g]$ and $\vec{y} = \vec{x}$. (D) follows directly from (iii) and the definition of $y^0_{gs} = 1 - \sum_{m \geq 1} y^m_{gs}$. It remains to show (E). To show (E), we divide into three cases.

a) (When there exists $m \geq 1$ s.t. $I^g_m(s, \vec{\lambda}_{-m}) > \lambda'_m$) From Corollary 3.7, there exists only one channel m^* such that $I^g_{m^*}(s, \vec{\lambda}_{-m^*}) > \lambda'_{m^*}$. By Definition 3.4-(i), this implies that $\mu^g_{m^*}(s, \vec{\lambda}) < \mu^g_u(s, \vec{\lambda}), \forall u \neq m^*, u \geq 0$. From Condition (iv), the optimal decision for the relaxed problem is then $x^{m^*}_{gs} = 1$ and $x^u_{gs} = 0, \forall u \neq m^*$. By definition (13), we must have $I^g_0(s, \vec{\lambda}) \leq 0$, i.e., $w^0_{gs} \leq 0$. Clearly, the decision of $y^{m^*}_{gs} = x^{m^*}_{gs} = 1$ and $y^u_{gs} = y^u_{gs} = 0, \forall u \neq m^*$ satisfies Condition (E).

b) (When there exists no $m \geq 1$ s.t. $I^g_m(s, \vec{\lambda}_{-m}) > \lambda_m$, but there exists some $m \geq 1$ s.t. $I^g_m(s, \vec{\lambda}_{-m}) = \lambda_m$) Define $J_{\text{eq}} \triangleq \{m \geq 1 | I^g_m(s, \vec{\lambda}_{-m}) = \lambda_m\}$. Then, for all other channels $m' \geq 1, m' \notin J_{\text{eq}}$, we have $I_{m'}(gs, \vec{\lambda}_{-m'}) < \lambda_{m'}$. For any m' such that $I^g_{m'}(s, \vec{\lambda}_{-m'}) < \lambda'_{m'}$, by Definition 3.4-(iii), we must have $\mu^g_{m'}(s, \vec{\lambda}') > \mu^g_u(s, \vec{\lambda}')$ for some $u \neq m'$. Thus, the optimal decision under $\boldsymbol{\pi}_{\text{rel}}$ for such channels is $x^{m'}_{gs} = 0$. Therefore, the optimal decision for relaxed problem can only use $m \in J_{\text{eq}}$ and possibly channel-0 (passive), i.e., we must have

$$\sum\nolimits_{u \geq 1: I^g_u(s, \vec{\lambda}_{-u}) = \lambda_u} x^u_{gs} = 1 \text{ if } x^0_{gs} = 0,$$

$$\sum\nolimits_{u \geq 0: I^g_u(s, \vec{\lambda}_{-u}) = \lambda_u} x^u_{gs} = 1 \text{ if } x^0_{gs} > 0.$$

Thus, regardless of the value of x_{gs}^0, setting $y_{gs}^u = x_{gs}^u, \forall u \geq 0$ also attains the maximum objective for (18). Thus, Condition (E) is satisfied.

c) (When $I_m^g(s, \vec{\lambda}_{-m}) < \lambda'_m$ for all $m \geq 1$) From Definition 3.4-(iii), for all $m \geq 1$, there exist some $u \geq 0$ such that $\mu_m^g(s, \vec{\lambda}) > \mu_u^g(s, \vec{\lambda})$. Clearly, the only possibility is that $\mu_0^g(s, \vec{\lambda}) < \mu_m^g(s, \vec{\lambda}), \forall m \geq 1$. As a result, the optimal solution satisfying Condition (iv) is $x_{gs}^0 = 1$ and $x_{gs}^m = 0, \forall m \geq 1$. In this case, $w_{gs}^0 = \nu_0 = 0$ and $w_{gs}^m < \nu_m, \forall m \geq 1$ for the MWM. Thus, the solution $[y_{gs}^u] = [x_{gs}^u], u \geq 0$ (i.e., $y_{gs}^0 = 1$ and $y_{gs}^m = 0, \forall m \geq 1$) also attains the maximum of (18). Thus, Condition (E) holds.

Combining the above three cases, we show that Condition (E) holds, which completes the proof for "$\Longrightarrow$" direction.

Next, we show "$\Longleftarrow$" direction. Similar to the above proof for the "$\Longrightarrow$" direction, we assume that the optimal solutions at the steady state of the MWM problem are given, i.e., variables $\vec{\nu}$, $[\vec{z}'_1, \ldots, \vec{z}'_G]$ and $[y_{gs}^u]$'s satisfy (16) and conditions (A)-(E). Our goal is to show that conditions (i)-(iv) are satisfied by letting $\vec{\lambda}^* = \vec{\nu} = \vec{\lambda}$, $[\vec{z}_g^*] = [\vec{z}'_g]$ and $[x_{gs}^u] = [y_{gs}^u]$. Clearly, condition (i) must hold due to condition (B). Condition (ii) follows from condition (C) directly, by our construction of $\vec{z}^*$ and $\vec{x}$. Since $y_{gs}^u \in [0, 1], \forall u \geq 0$, Condition (iii) is satisfied by condition (D). To show that condition (iv), we divide into two cases (note that Lemma 3.8 implies $\max_{u \geq 0}\{w_{gs}^u - \nu_u\} \geq 0$).

a) (When $\max_{u \geq 0}\{w_{gs}^u - \nu_u\} > 0$) From condition (E), we have $\sum_{u \in J^{\max}} y_{gs}^u = 1$, where $J^{\max}$ contains all actions $u \geq 0$ that attain the maximum of $\{w_{gs}^u - \nu_u\}$. By the definition in (13), channel-0's index is always non-positive. Hence, we must have $0 \notin J^{\max}$. Thus, $y_{gs}^0 = 0$. From Corollary 3.7, there can exist only one $u \geq 1$ such that $w_{gs}^u = I_u^g(s, \vec{\nu}_{-u}) > \nu_u$, in which case $w_{gs}^{u'} = I_{u'}^g(s, \vec{\nu}_{-u'}) < \nu_{u'}, \forall u' \neq u$. Define u^* as the unique index such that $u^* = \arg\max_u\{w_{gs}^u - \nu_u\}$. Thus, $y_{gs}^{u^*} = 1$. By Definition 3.4-(ii), we then have $\mu_{u^*}^g(s, \vec{\nu}) < \mu_{u'}^g(s, \vec{\nu}), \quad \forall u' \neq u^*, 0 \leq u' \leq M$. According to (10), $x_{gs}^{u^*} = y_{gs}^{u^*} = 1$ and $x_{gs}^{u'} = 0, \forall u' \neq u^*, u' \geq 0$ correspond to the optimal solutions for (8). Hence, condition (iv) holds for the relaxed problem.

b) (When $\max_{u \geq 0}\{w_{gs}^u - \nu_u\} = 0$) Again, from condition (E), we have $\sum_{u \in J^{\max}} y_{gs}^u = 1$. Thus, $\sum_{u \in J^{\max}} x_{gs}^u = 1$ by our construction. The definition of $J^{\max}$ implies that for all $u \in J^{\max}, u > 0$, we have $w_{gs}^u = I_u^g(s, \vec{\nu}_{-u}) = \nu_u$. Further, for all $u' \notin J^{\max}$, we have $w_{gs}^{u'} = I_{u'}^g(s, \vec{\nu}_{-u'}) < \nu_{u'}$. Thus, from Definition 3.4, every $u \in J^{\max}$ is an optimal action for the sub-problem (8), and every $u' \notin J^{\max}$ is not. The only question is whether the optimal action for (iv) should use action-0 or not. Next, we divide into two sub-cases. If $0 \notin J^{\max}$, then $I_0^g(s, \vec{\nu}) < 0$. By the definition in (13), $x_{gs}^0 = 0$ must hold for the sub-problem (8). Thus, the decisions $\sum_{u \in J^{\max}} x_{gs}^u = 1$ is optimal for (8). On the other hand, if $0 \in J^{\max}$, then $I_0^g(k, \vec{\nu}) = 0$. By definition of index (13), for any $\epsilon_{0u} > 0$, we must have $\mu_0^g(s, \vec{\nu}) - \epsilon_{0u} \leq \mu_u^g(s, \vec{\nu}), \forall u \geq 1$. Letting $\epsilon_{0u} \to 0$, we then have $\mu_0^g(s, \vec{\nu}) \leq \mu_u^g(s, \vec{\nu}), \forall u \geq 1$.

Thus, $x_{gs}^0 > 0$ is also optimal for the sub-problem (8). Combining the two sub-cases, the decision $\sum_{u \in J^{\max}} x_{gs}^u = 1$ is always optimal for the sub-problem (8). Thus, condition (iv) follows.

Combing the above cases, condition (iv) must hold for the fixed point of MWM. Hence, we have shown the "$\Longleftarrow$" direction. To conclude, we have shown that the fixed point for the relaxed problem is equivalent to the fixed point for the Algorithm 1. $\square$

Remark. Theorem 3.6 establishes an important connection between the fixed point of the relaxed problem and the fixed point of Algorithm 1. In contrast to the relaxed problem, the solution for MWM-PI naturally respects the instantaneous resource constraint (6a). Since its fixed point still achieves the optimal performance at the fluid limit, it provides useful guidance for proving the asymptotic optimality of our proposed SWIM policy.

Next, we evaluate the performance of Algorithm 1. We first state the following technical condition called "global attractor" [22].

Definition 3.9 (Global attractor)**.** *An equilibrum point $\vec{X}^*$ is a global attractor for a process $X(t)$ if, for any initial point $\vec{X}(0)$, the process $X(t)$ converges to $\vec{X}^*$.*

Next, we assume that a fixed point of Algorithm 1 satisfies the global attractor property. Notice that similar assumption has been made in [20], [21], [22]. As mentioned in [22], in general, it may be difficult to establish analytically that a fixed point is a global attractor for the process; thus, such property is only verified numerically. Our simulation results in Section 5-A indeed show that such convergence indeed happens for our proposed policy.

Based on this condition, we then show the asymptotic optimality of the SWIM policy. Specifically, we will consider the original MDP (6) in the following r-scaled system: we scale by r both the number of sources in each group, and the number of channels of each type, i.e., $N^r = rN$ and $\overline{C}^r = r\overline{C}$, while keeping $\alpha = \overline{C}^r/N^r$ a constant. The transition probabilities for each source remain unchanged. For such a r-scaled system, we define V_{SWIM}^r as the average cost per stage in the objective of (6) under our proposed SWIM policy. (Note that V_{SWIM}^r in (6) is already averaged by the number of sources N^r.)

Theorem 3.10 (Asymptotic Optimality)**.** *Suppose that Condition 3.5 holds for the sub-problem* (10)*. Suppose that a fixed point $(\vec{y}, \vec{z}', \vec{X}')$ of the policy $\boldsymbol{\pi}_{\text{SWIM}}$ in Algorithm 1 is a global attractor according to Definition 3.9. Then, $\boldsymbol{\pi}_{\text{SWIM}}$ is asymptotically optimal in minimizing the average cost per stage. Specifically, we have $\lim_{r\to\infty} V_{\text{SWIM}}^r = V^*$, where V^* is the optimal objective for the fluid relaxed problem* (15)*.*

Proof. The proof of Theorem 3.10 follows from the global attractor property and Theorem 3.6. See Appendix A. □

4. AoI Minimization in Heterogeneous Multi-channel Systems

In this section, we return to the setting of AoI minimization problem described in Section 2-A. Since the results in Section 3 is very general, we only need to verify that Definition 3.2 and Definition 3.4 indeed hold for the AoI setting. Then, the result of Theorem 3.10 and Algorithm 1 can be directly applied. As we will show soon, the verification of the indexability and the precise division property is highly non-trivial for sub-problem (8). Note that for single-channel systems, Whittle indexability has been verified for AoI minimization under the generate-at-will model [6]. However, the approach there is based on directly solving the value function, which appears to be infeasible for our heterogeneous multi-channel setting. Instead, we will develop new structural properties of the value function, based on which we will establish both partial indexability and the precise division property.

With this goal in mind, we assume that the dual costs $\vec{\lambda} = [\lambda_1, \ldots, \lambda_M]^T$ for all channels are given. Since all sub-problems (8) are independent, in the rest of the section, we omit the superscript g of the variables for the sub-problem (10) whenever no ambiguity occurs. As we mentioned in Sec. 2-B, the MDP of each sub-problem is an average cost per stage problem with infinite time horizon and countably infinite state space. For ease of notation, we define $\lambda_0 = 0$ and $p_0 = 0$. The corresponding Bellman Equation (10) for the AoI minimization problem described in Section 2-A can be written as, for any state (i.e., current AoI) $d \in \mathbb{N}_+$,

$$f(d) + J^* \stackrel{\Delta}{=} \min_{m \in \{0\} \cup \mathcal{M}} \left\{ \lambda_m + (1 - p_m)[d + f(d+1)] + p_m f(1) \right\}. \tag{19}$$

Here we slightly abuse notation, and use $\mu_m(d) \stackrel{\Delta}{=} \mu_m(d, \vec{\lambda})$ to denote each term in the minimization on the RHS of (19) when the parameter $\vec{\lambda}$ is given above. Recall that $\mu_m(d, \vec{\lambda})$ is the expected cost-to-go under channel costs $\vec{\lambda}$ if channel m is selected. Since $d = 1$ is a recurrent state, we can set $f(1) = 0$. Note that Bellman equation (19) cannot be solved in closed form due to multiple heterogeneous channels. Specifically, the optimal action for the current state depends on the value functions of future states, which possibly have different optimal actions. This complex dependency is in sharp contrast to [6], [9], which only consider a single channel.

Even though the exact solution is unavailable, we can still derive useful structure properties from (19). Next, we define a property called "multi-threshold-type" (MTT), and prove that the optimal policy for the sub-problem (8) is indeed MTT.

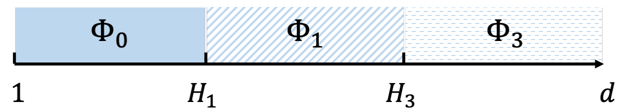


Fig. 2. An illustration for multi-threshold-type policy. Suppose a 3-channel system with $p_1 < p_2 < p_3$ and $\frac{\lambda_1}{p_1} < \frac{\lambda_3}{p_3} < \frac{\lambda_2}{p_2}$ (note that channel 2 is dominated by Channel 3).

Definition 4.1 (Multi-threshold-type)**.** *A channel selection policy for the sub-problem* (8) *is MTT if the followings hold:*

(1) (Threshold-based) For any channels $h, l \in \{0, 1, \ldots, M\}$ with $p_h > p_l$, there exists $H^{h,l} \geq 0$ such that $\mu_h(d) \leq \mu_l(d)$, for all $d \geq H^{h,l}$, and $\mu_h(d) > \mu_l(d)$, for all $d < H^{h,l}$.

(2) (Ordering of Channels) Suppose two states $d_1 < d_2$. Denote the optimal channels for d_1 and d_2 be $m^*(d_1)$ and $m^*(d_2)$, respectively. Then, $p_{m^*(d_1)} \leq p_{m^*(d_2)}$ must hold.

(3) (Channel Dominance) For any two channel types h and l with $p_h > p_l$, if $\frac{\lambda_h}{p_h} < \frac{\lambda_l}{p_l}$, then l is never the optimal channel for any state, i.e., $\mu_h(d) \leq \mu_l(d)$ for all state d whenever $\mu_l(d) \leq \mu_0(d)$.

In other words, Condition 1 states that a threshold exists for any pair of channels, such that a better-quality channel is always preferred to the worse one when d is above the threshold. Condition 2 specifies that, as d increases, the optimal decision increasingly prefer more reliable channels. Condition 3 means that if a channel type l is less reliable and also "more expensive" than the other channel type h, it should never be the optimal action. In that case, we say that channel l is dominated by channel h. The next lemma shows that the optimal policy for the sub-problem (8) is MTT.

Lemma 4.2. *Given the cost vector $\vec{\lambda}$, the optimal policy $\boldsymbol{\pi}^*$ satisfying* (19) *for the sub-problem* (8) *is MTT.*

Proof. Lemma 4.2 is intuitive because, when the state (i.e., age) is higher, it is more urgent for the source to use a more reliable channel. See the detailed proof in Appendix B. □

Fig. 2 illustrates a MTT policy in the state space $d \in \mathbb{N}_+$ of the sub-problem (8) with three channels. Define $\Phi_m \subset \mathbb{N}_+$ as the optimal decision region of m, i.e., m is the optimal channel type for all $d \in \Phi_m$. Thanks to Definition 4.1-(1), Φ_m must be contiguous for all m. We denote $H_m = \min_{d \in \Phi_m} d$ as the threshold for channel type m. Note that the optimal decision regions for some channels (Φ_2 is absent in Fig. 2) may be empty due to channel dominance. Before we proceed to the proof of partial indexability, we first prove the following lemma. Without loss of generality, we assume that all channels have distinct successful probabilities, and their qualities are arranged in an ascending order, i.e., $p_1 < \ldots < p_M$.

We note that the following presentation is different from our submitted conference version. Note that in our submitted version, Lemma 4.3, 4.4 and 4.6 are combined into one lemma. Instead, here we first state the upper bound in Lemma 4.3, and a special case of the lower bound (i.e., $d \geq H'_m$) in Lemma 4.4, based on which we can prove Prop. 4.5. Then, we state the full case of the lower bound in Lemma 4.6, which utilizes the result in Prop. 4.5. Lemma 4.3 and Lemma 4.6 here combined correspond to Lemma 4.3 in the conference version.

Lemma 4.3. *Given $\vec{\lambda}$, suppose $\vec{\lambda}' = [\lambda_1, \ldots, \lambda_m + \Delta, \ldots, \lambda_M], \Delta > 0$. Denote the optimal value functions in* (19) *under $\vec{\lambda}$ (with the optimal policy $\boldsymbol{\pi}$) and under $\vec{\lambda}'$ (with the optimal policy $\boldsymbol{\pi}'$) as $f(\cdot)$ and $f'(\cdot)$, respectively. Then, the difference between two value functions can be upper-bounded by*

$$f'(d) - f(d) < \frac{\Delta}{p_m}, \forall 1 \leq m \leq M.$$

Proof. To prove Lemma 4.3, we use the equivalence relationship between the average cost per stage problem and the stochastic shortest path problem [17]. See detailed proof in Appendix C. □

Next, we proceed to show the lower bound for $f'(d) - f(d)$ if the channel cost increases only on channel $m \neq M$. Notice that, to prove indexability in Proposition 4.5, we only need the lower bound for $d \geq H'_m$. Thus, we first show the case for $d \geq H'_m$ in Lemma 4.4. The proof for the case of $d < H'_m$ will utilize the partial indexability result in Proposition 4.5, which we will show in Lemma 4.6.

Lemma 4.4. *Denote $p_{M+1} \triangleq 1$. Under the same condition in Lemma 4.3, the difference of the two value functions can be lower-bounded by*

$$f'(d) - f(d) > -\frac{\Delta}{p_{m+1}}, \quad \text{for } d \geq H'_m, 1 \leq m \leq M,$$

where H'_m is the threshold for channel m under the optimal policy $\boldsymbol{\pi}'$ given $\vec{\lambda}'$.

Proof. See the proof in Appendix D. □

Proposition 4.5. *Given the cost vector $\vec{\lambda}$, the sub-problem* (8) *of heterogeneous multi-channel AoI minimization is partially indexable.*

Proof. See the proof in Appendix E. □

Now, we prove a stronger version of Lemma 4.4 that extends the lower bound in Lemma 4.4 to the entire state space $\mathcal{S} = \mathbb{N}_+$. The combined results of Lemma 4.3 and Lemma 4.6 correspond to Lemma 4.3 in the conference version.

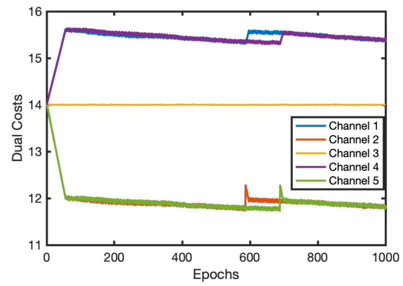


Fig. 3. Dual costs update of $\boldsymbol{\pi}_{\text{rel}}$ for the relaxed problem.

Lemma 4.6. *Denote* $p_{M+1} \stackrel{\Delta}{=} 1$. *Under the same condition in Lemma 4.3, the difference of the two value functions can be lower-bounded by*

$$f'(d) - f(d) > -\frac{\Delta}{p_{m+1}}. \quad \textit{for all } 1 \leq m \leq M, \tag{20}$$

Proof. See the proof in Appendix F. □

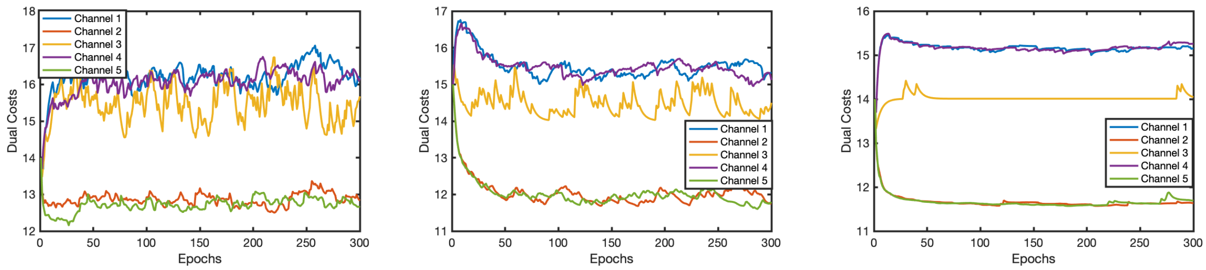
Proposition 4.7. *Given state space* $\mathcal{S}$ *and channel costs* $\vec{\lambda}$*, the sub-problem* (19) *satisfies the precise division property in Definition 3.4.*

Proof. The proof of Proposition 4.7 is similar to that of Proposition 4.5, and is available in Appendix G. □

To summarize this section, we briefly comment on the complexity of the SWIM policy. Note that the partial index incurs a higher computational complexity than Whittle's index [19], as it needs to be recomputed for every $\vec{\lambda}$. Nonetheless, given the number of states for each source (note that in practice one usually has to truncate the state space), the complexity is still linear in the number of sources and channels. Further, it is well-known that MWM incurs polynomial complexity [24]. In contrast, solving the original MDP using value iteration will incur exponential complexity. We leave for future work how to further reduce the complexity of the partial index computation.

5. Numerical Results

In this section, we present MATLAB simulation results to demonstrate the performance of our proposed SWIM policy in Algorithm 1. Specifically, we focus on the AoI minimization problem in Section 4 for heterogeneous multi-channel systems. We simulate an information update system with $N{=}50$ data sources, which are divided into $G{=}5$ groups, with $15, 5, 10, 15$ and 5 sources in each group 1 to 5, respectively. For the channels, we assume that there are $M{=}5$ types of channels, each of which is equipped with



(a) System scale $r = 1$ (b) System scale $r = 3$ (c) System scale $r = 10$

Fig. 4. Dual costs update of the proposed index-based policy $\boldsymbol{\pi}_{\text{SWIM}}$ for different system scales.

$\overline{C}$=2 identical instances. To model the heterogeneity and preferences between sources and channels, the channel success probabilities of source-channel pairs are different across groups. For simplicity, we assume that Group-g sources has the highest channel success probability on Type-g channels, where $g = 1, \ldots, 5$. In particular, the channel quality vector for group-1 users is $\vec{p_1} = [0.9, 0.7, 0.5, 0.3, 0.1]$. The channel quality vector $\vec{p_g}$ for group $g > 1$ is obtained by circularly shifting $\vec{p_1}$ by $g-1$ positions to the right. (Note that the entire system is not symmetric due to the uneven population of sources in each group.) To compare the scaling performance, we simulate on r-scaled systems that multiplies the number of the sources and channels by r, i.e., $N^r = rN$ and $\overline{C}^r = r\overline{C}$. The simulation time is divided into epochs, each of which consists of 50 discrete time-slots. To achieve a smoother update, we re-calculate the channel costs at the end of every epoch according to Line 8 of Algorithm 1 using the averaged $\overline{\nu}_m(t)$ over the current epoch.

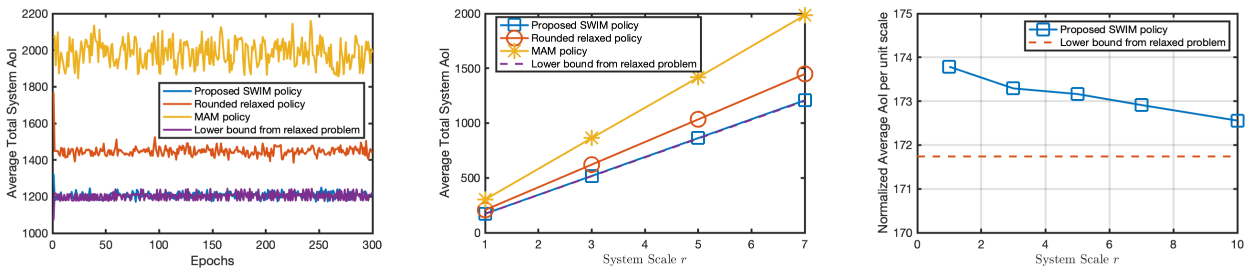
A. Convergence of the Channel-Cost Update

First, we compare the dynamics of the cost updates of the SWIM policy with that of the relaxed policy $\boldsymbol{\pi}_{\text{rel}}$, and verify that their fixed points indeed match. Fig. 3 illustrates the cost dynamics of all channels for the relaxed problem, with respect to the number of epochs. At each time, all sources independently compute their optimal actions (8) based on their states and current cost $\vec{\lambda}(t)$. At the end of each epoch, the BS performs a dual gradient update according to (9). We simulate for $T_{\text{rel}} = 1000$ epochs for $\boldsymbol{\pi}_{\text{rel}}$ to reach the fixed-point channel costs. In Fig. 3, all channels' costs converge to a small neighborhood of the optimal costs $\vec{\lambda}^*$ of the relaxed problem after $T_{\text{rel}} = 1000$ epochs.

Next, we verify that the channel-cost dynamics under $\boldsymbol{\pi}_{\text{SWIM}}$ approach that of the relaxed problem when the system scale is large. Specifically, we let $\beta = 0.2$ in Algorithm 1. Denote the fixed-point channel cost vector under SWIM policy for system scale r as $\vec{\lambda}_r^{\text{SWIM}}$. We simulate $\boldsymbol{\pi}_{\text{SWIM}}$ for $T_{\text{SWIM}} = 300$ epochs under different system scales. Fig. 4(a)-(c) show the channel-cost dynamics for the system at scale

$r = 1$, $r = 3$ and $r = 10$, respectively. Clearly, we can observe that, as the system scale r increases, $\vec{\lambda}_r^{\mathrm{SWIM}}$ approaches very close to the values of $\vec{\lambda}^*$ in Fig. 3. The convergence is more obvious for $r = 10$ (Fig. 4(c)), which confirms the result of Theorem 3.6, i.e., the fixed point solution $\vec{\lambda}'$ of $\boldsymbol{\pi}_{\mathrm{SWIM}}$ is equivalent to the optimal cost $\vec{\lambda}^*$ of the relaxed problem in the fluid limit.

B. Average System AoI



(a) Total system AoI evolution for scale $r = 7$.

(b) Average total system AoI v.s. system scales.

(c) Normalized AoI v.s. system scales.

Fig. 5. Performance comparison among different scheduling policies.

Next, we evaluate the average AoI performance of our proposed policy. We will use the solution for the relaxed problem as a performance lower bound for comparison. In addition, we will compare SWIM policy with the following scheduling policies that satisfy the instantaneous constraints in (6).

Rounded Relaxed Policy (RRP). As we discussed in Section 2-B, although the optimal solution for the relaxed problem under $\vec{\lambda}^*$ provides a lower bound for the original AoI minimization problem, $\boldsymbol{\pi}_{\mathrm{rel}}$ may violate the instantaneous hard constraints (6a). To satisfy feasibility, RRP is deduced from $\boldsymbol{\pi}_{\mathrm{rel}}$ with the following modification. (Note that we did not use the priority policy in [22], since it only works for single-channel systems and cannot be applied here.)

1. (Over-subscription) For any channel, if the number of transmitting sources exceeds (or equals to) the number of channel instance $\overline{C}^r$, RRP schedules $\overline{C}^r$ sources uniformly at random;
2. (Under-subscription) Otherwise, RRP schedules additional sources with largest AoI to reach $\overline{C}^r$ total sources for the channel.

Max-Age Matching (MAM) [15]. This policy was originally proposed for systems with multiple ON/OFF channels in [15]. As the name suggests, MAM attempts to serve sources with high AoI values at each time. The MAM scheduler in [15] requires knowledge of whether a channel is ON/OFF. For a fair comparison with our policy that does not require such knowledge, we take all channels with non-zero success probability as being ON. Then, we form a bipartite graph between all pairs of sources and the channel types, with the weights given by the current AoI of the sources. The scheduling problem is then mapped to a bipartite graph matching problem. Note that this policy ignores the exact channel success probability, and thus is expected to have poorer performance.

Fig. 5a shows the total system AoI dynamics under different policies for the system at scale $r = 7$. First, we can see that the total system AoI (about 1200) under our proposed SWIM policy is very close to the performance lower bound obtained from the relaxed problem (the lowest two curves). This observation verifies our result in Theorem 3.10 on the asymptotic optimality of our proposed policy. In contrast, the AoI of the rounded relaxed policy (RRP) is over 1400, which is about 20% worse than that of SWIM policy. This performance degradation suggests that it may not be efficient to use the solution from the relaxed problem even for medium-scaled systems. Finally, the AoI under the MAM policy is about 2000, which is about 65% worse than our SWIM policy. The result is not surprising, as the MAM policy simply ignores the channel heterogeneity.

Next, Fig. 5b shows the average total system AoI at different system scales. For all three policies, the average total system AoI scales almost linearly with the system scale r. Again, we observe that our proposed SWIM policy achieves close-to-optimal AoI performance under all simulated system scales. Finally, Fig. 5c shows the normalized average AoI, i.e., the total average system AoI divided by the scale parameter r, under our proposed SWIM policy. Clearly, as r increases, the normalized average AoI of our SWIM policy approaches closer to the lower bound obtained from the relaxed problem. This observation is again consistent with our result in Section 3-C on the asymptotic optimality of the SWIM policy.

6. Conclusion

In this work, we study the problem of minimizing AoI in heterogeneous multi-channel systems. We formulate the problem as an infinite-horizon constrained MDP. Existing results on Whilttle index cannot be applied to such a system with heterogeneous channels. Instead, we introduce a new notion of partial indexability, which is defined given the costs of all channels except one. Then, we propose a new scheduling policy, i.e., SWIM policy, based on partial index and MWM. Under suitable conditions, the SWIM policy asymptotically optimizes the total expected AoI of the system, when the system scale is large. To the best of our knowledge, we are the first in the literature to develop low-complexity and asymptotically optimal policies for weakly coupled MDP with multiple heterogeneous resources. The simulation results demonstrate near-optimal AoI performance that outperforms other multi-channel scheduling policies in the literature. For future work, we will study more-efficient computation of partial index, and other AoI problem settings, e.g., with stochastic arrivals.

REFERENCES

[1] M. A. Abd-Elmagid and H. S. Dhillon, "Average peak age-of-information minimization in UAV-assisted IoT networks," *IEEE TVT*, vol. 68, no. 2, pp. 2003–2008, 2019.

[2] Y. Gu, H. Chen, Y. Zhou, Y. Li, and B. Vucetic, "Timely status update in internet of things monitoring systems: An age-energy tradeoff," *IEEE IoT J.*, vol. 6, no. 3, pp. 5324–5335, 2019.

[3] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[4] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM 2012 - IEEE Conf. on Computer Comm.*, 2012, pp. 2731–2735.

[5] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, 2018.

[6] V. Tripathi and E. Modiano, "A Whittle index approach to minimizing functions of age of information," in *Allerton Conf. 2019*. IEEE, 2019, pp. 1160–1167.

[7] Y.-P. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Trans. Mobile Comput.*, 2019.

[8] Y.-P. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *Proc. IEEE Int. Symp. Inf. Theory*. IEEE, 2018, pp. 2634–2638.

[9] J. Sun, Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Closed-form Whittle's index-enabled random access for timely status update," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1538–1551, 2019.

[10] I. Kadota and E. Modiano, "Minimizing the age of information in wireless networks with stochastic arrivals," in *ACM Mobihoc '19*, 2019, p. 221–230.

[11] Z. Qian, F. Wu, J. Pan, K. Srinivasan, and N. B. Shroff, "Minimizing age of information in multi-channel time-sensitive information update systems," in *Proc. IEEE INFOCOM 2020 - IEEE Conf. on Comput. Commun.*, 2020.

[12] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[13] Y. Zou, K. T. Kim, X. Lin, M. Chiang, Z. Ding, R. Wichman, and J. Hämäläinen, "Low-overhead multi-antenna-enabled uplink random access for massive machine-type communications with low mobility," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, December 2019, pp. 1–6.

[14] ——, "Low-overhead joint beam-selection and random-access schemesfor massive internet-of-things with non-uniform channel and load," in *Proc. IEEE Conf. on Comput. Commun. (INFOCOM)*, July 2020, pp. 2126–2135.

[15] V. Tripathi and S. Moharir, "Age of information in multi-source systems," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.

[16] B. Sombabu and S. Moharir, "Age-of-information based scheduling for multi-channel systems," *IEEE Trans. on Wireless Comm.*, vol. 19, no. 7, pp. 4439–4448, 2020.

[17] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific Belmont, MA, 2007, vol. 2.

[18] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.

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

[20] R. R. Weber and G. Weiss, "On an index policy for restless bandits," *J. Appl. Probab.*, pp. 637–648, 1990.

[21] D. J. Hodge and K. D. Glazebrook, "On the asymptotic optimality of greedy index heuristics for multi-action restless bandits," *Adv. Appl. Probab.*, vol. 47, no. 3, pp. 652–667, 2015.

[22] I. M. Verloop, “Asymptotically optimal priority policies for indexable and nonindexable restless bandits,” *Ann. Appl. Probab.*, vol. 26, no. 4, pp. 1947–1995, 2016.

[23] A. Nedic and A. Ozdaglar, “Subgradient methods in network resource allocation: Rate analysis,” in *42nd Ann. Conf. Inform. Sci. and Sys.*, 2008, pp. 1189–1194.

[24] R. Duan and S. Pettie, “Linear-time approximation for maximum weight matching,” *J. ACM*, vol. 61, no. 1, Jan. 2014.

[25] D. R. Jiang and W. B. Powell, “An approximate dynamic programming algorithm for monotone value functions,” *Oper. Res.*, vol. 63, no. 6, pp. 1489–1511, 2015.

APPENDIX A
PROOF OF THEOREM 3.10

The proof of Theorem 3.10 consists of the following steps. Given the policy $\boldsymbol{\pi}_{\text{SWIM}}$ in Algorithm 1, since $(\vec{y}, \vec{z}', \vec{\lambda}')$ is a global attractor according to Definition 3.9, we have that the fixed point of fluid source-state process and channel costs converge to the fixed point of Algorithm 1, i.e., $\vec{z}(t) \to \vec{z}'$ and $\vec{\lambda}(t) \to \vec{\lambda}'$. From Theorem 3.6, since $\{\vec{y}, \vec{z}', \vec{\lambda}'\}$ is equivalent to the fixed point $\{\vec{x}, \vec{z}^*, \vec{\lambda}^*\}$ of the relaxed problem, we also have $\vec{z}(t) \to \vec{z}^*$, $\vec{\lambda}(t) \to \vec{\lambda}^*$. As a result, the scheduling decision $\vec{x}_{\text{SWIM}}(t) \to \vec{x}$ as well. Since $\{\vec{x}, \vec{z}^*, \vec{\lambda}^*\}$ are the optimal solutions for the fluid version of the relaxed problem (15) and incur cost value V^*, it implies that the fluid-scaled cost value under $\boldsymbol{\pi}_{\text{SWIM}}$ converges to V^*. By utilizing the result from [20, Prop.], the actual fractions of source-state vector $\vec{z}_{rN}(t) \to \vec{z}(t)$ and $V^r_{\text{SWIM}} \to V^*$ in the steady state as $r \to \infty$. As we mentioned in Section 3-C, V^* of the relaxed problem is a lower bound on the average cost per stage. Hence, we can conclude that the proposed policy is asymptotically optimal.

APPENDIX B
PROOF OF LEMMA 4.2

Suppose the policy $\boldsymbol{\pi}^*$ satisfies (19), i.e., $\boldsymbol{\pi}^*$ is the optimal policy under $\vec{\lambda}^*$. To show that $\boldsymbol{\pi}^*$ is MTT, we just need to verify (1)-(3) in Definition 4.1. Before that, we first show a lemma on the monotonicity of the optimal value function under $\boldsymbol{\pi}^*$.

Lemma B.1. *The optimal value function $f(d)$ in* (19) *under $\boldsymbol{\pi}^*$ is a non-decreasing function in $d, \forall d \geq 1$, i.e.,*

$$d \leq d' \quad \implies \quad f(d) \leq f(d') \tag{21}$$

Proof. The proof utilizes the result of [25, Prop. 3.1], which considers a finite T-stage total cost MDP problem. Recall that the sub-problem is described by state space $\mathcal{S}$, action space $\mathcal{U}$ and the transition probabilities (2) and (3). At each time t, let $\xi_t \in \Xi$ be a discrete time stochastic process that represents the perturbation due to random channel events. Similar to [25], we assume that there exists a state transition function $\Omega : \mathcal{S} \times \mathcal{U} \times \Xi \to \mathcal{S}$ that describes the system evolution, i.e., $s_{t+1} \triangleq \Omega(s_t, u_t, \xi_t)$. First, note that the sub-problem satisfies the following conditions. For $1 \leq d \leq d'$, $u \in \{0, 1, \ldots, M\}$ and $\xi \in \Xi$, we have

1) $\Omega(d, u, \xi) \leq \Omega(d', u, \xi)$. In other words, the order of the next states preserves because, at state d (and d'), the next states either increase to $d + 1$ (and $d' + 1$), or become exactly 1.
2) Let $g_u(d) = \lambda_u + (1 - p_u)d$ be the per stage cost in (19). Then, we must have $g_u(d) \leq g_u(d')$.
3) $\xi_{t+1} \in \Xi$ is independent of the state $d \in \mathcal{S}$.

Let $F_T(d)$ denote the value function for the T-stage total cost minimization problem. By [25, Prop. 3.1] and the above conditions, for any finite T, we have

$$d \leq d' \quad \Longrightarrow \quad F_T(d) \leq F_T(d').$$

Note that the value iteration (VI) algorithm for the infinite-stage average cost problem is guaranteed to converge [17]. Recall that increasing the number of iteration is equivalent to increasing the number of stages T. Thus, for any $\epsilon > 0$, there exists a T_{conv} such that, for all $T > T_{conv}$, we have $sp(F_T(\cdot) - F_{T-1}(\cdot)) < \epsilon$, where $sp(\cdot)$ is the span of a function. Thus, we have $|f(d) - (F_T(d) - F_T(1))| < \delta(\epsilon)$. Clearly, $F_T(d) - F_T(1)$ is non-decreasing in d, since $F_T(\cdot)$ is non-decreasing in d for any finite T. Hence, the lemma follows by taking $\epsilon \to 0$. ■

Now we proceed to verify the conditions in Definition 4.1. Since the following discussion are for a fixed $\vec{\lambda}$, we slightly abuse the notion by defining $\mu_m(d) \triangleq \mu_m(d, \vec{\lambda})$ (recall that we remove all superscript g in (19) and $\mu_m(d, \vec{\lambda}) = \lambda_m + (1 - p_m)[d + f(d + 1)]$). For Condition (1) of Definition 4.1 (i.e., “threshold-based”), for $d \geq 1$ and $h, l \in \{0, 1, \ldots, M\}$ such that $p_h > p_l$, from (19) we have

$$\begin{aligned}
&\mu_h(d) - \mu_l(d) \\
&= \lambda_h + (1-p_h)[d + f(d+1)] - [\lambda_l + (1-p_l)[d + f(d+1)] \\
&= \lambda_h - \lambda_l + (p_l - p_h)[d + f(d + 1)].
\end{aligned} \tag{22}$$

By Lemma B.1, we have that $d + f(d + 1)$ is strictly increasing in d. Noting $p_l - p_h < 0$, $\mu_h(d) - \mu_l(d)$ must be a strictly decreasing function of d. As a result, there must exist a threshold $H^{h,l}$ such that $\mu_h(d) > \mu_l(d)$ for $d < H^{h,l}$, and $\mu_h(d) \leq \mu_l(d)$ otherwise. Thus, Condition (1) of Definition 4.1 follows.

For Condition (2) of Definition 4.1 (i.e., “ordering of channels”), we prove by contradiction. Suppose $d_1 < d_2$, and $m^*(d_1) = h, m^*(d_2) = l$ as in the statement of this condition. Suppose in contrary that $p_h >$

h_l. For d_1, from (22) and $m^*(d_1) = h$, we have $\mu_h(d_1) - \mu_l(d_1) = \lambda_h - \lambda_l + (p_l - p_h)[d_1 + f(d_1+1)] \leq 0$. Now, for d_2, we have

$$\begin{aligned}\mu_h(d_2) - \mu_l(d_2) &= \lambda_h - \lambda_l + (p_l - p_h)[d_2 + f(d_2+1)] \\ &< \lambda_h - \lambda_l + (p_l - p_h)[d_1 + f(d_1+1)] \\ &\leq 0,\end{aligned}$$

where the strict inequality is from $p_l - p_h < 0$ and the fact that $d + f(d+1)$ is strictly increasing in d. Thus, $\mu_h(d_2) < \mu_l(d_2)$ contradicts with $m^*(d_2) = l$. Hence, Condition (2) of Definition 4.1 is true.

For Condition (3) of Definition 4.1 (i.e., "channel dominance"), suppose $\frac{\lambda_h}{p_h} < \frac{\lambda_l}{p_l}$ with $0 < p_l < p_h$. This implies that $\lambda_h p_l < \lambda_l p_h$ and

$$\frac{\lambda_h}{p_h} = \frac{\lambda_h(p_h - p_l)}{p_h(p_h - p_l)} > \frac{\lambda_h p_h - \lambda_l p_h}{p_h(p_h - p_l)} = \frac{\lambda_h - \lambda_l}{p_h - p_l}.$$

To show the condition, it suffices to show that, whenever $\mu_l(d) \leq \mu_0(d), \forall d \geq 1$, we have $\mu_h(d) < \mu_l(d)$. Notice that the expression for $\mu_m(d)$ also holds with $\lambda_0 \triangleq 0$ and $p_0 \triangleq 0$. From (22), we have $\mu_l(d) - \mu_0(d) = \lambda_l - p_l[d + f(d+1)] \leq 0$. Thus, we have

$$d + f(d+1) \geq \frac{\lambda_l}{p_l} > \frac{\lambda_h}{p_h} > \frac{\lambda_h - \lambda_l}{p_h - p_l}.$$

Combining the above equation with (22), we have $\mu_h(d) < \mu_l(d)$. Thus, channel l is never the optimal choice for any state $d \geq 1$. Hence, Condition (3) of Definition 4.1 also holds, and the result of Lemma 4.2 follows.

APPENDIX C
PROOF OF LEMMA 4.3

Recall that the Bellman equation (19) takes the equivalent form (10)

$$f(d) + J^*_{\vec{\lambda}} = \min_{u \in \{0,\ldots,M\}} \left[g_u(d, \vec{\lambda}) + \sum_{s=1}^{\infty} p_{ds}(u) f(s) \right], \tag{23}$$

where $g_u(d, \vec{\lambda}) = \lambda_u + d(1 - p_u)$ is the *stage cost* for state d on channel u in the sub-problem. Denote the optimal policy for the sub-problem under original costs $\vec{\lambda}$ and increased costs $\vec{\lambda}'$ as $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$, respectively. According to the equivalence between average cost per stage problem and stochastic shortest path (SSP) problem [17, p.176], $f(d)$ can be related to the cost of a SSP problem to a recurrent state ζ. In our case, we can choose this special state as $\zeta = 1$. Let $q^{\boldsymbol{\pi}}_{s\zeta}$ be the transition probability from state s to state ζ under policy $\boldsymbol{\pi}$, i.e., $q^{\boldsymbol{\pi}}_{s\zeta} = p_{\boldsymbol{\pi}(s)}$. Then, it is known that $f(d)$ can be written as [17]

$$f(d) = \min_{\hat{\boldsymbol{\pi}}} \Bigg[\mathbb{E}^{\hat{\boldsymbol{\pi}}}\{\text{cost to reach } \zeta \text{ from } d \text{ for the first time}\} \\ - \mathbb{E}^{\hat{\boldsymbol{\pi}}}\{\text{cost incurred from } d \text{ to } \zeta \text{ if the stage cost were } J^*_{\vec{\lambda}}\} \Bigg]. \quad (24)$$

Notice that such a trajectory from state d to state $\zeta = 1$ must first experience a number of time-slots with no successful transmission, in which case the state goes to $d+1, d+2, \dots$, until it reaches a state s' and experiences a successful transmission, in which case the state goes back to 1. For any state $s \geq d$, the probability that the trajectory will go through s (regardless of the future events after state s) is

$$\Pr\{\text{the trajectory from } d \text{ to } 1 \text{ goes through } s\} = \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}).$$

When this event occurs, the cost at state s, i.e., $g_{\boldsymbol{\pi}(s)}(s, \vec{\lambda})$, will be added to the first term of (24). Now, let $\mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda})$ and $N^{\boldsymbol{\pi}}_{d\zeta}$ be the expected cost and expected time, respectively, to reach ζ from d for the first time under $\boldsymbol{\pi}$. Then, given policy $\boldsymbol{\pi}$, we have

$$\mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}) = \sum_{s=d}^{\infty} g_{\boldsymbol{\pi}(s)}(s, \vec{\lambda}) \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}), \quad (25)$$

$$N^{\boldsymbol{\pi}}_{d\zeta} = \sum_{s=d}^{\infty} \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}). \quad (26)$$

Here in the above equations, when $s = d$, we use the convention that $\prod_{j=d}^{d-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) = 1$. Then, since $\boldsymbol{\pi}$ is the optimal policy that produces the minimum average cost of $J^*_{\vec{\lambda}}$, (24) is equivalent to

$$f(d) = \mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}) - N^{\boldsymbol{\pi}}_{d\zeta} \cdot J^*_{\vec{\lambda}}. \quad (27)$$

Similarly, we can write $f'(d)$ as

$$\begin{aligned} f'(d) &= \mathrm{Cost}^{\boldsymbol{\pi}'}_{d\zeta}(\vec{\lambda}') - N^{\boldsymbol{\pi}'}_{d\zeta} \cdot J^*_{\vec{\lambda}'} \\ &\leq \mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}') - N^{\boldsymbol{\pi}}_{d\zeta} \cdot J^*_{\vec{\lambda}'} && (28) \\ &\leq \mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}') - N^{\boldsymbol{\pi}}_{d\zeta} \cdot J^*_{\vec{\lambda}}. && (29) \end{aligned}$$

The first inequality (28) is from (24), as we replace the optimal policy $\boldsymbol{\pi}'$ under costs $\vec{\lambda}'$ by policy $\boldsymbol{\pi}$. The second inequality (29) comes from the fact that $J^*_{\vec{\lambda}'} \geq J^*_{\vec{\lambda}}$. Note that if channel m is the preferred choice under π for any state, then $J^*_{\vec{\lambda}'} > J^*_{\vec{\lambda}}$ and the inequality (29) becomes a strict inequality. Therefore, we have

$$\begin{aligned}
&f'(d) - f(d) \\
&\leq [\mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}') - N^{\boldsymbol{\pi}}_{d\zeta} \cdot J^*_{\vec{\lambda}}] - [\mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}) - N^{\boldsymbol{\pi}}_{d\zeta} \cdot J^*_{\vec{\lambda}}] \\
&= \mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}') - \mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}) \\
&= \sum_{s=d}^{\infty} g_{\boldsymbol{\pi}(s)}(s, \vec{\lambda}') \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) - \sum_{s=d}^{\infty} g_{\boldsymbol{\pi}(s)}(s, \vec{\lambda}) \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) \\
&= \sum_{s=d}^{\infty} [g_{\boldsymbol{\pi}(s)}(s, \vec{\lambda}') - g_{\boldsymbol{\pi}(s)}(s, \vec{\lambda})] \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) \\
&\qquad\qquad \text{(The term in } [\cdot] \text{ equals to } \Delta \text{ if policy } \boldsymbol{\pi} \text{ uses channel } m \text{ in state } s, \text{ and 0 otherwise.)} \\
&= \sum_{s:\boldsymbol{\pi}(s)=m} \Delta \cdot \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) \qquad (30)
\end{aligned}$$

Then, we divide into two cases. (1) If $m = M$, recall from the MTT property of $\boldsymbol{\pi}$ in Definition 4.1-(1) that channel M is always the preferred choice for some state no less than H_M. Thus, we must have $J^*_{\vec{\lambda}'} > J^*_{\vec{\lambda}}$, and strict inequality is taken for ((29)). Further, the summation in (30) has an infinite number of terms, i.e., we have

$$\begin{aligned}
f'(d) - f(d) &< \sum_{s:\boldsymbol{\pi}(s)=M} \Delta \cdot \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) \\
&\leq \Delta \cdot \sum_{s=H_M}^{\infty} (1 - p_M)^{s-d+1} \\
&\leq \Delta \cdot \sum_{s=H_M}^{\infty} (1 - p_M)^{s-H_M} = \frac{\Delta}{p_M}.
\end{aligned}$$

(Here H_M is the threshold for channel M. Note that $\prod_{j=d}^{H_M-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) \leq 1$.)

(2) If $m \neq M$, the sum in (30) has a finite number of terms, since π will eventually switch to channel M for some states no less than H_M. Thus, we have

$$\begin{aligned}
f'(d) - f(d) &\leq \sum\nolimits_{s:\boldsymbol{\pi}(s)=m} \Delta \cdot \prod_{j=d}^{s-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta}) \\
&\leq \Delta \cdot \sum\nolimits_{s=H_m}^{H_{m+1}-1} (1 - p_m)^{s-H_m} \\
&< \Delta \cdot \sum\nolimits_{s=H_m}^{\infty} (1 - p_m)^{s-H_m} = \frac{\Delta}{p_m}.
\end{aligned}$$

Combining (1) and (2), we have the result of the lemma.

Appendix D
Proof of Lemma 4.4

According to (27), we have

$$\begin{aligned}&f'(d)-f(d)\\&=\mathrm{Cost}_{d\zeta}^{\boldsymbol{\pi}'}(\vec{\lambda}')-N_{d\zeta}^{\boldsymbol{\pi}'}\cdot J_{\vec{\lambda}'}^{*}-[\mathrm{Cost}_{d\zeta}^{\boldsymbol{\pi}}(\vec{\lambda})-N_{d\zeta}^{\boldsymbol{\pi}}\cdot J_{\vec{\lambda}}^{*}].\end{aligned}$$

According to Lemma 4.2, the optimal policy for sub-problem given any cost vector is of multi-threshold type (MTT). Suppose $d \geq H'_m$, where H'_m is the threshold for channel m under policy $\boldsymbol{\pi}'$. By definition (24), we have

$$\begin{aligned}f(d)&=\mathrm{Cost}_{d\zeta}^{\boldsymbol{\pi}}(\vec{\lambda})-N_{d\zeta}^{\boldsymbol{\pi}}\cdot J_{\vec{\lambda}}^{*}\\&\leq \mathrm{Cost}_{d\zeta}^{\boldsymbol{\pi}'}(\vec{\lambda})-N_{d\zeta}^{\boldsymbol{\pi}'}\cdot J_{\vec{\lambda}}^{*}.\end{aligned} \tag{31}$$

Thus, we have

$$\begin{aligned}&f'(d)-f(d)\\&\geq \mathrm{Cost}_{d\zeta}^{\boldsymbol{\pi}'}(\vec{\lambda}')-N_{d\zeta}^{\boldsymbol{\pi}'}\cdot J_{\vec{\lambda}'}^{*}-[\mathrm{Cost}_{d\zeta}^{\boldsymbol{\pi}'}(\vec{\lambda})-N_{d\zeta}^{\boldsymbol{\pi}'}\cdot J_{\vec{\lambda}}^{*}]\\&=\sum_{s=d}^{\infty}\Big[g_{\boldsymbol{\pi}'(s)}(s,\vec{\lambda}')-J_{\vec{\lambda}'}^{*}-(g_{\boldsymbol{\pi}'(s)}(s,\vec{\lambda})-J_{\vec{\lambda}}^{*})\Big]\prod_{j=d}^{s-1}(1-q_{j\zeta}^{\boldsymbol{\pi}'})\end{aligned} \tag{32}$$

We now divide into two sub-cases. If $m = M$, since $d \geq H'_M$, we have $\boldsymbol{\pi}'(s) = M$ for all $s \geq d$. Therefore, the term in $[\cdot]$ of (32) equals to $\Delta - (J_{\vec{\lambda}'}^{*} - J_{\vec{\lambda}}^{*}) > 0$, and we immediately have

$$f'(d) - f(d) > 0, \quad \text{for } d \geq H'_M. \tag{33}$$

Thus, we only need to show the other sub-case when $m \neq M$. In this case, note that $\boldsymbol{\pi}'(s) = m$ for $H'_m \leq s < H'_{m+1}$. For those s such that $H'_m \leq s < H'_{m+1}$, the term in $[\cdot]$ of (32) also equals to $\Delta - (J_{\vec{\lambda}'}^{*} - J_{\vec{\lambda}}^{*}) > 0$. Further, when $s \geq H'_{m+1}$, we have $\boldsymbol{\pi}'(s) \geq m+1$ and thus $g_{\boldsymbol{\pi}'(s)}(s, \vec{\lambda}') = g_{\boldsymbol{\pi}'(s)}(s, \vec{\lambda})$, and the term in $[\cdot]$ of (32) equals to $J_{\vec{\lambda}}^{*} - J_{\vec{\lambda}'}^{*} < 0$. Now, we further divide into two sub-sub-cases based on the value of d.

- ($d \geq H'_{m+1}$) For all $s \geq d$, since $\boldsymbol{\pi}'(s) \geq m+1$, the summation term for each s in (32) is negative. Also, since $\boldsymbol{\pi}'$ is MTT, from Definition 4.1-(2), we have $q_{j\zeta}^{\boldsymbol{\pi}'} \geq p_{m+1}$ for $j \geq d$. Thus, we can obtain a lower bound for (32) as

$$(32) = \sum_{s=d}^{\infty}(-J_{\vec{\lambda}'}^{*} + J_{\vec{\lambda}}^{*})\prod_{j=d}^{s-1}(1 - q_{j\zeta}^{\boldsymbol{\pi}'}) = (-J_{\vec{\lambda}'}^{*} + J_{\vec{\lambda}}^{*})\sum_{s=d}^{\infty}\prod_{j=d}^{s-1}(1 - q_{j\zeta}^{\boldsymbol{\pi}'})$$

$$\geq (-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}) \sum_{s=d}^{\infty} \prod_{j=d}^{s-1} (1 - p_{m+1})$$
$$= (-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}) \sum_{s=d}^{\infty} (1 - p_{m+1})^{s-d}$$
$$= (-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}) \frac{1}{p_{m+1}}. \tag{34}$$

- ($H'_m \leq d < H'_{m+1}$) In this case, the summation in (32) contains some positive terms for $d \leq s \leq H'_{m+1} - 1$, and the negative terms for $s \geq H'_{m+1}$. Since $\boldsymbol{\pi}'$ is MTT, from Definition 4.1-(2), we have $q^{\boldsymbol{\pi}'}_{j\zeta} \geq p_{m+1}$ for $j \geq H'_{m+1}$. Thus, we can obtain a lower bound for (32) by omitting all the positive terms from $s = d$ to $s = H'_{m+1} - 1$, i.e.,

$$(32) \geq \sum_{s=H'_{m+1}}^{\infty} (-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}) \prod_{j=d}^{s-1} (1 - q^{\boldsymbol{\pi}'}_{j\zeta})$$
$$\geq \sum_{s=H'_{m+1}}^{\infty} (-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}) \prod_{j=H'_{m+1}}^{s-1} (1 - q^{\boldsymbol{\pi}'}_{j\zeta})$$
$$\geq \sum_{s=H'_{m+1}}^{\infty} (-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}) (1 - p_{m+1})^{s-H'_{m+1}} \tag{35}$$
$$= \frac{1}{p_{m+1}} (-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}). \tag{36}$$

Thus, in both sub-sub-cases, we have $f'(d) - f(d) > \frac{-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}}{p_{m+1}}$ Finally, when $m \neq M$, since there always exist some state $d' > H'_M$ whose stage cost does not increase, we must have $J^*_{\vec{\lambda}'} - J^*_{\vec{\lambda}} < \Delta$. Thus,

$$f'(d) - f(d) > \frac{-J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}}{p_{m+1}} > -\frac{\Delta}{p_{m+1}}. \tag{37}$$

Hence, combining (33) and (37), the result of the lemma holds for $d \geq H'_m$ and for all $1 \leq m \leq M$.

Appendix E
Proof of Proposition 4.5

To prove that the sub-problem (8) is partially indexable, it suffices to show the following two statements, i.e.,

(i) If $d \in \mathcal{P}_m(\vec{\lambda})$, then $d \in \mathcal{P}_m(\vec{\lambda}')$ must hold for $\vec{\lambda}' = [\lambda_1, \ldots, \lambda_m + \Delta, \ldots, \lambda_M]$, where $\Delta > 0$.

(ii) If $\lambda_m = \infty$, then $\mathcal{P}_m(\vec{\lambda}) = \mathcal{S}$.

Since the optimal policy for any sub-problem is MTT from Lemma 4.2, the passive sets under $\vec{\lambda}$ can be expressed as

$$\mathcal{P}_m(\vec{\lambda}) = \{1, \ldots, H_m{-}1\} \cup \{H_{m+1}, \ldots\} \quad \text{if } m < M, \tag{38}$$

and

$$\mathcal{P}_M(\vec{\lambda}) = \{1, \ldots, H_M - 1\}. \quad (39)$$

To show statement (i), it suffices to show $\mathcal{P}_m(\vec{\lambda}) \subseteq \mathcal{P}_m(\vec{\lambda}')$, which is equivalent to show

- $H_m \leq H'_m, \forall m \leq M$;
- $H_{m+1} \geq H'_{m+1}, \forall m < M$.

We first show $H_m \leq H'_m, \forall m \leq M$. For the purpose of contradiction, suppose $H_m > H'_m$. Then, at state H'_m, policy $\boldsymbol{\pi}$ prefers action $0 \leq l < m$ over channel m, i.e., $\mu_l(H'_m, \vec{\lambda}) \leq \mu_m(H'_m, \vec{\lambda})$. From (22), this implies that

$$(p_m - p_l)[H'_m + f(H'_m + 1)] \leq \lambda_m - \lambda_l. \quad (40)$$

Similarly, given $\vec{\lambda}'$, policy $\boldsymbol{\pi}'$ must prefer m over other channels at H'_m, i.e., $\mu_m(H'_m, \vec{\lambda}') \leq \mu_l(H'_m, \vec{\lambda}')$. This implies that

$$(p_m - p_l)[H'_m + f'(H'_m + 1)] \geq \lambda'_m - \lambda_l. \quad (41)$$

Recall that $\lambda'_m = \lambda_m + \Delta$. From (40) and (41), we have

$$f'(H'_m + 1) - f(H'_m + 1) \geq \frac{\Delta}{p_m - p_r} \geq \frac{\Delta}{p_m}.$$

However, this clearly contradicts with Lemma 4.3 that $f'(d) - f(d) < \frac{\Delta}{p_m}$. Thus, $H_m \leq H'_m, \forall 1 \leq m \leq M$ must hold.

Next, we show $H_{m+1} \geq H'_{m+1}$, if $m < M$. For the purpose of contradiction, suppose $H_{m+1} < H'_{m+1}$. Since $\boldsymbol{\pi}$ is MTT, at state $d' = H'_{m+1} - 1 \geq H_{m+1}$, policy $\boldsymbol{\pi}$ must prefer some channel $h > m$ over channel m, which implies that $d' \geq H_h \geq H_{m+1}$. From Lemma 4.2 and Definition 4.1-(1), we have $\mu_m(d', \vec{\lambda}) \geq \mu_{m+1}(d', \vec{\lambda})$, i.e.,

$$(p_{m+1} - p_m)[d' + f(d' + 1)] \geq \lambda_{m+1} - \lambda_m. \quad (42)$$

On the other hand, policy $\boldsymbol{\pi}'$ must prefer channel m at $d' = H'_{m+1} - 1$, i.e., $\mu_m(d', \vec{\lambda}') \leq \mu_{m+1}(d', \vec{\lambda}')$. This implies that

$$(p_{m+1} - p_m)[d' + f'(d' + 1)] \leq \lambda'_{m+1} - \lambda'_m. \quad (43)$$

Noting that $\lambda'_{m+1} = \lambda_{m+1}$ and $\lambda'_m = \lambda_m + \Delta$, subtracting (42) from (43), we have

$$(p_{m+1} - p_m)[f'(d' + 1) - f(d' + 1)] \leq \lambda_m - \lambda'_m = -\Delta.$$

Since $p_{m+1} - p_m > 0$, we have

$$f'(d' + 1) - f(d' + 1) \leq -\frac{\Delta}{p_{m+1} - p_m} < -\frac{\Delta}{p_{m+1}}.$$

However, since $d'+1 = H'_{m+1} \geq H'_m$, this contradicts with the result of Lemma 4.4 that $f'(d'+1) - f(d'+1) > -\frac{\Delta}{p_{m+1}}$.

So far, we have shown that $H_m \leq H'_m, \forall m \leq M$ and $H_{m+1} \geq H'_{m+1}, \forall m < M$. Thus, $\mathcal{P}_m(\vec{\lambda}) \subseteq \mathcal{P}_m(\vec{\lambda}')$, and statement (i) must hold. For Statement (ii), it is clear that $\mu_m(d,\vec{\lambda}) > \mu_0(d,\vec{\lambda})$ holds for any finite d when $\lambda_m = \infty$. Hence, we conclude that the sub-problem (8) is partially indexable.

APPENDIX F
PROOF OF LEMMA 4.6

According to (27), we have

$$\begin{aligned}
&f'(d) - f(d) \\
&= \mathrm{Cost}^{\boldsymbol{\pi}'}_{d\zeta}(\vec{\lambda}') - \mathrm{Cost}^{\boldsymbol{\pi}}_{d\zeta}(\vec{\lambda}) - [N^{\boldsymbol{\pi}'}_{d\zeta} \cdot J^*_{\vec{\lambda}'} - N^{\boldsymbol{\pi}}_{d\zeta} \cdot J^*_{\vec{\lambda}}].
\end{aligned}$$

According to Lemma 4.2, the optimal policy for sub-problem given any cost vector is of multi-threshold type (MTT). Next, we divide into three cases.

Case 1: $d \geq H'_m$

This case is shown in Lemma 4.4.

Case 2: $1 \leq d \leq H_m$

Recall that we have $H_m \leq H'_m$ from the proof of Prop. 4.5, where H_m and H'_m are the threshold for channel m under policy $\boldsymbol{\pi}$ for $\vec{\lambda}$ and that under $\boldsymbol{\pi}'$ for $\vec{\lambda}'$, respectively. Define the policy $\hat{\boldsymbol{\pi}}$ as the following: $\hat{\boldsymbol{\pi}}$ follows the decisions of policy $\boldsymbol{\pi}$ for $1 \leq s < d$, and follows policy $\boldsymbol{\pi}'$ otherwise. Then, by (24), we have

$$\begin{aligned}
&f'(1) \leq \mathrm{Cost}^{\hat{\boldsymbol{\pi}}}_{\zeta\zeta}(\vec{\lambda}') - N^{\hat{\boldsymbol{\pi}}}_{\zeta\zeta} \cdot J^*_{\vec{\lambda}'} \\
&= \sum_{s=1}^{d-1} \left[g_{u(\hat{\boldsymbol{\pi}})}(s,\vec{\lambda}') - J^*_{\vec{\lambda}'} \right] \prod_{j=1}^{s-1} (1 - q^{\hat{\boldsymbol{\pi}}}_{j\zeta}) + f'(d) \prod_{j=1}^{d-1} (1 - q^{\hat{\boldsymbol{\pi}}}_{j\zeta}) \\
&= \sum_{s=1}^{d-1} \left[g_{u(\boldsymbol{\pi})}(s,\vec{\lambda}') - J^*_{\vec{\lambda}'} \right] \prod_{j=1}^{s-1} (1 - q^{\boldsymbol{\pi}}_{j\zeta}) + f'(d) \prod_{j=1}^{d-1} (1 - q^{\boldsymbol{\pi}}_{j\zeta})
\end{aligned} \tag{44}$$

By definition, $f'(1) = 0$. Thus, we have

$$f'(d) \geq \frac{1}{\prod_{j=1}^{d-1}(1 - q^{\boldsymbol{\pi}}_{j\zeta})} \sum_{s=1}^{d-1} \left[J^*_{\vec{\lambda}'} - g_{u(\boldsymbol{\pi})}(s,\vec{\lambda}') \right] \prod_{j=1}^{s-1} (1 - q^{\boldsymbol{\pi}}_{j\zeta}) \tag{45}$$

Similarly, by expressing $f(1)$ in terms of $f(d)$, we have

$$f(d) = \frac{1}{\prod_{j=1}^{d-1}(1-q_{j\zeta}^{\boldsymbol{\pi}})} \sum_{s=1}^{d-1} \left[J_{\vec{\lambda}}^{*} - g_{u(\boldsymbol{\pi})}(s,\vec{\lambda}) \right] \prod_{j=1}^{s-1}(1-q_{j\zeta}^{\boldsymbol{\pi}}) \tag{46}$$

From (45) and (46), we have

$$\begin{aligned} & f'(d) - f(d) \\ \geq & \frac{\sum_{s=1}^{d-1} \left[J_{\vec{\lambda}'}^{*} - g_{u(\boldsymbol{\pi})}(s,\vec{\lambda}') - J_{\vec{\lambda}}^{*} + g_{u(\boldsymbol{\pi})}(s,\vec{\lambda}) \right] \prod_{j=1}^{s-1}(1-q_{j\zeta}^{\boldsymbol{\pi}})}{\prod_{j=1}^{d-1}(1-q_{j\zeta}^{\boldsymbol{\pi}})} \\ = & \frac{\sum_{s=1}^{d-1} \left[J_{\vec{\lambda}'}^{*} - J_{\vec{\lambda}}^{*} \right] \prod_{j=1}^{s-1}(1-q_{j\zeta}^{\boldsymbol{\pi}})}{\prod_{j=1}^{d-1}(1-q_{j\zeta}^{\boldsymbol{\pi}})} \geq 0. \end{aligned} \tag{47}$$

The first equality above is from $g_{u(\boldsymbol{\pi})}(s,\vec{\lambda}') = g_{u(\boldsymbol{\pi})}(s,\vec{\lambda})$, since $\boldsymbol{\pi}$ does not use channel m for state $s \leq d-1$, where $1 \leq d \leq H_m$. Hence, the lemma holds for $1 \leq d \leq H_m$.

Case 3: $H_m < d < H'_m$

For Case 3 to be non-empty, we must have $H'_m > H_m + 1$. In other words, policy $\boldsymbol{\pi}'$ uses some lower-rate channel $\boldsymbol{\pi}'(d) < m$, while policy $\boldsymbol{\pi}$ uses higher-rate channel $\boldsymbol{\pi}(d) \geq m$ for $H_m < d < H'_m$. In fact, from Prop. 4.5, $\boldsymbol{\pi}(d) = m$ must hold for $H_m < d < H'_m$ because $H'_m \leq H'_{m+1} \leq H_{m+1}$.

Now we focus on the range $H_m < d < H'_m$, where $\boldsymbol{\pi}$ uses channel m and $\boldsymbol{\pi}'$ uses channel $\boldsymbol{\pi}'(d)$ for state d. From Bellman equation (19), we have (recall that $f(1) = f'(1) = 0$ by definition)

$$f'(d) = g_{\boldsymbol{\pi}'(d)}(d,\vec{\lambda}') + (1-p_{\boldsymbol{\pi}'(d)})f'(d+1) - J_{\vec{\lambda}'}^{*},$$

and

$$f(d) = g_m(d,\vec{\lambda}) + (1-p_m)f(d+1) - J_{\vec{\lambda}}^{*},$$

where $g_{\boldsymbol{\pi}'(d)}(d,\vec{\lambda}') = \lambda'_{\boldsymbol{\pi}'(d)} + (1-p_{\boldsymbol{\pi}'(d)})d$ is the per stage cost under $\boldsymbol{\pi}'$ for state d. Denote $\Delta f(d)$ as

$$\begin{aligned} \Delta f(d) \triangleq & f'(d) - f(d) \\ = & g_{\boldsymbol{\pi}'(d)}(d,\vec{\lambda}') + (1-p_{\boldsymbol{\pi}'(d)})f'(d+1) - J_{\vec{\lambda}'}^{*} - \left[g_m(d,\vec{\lambda}) + (1-p_m)f(d+1) - J_{\vec{\lambda}}^{*} \right] \\ = & g_{\boldsymbol{\pi}'(d)}(d,\vec{\lambda}') + (1-p_{\boldsymbol{\pi}'(d)})f'(d+1) - \left[g_m(d,\vec{\lambda}) + \Delta + (1-p_m)f'(d+1) \right] \\ & + \Delta + (1-p_m)[f'(d+1) - f(d+1)] - J_{\vec{\lambda}'}^{*} + J_{\vec{\lambda}}^{*} \\ & \text{(By adding and subtracting } \Delta + (1-p_m)f'(d+1) \text{ on RHS.)} \\ = & \left[\mu_{\boldsymbol{\pi}'(d)}(d,\vec{\lambda}') - \mu_m(d,\vec{\lambda}') \right] + \Delta - J_{\vec{\lambda}'}^{*} + J_{\vec{\lambda}}^{*} + (1-p_m)\Delta f(d+1). \end{aligned} \tag{48}$$

Denote $\Delta\mu_{\boldsymbol{\pi}'(d),m}(d,\vec{\lambda}')\stackrel{\Delta}{=}\mu_{\boldsymbol{\pi}'(d)}(d,\vec{\lambda}') - \mu_m(d,\vec{\lambda}') \le 0$ for $H_m \le d < H'_m$. We next show that $\Delta\mu_{\boldsymbol{\pi}'(d),m}(d,\vec{\lambda}')$ is strictly increasing in d for $H_m \le d < H'_m$. To see this, note that $\Delta\mu_{\boldsymbol{\pi}'(d),m}(d,\vec{\lambda}') = \lambda_{\boldsymbol{\pi}'(d)} - \lambda_m + (p_m - p_{\boldsymbol{\pi}'(d)})(d + f'(d+1))$. Since $p_m > p_{\boldsymbol{\pi}'(d)}$ and $f'(d+1)$ is increasing in d by Lemma B.1, we have that $\Delta\mu_{\boldsymbol{\pi}'(d),m}(d,\vec{\lambda}')$ is strictly increasing in d for $H_m \le d < H'_m$. Denote the first few terms in (48) as

$$\theta(\boldsymbol{\pi}', d, \vec{\lambda}')\stackrel{\Delta}{=}\left[\mu_{\boldsymbol{\pi}'(d)}(d,\vec{\lambda}') - \mu_m(d,\vec{\lambda}')\right] + \Delta - J^*_{\vec{\lambda}'} + J^*_{\vec{\lambda}}. \tag{49}$$

Then, $\theta(\boldsymbol{\pi}', d, \vec{\lambda}')$ is also strictly increasing in d for $H_m \le d < H'_m$. Re-arranging (48) with (49), we have

$$\Delta f(d+1) = \frac{1}{1-p_m}\left[\Delta f(d) - \theta(\boldsymbol{\pi}', d, \vec{\lambda}')\right]. \tag{50}$$

We next show that the result of the lemma follows from (50) and the monotonicity of $\theta(\boldsymbol{\pi}', d, \vec{\lambda}')$ in d. At a high level, since $\theta(\boldsymbol{\pi}', d, \vec{\lambda}')$ is increasing in d, we can show that $\Delta f(d)$ may first increase in d from $d = H_m$, and then decrease in d until $d = H'_m$. However, since we have also shown that $\Delta f(H_m) \ge 0$ (i.e., Case 2) and $\Delta f(H'_m) \ge -\frac{\Delta}{p_{m+1}}$ (i.e., Case 1), we must have $\Delta f(d) \ge -\frac{\Delta}{p_{m+1}}$ for all $H_m \le d \le H'_m$. Specifically, we divide into three cases depending on the signs of $\theta(\boldsymbol{\pi}', d, \vec{\lambda}')$ and $\Delta f(d) - \theta(\boldsymbol{\pi}', d, \vec{\lambda}')$.

a. If $\theta(\boldsymbol{\pi}', d, \vec{\lambda}') \le 0$, since $\theta(\boldsymbol{\pi}', d, \vec{\lambda}')$ is increasing in d for $H_m \le d < H'_m$, we must have $\theta(\boldsymbol{\pi}', \underline{d}, \vec{\lambda}') \le 0$ for $H_m \le \underline{d} \le d$. Also, from the result of Case 2, we have $\Delta f(H_m) \ge 0$. Using induction in $\underline{d}$, we can then show that, for all $H_m \le \underline{d} \le d$, it must be true that $\Delta f(\underline{d}+1) = \frac{1}{1-p_m}\left[\Delta f(\underline{d}) - \theta(\boldsymbol{\pi}', \underline{d}, \vec{\lambda}')\right] \ge \Delta f(\underline{d}) - \theta(\boldsymbol{\pi}', \underline{d}, \vec{\lambda}') \ge \Delta f(\underline{d}) \ge 0$. Thus, $\Delta f(\underline{d})$ must be non-decreasing from H_m to $d+1$. We must then have $\Delta f(d) \ge 0$.
b. If $\theta(\boldsymbol{\pi}', d, \vec{\lambda}') > 0$ and $\Delta f(d) \ge \theta(\boldsymbol{\pi}', d, \vec{\lambda}')$, it holds trivially that $\Delta f(d) \ge 0$.
c. If $\theta(\boldsymbol{\pi}', d, \vec{\lambda}') > 0$ and $\Delta f(d) < \theta(\boldsymbol{\pi}', d, \vec{\lambda}')$, since $\theta(\boldsymbol{\pi}', d, \vec{\lambda}')$ is increasing in d, we must have $\theta(\boldsymbol{\pi}', \overline{d}, \vec{\lambda}') > 0$ for all $d \le \overline{d} < H'_m$. We now show by induction that $\Delta f(\overline{d}+1) < \Delta f(\overline{d})$ for all $d \le \overline{d} < H'_m$.
   - **Base case:** When $\overline{d} = d$, we have $\Delta f(d+1) = \frac{1}{1-p_m}\left[\Delta f(d) - \theta(\boldsymbol{\pi}', d, \vec{\lambda}')\right] < \Delta f(d) - \theta(\boldsymbol{\pi}', d, \vec{\lambda}') < \Delta f(d)$.
   - **Induction step:** Suppose that $\Delta f(d'+1) < \Delta f(d')$ for all $d \le d' \le \overline{d}$. We wish to show that $\Delta f(\overline{d}+2) < \Delta f(\overline{d}+1)$. Towards that end, from the hypothesis, we have $\Delta f(\overline{d}+1) < \Delta f(d)$. Since $\theta(\boldsymbol{\pi}', d, \vec{\lambda}')$ is increasing in d, we have $\Delta f(\overline{d}+1) - \theta(\boldsymbol{\pi}', \overline{d}+1, \vec{\lambda}') < \Delta f(d) - \theta(\boldsymbol{\pi}', d, \vec{\lambda}') < 0$. Thus, from (50), we must have $\Delta f(\overline{d}+2) = \frac{1}{1-p_m}[\Delta f(\overline{d}+1) - \theta(\boldsymbol{\pi}', \overline{d}+1, \vec{\lambda}')] < \Delta f(\overline{d}+1) - \theta(\boldsymbol{\pi}', \overline{d}+1, \vec{\lambda}') < \Delta f(\overline{d}+1)$.

Therefore, we have shown $\Delta f(\overline{d})$ is decreasing in $\overline{d}$ for all $d \leq \overline{d} < H'_m$ Since we have already shown in Case 1 that $\Delta f(H'_m) > -\frac{\Delta}{p_{m+1}}$ (recall that $p_{M+1} \triangleq 1$), it must then be true that $\Delta f(\overline{d}) > \Delta f(H'_m) > -\frac{\Delta}{p_{m+1}}$ for $d \leq \overline{d} < H'_m$.

From (a)-(c), we have shown that $\Delta f(d) > -\frac{\Delta}{p_{m+1}}$ for $H_m < d < H'_m$.

Combining above 3 cases, the result of the lemma follows.

APPENDIX G
PROOF OF PROPOSITION 4.7

Now, we proceed to show the three statements in Definition 3.4. Consider $1 \leq m \leq M$. First, to show *(i)*, it suffices to show that for any $\epsilon > 0$, $\mu_m(d, \vec{\lambda}) \leq \mu_k(d, \vec{\lambda}) + \epsilon, \forall k \neq m, k \geq 0$. From the definition in (12), for any $\delta > 0$, at the cost $\vec{\lambda}_\delta \triangleq [\lambda_1, \ldots, \lambda_m - \delta, \ldots, \lambda_M]$, we must have $\mu_m(d, \vec{\lambda}_\delta) \leq \mu_k(d, \vec{\lambda}_\delta), \forall k \neq m, k \geq 0$. Now consider the situation at cost $\vec{\lambda}$. We have

$$\begin{aligned}
&\mu_m(d, \vec{\lambda}) - \mu_k(d, \vec{\lambda}) \\
&= \lambda_m - \lambda_k + (p_k - p_m)[d + f(d+1, \vec{\lambda})] \\
&= \Big(\lambda_m - \delta - \lambda_k + (p_k - p_m)[d + f(d+1, \vec{\lambda}_\delta)]\Big) \\
&\quad + \delta + (p_k - p_m)[f(d+1, \vec{\lambda}) - f(d+1, \vec{\lambda}_\delta)] \\
&\leq \delta + (p_k - p_m)[f(d+1, \vec{\lambda}) - f(d+1, \vec{\lambda}_\delta)],
\end{aligned} \tag{51}$$

where the last inequality follows from $\mu_m(d, \vec{\lambda}_\delta) \leq \mu_k(d, \vec{\lambda}_\delta)$. Since $f(d+1, \vec{\lambda}) - f(d+1, \vec{\lambda}_\delta)$ is uniformly bounded by Lemma 4.3 and Lemma 4.6, we can choose $\delta(\epsilon) > 0$ such that $\delta(\epsilon) + (p_k - p_m)[f(d+1, \vec{\lambda}) - f(d+1, \vec{\lambda}_{\delta(\epsilon)})] \leq \epsilon$. Thus, for any $\epsilon > 0$, $\mu_m(d, \vec{\lambda}) \leq \mu_k(d, \vec{\lambda}) + \epsilon, \forall k \neq m, k \geq 0$ must hold. The result of *(i)* then follows.

To show *(ii)*, consider the cost $\vec{\lambda}_{I_m}(\delta) \triangleq [I_m(d, \vec{\lambda}_{-m}) - \delta, \vec{\lambda}_{-m}]$, where the m-th element of $\vec{\lambda}$ is replaced by $I_m(d, \vec{\lambda}_{-m}) - \delta$. By definition of partial index (12), for any $\delta > 0$, we have

$$\mu_m(d, \vec{\lambda}_{I_m}(\delta)) \leq \mu_k(d, \vec{\lambda}_{I_m}(\delta)), \forall k \neq m, k \geq 0. \tag{52}$$

Choose $\Delta = \delta = \frac{I_m(d, \vec{\lambda}_{-m}) - \lambda_m}{2} > 0$. Then, noting that $\lambda_{I_m(\delta), m} = \lambda_m + \Delta$, at cost $\vec{\lambda}$ we have

$$\begin{aligned}
&\mu_m(d, \vec{\lambda}) - \mu_k(d, \vec{\lambda}) \\
&= \lambda_m - \lambda_k + (p_k - p_m)[d + f(d+1, \vec{\lambda})] \\
&= \Big(\lambda_m + \Delta - \lambda_k + (p_k - p_m)[d + f(d+1, \vec{\lambda}_{I_m}(\delta))]\Big) \\
&\quad - \Delta + (p_k - p_m)[f(d+1, \vec{\lambda}) - f(d+1, \vec{\lambda}_{I_m}(\delta))]
\end{aligned}$$

$$\leq -\left(\Delta + (p_k - p_m)[f(d+1, \vec{\lambda}_{I_m}(\delta)) - f(d+1, \vec{\lambda})]\right), \tag{53}$$

where the last inequality follows from (52). If $p_k < p_m$, from Lemma 4.3, we have $\mu_m(d, \vec{\lambda}) - \mu_k(d, \vec{\lambda}) < -(\Delta + (p_k - p_m)\frac{\Delta}{p_m}) \leq 0$, which is exactly the conclusion in part (ii) of Definition 3.4. Note that if $m = M$, then $p_k < p_M$ for all k, and hence this is the only case that we need to consider. Thus, it only remains to show that $\mu_m(d, \vec{\lambda}) < \mu_k(d, \vec{\lambda})$ even if $m < M$ and $p_k > p_m$. Towards this end, suppose $m < M$ and $p_k > p_m$. Note that (52) must hold for $k = m+1$. From Lemma 4.6, we have $\mu_m(d, \vec{\lambda}) - \mu_{m+1}(d, \vec{\lambda}) < -(\Delta - (p_{m+1} - p_m)\frac{\Delta}{p_{m+1}}) < 0$. It then only remains to show that $\mu_m(d, \vec{\lambda}) < \mu_k(d, \vec{\lambda})$ for all $k > m+1$. Recall that $\boldsymbol{\pi}$ is MTT (Definition 4.1-(1)): since $\mu_m(d, \vec{\lambda}) < \mu_{m+1}(d, \vec{\lambda})$, then $d < H_{m+1} \leq H_k$ must hold. For the purpose of contradiction, suppose $\mu_k(d, \vec{\lambda}) \leq \mu_m(d, \vec{\lambda})$ for some $k > m+1$. Then, since $\mu_m(d, \vec{\lambda}) < \mu_{m+1}(d, \vec{\lambda})$, we must have $\mu_k(d, \vec{\lambda}) < \mu_{m+1}(d, \vec{\lambda})$. We now show that this also implies $\mu_k(d', \vec{\lambda}) < \mu_{m+1}(d', \vec{\lambda})$ for all $d' \geq d$. To see this, note that since $p_{m+1} < p_k$ and $f(d')$ is non-decreasing in d' from Lemma B.1, we must have that $\Delta\mu_{k,m+1}(d', \vec{\lambda}) = \mu_k(d', \vec{\lambda}) - \mu_{m+1}(d', \vec{\lambda}) = \lambda_k - \lambda_{m+1} + (p_{m+1} - p_k)[d' + f(d'+1)]$ is decreasing in d'. It must then be true that $\mu_k(d', \vec{\lambda}) < \mu_{m+1}(d', \vec{\lambda}), \forall d' > d$. In particular, when $d' = H_{m+1}$ (recall that $d < H_{m+1}$), we must have $\mu_k(H_{m+1}, \vec{\lambda}) < \mu_{m+1}(H_{m+1}, \vec{\lambda})$. However, by definition of H_{m+1}, we must have $\mu_{m+1}(H_{m+1}, \vec{\lambda}) \leq \mu_k(H_{m+1}, \vec{\lambda}), \forall k \neq m+1, k \geq 0$, which is a contradiction. Thus, $\mu_m(d, \vec{\lambda}) < \mu_k(d, \vec{\lambda})$ for $k > m+1$. Hence, (ii) holds.

Finally, if $I_m(d, \vec{\lambda}_{-m}) < \lambda_m$, by definition of the index, at cost $\vec{\lambda}$ there exist some channel k such that $\mu_m(d, \vec{\lambda}) > \mu_k(d, \vec{\lambda})$. (Otherwise, λ_m would have been the index for channel m by definition.) Thus, (iii) is true.

Hence, the result of the lemma follows.