

## Lingering issues in distributed scheduling

***Citation for published version (APA):***

Simatos, F., Bouman, N., & Borst, S. C. (2013). *Lingering issues in distributed scheduling*. (arXiv.org; Vol. 1302.2824 [cs.NI]). s.n.

***Document status and date:***

Published: 01/01/2013

***Document Version:***

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

***Please check the document version of this publication:***

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

***General rights***

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

***Take down policy***

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# Lingering Issues in Distributed Scheduling

Florian Simatos, Niek Bouman, Sem Borst<sup>\*</sup>  
Eindhoven University of Technology  
Den Dolech 2  
5612 AZ Eindhoven, The Netherlands  
{f.simatos, n.bouman, s.c.borst}@tue.nl

## ABSTRACT

Recent advances have resulted in queue-based algorithms for medium access control which operate in a distributed fashion, and yet achieve the optimal throughput performance of centralized scheduling algorithms. However, fundamental performance bounds reveal that the “cautious” activation rules involved in establishing throughput optimality tend to produce extremely large delays, typically growing exponentially in  $1/(1-\rho)$ , with  $\rho$  the load of the system, in contrast to the usual linear growth.

Motivated by that issue, we explore to what extent more “aggressive” schemes can improve the delay performance. Our main finding is that aggressive activation rules induce a lingering effect, where individual nodes retain possession of a shared resource for excessive lengths of time even while a majority of other nodes idle. Using central limit theorem type arguments, we prove that the idleness induced by the lingering effect may cause the delays to grow with  $1/(1-\rho)$  at a quadratic rate. To the best of our knowledge, these are the first mathematical results illuminating the lingering effect and quantifying the performance impact.

In addition extensive simulation experiments are conducted to illustrate and validate the various analytical results.

## Categories and Subject Descriptors

G.3 [Probability and Statistics]: Markov processes

## General Terms

Algorithms, Performance, Theory

## Keywords

Delay performance, distributed scheduling algorithms, heavy traffic scaling, wireless networks

<sup>\*</sup>This work was supported by Microsoft Research through its PhD Scholarship Programme and a TOP grant from NWO.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMETRICS'13, June 17–21, 2013, Pittsburgh, PA, USA.  
Copyright 2013 ACM 978-1-4503-1900-3/13/06 ...\$15.00.

## 1. INTRODUCTION

As networks continue to grow in size and complexity, they increasingly rely on scheduling algorithms for efficient allocation of shared resources and arbitration between users. As a result, the design and analysis of scheduling algorithms for complex network scenarios has attracted significant attention over the last several years.

One of the centerpieces in the scheduling literature is the celebrated MaxWeight algorithm as proposed in the seminal work [18, 19]. The MaxWeight algorithm provides throughput optimality and maximum queue stability in a variety of scenarios, and has emerged as a powerful paradigm in cross-layer control and resource allocation problems [5].

While not strictly optimal in terms of delay performance, MaxWeight algorithms do achieve so-called equivalent workload minimization and offer favorable scaling characteristics in heavy traffic conditions [13, 16]. As a further key appealing feature, MaxWeight algorithms only need information on the queue lengths and instantaneous service rates, and do not rely on any explicit knowledge of the underlying system parameters.

On the downside, solving the maximum-weight problem tends to be challenging and potentially NP-hard. This is exacerbated in a network setting, where a centralized control entity may be lacking or require global state information, creating a substantial communication overhead in addition to the computational burden. This concern is especially pertinent as the maximum-weight problem needs to be solved at a high pace, commensurate with the fast time scale on which scheduling algorithms typically need to operate.

This issue has provided a strong impetus for devising algorithms that entail *lower computational complexity and communication overhead*, but retain the *maximum stability and throughput guarantees* of the MaxWeight algorithm. Various approaches in that direction were proposed in [2, 10, 14, 15, 17, 20]. An exciting breakthrough in this quest was recently achieved in the design of random back-off schemes for wireless medium access control that seem to offer the best of both worlds. These schemes operate in a *distributed fashion, requiring no centralized control entity or global state information*, and yet, remarkably, provide the *capability of achieving throughput optimality and maximum stability*. More specifically, clever algorithms have been developed for finding the back-off rates that yield any given target throughput vector in the capacity region [8, 9]. In the same spirit, powerful algorithms have been devised for adapting the back-off rates based on queue length information, and been shown to guarantee maximum stability [7, 11, 12].

While the maximum-stability guarantees for the above-mentioned algorithms have strong appeal, they do not extend to performance metrics such as expected queue lengths or delays. In fact, fundamental performance bounds [1] indicate that the “cautious” back-off functions involved in establishing maximum stability tend to produce extremely large delays, typically growing *exponentially* in  $1/(1-\rho)$ , with  $\rho$  the load of the system, in contrast to the usual *linear* growth. More specifically, the bounds show that the expected queue lengths grow as  $\psi^{-1}(1-\rho)$  as  $\rho \uparrow 1$ . Here  $\psi^{-1}$  represents the inverse of the (decreasing) function  $\psi$ , specifying the probability of a node entering a back-off as a function of its current queue length. The bounds may be explained by noting that the queue lengths govern the fraction of back-off time through the function  $\psi$ . Since the fraction of back-off time cannot exceed the surplus capacity in order for the system to be stable, however, it is ultimately the amount of surplus  $1-\rho$  that dictates the queue lengths through the function  $\psi^{-1}$ . We note that maximum stability has been established under the condition that the function  $\psi(a)$  decays (no faster than) *inverse-logarithmically* as  $a \rightarrow \infty$ , i.e.,  $\psi(a) \sim 1/\log a$ . This entails that  $\psi^{-1}(s)$  grows (no slower than) *exponentially* in  $1/s$  as  $s \downarrow 0$ , yielding the stated exponential growth of  $\psi^{-1}(1-\rho)$  in  $1/(1-\rho)$  as  $\rho \uparrow 1$ .

The above lower bounds suggest that the delay performance may be improved when the function  $\psi$  decays faster, e.g., *inverse-polynomially*:  $\psi(a) \sim a^{-\beta}$ , with  $\beta > 0$ , so that  $\psi^{-1}(s) \sim s^{-1/\beta}$  as  $s \downarrow 0$ . The larger the value of the exponent  $\beta$ , the slower the growth of  $\psi^{-1}(1-\rho) \sim (1-\rho)^{-1/\beta}$  as  $\rho \uparrow 1$ . In particular, it might seem plausible that for  $\beta \geq 1$ , the expected queue lengths will only exhibit the usual linear growth in  $1/(1-\rho)$  as  $\rho \uparrow 1$ . Note that a larger value of  $\beta$  means that a node is more “aggressive”, in the sense that it is less likely to enter a back-off and more inclined to hold on to the medium, and hence the coefficient  $\beta$  will be referred to as the aggressiveness parameter. It is worth observing that maximum stability for the above back-off functions is not guaranteed by existing results, which do not apply for any  $\beta > 0$ . In fact, for  $\beta > 1$ , maximum stability has been shown not to hold in certain topologies [6].

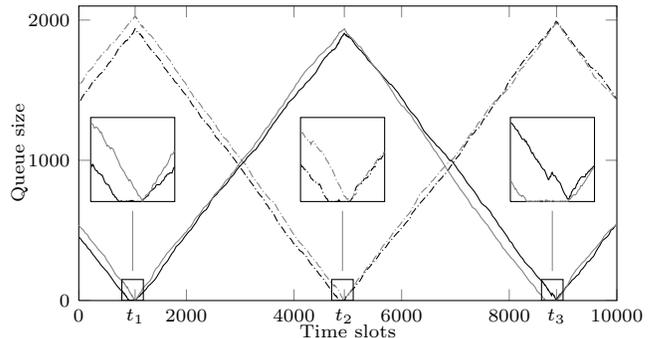
In the present paper we aim to gain fundamental insight whether a larger aggressiveness parameter can improve the delay performance. Our main finding is that for large values of  $\beta$ , a *lingering effect* can cause the mean stationary delay to increase in heavy traffic as  $1/(1-\rho)^2$ , and we focus on the simplest topology where this effect occurs. This topology, described in later sections, may seem at first sight rather restrictive in view of recent results [7, 11, 12] that apply to general topologies. We believe however that our results give insight into more general situations and discuss this in Section 7.

The remainder of the paper is organized as follows. We start in Section 2 by discussing an example that will explain the model that we consider as well as the lingering effect. In Section 3 we present a detailed model description. In Section 4 we provide an overview of the results and discuss the main performance implications. We conduct comprehensive simulation experiments to illustrate the various heavy traffic results in Section 5. In Section 6 we present a detailed asymptotic analysis and proof arguments, and we conclude in Section 7 by discussing the broader implications of our results.

## 2. ILLUSTRATIVE EXAMPLE

Consider a network consisting of four queues which are split into two groups, say groups 1 and 2, in such a way that if any queue of group 1 is transmitting a packet, no queue of group 2 may transmit. A group is said to be *active* if one of its queues is transmitting a packet, and a queue is said to be active if it belongs to the active group. The other group and queues are said to be *inactive*.

Active queues implement the following algorithm: after each transmission, each active queue flips a coin and *advertizes a release* with probability  $(1+a)^{-\beta}$ , with  $a$  the number of packets that this queue has to transmit and  $\beta > 0$ . If the two active queues advertise a release simultaneously, then active queues become inactive and vice-versa: such a time is called a *switching time*. This simple distributed algorithm gives rise to dynamics as illustrated in Figure 1, where the system is considered over three consecutive switching times  $t_1$ ,  $t_2$  and  $t_3$ . Between switching times, the backlog of active queues is drained while packets accumulate at inactive queues. The dynamics shown in Figure 1 are representative of the case  $\beta > 1$  where a switch does not occur until both active queues have emptied as will be established later.



**Figure 1: A typical sample path in the case  $\beta > 1$ . The three boxes zoom in to show the lingering effect. One queue hovers around zero while the other queue is yet to empty, resulting in an inefficient use of the resource.**

Let us now give a flavor of the lingering effect. Imagine that the two active queues start with initial queue lengths of the same order, say  $Q$ . As just mentioned, queues retain the shared resource until the time  $T^*$  at which the last active queue has emptied, thus preventing other queues from activating until this time. The law of large numbers suggests that  $T^*$  is of order  $Q$  (i.e., active queues are drained linearly as in Figure 1), but the central limit theorem suggests that up to time  $T^*$ , the first active queue to have emptied, while waiting for the other queue to empty, will be empty of the order of  $\sqrt{Q}$  units of time. This lingering effect is illustrated in Figure 1, which will be explained in greater detail in Section 5.

This  $(1/\sqrt{Q})$ -fraction of the time the shared resource is used inefficiently may at first sight seem negligible, and indeed queues seem to empty at the same time on the coarse time scale of Figure 1. However, we will actually establish that it has a significant impact in heavy traffic, causing queue lengths to grow at a rate  $1/(1-\rho)^2$  as  $\rho \uparrow 1$ , instead of the optimal  $1/(1-\rho)$ .

### 3. MODEL DESCRIPTION

#### 3.1 Informal description

Let us now give a more precise definition of our model. As mentioned in the introduction, in order to analyze the lingering effect in the simplest possible setting, we focus on a symmetric system consisting of two groups of  $R \geq 2$  queues. At any given point in time, one of the groups is *active* while the other is *inactive*.

Time is slotted and inactive queues have simple dynamics, driven by independent and identically distributed numbers of packet arrivals in each slot, so that they each simply grow according to random walks with step distribution denoted by  $\xi$ . During each time slot, active queues increase by independent amounts distributed as  $\xi$  as well, but, if at least one packet is present at the start of the slot, an active queue also flushes exactly one packet.

Moreover, at the end of each time slot, each active queue tosses a coin and advertizes a *momentary release* with probability  $\psi(a)$ , with  $a$  the number of packets in the queue at the end of the time slot. This (momentary) release gives inactive queues an opportunity to become active: if all the active queues *simultaneously* advertize a release, then inactive queues become active and active queues become inactive. Such a time is called a *switching time*. We will assume in the sequel that  $\psi(a) = (1+a)^{-\beta}$  for some parameter  $\beta > 0$ , called the aggressiveness parameter. In particular,  $\psi(a) \rightarrow 0$  as  $a \rightarrow +\infty$ , and so active queues are less likely to advertize a release when they are highly loaded; this mechanism thus gives priority to highly loaded queues in a distributed fashion. There is a cost associated with advertizing a release: each time an active queue advertizes a release and is not empty, it incurs an additional increase distributed according to some random variable  $\zeta$ .

The above-described model qualitatively resembles the canonical models for queue-based medium access control mechanisms. The main difference with continuous-time models in the literature is that in our model, back-off periods are infinitesimally short (hence, the releases are qualified as momentary), and the jump size  $\zeta$  represents the number of packets that would have arrived during a non-zero back-off period. The key advantage of our model is that it avoids the complication of requiring back-off periods to overlap to define a switching time. Moreover, simulation experiments show that our main results extend to models with non-infinitesimal back-off periods. We did not include these simulations to comply with page limitations.

**Remark.** The system may be interpreted as a set of  $R$  single-server two-queue polling systems, where servers are only allowed to switch between queues in a synchronized fashion. Such models with simultaneous service of several queues and synchronized switches are natural in applications (e.g., traffic lights at intersections), but appear not to have been considered.

#### 3.2 Parameters

The model described in the previous subsection is defined through four parameters: the number  $R \geq 2$  of queues in each group, two integer-valued random variables  $\xi, \zeta \in \mathbb{N} = \{0, 1, \dots\}$  and a real number  $\beta \in (0, \infty]$ , which defines the  $[0, 1]$ -valued sequence  $(\psi(a), a \geq 0)$  via  $\psi(a) = (1+a)^{-\beta}$  for  $a \geq 0$ , to be understood for  $\beta = +\infty$  as  $\psi(0) = 1$  and

$\psi(a) = 0$  for  $a > 0$ . Note that only the asymptotic behavior of  $\psi(\cdot)$  matters, and our results could easily be extended to any  $\psi(\cdot)$  with  $a^\beta \psi(a) \rightarrow \ell \in (0, \infty)$  as  $a \rightarrow +\infty$ .

We assume that  $\xi$  and  $\zeta$  have finite means, respectively  $\mathbb{E}\xi = \rho/2$  and  $\mathbb{E}\zeta = z$ , and that  $\xi$  has finite variance denoted by  $v = \mathbb{E}[(\xi - \rho/2)^2]$ . It will be argued that the system is stable if and only if  $\rho < 1$  and so we will refer to  $\rho$  as the load of the system (note that by symmetry, each queue is active half the time, which explains the factor 2 in the definition  $\rho = 2\mathbb{E}\xi$  of  $\rho$ ).

#### 3.3 Formal description

Because of the symmetry of the system, we do not need to label the queues individually, but only need to keep track of the state of active and inactive queues. We will consider the system embedded at switching times, and define  $Q_r^a(k)$  and  $Q_r^i(k)$  as the numbers of packets in the  $r$ th active and inactive queue, respectively, just after the  $k$ th switch occurred. We will be interested in the Markov chain  $(\mathbf{Q}(k), k \geq 0)$  which we also write as  $\mathbf{Q} = (\mathbf{Q}^a, \mathbf{Q}^i)$  with  $\mathbf{Q}^a = (Q_r^a, r \in \mathcal{R})$ ,  $\mathbf{Q}^i = (Q_r^i, r \in \mathcal{R})$ ,  $\mathcal{R} = \{1, \dots, R\}$  and we reserve in the sequel bold notation for vectors (of functions or numbers).

As informally described in Section 3.1, the dynamics of  $\mathbf{Q}$  in between two switching times are governed by two  $R$ -dimensional processes  $\mathbf{S} = (S_r, r \in \mathcal{R})$  and  $\mathbf{A} = (A_r, r \in \mathcal{R})$ :  $\mathbf{S}$  gives the increments of the inactive queues, while  $\mathbf{A}$  gives the state of the active queues. The dynamics are as follows:

- the  $2R$  processes  $A_r, S_r$  are independent;
- for each  $r \in \mathcal{R}$ ,  $(S_r(k), k \geq 0)$  is a random walk with step distribution  $\xi$ , started at 0;
- for each  $r \in \mathcal{R}$ ,  $(A_r(k), k \geq 0)$  is a space-inhomogeneous random walk with the following dynamics: for any  $a \in \mathbb{N}$  and any function  $f : \mathbb{N} \rightarrow [0, \infty)$ , we have

$$\begin{aligned} \mathbb{E}[f(A_r(1)) \mid A_r(0) = a] \\ = \mathbb{E}[f(Y(a) + \zeta \mathbf{1}_{\{Y(a) > 0, U < \psi(Y(a))\}})], \end{aligned} \quad (1)$$

where  $Y(a) = a + \xi - \mathbf{1}_{\{a > 0\}}$  is the number of packets at the end of the time slot and  $U$  is uniformly distributed in  $[0, 1]$ , and  $U, \xi$  and  $\zeta$  are independent.

The equation (1) describes the dynamics of an active queue and can be interpreted as follows. At each time slot, an active queue increases by  $\xi$  and, if not empty at the beginning of the time slot, flushes a packet, which brings the queue to state  $Y(a)$ . If  $U < \psi(Y(a))$ , we say that the active queue advertizes a release, which thus happens with the (conditional) probability  $\psi(Y(a))$  as described in Section 3.1. If at the end of the time slot the queue is not empty and it advertizes a release, i.e.,  $Y(a) > 0$  and  $U < \psi(Y(a))$ , then the active queue also incurs an additional increase by  $\zeta$ .

To define the  $2R$ -dimensional process  $\mathbf{Q} = (\mathbf{Q}^a, \mathbf{Q}^i)$  from  $\mathbf{S}$  and  $\mathbf{A}$ , it remains to adopt notation for the switching time, which we denote by  $T^*$ . Thus  $T^*$  is the first time at which all active queues advertize a release at the same time. Note that  $T^*$  and  $\mathbf{S}$  are independent. With these definitions, the dynamics of  $\mathbf{Q}$  as informally described in Section 3.1 obey the following equation: for any  $\mathbf{q} = (\mathbf{q}^a, \mathbf{q}^i) \in \mathbb{N}^R \times \mathbb{N}^R$  and any function  $f : \mathbb{N}^{2R} \rightarrow [0, \infty)$ ,

$$\begin{aligned} \mathbb{E}[f(\mathbf{Q}(1)) \mid \mathbf{Q}(0) = \mathbf{q}] \\ = \mathbb{E}[f(\mathbf{q}^i + \mathbf{S}(T^*), \mathbf{A}(T^*)) \mid \mathbf{A}(0) = \mathbf{q}^a]. \end{aligned} \quad (2)$$

The special case  $\beta = +\infty$  will be of particular importance. Indeed, we will show that it is representative of the system's behavior in the range  $\beta > 1$ . The case  $\beta = +\infty$  has been studied in [3], which refer to it as the random capture algorithm. When  $\beta = +\infty$ , active queues only advertize a release when they are empty (in which case there is no additional term  $\zeta$ ), and so  $\mathbf{A}(T^*) = \mathbf{0}$  and  $T^* = \inf \{k \geq 0 : \mathbf{A}(k) = \mathbf{0}\}$ .

### 3.4 Additional notation

In the remainder of the paper, and similarly as we have just done in (2), we will use the common symbol  $\mathbb{E}$  to denote expectation with respect to the laws of  $\mathbf{Q}$  and  $(\mathbf{A}, \mathbf{S})$ . Initial conditions will be denoted by a subscript, and it should always be clear from the context whether we consider initial conditions of  $\mathbf{Q}$ ,  $\mathbf{A}$  or some  $A_r$  (remember that  $\mathbf{S}(0)$  is always equal to  $\mathbf{0}$ ). For instance, (1) and (2) can be rewritten as follows:

$$\mathbb{E}_a [f(A_r(1))] = \mathbb{E} [f(Y(a) + \zeta \mathbf{I}_{\{Y(a) > 0, U < \psi(Y(a))\}})],$$

and

$$\mathbb{E}_{\mathbf{q}} [f(\mathbf{Q}(1))] = \mathbb{E}_{\mathbf{q}^a} [f(\mathbf{q}^i + \mathbf{S}(T^*), \mathbf{A}(T^*))].$$

The probability distributions corresponding to these various expectations are written as  $\mathbb{P}_a$ ,  $\mathbb{P}$ ,  $\mathbb{P}_{\mathbf{q}}$  and  $\mathbb{P}_{\mathbf{q}^a}$ . We also define  $\mathbb{P}_{\infty}$  with corresponding expectation  $\mathbb{E}_{\infty}$  as the laws of  $\mathbf{Q}$  and  $(\mathbf{A}, \mathbf{S})$  started in the stationary distribution of  $\mathbf{Q}$ , provided  $\mathbf{Q}$  is positive recurrent.

When  $\beta = +\infty$ , we see based on (2) that  $\mathbf{Q}^i(k) = \mathbf{0}$ , except maybe for  $k = 0$ . In particular, when  $\mathbf{Q}$  is positive recurrent,  $\mathbf{Q}^i(0)$  is  $\mathbb{P}_{\infty}$ -almost surely equal to  $\mathbf{0}$  and (2) therefore becomes

$$\mathbb{E}_{\infty} [f(\mathbf{Q}^a(0))] = \mathbb{E}_{\infty} [f(\mathbf{S}(T^*))] \quad (\beta = +\infty). \quad (3)$$

Further, in the remainder of the paper we let

$$\tau_r = \inf \{k \geq 0 : A_r(k) = 0\}, \quad \tau_{\max} = \max_{r \in \mathcal{R}} \tau_r,$$

and we define the  $\tau_{(r)}$ 's as the order statistics of the  $\tau_r$ 's, i.e.,  $\tau_{(1)} \leq \dots \leq \tau_{(R)}$  and  $\{\tau_{(r)}\} = \{\tau_r\}$ . Let finally  $|\cdot|$  be the  $L_{\infty}$  norm and  $\|\cdot\|$  be the  $L_1$  norm, i.e., if  $J \geq 1$  and  $\mathbf{x} \in \mathbb{R}^J$  then  $|\mathbf{x}| = \max_j |x_j|$  (which is just the absolute value for  $J = 1$ ) and  $\|\mathbf{x}\| = |x_1| + \dots + |x_J|$ .

## 4. OVERVIEW OF MAIN RESULTS

### 4.1 Two regimes

As stated in the introduction, we aim to gain fundamental insight in the impact of the function  $\psi(\cdot)$ , through the aggressiveness parameter  $\beta$ , on the system performance in terms of expected queue lengths and delays. We will demonstrate, based on a combination of heuristic arguments, simulation experiments and theoretical results, that the system's behavior changes as  $\beta$  increases from 0 to  $\infty$ . In the small  $\beta$  regime, typically  $\beta < 1/2$ , the system is characterized by a mean-reverting effect that induces (stationary) queue lengths of the order of  $1/(1-\rho)^{1/\beta}$ . This regime was already hinted upon in [1], where a corresponding lower bound was rigorously proved in a general setting. In the present paper we go one step further, by providing a heuristic argument in Section 4.2, explaining why this lower bound is sharp for

small  $\beta$ , and corroborating this by extensive simulation results in Section 5.2. However, when  $\beta$  increases, this mean-reverting effect vanishes, which causes the lower bound to become loose. For  $\beta > 1$ , we show that the system's performance is dominated by another phenomenon, which we call a lingering effect. The investigation of this latter effect constitutes the main contribution of the paper. We provide a heuristic explanation of this effect in Section 4.3, examine it via simulation experiments in Section 5.2, and prove various theoretical results in Section 6.

### 4.2 Small $\beta$ : a mean-reverting effect

Consider for a moment a given queue and let  $\Delta(a)$  be the mean increase of this queue starting from level  $a$ . Because of symmetry, the queue will be active half the time and inactive the other half of the time, so that

$$\Delta(a) \approx \frac{1}{2} \mathbb{E}_a [\text{increase} \mid \text{inactive}] + \frac{1}{2} \mathbb{E}_a [\text{increase} \mid \text{active}].$$

From the dynamics described in Section 3.3 we may deduce for large  $a$  (approximating  $Y(a) \approx a$  and neglecting the possibility of  $Y(a) = 0$ ) that  $2\Delta(a) \approx \mathbb{E}\xi + \mathbb{E}(\xi - 1 + \zeta \mathbf{I}_{\{U < \psi(a)\}})$ , which leads to

$$2\Delta(a) \approx \rho - 1 + z\psi(a).$$

This last approximation points to a mean-reverting effect: the average drift is negative for  $a > a^*$  and positive for  $a < a^*$ , where  $a^*$  is determined by the equation  $\rho - 1 + z\psi(a^*) = 0$ , which gives, since  $\psi(a) = (1+a)^{-\beta}$ ,  $a^* \approx 1/(1-\rho)^{1/\beta}$ . Using the convexity of  $\psi(\cdot)$  and Jensen's inequality, it is not difficult to convert the above back-of-the-envelope computations into a rigorous proof and show that  $1/(1-\rho)^{1/\beta}$  is a lower bound for the stationary mean number of packets in the system, as established in [1] in a continuous-time setting.

However, our explanation of this lower bound through this mean-reverting effect goes one step further, and shows that this lower bound should be sharp for small values of  $\beta$ . Indeed,  $\Delta(a)$  is the mean drift obtained by averaging over the active and inactive states. Roughly speaking, this quantity describes the drift experienced by a queue if the queue rapidly changes states, i.e., if the time needed for the queue to switch state is negligible compared to  $n$ , with  $n$  the length of the queue. Otherwise, the queue stays in the active or inactive state for a long period of time, during which it experiences (almost) constant drift, either positive or negative. This explanation bears some similarity with a random walk in a dynamic random environment, where the random walk essentially sees a constant environment if the random environment mixes rapidly.

To conclude our arguments, it remains to note that, at least intuitively, queues will switch more rapidly for smaller values of  $\beta$ . We will provide simulation experiments showing that the stationary mean number of packets grows as  $1/(1-\rho)^{1/\beta}$  for small values of  $\beta$ . As  $\beta$  increases however, this mean-reverting effect vanishes. The main contribution of this paper is to reveal and quantify a previously unknown effect, called the lingering effect, which explains the system's behavior in this case.

### 4.3 Large $\beta$ : the effect of lingering

As described in the introduction, when the function  $\psi(\cdot)$  decays sufficiently fast, once a queue gains possession of the

resource, it holds onto it, even when some or all of the other queues in the same group are empty, and it would be more efficient for the queues in the other group to receive the resource. This causes a lingering effect as illustrated in Figure 1 for a scenario with  $R = 2$ ,  $\beta = 2$ . It may appear that the two queues in the same group drain around the same time (as can indeed be shown to be the case on a “fluid scale”). When we zoom in, however, we see that there is actually a time period where one of the queues is already empty, while the other one clings to the resource and prevents the two queues in the other group from activating.

Let us now provide a somewhat more technical description of this phenomenon in the case  $\beta = +\infty$ , where it is easiest to see thanks to the simplification of the dynamics (3). Moreover, we will argue that the case  $\beta \in (1, +\infty)$  is essentially a perturbation of the case  $\beta = +\infty$ .

By applying (3) to  $f = \|\cdot\|$ , we obtain  $\mathbb{E}_\infty(\|\mathbf{Q}^a(0)\|) = \mathbb{E}_\infty(\|\mathbf{S}(T^*)\|)$  and since  $\|\mathbf{S}(\cdot)\|$  is a random walk with drift  $R\mathbb{E}\xi$  independent of  $T^*$  and  $\mathbf{A}(0) = \mathbf{Q}^a(0)$  (by definition), we obtain  $\mathbb{E}_\infty(\|\mathbf{A}(0)\|) = R\mathbb{E}(\xi)\mathbb{E}_\infty(T^*)$  and so by symmetry,

$$\mathbb{E}_\infty(A_1(0)) = \mathbb{E}(\xi)\mathbb{E}_\infty(T^*). \quad (4)$$

The goal is now to relate  $\mathbb{E}_\infty(T^*)$  to  $\mathbb{E}_\infty(A_1(0))$ . Remember that  $T^* = \inf\{k : \mathbf{A}(k) = \mathbf{0}\}$  when  $\beta = +\infty$ . It is not difficult to show that  $T^* \approx \tau_{\max}$ , essentially because once all queues have hit 0, it is only a matter of constant time for all queues to be empty simultaneously (this will be justified in Lemma 6.1). Then, the central limit theorem shows that  $\tau_r \approx A_r(0)/(1 - \mathbb{E}\xi) + A_r(0)^{1/2}$  (where we neglect multiplicative constants, possibly random, appearing in front of first- or second-order terms and that do not influence the order of magnitude of the final result), which leads to the approximation  $\tau_{\max} \approx |\mathbf{A}(0)|/(1 - \mathbb{E}\xi) + |\mathbf{A}(0)|^{1/2}$ . Since under  $\mathbb{P}_\infty$  queues are symmetric, we have  $|\mathbf{A}(0)| \approx A_1(0) + A_1(0)^{1/2}$  which finally leads to  $T^* \approx A_1(0)/(1 - \mathbb{E}\xi) + A_1(0)^{1/2}$ , i.e.,

$$\mathbb{E}_\infty(A_1(0)) \approx \frac{\mathbb{E}\xi}{1 - \mathbb{E}\xi} \mathbb{E}_\infty(A_1(0)) + \mathbb{E}_\infty[A_1(0)^{1/2}].$$

Thus upon a concentration-like result of the kind

$$\mathbb{E}_\infty[A_1(0)^{1/2}] \approx [\mathbb{E}_\infty(A_1(0))]^{1/2}$$

it is reasonable to expect

$$\left(1 - \frac{\mathbb{E}\xi}{1 - \mathbb{E}\xi}\right) \mathbb{E}_\infty(A_1(0)) \approx [\mathbb{E}_\infty(A_1(0))]^{1/2}.$$

Since  $1 - \mathbb{E}(\xi)/(1 - \mathbb{E}\xi) \approx 1 - \rho$  this shows that  $\mathbb{E}_\infty(A_1(0))$ , and hence  $\mathbb{E}_\infty(\|\mathbf{Q}(0)\|)$ , should grow as  $1/(1 - \rho)^2$ . While admittedly crude, the above heuristic arguments provide the correct estimates, and serve as a useful guide for a rigorous proof in Section 6.1.

As reflected in the above computations, the square factor really stems from the relation  $T^* \approx \tau_{(1)} + |\mathbf{A}(0)|^{1/2}$ , i.e.,  $T^*$  occurs somehow long after  $\tau_{(1)}$ , the time at which it would be optimal to switch in order to avoid inefficient use of the resource. But it is difficult to make the system switch exactly at  $\tau_{(1)}$  in a distributed fashion, and here the penalty incurred is a square root. Interestingly, the penalty may seem negligible, since it is only a square root, but this small inefficiency has a significant impact in heavy traffic.

We explain in the next subsection how we formalize the analysis of this square root effect.

## 4.4 Main results

In this subsection we present our main results. They are discussed in the following sections based on a combination of rigorous proofs, heuristic arguments and simulation results, and we explain our contributions in more detail at the end of this subsection. Our first main result introduces the notion of *scaling exponent*, which essentially determines the polynomial growth rate of the mean number of packets in stationarity.

**MAIN RESULT 1.**  *$\mathbf{Q}$  is positive recurrent if  $\rho < 1$  and transient if  $\rho > 1$ . Moreover, the limit*

$$\lim_{\rho \uparrow 1} \left( \frac{\log \mathbb{E}_\infty(\|\mathbf{Q}(0)\|)}{\log(1/(1 - \rho))} \right) \stackrel{\text{def.}}{=} \alpha \quad (5)$$

*exists and is called scaling exponent.*

**Remark.** The transience of  $\mathbf{Q}$  when  $\rho > 1$  is easy to see. Indeed, if  $L(k)$  is the number of packets at the beginning of the  $k$ th time slot,  $L$  is lower bounded by a random walk with drift  $R(\rho - 1)$ . Thus when  $\rho > 1$ , we have  $L(k) \rightarrow +\infty$  as  $k \rightarrow +\infty$  and since  $(\|\mathbf{Q}(k)\|, k \geq 0)$  is a subsequence of  $(L(k), k \geq 0)$  this proves the transience of  $\mathbf{Q}$ .

Stability for  $\rho < 1$  is very intuitive but more challenging to establish. If the active queues are in state  $\mathbf{a} = (a_r)$  with  $a_r > 0$  for every  $r$ , the variation of the mean number of packets in the system over the next time slot is equal to

$$-R \left( 1 - \rho - z \sum_{r=1}^R \psi(a_r) \right).$$

Since  $\psi(a) \rightarrow 0$  as  $a \rightarrow +\infty$ , the drift is negative, close to  $-R(1 - \rho)$ , when each  $a_r$  is large enough. The problem in formalizing this argument is twofold: first, in order to prove stability, one must be able to control every possible initial configuration, not only those where every  $a_r$  is large; second, this argument considers the system on the normal time scale, whereas we are interested in the system embedded at switching times.

Our interest in the scaling exponent comes from an expected polynomial growth of  $\mathbb{E}_\infty(\|\mathbf{Q}(0)\|)$  in heavy traffic. In general, we expect as  $\rho \uparrow 1$  a behavior of the kind

$$\mathbb{E}_\infty(\|\mathbf{Q}(0)\|) \approx \frac{C}{(1 - \rho)^\alpha} \quad (6)$$

for some finite constant  $C > 0$  (note that this would be stronger than (5)). The scaling exponent depends on the four model parameters  $R$ ,  $\xi$ ,  $\zeta$  and  $\beta$ . However, we will mostly be interested in the dependence of  $\alpha$  on  $\beta$  and thus write  $\alpha(\beta)$  when the other three parameters are kept fixed. In fact, our results suggest that  $\alpha$  does not depend on these other parameters, at least in the two extreme regimes we studied in detail. As explained in Section 4.3, the lingering makes  $\mathbb{E}_\infty(\|\mathbf{Q}(0)\|)$  grow as  $1/(1 - \rho)^2$  when  $\beta > 1$ , which can be formalized as follows.

**MAIN RESULT 2.** *If  $\beta > 1$ , then  $\alpha(\beta) = 2$ .*

It seems quite challenging to prove the existence of, and find a closed-form expression for,  $\alpha(\beta)$  when  $\beta \leq 1$ . Nonetheless, the heuristic arguments of Section 4.2 suggest that the lower bound  $\alpha(\beta) \geq 1/\beta$  of [1] is sharp as  $\beta \downarrow 0$ , which leads to the following result.

MAIN RESULT 3.  $\beta\alpha(\beta) \rightarrow 1$  as  $\beta \rightarrow 0$ .

We now explain in more detail how these three results are established in the rest of the paper. The case  $\beta = +\infty$  is treated rigorously: positive recurrence of  $\mathbf{Q}$  when  $\rho < 1$  and the result that  $\alpha(\infty) = 2$  are proved in Section 6.1. For  $\beta > 1$ , we explain in Section 6.2 why it can be seen as a perturbation of the case  $\beta = +\infty$ : we give some partial results toward a full proof in Sections 6.2.1 to 6.2.3, and heuristic arguments in Section 6.2.4 explaining the technical steps missing for a complete proof. These results are backed up by simulation results in Section 5. Finally, the main result 3 is discussed based on the simulation results of Section 5.2, which back up the heuristic arguments of Section 4.2.

## 5. SIMULATION EXPERIMENTS

In the previous section we provided an overview of the main results characterizing the heavy traffic behavior of the expected queue lengths. Before presenting detailed proof arguments in the case  $\beta > 1$  in the next section, we first discuss comprehensive simulation experiments that we conducted to illustrate the stated growth behavior.

The detailed asymptotic analysis and proofs in the next section will reveal that the value of  $R$  and the precise distributions of  $\xi$  and  $\zeta$  do not affect the stability of the system or the value of the scaling exponent. Throughout this section we therefore focus on the case where  $\xi$  is geometrically distributed with parameter  $2/(2+\rho)$  and  $\mathbb{P}(\zeta = 1) = 1$ . We ran simulations using different distributions for  $\xi$  and  $\zeta$  as well, including extreme cases such as distributions with infinite third moment (even infinite second moment for  $\zeta$ ). Because of page limitations we do not include these cases, but they yielded very similar results.

### 5.1 Simulations for the main result 1

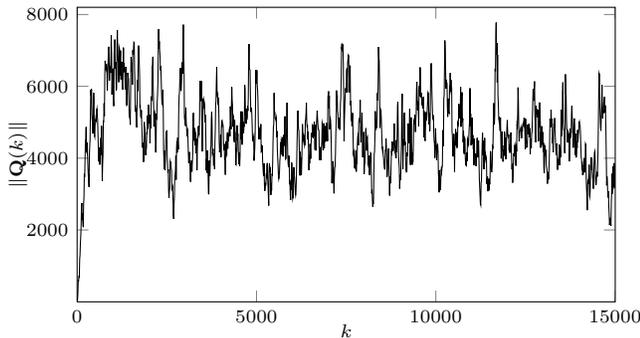


Figure 2:  $\|\mathbf{Q}(k)\|$  vs.  $k$  for  $R = \beta = 2$  and  $\rho = 0.99$ .

We now examine the case  $R = \beta = 2$  in detail. Figures 2 and 3 show the evolution of  $\|\mathbf{Q}(k)\|$ , starting at  $\mathbf{Q}(0) = \mathbf{0}$ , for  $\rho = 0.99$  and  $\rho = 1.01$  respectively. When  $\rho = 0.99$ , Figure 2 shows that  $\|\mathbf{Q}(k)\|$  fluctuates between 2000 and 8000 for  $k$  large enough, which strongly suggests that  $\mathbf{Q}$  is positive recurrent.

When  $\rho = 1.01$ , Figure 3 shows that  $\|\mathbf{Q}(k)\|$  increases until we stop the simulation when  $1.5 \cdot 10^6$  packets are present in the system. Note that for a transient system we would expect that, when the queues are large, the total queue size grows by a constant amount on average in every time slot

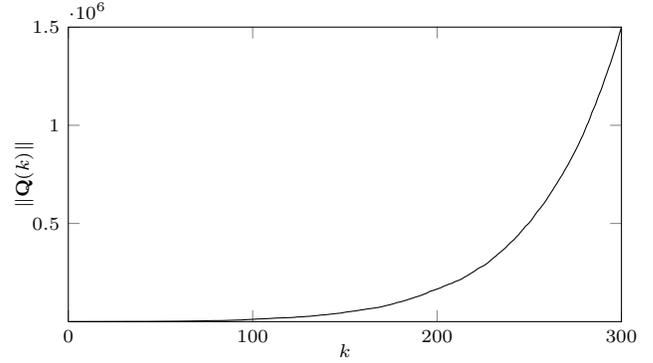


Figure 3:  $\|\mathbf{Q}(k)\|$  vs.  $k$  for  $R = \beta = 2$  and  $\rho = 1.01$ .

and that the time between two consecutive switching times increases. This explains the super-linear growth of  $\|\mathbf{Q}(k)\|$  as we consider the system at switching times. In fact, the reasoning in Section 4.3 suggests that when  $\|\mathbf{Q}(k)\|$  is large we have  $\|\mathbf{Q}(k+1)\|/\|\mathbf{Q}(k)\| \approx \mathbb{E}\xi/(1 - \mathbb{E}\xi)$ . This gives a heuristic explanation for the exponential growth (at rate  $\rho/(2 - \rho) > 1$ ) observed in Figure 3. All in all, this suggests that  $\mathbf{Q}$  is transient for  $\rho = 1.01$ .

Observe moreover that it may in general be difficult to distinguish between positive recurrent and transient systems based on simulation results. Here however,  $\mathbf{Q}$  obeys two clearly distinguishable types of behavior: stochastic fluctuations when  $\rho < 1$ , and almost deterministic exponential growth when  $\rho > 1$ .

### 5.2 Simulations for the main results 2 and 3

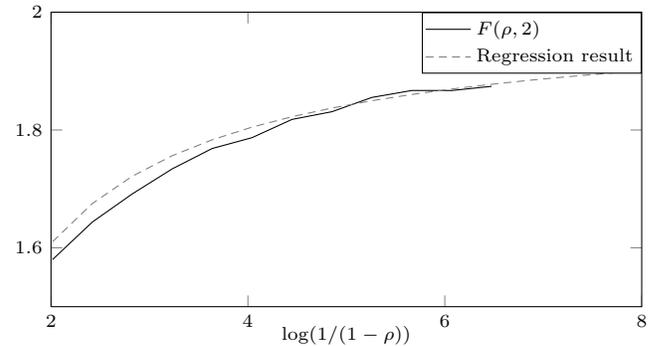


Figure 4: Approximating  $\alpha(2)$  for  $R = 2$ .

From the results in Figure 2 we find by averaging over time  $\mathbb{E}_\infty(\|\mathbf{Q}(0)\|) \approx 4700$  for  $\rho = 0.99$ , corresponding in view of the definition (5) of the scaling exponent to the estimate  $\alpha(2) \approx \log 4700 / \log 100 \approx 1.84$ . To simplify the discussion define in the sequel

$$F(\rho, \beta) = \frac{\log \mathbb{E}_\infty(\|\mathbf{Q}(0)\|)}{\log(1/(1-\rho))}, \quad \beta \in (0, \infty], \rho < 1. \quad (7)$$

Then, the scaling exponent  $\alpha(\beta)$  is defined in (5) via a limiting procedure, namely  $\alpha(\beta) = F(1-, \beta) = \lim_{\rho \uparrow 1} F(\rho, \beta)$ . Using the results of Figure 2 to estimate  $\alpha(2)$  amounts to using the approximation  $F(1-, 2) \approx F(0.99, 2)$ . In order to check whether this approximation is valid, we performed the same simulation for different values of  $\rho$ . In Figure 4

the value of  $F(\rho, 2)$  is plotted for different values of  $\rho \in (0.87, 0.999)$ .  $F(\rho, 2)$  is plotted versus  $\log(1/(1-\rho))$  in order to “dilate” time around the value  $\rho = 1$  that we are interested in (the quantity  $\beta\alpha(\beta)$  is plotted versus  $1/\beta$  in Figure 6 for the same reason), and also because it is natural to regress  $F(\rho, 2)$ , as function of  $\rho$ , against  $\log(1/(1-\rho))$  (see forthcoming discussion).

Figure 4 shows that the limit  $F(1-, 2)$  seems to exist, but  $F(\rho, 2)$  is still significantly increasing for  $\rho = 0.999$ . Thus  $F(0.999, 2)$ , and in particular  $F(0.99, 2)$ , cannot be used as an estimate of  $\alpha(2)$ . It is numerically difficult to run a simulation for even higher values of  $\rho$ , and so to circumvent this problem we use the simulation results displayed in Figure 4 to find the asymptotic value of  $F(\rho, 2)$ . To do so, we use the approximation (6) to infer the form of  $F(\rho, \beta)$ , namely

$$F(\rho, \beta) \approx \alpha(\beta) + \frac{\log C}{\log(1/(1-\rho))}, \quad (8)$$

which suggests, as mentioned above, to regress  $F(\rho, 2)$  against  $a + b/x$  in the scale  $x = \log(1/(1-\rho))$ . We performed this regression for the curve displayed in Figure 4 and found an optimal value  $a = 1.9984 \approx \alpha(2)$ .

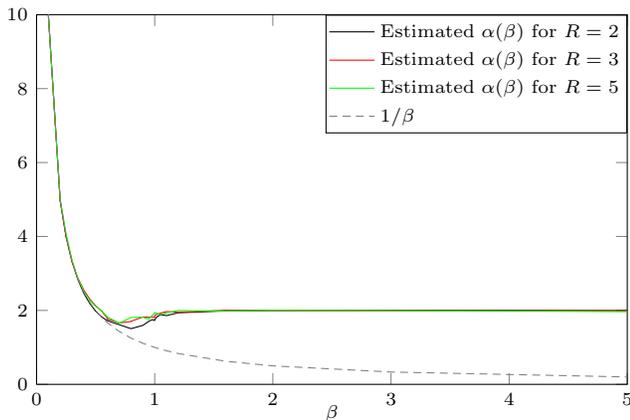


Figure 5:  $\alpha(\beta)$  vs.  $\beta$  for  $R \in \{2, 3, 5\}$ .

Applying the same approach, we can find an estimate for  $\alpha(\beta)$  for any value of  $\beta$ . The results for  $R = 2, 3$  and  $5$  are given in Figure 5, and confirm that  $R$  does not seem to influence  $\alpha(\beta)$ . Further, the approximation  $\alpha(\beta) \approx 2$  appears to be very good for any  $\beta > 1.2$ , namely, the estimated value of  $\alpha(\beta)$  is at most 3% away from 2 for any  $\beta > 1.2$  and any  $R = 2, 3, 5$ , in very good agreement with the main result 2. In Section 6.2.4 we will discuss in greater detail the case where  $\beta > 1$  is close to 1.

Finally, for small values of  $\beta$ , we observe that  $\alpha(\beta)$  is close to the lower bound  $1/\beta$ . To see how close  $\alpha(\beta)$  is to the lower bound, we made a plot for  $\beta\alpha(\beta)$  in Figure 6. We observe that  $\beta\alpha(\beta) \in (0.98, 1.02)$  for any  $\beta < 0.45$ , which supports our claim made in the main result 3 and the heuristic argument given in Section 4.2.

## 6. ANALYSIS OF THE CASE $\beta \in (1, \infty]$

### 6.1 The case $\beta = +\infty$

We study here in detail the case  $\beta = \infty$ . In Section 6.2 we explain why the case  $\beta > 1$  can be seen as a perturbation of

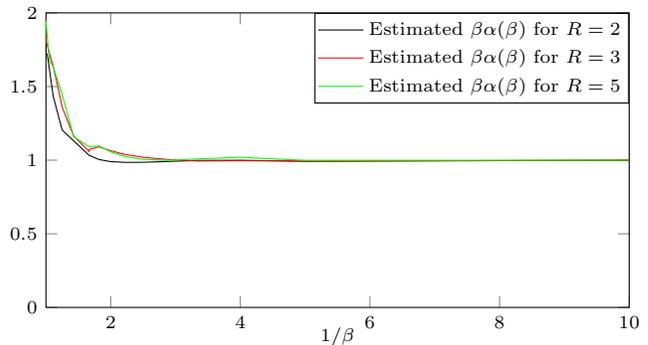


Figure 6:  $\beta\alpha(\beta)$  vs.  $1/\beta$  for  $R \in \{2, 3, 5\}$ .

this case. In the remainder of this subsection we assume that  $\beta = \infty$ , so that  $\mathbf{A}(T^*) = \mathbf{0}$  and  $T^* = \inf\{k \geq 0 : \mathbf{A}(k) = \mathbf{0}\}$ .

#### 6.1.1 Control of $T^*$

We first prove that  $T^* \approx \tau_{\max}$ , i.e., the time at which the  $R$  independent random walks  $A_r$  simultaneously hit 0, is close to the first time at which each process has visited 0 at least once.

LEMMA 6.1.  $\sup_{\mathbf{a}, \rho} \mathbb{E}_{\mathbf{a}}(T^* - \tau_{\max})$  is finite, where the supremum is taken over  $\mathbf{a} \in \mathbb{N}^R$  and  $\rho \leq 1$ .

PROOF. By monotonicity of  $T^*$  in  $\rho$ , it is enough to show that  $\sup_{\mathbf{a}} \mathbb{E}_{\mathbf{a}}(T^* - \tau_{\max})$  is finite for  $\rho = 1$ . So assume in the remainder of the proof  $\rho = 1$ : then each active queue is stable (for that we only need  $\rho < 2$ ) and in particular,  $T^*$  is almost surely finite. Thus the strong Markov property at time  $\tau_{\max} \leq T^*$  gives

$$\mathbb{E}_{\mathbf{a}}(T^* - \tau_{\max}) = \mathbb{E}_{\mathbf{a}}[\mathbb{E}_{\mathbf{A}(\tau_{\max})}(T^*)].$$

When  $A_r$  hits 0 for the first time at time  $\tau_r$ , we can couple it with a stationary version  $\hat{A}_r$  of  $A_1$  in such a way that  $\hat{A}_r(k) \geq A_r(k)$  for every  $k \geq \tau_r$  (this is the usual coupling between two processes starting from different initial conditions and sharing the same stochastic primitives). By monotonicity and since queues are independent, we obtain

$$\mathbb{E}_{\mathbf{a}}(T^* - \tau_{\max}) \leq \mathbb{E}_{\varrho}(T^*),$$

where for any  $r \geq 2$ , we define  $\varrho = (X_1, \dots, X_R) \in \mathbb{N}^R$  with  $(X_i, i \geq 1)$  i.i.d. with common distribution the stationary distribution  $X$  of  $A_r$ .

Define  $\sigma(0) = 0$  and for  $k \geq 1$ ,  $\tau(k) = \tau_{\max} \circ \theta_{\sigma(k-1)}$  and  $\sigma(k) = \sigma(k-1) + \tau(k)$ , where  $\theta$  is the shift operator. In words,  $\sigma(k+1)$  is the smallest time after  $\sigma(k)$  such that every queue will have visited 0 at least once after (and including) time  $\sigma(k)$ . Note that if  $\mathbf{A}(\sigma(k)) = \mathbf{0}$  for some  $k$ , then  $\sigma(k') = \sigma(k)$  for  $k' \geq k$ . In particular, the limit  $\sigma(\infty)$  as  $k \rightarrow +\infty$  exists, and is by construction equal to  $T^*$ . Since  $\tau(k) = \sigma(k) - \sigma(k-1)$  this can be rewritten as  $T^* = \sum_{k \geq 1} \tau(k)$ .

By construction,  $A_r$  takes the value 0 at least once between  $\sigma(k)$  and  $\sigma(k+1)$ : at this time we can couple it with a stationary version of itself that stays above it. Following this stationary process from this time until the time  $\sigma(k+1)$ , this gives the existence of i.i.d. random variables  $(X_r(k), r \in \mathcal{R}, k \geq 0)$  with common distribution  $X$  such that  $A_r(\sigma(k)) \leq X_r(k)$  for every  $r$  and  $k$ .

Considering the time needed for every queue to visit 0 at least once starting from  $(X_r(k), r \in \mathcal{R})$  gives an upper bound on  $\tau(k)$ . Each time we do so, there is a probability  $p = \mathbb{P}_\rho(\mathbf{A}(\tau_{\max}) = \mathbf{0}) \in (0, 1)$  that we reach  $\mathbf{0}$  at time  $\tau_{\max}$ . Thus,  $T^*$  is (stochastically) upper bounded by  $\tau'_1 + \dots + \tau'_G$  with  $G$  a geometric random variable with parameter  $1 - p$  and  $(\tau'_i)$  i.i.d., independent from  $G$  and with common distribution  $\tau_{\max}$  under  $\mathbb{P}_\rho$ . This implies

$$\mathbb{E}_\rho(T^*) \leq \mathbb{E}(\tau'_1 + \dots + \tau'_G) = \frac{\mathbb{E}_\rho(\tau_{\max})}{\mathbb{P}_\rho(\mathbf{A}(\tau_{\max}) = \mathbf{0})}.$$

Since  $\tau_{\max} \leq \tau_1 + \dots + \tau_R$ ,  $\mathbb{E}_{a_r}(\tau_r) = a_r/(1 - \mathbb{E}\xi)$  and  $X$  has finite mean (because  $\xi$  is assumed to have finite second moment),  $\mathbb{E}_\rho(\tau_{\max})$  is finite and so the proof is complete.  $\square$

The previous lemma justifies the approximation  $T^* \approx \tau_{\max}$ , and to further control  $\tau_{\max}$ , we will use the fact that  $(\tau_r, r \in \mathcal{R})$  under  $\mathbb{P}_\mathbf{a}$  is equal in distribution to  $(V_r(a_r), r \in \mathcal{R})$ , where  $(V_r, r \in \mathcal{R})$  are i.i.d. random walks started at 0 and with step distribution  $\delta$ , equal in distribution to  $\tau_1$  under  $\mathbb{P}_1$  (i.e.,  $\delta$  is the time needed for a random walk with step distribution  $\xi - 1$  to go from 1 to 0). Then  $\delta$  has finite mean  $1/(1 - \mathbb{E}\xi)$  and also finite variance, which we denote by  $\nu$ . To control the maximum of random walks, we will use the following result.

LEMMA 6.2. *If  $(W_r)$  are i.i.d. random walks started at 0 with step distribution having mean  $m$  and variance  $w$ , then for any  $\mathbf{x} = (x_r) \in \mathbb{N}^R$  it holds that*

$$\mathbb{E}\left(\max_r W_r(x_r)\right) \leq m|\mathbf{x}| + R(w|\mathbf{x}|)^{1/2}.$$

PROOF. Defining  $Y_r = (W_r(x_r) - mx_r)/(wx_r)^{1/2}$ , we have  $\max_r W_r(x_r) = \max_r(mx_r + (wx_r)^{1/2}Y_r)$  so that

$$\max_r W_r(x_r) \leq m|\mathbf{x}| + (w|\mathbf{x}|)^{1/2}\|\mathbf{Y}\|$$

which proves the result since  $\mathbb{E}(\|\mathbf{Y}\|) = \sum_r \mathbb{E}(|Y_r|) \leq R$ .  $\square$

### 6.1.2 Proof of $\alpha(\infty) = 2$

Before proving the main result 2 in the case  $\beta = +\infty$ , we first need to prove stability and finiteness of the stationary mean.

PROPOSITION 6.3. *Assume that  $\beta = +\infty$  and that  $\rho < 1$ . Then  $\mathbf{Q}$  is positive recurrent and  $\mathbb{E}_\infty(\|\mathbf{Q}(0)\|) < +\infty$ .*

PROOF. Let  $\Phi(\mathbf{q}) = |\mathbf{q}^a| + |\mathbf{q}^i|$ : to prove Proposition 6.3, it is enough to prove that

$$\lim_{K \rightarrow +\infty} \sup_{\mathbf{q}: \Phi(\mathbf{q}) \geq K} \left( \frac{\mathbb{E}_\mathbf{q}[\Phi(\mathbf{Q}(2)) - \Phi(\mathbf{q})]}{\Phi(\mathbf{q})} \right) < 0. \quad (9)$$

Indeed, this shows that  $\Phi$  is a Lyapunov function, which implies positive recurrence of  $\mathbf{Q}$  using for instance the Foster-Lyapunov criterion. But it shows more than that: in the terminology of [4] it implies that  $\Phi$  is a geometric Lyapunov function, and Theorem 5 in [4] shows that (9) implies that  $\mathbb{E}_\infty[\Phi(\mathbf{Q}(0))]$ , and in particular  $\mathbb{E}_\infty(\|\mathbf{Q}(0)\|)$ , is finite. Thus we only have to prove (9).

Since  $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$ , (2) implies that

$$\mathbb{E}_\mathbf{q}[\Phi(\mathbf{Q}(1))] \leq |\mathbf{q}^i| + \mathbb{E}_{\mathbf{q}^a}(|\mathbf{S}(T^*)|) + \mathbb{E}_{\mathbf{q}^a}(|\mathbf{A}(T^*)|).$$

Since  $\mathbf{S}$  and  $T^*$  are independent, Lemma 6.2 gives that  $\mathbb{E}_\mathbf{a}(|\mathbf{S}(T^*)|) \leq \mathbb{E}(\xi)\mathbb{E}_\mathbf{a}(T^*) + R\mathbb{E}_\mathbf{a}((vT^*)^{1/2})$  for any  $\mathbf{a} \in \mathbb{N}^R$

(recall that  $v$  is the variance of  $\xi$ , assumed to be finite). Thus after rearranging the terms, we end up with the bound

$$\mathbb{E}_\mathbf{q}[\Phi(\mathbf{Q}(1)) - \Phi(\mathbf{q})] \leq -(1 - \rho)\mathbb{E}_{\mathbf{q}^a}(T^*) + \Psi(\mathbf{q}^a), \quad (10)$$

where

$$\Psi(\mathbf{a}) = \mathbb{E}_\mathbf{a}\left(|\mathbf{A}(T^*)| + (1 - \mathbb{E}\xi)T^* - |\mathbf{A}(0)| + R(vT^*)^{1/2}\right).$$

We now argue that  $\Psi(\mathbf{a}) \leq c[\mathbb{E}_\mathbf{a}(T^*)]^{1/2}$  for some finite constant  $c$  independent of  $\mathbf{a}$ . Since we are considering the case  $\beta = +\infty$ , we have  $|\mathbf{A}(T^*)| = 0$  and so we only have to show that  $(1 - \mathbb{E}\xi)\mathbb{E}_\mathbf{a}(T^*) - |\mathbf{a}| \leq c[\mathbb{E}_\mathbf{a}(T^*)]^{1/2}$ . By Lemma 6.1, we have  $\mathbb{E}_\mathbf{a}(T^*) \leq \mathbb{E}_\mathbf{a}(\tau_{\max}) + c'$  for some finite constant  $c'$  independent of  $\mathbf{a}$ . Further, since  $\tau_{\max}$  is equal in distribution to  $\max_r V_r(a_r)$ , Lemma 6.2 gives (recall that  $\nu$  is the variance of the step distribution of the  $V_r$ 's)

$$\mathbb{E}_\mathbf{a}(T^*) \leq \frac{|\mathbf{a}|}{1 - \mathbb{E}\xi} + R(\nu|\mathbf{a}|)^{1/2} + c',$$

which implies the existence of the desired constant  $c$  such that  $\Psi(\mathbf{a}) \leq c[\mathbb{E}_\mathbf{a}(T^*)]^{1/2}$ . Defining

$$\Gamma(\mathbf{a}) = (1 - \rho)\mathbb{E}_\mathbf{a}(T^*) - c[\mathbb{E}_\mathbf{a}(T^*)]^{1/2}, \quad (11)$$

(10) can be rewritten as  $\mathbb{E}_\mathbf{q}[\Phi(\mathbf{Q}(1)) - \Phi(\mathbf{q})] \leq -\Gamma(\mathbf{q}^a)$ . Using the Markov property and (2), this gives

$$\begin{aligned} \mathbb{E}_\mathbf{q}[\Phi(\mathbf{Q}(2)) - \Phi(\mathbf{q})] \\ \leq -\mathbb{E}_{\mathbf{q}^a}\left(\Gamma(\mathbf{q}^a) + \Gamma(\mathbf{q}^i + \mathbf{S}(T^*))\right). \end{aligned} \quad (12)$$

When  $\Phi(\mathbf{q}) = |\mathbf{q}^a| + |\mathbf{q}^i|$  is large, at least one of the  $2R$  coordinates of  $\mathbf{q}$  must be large. Since  $\mathbb{E}_\mathbf{a}(T^*) \geq a_r/(1 - \mathbb{E}\xi)$  (as a consequence of  $T^* \geq \tau_r$  and  $\mathbb{E}_\mathbf{a}(\tau_r) = a_r/(1 - \mathbb{E}\xi)$ ), it is not hard to show that

$$\lim_{K \rightarrow +\infty} \inf_{\mathbf{q}: \Phi(\mathbf{q}) \geq K} \frac{\mathbb{E}_{\mathbf{q}^a}(\Gamma(\mathbf{q}^a) + \Gamma(\mathbf{q}^i + \mathbf{S}(T^*)))}{\Phi(\mathbf{q})} > 0,$$

which completes the proof of the result.  $\square$

Now that we have stability and finiteness of the stationary mean, we prove that  $\alpha(\infty) = 2$ . The proof will make use of the following result.

LEMMA 6.4. *If  $(W_j)$  are  $J$  i.i.d. random walks started at 0 with non-negative step distribution having mean  $m$  and variance  $w$ , then for any  $\mathbf{x} = (x_j) \in \mathbb{N}^J$  it holds that,*

$$\mathbb{E}\left[\left(\max_j W_j(x_j)\right)^{1/2}\right] \geq (m|\mathbf{x}|)^{1/2} - (w/m^{3/2})|\mathbf{x}|^{-1/2}.$$

PROOF. Since  $\mathbb{E}(\max_j W_j(x_j)) \geq \max_j \mathbb{E}(W_j(x_j))$  it is enough to prove the result for  $J = 1$ . Fix  $k \geq 0$  and let  $Y = (W_1(k) - mk)/(mk)^{1/2}$  and

$$f(y) = \frac{1 + y/2 - (1 + y)^{1/2}}{y^2}, \quad y \geq -1.$$

Since  $\mathbb{E}Y = 0$  and  $Y \geq -(km)^{1/2}$ , we can rewrite after some algebra

$$\mathbb{E}(W_1(k)^{1/2}) = (km)^{1/2} - (km)^{-1/2}\mathbb{E}\left[Y^2 f((km)^{-1/2}Y)\right],$$

and since  $\sup f = 1/2$  and  $\mathbb{E}(Y^2) = w/m$  this gives the result.  $\square$

THEOREM 6.5. *If  $\beta = +\infty$ , then*

$$0 < \liminf_{\rho \uparrow 1} [(1 - \rho)^2 \mathbb{E}_\infty (\|\mathbf{Q}(0)\|)] \leq \limsup_{\rho \uparrow 1} [(1 - \rho)^2 \mathbb{E}_\infty (\|\mathbf{Q}(0)\|)] < +\infty. \quad (13)$$

*In particular,  $\alpha(\infty) = 2$ .*

PROOF. Since  $\beta = +\infty$ , we have

$$\mathbb{E}_\infty (\|\mathbf{Q}(0)\|) = \mathbb{E}_\infty (\|\mathbf{A}(0)\|) = R \mathbb{E}_\infty (A_1(0)).$$

Thus we only need to prove

$$\limsup_{\rho \uparrow 1} [(1 - \rho)^2 \mathbb{E}_\infty (A_1(0))] < +\infty, \quad (14)$$

which implies the lower bound in (13), and

$$\liminf_{\rho \uparrow 1} \left\{ (1 - \rho) \mathbb{E}_\infty [A_1(0)^{1/2}] \right\} > 0 \quad (15)$$

which by Jensen's inequality implies the upper bound in (13).

*Proof of (14).* Starting from (4), using that  $\mathbb{E}_a(\tau_1) = a/(1 - \mathbb{E}\xi)$ , subtracting on both sides  $\mathbb{E}(\xi)\mathbb{E}_\infty(\tau_1)$  (for this precise operation we need the finiteness of the stationary first moment, to avoid doing  $\infty - \infty$ ) and dividing by  $\mathbb{E}\xi$ , we end up with

$$\frac{1}{\mathbb{E}\xi} \left( 1 - \frac{\mathbb{E}\xi}{1 - \mathbb{E}\xi} \right) \mathbb{E}_\infty (A_1(0)) = \mathbb{E}_\infty (T^* - \tau_1).$$

Then, adding and subtracting  $\tau_{\max}$  in the right hand side, and using that  $\mathbb{E}\xi = \rho/2$ , we obtain

$$g_\rho(1 - \rho) \mathbb{E}_\infty (A_1(0)) = \mathbb{E}_\infty (\tau_{\max} - \tau_1) + \mathbb{E}_\infty (T^* - \tau_{\max}),$$

with  $g_\rho = 4/(\rho(2 - \rho))$ . Thus in view of Lemma 6.1, to prove (14) we only have to show that

$$\limsup_{\rho \uparrow 1} \left( \frac{\mathbb{E}_\infty (\tau_{\max} - \tau_1)}{[\mathbb{E}_\infty (A_1(0))]^{1/2}} \right) < +\infty. \quad (16)$$

Applying Lemma 6.2 to  $\tau_{\max}$  under  $\mathbb{P}_a$  (equal in distribution to  $\max_r V_r(a_r)$ ), we obtain, denoting temporarily  $\mu = 1/(1 - \mathbb{E}\xi)$ ,  $\mathbb{E}_a(\tau_{\max} - \tau_1) \leq \mu(|a| - a_1) + R\nu|a|^{1/2}$ . Integrating over the stationary distribution of  $\mathbf{Q}$  and using Jensen's inequality, we obtain

$$\mathbb{E}_\infty (\tau_{\max} - \tau_1) \leq \mu \mathbb{E}_\infty (|\mathbf{A}(0)| - A_1(0)) + R\nu [\mathbb{E}_\infty (A_1(0))]^{1/2}.$$

In particular, to prove (16) it is enough to show that

$$\limsup_{\rho \uparrow 1} \left( \frac{\mathbb{E}_\infty (|\mathbf{A}(0)| - A_1(0))}{[\mathbb{E}_\infty (A_1(0))]^{1/2}} \right) < +\infty. \quad (17)$$

We have already shown in the proof of Proposition 6.3 that  $\mathbb{E}_\infty (|\mathbf{A}(0)|) = \mathbb{E}_\infty (|\mathbf{S}(T^*)|) \leq \mathbb{E}(\xi)\mathbb{E}_\infty (T^*) + R\nu [\mathbb{E}_\infty (T^*)]^{1/2}$ , and since  $\mathbb{E}_\infty (A_1(0)) = \mathbb{E}(\xi)\mathbb{E}_\infty (T^*)$  this gives

$$\mathbb{E}_\infty (|\mathbf{A}(0)| - A_1(0)) \leq R\nu [\mathbb{E}_\infty (A_1(0))/\mathbb{E}\xi]^{1/2}.$$

This proves (17) which completes the proof of (14).

*Proof of (15).* We have

$$\mathbb{E}_\infty [(Q_1^a(0))^{1/2}] = \mathbb{E}_\infty [(S_1(T^*))^{1/2}] \geq \mathbb{E}_\infty [(S_1(\tau_{\max}))^{1/2}].$$

Applying Lemma 6.4 to  $S_1(\tau_{\max})$  by using the independence between  $S_1$  and  $\tau_{\max}$ , we obtain

$$\mathbb{E}_\infty [(Q_1^a(0))^{1/2}] \geq (\mathbb{E}\xi)^{1/2} \mathbb{E}_\infty [(\tau_{\max})^{1/2}] - c \mathbb{E}_\infty [(\tau_{\max})^{-1/2}]$$

for some finite constant  $c$  independent of  $\rho$ . Applying again Lemma 6.4 to  $\tau_{\max}$  we obtain

$$\mathbb{E}_\infty [(Q_1^a(0))^{1/2}] \geq (\mu \mathbb{E}\xi)^{1/2} \mathbb{E}_\infty (|\mathbf{Q}^a(0)|^{1/2}) - c \mathbb{E}_\infty (|\mathbf{Q}^a(0)|^{-1/2}) - c \mathbb{E}_\infty [(\tau_{\max})^{-1/2}].$$

Subtracting  $(\mu \mathbb{E}\xi)^{1/2} \mathbb{E}_\infty (Q_1^a(0)^{1/2})$  on both sides we finally end up with

$$h_\rho(1 - \rho) \mathbb{E}_\infty [(Q_1^a(0))^{1/2}] \geq (\mu \mathbb{E}\xi)^{1/2} \mathbb{E}_\infty (|\mathbf{Q}^a(0)|^{1/2} - (Q_1^a(0))^{1/2}) - c \mathbb{E}_\infty (|\mathbf{Q}^a(0)|^{-1/2}) - c \mathbb{E}_\infty [(\tau_{\max})^{-1/2}],$$

with  $h_\rho = (1 - (\rho/(2 - \rho))^{1/2})/(1 - \rho) \rightarrow 1$  as  $\rho \rightarrow 1$ . The two last terms of the previous lower bound vanish as  $\rho \uparrow 1$ . As for the first term, it is not hard based on (2) to show that the  $R$ -dimensional vector  $((Q_1^a(0) + 1)^{-1/2}(Q_r^a(0) - Q_1^a(0)), r \in \mathcal{R})$  under  $\mathbb{P}_\infty$  converges weakly as  $\rho \uparrow 1$ , from which one readily deduces thanks to the continuous mapping theorem that  $|\mathbf{Q}^a(0)|^{1/2} - (Q_1^a(0))^{1/2}$  also converges weakly as  $\rho \uparrow 1$  to a random variable which is not identically zero. Using Fatou's lemma, this implies that

$$\liminf_{\rho \uparrow 1} \mathbb{E}_\infty (|\mathbf{Q}^a(0)|^{1/2} - (Q_1^a(0))^{1/2}) > 0,$$

which concludes the proof.  $\square$

## 6.2 The case $\beta > 1$

### 6.2.1 Main result and a perturbation argument

In the previous subsection, we analyzed in detail the case  $\beta = \infty$  and proved  $\alpha(\infty) = 2$ . On the other hand, the simulation results, see Figure 5, strongly suggest that  $\alpha(\beta) = 2$  for any  $\beta > 1$ , and the goal of this subsection is to explain this result. The key to understand the behavior of  $\mathbf{Q}$  for  $\beta > 1$  is Proposition 6.6.

This proposition 6.6 is only concerned with the behavior of an active queue, i.e., a queue subject to the dynamic (1). Thus in order to prove Proposition 6.6, we only need to assume  $\rho < 2$ , which is the stability condition for an active queue. It is easy to see that these results actually hold uniformly in  $\rho \leq 1$ , which is what would be needed in order to go in heavy traffic for the full model.

In the sequel, we let  $T_1$  be the time at which  $A_1$  advertizes a release for the first time.

PROPOSITION 6.6. *If  $\beta > 1$ , then  $(A_1(T_1), T_1 - \tau_1)$  under  $\mathbb{P}_{a_1}$  converges weakly as  $a_1 \rightarrow +\infty$  to a non-degenerate random variable.*

By non-degenerate, we mean a random variable  $(X, Y) \in \mathbb{N} \times \mathbb{Z}$  such that both  $X$  and  $Y$  are non-deterministic and almost surely finite (an explicit expression for the weak limit of  $(A_1(T_1), T_1 - \tau_1)$  is given in Lemma 6.9). Note that from Proposition 6.6, which is concerned with the behavior of one active queue, one can easily deduce the behavior of  $\mathbf{A}(T^*)$

and  $T^*$  (the proof of the following result is only sketched in order to comply with page limitations).

**COROLLARY 6.7.** *Assume that  $\beta > 1$  and consider any sequence of initial states  $\mathbf{a}_n = (a_{n,r})$  such that  $\min_r a_{n,r} \rightarrow +\infty$ : then  $(\mathbf{A}(T^*), T^* - \tau_{\max})$  under  $\mathbb{P}_{\mathbf{a}_n}$  converges weakly as  $n \rightarrow +\infty$  to a non-degenerate random variable.*

**PROOF (SKETCH).** At time  $T^*$ , each active queue needs to have advertized a release at least once. At this time, the active queue was of order one by Proposition 6.6 and since it is stable, it remains of order one at time  $T^*$ .  $\square$

In the extreme case  $\beta = +\infty$  we have by definition  $\mathbf{A}(T^*) = \mathbf{0}$  and (by Lemma 6.1)  $T^* \approx \tau_{\max}$ . By showing that both  $\mathbf{A}(T^*)$  and  $T^* - \tau_{\max}$  are of order one when  $\beta > 1$ , Corollary 6.7 therefore justifies seeing the case  $\beta > 1$  as a perturbation of the case  $\beta = +\infty$ . In particular, treating  $\mathbb{E}_\infty(\mathbf{A}(T^*))$  and  $\mathbb{E}_\infty(T^* - \tau_{\max})$  as constants, the heuristic reasoning outlined in Section 4.3 goes through and again predicts a scaling exponent  $\alpha(\beta) = 2$  for any  $\beta > 1$ . This is indeed in very good agreement with the simulation results discussed in Section 5.2 and was stated as our main result 2.

The remainder of this subsection is devoted to the proof of Proposition 6.6. The proof proceeds in two steps: in the first step we show that the proof of Proposition 6.6 reduces to proving a simpler property of some particular random walk (see (18)). In the second step we prove that this property holds when  $\beta > 1$ .

### 6.2.2 First step

Since before time  $T_1$ , the active queue  $A_1$  does by definition not advertize any release, it is enough to prove Proposition 6.6 in the case  $\zeta = 0$ , which we assume in the remainder of this subsection. In particular,  $A_1$  is a random walk with step distribution  $\xi - 1$  reflected at 0. The reduction of the proof of Proposition 6.6 to proving (18) relies on the following coupling of the processes  $A_1$  for all possible initial states  $a \geq 0$ .

Let  $V$  and  $W^\uparrow$  be two independent processes with the following distribution. Let  $V$  be a version of  $A_1$  under  $\mathbb{P}_0$ , i.e.,  $(V(k), k \geq 0)$  is a random walk started at 0, with step distribution  $\xi - 1$  and reflected at 0.

Let  $W^\uparrow$  be a random walk started at 0, with step distribution  $1 - \xi$  and conditioned on never visiting 0 after time 0: since  $\mathbb{E}(1 - \xi) > 0$  this conditioning is well-defined. Let moreover  $\kappa(a) = \max\{k \geq 0 : W^\uparrow(k) = a\}$  be the time of the last visit to  $a \in \{0, 1, \dots, \infty\}$  (to be understood as  $\kappa(a) = +\infty$  for  $a = +\infty$ ), so that  $\kappa(a)$  is almost surely finite if  $a$  is finite. Let finally  $W_a^\uparrow$  be the process  $W^\uparrow$  stopped at  $\kappa(a)$ , i.e.,  $W_a^\uparrow(k) = W^\uparrow(k)$  if  $k \leq \kappa(a)$  and  $W_a^\uparrow(k) = W^\uparrow(\kappa(a)) = a$  if  $k \geq \kappa(a)$ .

**LEMMA 6.8.** *Extend  $A_1$  on  $\mathbb{Z}$  by setting  $A_1(k) = A_1(0)$  for  $k \leq 0$ , and let  $A^+ = (A_1(\tau_1 + k), k \geq 0)$  and  $A^- = (A_1(\tau_1 - k), k \geq 0)$ . Then for any finite  $a \geq 0$ ,  $(A^+, A^-)$  under  $\mathbb{P}_a$  is equal in distribution to  $(V, W_a^\uparrow)$ .*

*In particular, as  $a \rightarrow +\infty$ ,  $(A^+, A^-)$  under  $\mathbb{P}_a$  converges weakly to  $(V, W^\uparrow)$ .*

**PROOF.** That  $A^+$  is equal in distribution to  $V$  and is independent from  $A^-$  follows from the strong Markov property at time  $\tau_1$ . That  $A^-$  is equal in distribution to  $W_a^\uparrow$  comes from duality. The weak convergence result then follows from the fact that  $\kappa(a) \rightarrow +\infty$  almost surely as  $a \rightarrow +\infty$ , so that  $(V, W_a^\uparrow) \rightarrow (V, W^\uparrow)$  almost surely as  $a \rightarrow +\infty$ .  $\square$

Essentially, this representation of  $A_1$  shifts the origin of time at  $\tau_1$ :  $A^+$  looks at  $A_1$  from time  $\tau_1$  forward in time, while  $A^-$  looks at  $A_1$  from  $\tau_1$  backward in time. Moreover, this representation couples all the processes  $A_1$  with different initial states on the same probability space, which yields a simple representation for the law of  $(A_1(T_1), T_1 - \tau_1)$ . Let in the sequel  $Z = (Z(k), k \in \mathbb{Z})$  be the following process (indexed by  $\mathbb{Z}$ ):  $Z(k) = V(k)$  if  $k \geq 0$  and  $Z(k) = W^\uparrow(-k)$  if  $k \leq 0$ . The previous coupling immediately implies the following result.

**LEMMA 6.9.** *Let  $(U_k, k \in \mathbb{Z})$  be i.i.d., uniformly distributed in  $[0, 1]$ , independent from  $Z$ , and for  $0 \leq a \leq +\infty$  and  $k \in \mathbb{Z}$  let*

$$D_{a,k} = \begin{cases} 0 & \text{if } k < -\kappa(a), \\ \mathbb{1}_{\{U_k < \psi(Z(k))\}} & \text{else,} \end{cases}$$

and  $T_a^Z = \inf\{k \in \mathbb{Z} : D_{a,k} = 1\}$ . Then for any finite  $a \geq 0$ ,  $(A_1(T_1), T_1 - \tau_1)$  under  $\mathbb{P}_a$  is equal in distribution to  $(Z(T_a^Z), T_a^Z)$  and in particular, it converges weakly as  $a \rightarrow +\infty$  to  $(Z(T_\infty^Z), T_\infty^Z)$  (with  $Z(T_\infty^Z) = +\infty$  if  $T_\infty^Z = -\infty$ ).

**PROOF.** The equality in law between  $(A_1(T_1), T_1 - \tau_1)$  and  $(Z(T_a^Z), T_a^Z)$  is clear from the construction, and not difficult (although a bit heavy in notation) to formalize. Moreover,  $T_a^Z$  is by construction decreasing in  $a$  and so its limit as  $a \rightarrow +\infty$  exists. It is not hard to show that its limit is exactly  $T_\infty^Z$  and by continuity we deduce that  $Z(T_a^Z) \rightarrow Z(T_\infty^Z)$  as  $a \rightarrow +\infty$ , which implies the result.  $\square$

Thus to prove Proposition 6.6, we only have to establish that  $|T_\infty^Z|$  is (almost surely) finite. Since  $V$  is positive recurrent and starts at 0, it is clear that  $\min\{k \geq 0 : D_{\infty,k} = 1\}$  is finite and so to prove that  $|T_\infty^Z|$  is finite, we only have to demonstrate that  $\inf\{k \leq 0 : D_{\infty,k} = 1\}$  is finite. In other words, we have to prove that  $\sup\{k \geq 0 : U_{-k} < \psi(W^\uparrow(k))\}$  is finite, which informally means that  $W^\uparrow$  advertizes a release only finitely many times.

So in the sequel, we consider  $(U_k, k \geq 0)$  i.i.d. random variables, uniformly distributed on  $[0, 1]$  and independent of  $W^\uparrow$ , and we define

$$N = \sum_{k \geq 0} \mathbb{1}_{\{U_k < \psi(W^\uparrow(k))\}}$$

as the number of times  $W^\uparrow$  advertizes a release. The proof of Proposition 6.6 will thus be complete if we can prove that

$$\mathbb{P}(N < +\infty) = 1. \quad (18)$$

### 6.2.3 Second step

We now assume that  $\beta > 1$  and we prove that  $\mathbb{P}(N > n) \rightarrow 0$  as  $n \rightarrow +\infty$ , which will prove (18). By definition,

$$\mathbb{P}(N = 0) = \mathbb{P}\left(U_k > \psi(W^\uparrow(k)), k \geq 0\right),$$

and since  $W^\uparrow$  and the  $U_k$ 's are independent this gives

$$\mathbb{P}(N = 0) = \mathbb{E}\left[\prod_{k \geq 0} (1 - \psi(W^\uparrow(k)))\right].$$

Let  $a \geq 0$ : introducing  $\varphi(a) = -\log(1 - \psi(a))$  and  $L^\uparrow(a) =$

$\sum_{k \geq 0} I_{\{W^\uparrow(k)=a\}}$ , the local time at level  $a$ , we obtain

$$\mathbb{P}(N = 0) = \mathbb{E} \left[ \exp \left( - \sum_{a \geq 0} \varphi(a) L^\uparrow(a) \right) \right]. \quad (19)$$

LEMMA 6.10. *The quantity  $\sup_{a \geq 0} \mathbb{E}(L^\uparrow(a))$  is finite. In particular,  $\mathbb{P}(N = 0) > 0$ .*

PROOF. Let  $W^-$  be a random walk with step distribution  $\xi - 1$ , started at 0 and independent from  $W^\uparrow$ , and for  $k \in \mathbb{Z}$  define  $W^*(k) = W^\uparrow(k)$  if  $k \geq 0$  and  $W^*(k) = W^-(-k)$  if  $k \leq 0$ . Thus, defining  $L^*(a) = \sum_{k \in \mathbb{Z}} I_{\{W^*(k)=a\}}$  we have the obvious inequality  $L^\uparrow(a) \leq L^*(a)$  and so we only have to prove that  $\sup_{a \in \mathbb{Z}} \mathbb{E}(L^*(a))$  is finite.

It is clear that  $L^*$  stays the same if  $W^*$  is shifted in time, and that shifting  $L^*$  in time amounts to shifting  $W^*$  in space. Moreover, for any  $w \in \mathbb{Z}$  the process  $(W^*(k) + w, k \in \mathbb{Z})$  shifted at the time of last visit to 0 is equal in distribution to  $W^*$ . Combining these facts, we see that  $L^*$  is a stationary sequence and in particular,  $\sup_{a \in \mathbb{Z}} \mathbb{E}(L^*(a)) = \mathbb{E}(L^*(0))$ . But by the strong Markov property, it is clear that  $L^*(0)$  is a geometric random variable, in particular it has finite first moment. This proves the finiteness of  $\sup_a \mathbb{E}(L^\uparrow(a))$ .

As for  $\mathbb{P}(N > 0)$ , we have

$$\mathbb{E} \left( \sum_{a \geq 0} \varphi(a) L^\uparrow(a) \right) \leq \sup_a \mathbb{E}(L^\uparrow(a)) \sum_{a \geq 0} \varphi(a),$$

and since  $\varphi(a) \sim a^{-\beta}$  as  $a \rightarrow +\infty$  and  $\beta > 1$ , the sum  $\sum_a \varphi(a)$  is finite which implies, in view of the last display, that the random variable  $\sum_a \varphi(a) L^\uparrow(a)$  is almost surely finite. This proves  $\mathbb{P}(N = 0) > 0$  in view of (19) and concludes the proof of the lemma.  $\square$

We now prove that  $\mathbb{P}(N > n) \rightarrow 0$  as  $n \rightarrow +\infty$ . Let  $W$  be a random walk with step distribution  $1 - \xi$  and  $I = \inf_{k \geq 1} W(k)$ . Then,  $W^\uparrow$  is by definition equal in distribution to  $W$  under  $\mathbb{P}_0(\cdot | I \geq 1)$  (where the subscript refers to the initial state of  $W$ ). Let moreover  $B_n$  be the time at which  $W$  advertizes a release for the  $n$ th time, so that

$$\mathbb{P}(N = n) = b \mathbb{P}_0(B_n < +\infty, B_{n+1} = +\infty, I \geq 1),$$

with  $b = 1/\mathbb{P}_0(I \geq 1)$ . Writing the event  $\{I \geq 1\}$  as the union between the two events  $\{\inf_{1 \leq k \leq B_n} W(k) \geq 1\}$  and  $\{\inf_{k > B_n} W(k) \geq 1\}$ , the strong Markov property at time  $B_n$  entails

$$\mathbb{P}(N = n) = b \mathbb{E}_0 \left( p(W(B_n)); B_n < +\infty, \inf_{1 \leq k \leq B_n} W(k) \geq 1 \right),$$

with  $p(w) = \mathbb{P}_w(N = 0, I \geq 1)$ . Coupling  $W$  under  $\mathbb{P}_w$  with a version of  $W$  under  $\mathbb{P}_0$  that stays below it, it is easy to see that  $p(w)$  is increasing in  $w$  and so

$$\begin{aligned} \mathbb{P}(N = n) &\geq b p(0) \mathbb{P}_0 \left( B_n < +\infty, \inf_{1 \leq k \leq B_n} W(k) \geq 1 \right) \\ &\geq b p(0) \mathbb{P}_0(B_n < +\infty, I \geq 1). \end{aligned}$$

This last lower bound is equal to  $p(0) \mathbb{P}_0(B_n < +\infty | I \geq 1)$  which by definition is equal to  $p(0) \mathbb{P}(N > n)$ . Since  $p(0) = \mathbb{P}(N = 0)/b$  is  $> 0$  by Lemma 6.10, dividing by  $p(0)$  leads to

$$\mathbb{P}(N > n) \leq \frac{b \mathbb{P}(N = n)}{\mathbb{P}(N = 0)}.$$

Since  $\mathbb{P}(N = n) \rightarrow 0$ , this finally proves (18) and hence Proposition 6.6.

#### 6.2.4 More on the case $\beta > 1$ , $\beta \approx 1$

The simulation results in Section 5 show a rather fuzzy behavior of  $\alpha(\beta)$  for  $\beta$  close to 1. Indeed, the curves shown in Figure 5 are smooth for small and large values (say,  $\beta < 1/2$  and  $\beta > 1.2$ ) but for  $\beta$  close to one it is difficult to obtain stable numerical results. Our goal here is to discuss potential interesting phenomena arising for  $\beta > 1$  close to one.

The shape of the function  $\rho \mapsto F(\rho, \beta)$  depicted in Figure 4 is typical for large values of  $\beta$ , say  $\beta > 1.2$ . In particular, this function is increasing which makes the regression of  $F$  against  $1/\log(1/(1 - \rho))$ , such as in (8), reasonable. However, as  $\beta$  gets closer to one, the shape of this function changes. For instance, Figure 7 shows simulation results for  $F(\rho, 1.2)$  which are representative of  $F(\rho, \beta)$  for small  $\beta$ , say  $1 < \beta < 1.2$ . Noticeably, the function  $F(\rho, 1.2)$  is not monotone in  $\rho$  and so the approximation (8) cannot be valid for every  $\rho$ . Rather, we find that  $F(\rho, \beta)$  decreases and then increases, and that the regression against  $1/\log(1/(1 - \rho))$  is only accurate past the minimum.

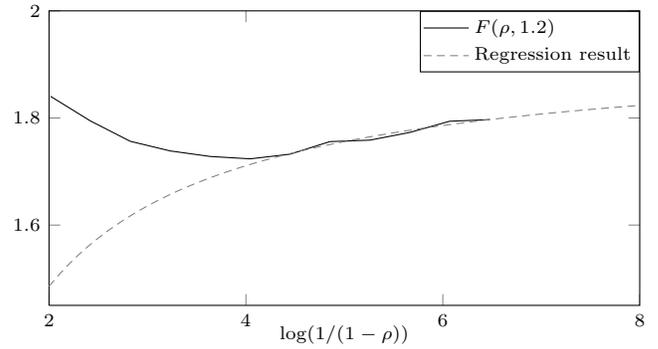


Figure 7: Approximating  $\alpha(1.2)$  for  $R = 2$ .

Regressing the curve obtained in Figure 7 past the minimum leads to the approximation  $\alpha(1.2) \approx 1.94$ , which is still very much in line with our theoretical result  $\alpha(1.2) = 2$ . However, the point where the minimum of the function  $\rho \mapsto F(\rho, \beta)$  is attained, shifts to the right when  $\beta$  gets closer to 1, leaving us with less points against which to regress. In particular, for  $\beta < 1.2$  one would need to simulate the system at loads higher than 0.999 to get accurate results.

We suspect that this numerical instability is caused by heavy tails phenomena that seem to appear for  $\beta < 2$ . More precisely, inspecting the proofs of Proposition 6.3 and Theorem 6.5, one sees that Corollary 6.7 is not strong enough for the proofs of the case  $\beta = +\infty$  to go through directly. Indeed, instead of controlling the behavior of  $\mathbf{A}(T^*)$  and  $T^*$  in distribution, one needs (at least, with the proposed proof strategy) to control their mean behavior. Let us do a small computation: let  $T_\infty^Z$  be the random variable introduced in Lemma 6.9, which is the weak limit of  $T_1 - \tau_1$ , and let  $B^\uparrow$  be the first time the process  $W^\uparrow$  advertizes a release. Then

$$\mathbb{E}(|T_\infty^Z|; T_\infty^Z \leq 0) = \mathbb{E}(B^\uparrow; B^\uparrow < +\infty) = \sum_{k \geq 0} k \mathbb{P}(B^\uparrow = k),$$

and, as before, we have

$$\mathbb{P}(B^\dagger = k) = \mathbb{E} \left[ \psi(W^\dagger(k)) \prod_{i < k} (1 - \psi(W^\dagger(i))) \right].$$

Thus for large  $k$ , we should have  $\mathbb{P}(B^\dagger = k) \approx k^{-\beta}$  which suggests that  $B^\dagger$ , and in particular  $|T_\infty^Z|$ , has infinite mean for  $\beta \leq 2$ , although both random variables are almost surely finite for  $\beta > 1$ . It is possible to use this result to show that the (almost surely finite) weak limits of  $\mathbf{A}(T^*)$  and  $T^* - \tau_{\max}$  have infinite first moment. This potentially invalidates the back-of-the-envelope computations in Section 4.3, and so more care is needed. For instance, simulation experiments for  $\beta = 1.2$  and  $R = 2$  suggest that  $\mathbb{E}_{a_1}(A_1(T^*))$  grows as  $a_1^{0.4}$  as  $a_1 \rightarrow +\infty$ .

## 7. BROADER IMPLICATIONS

Motivated by the poor heavy traffic delay performance of “cautious” activation rules in queue-based schemes for distributed medium access control, we investigated more aggressive schemes. Our main contribution lies in highlighting a new effect that we called the lingering effect and in studying the performance ramifications of this effect for a special topology. In this section we explain and discuss various directions in which our framework could possibly be extended.

First of all, it would be straightforward to extend our results to the following asymmetric case: instead of having  $R$  queues in each group with identically distributed arrival processes across queues, we have two groups of  $R_1$  and  $R_2$  queues and the arrivals into the  $k$ th queue of group  $g$  have distribution  $\xi_{g,k}$ . We chose to study a symmetric scenario for technical reasons, since then there is no need to label queues individually. In this setting, the lingering effect will occur whenever, informally speaking, the two dominant queues of at least one of the two groups have the same arrival rate. For instance, the delay will scale like  $1/(1-\rho)^2$  with  $\rho = \max_k \mathbb{E}(\xi_{1,k}) + \max_k \mathbb{E}(\xi_{2,k})$  if the condition  $\mathbb{E}(\xi_{1,1}) = \mathbb{E}(\xi_{1,2}) \geq \max_{g,k} \mathbb{E}(\xi_{g,k})$  is satisfied.

We believe that the insights provided by the complete bipartite constraint graph carry over to more general topologies. Note that for a general topology, our model needs to be changed, since one needs then to specify more precisely how queues become active. One can for instance think of queues going into back-off, and then trying to grab the channel at some rate. We conjecture that whenever the constraint graph is not complete and thus contains an independent set of several nodes, the lingering effect can rear its head provided some algebraic condition between the arrival rates at the various queues is satisfied. It would be very interesting to be able to formulate a precise and formal conjecture reflecting this intuition, and most probably even more challenging to prove it.

## 8. REFERENCES

- [1] N. Bouman, S. Borst, and J. van Leeuwen. Achievable delay performance in CSMA networks. In *Proc. Allerton Conf.*, 2011.
- [2] P. Chaporkar and S. Sarkar. Stable scheduling policies for maximizing throughput in generalized constrained queueing. In *Proc. Infocom*, 2006.
- [3] M. Feuillet, A. Proutiere, and P. Robert. Random capture algorithms: Fluid limits and stability. In *Proc. ITA Workshop*, 2010.
- [4] D. Gamarnik and A. Zeevi. Validity of heavy traffic steady-state approximation in generalized Jackson networks. *Ann. Appl. Probab.*, 16(1):56–90, 2006.
- [5] L. Georgiadis, M. Neely, and L. Tassiulas. Resource allocation and cross-layer control in wireless networks. *Found. Trends Netw.*, 1:1–144, 2006.
- [6] J. Ghaderi, S. Borst, and P. Whiting. Backlog-based random-access in wireless networks: Fluid limits and instability issues. In *Proc. IEEE WiOpt Conf.*, 2012.
- [7] L. Jiang, D. Shah, J. Shin, and J. Walrand. Distributed random access algorithm: scheduling and congestion control. *IEEE Trans. Inform. Theory*, 56(12):6182–6207, 2010.
- [8] L. Jiang and J. Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In *Proc. Allerton Conf.*, 2008.
- [9] P. Marbach and A. Eryilmaz. A backlog-based CSMA mechanism to achieve fairness and throughput-optimality in multihop wireless networks. In *Proc. Allerton Conf.*, 2008.
- [10] E. Modiano, D. Shah, and G. Zussman. Maximizing throughput in wireless networks via gossiping. In *Proc. ACM SIGMETRICS/Performance Conf.*, 2006.
- [11] S. Rajagopalan, D. Shah, and J. Shin. Network adiabatic theorem: An efficient randomized protocol for content resolution. In *Proc. ACM SIGMETRICS/Performance Conf.*, 2009.
- [12] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *Ann. Appl. Probab.*, 22(1):128–171, 2012.
- [13] D. Shah and D. Wischik. Switched networks with maximum weight policies: Fluid approximation and multiplicative state space collapse. *Ann. Appl. Probab.*, 22(1):70–127, 2012.
- [14] G. Sharma, R. Mazumdar, and N. Shroff. On the complexity of scheduling in wireless networks. In *Proc. MobiCom*, 2006.
- [15] G. Sharma, N. Shroff, and R. Mazumdar. Joint congestion control and distributed scheduling for throughput guarantees in wireless networks. In *Proc. Infocom*, 2007.
- [16] A. Stolyar. Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *Ann. Appl. Probab.*, 14:1–53, 2004.
- [17] L. Tassiulas. Linear complexity algorithms for maximum throughput in radio networks and input queued switches. In *Proc. Infocom*, 1998.
- [18] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Aut. Contr.*, 37:1936–1948, 1992.
- [19] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Trans. Inf. Theory*, 39:466–478, 1993.
- [20] X. Wu and R. Srikant. Scheduling efficiency of distributed greedy scheduling algorithms in wireless networks. In *Proc. Infocom*, 2006.