



**HAL**  
open science

# On the Global Optimality of Whittle's Index Policy for Minimizing the Age of Information

Saad Kriouile, Mohamad Assaad, Ali Maatouk

► **To cite this version:**

Saad Kriouile, Mohamad Assaad, Ali Maatouk. On the Global Optimality of Whittle's Index Policy for Minimizing the Age of Information. IEEE Transactions on Information Theory, 2021, 10.1109/tit.2021.3121257 . hal-03427291

**HAL Id: hal-03427291**

**<https://hal.science/hal-03427291>**

Submitted on 13 Nov 2021

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

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

# On the Global Optimality of Whittle's Index Policy for Minimizing the Age of Information

Saad Kriouile, Mohamad Assaad, and Ali Maatouk

Laboratoire des Signaux et Systèmes, CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette, France

**Abstract**—This paper examines the average age minimization problem where only a fraction of the network users can transmit simultaneously over unreliable channels. Finding the optimal scheduling scheme, in this case, is known to be challenging. Accordingly, the Whittle's index policy was proposed in the literature as a low-complexity heuristic to the problem. Although simple to implement, characterizing this policy's performance is recognized to be a notoriously tricky task. In the sequel, we provide a new mathematical approach to establish its optimality in the many-users regime for specific network settings. Contrary to previous works in the literature that use restrictive mathematical assumptions, which can be challenging to verify, our novel approach is based on intricate techniques and it is free of any strong mathematical assumptions. These findings showcase that the Whittle's index policy has analytically provable asymptotic optimality for the AoI minimization problem. Finally, we lay out numerical results that corroborate our theoretical findings and demonstrate the policy's notable performance in the many-users regime.

## I. INTRODUCTION

Technological advances in wireless communications and the cheap cost of hardware have led to the emergence of real-time monitoring services. In these systems, an entity is interested in knowing the status of one or multiple processes observed by a remote source. Accordingly, the source sends packets to the monitor to provide information about the process/processes of interest. The main goal in these applications is to keep the monitor up to date. In fact, in such applications, information has the highest value when it is fresh since the outcome of the monitor's tasks is better when it is based on new rather than outdated data. To quantify this notion of freshness, the Age of Information (**AoI**) was introduced in [1]. Ever since, the AoI has become a hot research topic, and a considerable number of research works have been published on the subject [2]–[9].

Among the most fundamental issues that the research community aimed to address is age-based resource allocation. In most real-time applications, numerous sources share the same transmission channel where the available resources are scarce. The scarcity can be a consequence of battery considerations for the devices involved or physical interference that may limit the number of simultaneous transmissions. Consequently, a smart resource allocation scheme has to be adopted to minimize the AoI and attain the desired timeliness objective. In [10], the authors proposed both age-optimal and near age-optimal scheduling policies for the single and multi-server cases, respectively. In particular, they have shown that a greedy policy is age-optimal under certain assumptions in the single

exponential server case. In [11], the authors examined a single-source scenario where the source's update rate cannot exceed a predefined limit due to battery considerations. In this case, they were able to propose an age-optimal scheduling policy when the channel exhibit possible decoding errors. Age-optimal policies were also proposed in various network settings such as distributed scheduling and random access environments [12]–[14].

Among the scheduling problems investigated in the literature, we cite the following: consider  $N$  users communicating with a central entity over unreliable channels where, at most,  $M < N$  users can transmit simultaneously. What is the age-optimal strategy in this case? The wide range of applications that this problem encompasses let it emerge as a fundamental one that needs to be investigated. Unfortunately, this problem belongs to the family of Restless Multi-Armed Bandit (**RMAB**) problems, which are generally difficult to solve optimally. To address this difficulty, the authors in [15] have examined this problem and proved that a greedy algorithm is optimal when users have identical channel statistics. For the asymmetric case, the authors proposed a sub-optimal policy, known as the Whittle's index policy. The Whittle's index policy has been embraced by many works in various frameworks [16]–[25] as it is recognized for its low complexity and its notable performance. For example, in [17], the Whittle's index policy was adopted to minimize the average delay of queues. In another line of work, the authors in [22] employed a Whittle's index-based policy to maximize the average throughput over Markovian channels. Although it is simple to implement, the main challenge that arises when adopting this policy is characterizing its performance since its analysis is known to be notoriously difficult. To attend to this difficulty, the authors in [24] provided a sufficient condition, dubbed as Weber's condition, for the Whittle's index policy's asymptotic optimality in the many-users regime. However, this condition requires ruling out the existence of both closed orbits and chaotic behavior of a high-dimensional non-linear differential equation, which is extremely difficult to verify even numerically. To further facilitate the analysis of the policy, the works in [17], [22] have provided an approach based on a fluid limit model for the delay minimization and throughput maximization frameworks. By leveraging this model, they proved the asymptotic optimality of the Whittle's index policy in these frameworks under a recurrence assumption that is easier than Weber's condition but still requires simulations to be verified. Following the same footsteps, the present authors adopted the fluid limit model and provided proof of the asymptotic optimality of the

Whittle's index policy in the AoI framework under similar assumptions [21] that can be verified numerically by simulations. This raises the following important question: can we prove the Whittle's index policy's asymptotic age-optimality in specific network settings without recouring to **any** strong assumptions? Answering this question is extremely difficult and has yet to be answered even for the standard delay and throughput metrics. In this paper, we examine this question in the AoI framework, and we provide rigorous theoretical results that showcase the validity of the Whittle's index asymptotic optimality in *certain* network settings without imposing any strong mathematical assumptions. Note that the importance of the asymptotic many-users regime stems from the astronomical growth in the number of interconnected devices. For example, machine-type communications and the IoT in 5G networks require supporting tens of thousands of connected devices in a *single* cell. To that end, we summarize in the following the structure of the paper along with its key contributions:

- We start by formulating the problem of minimizing the average age of a network where  $M$  out of  $N$  users can communicate simultaneously with the central entity. As previously explained, this problem belongs to the class of RMAB problems, which are known to be notoriously difficult to solve. Accordingly, the Whittle's index policy has been proposed in previous works as a low-complexity solution, which is the main focus of our work. To establish the Whittle's index policy, the following steps have to be taken:
  - 1) Provide a relaxed version of the original problem and tackle it through a Lagrangian approach.
  - 2) Prove the indexability property of the relaxed problem and derive the Whittle's index expressions.

These steps have been carried out in previous works by the authors in [15], and their main results are reported in our paper for completeness.

- Next, we present a fluid limit model that approximates the Whittle's index policy behavior. In the many-users regime, we prove that the fluid limit can be made arbitrarily close to the actual network's evolution. Therefore, we mainly focus on the evolution of the fluid limit vector in our optimality analysis. The method previously carried out in the literature to establish the Whittle's index policy's asymptotic optimality follows a spectral analysis approach [22]. However, this approach is highly contingent on the initial state of the system. Accordingly, to extend their results to any random initial state, the authors imposed a restrictive assumption, which can only be verified numerically. In our paper, we take a different approach to analyze the fluid model. Specifically, we propose a novel method based on intricate techniques (e.g., Cauchy criterion) to prove the fluid model's convergence to a fixed point. We stress that this step's technical details are intricate and constitute our paper's main technical contribution. Note that, even for the standard delay and throughput metrics, such proof was not provided in the literature, which further highlights our approach's novelty. Afterwards, we establish the global optimality of

Whittle's index policy leveraging the fact that the aforementioned fixed point is nothing but the optimal system's operating point in the many-users regime. Finally, we provide numerical results that corroborate the theoretical results and highlight the Whittle's index policy's notable performance in the many-users regime.

The rest of the paper is organized as follows: Section II is devoted to the system model and the problem formulation. Section III is dedicated to the establishment of the Whittle's index policy. In Section IV-B, we provide our main results where we prove the asymptotic optimality of the Whittle's index policy. Numerical results that corroborate our theoretical findings are given in Section V while Section VI concludes the paper.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. System Model

We consider a time-slotted system with one base station,  $M$  uncorrelated channels, and  $N$  users ( $N > M$ ). Time is considered to be normalized to the slot duration (i.e.,  $t = 1, 2, \dots$ ). We suppose that any of the  $M$  channels can be allocated to at most one user. Hence, at most  $M$  users will be able to transmit in each time slot  $t$ . If a user is scheduled at time  $t$ , it generates a fresh new packet and sends it to the base station. This packet is successfully decoded by the base station at time  $t+1$  with a certain success probability. We consider that if a decoding error takes place, the packet is discarded (i.e., users are not equipped with buffers). In practice, users may share similar channel conditions. Accordingly, we suppose that the users can be partitioned into  $K = 2$  different classes such that users within the same class share the same decoding success probability. In other words, each user  $i$  belonging to class  $k \in \{1, 2\}$  has a decoding success probability  $p_k$ , which is assumed to be known by the scheduler. We let  $\gamma_k$  be the proportion of users belonging to class  $k$ . To that end, the following always holds:  $\gamma_1 + \gamma_2 = 1$ .

A scheduling policy  $\pi$  is defined as a sequence of actions  $\pi = (\mathbf{a}^\pi(0), \mathbf{a}^\pi(1), \dots)$  where  $\mathbf{a}^\pi(t) = (a_1^{1,\pi}(t), a_2^{1,\pi}(t), \dots, a_{\gamma_1 N}^{1,\pi}(t), a_1^{2,\pi}(t), a_2^{2,\pi}(t), \dots, a_{\gamma_2 N}^{2,\pi}(t))$  is a binary vector such that  $a_i^{k,\pi}(t) = 1$  if user  $i$  of class  $k$  is scheduled at time  $t$ . We also let the binary random variable  $c_i^k(t)$  denote the channel state of user  $i$  of class  $k$  such that  $c_i^k(t) = 1$  if no decoding error takes place. As per our system model, we always have  $\Pr(c_i^k(t) = 1) = p_k$  and  $\Pr(c_i^k(t) = 0) = 1 - p_k$  for any user  $i$  of class  $k$ . We let  $B_i^{k,\pi}(t)$  denote the time-stamp of the freshest packet delivered by user  $i$  of class  $k$  to the base station at time  $t$  under the scheduling policy  $\pi$ . The age of information, or simply the age, of user  $i$  of class  $k$  is defined as [1]:

$$s_i^k(t) = t - B_i^k(t) \quad (1)$$

By taking into account the variables defined, the age of this user under policy  $\pi$  evolves as follows:

$$s_i^{k,\pi}(t+1) = \begin{cases} 1 & \text{if } a_i^{k,\pi}(t) = 1, c_i^k(t) = 1 \\ s_i^{k,\pi}(t) + 1 & \text{if } a_i^{k,\pi}(t) = 1, c_i^k(t) = 0 \\ s_i^{k,\pi}(t) + 1 & \text{if } a_i^{k,\pi}(t) = 0, \end{cases} \quad (2)$$

We let  $\mathbf{s}^\pi(t)$  denote the vector of all users' age  $\mathbf{s}^\pi(t) = (s_1^{1,\pi}(t), \dots, s_{\gamma_1 N}^{1,\pi}(t), s_1^{2,\pi}(t), \dots, s_{\gamma_2 N}^{2,\pi}(t))$  under policy  $\pi$ . With all these notations in mind, we can formulate the optimization problem that we focus on in our paper.

### B. Problem Formulation

In this paper, we are interested in minimizing the total expected average age of information of the network under the constraint on the number of users scheduled at each time slot  $t$ . The latter must be less than the total number of channels  $\alpha N$  where  $\alpha$  is equal to  $\frac{M}{N}$ . We let  $\Pi$  denote the set of all causal scheduling policies in which the scheduling decisions are made based on the history and current states of the system. To that end, and given an initial system state  $\mathbf{s}(0) = (s_1^1(0), \dots, s_{\gamma_1 N}^1(0), s_1^2(0), \dots, s_{\gamma_2 N}^2(0))$ , our problem can be formulated as follows:

$$\begin{aligned} \min_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\pi \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{\gamma_k N} s_i^{k,\pi}(t) \mid \mathbf{s}(0) \right] \\ \text{s.t. } \sum_{k=1}^2 \sum_{i=1}^{\gamma_k N} a_i^{k,\pi}(t) \leq \alpha N, \quad t = 0, 1, 2, \dots \end{aligned} \quad (3)$$

This problem belongs to the family of RMAB problems, which are generally difficult to solve optimally (see Papadimitriou et al. [26]). For this reason, one should aim to develop a well-performing sub-optimal policy. As it has been mentioned, the low-complexity scheduling policy that we are interested in throughout this paper is the Whittle's index policy. To establish this policy and derive the Whittle's indices expressions, one has to follow the steps below:

- 1) Provide a relaxed version of the original problem and tackle it through a Lagrangian approach.
- 2) Prove the indexability property of the problem and derive the Whittle's index expressions.

As previously mentioned, these steps have been carried out in previous works by the authors in [15]. For completeness, and as we will use these steps later in our optimality analysis, we report them along with the main results of [15] in the following section.

## III. RELAXED PROBLEM AND WHITTLE'S INDEX POLICY

### A. Relaxed Problem

The first step toward establishing the Whittle's index policy consists of relaxing the constraint on the number of scheduled users of the problem in (3). Specifically, instead of having the constraint satisfied at each time slot, we consider that it has to be satisfied on average. Therefore, the relaxed problem can be formulated as follows:

$$\begin{aligned} \min_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\pi \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{\gamma_k N} s_i^{k,\pi}(t) \mid \mathbf{s}(0) \right] \\ \text{s.t. } \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\pi \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{\gamma_k N} a_i^{k,\pi}(t) \right] \leq \alpha N \end{aligned} \quad (4)$$

To study this problem, one has to introduce a Lagrangian approach to transform the problem into an unconstrained one as will be detailed in the sequel.

### B. Dual Problem

To circumvent the difficulty of studying the constrained problem in (4), a Lagrangian approach has to be adopted. In particular, let us denote by  $\lambda \geq 0$  the Lagrangian parameter. For a fixed  $\lambda$ , the Lagrangian function of the relaxed problem is:

$$F(\lambda, \pi) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\pi \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{\gamma_k N} s_i^{k,\pi}(t) + \lambda (a_i^{k,\pi}(t) - \alpha) \mid \mathbf{s}(0) \right] \quad (5)$$

Based on the dual approach, the next step consists of finding the policy  $\pi$  that minimizes  $F(\lambda, \pi)$ . Note that the term  $\frac{1}{T} \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{\gamma_k N} \lambda \alpha$ , which is equal to  $N \lambda \alpha$ , doesn't depend on  $\pi$ . Therefore, the policy that minimizes the above function  $F(\lambda, \pi)$  also minimizes the following function:

$$f(\lambda, \pi) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\pi \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{\gamma_k N} s_i^{k,\pi}(t) + \lambda a_i^{k,\pi}(t) \mid \mathbf{s}(0) \right] \quad (6)$$

Then, we can formulate the dual problem as follows:

$$\min_{\pi \in \Pi} f(\lambda, \pi) \quad (7)$$

### C. Structural Results

To solve the problem in (7), it can be shown that this  $N$ -dimensional problem can be decomposed into  $N$  one-dimensional problems that can be solved independently [15]. Therefore, we can drop the  $i$  and  $k$  indices from (6) and simply investigate the following one-dimensional problem:

$$\min_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\pi \left[ \sum_{t=0}^{T-1} s^\pi(t) + \lambda a^\pi(t) \mid s(0) \right] \quad (8)$$

It turns out that the above one dimensional problem can be cast into an infinite horizon average cost Markov Decision Process (MDP) that is defined as follows:

- **States:** The state of the MDP at time  $t$  is the age of the user  $s(t)$  that can take any integer value strictly higher than 0. Therefore, the considered state space is countable and infinite.
- **Actions:** The action at time  $t$ , denoted by  $a(t)$ , indicates if a transmission is attempted (value 1) or the user remains idle (value 0).
- **Transitions probabilities:** The transitions probabilities between the different states have been previously detailed in Section II.
- **Cost:** The cost function at time  $t$  is designated by  $C(s(t), a(t)) = s(t) + \lambda a(t)$ .

To solve this MDP, the authors in [15] have leveraged the Bellman equation and studied the characteristics of the value function involved. Based on the particularity of the value function, the following result was found:

**Proposition 1.** *The optimal policy that solves problem (8) is of a threshold nature.*

*Proof.* See [15, Proposition 14].  $\square$

The above results tell us that there exists an integer  $l_k \in \mathbb{N}^*$  such that by only letting users of class  $k$  with an age larger or equal to  $l_k$  to transmit, we attain the optimal operating point of (8). These results are pivotal to proceed with establishing the Whittle's index policy.

#### D. Indexability and Whittle's Index Expressions

To proceed toward our goal, one has to analyze the behavior of the MDP when a threshold policy is adopted. To that end, we note that for any fixed threshold  $n$ , the MDP can be modeled through a Discrete Time Markov Chain (DTMC) where:

- The state is the age  $s(t)$ .
- For any state  $s(t) < n$ , the user is idle. On the other hand, when  $s(t) \geq n$ , the user is scheduled.

The DTMC is reported in Fig. 1. To be able to prove the

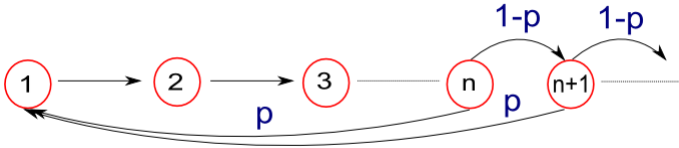


Figure 1: The states transition when a threshold policy is adopted

indexability property and find the Whittle's index expression, one has to find the average objective function in (8) when a threshold policy is adopted. To that end, we provide the following propositions.

**Proposition 2.** For a fixed threshold  $n$ , the stationary distribution  $u^n$  of the DMTC when the decoding success probability is equal to  $p$  is:

$$u^n(i) = \begin{cases} \frac{p}{np+1-p} & \text{if } 1 \leq i \leq n \\ (1-p)^{i-n} \frac{p}{np+1-p} & \text{if } i \geq n \end{cases} \quad (9)$$

*Proof.* The results can be easily obtained by solving the full balance equations.  $\square$

The next step consists of calculating the average objective function in (8) when a threshold policy is employed.

**Proposition 3.** For a fixed threshold  $n$ , the average cost of the threshold policy of the problem (8) is:

$$\bar{C}(n, \lambda) = \frac{[(n-1)^2 + (n-1)]p^2 + 2p(n-1) + 2}{2p((n-1)p+1)} + \frac{\lambda}{np+1-p} \quad (10)$$

*Proof.* The results can be concluded by leveraging the stationary distribution expressions and the fact that  $\bar{C}(n, \lambda) = \sum_{i=1}^{+\infty} i u^n(i) + \lambda \sum_{i=n}^{+\infty} u^n(i)$ .  $\square$

Using the stationary distribution, and the average cost, one can then prove the indexability property of the problem, which ensures the existence of the Whittle's indices. Before providing these results, we first lay out the definition of the aforementioned property.

**Definition 1 (Indexability).** For a fixed  $\lambda$ , consider the vector  $l(W) = (l_1(\lambda), l_2(\lambda))$  where  $l_k(\lambda)$  is the optimal threshold for the problem in (8) for each user of class  $k$ . We define  $D^k(\lambda) = \{s \in \mathbb{N}^* : s < l_k(\lambda)\}$  as the set of states for which the optimal action is to not schedule the users belonging to class  $k$ . The one-dimensional problem associated with these users is said to be indexable if  $D^k(\lambda)$  is increasing in  $\lambda$ . More specifically, the following should hold:

$$\lambda' \leq \lambda \Rightarrow D^k(\lambda') \subseteq D^k(\lambda) \quad (11)$$

The indexability property for the problem in (8) was established by the authors in [15]. With the Whittle's indices ensured to exist, one can then leverage the stationary distribution and the average cost reported in Proposition 2 and 3 to derive the Whittle's index expressions as previously done in [15] and [21].

**Proposition 4.** For any given class  $k$ , the Whittle's index expression of state  $i$  is:

$$W^k(i) = \frac{(i-1)p_k i}{2} + i \quad (12)$$

*Proof.* See [15, pp. 10].  $\square$

With the Whittle's index expression derived, we can now establish the Whittle's index scheduling policy. This can be summarized in the following algorithm description.

#### Algorithm 1 Whittle's index scheduling policy

- 1: At each time slot  $t$ , calculate the Whittle's index of all users in the network using (4).
- 2: Schedule the  $M$  users having the highest Whittle's index values at time  $t$ , with ties broken arbitrarily.

Although the above scheduling policy is easy to implement, it remains sub-optimal. Accordingly, characterizing its performance compared to the optimal policy is important. Equipped with the above results and notations, we can now tackle the main issue that we aim to address in our paper: the asymptotic optimality of this policy.

## IV. ASYMPTOTIC OPTIMALITY OF THE WHITTLE'S INDEX POLICY

### A. Optimal Solution of the Relaxed Problem

To be able to prove the asymptotic optimality of the Whittle's index policy, one has to compare its performance to the optimal policy that solves (3). However, as previously explained, the optimal policy of (3) is not known. To circumvent this, and to have a benchmark performance to compare to, we note that the following always holds:

$$\frac{C^{RP,N}}{N} \leq \frac{C^{OP,N}}{N} \leq \frac{C^{WIP,N}}{N} \quad (13)$$

where  $\frac{C^{WIP,N}}{N}$  is the average age per-user under the Whittle's index policy,  $\frac{C^{OP,N}}{N}$  is the optimal expected average age per-user of the original problem (3), and  $\frac{C^{RP,N}}{N}$  is the optimal

average age per-user of the relaxed problem (4). Thus, in order to show the asymptotic optimality, it is sufficient to prove that for a large number of users  $N$ ,  $\frac{C^{WIP,N}}{N}$  converges to  $\frac{C^{RP,N}}{N}$ . To that end, the next task is to find an expression of  $C^{RP} = \frac{C^{RP,N}}{N}$ . For this purpose, we provide the following proposition.

**Proposition 5.** *The optimal solution of the relaxed problem is of type threshold for each class. More precisely, it is a linear combination between two threshold vectors  $(l_1^1, l_2^1)$  and  $(l_1^2, l_2^2)$  such that:*

- There exists a unique real value  $W^* \in \mathbb{R}$ , a class  $m$  and state  $p$  such that  $W^* = W^m(p)$ .
- The expressions of  $l_k^1$  and  $l_k^2$  are as follows:

$$l_k^1 = \underset{i \in \mathbb{N}^*}{\operatorname{argmax}} \{W^k(i) : W^k(i) \leq W^*\} + 1 \quad \forall k \in \{1, 2\}$$

$$l_k^2 = \underset{i \in \mathbb{N}^*}{\operatorname{argmax}} \{W^k(i) : W^k(i) < W^*\} + 1 \quad \forall k \in \{1, 2\} \quad (14)$$

- There exists a unique  $0 < \theta \leq 1$  that satisfies  $\theta \sum_{k=1}^2 \gamma_k \sum_{i=l_k^1}^{+\infty} u_k^1(i) + (1 - \theta) \sum_{k=1}^2 \gamma_k \sum_{i=l_k^2}^{+\infty} u_k^2(i) = \alpha$ , where  $u_k^n$  is the stationary distribution of the age given a threshold  $n$  for class  $k$ .

*Proof.* See [21, Proposition 5].  $\square$

Thanks to this proposition, we can conclude that the optimal per-user cost of the relaxed problem has the following expression:

$$C^{RP} = \sum_{k=1}^2 \gamma_k \sum_{i=1}^{+\infty} [\theta u_k^1(i) + (1 - \theta) u_k^2(i)] i \quad (15)$$

By leveraging these results, we can proceed with characterizing the performance of the Whittle's index policy.

### B. Global Optimality of the Whittle's index policy

This section constitutes the main contribution of the paper where we show the asymptotic optimality of the Whittle's index policy. The idea is to show that the performance of this policy converges to  $C^{RP}$  when  $N$  is large and the ratio  $\alpha = \frac{M}{N}$  is kept constant.

We let  $Z_i^{k,N}(t)$  denote the proportion of users belonging to class  $k$  in state  $i$  at time  $t$ . In other words, it denotes the ratio of the number of users in class  $k$  having an age equal to  $i$  to the total number of users  $N$ . We have that  $\mathbf{Z}^N(t) = (Z^{1,N}(t), Z^{2,N}(t))$  with  $Z^{k,N}(t) = (Z_1^{k,N}(t), \dots, Z_{m^k(t)}^{k,N}(t))$ , where  $m^k(t)$  is the highest state at time  $t$  in class  $k$  and  $\sum_{i=0}^{m^k(t)} Z_i^{k,N}(t) = \gamma_k$  for each class  $k$ . We also denote by  $\mathbf{z}^*$  the proportion corresponding to the optimal policy of the relaxed problem. Thus, the elements of the vector  $\mathbf{z}^*$  are exactly the set  $\{\gamma_k (\theta u_k^1(i) + (1 - \theta) u_k^2(i))\}_{\substack{1 \leq k \leq 2 \\ 1 \leq i}}$  where  $i$  and  $k$  refer to the user  $i$  and class  $k$  respectively. This can be easily concluded from the results previously laid out in eq. (15).

Before tackling our novel approach, we give in the following two common methods used in the literature to study the convergence of Whittle's index policy while emphasizing their shortcomings:

- 1) Method based on fluid approximation technique used in [21], [22], [19] for the local asymptotic optimality of Whittle's index policy. This method consists basically of:
  - Showing that the approximated vector of  $\mathbf{Z}^N(t)$  denoted by  $\mathbf{z}(t)$  converges to  $\mathbf{z}^*$  starting from any initial fluid approximation vector  $\mathbf{z}(0)$ .
  - Establishing that  $\mathbf{Z}^N(t)$  converges in probability to  $\mathbf{z}^*$  when  $N$  scales.
  - Proving that  $\frac{C^{WIN,N}}{N}$  converges to  $C^{RP}$  when  $N$  scales.

The difficulty encountered by adopting this aforementioned method, is to find a simple equation satisfied by  $\mathbf{z}(t)$  in order to demonstrate the first point which is the convergence of  $\mathbf{z}(t)$  to  $\mathbf{z}^*$ . To circumvent this problem, the authors in [21] assume the following:

#### Assumption 1. Local optimality:

*The initial state  $\mathbf{z}(0)$  is within a very tight neighborhood of  $\mathbf{z}^*$ .*

Leveraging this assumption, the authors in [21] proved the convergence of  $\mathbf{z}(t)$  by establishing the linear equation  $\mathbf{z}(t+1) = \mathbf{Q}\mathbf{z}(t) + \mathbf{c}$  (which can be obtained only under the aforementioned assumption) and by showing that the spectral value of  $\mathbf{Q}$  is strictly less than 1

- 2) Method based on the stationary distribution of  $\mathbf{Z}^N(t)$  for asymptotic global optimality. This method consists of:

- Establishing the existence of the stationary distribution of  $\mathbf{Z}^N(t)$  denoted by  $\mathbf{Z}^N(\infty)$  for fixed  $N$ .
- Showing that  $\mathbf{Z}^N(\infty)$  converges in probability to any neighborhood of  $\mathbf{z}^*$  when  $N$  scales under the following assumption:

#### Assumption 2. Recurrence assumption:

*The expected time to enter a neighborhood of  $\mathbf{z}^*$  from any initial state  $\mathbf{x}$  does not depend on the number of users  $N$ . In other words, for all  $N$  the time when  $\mathbf{Z}^N(t) \in \Omega_\epsilon(\mathbf{z}^*)$  denoted by  $\Gamma_{\mathbf{x}}^N(\epsilon)$  is bounded by a constant  $T_{b_\epsilon}$ .*

- Proving that  $\frac{C^{WIN,N}}{N}$  converges to  $C^{RP}$  when  $N$  scales.

This latter assumption simplifies the analysis significantly, as one can see in [21]. Accordingly, proving the asymptotic global optimality is based essentially on this assumption which is a very strong assumption that, not only may not be satisfied in all scenarios, but that can only be verified numerically by simulating the system. Proving that this assumption holds is not at all a simple task. One should note that this aforementioned assumption was also considered in [22] and [17] to establish the global optimality of WIP but for totally different resource allocation problems. Therefore, proving the global asymptotic optimality of

WIP without the aforementioned assumption is one of the main challenging tasks that need to be tackled, not only for age-oriented scheduling problems but for many other resource allocation problems. In this paper, we removed the aforementioned assumption and established the asymptotic global optimality without requiring any strong assumption. We apply for that the fluid approximation used in [21] to prove local optimality as presented before. However, the approach here is different from the fluid approximation in [21] as follows. In [21], it is assumed that the initial state  $z(0)$  is within a tight neighborhood of  $z^*$ , which allows to obtain a linear equation between  $z(t+1)$  and  $z(t)$  (obtaining the linear equation is conditioned on this assumption of the initial state of  $z(0)$  as has been considered in [1]). In fact, the general relation between  $z(t+1)$  and  $z(t)$  regardless of the initial state is a nonlinear equation  $z(t+1) = \mathbf{Q}(z(t))z(t) + c(t)$ . Besides that, since the AoI evolves in an infinite state space, the matrix  $\mathbf{Q}(z(t))$  is not square. Therefore, the spectral value of  $\mathbf{Q}(z(t))$  is not defined. This explains why the fluid approximation approach in [21] cannot be used to prove the asymptotic global optimality of WIP. To that extent, the main technical challenge to prove the global optimality of WIP is to find a simple relation between the elements of the vector  $z(t+1)$  and those of the vector  $z(t)$ . In this paper, instead of having a complete relation between  $z(t+1)$  and  $z(t)$ , we show that it is sufficient to have a partial relation between  $z(t+1)$  and  $z(t)$  to establish the global asymptotic optimality of WIP. Specifically, the basis of our new method in this paper is as follows:

- Defining  $\alpha_1(t)$  and  $\alpha_2(t)$  as users' proportions that refer to the scheduled users' proportion at time  $t$  in class 1 and 2 respectively ( $\alpha_1(t) + \alpha_2(t) = \alpha$ )
- Expressing each proportion  $z_i^k(t)$  in function of one term of  $\{\alpha_k(t')\}_{t' \leq t}$
- Proving that  $\alpha_k(t)$  converges for  $k = 1, 2$  and conclude that  $z(t)$  also converges.

The hardest task among the three steps above is the last one since we do not dispose of an explicit expression of  $\alpha_k(t)$  in function of  $t$ . That is why we proceed with non trivial and intricate mathematical analyses depicted in Propositions 11,12,13 and Theorem 1 to describe the evolution of  $\alpha_k(t)$  and its convergence when  $t$  grows. In more detail, the main steps of our new approach are:

- Since the relation between  $z(t+1)$  and  $z(t)$  is not linear, our approach to establish the convergence of  $z(t)$  involves two terms:  $\alpha_1(t)$  and  $\alpha_2(t)$ . As mentioned above, these two proportions are nothing but the scheduled proportion at time  $t$  of class 1 and 2, respectively. Note that we always have  $\alpha_1(t) + \alpha_2(t) = \alpha$ . Based on Lemma 1, we show that for a large enough time  $t$ , there exists  $T_t$  such that we can find a partial relation between each element of the vector  $z(t+T_t)$  and terms of the sequence  $\{\alpha_k(t')\}_{\substack{k=1,2 \\ t' \leq t+T_t}}$ . More precisely, we prove that for  $T_t$ , we can express each proportion that is not scheduled at time

$t + T_t$  in function of one term of  $\{\alpha_k(t')\}_{\substack{k=1,2 \\ t' \leq t+T_t}}$ . This allows us to obtain  $1 - \alpha$  as a linear combination between the terms of  $\{\alpha_k(t')\}_{\substack{k=1,2 \\ t' \leq t+T_t}}$  at time  $t + T_t$ .

- Subsequently, we introduce in Definition 4,  $T_{\max}$  that satisfies these two following properties proven in Propositions 7 and 8 using Lemma 2: the Whittle's index alternates (the definition of index alternation will be given later on in this section) between the two classes from state 1 to  $T_{\max}$  under a given assumption on  $\alpha$ ; the instantaneous thresholds  $l_1(\cdot)$  and  $l_2(\cdot)$  are bounded by  $T_{\max}$  at time  $t + T_t$ .
- Based on that, we derive the relation between the instantaneous thresholds at time  $t + T_t$  in Proposition 9. Taking as initial time  $t + T_t = T_0$ , we show by induction in Proposition 10 that, for all  $T \geq T_0$ , the instantaneous thresholds are less than  $T_{\max}$  and that all none scheduled proportions can be expressed in function of terms of the sequence  $\{\alpha_k(t')\}_{\substack{k=1,2 \\ t' \leq T}}$ . Next, we define for each class  $k$  a vector  $A_k(T)$  composed by  $\alpha_k(T)$  (the scheduled proportion at time  $T$ ) plus the finite subset of the sequence  $\{\alpha_k(t')\}_{t' \leq T}$  such that for all proportion of users in class  $k$  at a given state at time  $T$  that is not scheduled can be expressed by one element belonging to this subset. After that, we provide the relation between the elements of the vectors  $A_k(T)$  and  $A_k(T+1)$  in Propositions 11 and 12
- As was mentioned in the introduction, our proof is based on Cauchy criterion which states that in the real number space  $\mathbb{R}$ , a given sequence  $h(T)$  is convergent if and only if its terms become closer together as  $T$  increases. To that extent, we show that the elements of the vector  $A_k(\cdot)$  which are nothing but the terms of the sequence  $\alpha_k(\cdot)$  are getting closer when  $T$  increases. For that purpose, we prove that the highest and the smallest element of  $A_k(T)$  converge to the same limit when  $T$  grows. To that end, we start by establishing the convergence of the highest and the smallest element of  $A_k(T)$  in Theorem 1. Then, we demonstrate by contradiction that the highest and the smallest element of  $A_k(T)$  must converge to the same limit in Proposition 13. This last result implies that  $\alpha_k(t)$  converges when  $t$  scales. In light of that fact, we prove that  $z(t)$  converges to  $z^*$  in Proposition 14. Finally, using Kurth theorem, we show in Proposition 15 that  $\mathbf{Z}^N(t)$  converges to  $z^*$  in probability, and finally we establish in Proposition 16 the convergence of  $\frac{C^{WIP,N}}{N}$  to  $C^{RP}$ .

With the steps of our approach clarified, we can proceed with introducing the fluid limit approximation. The fluid limit technique consists of analyzing the evolution of the expectation of  $\mathbf{Z}^N(t)$  under the Whittle's index policy. For that, we define the vector  $z(t)$  as follows:

$$z(t+1) - z(t)|_{z(t)=z} = \mathbb{E} \left[ \mathbf{Z}^N(t+1) - \mathbf{Z}^N(t) | \mathbf{Z}^N(t) = z \right] \quad (16)$$

This above equation reveals to us that we have a sequence  $z(t)$  defined by recurrence for a fixed initial state  $z(0)$  that we should study its behavior when  $t$  is very large. Hence, we end up with a function  $z(t)$  that depends on two variables,  $t$  and

the initial value  $z(0)$ . To that extent, our aim is to prove that  $z(t)$  converges to  $z^*$  regardless of the initial state  $z(0)$ . We let  $z(t) = (z^1(t), z^2(t))$  with  $z^k(t) = (z_1^k(t), \dots, z_{m^k}^k(t))$  where  $z_i^k(t)$  is the expected proportion of users at state  $i$  in class  $k$  at time  $t$  with respect to the equation (16). Accordingly we have that  $\sum_{i=0}^{m^k(t)} z_i^k(t) = \gamma_k$  for each class  $k$ . We can also formulate the fluid limit in Equation (16) as follows:

$$z(t+1) = \mathbb{E} \left[ \mathbf{Z}^N(t+1) \middle| \mathbf{Z}^N(t) = z(t) \right] \quad (17)$$

One can notice that  $z^*$  is a particular vector with respect to the equation (17).

**Proposition 6.**  $z^*$  is the *unique* fixed point of the fluid approximation equation, or equivalently the unique solution of the following equation:

$$z = \mathbb{E} \left[ \mathbf{Z}^N(t+1) \middle| \mathbf{Z}^N(t) = z \right]$$

In other words,  $z(t) = z(t+1)$ , if and only if  $z(t) = z^*$ .

*Proof.* The proof follows the same methodology as in [22, Lemma 9]  $\square$

According to this proposition, it is sufficient to show that  $z(t)$  converges starting from any initial state  $z(0)$ , as the only eventual finite limit of  $z(t)$  when  $t$  tends to  $+\infty$  is the fixed point of the equation (16),  $z^*$ .

**Remark 1.** We highly emphasize that the proportion  $\alpha_k(t)$  and  $1 - \alpha$  refer to the scheduled users' proportion at time  $t$  in class  $k$  and the non scheduled users' proportion either for class 1 or 2 respectively. Meanwhile, for any other proportion  $A$ , it refers only to the number of users in this proportion over the total users' number of the system whatever the different states of users that contains. Having said that,  $A = B$  means that they are equal in terms of proportion, while they can contain users in different states.

In the following, we prove that the fluid approximation vector of  $\mathbf{Z}^N(t)$ ,  $z(t)$  under Whittle Index Policy converges starting from any initial state. We prove this result for 2 different classes of users where  $p_1$  and  $p_2$  are the successful transmission probabilities of the class 1 and 2 respectively ( $p_1 > p_2$ ), given a sufficient condition on  $\alpha$ . Throughout this section, we denote by  $w_1(n)$  and  $w_2(n)$ , the Whittle's index, whose expression is given in Proposition 4, of state  $n$  in class 1 and class 2 respectively. We need to prove that  $z_i^k(t)$  converges for each state  $i$  in class  $k$ .

Now, focusing on the Whittle index policy, we can see it as an instantaneous threshold policy for each class, where the thresholds vary over time  $t$ . Moreover, under the Whittle index policy, the proportion of users that are scheduled at each time slot  $t$  is fixed and equals to  $\alpha$  since the number of scheduled users at each time slot  $t$  is  $\alpha N$ . This proportion  $\alpha$  contains the users with the highest Whittle index values. In that respect, we define  $\alpha_1(t)$  and  $\alpha_2(t)$  the proportion of users in class 1 and class 2 respectively at time  $t$  with the highest Whittle

index values such that  $\alpha_1(t) + \alpha_2(t) = \alpha$ . The remaining proportion of users which are not scheduled at each time slot  $t$ , which is equal to  $1 - \alpha$ , contains the users with the smallest Whittle index values. Now, regarding this proportion, we give its decomposition into proportions of users at different states in different classes. Denoting by  $l_1(t)$  and  $l_2(t)$  at time  $t$  the instantaneous threshold integers under Whittle index policy, then there exists two real values between 0 and 1,  $\beta(t)$  and  $\gamma(t)$ , with  $\gamma(t) = 1$  and  $0 < \beta(t) \leq 1$ , or  $0 < \gamma(t) \leq 1$  and  $\beta(t) = 1$ , such that:

$$\sum_{i=1}^{l_1(t)-1} z_i^1(t) + \sum_{i=1}^{l_2(t)-1} z_i^2(t) + \beta(t) z_{l_1(t)}^1(t) + \gamma(t) z_{l_2(t)}^2(t) = 1 - \alpha \quad (18)$$

and  $\{z_i^1(t)\}_{1 \leq i \leq l_1(t)} \cup \{z_i^2(t)\}_{1 \leq i \leq l_2(t)}$  is exactly the set  $\{z_i^k(t) : w_k(i) \leq \max(w_1(l_1(t)), w_2(l_2(t)))\}$ .

In [21], as mentioned before, in order to prove the convergence of  $z(t)$ , it is assumed that  $z(0)$  is within a precise neighborhood of  $z^*$  and that the number of states is finite. These assumptions allow to find an easy linear relation between  $z(t)$  and  $z(t+1)$  ( $z(t+1) = \mathbf{Q}z(t) + c$  see [21, Section IV-C]), and then to deduce the convergence of  $z$  by establishing that the spectral value of  $\mathbf{Q}$  is strictly less than one. In our case, as we aim to prove the convergence of  $z$  from any initial state, the relation between  $z(t+1)$  and  $z(t)$  is as follows

$$z(t+1) = \mathbf{Q}(z(t))z(t) + c(t) \quad (19)$$

This equation is not linear which makes studying the evolution of  $z(\cdot)$  a hard task. Moreover, as the number of state is infinite, then the dimensions of  $z(t)$  varies per time. Therefore, the matrix  $\mathbf{Q}(z(t))$  is not square. Hence we can not apply the same method as in [21] since the spectral values are not defined for a non square matrix. For these reasons, we proceed differently than [21]. Our method consists in fact in expressing each proportion  $z_i^k(t)$  that belongs to a non scheduled users' proportion at time  $t$  in function of a term of  $\alpha_k(\cdot)$  at a given time less than  $t$ . By this way, we will obtain a part of the vector  $z(t)$  in function of  $\{\alpha_k(t')\}_{t' \leq t, k \in \{1,2\}}$ , and the sum of the other part is equal to  $\alpha$ . Then, we show that  $\alpha_k(\cdot)$  converges for  $k = 1, 2$ . We will see later that it is sufficient to show that  $\alpha_k(\cdot)$  converges in order to conclude for the convergence of  $z(\cdot)$ . To find the partial relation between  $z(t)$  and  $\{\alpha_k(t')\}_{t' \leq t, k \in \{1,2\}}$ , we will be in need of Equation

(2) that illustrates the evolution of AoI. In fact, according to this equation, if user  $i$  in class  $k$  is scheduled at time  $t$  ( $a_i^k(t) = 1$ ), the AoI will move to state 1 with probability  $p_k$  (the channel state is  $c_i^k(t) = 1$ ). Otherwise, the value of AoI will be incremented by one with probability  $1 - p_k$  (the channel state is  $c_i^k(t) = 0$ ). When user  $i$  is not scheduled, according to Equation (2), the value of AoI will be incremented by one with probability one. This implies that, considering our system model that contains many users, if the whole proportion  $z_i^k(t)$  that represents the users in class  $k$  at state  $i$  is scheduled, on average  $p_k z_i^k$  from the proportion  $z_i^k$  will move to state 1. Similarly, if  $z_i^k$  is not scheduled, the state of all users in this proportion will be incremented by one. To that extent, we give



the following lemma that illustrates the evolution of  $z(\cdot)$  under the Whittle Index Policy.

**Lemma 1.** *Knowing  $z^k(t)$ ,  $\alpha_k(t)$  and  $l_k(t)$ , we have that:*

For  $i = 1$ :

$$z_1^k(t+1) = p_k \alpha_k(t).$$

For  $1 \leq i < l_k(t)$ :

$$z_{i+1}^k(t+1) = z_i^k(t).$$

*Proof.* See Appendix A.  $\square$

According to Lemma 1, after scheduling under the Whittle's Index Policy, we get at time  $t+1$ , a proportion of  $p_1 \alpha_1(t)$  of users at state 1 in class 1 and  $p_2 \alpha_2(t)$  of users at state 1 in class 2 respectively (i.e.  $z_1^1(t+1) = p_1 \alpha_1(t)$  and  $z_1^2(t+1) = p_2 \alpha_2(t)$ ).

According to the same lemma, at time  $t+2$ , a proportion of  $p_1 \alpha_1(t)$  and  $p_2 \alpha_2(t)$  of users will go to state 2 in class 1 and class 2 respectively and  $p_1 \alpha_1(t+1)$ ,  $p_2 \alpha_2(t+1)$  of users will move to state 1 in class 1 and class 2 respectively (i.e.  $z_1^1(t+2) = p_1 \alpha_1(t+1)$ ,  $z_2^2(t+2) = p_2 \alpha_2(t+1)$ ,  $z_1^2(t+2) = p_1 \alpha_1(t)$  and  $z_2^1(t+2) = p_2 \alpha_2(t)$ ).

At time  $t+3$ , a proportion of  $p_1 \alpha_1(t)$  and  $p_2 \alpha_2(t)$  of users will go to state 3 in class 1 and class 2 respectively,  $p_1 \alpha_1(t+1)$ ,  $p_2 \alpha_2(t+1)$  of users will move to state 2 in class 1 and class 2 respectively,  $p_1 \alpha_1(t+2)$  and  $p_2 \alpha_2(t+2)$  of users will move to state 1 in class 1 and class 2 respectively, (i.e.  $z_1^1(t+3) = p_1 \alpha_1(t+2)$ ,  $z_2^2(t+3) = p_2 \alpha_2(t+2)$ ,  $z_2^1(t+3) = p_1 \alpha_1(t+1)$  and  $z_3^2(t+3) = p_2 \alpha_2(t+1)$ ,  $z_3^1(t+3) = p_1 \alpha_1(t)$ ,  $z_3^2(t+3) = p_2 \alpha_2(t)$ ).

Thereby, at time  $t+t_0$  where the instantaneous threshold  $l_k(t+t_0) \geq t_0$ , we get a set of proportions  $\{p_1 \alpha_1(t), p_2 \alpha_2(t), \dots, p_1 \alpha_1(t+t_0-1), p_2 \alpha_2(t+t_0-1)\}$  that belong to the proportion  $1-\alpha$  of users with the lowest Whittle index values, such that  $z_1^1(t+t_0) = p_1 \alpha_1(t+t_0-1)$ ,  $z_2^2(t+t_0) = p_2 \alpha_2(t+t_0-1)$ ,  $\dots$ ,  $z_{t_0}^1(t+t_0) = p_1 \alpha_1(t)$  and  $z_{t_0}^2(t+t_0) = p_2 \alpha_2(t)$ . Hence, we obtain a  $z_i^k(t+1)$  which is well expressed in function of terms of  $\alpha_k(\cdot)$  ( $k = 1, 2$ ) for  $i \in [1, t_0]$ ,  $k = 1, 2$ .

**Remark 2.** *Considering Whittle index policy framework, the order of the different users' proportions with respect to their Whittle index values must be taking into account throughout this analysis. In fact, as we have already mentioned, we need to give the expression of the non scheduled users' proportions in function of the terms of  $\alpha_k(\cdot)$  for  $k = 1, 2$ , which can not be done only if we consider the order of the Whittle index values. To that extent, since the set of the non scheduled users' proportions, according to the Whittle's index policy, is exactly the set of users' proportions with the lowest Whittle index values among all the different users' proportions of the system, then the form at time  $t$  of this specific set will be  $\{z_i^k(t) : w_k(i) \leq w_m(n)\}$  for a given  $m$  and  $n$  that vary with  $t$ .*

Based on this remark above, we need to find at time  $t+t_0$ , a set of the form  $\{z_i^k(t+t_0) : w_k(i) \leq w_m(n)\}$  for a given class  $m$  and state  $n$ , such that all the elements of this set are well expressed in function of  $\alpha_k(\cdot)$ . We show in the sequel that the highest Whittle index of this set could be  $w_2(t_0)$ .

Indeed, given that the Whittle index function is increasing with  $n$  where  $n$  refers to a given age of information state, then for any state in class 2 with Whittle index less than  $w_2(t_0)$ , belongs to  $[1, t_0]$ . Moreover, considering the state  $q$  in class 1 such that  $w_1(q) \leq w_2(t_0) \leq w_1(t_0)$  ( $p_1 > p_2$ ), then  $w_1(q) \leq w_1(t_0)$ , which means that  $q \in [1, t_0]$ . Hence, for any element in  $\{z_i^k(t+t_0) : w_k(i) \leq w_2(t_0)\}$ , can be expressed in function of terms of  $\alpha_k(\cdot)$  ( $k = 1, 2$ ). Accordingly,  $\{z_i^k(t+t_0) : w_k(i) \leq w_2(t_0)\}$  equals to the set  $\{p_2 \alpha_2(t), \dots, p_2 \alpha_2(t+t_0-1), p_1 \alpha_1(t+t_0-l(t+t_0)), \dots, p_1 \alpha_1(t+t_0-1)\}$ , where  $l(t+t_0)$  is the greatest state  $q$  in class 1 such that  $w_1(q) \leq w_2(t_0)$ . We note that  $l(t+t_0) \leq t_0$  because  $w_2(l(t+t_0)) \leq w_1(l(t+t_0)) \leq w_2(t_0)$ . Therefore, in that regards, for a fixed  $t$ , we associate for each  $t_0$  the corresponding sum  $\sum_{j=1}^{l(t+t_0)} z_j^1(t+t_0) + \sum_{j=1}^{t_0} z_j^2(t+t_0) = \sum_{j=1}^{l(t+t_0)} p_1 \alpha_1(t+t_0-j) + \sum_{j=1}^{t_0} p_2 \alpha_2(t+t_0-j)$ . To that extent, we define in the following the time  $t_0$  when this aforementioned sum exceeds  $1-\alpha$ .

**Definition 2.** *Starting at time  $t$ , we define  $T_t$  such that  $t+T_t$  is the first time that verifies:*

$$\sum_{j=1}^{l(t+T_t)} p_1 \alpha_1(t+T_t-j) + \sum_{i=1}^{T_t} p_2 \alpha_2(t+T_t-j) \geq 1-\alpha \quad (20)$$

*In other words, the first time when  $\sum_{j=1}^{l(t+t_0)} p_1 \alpha_1(t+t_0-j) + \sum_{i=1}^{t_0} p_2 \alpha_2(t+t_0-j)$  exceeds  $1-\alpha$  is  $t+t_0 = t+T_t$ .*

Then, at time  $t+T_t$ , there exists  $l'_1(t+T_t) \leq l(t+T_t)$ ,  $l'_2(t+T_t) \leq T_t$ , such that the set  $\{z_i^1(t+T_t)\}_{1 \leq i \leq l'_1(t+T_t)} \cup \{z_i^2(t+T_t)\}_{1 \leq i \leq l'_2(t+T_t)}$  is exactly the set  $\{z_i^k(t+T_t) : w_k(i) \leq \max(w_1(l'_1(t+T_t)), w_2(l'_2(t+T_t)))\}$ <sup>1</sup>, and  $\gamma(t+T_t) = 1$  and  $0 < \beta(t+T_t) \leq 1$ , or  $0 < \gamma(t+T_t) \leq 1$  and  $\beta(t+T_t) = 1$  such that:

$$\begin{aligned} & \sum_{j=1}^{l'_1(t+T_t)-1} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{l'_2(t+T_t)-1} p_2 \alpha_2(t+T_t-j) \\ & + \beta(t+T_t) p_1 \alpha_1(t+T_t-l'_1(t+T_t)) \\ & + \gamma(t+T_t) p_2 \alpha_2(t+T_t-l'_2(t+T_t)) \\ & = 1-\alpha, \end{aligned} \quad (21)$$

with  $l'_1(t+T_t)$  and  $l'_2(t+T_t)$  being the instantaneous thresholds in class 1 and 2 respectively at time  $t+T_t$ .  $\alpha_1(t+T_t)$  and  $\alpha_2(t+T_t)$  are the users' proportions with the highest Whittle index values, and their sum is equal to  $\alpha$ . Without loss of generality, we let  $l'_k(t+T_t) = l_k(t+T_t)$ .

As we can see, at time  $t+T_t$ , all the expressions of the users' proportions that belong to the  $1-\alpha$  of users with the smallest Whittle index values, are in function of  $\alpha_1(t)$  or  $\alpha_2(t)$  at various time. In fact, at time  $t+T_t$ , we end up with  $z_1^1(t+T_t) = p_1 \alpha_1(t+T_t-1)$ ,  $z_2^2(t+T_t) = p_2 \alpha_2(t+T_t-1)$ ,  $\dots$ ,  $z_{l'_1(t+T_t)}^1(t+T_t) = p_1 \alpha_1(t+T_t-l_1(t+T_t))$  and  $z_{l'_2(t+T_t)}^2(t+T_t) = p_2 \alpha_2(t+T_t-l_2(t+T_t))$ , and the rest of the proportions belongs to  $\alpha_1(t+T_t)$  for class 1 and  $\alpha_2(t+T_t)$  for class 2.

<sup>1</sup>According to Remark 2, the form of this set means that it contains the users' proportions with the lowest Whittle index values among all users' proportions of the system

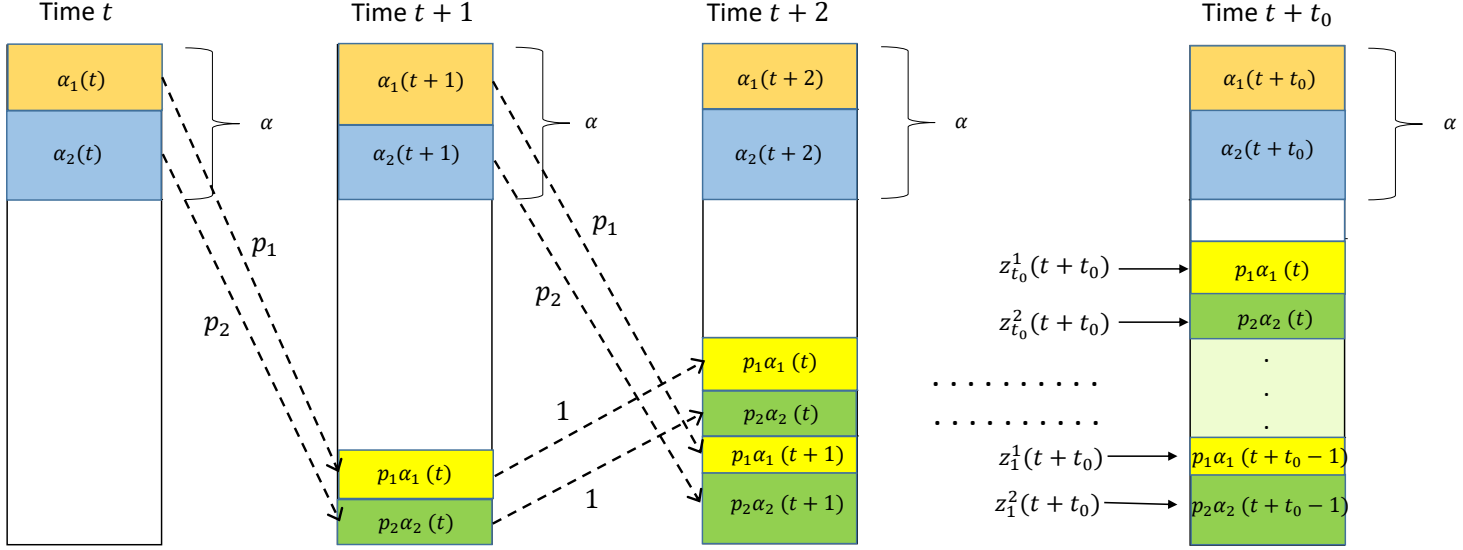


Figure 2: Evolution of  $\alpha_1(t)$  and  $\alpha_2(t)$  under the Whittle Index Policy (the green and the yellow colors refer to class 1 and 2 respectively)

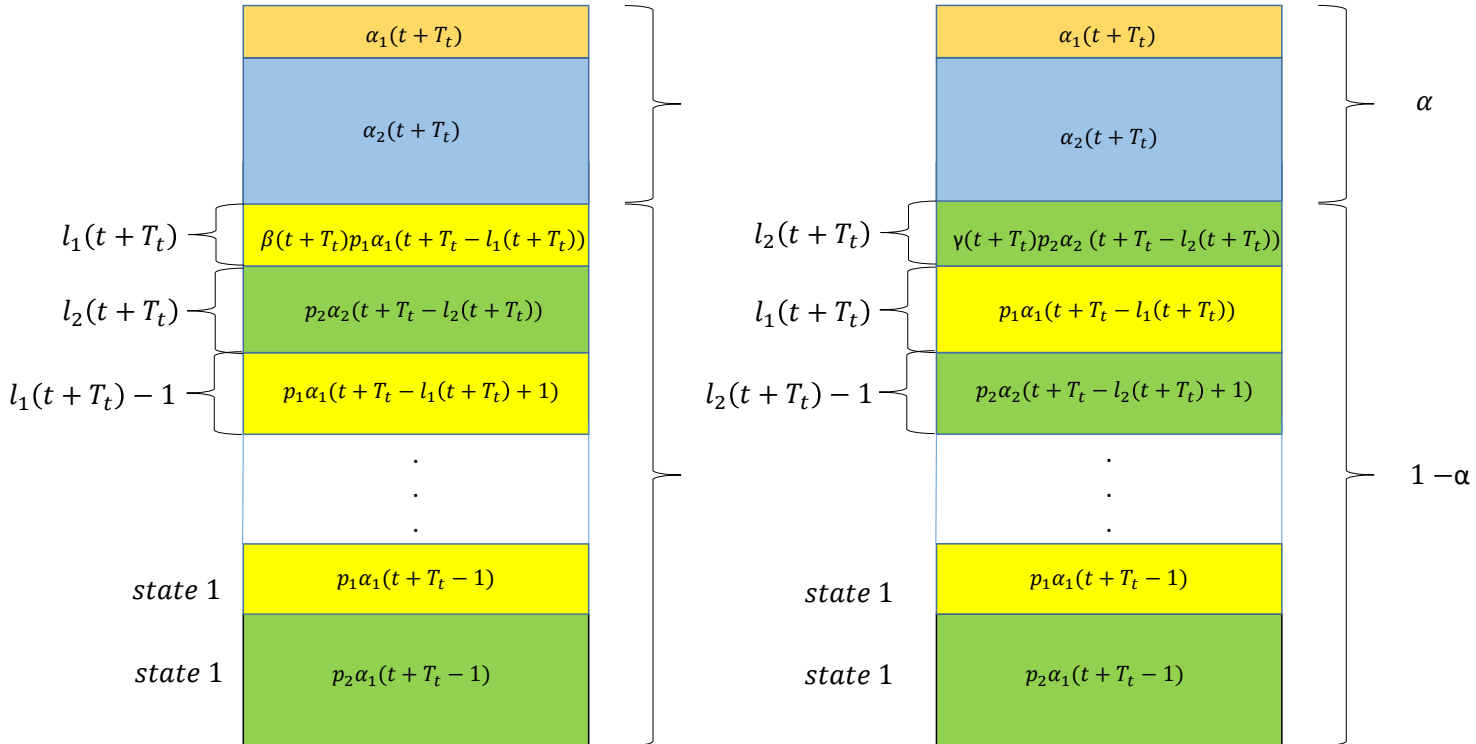


Figure 3: The proportions of users at different states at time  $t+T_t$  when  $\gamma(t+T_t) = 1$  and  $0 < \beta(t+T_t) \leq 1$

Figure 4: The proportions of users at different states at time  $t+T_t$  when  $\beta(t+T_t) = 1$  and  $0 < \gamma(t+T_t) \leq 1$

For this reason, we work only with  $\alpha_1(\cdot)$  and  $\alpha_2(\cdot)$  in order to prove the convergence.

As we have mentioned earlier, the proof of the optimality is valid under the **alternation condition** that we define as follows:

**Definition 3.** We say that the order of the Whittle index strictly alternates between the two classes in  $[1, n]$  or from state 1 to  $n$ , if we have  $w_2(1) < w_1(1) < w_2(2) < w_1(2) < w_2(3) < w_1(3) < \dots < w_2(n) < w_1(n)$ .

The proof of  $\alpha_k(\cdot)$  convergence is feasible when the alternation condition is satisfied from 1 to  $l_k(t + T_t) + 1$  for all  $t$ . We note that this condition will be relevant in the proof of Proposition 12. To that end, we start by introducing the assumption on  $\alpha$ . Then, we demonstrate effectively that under this assumption the condition of alternation is satisfied from 1 to  $l_k(t + T_t) + 1$ .

**Assumption 3.** Denoting  $\frac{1}{p_1 - p_2} \left( \frac{p_1 + p_2}{2} + \sqrt{2(p_1 - p_2) + \frac{(p_1 + p_2)^2}{4}} \right)$  by  $D$ . Then, the users' proportion scheduled at each time  $\alpha$  satisfies:

$$\alpha > \frac{1}{1 + (D - 2)p_2} \quad (22)$$

If we are able to find an upper bound of  $l_k(t + T_t)$  that we denote by  $l_{\max}$ , then it is sufficient to prove that the hypothesis of the Whittle index alternation is satisfied from 1 to  $l_{\max} + 1$  in order to prove our desired result. To that end, we start first by defining certain constant  $T_{\max}$ . Then under Assumption 3, we show that the order of Whittle index alternates between the two classes in the set  $[1, T_{\max} + 1]$  and establish that  $T_{\max}$  is effectively an upper bound of  $l_k(t + T_t)$ , i.e.  $T_{\max} = l_{\max}$ .

First of all, we give a lemma which will be useful to prove the propositions 8, 9 and 10.

**Lemma 2.** There exists a time  $t_f$  such that for all  $t \geq t_f$ ,  $\alpha_1(t) > 0$ .

*Proof.* See appendix B.  $\square$

In this following definition, we define  $T_{\max}$ , and we check later that it coincides with the upper bound of  $l_k(t + T_t)$  for  $k = 1, 2$ .

**Definition 4.** We define the integer  $T_{\max}$  as:  $T_{\max} = \lfloor \frac{1 - \alpha}{p_2 \alpha} + 1 \rfloor$ , where  $\lfloor x \rfloor$  refers to the integer part of  $x$ .

**Proposition 7.** Under Assumption 3, the order of the Whittle index alternates between the two classes from state 1 to state  $T_{\max} + 1$ .

*Proof.* See appendix C.  $\square$

Now we prove that the instantaneous thresholds of the two classes can not exceed  $T_{\max}$ .

**Proposition 8.** Denoting by  $l_{\max}$  the highest instantaneous threshold in the sense that  $\forall t \geq t_f, \max(l_1(t + T_t), l_2(t + T_t)) \leq l_{\max}$ , then  $T_{\max} = l_{\max}$ .

*Proof.* See appendix D  $\square$

According to the last proposition,  $T_{\max}$  is truly the upper bound of  $l_k(t + T_t)$  for all  $t$  and  $k = 1, 2$ . As a consequence, the order of the Whittle index alternates between the two classes in the set  $[1, l_k(t + T_t) + 1]$ . The next goal is to find a relation between  $l_1(t + T_t)$  and  $l_2(t + T_t)$ . To do so, we recall that we have at time  $t$ :

$$\sum_{i=1}^{l_1(t)-1} z_i^1(t) + \sum_{i=1}^{l_2(t)-1} z_i^2(t) + \beta(t) z_{l_1(t)}^1(t) + \gamma(t) z_{l_2(t)}^2(t) = 1 - \alpha \quad (23)$$

with  $l_1(t)$  and  $l_2(t)$  being the thresholds in class 1 and 2 respectively at time  $t$ , and  $\beta(t) = 1$  and  $0 < \gamma(t) \leq 1$ , or  $\gamma(t) = 1$  and  $0 < \beta(t) \leq 1$ . Thereby, the first step consists of establishing the relationship between  $l_1(t)$  and  $l_2(t)$  when  $\max(l_1(t), l_2(t)) \leq T_{\max}$  depending on two different cases that we will explain thereafter in order to give a generalized expression of the aforementioned equation (23).

**Remark 3.** It is worth mentioning that, as we have defined  $l_{\max}$  in Proposition 8, it refers to the highest value that can be attained by the thresholds of the class 1 or 2 at time  $t + T_t$  for  $t > t_f$  where  $t_f$  is a given in Lemma 2. Whereas, at any time  $t > t_f$ ,  $\max(l_1(t), l_2(t)) \leq l_{\max}$  might not be true since we don't have necessary a given  $t'$  such that  $t' + T_{t'} = t$  for any  $t' > t_f$ .

**Proposition 9.** At any time  $t > t_f$ , if  $\max(l_1(t), l_2(t)) \leq T_{\max} = l_{\max}$ , then there exists  $l(t) \leq l_{\max}$  and,  $\beta(t) = 0$  and  $0 < \gamma(t) \leq 1$ , or  $0 < \beta(t) \leq 1$  and  $\gamma(t) = 1$  such that:

$$\sum_{i=1}^{l(t)-1} z_i^1(t) + \sum_{i=1}^{l(t)-1} z_i^2(t) + \beta(t) z_{l(t)}^1(t) + \gamma(t) z_{l(t)}^2(t) = 1 - \alpha \quad (24)$$

*Proof.* See appendix F.  $\square$

Starting at time  $t \geq t_f$ , we have that at time  $t + T_t$ , the thresholds  $l_1(t + T_t)$  and  $l_2(t + T_t)$  are less than  $l_{\max}$ . Hence, according to Proposition (9), there exists  $l(t + T_t)$  such that:

$$\begin{aligned} \sum_{j=1}^{l(t+T_t)-1} p_1 \alpha_1(t + T_t - j) + \sum_{j=1}^{l(t+T_t)-1} p_2 \alpha_2(t + T_t - j) \\ + \beta(t + T_t) p_1 \alpha_1(t + T_t - l(t + T_t)) \\ + \gamma(t + T_t) p_2 \alpha_2(t + T_t - l(t + T_t)) \\ = 1 - \alpha \end{aligned} \quad (25)$$

where  $\beta(t + T_t) = 0$  and  $0 \leq \gamma(t + T_t) < 1$ , or  $0 \leq \beta(t + T_t) < 1$  and  $\gamma(t + T_t) = 1$ .

Denoting  $t + T_t$  by  $T_0$ , we obtain:

$$\begin{aligned} \sum_{j=1}^{l(T_0)-1} p_1 \alpha_1(T_0 - j) + \sum_{j=1}^{l(T_0)-1} p_2 \alpha_2(T_0 - j) \\ + \beta(T_0) p_1 \alpha_1(T_0 - l(T_0)) + \gamma(T_0) p_2 \alpha_2(T_0 - l(T_0)) \\ = 1 - \alpha \end{aligned} \quad (26)$$

where  $\beta(T_0) = 0$  and  $0 < \gamma(T_0) \leq 1$ , or  $0 < \beta(T_0) \leq 1$  and  $\gamma(T_0) = 1$ .

Now, we prove by induction that if this latter expression is valid for all  $T \geq T_0$ , and that  $l(T)$ , the instantaneous threshold at time  $T$ , is less than  $l_{\max}$ .

**Proposition 10.** *For all  $T \geq T_0$ , there exists  $l(T) \leq l_{\max}$ ,  $\beta(T)$  and  $\gamma(T)$ , such that:*

$$\begin{aligned} & \sum_{j=1}^{l(T)-1} p_1 \alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2 \alpha_2(T-j) \\ & + \beta(T) p_1 \alpha_1(T-l(T)) + \gamma(T) p_2 \alpha_2(T-l(T)) \\ & = 1 - \alpha \end{aligned} \quad (27)$$

where  $\beta(T) = 0$  and  $0 < \gamma(T) \leq 1$ , or  $0 < \beta(T) \leq 1$  and  $\gamma(T) = 1$ .

*Proof.* See appendix G.  $\square$

According to the latter proposition, we can now define at each time  $T \geq T_0$ , for each class  $k$ , the vector  $\mathbf{A}_k(T) = (\alpha_k(T), \alpha_k(T-1), \dots, \alpha_k(T-l(T)))$ , such that, there exists  $\beta(T)$  and  $\gamma(T)$ :

$$\begin{aligned} & \sum_{j=1}^{l(T)-1} p_1 \alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2 \alpha_2(T-j) \\ & + \beta(T) p_1 \alpha_1(T-l(T)) + \gamma(T) p_2 \alpha_2(T-l(T)) = 1 - \alpha \end{aligned} \quad (28)$$

where  $\beta(T) = 0$  and  $0 < \gamma(T) \leq 1$ , or  $0 < \beta(T) \leq 1$  and  $\gamma(T) = 1$ . We note that as we have explained previously, the relation between  $\mathbf{A}_k(T)$  and  $\mathbf{z}^k(T)$  is:  $p_k \alpha_k(T-1) = z_1^k(T)$ ,  $p_k \alpha_k(T-2) = z_2^k(T)$ ,  $\dots$ ,  $p_k \alpha_k(T-l(T)) = z_{l(T)}^k(T)$ .

**Remark 4.** *We emphasize that in the following analysis,  $T$  is always considered greater than  $T_0$ .*

We prove in the sequel that  $\max \mathbf{A}_k(T)$  is decreasing and  $\min \mathbf{A}_k(T)$  is increasing (with the max and min referring to the element of the vector with the greatest value, and the smallest value respectively). After that, we conclude the convergence of  $\max \mathbf{A}_k(T)$  and  $\min \mathbf{A}_k(T)$  when  $T$  tends to  $+\infty$ . Then, we prove that they must converge to the same real number. In order to prove that  $\max \mathbf{A}_k(T)$  is decreasing and  $\min \mathbf{A}_k(T)$  is increasing, we first demonstrate this following proposition.

**Proposition 11.** *All the elements of the vector  $\mathbf{A}_k(T+1)$  belong to the elements of the vector  $\mathbf{A}_k(T)$  except  $\alpha_k(T+1)$ .*

*Proof.* See appendix H.  $\square$

With the intention of proving the monotony of  $\max \mathbf{A}_1(T)$  and  $\min \mathbf{A}_1(T)$ , we still need to prove that the value of  $\alpha_1(T+1)$  must be less than  $\max \mathbf{A}_1(T)$  and greater than  $\min \mathbf{A}_k(T)$ . For that, we introduce the following proposition. Before doing that, we note that, as  $\alpha_1(t) + \alpha_2(t) = \alpha$  at each time slot  $t$ , then it is sufficient for us to prove that  $\alpha_1(\cdot)$  is converging. To that extent, we study only the vector function  $\mathbf{A}_1(T)$  in order to prove the convergence.

**Proposition 12.** *Under assumption 3, for a given vector  $\mathbf{A}_1(T) = (\alpha_1(T), \alpha_1(T-1), \dots, \alpha_1(T-l(T))) (T \geq T_0)$ , we have four possible cases of inequalities:*

$$\alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T))$$

$$\alpha_1(T-l(T)) \leq \alpha_1(T+1) \leq \alpha_1(T)$$

$$\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T)$$

$$\alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T)+1)$$

*Moreover: If  $\alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T))$ , then:*

$$\alpha_1(T+1) - \alpha_1(T) \leq p_1(\alpha_1(T-l(T)) - \alpha_1(T))$$

*If  $\alpha_1(T-l(T)) \leq \alpha_1(T+1) \leq \alpha_1(T)$ , then:*

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)))$$

*If  $\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T)$ , then:*

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1))$$

*If  $\alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T)+1)$ , then:*

$$\alpha_1(T+1) - \alpha_1(T) \leq p_1(\alpha_1(T-l(T)+1) - \alpha_1(T))$$

*Proof.* See appendix I.  $\square$

**Theorem 1.**  *$\min \mathbf{A}_1(T)$  and  $\max \mathbf{A}_1(T)$  converge and we denote their limits respectively by  $l_1$  and  $l_2$ .*

*Proof.* According to Proposition 11, the elements of the vector  $\mathbf{A}_1(T+1)$  except the first element which is  $\alpha_1(T+1)$  belong to the elements of the vector  $\mathbf{A}_1(T)$ . Hence, the values of these elements (except the first element of  $\mathbf{A}_1(T+1)$ ) is less than  $\max \mathbf{A}_1(T)$  and greater than  $\min \mathbf{A}_1(T)$ . According to the first result of Proposition 12, we deduce that  $\alpha_1(T+1)$  is between two values of two elements of the vector  $\mathbf{A}_1(T)$ . Hence, combining the results of Proposition 11 and 12,  $\max \mathbf{A}_1(T+1) \leq \max \mathbf{A}_1(T)$  and  $\min \mathbf{A}_1(T+1) \geq \min \mathbf{A}_1(T)$ . Then  $\max \mathbf{A}_1(T)$  is decreasing with  $T$  and  $\min \mathbf{A}_1(T)$  is increasing with  $T$ . Given that for all  $T$ ,  $0 \leq \alpha_1(T) \leq \alpha$ , then  $\max \mathbf{A}_1(T)$  and  $\min \mathbf{A}_1(T)$  are bounded by 0 and  $\alpha$ . Therefore, we can conclude that  $\min \mathbf{A}_1(T)$  and  $\max \mathbf{A}_1(T)$  converge and we denote their limits by  $l_1$  and  $l_2$  respectively. Moreover  $\max \mathbf{A}_1(T)$  is lower bounded by  $l_2$  and  $\min \mathbf{A}_1(T)$  is upper bounded by  $l_1$ .  $\square$

However, in order to have  $\alpha_1(T)$  converges to a unique point, we need to establish that  $\max \mathbf{A}_1(T)$  and  $\min \mathbf{A}_1(T)$  converge to the same limit. In other words, we need to prove that  $l_1 = l_2$ . For that, we will use the second result of Proposition 12. To that extent, we proceed by contradiction, i.e. we suppose that  $l_1 \neq l_2$ . More specifically, given that  $l_1 \leq l_2$  by definition, the two possible cases satisfied by  $l_1$  and  $l_2$  are:  $l_1 < l_2$  or  $l_1 = l_2$ , then to show that  $l_1 = l_2$ , it is sufficient to find a contradiction considering  $l_1 < l_2$ .

In fact, we prove that if  $l_1 < l_2$ , there exists  $T_d$  such that all the elements of  $\mathbf{A}_1(T_d)$  are strictly less than  $l_2$ , that contradicts with the fact that  $\max \mathbf{A}_1(T)$  is lower bounded by  $l_2$ .

As  $\max \mathbf{A}_1(T)$  converges to  $l_2$ , then for a given  $\epsilon > 0$ , there exists a given time slot that we denote by  $T_\epsilon \geq T_0$  such that

for all  $T \geq T_\epsilon$ ,  $\max \mathbf{A}_1(T) < l_2 + \epsilon$ . Our proof consists of showing that for a small enough  $\epsilon$ , there exists  $T \geq T_\epsilon$ ,  $\max \mathbf{A}_1(T)$  is strictly less than  $l_2$ . We need first to determine an upper bound on the number of the elements of the vector  $\mathbf{A}_1(T)$  whatever  $T$ . In fact, as we have demonstrated that at each time  $T$ , the instantaneous threshold  $l(T)$  is less than  $l_{\max}$ . Then the number of the elements of  $\mathbf{A}_1(T)$  will not exceed  $l_{\max} + 1$ . In the following proof, we denote  $l_{\max}$  by  $L$ .

**Proposition 13.** *If  $l_1 < l_2$ , for  $\epsilon \leq (l_2 - l_1) \frac{(1-p_1)^L}{1-(1-p_1)^L}$ , there exist  $T_d \geq T_\epsilon$  such that all the elements of  $\mathbf{A}_1(T_d)$  are strictly less than  $l_2$ .*

*Proof.* See appendix J.  $\square$

Providing that  $l_2$  is a lower bound of  $\max \mathbf{A}_1(T)$  which contradicts with the result of the above proposition. Hence, the supposition of  $l_1 \neq l_2$  is not valid. Therefore,  $l_1 = l_2$ . Consequently,  $\max \mathbf{A}_1(T)$  and  $\min \mathbf{A}_1(T)$  converge to the same limit denoted  $\alpha_1^*$ . Given that  $\min \mathbf{A}_1(T) \leq \alpha_1(T) \leq \max \mathbf{A}_1(T)$  for all  $T$ , then  $\alpha_1(T)$  also converges to  $\alpha_1^*$ . Similarly,  $\alpha_2(T)$  converges to  $\alpha - \alpha_1^* = \alpha_2^*$ . In the following proposition, we prove that  $z(t)$  converges.

**Proposition 14.** *If  $\alpha_k(t)$  converges to  $\alpha_k^*$ , then for each state  $i$  and class  $k$ ,  $z_i^k(t)$  converges to  $z_i^{k,*}$ .*

*Proof.* See appendix K.  $\square$

However, we still have to establish that the stochastic vector  $\mathbf{Z}^N(t)$  converges to  $z^*$  in probability when  $N$  scales. For that, we introduce the following proposition inspired from the discrete-time version of Kurtz Theorem in [27]. Before that, knowing that the norms on the infinite dimension vector space are not equivalent, we work only with a specific norm which will be useful to show the optimality of the Whittle index's policy. Accordingly, we define  $\|\cdot\|$  as follows:

$$\|v\| = \sum_{i=1}^{+\infty} |v_i^1| i + \sum_{i=1}^{+\infty} |v_i^2| i \quad (29)$$

where  $v_i^k$  is the  $i$ -th component in the class  $k$  of the vector  $v$ . The reason behind chosen such a norm will be revealed in the proof of Proposition 16.

**Proposition 15.** *For any  $\mu > 0$  and finite time horizon  $T$ , there exists positive constant  $C$  such that*

$$P_x \left( \sup_{0 \leq t < T} \|\mathbf{Z}^N(t) - z(t)\| \geq \mu \right) \leq \frac{C}{N}$$

where  $P_x$  denotes the probability conditioned on the initial state  $\mathbf{Z}^N(0) = z(0) = x$ . Furthermore,  $C$  is independent of  $N$ .

*Proof.* See appendix L.  $\square$

One should note that a nearly similar proposition has been provided in [24] but for continuous time model. Specifically, the authors of [24] state that on average  $\mathbf{Z}^N(t)$  does not much differ from a  $z^*(t)$  when  $N$  scales. No proof of the claim was provided in [24]. In this paper, we consider a

time-slotted system and, as one can notice in the proposition above, we state that  $\mathbf{Z}^N(t)$  converges to  $z^*(t)$  when  $N$  scales and when  $t$  is upper bounded by a fixed constant  $T$  and we further give a detailed proof of our statement. According to this Proposition, the system state  $\mathbf{Z}^N(t)$  behaves very close to the fluid approximation model  $z(t)$  when the number of users  $N$  is large and starting from any initial state. To that extent, in order to establish the optimality of Whittle's index policy, we give first this following lemma which is a consequence of the Proposition 15.

**Lemma 3.** *For any  $\mu > 0$ , there exists a time  $T_0$  such that for each  $T > T_0$ , there exists a positive constant  $s$  with,*

$$P_x \left( \sup_{T_0 \leq t < T} \|\mathbf{Z}^N(t) - z^*\| \geq \mu \right) \leq \frac{s}{N}$$

*Proof.* See appendix M  $\square$

We remind that starting from an initial state  $x$ , our objective is to compare the total expected average age per user under the Whittle index policy which can be expressed as  $\frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} Z_i^{k,N}(t) i \mid \mathbf{Z}^N(0) = x \right]$  where  $\mathbf{Z}^N(t)$  evolves under the Whittle index policy, with the optimal age of the relaxed problem per user whose expression in function of  $z^*$  is,  $C^{RP} = \frac{C^{RP,N}}{N} = \sum_1^2 \sum_{i=1}^{+\infty} z_i^{k,*} i$ , when the number of users  $N$  as well as the time duration  $T$  grow.

According to Lemma 3, we are ready now to establish the asymptotic optimality of the Whittle index policy.

**Proposition 16.** *Starting from a given initial state  $\mathbf{Z}^N(0) = z(0) = x$ , then:*

$$\begin{aligned} \lim_{T \rightarrow +\infty} \lim_{N \rightarrow \infty} \frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} Z_i^{k,N}(t) i \mid \mathbf{Z}^N(0) = x \right] \\ = \sum_{k=1}^2 \sum_{i=1}^{+\infty} z_i^{k,*} i \end{aligned} \quad (30)$$

*Proof.* See appendix N.  $\square$

## V. NUMERICAL RESULTS

### A. Verification of assumption 3

In this section, we compute the value of the lower bound on  $\alpha$  given in Assumption 3. We denote this lowerbound by  $B_\alpha$ . For a wide range of parameters  $p_1$  and  $p_2$ , we provide an exhaustive table that represents the lower bound on  $\alpha$  in function of  $p_1$  and  $p_2$ . As can be seen, the lowerbound decreases when  $p_1$  and  $p_2$  are close one to the other. Moreover, it grows even smaller when  $p_1$  and  $p_2$  have relatively high values. According to table I, we can notice that in most cases of  $(p_1, p_2)$ , the lower bound of  $\alpha$  doesn't exceed 0.5. This implies that the interval of  $\alpha$  where the assumption 3 is satisfied, is enough wide for different values of  $p_1$  and  $p_2$ .

### B. Implementation of the Whittle's index policy

In this section, we evaluate the Whittle's index policy's performance by comparing the per-user average age of the Whittle's index policy to the optimal per-user average age of the relaxed problem  $C^{rp}$ . To that extent, we let the number of

$p_2$	$p_1$	$B_\alpha$
0.1	0.2	0.7034
0.2	0.4	0.6250
0.3	0.5	0.4711
0.4	0.6	0.3556
0.4	0.8	0.5328
0.5	0.8	0.3612
0.5	1	0.5
0.6	0.9	0.2893
0.7	0.9	0.1675
0.8	0.9	0.1351

Table I: Evaluation of  $B_\alpha$  for a wide range of channel statics

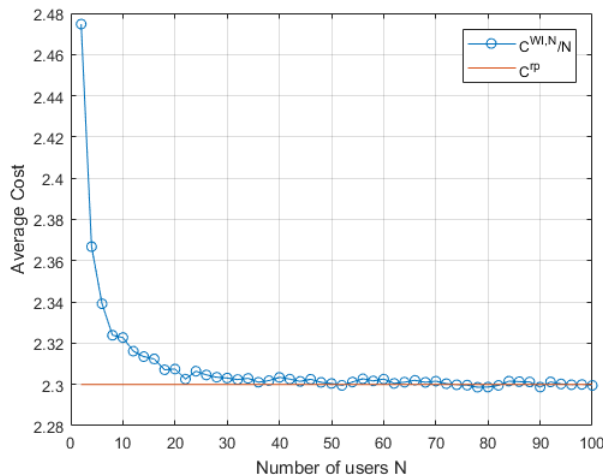


Figure 5: Average age per-user under the Whittle's index policy

users in class 1 and class 2 be equal to  $\frac{N}{2}$ . The probability of successful transmission of class 1 and class 2 are set to 0.8 and 0.5, respectively. At each time slot  $t$ , at most,  $M = \frac{N}{2}$  of users can be scheduled per each time slot, i.e.,  $\alpha = \frac{M}{N} = \frac{1}{2}$ . As seen in Figure 5, the gap between the two policies tightens as the number of users  $N$  grows. Indeed, these numerical results corroborate our theoretical analysis and show that the Whittle's index policy is effectively globally asymptotically optimal.

## VI. CONCLUSION

In this paper, we have examined the average age minimization problem where only a fraction of the network users can transmit simultaneously over unreliable channels. We presented and derived a novel method based Cauchy criterion to prove the optimality of the Whittle's index policy in the many-users regime. Contrary to the state of the art methods, in which strong assumptions that cannot be verified easily are used, our approach does not require imposing any restrictive mathematical assumptions. We also provided numerical results that corroborate our theoretical findings and highlight the performance of the Whittle's index policy. Moving forward, the next research direction is to extend our proof to various other scheduling problems under different system models and objective functions.

## REFERENCES

- [1] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *2012 Proceedings IEEE INFOCOM*, March 2012, pp. 2731–2735.
- [2] B. Buyukates and S. Ulukus, "Timely distributed computation with stragglers. october 2019," *Available on*.
- [3] S. Farazi, A. G. Klein, J. A. McNeill, and D. R. Brown, "On the age of information in multi-source multi-hop wireless status update networks," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2018, pp. 1–5.
- [4] A. Maatouk, M. Assaad, and A. Ephremides, "The age of updates in a simple relay network," in *2018 IEEE Information Theory Workshop (ITW)*. IEEE, 2018, pp. 1–5.
- [5] P. Zou, O. Ozel, and S. Subramaniam, "Waiting before serving: A companion to packet management in status update systems," *IEEE Transactions on Information Theory*, 2019.
- [6] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksall, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.
- [7] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2017, pp. 486–493.
- [8] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Age-optimal information updates in multihop networks," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 576–580.
- [9] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age of information performance of multiaccess strategies with packet management," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 244–255, 2019.
- [10] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, "Age-optimal updates of multiple information flows," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2018, pp. 136–141.
- [11] E. T. Ceran, D. Gündüz, and A. György, "Average age of information with hybrid arq under a resource constraint," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1900–1913, 2019.
- [12] A. Maatouk, M. Assaad, and A. Ephremides, "On the age of information in a csma environment," *IEEE/ACM Transactions on Networking*, vol. 28, no. 2, pp. 818–831, 2020.
- [13] Z. Jiang, B. Krishnamachari, X. Zheng, S. Zhou, and Z. Niu, "Timely status update in massive iot systems: Decentralized scheduling for wireless uplinks," *arXiv preprint arXiv:1801.03975*, 2018.
- [14] R. Talak, S. Karaman, and E. Modiano, "Distributed scheduling algorithms for optimizing information freshness in wireless networks," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2018, pp. 1–5.
- [15] 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 Transactions on Networking*, vol. 26, no. 6, pp. 2637–2650, 2018.
- [16] P. Ansell, K. D. Glazebrook, J. Niño-Mora, and M. O'Keefe, "Whittle's index policy for a multi-class queueing system with convex holding costs," *Mathematical Methods of Operations Research*, vol. 57, no. 1, pp. 21–39, 2003.
- [17] S. Kriouile, M. Larranaga, and M. Assaad, "Asymptotically optimal delay-aware scheduling in wireless networks," *arXiv preprint arXiv:1807.00352*, 2018.
- [18] —, "Whittle index policy for multichannel scheduling in queueing systems," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019, pp. 2524–2528.
- [19] M. Larrañaga, M. Assaad, A. Destounis, and G. S. Paschos, "Asymptotically optimal pilot allocation over markovian fading channels," *IEEE Transactions on Information Theory*, 2017.
- [20] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access," *IEEE*

*Transactions on Information Theory*, vol. 56, no. 11, pp. 5547–5567, 2010.

- [21] A. Maatouk, S. Kriouile, M. Assad, and A. Ephremides, “On the optimality of the Whittle’s index policy for minimizing the age of information,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1263–1277, 2020.
- [22] W. Ouyang, A. Eryilmaz, and N. B. Shroff, “Downlink scheduling over Markovian fading channels,” *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1801–1812, 2015.
- [23] K. P. Papadaki and W. B. Powell, “Exploiting structure in adaptive dynamic programming algorithms for a stochastic batch service problem,” *European Journal of Operational Research*, vol. 142, no. 1, pp. 108–127, 2002.
- [24] R. R. Weber and G. Weiss, “On an index policy for restless bandits,” *Journal of Applied Probability*, vol. 27, no. 3, pp. 637–648, 1990.
- [25] P. Whittle, “Restless bandits: Activity allocation in a changing world,” *Journal of applied probability*, vol. 25, no. A, pp. 287–298, 1988.
- [26] C. H. Papadimitriou and J. N. Tsitsiklis, “The complexity of optimal queueing network control,” *Mathematics of Operations Research*, vol. 24, no. 2, pp. 293–305, 1999.
- [27] T. G. Kurtz, “Strong approximation theorems for density dependent Markov chains,” *Stochastic Processes and their Applications*, vol. 6, no. 3, pp. 223–240, 1978.

**Saad Kriouile** Saad Kriouile received his M.Sc and PhD degrees both in Telecommunications from CentraleSupélec, Gif-sur-Yvette, France, in 2017 and 2021, respectively. His PhD thesis was on the design of asymptotically optimal scheduling in large networks and has been supported by the TCL 5G chair at CentraleSupélec. He is currently a postdoctoral researcher at the Laboratory of Signals and Systems (L2S), CentraleSupélec and University of Paris-Saclay. His research interest includes Age of Information, stochastic resource optimization and design of optimal scheduling in wireless and queueing networks.

**Mohamad Assaad** Mohamad Assaad received the M.Sc and PhD degrees, both in telecommunications, from Telecom ParisTech, Paris, France, in 2002 and 2006, respectively. Since 2006, he has been with the Telecommunications Department at CentraleSupélec, where he is currently a professor. He is also a researcher at the Laboratoire des Signaux et Systèmes (L2S, CNRS) and holds the 5G Chair. He has co-authored 1 book and more than 120 publications in journals and conference proceedings and has served as TPC member or TPC co-chair for top-tier international conferences, including TPC co-chair for IEEE WCNC’21, IEEE Globecom’20 Mobile and Wireless Networks Symposium Co-chair, etc. He is an Editor for the IEEE Wireless Communications Letters and the Journal of Communications and Information Networks. He served also as a guest co-editor for a special issue of the IEEE transactions on Network Science and Engineering. He has given in the past successful tutorials on several topics related to 5G systems, and Age of Information at various conferences including IEEE ISWCS’15, IEEE WCNC’16, and IEEE ICC’21 conferences. His research interests include 5G and beyond systems, fundamental networking aspects of wireless systems, Age of Information, resource optimization and Machine Learning in wireless networks.

**Ali Maatouk** Ali Maatouk received his M.Sc. and Ph.D. degrees both from CentraleSupélec, Gif-sur-yvette, France in 2017 and 2020 respectively. He is the recipient of the Labex Digicosme scholarship for academic excellence in 2016. He is also the recipient of both the regional and national André Blanc-Lapierre awards for best master’s student in France in the domain of electricity, information technology, and communication. For his Ph.D., he received the Paris-Saclay Ph.D. award for one of the best scientific contributions. Between 2018 and 2020, he spent some time at the University of Maryland, College Park, as a visiting PhD student. He is currently a researcher at Huawei Technologies in Boulogne-Billancourt, France. His research interests lie in the optimization of wireless systems, 5G and beyond systems, and the notion of age of information.

## APPENDIX A PROOF OF LEMMA 1

At time  $t + 1$ , applying Whittle index policy, on average exactly a proportion of  $p_k \alpha_k(t)$  of users will be at state one since  $\alpha_k(t)$  refers to the proportion of users in class  $k$  that are scheduled. Accordingly,  $z_1^k(t + 1) = p_k \alpha_k(t)$ . While for  $1 \leq i < l_k(t)$ , the users’ proportion  $z_i^k(t)$  is not scheduled. Therefore at time  $t + 1$ , since prescribing idle action to a given user implies that its state will be increased by 1, the proportion  $z_i^k(t)$  at state  $i$  in class  $k$  will be at state  $i + 1$ . Thus,  $\mathbb{E} \left[ Z_{i+1}^{N,k}(t + 1) \mid Z^N(t) = \mathbf{z}(t) \right] = z_{i+1}^k(t + 1) = z_i^k(t)$ .

## APPENDIX B PROOF OF LEMMA 2

First of all, we provide an useful lemma.

**Lemma 4.** *We have for all integer  $i$  and for  $k = 1, 2$ :*

$$w_k(i + 1) - w_k(i) = ip_k + 1$$

*Proof.* The result can be obtained directly by replacing  $w_k(i)$  by its expression. ■

In order to prove the present lemma, we proceed in two steps:

- We prove first by contradiction that there exists a given time  $t_f$  such that  $\alpha_1(t_f) > 0$ .
- We prove that if  $\alpha_1(t_f) > 0$ , then  $\alpha_1(t) > 0$  for all  $t \geq t_f$ .

1) For the first point, we suppose that for all  $t$ , we have that  $\alpha_1(t) = 0$ . Consequently, we get that  $z_1^1(t + T_t) = 0, \dots, z_{l_1(t+T_t)}^1(t + T_t) = 0$ , and  $\alpha_1(t + T_t) = 0$ . This means that, the proportion of all users in class 1 is equal to 0. However, the users’ proportion of class 1 is  $\gamma_1 \neq 0$ . That is, there exists a given time  $t_f$  such  $\alpha_1(t_f) > 0$ .

2) Before addressing the second point, we recall that  $\alpha_1(t)$  refers to the scheduled users’ proportion in the class 1. Thereby,  $\alpha_1(t)$  contains all users with the highest Whittle index values among all users in class 1. To that extent, at time  $t_f$ , the Whittle index of  $\alpha_1(t_f)$  is greater than the Whittle index of the users’ proportion  $1 - \alpha$  that we denote by  $C$ . We let  $\mathcal{S}_{t_f}(C)$  be the set of pair (state,class) at time  $t_f$  in the users’ proportion  $C$ . Denoting by  $q$  the smallest state of  $\alpha_1(t_f)$ ,  $n$  and  $m$  a given state and class respectively such that  $z_n^m(t)$  belongs to  $C$  at time  $t_f$ , then  $w_m(n) \leq w_1(q)$ . Under the Whittle index policy, at time  $t_f + 1$ , the states of a users’ proportion that equals to  $(1 - p_1)\alpha_1(t_f)$  among the users’ proportion  $\alpha_1(t_f)$ , will be increased by one in comparison with the time slot  $t_f$ , as well as the users’ proportion  $C$ . Accordingly, the smallest state of the proportion  $(1 - p_1)\alpha_1(t_f)$ , is  $q + 1$ .  $\mathcal{S}_{t_f+1}(C)$  is shifted of one with respect to  $\mathcal{S}_{t_f}(C)$ , i.e.,  $(n, m) \in \mathcal{S}_{t_f}(C) \Leftrightarrow (n + 1, m) \in \mathcal{S}_{t_f+1}(C)$ . We compare  $w_1(q + 1)$  with the Whittle index of  $n$  in class  $m$  such that  $(n, m) \in \mathcal{S}_{t_f+1}(C)$ . In that direction, we

let  $(n, m) \in \mathcal{S}_{t_f+1}(C)$ , and we distinguish between two cases:

- $m = 1$ : Leveraging the fact that  $(n-1, m) \in \mathcal{S}_{t_f}(C)$ , then  $w_1(q) \geq w_1(n-1)$ . That implies that  $n-1 \leq q$  since  $w_k(\cdot)$  is increasing. Hence  $n \leq q+1$ . As consequence,  $w_1(n) \leq w_1(q+1)$
- $m = 2$ : Again we distinguish between two case:
  - If  $n-1 \leq q$ , then  $w_2(n) < w_1(n) \leq w_1(q+1)$ . Therefore, we obtain our desired result for the first case.
  - If  $n-1 > q$ :  
We have that:

$$\begin{aligned} w_1(q+1) - w_2(n) &= (w_1(q+1) - w_1(q)) \\ &\quad - (w_2(n) - w_2(n-1)) \\ &\quad + w_1(q) - w_2(n-1) \end{aligned} \quad (31)$$

Applying Lemma 4, we obtain:  $(w_1(q+1) - w_1(q)) - (w_2(n) - w_2(n-1)) = qp_1 - (n-1)p_2$ . Given that  $w_2(n-1) \leq w_1(q)$ , therefore replacing by their expressions we get:

$$(n-2)(n-1)p_2/2 + n-1 \leq (q-1)qp_1/2 + q$$

As  $n-1 > q$ , then:

$$(n-2)(n-1)p_2/2 \leq (q-1)qp_1/2$$

Hence:

$$(n-1)p_2 \leq qp_1$$

Therefore,  $(w_1(q+1) - w_1(q)) - (w_2(n) - w_2(n-1)) \geq 0$ . Hence, knowing that  $w_1(q) - w_2(n-1) \geq 0$ , we end up with our desired result for this case, i.e.  $w_1(q+1) - w_2(n) \geq 0$ .

Thus, we have proved that at time  $t_f + 1$ , all the users' proportions in  $C$  whose sum is equal to  $1 - \alpha$  have a Whittle index less than that of  $(1-p_1)\alpha_1(t_f)$  defined in the beginning of this proof. That means that there exists at least a users' proportion that equals to  $1 - \alpha$  with Whittle index values less than those of the states of the users' proportion  $(1-p_1)\alpha_1(t_f)$ . Then surely, the users' proportion  $(1-p_1)\alpha_1(t_f)$  that is different from 0 belongs to the users' proportion  $\alpha$  with the highest Whittle index values. This implies that surely at time  $t_f + 1$ , there will be at least one queue in class 1 belonging to  $\alpha$  with the highest Whittle index values. Therefore, we have that  $\alpha_1(t_f + 1) > 0$ . This result can be generalized for all  $t \geq t_f$ . In other words, we have for all  $t \geq t_f$ ,  $\alpha_1(t) > 0$ .

#### APPENDIX C

##### PROOF OF PROPOSITION 7

We have that  $w_1(n) = \frac{(n-1)p_1n}{2} + n$ , and  $w_2(n) = \frac{(n-1)p_2n}{2} + n$ . We start first by finding the set of states for which the Whittle index alternate between the two classes. As we can see from the expression of the Whittle index, for a

given state  $n$ ,  $w_2(n) < w_1(n)$  as  $p_2 < p_1$ . In order to have the condition of alternation strictly satisfied for any given state  $n$ , we must have  $w_1(n) < w_2(n+1)$ . Hence, denoting by  $f(n)$  the difference  $w_2(n+1) - w_1(n)$ , we study the sign of  $f(n)$  to see for which  $n$   $f$  is strictly positive.

**Lemma 5.** For all  $n \in [1, D[$ ,  $f(n) > 0$

*Proof.* We have that:

$$f(n) = \frac{n^2}{2}(p_2 - p_1) + \frac{n}{2}(p_1 + p_2) + 1 \quad (32)$$

Hence:

$$f'(n) = n(p_2 - p_1) + \frac{p_1 + p_2}{2} \quad (33)$$

The derivative is equal to zero for  $n = \frac{p_1 + p_2}{2(p_1 - p_2)}$ , which is greater strictly than 0. This means that  $f$  is strictly increasing in  $[0, \frac{p_1 + p_2}{2(p_1 - p_2)}]$  since  $f'(n) > 0$  in  $[0, \frac{p_1 + p_2}{2(p_1 - p_2)}[$ . Providing that  $f(0) = 1$ , then surely  $f$  is strictly positive in  $[0, \frac{p_1 + p_2}{2(p_1 - p_2)}]$ . This means that, the unique positive solution for  $f(n) = 0$  must be in the interval  $[\frac{p_1 + p_2}{2(p_1 - p_2)}, +\infty[$ , as  $\lim_{n \rightarrow +\infty} f(n) = -\infty$ . Indeed, the unique solution  $n_0$  of  $f(n) = 0$  in  $[\frac{p_1 + p_2}{2(p_1 - p_2)}, +\infty[$  is the biggest root of the polynomial (32) which is exactly the value  $D$  introduced in Assumption 3. As the function  $f$  is decreasing in  $[\frac{p_1 + p_2}{2(p_1 - p_2)}, +\infty[$ , then  $f$  is strictly positive in  $[0, D[$ . Therefore,  $f(n) > 0$  for  $n \in [1, D[$ , which concludes the proof. ■

According to Lemma 5, the order of the Whittle index strictly alternates between the two states when  $n \in [1, D[$ . Therefore, we need to prove that  $T_{\max} + 1$  is upper bounded by  $D$  in order to prove that the alternation condition is satisfied from state 1 to  $T_{\max} + 1$ .

Indeed, as we have  $T_{\max} \leq \frac{1-\alpha}{p_2\alpha} + 1$ , we just need to prove that  $\frac{1-\alpha}{p_2\alpha} + 2$  is strictly less than  $D$ .

Under Assumption (3), we have that:

$$\alpha > \frac{1}{1 + (D-2)p_2} \quad (34)$$

$$\alpha(1 + p_2(D-2)) > 1 \quad (35)$$

$$\alpha p_2(D-2) > 1 - \alpha \quad (36)$$

$$D-2 > \frac{1-\alpha}{p_2\alpha} \quad (37)$$

$$D > \frac{1-\alpha}{p_2\alpha} + 2 \quad (38)$$

Hence, from state 1 to  $T_{\max} + 1$ , the order of the Whittle index strictly alternates between the two classes. Accordingly, the proof is concluded.

#### APPENDIX D

##### PROOF OF PROPOSITION 8

We present first a lemma which will be helpful in proving this proposition as well as the next ones.

**Lemma 6.** For any state  $q$ , at any time  $t$ , we have that:

$$w_1(q) \leq w_2(l_2(t)) \Rightarrow w_1(q) \leq w_1(l_1(t))$$

and

$$w_2(q) \leq w_1(l_1(t)) \Rightarrow w_2(q) \leq w_2(l_2(t))$$



*Proof.* See appendix E

We consider  $t \geq t_f$ . After time  $T_t$ , we have that:

$$\sum_{j=1}^{l(t+T_t)} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{T_t} p_2 \alpha_2(t+T_t-j) \geq 1 - \alpha \quad (39)$$

Then, as it has been showcased, at time  $t + T_t$ , there exists  $l_1(t + T_t) \leq l(t + T_t)$ ,  $l_2(t + T_t) \leq T_t$ ,  $\gamma(t + T_t) = 1$  and  $0 < \beta(t + T_t) \leq 1$ ; or  $0 < \gamma(t + T_t) \leq 1$  and  $\beta(t + T_t) = 1$  such that:

$$\begin{aligned} & \sum_{j=1}^{l_1(t+T_t)-1} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{l_2(t+T_t)-1} p_2 \alpha_2(t+T_t-j) \\ & + \beta(t+T_t) p_1 \alpha_1(t+T_t-l_1(t+T_t)) \\ & + \gamma(t+T_t) p_2 \alpha_2(t+T_t-l_2(t+T_t)) \\ & = 1 - \alpha \end{aligned} \quad (40)$$

with  $l_1(t+T_t)$  and  $l_2(t+T_t)$  being the instantaneous thresholds in class 1 and 2 respectively at time  $t + T_t$ .

Now, we prove by contradiction that  $\max(l_1(t + T_t), l_2(t + T_t)) \leq T_{max}$ .

We prove first that  $l_2(t + T_t)$  is greater than  $l_1(t + T_t)$ .

As we have that  $w_2(l_1(t + T_t)) < w_1(l_1(t + T_t))$ , then according to lemma 6,  $w_2(l_1(t + T_t)) \leq w_2(l_2(t + T_t))$ . This implies that  $l_2(t + T_t)$  is greater than  $l_1(t + T_t)$ .

Reasoning by contradiction, we suppose that  $l_2(t+T_t) > T_{max}$  ( $l_2(t + T_t) = \max(l_1(t + T_t), l_2(t + T_t)) > T_{max}$ ). Based on this, we have that  $w_1(T_{max}) < w_2(l_2(t + T_t))$  because  $w_1(T_{max}) < w_2(T_{max} + 1) \leq w_2(l_2(t + T_t))$  since the order of the Whittle index alternates between the two classes as it has been proved in Proposition 8. To that extent, we distinguish between two cases:

1) **First case:** If  $\beta(t + T_t) = 1$ :

We have that  $w_1(T_{max}) < w_2(l_2(t + T_t))$ . Then, according to Lemma 6, we have that  $w_1(T_{max}) \leq w_1(l_1(t + T_t))$ . Hence, we can conclude that  $T_{max} \leq l_1(t + T_t)$  as  $w_1$  is an increasing function with the age of information.

Moreover, since we have that  $p_1 \alpha_1(t + T_t - j) + p_2 \alpha_2(t + T_t - j) > p_2 \alpha$  (the strict inequality is due to the fact that  $\alpha_1(t) > 0$  as  $t \geq t_f$  according to Lemma 2), then we obtain:

$$\begin{aligned} & \sum_{j=1}^{l_1(t+T_t)-1} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{l_2(t+T_t)-1} p_2 \alpha_2(t+T_t-j) \\ & + \beta(t+T_t) p_1 \alpha_1(t+T_t-l_1(t+T_t)) \\ & + \gamma(t+T_t) p_2 \alpha_2(t+T_t-l_2(t+T_t)) \\ & = 1 - \alpha \\ & = \sum_{j=1}^{l_1(t+T_t)} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{l_2(t+T_t)-1} p_2 \alpha_2(t+T_t-j) \\ & + \gamma(t+T_t) p_2 \alpha_2(t+T_t-l_2(t+T_t)) \\ & \geq \sum_{j=1}^{T_{max}} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{T_{max}} p_2 \alpha_2(t+T_t-j) \\ & > T_{max} p_2 \alpha \geq 1 - \alpha \end{aligned} \quad (41)$$

■ The last inequality comes from the fact that  $T_{max} \geq \frac{1-\alpha}{p_2 \alpha}$ . This implies that:

$$1 - \alpha > 1 - \alpha \quad (42)$$

This gives us an illogical statement. Consequently, in this case, the assumption  $l_2(t + T_t) > T_{max}$  is not true.

2) **Second case:** If  $\beta(t + T_t) < 1$ :

As we have that  $\beta(t + T_t) < 1$ , then  $\gamma(t + T_t)$  should be equal to 1. Therefore, all users at state  $l_2(t + T_t)$  in class 2 are in the users' proportion  $1 - \alpha$  with the smallest Whittle index values. However, there exists users in state  $l_1(t + T_t)$  in class 1 in the users' proportion  $\alpha$  that has the highest Whittle index values. That is, we have  $w_1(l_1(t + T_t)) \geq w_2(l_2(t + T_t))$ . As it has been established before tackling the first case,  $w_1(T_{max}) < w_2(l_2(t + T_t))$ , then  $w_1(T_{max}) < w_1(l_1(t + T_t))$ . This means that  $l_1(t + T_t) > T_{max}$ . Therefore, we have that:

$$\begin{aligned} & \sum_{j=1}^{l_1(t+T_t)-1} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{l_2(t+T_t)-1} p_2 \alpha_2(t+T_t-j) \\ & + \beta(t+T_t) p_1 \alpha_1(t+T_t-l_1(t+T_t)) \\ & + \gamma(t+T_t) p_2 \alpha_2(t+T_t-l_2(t+T_t)) \\ & = 1 - \alpha \\ & \geq \sum_{j=1}^{T_{max}} p_1 \alpha_1(t+T_t-j) + \sum_{j=1}^{T_{max}} p_2 \alpha_2(t+T_t-j) \\ & > T_{max} p_2 \alpha \geq 1 - \alpha \end{aligned} \quad (43)$$

This implies that:

$$1 - \alpha > 1 - \alpha \quad (44)$$

Consequently, in this case, the assumption  $l_2(t + T_t) > T_{max}$  is not true.

Hence, in both cases,  $l_2(t + T_t)$  must be less than  $T_{max}$ , i.e.  $\max(l_1(t + T_t), l_2(t + T_t)) \leq T_{max}$  for all  $t$ .

Thus, we end up with  $T_{max} = l_{max}$ , which concludes our proof.

## APPENDIX E PROOF OF LEMMA 6

We prove only the first statement as the proof steps for both cases are exactly the same. By definition of  $l_1(t)$  and  $l_2(t)$ , we have that  $\{z_i^1(t)\}_{1 \leq i \leq l_1(t)} \cup \{z_i^2(t)\}_{1 \leq i \leq l_2(t)}$  is exactly the set  $\{z_i^k(t) : w_k(i) \leq \max(w_1(l_1(t)), w_2(l_1(t)))\}$ . Hence, if a given  $q$  verifies  $w_1(q) \leq w_2(l_2(t))$ , then  $w_1(q) \leq \max(w_1(l_1(t)), w_2(l_2(t)))$ , that implies that  $z_q^1(t) \in \{z_i^k(t) : w_k(i) \leq \max(w_1(l_1(t)), w_2(l_2(t)))\} = \{z_i^1(t)\}_{1 \leq i \leq l_1(t)} \cup \{z_i^2(t)\}_{1 \leq i \leq l_2(t)}$ . Knowing that the highest users' proportion's state of the aforementioned set in class 1 is  $l_1(t)$ , then  $q \leq l_1(t)$ . Therefore as  $w_1(\cdot)$  is increasing,  $w_1(q) \leq w_1(l_1(t))$ .

## APPENDIX F PROOF OF PROPOSITION 9

We have that:

$$\sum_{i=1}^{l_1(t)-1} z_i^1(t) + \sum_{i=1}^{l_2(t)-1} z_i^2(t) + \beta(t) z_{l_1(t)}^1(t) + \gamma(t) z_{l_2(t)}^2(t) = 1 - \alpha \quad (45)$$

with  $l_1(t)$  and  $l_2(t)$  being the thresholds in class 1 and 2 respectively at time  $t$ , and  $\beta(t) = 1$  and  $0 < \gamma(t) \leq 1$ , or  $\gamma(t) = 1$  and  $0 < \beta(t) \leq 1$ .

Our aim in this proof is to show that there is a link between  $l_1(t)$  and  $l_2(t)$  when they are less than  $T_{\max}$ . By doing so, we find a general form of the aforementioned equation. To that end, we prove first that  $l_1(t)$  is less than  $l_2(t)$ .

Indeed, as we have  $w_2(l_1(t)) < w_1(l_1(t))$ , then according to lemma 6,  $w_2(l_1(t)) \leq w_2(l_2(t))$ . Consequently, we can conclude that  $l_1(t) \leq l_2(t)$ .

Secondly, we prove that  $l_2(t) \leq l_1(t) + 1$ . As the order of the Whittle indices alternates between the two classes from state 1 to state  $T_{\max} + 1$ ,  $w_1(l_2(t) - 1) < w_2(l_2(t))$ . Hence, according to lemma 6, we have that  $w_1(l_2(t) - 1) \leq w_1(l_1(t))$ . Consequently,  $l_2(t) - 1 \leq l_1(t)$ .

Given that  $l_1(t) \leq l_2(t) \leq l_1(t) + 1$ , then  $l_1(t)$  can be either  $l_2(t)$  or  $l_2(t) - 1$ .

The second step consists of deriving the value of  $\beta(t)$  or  $\gamma(t)$  depending on the value of  $l_1(t)$  and  $l_2(t)$ .

- If  $l_1(t) = l_2(t)$ :

We prove that  $\gamma(t) = 1$  if  $z_{l_2(t)}^2 > 0$ . Indeed, if  $\gamma(t) \neq 1$  and  $z_{l_2(t)}^2 > 0$ , thus there is at least a non empty set of users in class 2 at state  $l_2(t)$  that belongs to the users' proportion  $\alpha$  with the highest Whittle index values. However there exists always a non empty set of users in class 1 at state  $l_1(t)$  that belong to  $1 - \alpha$  users' proportion with the least Whittle index values, since  $\beta(t) > 0$ . Then, we have that  $w_2(l_2(t)) \geq w_1(l_1(t))$ . However, we know that  $w_2(l_2(t)) = w_2(l_1(t)) < w_1(l_1(t))$ . This later inequality contradicts with what precedes. Thus, the statement that  $\gamma(t) \neq 1$  is not true, i.e.  $\gamma(t) = 1$ .

In this case we denote  $l(t) = l_1(t) = l_2(t)$ .

We end up:

$$\sum_{j=1}^{l(t)-1} z_j^1(t) + \sum_{i=1}^{l(t)-1} z_i^2(t) + \beta(t)z_{l(t)}^1(t) + z_{l(t)}^2(t) = 1 - \alpha \quad (46)$$

If  $z_{l_2(t)}^2 = 0$ , the last equation still valid since  $z_{l_2(t)}^2 = 0$  whatever the value of  $\gamma(t)$ , namely when  $\gamma(t) = 1$ .

- If  $l_1(t) + 1 = l_2(t)$ :

We prove that  $\beta(t) = 1$  if  $z_{l_1(t)}^1 > 0$ . Indeed, if  $\beta(t) \neq 1$  and  $z_{l_1(t)}^1 > 0$ , there is at least a set of users in class 1 in state  $l_1(t)$  that belongs to the users' proportion  $\alpha$  with the highest Whittle index values. However there is always a set of users in class 2 at state  $l_2(t)$  that belong to  $1 - \alpha$  users' proportion with the least Whittle index values, since  $\gamma(t) > 0$ . Then, we have that  $w_1(l_1(t)) \geq w_2(l_2(t))$ . However, we know that  $w_2(l_2(t)) = w_2(l_1(t) + 1) > w_1(l_1(t))$  since the order of Whittle index alternates between the two classes from state 1 to  $T_{\max} + 1$  according to Proposition 8. Thus,  $w_2(l_1(t) + 1) > w_1(l_1(t)) \geq w_2(l_1(t) + 1)$ , which gives us an obvious contradiction. Therefore, we can assert that  $\beta(t) = 1$ .

In this case, we consider that  $l(t) = l_1(t) + 1 = l_2(t)$  and

we get:

$$\sum_{i=1}^{l(t)-1} z_i^1(t) + \sum_{i=1}^{l(t)-1} z_i^2(t) + \gamma(t)z_{l(t)}^2(t) = 1 - \alpha \quad (47)$$

Similarly to the first case, if  $z_{l_1(t)}^1 = 0$ , the last equation still valid since  $z_{l_1(t)}^1 = 0$  whatever the value of  $\beta(t)$ , namely when  $\beta(t) = 1$ . Subsequently, combining the two cases, there exists  $l(t)$  such that:

$$\sum_{i=1}^{l(t)-1} z_i^1(t) + \sum_{i=1}^{l(t)-1} z_i^2(t) + \beta(t)z_{l(t)}^1(t) + \gamma(t)z_{l(t)}^2(t) = 1 - \alpha \quad (48)$$

where  $\beta(t) = 0$  and  $0 < \gamma(t) \leq 1$ , or  $0 < \beta(t) \leq 1$  and  $\gamma(t) = 1$ .

## APPENDIX G PROOF OF PROPOSITION 10

We prove the Proposition by induction:

- For  $T = T_0$ , we have already proved our claim.
- We suppose that the statement is valid for a given  $T$ , i.e. there exists  $l(T)$ ,  $\beta(T)$  and  $\gamma(T)$  such that:

$$\sum_{j=1}^{l(T)-1} p_1\alpha_1(T - j) + \sum_{j=1}^{l(T)-1} p_2\alpha_2(T - j) + \beta(T)p_1\alpha_1(T - l(T)) + \gamma(T)p_2\alpha_2(T - l(T)) = 1 - \alpha \quad (49)$$

where  $\beta(T) = 0$  and  $0 < \gamma(T) \leq 1$ , or  $0 < \beta(T) \leq 1$  and  $\gamma(T) = 1$ . Then, at the next time slot, among the users' proportion scheduled,  $\alpha$ , exactly  $p_1\alpha_1(T)$  and  $p_2\alpha_2(T)$  will go to state one for each class, while for the rest, their states will be incremented by one. Likewise, for the other users for which the action taken is passive, their states will be incremented by one. As consequence, the decreasing order according to the Whittle index value for these proportions of users at the next slot is  $\beta(T)p_1\alpha_1(T - l(T))$ ,  $\gamma(T)p_2\alpha_2(T - l(T))$ ,  $p_1\alpha_1(T - l(T) + 1)$ ,  $p_2\alpha_2(T - l(T) + 1)$ ,  $p_1\alpha_1(T - l(T) + 2)$ ,  $p_2\alpha_2(T - l(T) + 2)$ ,  $p_1\alpha_1(T - l(T) + 3)$ ,  $p_2\alpha_2(T - l(T) + 3)$ ,  $\dots$ ,  $p_1\alpha_1(T)$ ,  $p_2\alpha_2(T)$  (As we have mentioned before, the order of the Whittle indices alternates between the two classes because  $l(T) + 1 \leq l_{\max} + 1$ ). Moreover, the states of the users' proportion  $(1 - p_1)\alpha_1(t)$  and  $(1 - p_2)\alpha_2(t)$ ; which are scheduled but they don't transit to the state 1 with respect to their classes; will be increased by one. Leveraging the above results, we provide the decreasing order of all users' proportions according to the Whittle index value depending on two cases of  $\beta(t)$ .

If  $\beta(T) = 0$ , then the smallest state's value among the users' proportions  $(1 - p_1)\alpha_1(t)$  and  $(1 - p_2)\alpha_2(t)$  at time  $T + 1$  is  $l(T) + 1$ . Hence, their Whittle index values will be higher than  $w_2(l(T) + 1)$ , and consequently, they will be higher than those of users' proportion of  $\gamma(T)p_2\alpha_2(T - l(T))$  at state  $l(T) + 1$  in class 2.

If  $\beta(T) \neq 1$ , the smallest state value among the users'

proportions  $(1 - p_1)\alpha_1(t)$  and  $(1 - p_2)\alpha_2(t)$  at time  $T + 1$  is respectively  $l(T) + 1$  and  $l(T) + 2$ . Then, their Whittle index values will be higher than  $w_1(l(T) + 1)$  ( $w_1(l(T) + 1) < w_2(l(T) + 2)$ ) as the alternation condition is satisfied from 1 until  $l_{\max} + 1$ ). Consequently, their Whittle index values will be higher than the Whittle index of users' proportion  $\beta(T)p_1\alpha_1(T - l(T))$  at state  $l(T) + 1$  in class 1.

Thus, the decreasing order of all users' proportions according to the Whittle index value whatever the value of  $\beta(T)$  at  $T + 1$  is:  $(1 - p_1)\alpha_1(t), (1 - p_2)\alpha_2(t), \beta(T)p_1\alpha_1(T - l(T)), \gamma(T)p_2\alpha_2(T - l(T)), p_1\alpha_1(T - l(T) + 1), p_2\alpha_2(T - l(T) + 1), p_1\alpha_1(T - l(T) + 2), p_2\alpha_2(T - l(T) + 2), p_1\alpha_1(T - l(T) + 3), p_2\alpha_2(T - l(T) + 3), \dots, p_1\alpha_1(T), p_2\alpha_2(T)$ .

As we have that  $(1 - p_1)\alpha_1(t) + (1 - p_2)\alpha_2(t) \leq \alpha$ , then surely the thresholds at time  $T + 1$  in class 1 and in class 2 are less than the state of the users' proportion  $\beta(T)p_1\alpha_1(T - l(T))$  and  $\gamma(T)p_2\alpha_2(T - l(T))$  respectively. Therefore, there exists  $l_1(T + 1), l_2(T + 1), \beta(T + 1)$  and  $\gamma(T + 1)$  such that  $0 < \beta(T + 1) \leq 1$  and  $\gamma(T + 1) = 1$ , or  $\beta(T + 1) = 1$  and  $0 < \gamma(T + 1) \leq 1$ :

$$\begin{aligned} & \sum_{j=1}^{l_1(T+1)-1} p_1\alpha_1(T+1-j) + \sum_{j=1}^{l_2(T+1)-1} p_2\alpha_2(T+1-j) \\ & + \beta(T+1)p_1\alpha_1(T+1-l_1(T+1)) \\ & + \gamma(T)p_2\alpha_2(T+1-l_2(T+1)) \\ & = 1 - \alpha \end{aligned} \quad (50)$$

Now we prove by contradiction that  $\max(l_1(T + 1), l_2(T + 1)) \leq T_{\max}$ .

We prove first that  $l_2(T + 1)$  is greater than  $l_1(T + 1)$ .

As  $w_2(l_1(T + 1)) < w_1(l_1(T + 1))$ , that means according to lemma 6,  $l_2(T + 1)$  is greater than  $l_1(T + 1)$  ( $w_2(l_1(T + 1)) < w_2(l_2(T + 1))$ ).

Reasoning by contradiction, if  $l_2(T + 1) > T_{\max}$ , then we distinguish between two cases:

- First case: If  $\beta(T + 1) = 1$ :

we have that  $w_1(T_{\max}) < w_2(l_2(T + 1))$  ( $w_1(T_{\max}) < w_2(T_{\max} + 1)$ ) as the alternation condition is satisfied in  $[1, T_{\max} + 1]$ , i.e., according to lemma 6, we have that  $l_{\max} \leq l_1(T + 1)$ . Hence, according to lemma 2, we have that:

$$\begin{aligned} & \sum_{j=1}^{l_1(T+1)-1} p_1\alpha_1(T+1-j) \\ & + \sum_{j=1}^{l_2(T+1)-1} p_2\alpha_2(T+1-j) \\ & + \beta(T+1)p_1\alpha_1(T+1-l_1(T+1)) \\ & + \gamma(T+1)p_2\alpha_2(T+1-l_2(T+1)) = 1 - \alpha \\ & \geq \sum_{j=1}^{T_{\max}} p_1\alpha_1(T+1-j) + \sum_{j=1}^{T_{\max}} p_2\alpha_2(T+1-j) \\ & > T_{\max}p_2\alpha \geq 1 - \alpha \end{aligned} \quad (51)$$

Therefore we end up with:

$$1 - \alpha > 1 - \alpha \quad (52)$$

Hence, the assumption that  $l_2(T + 1) > T_{\max}$  leads us to an illogical statement. Consequently, the hypothesis of  $l_2(T + 1) > l_{\max}$  is not valid for the first case.

- Second case: If  $\beta(T + 1) < 1$ :

Then we have that  $\gamma(T + 1) = 1$ . Therefore, all users at state  $l_2(T + 1)$  in class 2 are in the proportion  $1 - \alpha$  with the smallest Whittle index values. However, there are users in state  $l_1(T + 1)$  in class 1 of the  $\alpha$  proportion with the highest Whittle index values. In other words,  $w_1(l_1(T + 1)) \geq w_2(l_2(T + 1)) > w_1(T_{\max})$ . This means that  $l_1(T + 1) > T_{\max}$ . Therefore, according to lemma 2:

$$\begin{aligned} & \sum_{j=1}^{l_1(T+1)-1} p_1\alpha_1(T+1-j) \\ & + \sum_{j=1}^{l_2(T+1)-1} p_2\alpha_2(T+1-j) \\ & + \beta(T+1)p_1\alpha_1(T+1-l_1(T+1)) \\ & + \gamma(T+1)p_2\alpha_2(T+1-l_2(T+1)) = 1 - \alpha \\ & \geq \sum_{j=1}^{T_{\max}} p_1\alpha_1(T+1-j) + \sum_{j=1}^{T_{\max}} p_2\alpha_2(T+1-j) \\ & > T_{\max}p_2\alpha \geq 1 - \alpha > 1 - \alpha \end{aligned} \quad (53)$$

Therefore, the hypothesis of  $l_2(T + 1) > T_{\max}$  is not valid for the second case.

Consequently, we have that  $l_2(T + 1) \leq T_{\max}$ , i.e.  $\max(l_1(T + 1), l_2(T + 1)) \leq T_{\max}$ . Then, according to Proposition 9, there exists  $l(T + 1)$ , and  $\gamma(T + 1)$  and  $\beta(T + 1)$  such that:

$$\begin{aligned} & \sum_{j=1}^{l(T+1)-1} p_1\alpha_1(T+1-j) + \sum_{j=1}^{l(T+1)-1} p_2\alpha_2(T+1-j) \\ & + \beta(T+1)p_1\alpha_1(T+1-l(T+1)) \\ & + \gamma(T+1)p_2\alpha_2(T+1-l(T+1)) = 1 - \alpha \end{aligned} \quad (54)$$

where  $\beta(T + 1) = 0$  and  $0 < \gamma(T + 1) \leq 1$ , or  $0 < \beta(T + 1) \leq 1$  and  $\gamma(T + 1) = 1$ .

To conclude, we have proved by induction, that for all  $T \geq T_0$ , there exists  $l(T)$ ,  $\beta(T)$  and  $\gamma(T)$ , such that:

$$\begin{aligned} & \sum_{j=1}^{l(T)-1} p_1\alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2\alpha_2(T-j) + \beta(T)p_1\alpha_1(T-l(T)) \\ & + \gamma(T)p_2\alpha_2(T-l(T)) = 1 - \alpha \end{aligned} \quad (55)$$

where  $\beta(T) = 0$  and  $0 < \gamma(T) \leq 1$ , or  $0 < \beta(T) \leq 1$  and  $\gamma(T) = 1$ , which concludes our proof.

APPENDIX H  
PROOF OF PROPOSITION 11

We proceed by the same method used to prove the Proposition 10.

We consider at time  $T$ :

$$\sum_{j=1}^{l(T)-1} p_1 \alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2 \alpha_2(T-j) + \beta(T) p_1 \alpha_1(T-l(T)) + \gamma(T) p_2 \alpha_2(T-l(T)) = 1 - \alpha, \quad (56)$$

where  $\beta(T) = 0$  and  $0 \leq \gamma(T) < 1$ , or  $0 \leq \beta(T) < 1$  and  $\gamma(T) = 1$ . Among the users' proportion scheduled  $\alpha$ , exactly  $p_1 \alpha_1(T)$  and  $p_2 \alpha_2(T)$  will go to state one for each classes, and  $(1-p_1)\alpha_1(T)$  and  $(1-p_2)\alpha_2(T)$  will go to the next state.

For the other users for which the action taken is passive, their states will be increased by one, then the decreasing order according to the Whittle index value at the next time slot is  $\beta(T)p_1\alpha_1(T-l(T)), \gamma(T)p_2\alpha_2(T-l(T)), p_1\alpha_1(T-l(T)+1), p_2\alpha_2(T-l(T)+1), p_1\alpha_1(T-l(T)+2), p_2\alpha_2(T-l(T)+2) \cdots p_1\alpha_1(T), p_2\alpha_2(T)$  (As we said before that the order based on the value of the Whittle indices, alternate between the two classes from state 1 to  $l(T) \leq l_{\max} + 1$ ). Moreover, the users' proportion scheduled  $(1-p_1)\alpha_1(T)$  and  $(1-p_2)\alpha_2(T)$  will be at states that have Whittle index values higher than those of  $\beta(T)p_1\alpha_1(T-l(T))$  and  $\gamma(T)p_2\alpha_2(T-l(T))$  (as we have explained in the proof of Proposition 10).

Hence, the global decreasing order according to the Whittle index value is  $(1-p_1)\alpha_1(T), (1-p_1)\alpha_2(T), \beta(T)p_1\alpha_1(T-l(T)), \gamma(T)p_2\alpha_2(T-l(T)), p_1\alpha_1(T-l(T)+1), p_2\alpha_2(T-l(T)+1), p_1\alpha_1(T-l(T)+2), p_2\alpha_2(T-l(T)+2) \cdots p_1\alpha_1(T), p_2\alpha_2(T)$ .

Providing that  $(1-p_1)\alpha_1(t) + (1-p_2)\alpha_2(t) \leq \alpha$ , then at time  $T+1$ :

$$\sum_{j=1}^{l(T)-1} p_1 \alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2 \alpha_2(T-j) + \beta(T) p_1 \alpha_1(T-l(T)) + \gamma(T) p_2 \alpha_2(T-l(T)) + p_1 \alpha_1(T) + p_2 \alpha_2(T) \geq 1 - \alpha \quad (57)$$

Then, there exists  $\beta = 0$  and  $0 < \gamma \leq 1$ , or  $0 < \beta \leq 1$  and  $\gamma = 1$ , and sub-set  $\{\alpha_1(T), \alpha_2(T), \alpha_1(T-1), \alpha_2(T-1) \cdots \alpha_1(T-m), \alpha_2(T-m)\} \subset \{\alpha_1(T-l(T)), \alpha_2(T-l(T)), \alpha_1(T-l(T)+1), \alpha_2(T-l(T)+1), \alpha_1(T-l(T)+2), \alpha_2(T-l(T)+2) \cdots \alpha_1(T), \alpha_2(T)\}$ , such that:

$$\sum_{j=1}^{(m+1)-1} p_1 \alpha_1(T+1-j) + \sum_{j=1}^{(m+1)-1} p_2 \alpha_2(T+1-j) + \beta p_1 \alpha_1(T+1-(m+1)) + \gamma p_2 \alpha_2(T+1-(m+1)) = 1 - \alpha \quad (58)$$

Indeed,  $m+1$  is effectively  $l(T+1)$ ,  $\beta = \beta(T+1)$ ,  $\gamma = \gamma(T+1)$ , and the elements of the set  $\{\alpha_1(T), \alpha_1(T-1), \cdots, \alpha_1(T-m)\} \cup \{\alpha_1(T+1)\}$  and the set  $\{\alpha_2(T), \alpha_2(T-1), \cdots, \alpha_2(T-m)\} \cup \{\alpha_2(T+1)\}$  are exactly the elements of the vectors  $A_1(T+1)$  and  $A_2(T+1)$  respectively. Given that  $\{\alpha_1(T), \alpha_1(T-1), \cdots, \alpha_1(T-m)\}$  and  $\{\alpha_2(T), \alpha_2(T-$

$1), \cdots, \alpha_2(T-m)\}$  are included in the set of elements of the vector  $A_1(T)$  and  $A_2(T)$  respectively, then for  $k = 1, 2$ , all the elements of the vector  $A_k(T+1)$  except  $\alpha_k(T+1)$  belong to the elements of vector  $A_k(T)$ .

APPENDIX I  
PROOF OF PROPOSITION 12

According to Proposition 10, the elements of the vectors  $A_1(T)$  and  $A_2(T)$  satisfy:

$$\sum_{j=1}^{l(T)-1} p_1 \alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2 \alpha_2(T-j) + \beta(T) p_1 \alpha_1(T-l(T)) + \gamma(T) p_2 \alpha_2(T-l(T)) = 1 - \alpha, \quad (59)$$

where  $0 < \beta(T) \leq 1$  and  $\gamma(T) = 1$ , or  $\beta(T) = 0$  and  $0 < \gamma(T) \leq 1$ . We distinguish between two cases depending on the values of  $\beta$  and  $\gamma$  (we drop the index  $T$  on  $\beta(T)$  and  $\gamma(T)$  to ease the notation):

- First case:  $0 < \beta \leq 1$ , and  $\gamma = 1$ :

Hence:

$$\sum_{j=1}^{l(T)-1} p_1 \alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2 \alpha_2(T-j) + \beta p_1 \alpha_1(T-l(T)) + p_2 \alpha_2(T-l(T)) = 1 - \alpha \quad (60)$$

Our aim is to derive the expression of  $\alpha_k(T+1)$  for class 1 and class 2. Among the users' proportion scheduled  $\alpha$ , exactly  $p_1 \alpha_1(T)$  and  $p_2 \alpha_2(T)$  will go to state one for each class, and the rest will go to the next state. Hence:

$$\alpha_1(T+1) = (1-p_1)\alpha_1(T) + B_1(T) \quad (61)$$

$$\alpha_2(T+1) = (1-p_2)\alpha_2(T) + B_2(T) \quad (62)$$

such that  $B_1(T) + B_2(T) = p_1 \alpha_1(T) + p_2 \alpha_2(T)$ .

At time  $T+1$ , the decreasing order according to the Whittle index value is  $(1-p_1)\alpha_1(T), (1-p_2)\alpha_2(T), \beta p_1 \alpha_1(T-l(T)), p_2 \alpha_2(T-l(T)), p_1 \alpha_1(T-l(T)+1), p_2 \alpha_2(T-l(T)+1), p_1 \alpha_1(T-l(T)+2), p_2 \alpha_2(T-l(T)+2), \cdots, p_1 \alpha_1(T), p_2 \alpha_2(T)$ .

In order to get  $B_1(T)$  and  $B_2(T)$ , we sum the users' proportions at different states starting from the users' proportion  $\beta p_1 \alpha_1(T-l(T))$  following the decreasing order of the Whittle index until we get the sum that equals to  $p_1 \alpha_1(T) + p_2 \alpha_2(T)$ . We distinguish between six sub-cases and for each sub-case, we prove that  $\alpha_k(T+1)$  is surely between two elements of the vector  $A_k(T)$ . In fact, if we prove it just for one class, the result will be true for the other one, since  $\alpha_1(T) + \alpha_2(T) = \alpha$  for all  $T$ . In the following, we derive the expression of  $\alpha_k(T+1)$  for  $k = 1, 2$ , in function of the elements of the vector  $A_1(T)$  and  $A_2(T)$  and we show that  $\alpha_1(T+1)$  is surely between two elements of the vector  $A_1(T)$ .

1) If  $p_1 \alpha_1(T) + p_2 \alpha_2(T) \leq p_1 \beta \alpha_1(T-l(T))$ :

In this case  $p_1 \alpha_1(T) + p_2 \alpha_2(T)$  is less than  $p_1 \beta \alpha_1(T-l(T))$ . Therefore, we will take a proportion of users from  $p_1 \beta \alpha_1(T-l(T))$  that equals to  $p_1 \alpha_1(T) + p_2 \alpha_2(T)$

denoted by  $C$ . This users' proportion exactly equals to  $B_1(T) + B_2(T)$  that we add to  $(1 - p_1)\alpha_1(T)$  and  $(1 - p_2)\alpha_2(T)$ . Thus,  $B_1(T) + B_2(T) = C$ . However, since all the users of the proportion  $C$  belong to  $p_1\beta\alpha_1(T - l(T))$ , then  $C$  contains only the users of the class 1. Consequently,  $B_1(T) = C$  and  $B_2(T) = 0$ . Hence:

$$\alpha_2(T + 1) = (1 - p_2)\alpha_2(T) \quad (63)$$

As  $\alpha_1(T + 1) + \alpha_2(T + 1) = \alpha$ , then:

$$\alpha_1(T + 1) = \alpha - \alpha_2(T + 1) \quad (64)$$

Now we find the upper bound of  $\alpha_2(T) - \alpha_2(T + 1)$ :

$$\alpha_2(T) - \alpha_2(T + 1) = p_2\alpha_2(T) \quad (65)$$

$$\leq \beta\alpha_1(T - l(T))p_1 - \alpha_1(T)p_1 \quad (66)$$

$$\leq p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \quad (67)$$

$$= p_1(\alpha_2(T) - \alpha_2(T - l(T))) \quad (68)$$

The first inequality comes from the fact that  $p_1\alpha_1(T) + p_2\alpha_2(T) \leq p_1\beta\alpha_1(T - l(T))$  and the second one comes from the fact that  $\beta \leq 1$ .

Given that  $\alpha_2(i) - \alpha_2(j) = \alpha_1(j) - \alpha_1(i)$  for all integers  $i$  and  $j$ , thus:

$$\alpha_1(T + 1) - \alpha_1(T) \leq p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \quad (69)$$

Moreover, we have that  $\alpha_1(T + 1) - \alpha_1(T) \geq 0$  because  $\alpha_2(T + 1) - \alpha_2(T) \leq 0$ . Therefore,  $\alpha_1(T) \leq \alpha_1(T + 1)$ . On the other hands, as  $p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \geq \alpha_1(T + 1) - \alpha_1(T) \geq 0$  then  $\alpha_1(T - l(T)) - \alpha_1(T) \geq \alpha_1(T + 1) - \alpha_1(T)$ . This means that  $\alpha_1(T - l(T)) \geq \alpha_1(T + 1)$ . Consequently, we end up with:

$$\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T)) \quad (70)$$

2) If  $\beta\alpha_1(T - l(T))p_1 \leq p_1\alpha_1(T) + p_2\alpha_2(T) \leq \beta\alpha_1(T - l(T))p_1 + \alpha_2(T - l(T))p_2$ :

Hence:

$$\alpha_1(T + 1) = (1 - p_1)\alpha_1(T) + \beta p_1\alpha_1(T - l(T)) \quad (71)$$

$$\alpha_2(T + 1) = \alpha - \alpha_1(T + 1) \quad (72)$$

Then:

$$\alpha_1(T + 1) - \alpha_1(T) = \beta p_1\alpha_1(T - l(T)) - p_1\alpha_1(T) \quad (73)$$

$$\leq p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \quad (74)$$

On the other hand, we have according to the right inequality of sub-case's assumption:

$$\alpha_1(T + 1) - \alpha_1(T) = \beta p_1\alpha_1(T - l(T)) - p_1\alpha_1(T) \quad (75)$$

$$\geq p_2\alpha_2(T) - p_2\alpha_2(T - l(T)) \quad (76)$$

$$= p_2(\alpha_1(T - l(T)) - \alpha_1(T)) \quad (77)$$

Hence :

$$\begin{aligned} p_2(\alpha_1(T - l(T)) - \alpha_1(T)) &\leq \alpha_1(T + 1) - \alpha_1(T) \\ &\leq p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \end{aligned} \quad (78)$$

Knowing that  $p_2 < p_1$ , the later inequalities imply that  $\alpha_1(T - l(T)) - \alpha_1(T) \geq 0$ .

As a result we have that:

$$\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T)) \quad (79)$$

And

$$\alpha_1(T + 1) - \alpha_1(T) \leq p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \quad (80)$$

3) If  $\beta\alpha_1(T - l(T))p_1 + \alpha_2(T - l(T))p_2 \leq p_1\alpha_1(T) + p_2\alpha_2(T) \leq \beta\alpha_1(T - l(T))p_1 + \alpha_2(T - l(T))p_2 + p_1\alpha_1(T - l(T) + 1)$ :

Hence:

$$\alpha_2(T + 1) = (1 - p_2)\alpha_2(T) + p_2\alpha_2(T - l(T)) \quad (81)$$

$$\alpha_1(T + 1) = \alpha - \alpha_2(T + 1) \quad (82)$$

Therefore:

$$\alpha_2(T + 1) - \alpha_2(T) = p_2(\alpha_2(T - l(T)) - \alpha_2(T)) \quad (83)$$

And:

$$\alpha_1(T) - \alpha_1(T + 1) = p_2(\alpha_1(T) - \alpha_1(T - l(T))) \quad (84)$$

This means that if  $\alpha_1(T) \leq \alpha_1(T + 1)$ :

$$\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T)) \quad (85)$$

And

$$\alpha_1(T + 1) - \alpha_1(T) \leq p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \quad (86)$$

If  $\alpha_1(T + 1) \leq \alpha_1(T)$ :

$$\alpha_1(T - l(T)) \leq \alpha_1(T + 1) \leq \alpha_1(T) \quad (87)$$

And

$$\alpha_1(T) - \alpha_1(T + 1) \leq p_1(\alpha_1(T) - \alpha_1(T - l(T))) \quad (88)$$

4) If  $\beta\alpha_1(T - l(T))p_1 + \alpha_2(T - l(T))p_2 + p_1\alpha_1(T - l(T) + 1) \leq p_1\alpha_1(T) + p_2\alpha_2(T) \leq \beta\alpha_1(T - l(T))p_1 + \alpha_2(T - l(T))p_2 + p_1\alpha_1(T - l(T) + 1) + p_2\alpha_2(T - l(T) + 1)$  :

Hence:

$$\begin{aligned} \alpha_1(T + 1) &= (1 - p_1)\alpha_1(T) + p_1\beta\alpha_1(T - l(T)) \\ &\quad + p_1\alpha_1(T - l(T) + 1) \end{aligned} \quad (89)$$

$$\alpha_2(T + 1) = \alpha - \alpha_1(T + 1) \quad (90)$$

Therefore:

$$\begin{aligned} \alpha_1(T + 1) - \alpha_1(T) &= -p_1\alpha_1(T) + p_1\beta\alpha_1(T - l(T)) \\ &\quad + p_1\alpha_1(T - l(T) + 1) \end{aligned} \quad (91)$$

$$(92)$$

According to the left inequality of the assumption of this case, we have that:

$$\alpha_1(T+1) - \alpha_1(T) \leq p_2\alpha_2(T) - p_2\alpha_2(T-l(T)) \quad (93)$$

$$= p_2(\alpha_1(T-l(T)) - \alpha_1(T)) \quad (94)$$

On the other hand, we have that:

$$\alpha_1(T+1) - \alpha_1(T) = -p_1\alpha_1(T) + p_1\beta\alpha_1(T-l(T)) + p_1\alpha_1(T-l(T)+1) \quad (95)$$

$$\geq p_1(\alpha_1(T-l(T)+1) + \alpha_1(T)) \quad (96)$$

$$(97)$$

Hence:

$$\begin{aligned} p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) &\leq \alpha_1(T+1) - \alpha_1(T) \\ &\leq p_2(\alpha_1(T-l(T)) - \alpha_1(T)) \end{aligned} \quad (98)$$

Thus:

If  $\alpha_1(T) \leq \alpha_1(T+1)$ :

$$\alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T)) \quad (99)$$

And

$$\begin{aligned} \alpha_1(T+1) - \alpha_1(T) &\leq p_2(\alpha_1(T-l(T)) - \alpha_1(T)) \\ &\leq p_1(\alpha_1(T-l(T)) - \alpha_1(T)) \end{aligned} \quad (100)$$

$$(101)$$

If  $\alpha_1(T+1) \leq \alpha_1(T)$ :

$$\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (102)$$

And

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \quad (103)$$

5) If there exists  $m \geq 1$  such that:

$$\begin{aligned} \beta\alpha_1(T-l(T))p_1 + \alpha_2(T-l(T))p_2 + p_1\alpha_1(T-l(T)+1) + \dots + p_1\alpha_1(T-l(T)+m) + p_2\alpha_2(T-l(T)+m) &\leq \\ p_1\alpha_1(T) + p_2\alpha_2(T) &\leq \beta\alpha_1(T-l(T))p_1 + \alpha_2(T-l(T))p_2 + p_1\alpha_1(T-l(T)+1) + \dots + p_1\alpha_1(T-l(T)+m) + p_2\alpha_2(T-l(T)+m) + p_1\alpha_1(T-l(T)+m+1) : \end{aligned}$$

This means that:

$$\alpha_2(T+1) = (1-p_2)\alpha_2(T) + p_2\alpha_2(T-l(T)) + \dots + p_2\alpha_2(T-l(T)+m) \quad (104)$$

$$\alpha_1(T+1) = \alpha - \alpha_2(T+1) \quad (105)$$

We have that:

$$\begin{aligned} \alpha_2(T+1) - \alpha_2(T) &= -p_2\alpha_2(T) + p_2\alpha_2(T-l(T)) \\ &\quad + p_2\alpha_1(T-l(T)+1) + \dots \\ &\quad + p_2\alpha_2(T-l(T)+m) \\ &\geq p_2(\alpha_2(T-l(T)+1) - \alpha_2(T)) \end{aligned} \quad (106)$$

On the other hand:

$$\begin{aligned} \alpha_2(T+1) - \alpha_2(T) &= -p_2\alpha_2(T) + p_2\alpha_2(T-l(T)) \\ &\quad + p_2\alpha_2(T-l(T)+1) + \dots \\ &\quad + p_2\alpha_2(T-l(T)+m) \\ &\leq p_1\alpha_1(T) - \beta p_1\alpha_1(T-l(T)) \\ &\quad - \sum_{i=1}^m p_1\alpha_1(T-l(T)+i) \\ &\leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \\ &= p_1(\alpha_2(T-l(T)+1) - \alpha_2(T)) \end{aligned} \quad (107)$$

Thus:

$$\begin{aligned} p_2(\alpha_1(T) - \alpha_1(T-l(T)+1)) &\leq \alpha_1(T) - \alpha_1(T+1) \\ &\leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \end{aligned} \quad (108)$$

Therefore:

$$\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (109)$$

And:

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \quad (110)$$

6) If there exists  $m \geq 1$  such that:

$$\begin{aligned} \beta\alpha_1(T-l(T))p_1 + \alpha_2(T-l(T))p_2 + p_1\alpha_1(T-l(T)+1) + p_2\alpha_2(T-l(T)+1) + \dots + p_1\alpha_1(T-l(T)+m) + p_2\alpha_2(T-l(T)+m) + p_1\alpha_1(T-l(T)+m+1) &\leq \\ p_1\alpha_1(T) + p_2\alpha_2(T) &\leq \beta\alpha_1(T-l(T))p_1 + \alpha_2(T-l(T))p_2 + p_1\alpha_1(T-l(T)+1) + \dots + p_1\alpha_1(T-l(T)+m) + p_2\alpha_2(T-l(T)+m) + p_1\alpha_1(T-l(T)+m+1) + p_2\alpha_2(T-l(T)+m+1) : \end{aligned}$$

Hence:

$$\alpha_1(T+1) = (1-p_1)\alpha_1(T) + p_1\beta\alpha_1(T-l(T)) + \dots + p_1\alpha_1(T-l(T)+m+1) \quad (111)$$

$$\alpha_2(T+1) = \alpha - \alpha_1(T+1) \quad (112)$$

We have that:

$$\begin{aligned} \alpha_1(T+1) - \alpha_1(T) &= -p_1\alpha_1(T) + p_1\beta\alpha_1(T-l(T)) \\ &\quad + p_1\alpha_1(T-l(T)+1) + \dots \\ &\quad + p_1\alpha_1(T-l(T)+m+1) \\ &\geq p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) \end{aligned} \quad (113)$$

On the other hand:

$$\begin{aligned} \alpha_1(T+1) - \alpha_1(T) &= -p_1\alpha_1(T) + p_1\beta\alpha_1(T-l(T)) \\ &\quad + p_1\alpha_1(T-l(T)+1) + \dots \\ &\quad + p_1\alpha_1(T-l(T)+m+1) \\ &\leq p_2\alpha_2(T) - \sum_{i=0}^m p_2\alpha_2(T-l(T)+i) \\ &\leq p_2(\alpha_2(T) - \alpha_2(T-l(T)+1)) \\ &= p_2(\alpha_1(T-l(T)+1) - \alpha_1(T)) \end{aligned} \quad (114)$$

Thus:

$$\begin{aligned} p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) &\leq \alpha_1(T+1) - \alpha_1(T) \\ &\leq p_2(\alpha_1(T-l(T)+1) - \alpha_1(T)) \end{aligned} \quad (115)$$

Therefore:

$$\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (116)$$

And:

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \quad (117)$$

- Second case:  $\beta = 0$  and  $0 < \gamma \leq 1$ :

Hence, we have that:

$$\begin{aligned} \sum_{j=1}^{l(T)-1} p_1\alpha_1(T-j) + \sum_{j=1}^{l(T)-1} p_2\alpha_2(T-j) + \gamma p_2\alpha_2(T-l(T)) \\ = 1 - \alpha \end{aligned} \quad (118)$$

Then, at time  $T+1$ , the decreasing order according to the Whittle index value is  $(1-p_1)\alpha_1(T), (1-p_2)\alpha_2(T), \gamma p_2\alpha_2(T-l(T)), p_1\alpha_1(T-l(T)+1), p_2\alpha_2(T-l(T)+1), p_1\alpha_1(T-l(T)+2), p_2\alpha_2(T-l(T)+2), \dots, p_1\alpha_1(T), p_2\alpha_2(T)$ . In order to obtain  $B_1(T)$  and  $B_2(T)$ , we sum the users' proportions at different states starting from the users' proportion  $\gamma p_2\alpha_2(T-l(T))$  following the decreasing order of the Whittle index until we get the sum that equals to  $p_1\alpha_1(T) + p_2\alpha_2(T)$ . For this case, we distinguish between five sub-cases, and for each sub-case, we prove that  $\alpha_1(T+1)$  is surely between two elements of the vector  $A_1(T)$ .

1) If  $p_1\alpha_1(T) + p_2\alpha_2(T) \leq \gamma\alpha_2(T-l(T))p_2$ :

Hence:

$$\alpha_1(T+1) = (1-p_1)\alpha_1(T) \quad (119)$$

$$\alpha_2(T+1) = \alpha - \alpha_1(T+1) \quad (120)$$

We have that:

$$\alpha_1(T) - \alpha_1(T+1) = p_1\alpha_1(T) \quad (121)$$

$$\leq \gamma\alpha_2(T-l(T))p_2 - \alpha_2(T)p_2 \quad (122)$$

$$\leq p_2(\alpha_2(T-l(T)) - \alpha_2(T)) \quad (123)$$

$$= p_2(\alpha_1(T) - \alpha_1(T-l(T))) \quad (124)$$

Thus:

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T))) \quad (125)$$

And:

$$\alpha_1(T-l(T)) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (126)$$

2) If  $\gamma\alpha_2(T-l(T))p_2 \leq p_1\alpha_1(T) + p_2\alpha_2(T) \leq \gamma\alpha_2(T-l(T))p_2 + \alpha_1(T-l(T)+1)p_1$

Consequently:

$$\alpha_2(T+1) = (1-p_2)\alpha_2(T) + \gamma p_2\alpha_2(T-l(T)) \quad (127)$$

$$\alpha_1(T+1) = \alpha - \alpha_2(T+1) \quad (128)$$

Hence:

$$\alpha_2(T+1) - \alpha_2(T) = -p_2\alpha_2(T) + \gamma p_2\alpha_2(T-l(T)) \quad (129)$$

$$\leq p_2(\alpha_2(T-l(T)) - \alpha_2(T)) \quad (130)$$

$$(131)$$

On the other hand, according to the right inequality of the assumption of this case, we have that:

$$\alpha_2(T+1) - \alpha_2(T) = -p_2\alpha_2(T) + \gamma p_2\alpha_2(T-l(T)) \quad (132)$$

$$\geq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \quad (133)$$

$$= p_1(\alpha_2(T-l(T)+1) - \alpha_2(T)) \quad (134)$$

That means:

$$\begin{aligned} p_1(\alpha_2(T-l(T)+1) - \alpha_2(T)) &\leq \alpha_2(T+1) - \alpha_2(T) \\ &\leq p_2(\alpha_2(T-l(T)) - \alpha_2(T)) \end{aligned} \quad (135)$$

i.e.

$$\begin{aligned} p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) &\leq \alpha_1(T) - \alpha_1(T+1) \\ &\leq p_2(\alpha_1(T) - \alpha_1(T-l(T))) \end{aligned} \quad (136)$$

Therefore:

If  $\alpha_1(T) \leq \alpha_1(T+1)$ :

$$\alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T)+1) \quad (137)$$

And:

$$\alpha_1(T+1) - \alpha_1(T) \leq p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) \quad (138)$$

If  $\alpha_1(T+1) \leq \alpha_1(T)$ :

$$\alpha_1(T-l(T)) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (139)$$

And:

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T))) \quad (140)$$

3) If  $\gamma\alpha_2(T-l(T))p_2 + \alpha_1(T-l(T)+1)p_1 \leq p_1\alpha_1(T) + p_2\alpha_2(T) \leq \gamma\alpha_2(T-l(T))p_2 + \alpha_1(T-l(T)+1)p_1 + p_2\alpha_2(T-l(T)+1)$ .

Hence:

$$\alpha_1(T+1) = (1-p_1)\alpha_1(T) + p_1\alpha_1(T-l(T)+1) \quad (141)$$

$$\alpha_2(T+1) = \alpha - \alpha_1(T+1) \quad (142)$$

We have that:

$$\alpha_1(T+1) - \alpha_1(T) = p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) \quad (143)$$

If  $\alpha_1(T) \leq \alpha_1(T+1)$ :

$$\alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T)+1) \quad (144)$$

And:

$$\alpha_1(T+1) - \alpha_1(T) \leq p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) \quad (145)$$

If  $\alpha_1(T+1) \leq \alpha_1(T)$ :

$$\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (146)$$

And:

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \quad (147)$$

4) If there exists  $m \geq 1$  such that:

$$\gamma\alpha_2(T-l(T))p_2 + \dots + \alpha_1(T-l(T)+m)p_1 + p_2\alpha_2(T-l(T)+m) \leq p_1\alpha_1(T) + p_2\alpha_2(T) \leq \gamma\alpha_2(T-l(T))p_2 + \dots + \alpha_1(T-l(T)+m)p_1 + p_2\alpha_2(T-l(T)+m) + p_1\alpha_1(T-l(T)+m+1):$$

Hence:

$$\alpha_2(T+1) = (1-p_2)\alpha_2(T) + p_2\gamma\alpha_2(T-l(T)) + \dots + p_2\alpha_2(T-l(T)+m) \quad (148)$$

$$\alpha_1(T+1) = \alpha - \alpha_2(T+1) \quad (149)$$

$$\begin{aligned} \alpha_2(T+1) - \alpha_2(T) &= -p_2\alpha_2(T) + p_2\gamma\alpha_2(T-l(T)) \\ &\quad + p_2\alpha_2(T-l(T)+1) + \dots \\ &\quad + p_2\alpha_2(T-l(T)+m) \\ &\geq p_2(\alpha_2(T-l(T)+1) - \alpha_2(T)) \end{aligned} \quad (150)$$

On the other hand:

$$\begin{aligned} \alpha_2(T+1) - \alpha_2(T) &= -p_2\alpha_2(T) + p_2\gamma\alpha_2(T-l(T)) \\ &\quad + p_2\alpha_2(T-l(T)+1) + \dots \\ &\quad + p_2\alpha_2(T-l(T)+m) \\ &\leq p_1\alpha_1(T) - \sum_{i=1}^m p_1\alpha_1(T-l(T)+i) \\ &\leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \\ &= p_1(\alpha_2(T-l(T)+1) - \alpha_2(T)) \end{aligned} \quad (151)$$

Thus:

$$\begin{aligned} p_2(\alpha_1(T) - \alpha_1(T-l(T)+1)) &\leq \alpha_1(T) - \alpha_1(T+1) \\ &\leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \end{aligned} \quad (152)$$

Therefore:

$$\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (153)$$

And:

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \quad (154)$$

5) If there exists  $m \geq 1$  such that:

$$\begin{aligned} \gamma\alpha_2(T-l(T))p_2 + \dots + \alpha_1(T-l(T)+m)p_1 + p_2\alpha_2(T-l(T)+m) + p_1\alpha_1(T-l(T)+m+1) &\leq p_1\alpha_1(T) + p_2\alpha_2(T) \\ &\leq \gamma\alpha_2(T-l(T))p_2 + \dots + \alpha_1(T-l(T)+m)p_1 + p_2\alpha_2(T-l(T)+m) + p_1\alpha_1(T-l(T)+m+1) + p_2\alpha_2(T-l(T)+m+1): \end{aligned}$$

That implies that:

$$\begin{aligned} \alpha_1(T+1) &= (1-p_1)\alpha_1(T) + \dots + p_1\alpha_1(T-l(T)+m) \\ &\quad + p_1\alpha_1(T-l(T)+m+1) \end{aligned} \quad (155)$$

$$\alpha_2(T+1) = \alpha - \alpha_1(T+1) \quad (156)$$

$$\begin{aligned} \alpha_1(T+1) - \alpha_1(T) &= -p_1\alpha_1(T) + p_1\alpha_1(T-l(T)+1) \\ &\quad + \dots + p_1\alpha_1(T-l(T)+m+1) \\ &\geq p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) \end{aligned} \quad (157)$$

On the other hand:

$$\begin{aligned} \alpha_1(T+1) - \alpha_1(T) &= -p_1\alpha_1(T) + p_1\alpha_1(T-l(T)+1) \\ &\quad + \dots + p_1\alpha_1(T-l(T)+m+1) \\ &\leq p_2\alpha_2(T) - \gamma p_2\alpha_2(T-l(T)) \\ &\quad - \sum_{i=1}^m p_2\alpha_2(T-l(T)+i) \\ &\leq p_2(\alpha_2(T) - \alpha_2(T-l(T)+1)) \\ &= p_2(\alpha_1(T-l(T)+1) - \alpha_1(T)) \end{aligned} \quad (158)$$

Thus:

$$\begin{aligned} p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)) &\leq \alpha_1(T+1) - \alpha_1(T) \\ &\leq p_2(\alpha_1(T-l(T)+1) - \alpha_1(T)) \end{aligned} \quad (159)$$

Therefore:

$$\alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T) \quad (160)$$

And:

$$\alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)) \quad (161)$$

In conclusion, all these six sub-cases when  $\gamma = 1$  and  $0 < \beta \leq 1$ , plus the five sub-cases when  $\beta = 0$  and  $0 < \gamma \leq 1$ , can be summarized in four cases:

$$1) \alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T)), \text{ and } \alpha_1(T+1) - \alpha_1(T) \leq p_1(\alpha_1(T-l(T)) - \alpha_1(T)).$$

$$2) \alpha_1(T-l(T)) \leq \alpha_1(T+1) \leq \alpha_1(T), \text{ and } \alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T))).$$

$$3) \alpha_1(T-l(T)+1) \leq \alpha_1(T+1) \leq \alpha_1(T), \text{ and } \alpha_1(T) - \alpha_1(T+1) \leq p_1(\alpha_1(T) - \alpha_1(T-l(T)+1)).$$

$$4) \alpha_1(T) \leq \alpha_1(T+1) \leq \alpha_1(T-l(T)+1), \text{ and } \alpha_1(T+1) - \alpha_1(T) \leq p_1(\alpha_1(T-l(T)+1) - \alpha_1(T)).$$

Thus, the proof is concluded.



APPENDIX J  
PROOF OF PROPOSITION 13

In all the proof, we consider that  $\epsilon \leq (l_2 - l_1) \frac{(1-p_1)^L}{1-(1-p_1)^L}$ . Before tackling the proof, we give a brief insight about the procedure adopted to establish the desired result: We start by finding a given time denoted  $T_2 \geq T_\epsilon$  where  $\alpha_1(T_2)$  is less than  $l_1$ . Then, we show that  $\alpha_1(T_2), \dots, \alpha_1(T_2 + L)$  are strictly less than  $l_2$ . To that end, we start first by defining a relevant sequence  $u_n$  in function of  $\epsilon, l_1, l_2$  and  $p_1$  when  $n \in [0, L]$ . After that, we prove that  $u_n$  is increasing with  $n$  and strictly less than  $l_2$ . Next, we establish that  $u_n$  is an upper bound of  $\alpha_1(\cdot)$  in  $[T_2, T_2 + L]$ . More precisely, we show that  $\alpha_1(T_2 + n) \leq u_n$  for  $n \in [0, L]$ . For that purpose, we proceed with two following steps: The first one consists of deriving an inequality verified by two consecutive terms of the sequence  $\alpha_1(\cdot)$ , namely  $\alpha_1(T)$  and  $\alpha_1(T + 1)$  using the Proposition 12 given that  $T \geq T_\epsilon$ . As for the second step, we use essentially the aforementioned result to demonstrate by induction that  $u_n$  is indeed an upper bound of  $\alpha_1(T_2 + n)$ . Finally, based on these results, we show that there exists  $T_d$  such that  $\max A_1(T_d) < l_2$ .

To find a time  $T_2 \geq T_\epsilon$  such that  $\alpha_1(T_2)$  is less than  $l_1$ , we use the fact that  $\min A_1(t) \leq l_1$  for all  $t$ . At time  $T_\epsilon + L$ , we have the vector  $A_1(T_\epsilon + L) = (\alpha_1(T_\epsilon + L), \alpha_1(T_\epsilon + L - 1), \dots, \alpha_1(T_\epsilon + L - l(T_\epsilon + L)))$ . Providing that  $\min A_1(T_\epsilon + L) \leq l_1$ , then there exists an element from the vector  $A_1(T_\epsilon + L)$  less than  $l_1$  denoted by  $\alpha_1(T_2)$ . According to 10, we have for all  $T \geq T_0$ ,  $l(T) \leq l_{\max} = L$ , then  $l(T_\epsilon + L) \leq L$ . That is,  $T_2$  is greater than  $T_\epsilon$  since  $T_2 \geq T_\epsilon + L - l(T_\epsilon + L) \geq T_\epsilon$ . Therefore, we find an element of the sequence  $\alpha_1(\cdot)$  at time  $T_2 \geq T_\epsilon$  such that  $\alpha_1(T_2) \leq l_1$ . To that extent, we are interested in proving that  $\alpha_1(T_2), \dots, \alpha_1(T_2 + L)$  are strictly less than  $l_2$ .

To do so, we define a sequence  $u_n$  which will constitute an upper bound of the function  $\alpha_1(T)$ .

**Definition 5.** We define a sequence  $u_n$  by induction:

$$\begin{cases} u_0 = l_1 & \text{if } n = 0 \\ u_{n+1} = p_1(l_2 + \epsilon) + (1 - p_1)u_n & \text{if } n > 0 \end{cases} \quad (162)$$

Next, we prove that the  $L$  first terms of this sequence are strictly less than  $l_2$ . We detail this in the following.

**Lemma 7.** For  $n \in [0, L]$ ,  $u_n < l_2$

*Proof.* In fact, the sequence  $u_n$  satisfies for all  $n$ :

$$u_n = \lambda(1 - p_1)^n + (l_2 + \epsilon) \quad (163)$$

where  $\lambda = -(\epsilon + l_2 - l_1)$ .

$u_n$  is clearly increasing with  $n$ , then for all  $n \in [0, L]$ :

$$\begin{aligned} u_n &\leq u_L = \lambda(1 - p_1)^L + (l_2 + \epsilon) \\ &= \epsilon(1 - (1 - p_1)^L) + l_2 - (l_2 - l_1)(1 - p_1)^L \end{aligned} \quad (164)$$

We have that:

$$\epsilon < (l_2 - l_1) \frac{(1 - p_1)^L}{1 - (1 - p)^L} \quad (165)$$

Given that  $1 - (1 - p_1)^L \geq 0$ , then:

$$(1 - (1 - p_1)^L)\epsilon < (l_2 - l_1)(1 - p_1)^L \quad (166)$$

$$(1 - (1 - p_1)^L)\epsilon + l_2 - (l_2 - l_1)(1 - p_1)^L < l_2 \quad (167)$$

Therefore,  $u_L < l_2$ . ■

Based on the lemma above, we prove that for any element of the set  $\{\alpha_1(T_2), \dots, \alpha_1(T_2 + L)\}$  must be less than  $u_L$ . For that, we introduce a useful Lemma:

**Lemma 8.** If for  $T \in [T_2, T_2 + L - 1]$ , we have that:

$$\alpha_1(T) \leq \alpha_1(T + 1) \quad (168)$$

Then, we have that:

$$\alpha_1(T + 1) \leq p_1(l_2 + \epsilon) + (1 - p_1)\alpha_1(T) \quad (169)$$

*Proof.* Before starting the proof, we recall that, according to the first result of Proposition 12, the four possible inequalities satisfied by  $\alpha_1(T), \alpha_1(T + 1), \alpha_1(T - l(T)), \alpha_1(T - l(T) + 1)$  are:

$$\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T)) \quad (170)$$

$$\alpha_1(T - l(T)) \leq \alpha_1(T + 1) \leq \alpha_1(T) \quad (171)$$

$$\alpha_1(T - l(T) + 1) \leq \alpha_1(T + 1) \leq \alpha_1(T) \quad (172)$$

$$\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T) + 1) \quad (173)$$

Therefore, the two cases for which  $\alpha_1(T) \leq \alpha_1(T + 1)$  are:

- $\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T))$ .
- $\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T) + 1)$ .

Hence, according to the results of Proposition 12, the inequalities satisfied by  $\alpha_1(T + 1) - \alpha_1(T)$  are:

If  $\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T))$ , then:

$$\alpha_1(T + 1) - \alpha_1(T) \leq p_1(\alpha_1(T - l(T)) - \alpha_1(T)) \quad (174)$$

If  $\alpha_1(T) \leq \alpha_1(T + 1) \leq \alpha_1(T - l(T) + 1)$ , then:

$$\alpha_1(T + 1) - \alpha_1(T) \leq p_1(\alpha_1(T - l(T) + 1) - \alpha_1(T)) \quad (175)$$

Since, by assumption of the Lemma,  $T \geq T_2 \geq T_\epsilon$ , then  $\max A_1(T) \leq l_2 + \epsilon$ . As a consequence,  $\alpha_1(T - l(T) + 1)$  and  $\alpha_1(T - l(T))$  which are elements of the vector  $A_1(T)$ , are less than  $l_2 + \epsilon$ .

Hence, for  $T \in [T_2, T_2 + L - 1]$ :

$$\alpha_1(T + 1) - \alpha_1(T) \leq p_1(l_2 + \epsilon - \alpha_1(T)) \quad (176)$$

Therefore:

$$\alpha_1(T + 1) \leq p_1(l_2 + \epsilon) + (1 - p_1)\alpha_1(T) \quad (177)$$

Now we should prove that for all possible sequences of  $\alpha_1$  in  $[T_2, T_2 + L]$ , their values can not exceed  $\lambda(1 - p_1)^L + (l_2 + \epsilon) = u_L$ . ■

**Lemma 9.** For all sequences of  $\alpha_1$  when  $T \in [T_2, T_2 + L]$ ,  $\alpha_1(T) \leq u_{T-T_2}$

*Proof.* We prove this result by induction.

For  $T = T_2$ , we have that:

$$\alpha_1(T_2) \leq l_1 = u_0 \quad (178)$$

We suppose that at time  $T$ ,  $\alpha_1(T) \leq u_{T-T_2}$ , then at time  $T+1$ :

If  $\alpha_1(T+1) \leq \alpha_1(T)$ :

Then as  $u_{T-T_2}$  is increasing in  $T$ :

$$\alpha_1(T+1) \leq u_{T-T_2} \leq u_{T-T_2+1} \quad (179)$$

If  $\alpha_1(T+1) \geq \alpha_1(T)$ :

Then, according to Lemma 8:

$$\alpha_1(T+1) \leq p_1(l_2 + \epsilon) + (1-p_1)\alpha_1(T) \quad (180)$$

$$\leq p_1(l_2 + \epsilon) + (1-p_1)u_{T-T_2} \quad (181)$$

$$= u_{T-T_2+1} \quad (182)$$

Therefore,  $\alpha_1(T+1) \leq u_{T-T_2+1}$ .

Hence, we have proved by induction that for all  $T \in [T_2, T_2 + L]$ ,  $\alpha_1(T) \leq u_{T-T_2}$  ■

As  $u_{T-T_2}$  is less than  $u_L$  for  $T \in [T_2, T_2 + L]$ , then according to Lemma 9, the elements  $\alpha_1(T_2 + 1), \dots, \alpha_1(T_2 + L)$  are less than  $u_L < l_2$ .

Thus, we have found  $T_2 \geq T_\epsilon$  such that  $\alpha_1(T_2), \alpha_1(T_2 + 1), \dots, \alpha_1(T_2 + l_{\max})$  are strictly less than  $l_2$ . We denote  $T_2 + l_{\max}$  by  $T_d$  and we verify that  $\max A_1(T_d) < l_2$ . Indeed, we now that  $T_d - l(T_d) \geq T_d - l_{\max} = T_2$ , then the elements of the vector  $A_1(T_d)$  are included in the set of elements  $\{\alpha_1(T_2), \alpha_1(T_2 + 1), \dots, \alpha_1(T_2 + l_{\max})\}$ . That is  $\max A_1(T_d) < l_2$ .

Hence, we have found  $T_d \geq T_\epsilon$ , such that  $\max A_1(T_d) < l_2$ .

## APPENDIX K

### PROOF OF PROPOSITION 14

In this proof, we show that for each state  $i$  in class  $k$ ,  $z_i^k(t)$  converges. To that end, we start first by specifying the eventual limit of  $z_i^k(t)$  for each  $i$ . To do so, we decompose  $1 - \alpha$  as follows:

$$l(p_1\alpha_1^* + p_2\alpha_2^*) + \gamma p_2\alpha_2^* + \beta p_1\alpha_1^* = 1 - \alpha \quad (183)$$

where  $l$  is the biggest integer such that:  $l(p_1\alpha_1^* + p_2\alpha_2^*) < 1 - \alpha$ , and  $0 < \gamma \leq 1$  and  $\beta = 0$ ; or  $\gamma = 1$  and  $0 < \beta \leq 1$ .

Then, we proceed with these following steps:

- We prove by induction that for all states  $1 \leq i \leq l+1$ ,  $z_i^k(t)$  converges to  $p_k\alpha_k^*$ .
- Based on the theoretical findings of the first step, we prove that  $z_{l+2}^1(t)$  converges to  $(\beta + (1-p_1)(1-\beta))p_1\alpha_1^*$  and  $z_{l+2}^2(t)$  converges to  $(\gamma + (1-p_2)(1-\gamma))p_2\alpha_2^*$ .
- Finally, we show that for all states  $i > l+2$ ,  $z_i^1(t)$  converges to  $(1-p_1)^{i-l-2}(\beta + (1-p_1)(1-\beta))p_1\alpha_1^*$  and  $z_i^2(t)$  converges to  $(1-p_2)^{i-l-2}(\gamma + (1-p_2)(1-\gamma))p_2\alpha_2^*$

1) For all states  $1 \leq i \leq l+1$ ,  $z_i^k(t) \rightarrow p_k\alpha_k^*$ :

We prove this result by induction

- For  $i = 1$ , we have that  $z_1^k(t) = p_k\alpha_k(t-1)$ . Therefore,  $z_1^k(t)$  converge to  $p_k\alpha_k^*$  as  $\alpha_k(t)$  converges to  $\alpha_k^*$ .

- We consider that for a certain  $j \leq l$ , for each  $1 \leq i \leq j$ ,  $z_i^k(t)$  converges to  $p_k\alpha_k^*$  and we show that  $z_{j+1}^k(t)$  converges also to  $p_k\alpha_k^*$ . Given that  $j \leq l$ :

$$j(p_1\alpha_1^* + p_2\alpha_2^*) < 1 - \alpha$$

We consider  $0 < \epsilon \leq 1 - \alpha - j(p_1\alpha_1^* + p_2\alpha_2^*)$ . Providing that  $z_i^k(t)$  converges to  $p_k\alpha_k^*$  for all  $1 \leq i \leq j$ , that means there exists  $t_j$  such that for  $t \geq t_j$ , for  $1 \leq i \leq j$ :

$$|z_i^k(t) - p_k\alpha_k^*| < \frac{\epsilon}{2^j}$$

Hence:

$$\sum_{i=1}^j |z_i^1(t) - p_1\alpha_1^*| + \sum_{i=1}^j |z_i^2(t) - p_2\alpha_2^*| < \epsilon$$

That is,

$$\sum_{i=1}^j z_i^1(t) + \sum_{i=1}^j z_i^2(t) < \epsilon + j(p_1\alpha_1^* + p_2\alpha_2^*)$$

As consequence, for all  $t \geq t_j$ , we have that:

$$\sum_{i=1}^j z_i^1(t) + \sum_{i=1}^j z_i^2(t) < 1 - \alpha$$

Thus, for all  $t \geq t_j$ , the action prescribed to the users' proportion  $z_j^k(t)$  is the passive action<sup>2</sup>. Then, for all  $t \geq t_j$ :

$$z_{j+1}^k(t+1) = z_j^k(t)$$

Therefore,  $z_{j+1}^k(t)$  converges to  $p_k\alpha_k^*$ .

Consequently, we prove by induction that for all  $1 \leq i \leq l+1$ ,  $z_i^k(t)$  converges to  $p_k\alpha_k^*$ .

2)  $z_{l+2}^1(t) \rightarrow (\beta + (1-p_1)(1-\beta))p_1\alpha_1^*$  and  $z_{l+2}^2(t) \rightarrow (\gamma + (1-p_2)(1-\gamma))p_2\alpha_2^*$ .

To avoid redundancy, we will be limited to the first case when  $0 < \gamma \leq 1$  and  $\beta = 0$ , since the proof's steps for both cases are exactly the same. We have that:

$$l(p_1\alpha_1^* + p_2\alpha_2^*) + \gamma p_2\alpha_2^* = 1 - \alpha$$

As  $\sum_{i=1}^l z_i^1(t) + \sum_{i=1}^l z_i^2(t)$  converges to  $l(p_1\alpha_1^* + p_2\alpha_2^*)$  which is strictly less than  $1 - \alpha$ , then there exists  $t_l$  such that for all  $t \geq t_l$ , we have that:

$$\sum_{i=1}^l z_i^1(t) + \sum_{i=1}^l z_i^2(t) < 1 - \alpha$$

<sup>2</sup>Knowing that the order of the proportions of the users according to the Whittle's index value alternates between the two classes in the set  $[1, l_{\max} + 1]$  as was established in Proposition 7, then for all integer  $b \in [1, l_{\max}]$ , the set  $\{z_i^k : k = 1, 2; 1 \leq i \leq b\}$  is the set of users with the lowest Whittle's index value. Therefore,  $\sum_{i=1}^b z_i^1(t) + \sum_{i=1}^b z_i^2(t) < 1 - \alpha$  implies that the actions prescribed to the users belonging to the set  $\{z_i^k : k = 1, 2; 1 \leq i \leq b\}$  is the passive action. By definition of  $l$ ,  $l < \frac{1-\alpha}{p_2\alpha_2^*}$ , then,  $l < l_{\max}$  (see Definition 4). Hence, the above reasoning can be applied as well when  $b = l$ .

As  $\sum_{i=1}^{l+1} z_i^1(t) + \sum_{i=1}^{l+1} z_i^2(t)$  converges to  $(l+1)(p_1\alpha_1^* + p_2\alpha_2^*)$  which is strictly greater than  $1 - \alpha$ , then there exists  $t_{l+1}$  such that for all  $t \geq t_{l+1}$ , we have that:

$$\sum_{i=1}^{l+1} z_i^1(t) + \sum_{i=1}^{l+1} z_i^2(t) > 1 - \alpha$$

For  $t \geq \max\{t_l, t_{l+1}\}$ , we have that:

$$\sum_{i=1}^l z_i^1(t) + \sum_{i=1}^l z_i^2(t) < 1 - \alpha < \sum_{i=1}^{l+1} z_i^1(t) + \sum_{i=1}^{l+1} z_i^2(t)$$

Denoting  $\gamma(t)$  and  $\beta(t)$  the users' proportion of  $z_{l+1}^2(t)$  and  $z_{l+1}^1(t)$  respectively which are not scheduled, therefore, the relation that links  $z_{l+2}^1(t+1)$  and  $z_{l+2}^2(t+1)$  to  $z_{l+1}^1(t)$  and  $z_{l+1}^2(t)$  when  $t \geq \max\{t_l, t_{l+1}\}$ :

$$\begin{aligned} z_{l+2}^1(t+1) &= \beta(t)z_{l+1}^1(t) + (1-p_1)(1-\beta(t))z_{l+1}^1(t) \\ z_{l+2}^2(t+1) &= \gamma(t)z_{l+1}^2(t) + (1-p_2)(1-\gamma(t))z_{l+1}^2(t) \end{aligned}$$

with  $0 < \gamma(t) \leq 1$  and  $\beta(t) = 0$ ; or  $\gamma(t) = 1$  and  $0 < \beta(t) \leq 1$ . To that extent, we show that  $\beta(t)$  tends to  $\beta = 0$  and  $\gamma(t)$  tends to  $\gamma$ . For that purpose, we give the following equation which is always satisfied when  $t \geq \max\{t_l, t_{l+1}\}$ :

$$\sum_{i=1}^l z_i^1(t) + \sum_{i=1}^l z_i^2(t) + \gamma(t)z_{l+1}^2(t) + \beta(t)z_{l+1}^1(t) = 1 - \alpha \quad (184)$$

Tending  $t$  to  $+\infty$  in the equation 184, we obtain:

$$\lim_{t \rightarrow +\infty} [\gamma(t)z_{l+1}^2(t) + \beta(t)z_{l+1}^1(t)] = \gamma p_2 \alpha_2^*$$

We consider the set  $\{t : \beta(t) \neq 0\}$ . If this set is infinite, then there exists a strictly increasing function  $n(\cdot)$  from  $\mathbf{N}$  to  $\{t \in \mathbf{N} : \beta(t) \neq 0\}$ , such that  $\beta(n(t))$  is a subsequence of  $\beta(t)$ . As  $\beta(n(t)) \neq 0$ , then  $\gamma(n(t)) = 1$ . Therefore, we get:

$$\lim_{t \rightarrow +\infty} [z_{l+1}^2(n(t)) + \beta(n(t))z_{l+1}^1(n(t))] = \gamma p_2 \alpha_2^*$$

Since  $z_{l+1}^2(n(t))$  converges to  $p_2 \alpha_2^*$ , then:

$$\lim_{t \rightarrow +\infty} [\beta(n(t))z_{l+1}^1(n(t))] = (\gamma - 1)p_2 \alpha_2^*$$

$(\gamma - 1)p_2 \alpha_2^*$  is less than 0, and  $\beta(n(t))z_{l+1}^1(n(t))$  is greater than 0 for all  $t$ . Thus:

$$\lim_{t \rightarrow +\infty} [\beta(n(t))z_{l+1}^1(n(t))] = (\gamma - 1)p_2 \alpha_2^* = 0$$

This implies that  $\gamma = 1 = \gamma(n(t))$ , and  $\lim_{t \rightarrow +\infty} \beta(n(t)) = 0$  because  $z_{l+1}^1(n(t))$  converges to  $p_1 \alpha_1^* \neq 0$ . Hence  $\lim_{t \rightarrow +\infty} \beta(t) = 0 = \beta$ , i.e.  $\lim_{t \rightarrow +\infty} \gamma(t) = \gamma = 1$ . If  $\{t : \beta(t) \neq 0\}$  is finite, then there exists  $t_e$  such that for all  $t \geq t_e$ ,  $\beta(t) = 0$ . Therefore, for all  $t \geq t_e$ , we have that:

$$\lim_{t \rightarrow +\infty} [\gamma(t)z_{l+1}^2(t)] = \gamma p_2 \alpha_2^*$$

That means  $\lim_{t \rightarrow +\infty} \beta(t) = 0$ , and  $\lim_{t \rightarrow +\infty} \gamma(t) = \gamma$ . Hence, in both cases,  $\beta(t) \rightarrow \beta = 0$  and  $\gamma(t) \rightarrow \gamma$ .

Consequently, combining the last result with the one derived in the first step, we conclude that  $z_{l+2}^1(t)$  converges to  $(\beta + (1-p_1)(1-\beta))p_1\alpha_1^*$  and  $z_{l+2}^2(t)$  converges to  $(\gamma + (1-p_2)(1-\gamma))p_2\alpha_2^*$ . Similar analysis can be applied to come with the aforementioned result when  $\gamma(t) = 1$  and  $0 < \beta(t) \leq 1$ .

- 3) For  $i > l+2$ ,  $z_i^1(t) \rightarrow (1-p_1)^{i-l-2}(\beta + (1-p_1)(1-\beta))p_1\alpha_1^*$  and  $z_i^2(t) \rightarrow (1-p_2)^{i-l-2}(\gamma + (1-p_2)(1-\gamma))p_2\alpha_2^*$ :

For  $t \geq \max\{t_l, t_{l+1}\}$ , we are sure that the action prescribed to  $z_i^k(t)$  for all  $i \geq l+2$  is the active action. As consequence,  $z_{i+1}^k(t+1)$  satisfies:

$$z_{i+1}^k(t+1) = (1-p_k)z_i^k(t)$$

Therefore, as  $z_{l+2}^1(t)$  converges to  $(\beta + (1-p_1)(1-\beta))p_1\alpha_1^*$  and  $z_{l+2}^2(t)$  converges to  $(\gamma + (1-p_2)(1-\gamma))p_2\alpha_2^*$ , one can easily establish by induction that  $z_i^1(t)$  converges to  $(1-p_1)^{i-l-2}(\beta + (1-p_1)(1-\beta))p_1\alpha_1^*$  and  $z_i^2(t)$  converges to  $(1-p_2)^{i-l-2}(\gamma + (1-p_2)(1-\gamma))p_2\alpha_2^*$  for all  $i > l+2$ .

We conclude that for all states  $i$  and  $k = 1, 2$ ,  $z_i^k(t)$  converges. On the other hands, according to Proposition 6, the only possible limit of  $z(t)$  is  $z^*$ . As consequence, for each  $k$  and  $i$ ,  $z_i^k(t)$  converges to  $z_i^{k,*}$ .

## APPENDIX L

### PROOF OF PROPOSITION 15

For a given  $z$ , let  $m_1(z)$  and  $m_2(z)$  be the highest states of the class 1 and the class 2 respectively and  $l_1(z)$  and  $l_2(z)$  be the thresholds of class 1 and 2 respectively at time  $t$  when  $\mathbf{Z}^N(t) = z$ . Given that, we introduce the following lemma.

**Lemma 10.** *For any  $\mu$ , there exists positive constant  $C(z)$  such that:*

$$P(\|\mathbf{Z}^N(t+1) - z'\| \geq \mu | \mathbf{Z}^N(t) = z) \leq \frac{C(z)}{N} \quad (185)$$

where  $C(z)$  is independent of  $N$  and  $z' = Q(z)z = \mathbf{E}(\mathbf{Z}^N(t+1) | \mathbf{Z}^N(t) = z)$

*Proof.* By definition of  $m_1(z)$  and  $m_2(z)$ , we have that  $z = (z_1^1, \dots, z_{m_1(z)}^1, z_1^2, \dots, z_{m_2(z)}^2)$ . One can easily show that  $m_1(z') = m_1(z) + 1$  and  $m_2(z') = m_2(z) + 1$  since the users' proportions at states  $m_1(z)$  and  $m_2(z)$  in class 1 and class 2 will become at states  $m_1(z) + 1$  and  $m_2(z) + 1$  at the next time slot respectively. To prove this lemma, we use the Chebychev inequality presented as follows:

$$P(|X - \mathbb{E}(X)| > \mu) \leq \frac{\text{Var}(X)}{\mu^2} \quad (186)$$

for any  $\mu > 0$  and random variable  $X$ . As  $z' = \mathbf{E}(\mathbf{Z}^N(t+1) | \mathbf{Z}^N(t) = z)$ , we can apply the Chebychev inequality. However we need to find the distribution of  $\mathbf{Z}^N(t+1)$  knowing  $\mathbf{Z}^N(t) = z$  in order to derive the expression of  $\text{Var}(\mathbf{Z}^N(t+1) | \mathbf{Z}^N(t) = z)$ . It is more simple to study the parameters of one dimensional random variable than multi-dimensional random variable. Hence, instead of

investigating  $\mathbf{Z}^N(t+1)$ , we look into  $Z_i^{N,k}$ . In this regard, we have that:

$$\begin{aligned} & \{\mathbf{Z}^N(t+1) : \|\mathbf{Z}^N(t+1) - \mathbf{z}'\| \geq \mu\} \\ & \subset \bigcup_{k,i} \{\mathbf{Z}^N(t+1) : \|Z_i^{N,k}(t+1) - z_i^{\prime k}\| > \frac{\mu}{m_1(\mathbf{z}') + m_2(\mathbf{z}')}\} \end{aligned} \quad (187)$$

Therefore:

$$\begin{aligned} & P(\|\mathbf{Z}^N(t+1) - \mathbf{z}'\| \geq \mu | \mathbf{Z}^N(t) = \mathbf{z}) \\ & \leq P(\bigcup_{k,i} \{\|Z_i^{N,k}(t+1) - z_i^{\prime k}\| > \frac{\mu}{m_1(\mathbf{z}') + m_2(\mathbf{z}')}\} | \mathbf{Z}^N(t) = \mathbf{z}) \\ & \leq \sum_{k,i} P(\{\|Z_i^{N,k}(t+1) - z_i^{\prime k}\| > \frac{\mu}{m_1(\mathbf{z}') + m_2(\mathbf{z}')}\} | \mathbf{Z}^N(t) = \mathbf{z}) \end{aligned} \quad (188)$$

Now, we look for the distribution of  $Z_i^{N,k}(t+1)$  knowing  $\mathbf{Z}^N(t) = \mathbf{z}$ .

For  $2 \leq i \leq l_k(\mathbf{z})$ , as all the users at state  $i-1$  strictly less than  $l_k(\mathbf{z})$  will transit to the state  $i$  at the next time slot, then we have  $Z_i^{N,k}(t+1) = z_{i-1}^k = z_i^{\prime k}$ . This implies that:

$$P(\{\|Z_i^{N,k}(t+1) - z_i^{\prime k}\| > \frac{\mu}{m_1(\mathbf{z}') + m_2(\mathbf{z}')}\} | \mathbf{Z}^N(t) = \mathbf{z}) = 0 \quad (189)$$

For  $i = 1$ , defining  $\alpha_1(\mathbf{z})$  and  $\alpha_2(\mathbf{z})$  as the proportions of the scheduled users in class 1 and class 2 respectively when  $\mathbf{Z}^N(t) = \mathbf{z}$ , then  $NZ_1^{N,k}(t+1) | \mathbf{Z}^N(t) = \mathbf{z}$  follows a binomial distribution with parameters  $p_k$  and  $\alpha_k(\mathbf{z})N$ . Therefore,  $\text{Var}(NZ_1^{N,k}(t+1) | \mathbf{Z}^N(t) = \mathbf{z}) = p_k(1-p_k)\alpha_k(\mathbf{z})N$ , which means that  $\text{Var}(Z_1^{N,k}(t+1) | \mathbf{Z}^N(t) = \mathbf{z}) = \frac{p_k(1-p_k)\alpha_k(\mathbf{z})}{N}$ . As a result, according to Chebychev inequality, we have that:

$$\begin{aligned} & P(\{\|Z_1^{N,k}(t+1) - z_1^{\prime k}\| > \frac{\mu}{m_1(\mathbf{z}') + m_2(\mathbf{z}')}\} | \mathbf{Z}^N(t) = \mathbf{z}) \\ & \leq \frac{p_k(1-p_k)\alpha_k(\mathbf{z})}{N\mu^2} (m_1(\mathbf{z}') + m_2(\mathbf{z}'))^2 \end{aligned} \quad (190)$$

For  $i \geq l_k(\mathbf{z}) + 2$ ,  $NZ_i^{N,k}(t+1) | \mathbf{Z}^N(t) = \mathbf{z}$  follows a binomial distribution with parameters  $1-p_k$  and  $z_{i-1}^k N$ . Hence,  $\text{Var}(Z_i^{N,k}(t+1) | \mathbf{Z}^N(t) = \mathbf{z}) = \frac{p_k(1-p_k)z_{i-1}^k}{N}$ . Thus:

$$\begin{aligned} & P(\{\|Z_i^{N,k}(t+1) - z_i^{\prime k}\| > \frac{\mu}{i(m_1(\mathbf{z}') + m_2(\mathbf{z}'))}\} | \mathbf{Z}^N(t) = \mathbf{z}) \\ & \leq \frac{p_k(1-p_k)z_{i-1}^k}{N\mu^2} (m_1(\mathbf{z}') + m_2(\mathbf{z}'))^2 i^2 \end{aligned} \quad (191)$$

Denoting  $\beta_k(\mathbf{z})$  the users' proportion of  $z_{l_k(\mathbf{z})}^k$  that will not be transmitted, then for  $i = l_k(\mathbf{z}) + 1$ ,  $NZ_i^{N,k}(t+1) | (\mathbf{Z}^N(t) = \mathbf{z}) = \beta_k(\mathbf{z})Nz_{i-1}^k + X$ , where  $X$  follows a binomial distribution with parameters  $1-p_k$  and  $(1-\beta_k(\mathbf{z}))z_{i-1}^k N$ , then:

$$\begin{aligned} & P(\{\|Z_i^{N,k}(t+1) - z_i^{\prime k}\| > \frac{\mu}{i(m_1(\mathbf{z}') + m_2(\mathbf{z}'))}\} | \mathbf{Z}^N(t) = \mathbf{z}) \\ & \leq \frac{p_k(1-p_k)(1-\beta_k(\mathbf{z}))z_{i-1}^k}{N\mu^2} (m_1(\mathbf{z}') + m_2(\mathbf{z}'))^2 i^2 \end{aligned} \quad (192)$$

We end up with:

$$\begin{aligned} & P(\|\mathbf{Z}^N(t+1) - \mathbf{z}'\| \geq \mu | \mathbf{Z}^N(t) = \mathbf{z}) \\ & \leq (m_1(\mathbf{z}') + m_2(\mathbf{z}'))^2 \cdot \left[ \frac{p_1(1-p_1)\alpha_1(\mathbf{z})}{N\mu^2} + \frac{p_2(1-p_2)\alpha_2(\mathbf{z})}{N\mu^2} \right. \\ & \quad + \sum_{i \geq l_1(\mathbf{z})+2} \frac{p_1(1-p_1)i^2 z_{i-1}^1}{N\mu^2} + \sum_{i \geq l_2(\mathbf{z})+2} \frac{p_2(1-p_2)i^2 z_{i-1}^2}{N\mu^2} \\ & \quad + \frac{p_1(1-p_1)(l_1(\mathbf{z})+1)^2(1-\beta_1(\mathbf{z}))z_{l_1(\mathbf{z})}^1}{N\mu^2} \\ & \quad \left. + \frac{p_1(1-p_2)(l_2(\mathbf{z})+1)^2(1-\beta_2(\mathbf{z}))z_{l_2(\mathbf{z})}^2}{N\mu^2} \right] \end{aligned} \quad (193)$$

Knowing that  $\alpha_k(\mathbf{z}) \leq 1$ ,  $\sum_{i \geq l_k(\mathbf{z})} z_i^k \leq 1$ ,  $1 - \beta_k(\mathbf{z}) \leq 1$ , and for all state  $i$  in the vector  $\mathbf{z}'$ ,  $i \leq m_1(\mathbf{z}') + m_2(\mathbf{z}')$  then:

$$\begin{aligned} & P(\|\mathbf{Z}^N(t+1) - \mathbf{z}'\| \geq \mu | \mathbf{Z}^N(t) = \mathbf{z}) \\ & \leq \frac{(m_1(\mathbf{z}') + m_2(\mathbf{z}'))^4}{\mu^2 N} [2p_1(1-p_1) + 2p_2(1-p_2)] \end{aligned} \quad (194)$$

Hence, denoting by  $C(\mathbf{z})$ ,  $\frac{(m_1(\mathbf{z}') + m_2(\mathbf{z}'))^4}{\mu^2} [2p_1(1-p_1) + 2p_2(1-p_2)] = \frac{(m_1(\mathbf{z}) + 1 + m_2(\mathbf{z}) + 1)^4}{\mu^2} [2p_1(1-p_1) + 2p_2(1-p_2)]$ , we obtain as a result:

$$P(\|\mathbf{Z}^N(t+1) - \mathbf{z}'\| \geq \mu | \mathbf{Z}^N(t) = \mathbf{z}) \leq \frac{C(\mathbf{z})}{N} \quad (195)$$

■

Now, we give a lemma that bounds the probability knowing the initial state  $\mathbf{z}(0) = \mathbf{x}$ . One can easily verify that  $m_1(\mathbf{z}(t)) = m_1(\mathbf{x}) + t$  and  $m_2(\mathbf{z}(t)) = m_2(\mathbf{x}) + t$  by induction. Without loss of generality, we let  $m_k(\mathbf{z}(t)) = m_k(t)$  for  $k = 1, 2$ .

**Lemma 11.** For any  $\mu$ , there exists positive constant  $C(t+1)$  such that:

$$P_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu) \leq \frac{C(t+1)}{N} \quad (196)$$

where  $C(t+1)$  is independent of  $N$ .

*Proof.* We recall from Lemma 10 that for any  $\mu > 0$ , there exists a constant  $C(\mathbf{z})$  independent of  $N$  such that:

$$P(\|\mathbf{Z}^N(t+1) - Q(\mathbf{z})\mathbf{z}\| \geq \mu | \mathbf{Z}^N(t) = \mathbf{z}) \leq \frac{C(\mathbf{z})}{N} \quad (197)$$

Before proving the present lemma, we give an important lemma that will help us in the later analysis.

**Lemma 12.** For any proportion vector  $\mathbf{z}(t)$ , there exists  $\sigma > 0$  such that if  $\|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \leq \sigma$ , then  $Q(\mathbf{Z}^N(t)) = Q(\mathbf{z}(t))$ .

*Proof.* One can deduce from the analysis done in [21, Section IV-C] that there exists  $\sigma > 0$  such that if  $\mathbf{Z}^N(t) \in \Omega_{\sigma}(\mathbf{z})$ ,  $Q(\mathbf{Z}^N(t))$  is constant and doesn't depend on  $\mathbf{Z}^N(t)$ . In our case, given that  $m_k(\mathbf{Z}^N(t)) = m_k(\mathbf{z}(t))$ , we can as well affirm that there exists  $\sigma > 0$  such that  $Q(\mathbf{Z}^N(t)) = Q(\mathbf{z}(t))$ . That concludes the proof. ■

**Corollary 1.** For any  $v > 0$ , there exists  $\rho$  such that  $\|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \leq \rho \Rightarrow \|\mathbf{Q}(\mathbf{Z}^N(t))\mathbf{Z}^N(t) - \mathbf{Q}(\mathbf{z}(t))\mathbf{z}(t)\| \leq v$

*Proof.* According to the previous lemma, if  $\|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \leq \sigma$ , then  $\mathbf{Q}(\mathbf{Z}^N(t)) = \mathbf{Q}(\mathbf{z}(t))$ . This implies that  $\|\mathbf{Q}(\mathbf{Z}^N(t))\mathbf{Z}^N(t) - \mathbf{Q}(\mathbf{z}(t))\mathbf{z}(t)\| = \|\mathbf{Q}(\mathbf{z}(t))\mathbf{Z}^N(t) - \mathbf{Q}(\mathbf{z}(t))\mathbf{z}(t)\| \leq \|\mathbf{Q}(\mathbf{z}(t))\| \|\mathbf{Z}^N(t) - \mathbf{z}(t)\|$ . That is, choosing  $\rho = \min\{\frac{v}{\|\mathbf{Q}(\mathbf{z}(t))\|}, \sigma\}$ , we get  $\|\mathbf{Q}(\mathbf{Z}^N(t))\mathbf{Z}^N(t) - \mathbf{Q}(\mathbf{z}(t))\mathbf{z}(t)\| \leq v$ . ■

With the above corollary being laid out, we prove the statement by a mathematical induction.

For  $t = 1$ , applying Lemma 10, the following holds:

$$\begin{aligned} & \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(1) - \mathbf{z}(1)\| \geq \mu) \\ &= P(\|\mathbf{Z}^N(t+1) - Q(\mathbf{x})\mathbf{x}\| \geq \mu | \mathbf{Z}^N(t) = \mathbf{x}) \\ &\leq \frac{C(\mathbf{x})}{N} = \frac{C(1)}{N} \end{aligned} \quad (198)$$

and the desired result holds for  $t = 1$  by simply choosing  $C(1) = \frac{(m_1(\mathbf{x})+1+m_2(\mathbf{x})+1)^4}{\mu^2} [2p_1(1-p_1)+2p_2(1-p_2)]$ . Let us suppose that the statement holds for any  $t \geq 1$ . We investigate the property for  $t + 1$ . To that end, let us consider  $\nu < \mu$ . Therefore, according to Corollary 1, there exists  $\rho$  such that:

$$\|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \leq \rho \Rightarrow \|\mathbf{Q}(\mathbf{Z}^N(t))\mathbf{Z}^N(t) - \mathbf{Q}(\mathbf{z}(t))\mathbf{z}(t)\| \leq v \quad (199)$$

Bearing that in mind, we have that:

$$\begin{aligned} & \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu) \\ &= \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \rho) \\ & \quad \times \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \rho) \\ &+ \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \\ & \quad \times \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \\ &\leq^{(a)} \frac{C'(t)}{N} \\ & \quad + \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \end{aligned} \quad (200)$$

where (a) follows from  $\Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \rho) \leq 1$  and  $C'(t)$  being the constant related to the statement holding for  $t$  and for  $\rho$ . Next, we

tackle the second term of the inequality in (200):

$$\begin{aligned} & \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \\ &= \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - Q(\mathbf{Z}^N(t))\mathbf{Z}^N(t) \\ & \quad + Q(\mathbf{Z}^N(t))\mathbf{Z}^N(t) - \mathbf{z}(t+1)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \\ &\leq^{(a)} \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - Q(\mathbf{Z}^N(t))\mathbf{Z}^N(t)\| \\ & \quad + \|\mathbf{Q}(\mathbf{Z}^N(t))\mathbf{Z}^N(t) - \mathbf{Q}(\mathbf{z}(t))\mathbf{z}(t)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \\ &\leq^{(b)} \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - Q(\mathbf{Z}^N(t))\mathbf{Z}^N(t)\| \geq \mu - \nu \\ & \quad \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \\ &= \sum_{\substack{\mathbf{z} \in \Omega_{\rho}(\mathbf{z}(t)) \\ m_k(\mathbf{z}) \leq m_k(\mathbf{z}(t)) \\ k=1,2}} \left[ \Pr_{\mathbf{x}}(\mathbf{Z}^N(t) = \mathbf{z} \mid \mathbf{Z}^N(t) \in \Omega_{\rho}(\mathbf{z}(t))) \right. \\ & \quad \times \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - Q(\mathbf{z})\mathbf{z}\| \geq \mu - \nu \mid \mathbf{Z}^N(t) = \mathbf{z}) \left. \right] \\ & \quad + \sum_{\substack{\mathbf{z} \in \Omega_{\rho}(\mathbf{z}(t)) \\ m_1(\mathbf{z}) > m_1(\mathbf{z}(t)) \\ \text{or} \\ m_2(\mathbf{z}) > m_2(\mathbf{z}(t))}} \left[ \Pr_{\mathbf{x}}(\mathbf{Z}^N(t) = \mathbf{z} \mid \mathbf{Z}^N(t) \in \Omega_{\rho}(\mathbf{z}(t))) \right. \\ & \quad \times \Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - Q(\mathbf{z})\mathbf{z}\| \geq \mu - \nu \mid \mathbf{Z}^N(t) = \mathbf{z}) \left. \right] \end{aligned} \quad (201)$$

where (a) and (b) follows from the triangular inequality and the relationship in (199). One can notice that at any time slot  $t$ ,  $m_k(\mathbf{Z}^N(t)) \leq m_k(\mathbf{z}(t))$ . In light of that fact, the second term of the equation (201) is equal to 0. Bearing that in mind, We have for  $\mathbf{z} \in \Omega_{\rho}(\mathbf{z}(t))$  such that  $m_k(\mathbf{z}) \leq m_k(\mathbf{z}(t))$ :

$$\Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - Q(\mathbf{z})\mathbf{z}\| \geq \mu - \nu \mid \mathbf{Z}^N(t) = \mathbf{z}) \leq \frac{C_1(\mathbf{z}(t))}{N} \quad (202)$$

where  $C_1(t) = \frac{(m_1(\mathbf{z}(t))+m_2(\mathbf{z}(t))+2)^4}{(\mu-\nu)^2} [2p_1(1-p_1) + 2p_2(1-p_2)] = \frac{(m_1(t)+m_2(t)+2)^4}{(\mu-\nu)^2} [2p_1(1-p_1) + 2p_2(1-p_2)]$ . By substituting the above results in (201), we get:

$$\Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu \mid \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| < \rho) \leq \frac{C_1(t)}{N} \quad (203)$$

Combining this with (200), we can conclude that there exists a constant  $C(t+1)$  such that:

$$\Pr_{\mathbf{x}}(\|\mathbf{Z}^N(t+1) - \mathbf{z}(t+1)\| \geq \mu) \leq \frac{C(t+1)}{N} \quad (204)$$

which concludes our inductive proof. ■

Knowing that:

$$P_{\mathbf{x}}\left(\sup_{0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \mu\right) \leq \sum_{t=0}^{T-1} P_{\mathbf{x}}(\|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \mu)$$

Therefore, from Lemma 11, there exists a constant  $C$  which doesn't depend on  $N$  such that:

$$P_{\mathbf{x}}\left(\sup_{0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \mu\right) \leq \frac{C}{N}$$

Which concludes the proof.

APPENDIX M  
PROOF OF LEMMA 3

We show first of all that  $\mathbf{z}(t)$  converges to  $\mathbf{z}^*$  with respect to our considered norm, i.e.  $\lim_{t \rightarrow +\infty} \sum_{i=1}^{+\infty} |z_i^k(t) - z_i^{k,*}|i = 0$  for  $k = 1, 2$ . For that purpose, we use the limit inversion theorem which states that:

- If the series  $\sum_i f_i(t)$  is uniformly convergent on  $\mathbb{R}^+$
- If for each integer  $i$ ,  $f_i(t)$  admits a finite limit  $r_i$  when  $t$  tends to  $+\infty$ .

Therefore,  $\lim_{t \rightarrow +\infty} \sum_{i=1}^{+\infty} f_i(t) = \sum_{i=1}^{+\infty} \lim_{t \rightarrow +\infty} f_i(t) = \sum_{i=1}^{+\infty} r_i$ .

By letting  $f_i(t)$  denotes  $|z_i^k(t) - z_i^{k,*}|i$  for a given  $k$ , proving the result above is equivalent to establish that:

$$\lim_{t \rightarrow +\infty} \sum_{i=1}^{+\infty} |z_i^k(t) - z_i^{k,*}|i = \sum_{i=1}^{+\infty} \lim_{t \rightarrow +\infty} |z_i^k(t) - z_i^{k,*}|i$$

To that extent, we check if the aforementioned conditions are satisfied for this specific function  $f_i(t) = |z_i^k(t) - z_i^{k,*}|i$ .

- **Uniform convergence:** According to Weierstrass criterion,  $\sum_i f_i(t)$  is uniformly convergent if for each  $i$  the function  $f_i(t)$  is bounded by a constant  $c_i$  such that  $\sum_i c_i$  is convergent. Based on the proof of the Proposition 14, one can deduce that for large enough  $t$  denoted by  $t_l$ , the following induction relation always holds for  $t \geq t_l$  and  $i \geq l_{\max} + 1$ :

$$z_{i+1}^k(t+1) = p_k z_i^k(t)$$

That is, choosing  $t_0$  greater than  $t_l$ , and denoting by  $i_0 = m_k(t_0)$  the highest state of the vector  $\mathbf{z}(t_0)$  which is greater than  $l_{\max} + 1$ , we have that for each  $i > i_0$ :

$$z_i^k(t) = \begin{cases} 0 & \text{if } t_0 \leq t < t_0 + i - i_0 \\ p_k^{i-i_0} z_{i_0}^k(t - (i - i_0)) & \text{if } t \geq t_0 + i - i_0 \end{cases} \quad (205)$$

Based on the above equation, for each  $i > i_0$ ,  $z_i^k(t)$  is less than  $p_k^{i-i_0}$  for all  $t \geq t_0$ . To that extent, we investigate the evolution of the series of interest only when  $t \geq t_0$  (the limit inversion theorem still applicable since  $+\infty > t_0$ ). Moreover, we have that for all  $t \geq t_0$ :

$$\begin{aligned} \sum_i |z_i^k(t) - z_i^{k,*}|i &= \sum_{i=1}^{i_0} |z_i^k(t) - z_i^{k,*}|i + \sum_{i=i_0+1}^{+\infty} |z_i^k(t) - z_i^{k,*}|i \\ &\leq i_0^2 + \sum_{i=i_0+1}^{+\infty} (p_k^{i-i_0} i + z_i^{k,*} i) \end{aligned} \quad (206)$$

This last sum is known to be a finite sum since  $\sum_{i=1}^{+\infty} z_i^{k,*} i$  is the optimal average age of the relaxed problem for the class  $k$  which is finite, and  $\sum_{i=1}^{+\infty} p^i i$  is a finite sum for any  $0 \leq p < 1$ . Hence, the uniform convergence can be accordingly concluded.

- **Existence of the limit of  $f_i(t) = |z_i^k(t) - z_i^{k,*}|i$ :** According to the result of Proposition 14, we have  $\lim_{t \rightarrow +\infty} |z_i^k(t) - z_i^{k,*}|i = 0$  which is finite. Therefore, the second condition is satisfied.

Leveraging these findings, we can inverse the order between the limit and the sum. Subsequently:

$$\lim_{t \rightarrow +\infty} \sum_{i=1}^{+\infty} |z_i^k(t) - z_i^{k,*}|i = \sum_{i=1}^{+\infty} \lim_{t \rightarrow +\infty} |z_i^k(t) - z_i^{k,*}|i = 0$$

In other words, for  $k = 1, 2$ ,  $\sum_{i=1}^{+\infty} |z_i^k(t) - z_i^{k,*}|i$  tends to 0 when  $t$  grows. Consequently,  $\mathbf{z}(t)$  converges to  $\mathbf{z}^*$  with respect to our defined norm.

Therefore, for  $0 < \nu < \mu$ , there exists  $T_0$  such that for any  $t \geq T_0$ :

$$\|\mathbf{z}(t) - \mathbf{z}^*\| \leq \nu \quad (207)$$

By leveraging Proposition 15, we have:

$$\begin{aligned} &\Pr_{\mathbf{x}} \left( \sup_{T_0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}^*\| \geq \mu \right) \\ &\leq \Pr_{\mathbf{x}} \left( \sup_{T_0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| + \|\mathbf{z}(t) - \mathbf{z}^*\| \geq \mu \right) \\ &\leq \Pr_{\mathbf{x}} \left( \sup_{T_0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \mu - \nu \right) \\ &\leq \Pr_{\mathbf{x}} \left( \sup_{0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}(t)\| \geq \mu - \nu \right) \leq \frac{s}{N} \end{aligned} \quad (208)$$

which concludes the proof.

APPENDIX N  
PROOF OF PROPOSITION 16

We have that:

$$\begin{aligned} &\left| \frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} Z_i^{k,N}(t) i \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] - \sum_{k=1}^2 \sum_{i=1}^{+\infty} z_i^{k,*} i \right| \\ &= \left| \frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} (Z_i^{k,N}(t) i - z_i^{k,*} i) \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \\ &\leq \left| \frac{1}{T} \sum_{t=0}^{T_0-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i - z_i^{k,*} i \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \end{aligned} \quad (209)$$

$$+ \left| \frac{1}{T} \sum_{t=T_0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i - z_i^{k,*} i \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \quad (210)$$

We start by bounding (209). We have that:

$$\begin{aligned} &\left| \frac{1}{T} \sum_{t=0}^{T_0-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i - z_i^{k,*} i \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \\ &\leq \frac{1}{T} \sum_{t=0}^{T_0-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ |Z_i^{k,N}(t) i - z_i^{k,*} i| \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] \\ &\leq \frac{1}{T} \sum_{t=0}^{T_0-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] + \frac{1}{T} \sum_{t=0}^{T_0-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} z_i^{k,*} i \\ &= \frac{1}{T} \sum_{t=0}^{T_0-1} \sum_{k=1}^2 \sum_{i=1}^{\max\{m_1(t), m_2(t)\}} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i \middle| \mathbf{Z}^N(0) = \mathbf{x} \right] \\ &\quad + \frac{1}{T} \sum_{t=0}^{T_0-1} C^{RP} \end{aligned} \quad (211)$$

As  $m_k(\cdot)$  is increasing with  $t$ , then denoting  $m(t) = \max\{m_1(t), m_2(t)\}$ , we get:

$$\begin{aligned} & \frac{1}{T} \sum_{t=0}^{T_0-1} \sum_{k=1}^2 \sum_{i=1}^{\max\{m_1(t), m_2(t)\}} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i \mid \mathbf{Z}^N(0) = \mathbf{x} \right] \\ & + \frac{1}{T} \sum_{t=0}^{T_0-1} C^{RP} \leq \frac{(m(T_0) + C^{RP})T_0}{T} \end{aligned} \quad (212)$$

We denote  $Y_N$  the event  $\sup_{T_0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}^*\| \geq \mu$ , and we proceed to bound the second term (210).

$$\begin{aligned} & \left| \frac{1}{T} \sum_{t=T_0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i - z_i^{k,*} i \mid \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \\ & = P_{\mathbf{x}}(Y_N) \left| \frac{1}{T} \sum_{t=T_0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i - z_i^{k,*} i \mid Y_N, \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \\ & + (1 - P_{\mathbf{x}}(Y_N)) \left| \frac{1}{T} \sum_{t=T_0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i - z_i^{k,*} i \mid \overline{Y_N}, \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \\ & \stackrel{(a)}{\leq} \frac{(T - T_0)(m(T) + C^{RP})}{T} P_{\mathbf{x}}(Y_N) + (1 - P_{\mathbf{x}}(Y_N)) \mu \end{aligned} \quad (213)$$

where (a) results from:

$$\begin{aligned} & \left| \frac{1}{T} \sum_{t=T_0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} \mathbb{E}^{wi} \left[ Z_i^{k,N}(t) i - z_i^{k,*} i \mid \overline{Y_N}, \mathbf{Z}^N(0) = \mathbf{x} \right] \right| \\ & \leq \sup_{T_0 \leq t < T} \mathbb{E}^{wi} \left[ \sum_{k=1}^2 \sum_{i=1}^{+\infty} |Z_i^{k,N}(t) i - z_i^{k,*} i| \mid \overline{Y_N}, \mathbf{Z}^N(0) = \mathbf{x} \right] \\ & = \mathbb{E}^{wi} \left[ \sup_{T_0 \leq t < T} \|\mathbf{Z}^N(t) - \mathbf{z}^*\| \mid \overline{Y_N}, \mathbf{Z}^N(0) = \mathbf{x} \right] < \mu \end{aligned} \quad (214)$$

According to Lemma 3, we have  $\lim_{N \rightarrow \infty} P_{\mathbf{x}}(Y_N) = 0$ . Thus, combining the result (212) and (213), we obtain:

$$\begin{aligned} & \lim_{N \rightarrow \infty} \left| \frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} Z_i^{k,N}(t) i \mid \mathbf{Z}^N(0) = \mathbf{x} \right] - \sum_{k=1}^2 \sum_{i=1}^{+\infty} z_i^{k,*} i \right| \\ & \leq \frac{T_0(m(T_0) + C^{RP})}{T} + \mu \end{aligned} \quad (215)$$

This inequality is true for all  $\mu > 0$ , then:

$$\begin{aligned} & \lim_{N \rightarrow \infty} \left| \frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} Z_i^{k,N}(t) i \mid \mathbf{Z}^N(0) = \mathbf{x} \right] - \sum_{k=1}^2 \sum_{i=1}^{+\infty} z_i^{k,*} i \right| \\ & \leq \frac{T_0(m(T_0) + C^{RP})}{T} \end{aligned} \quad (216)$$

Finally we have:

$$\begin{aligned} & \lim_{T \rightarrow \infty} \lim_{N \rightarrow \infty} \left| \frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} Z_i^{k,N}(t) i \mid \mathbf{Z}^N(0) = \mathbf{x} \right] \right. \\ & \quad \left. - \sum_{k=1}^2 \sum_{i=1}^{+\infty} z_i^{k,*} i \right| = 0 \end{aligned} \quad (217)$$

As consequence:

$$\begin{aligned} & \lim_{T \rightarrow +\infty} \lim_{N \rightarrow \infty} \frac{1}{T} \mathbb{E}^{wi} \left[ \sum_{t=0}^{T-1} \sum_{k=1}^2 \sum_{i=1}^{+\infty} Z_i^{k,N}(t) i \mid \mathbf{Z}^N(0) = \mathbf{x} \right] \\ & = \sum_{k=1}^2 \sum_{i=1}^{+\infty} z_i^{k,*} i \end{aligned} \quad (218)$$