

## INTERACTING BRANCHING PROCESSES AND LINEAR FILE-SHARING NETWORKS

LASSE LESKELÄ,\* *Helsinki University of Technology*

PHILIPPE ROBERT\*\* \*\*\*\* AND

FLORIAN SIMATOS,\*\* \*\*\*\* *INRIA Paris-Rocquencourt*

### Abstract

File-sharing networks are distributed systems used to disseminate files among nodes of a communication network. The general simple principle of these systems is that once a node has retrieved a file, it may become a server for this file. In this paper, the capacity of these networks is analyzed with a stochastic model when there is a constant flow of incoming requests for a given file. It is shown that the problem can be solved by analyzing the asymptotic behavior of a class of interacting branching processes. Several results of independent interest concerning these branching processes are derived and then used to study the file-sharing systems.

*Keywords:* Interacting branching process; peer-to-peer algorithm

2010 Mathematics Subject Classification: Primary 60J80; 60J85  
Secondary 90B18

### 1. Introduction

File-sharing networks are distributed systems used to disseminate information among some nodes of a communication network. The general simple principle is the following: once a node has retrieved a file it becomes a server for this file. An improved version of this principle consists of splitting the original file into several pieces (called ‘chunks’) so that a given node can retrieve simultaneously several chunks of the same file from different servers. In this case, the rate to get a given file may thus increase significantly, as well as the global capacity of the file-sharing system: a node becomes a server of a chunk as soon as it has retrieved it and not only when it has the whole file. These schemes disseminate information efficiently as long as the number of servers grows rapidly.

In this paper we investigate the maximal throughput of these file-sharing networks, i.e. whether the system can or cannot cope with a constant flow of incoming requests for a given file. Two cases are considered, either the file consists of one chunk or the file is split into two chunks that are retrieved sequentially, which we will refer to as a linear file-sharing network. It is assumed that arrival times are Poisson and chunk transmission times are exponentially distributed. In this setting, the stability property of the file-sharing system is expressed as

---

Received 6 July 2009; revision received 16 April 2010.

Work partially supported by the SCALP project funded by the EEC Network of Excellence Euro-FGI, and the Academy of Finland.

\* Current address: Aalto University, Department of Mathematics and Systems Analysis, PO Box 11100, 00076 Aalto, Finland. Email address: lasse.leskela@iki.fi

\*\* Postal address: INRIA Paris-Rocquencourt, Domaine de Voluceau, BP 105, 78153 Le Chesnay, France.

\*\*\* Email address: philippe.robert@inria.fr

\*\*\*\* Email address: florian.simatos@inria.fr

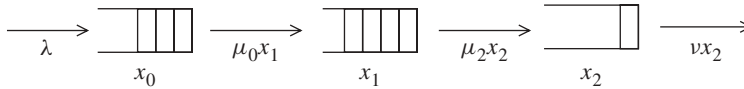


FIGURE 1: Transition rates outside the boundaries of the file-sharing network with two chunks.

the ergodicity property of the associated Markov process. Even in this simple framework, mathematical studies are quite scarce; see [13], [17], [18], and the references therein.

The main technical difficulty in proving stability/instability results for this class of stochastic networks is that, except for the input, the Markov process has unbounded jump rates, in fact proportional to one of the coordinates of the current state. If  $x$  servers have a chunk, since each of them can deliver this chunk, they globally provide this chunk at a rate proportional to  $x$ . For this reason, the classical tools related to fluid limits cannot be used easily in this setting. See [4], [5], and [14] for example. However, this class of processes is close to another important class of stochastic processes, namely branching processes, where a population of size  $x$  evolves at a rate proportional to  $x$ . As will be seen, a file-sharing system that splits files into several chunks can be represented as a Markov process associated to multitype branching processes with interaction.

### 1.1. Interacting branching processes

Consider the case of a file split into two chunks, and suppose that a new user arriving in the network first requests chunk number 1 and then requests chunk number 2. Assume further that a server of type 1, i.e. having chunk 1 only, distributes it at rate  $\mu_1$ , while a server of type 2, having chunks 1 and 2, distributes only chunk 2 at rate  $\mu_2$ . For  $i = 0, 1, 2$ , let  $X_i(t)$  be the number of type- $i$  servers at time  $t$ . A type-0 server is simply a user without any chunk. The crucial observation is that as long as there are requests for chunk  $i \in \{1, 2\}$ , then  $(X_i(t))$  grows similarly as a branching process where individuals give birth to one child at rate  $\mu_i$ , since each server offers a capacity  $\mu_i$  for the requested chunk. On the other hand, a new arrival in  $X_2$  corresponds to a departure in  $X_1$ , since the new type-2 server was a type-1 server; thus deaths in  $X_1$  are governed by births in  $X_2$ . See Figure 1.

The file-sharing network under consideration can thus be seen as a system of interacting branching processes where the births and the deaths of individuals are correlated. In the simpler case of one chunk, a description as a branching process (without interaction) has been used to analyze the transient behavior of these file-sharing systems. See [17], [19], and [21].

In Sections 2 and 3 we present results of independent interest concerning branching processes where individuals are killed at the instants  $(\sigma_n)$  of a random point process. In Section 2 a criterion for the extinction of the branching process is obtained in terms of the sequence  $(\sigma_n)$ , and an asymptotic result is derived in this case. In Section 3 we consider the case where  $(\sigma_n)$  is the sequence of birth instants of another independent branching process. Several useful estimates are derived in this setting. These results are used to establish the stability results concerning file-sharing systems with two chunks. The stability properties of a network with a single-chunk file are analyzed in detail in Section 4. The case of file-sharing networks with two chunks is detailed in Section 5.

## 2. Yule processes with killing

A Yule process  $(Y(t))$  with rate  $\mu > 0$  is a Markovian branching process with  $Q$ -matrix

$$q_Y(x, x + 1) = \mu x \quad \text{for all } x \geq 0. \tag{2.1}$$

A Yule process is simply a pure-birth process, where each individual gives birth to a child at rate  $\mu$ .

Yule processes are the basic branching processes that appear in analyzing the two-chunk network of Section 5. Actually, a variant of this stochastic model will be needed, where some individuals are killed. In this section we study this model when killings are given by an exogenous process and occur at fixed (random or deterministic) epochs; in Section 3 killings result from the interaction with another branching process.

In terms of branching processes, this killing procedure amounts to pruning the tree, i.e. to cut some edges of the tree, and the subtree attached to it. This procedure is fairly common for branching processes, in the Crump–Mode–Jagers model for example; see [6]. See also [1] or [11].

**2.1. A Yule process killed at fixed instants**

Until the end of this section, it is assumed that, provided that the population is nonempty, at epochs  $\sigma_n, n \geq 1$ , an individual is removed from the population of an ordinary Yule process  $(Y(t))$  with rate  $\mu_W$  starting with  $Y(0) = w \in \mathbb{N}$  individuals. It is assumed that  $(\sigma_n)$  is some fixed nondecreasing sequence. It will be shown that the process  $(W(t))$  obtained by killing one individual of  $(Y(t))$  at each of the successive instants  $(\sigma_n)$  survives with positive probability when the series with general term  $(e^{-\mu_W \sigma_n})$  converges.

We define

$$\kappa = \inf\{n \geq 1: W(\sigma_n) = 0\}.$$

The process  $(W(t))$  can be represented by

$$W(t) = Y(t) - \sum_{i=1}^{\kappa} X_i(t) \mathbf{1}_{\{\sigma_i \leq t\}}, \tag{2.2}$$

where, for  $1 \leq i \leq \kappa$  and  $t \geq \sigma_i$ ,  $X_i(t)$  is the total number of children at time  $t$  in the original Yule process of the  $i$ th individual killed at time  $\sigma_i$ . In terms of trees,  $(W(t))$  can be seen as a subtree of  $(Y(t))$ : for  $1 \leq i \leq \kappa$ ,  $(X_i(t))$  is the subtree of  $(Y(t))$  associated with the  $i$ th particle killed at time  $\sigma_i$ .

It is easily checked that  $(X_i(t - \sigma_i), t \geq \sigma_i)$  is a Yule process starting with one individual and, since a killed individual cannot have one of his descendants killed, that the processes

$$(\tilde{X}_i(t)) = (X_i(t + \sigma_i), t \geq 0), \quad 1 \leq i \leq \kappa,$$

are independent Yule processes.

For any process  $(U(t))$ , we define

$$(M_U(t)) := (e^{-\mu_W t} U(t)). \tag{2.3}$$

If  $(\tilde{X}(t))$  is a Yule process with rate  $\mu_W$ , the martingale  $(M_{\tilde{X}}(t))$  converges almost surely and in  $L_2$  to a random variable  $M_{\tilde{X}}(\infty)$  with an exponential distribution with mean  $\tilde{X}(0)$ , and, by Doob’s inequality,

$$\mathbb{E}\left(\sup_{t \geq 0} M_{\tilde{X}}(t)^2\right) \leq 2 \sup_{t \geq 0} \mathbb{E}(M_{\tilde{X}}(t)^2) < +\infty.$$

See [3]. Consequently,

$$e^{-\mu_W t} W(t) = M_Y(t) - \sum_{i=1}^{\kappa} e^{-\mu_W \sigma_i} M_{\tilde{X}_i}(t - \sigma_i) \mathbf{1}_{\{\sigma_i \leq t\}},$$

and, for any  $t \geq 0$ ,

$$\sum_{i=1}^{\kappa} e^{-\mu w \sigma_i} M_{\tilde{X}_i}(t - \sigma_i) \mathbf{1}_{\{\sigma_i \leq t\}} \leq \sum_{i=1}^{\kappa} e^{-\mu w \sigma_i} \sup_{s \geq 0} M_{\tilde{X}_i}(s).$$

Assume now that  $\sum_{i \geq 1} e^{-\mu w \sigma_i} < +\infty$ : then the last expression is integrable, and Lebesgue’s theorem implies that  $(M_W(t)) = (e^{-\mu w t} W(t))$  converges almost surely and in  $L_2$  to

$$M_W(\infty) = M_Y(\infty) - \sum_{i=1}^{\kappa} e^{-\mu w \sigma_i} M_{\tilde{X}_i}(\infty).$$

Clearly, for some large enough  $w^*$  and then for any  $w \geq w^*$ , we have

$$E_w(M_W(\infty)) \geq w - \sum_{i=1}^{+\infty} e^{-\mu w \sigma_i} > 0,$$

in particular  $P_w(M_W(\infty) > 0) > 0$  and  $P_w(W(t) \geq 1 \text{ for all } t \geq 0) > 0$ . If  $Y(0) = w < w^*$  and  $\sigma_1 > 0$ , then  $P_w(Y(\sigma_1) \geq w^* + 1) > 0$ , and, therefore, by translation at time  $\sigma_1$ , the same conclusion holds when the sequence  $(e^{-\mu w \sigma_i})$  has a finite sum. The following proposition has thus been proved.

**Proposition 2.1.** *Let  $(W(t))$  be a process growing as a Yule process with rate  $\mu_W$  and for which individuals are killed at nondecreasing instants  $(\sigma_n)$  with  $\sigma_1 > 0$ . If*

$$\sum_{i=1}^{+\infty} e^{-\mu w \sigma_i} < +\infty$$

*then as  $t$  gets large, and for any  $w \geq 1$ , the variable  $(e^{-\mu w t} W(t))$  converges  $P_w$ -almost surely and in  $L_2$  to a finite random variable  $M_W(\infty)$  such that  $P_w(M_W(\infty) > 0) > 0$ .*

Proposition 2.1 establishes the minimal results needed in Section 5. However, Kolmogorov’s three-series, see [20, p. 115], can be used in conjunction with Fatou’s lemma to show that  $(W(t))$  dies out almost surely when the series with general term  $(e^{-\mu w \sigma_n})$  diverges.

### 3. Interacting branching processes

In this section we study the Yule process of the previous section when killing times correspond to the birth times of some other branching process. The other branching process can be seen as a renewing Bellman–Harris process.

#### 3.1. Renewing Bellman–Harris process

In the rest of this section,  $\mu_Z, \nu > 0$  are fixed and  $(Z(t))$  is a birth-and-death process whose  $Q$ -matrix,  $Q_Z$ , is given by

$$q_Z(z, z + 1) = \mu_Z(z \vee 1) \quad \text{and} \quad q_Z(z, z - 1) = \nu z. \tag{3.1}$$

In the rest of the paper  $n \vee m$  denotes  $\max(n, m)$  for  $n, m \in \mathbb{N}$ . This process can be described equivalently as a time-changed M/M/1 queue (see Proposition 3.1) or as a sequence of independent branching processes (see Proposition 3.2). The time change is the discrete analog

of the Lamperti transform between continuous-state branching processes and Lévy processes; see [7]. As will be seen, these two viewpoints are complementary.

Let  $(\sigma_n)$  be the sequence of birth instants (i.e. positive jumps) of  $(Z(t))$ , and let  $(B_\sigma(t))$ , the corresponding counting process of  $(\sigma_n)$ , be given by, for  $t \geq 0$ ,

$$B_\sigma(t) = \sum_{i \geq 1} \mathbf{1}_{\{\sigma_i \leq t\}}.$$

**Proposition 3.1.** (Queueing representation.) *If  $Z(0) = z \in \mathbb{N}$  then*

$$(Z(t), t \geq 0) \stackrel{D}{=} (L(C(t)), t \geq 0), \tag{3.2}$$

where  $(L(t))$  is the process of the number of jobs of an M/M/1 queue with input rate  $\mu_Z$  and service rate  $v$ , and with  $L(0) = z$  and  $C(t) = \inf\{s > 0 : A(s) > t\}$ , where

$$A(t) = \int_0^t \frac{1}{1 \vee L(u)} du.$$

*Proof.* It is not difficult to check that the process  $(M(t)) := (L(C(t)))$  has the Markov property. Let  $Q_M$  be its  $Q$ -matrix. For  $z \geq 0$ ,

$$P(L(C(h)) = z + 1 \mid L(0) = z) = \mu_Z E(C(h)) + o(h) = \mu_Z(z \vee 1)h + o(h);$$

hence,  $q_M(z, z + 1) = \mu_Z(z \vee 1)$ . Similarly,  $q_M(z, z - 1) = v z$ . The proposition is proved.

**Corollary 3.1.** *For any  $\gamma > (\mu_Z - v) \vee 0$  and  $z = Z(0) \in \mathbb{N}$ ,*

$$E_z \left( \sum_{n=1}^{+\infty} e^{-\gamma \sigma_n} \right) < +\infty.$$

*Proof.* Proposition 3.1 shows that, in particular, the sequences of positive jumps of  $(Z(t))$  and of  $(L(C(t)))$  have the same distribution. Hence, if  $\mathcal{N}_{\mu_Z} = (t_n)$  is the arrival process of the M/M/1 queue, a Poisson process with parameter  $\mu_Z$ , then, using the notation of the above proposition, the relation

$$(\sigma_n) \stackrel{D}{=} (A(t_n))$$

holds. By using standard martingale properties of stochastic integrals with respect to Poisson processes, see [15], we obtain, for  $t \geq 0$ ,

$$\begin{aligned} E_z \left( \sum_{n \geq 1} e^{-\gamma A(t_n)} \right) &= E_z \left( \int_0^\infty e^{-\gamma A(s)} \mathcal{N}_{\mu_Z}(ds) \right) \\ &= \mu_Z E_z \left( \int_0^\infty e^{-\gamma A(s)} ds \right) \\ &= \mu_Z \int_0^\infty e^{-\gamma u} E_z(Z(u) \vee 1) du, \end{aligned} \tag{3.3}$$

where relation (3.2) has been used for the last equality. Kolmogorov’s equation for the process  $(Z(t))$  gives

$$\begin{aligned} \phi(t) &:= E_z(Z(t)) \\ &= \mu_Z \int_0^t E_z(Z(u) \vee 1) du - v \int_0^t E_z(Z(u)) du \\ &\leq (\mu_Z - v) \int_0^t \phi(u) du + \mu_Z t; \end{aligned}$$

therefore, by Gronwall’s lemma,

$$\phi(t) \leq \phi(0) + \mu_Z \int_0^t u e^{(\mu_Z - \nu)u} du \leq z + \frac{\mu_Z}{\mu_Z - \nu} t e^{(\mu_Z - \nu)t}.$$

From (3.3), we conclude that

$$E_z \left( \sum_n e^{-\gamma \sigma_n} \right) = E_z \left( \sum_n e^{-\gamma A(t_n)} \right) < +\infty.$$

The proposition is proved.

Before hitting 0,  $(Z(t))$  can be seen as a Bellman–Harris branching process with Malthusian parameter  $\alpha = \mu_Z - \nu$ ; see [3]. This Bellman–Harris branching process describes the evolution of a population of independent particles: each particle, at rate  $\lambda := \mu_Z + \nu$ , either splits into two particles with probability  $p := \mu_Z / (\mu_Z + \nu)$  or dies with probability  $1 - p$ . These processes will be referred to as  $(p, \lambda)$ -branching processes in the sequel.

A  $(p, \lambda)$ -branching process survives with positive probability only when  $p > \frac{1}{2}$ , in which case the probability of extinction is equal to  $q = (1 - p)/p = \nu / \mu_Z$ . The process  $(Z(t))$  only differs from a  $(p, \lambda)$ -branching process insofar as it regenerates after hitting 0, after a time exponentially distributed. When it regenerates, it again behaves as a  $(p, \lambda)$ -branching process (started with one particle), until it hits 0 again.

**Proposition 3.2.** (Branching representation.) *If  $Z(0) = z \in \mathbb{N}$  and  $(\tilde{Z}(t))$  is a  $(p, \lambda)$ -branching process started with  $z \in \mathbb{N}$  particles and  $\tilde{T}$  its extinction time, then*

$$(Z(t), 0 \leq t \leq T) \stackrel{D}{=} (\tilde{Z}(t), 0 \leq t \leq \tilde{T}),$$

where  $T = \inf\{t \geq 0 : Z(t) = 0\}$  is the hitting time of 0 by  $(Z(t))$ .

**Corollary 3.2.** *Suppose that  $\mu_Z > \nu$ . Then,  $P_z$ -almost surely for any  $z \geq 0$ , there exists a finite random variable  $Z(\infty)$  such that*

$$\lim_{t \rightarrow +\infty} e^{-(\mu_Z - \nu)t} Z(t) = Z(\infty) \quad \text{and} \quad Z(\infty) > 0.$$

*Proof.* When  $\mu_Z \geq \nu$ , the process  $(Z(t))$  couples in finite time with a supercritical  $(p, \lambda)$ -branching process  $(\tilde{Z}(t))$  conditioned on nonextinction; this follows readily from Proposition 3.2 (or see Appendix A for details). Since, for any supercritical  $(p, \lambda)$ -branching process,  $(\exp(-(\mu_Z - \nu)t)\tilde{Z}(t))$  converges almost surely to a finite random variable  $\tilde{Z}(\infty)$ , positive on the event of nonextinction (see [10]), we obtain the desired result.

Due to its technicality, the proof of the following result is postponed to Appendix A; this result is used in the proof of Proposition 3.4.

**Proposition 3.3.** *Suppose that  $\mu_Z > \nu$ . If*

$$\eta^*(x) = \frac{2 - x - \sqrt{x(4 - 3x)}}{2(1 - x)}, \quad 0 < x < 1, \tag{3.4}$$

then, for any  $0 < \eta < \eta^*(\nu / \mu_Z)$ ,

$$\sup_{z \geq 0} \left[ E_z \left( \sup_{t \geq \sigma_1} (e^{\eta(\mu_Z - \nu)t} B_\sigma(t)^{-\eta}) \right) \right] < +\infty.$$

**3.2. A Yule process killed at the birth instants of a renewing Bellman–Harris process**

We now consider the main interacting branching processes: in addition to  $(Z(t))$ , we consider an independent Yule process  $(Y(t))$  with parameter  $\mu_W$  (its  $Q$ -matrix is defined by (2.1) with  $\mu = \mu_W$ ). We proceed similarly as in Section 2: a process  $(W(t))$  is defined by killing one individual of  $(Y(t))$  at each of the birth instants  $(\sigma_n)$  of  $(Z(t))$  (see [2] and the references therein for related models). Recall that  $(W(t))$  is given by (2.2). The following results are key to analyzing the two-chunk network of Section 5.

**Proposition 3.4.** *Assume that  $\mu_Z - \nu > \mu_W$ , and let  $H_0$  be the extinction time of  $(W(t))$ , i.e.*

$$H_0 = \inf\{t \geq 0: W(t) = 0\}.$$

*Then the random variable  $H_0$  is almost surely finite and*

- (i)  $Z(H_0) - Z(0) \leq e^{\mu_W H_0} M_Y^*$ , where

$$M_Y^* = \sup_{t \geq 0} (e^{-\mu_W t} Y(t)),$$

- (ii) *there exists a finite constant  $C$  such that, for any  $z \geq 0$  and  $w \geq 1$ ,*

$$E_{(w,z)}(H_0) \leq C(\log(w) + 1). \tag{3.5}$$

In (3.5) the subscript  $(w, z)$  refers to the initial state of the Markov process  $(W(t), Z(t))$ . More generally, in the rest of the paper we use the convention that if  $(U(t))$  is a Markov process then the index  $u$  of  $P_u$  and  $E_u$  will refer to the initial condition of this Markov process.

*Proof of Proposition 3.4.* Define  $\alpha = \mu_Z - \nu$ . Concerning the almost-sure finiteness of  $H_0$ , note that (2.2) entails that  $W(t) \leq Y(t) - B_\sigma(t)$  for all  $t \geq 0$  on the event  $\{H_0 = +\infty\}$ . As  $t$  goes to  $\infty$ , both  $\exp(-\mu_W t)Y(t)$  and  $\exp(-\alpha t)B_\sigma(t)$  converge almost surely to positive and finite random variables (see [10]), which implies, when  $\alpha = \mu_Z - \nu > \mu_W$ , that  $W(t)$  converges to  $-\infty$  on  $\{H_0 = +\infty\}$ , and so this event is necessarily of probability 0.

Point (i) of the proposition comes from identity (2.2) evaluated at  $t = H_0$ :

$$Z(H_0) - Z(0) \leq B_\sigma(H_0) \leq Y(H_0) \leq e^{\mu_W H_0} M_Y^*. \tag{3.6}$$

By using the relation  $\exp(x) \geq x$ , (3.5) follows from the following bound: for any  $\eta < \eta^*(\nu/\mu_Z)$  (recall that  $\eta^*$  is given by (3.4)),

$$\sup_{w \geq 1, z \geq 0} [w^{-\eta} E_{(w,z)}(e^{\eta(\alpha - \mu_W)H_0})] < +\infty. \tag{3.7}$$

So all that is left to prove is this bound. Under  $P_{(w,z)}$ ,  $(Y(t))$  can be represented as the sum of  $w$  independent and identically distributed (i.i.d.) Yule processes, and so  $M_Y^* \leq M_{Y,1}^* + \dots + M_{Y,w}^*$  with  $(M_{Y,i}^*)$  i.i.d. distributed like  $M_Y^*$  under  $P_{(1,z)}$ ; inequality (3.6) then entails that

$$e^{(\alpha - \mu_W)H_0} \leq \left( \sum_{i=1}^w M_{Y,i}^* \right) \sup_{t \geq \sigma_1} \left( \frac{e^{\alpha t}}{B_\sigma(t)} \right).$$

By the independence of  $(M_{Y,i}^*)$  and  $(B_\sigma(t))$ , Jensen’s inequality gives, for any  $\eta < 1$ ,

$$E_{(w,z)}(e^{\eta(\alpha - \mu_W)H_0}) \leq w^\eta (E(M_{Y,1}^*))^\eta E_z \left( \sup_{t \geq \sigma_1} (e^{\eta \alpha t} B_\sigma(t)^{-\eta}) \right);$$

hence, bound (3.7) follows from Proposition 3.3.

We conclude this section with a Markov chain which will be used in Section 5. Define recursively the sequence  $(V_n)$  by  $V_0 = v$  and

$$V_{n+1} = \sum_{k=1}^{A_n(V_n)} I_{n,k}, \quad n \geq 0, \tag{3.8}$$

where, for each  $n$ ,  $(I_{n,k}, k \geq 1)$  are identically distributed integer-valued random variables independent of  $V_n$  and  $A_n(V_n)$ , and such that  $E(I_{n,1}) = p$  for some  $p \in (0, 1)$ . For  $v > 0$ ,  $A_n(v)$  is an independent random variable with the same distribution as  $Z(H_0)$  under  $P_{(1,v)}$ , i.e. with the initial condition  $(W(0), Z(0)) = (1, v)$ .

Equation (3.8) can be interpreted as a branching process with immigration, see [16], or also as an autoregressive model.

**Proposition 3.5.** *Under the condition  $\mu_Z - v > \mu_W$ , if  $(V_n)$  is the Markov chain defined by (3.8) and, for  $K \geq 0$ ,*

$$N_K = \inf\{n \geq 0: V_n \leq K\},$$

*then there exist  $\gamma > 0$  and  $K \in \mathbb{N}$  such that*

$$E(N_K \mid V_0 = v) \leq \frac{1}{\gamma} \log(1 + v) \quad \text{for all } v \geq 0. \tag{3.9}$$

*The Markov chain  $(V_n)$  is in particular positive recurrent.*

*Proof.* For  $V_0 = v \in \mathbb{N}$ , Jensen’s inequality and (3.8) give the relation

$$E_v \log\left(\frac{1 + V_1}{1 + v}\right) \leq E_{(1,v)} \log\left[\frac{1 + pZ(H_0)}{1 + v}\right]. \tag{3.10}$$

From Proposition 3.4 and by using the same notation, we obtain, under  $P_{(1,v)}$ ,

$$Z(H_0) \leq v + e^{\mu_W H_0} M_Y^*,$$

where  $(Y(t))$  is a Yule process starting with one individual. By looking at the birth instants of  $(Z(t))$ , it is easily checked that the random variable  $H_0$  under  $P_{(1,v)}$  is stochastically bounded by  $H_0$  under  $P_{(1,0)}$ . It follows from the integrability of  $H_0$  under  $P_{(1,0)}$  (proved in Proposition 3.4) and of  $M_Y^*$  that the expression

$$\log\left(\frac{1 + p(v + e^{\mu_W H_0} M_Y^*)}{1 + v}\right)$$

bounding the right-hand side of (3.10) is also an integrable random variable under  $P_{(1,0)}$ . Lebesgue’s theorem therefore gives

$$\limsup_{v \rightarrow +\infty} \left[ E_v \log\left(\frac{1 + V_1}{1 + v}\right) \right] \leq \log p < 0.$$

Consequently, we conclude that  $v \mapsto \log(1 + v)$  is a Lyapunov function for the Markov chain  $(V_n)$ , i.e. if  $\gamma = -(\log p)/2$ , there exists  $K$  such that, for  $v \geq K$ ,

$$E_v \log(1 + V_1) - \log(1 + v) \leq -\gamma.$$

Foster’s criterion, see Theorem 8.6 of [14], implies that  $(V_n)$  is indeed ergodic and that (3.9) holds.



### 4. The single-chunk network

This section is devoted to the study of a class of two-dimensional Markov jump processes  $(X_0(t), X_1(t))$ , with corresponding  $Q$ -matrix  $\Omega_r$  given, for  $x = (x_0, x_1) \in \mathbb{N}^2$ , by

$$\begin{cases} \Omega_r[(x_0, x_1), (x_0 + 1, x_1)] = \lambda, \\ \Omega_r[(x_0, x_1), (x_0 - 1, x_1 + 1)] = \mu r(x_0, x_1)(x_1 \vee 1) \mathbf{1}_{\{x_0 > 0\}}, \\ \Omega_r[(x_0, x_1), (x_0, x_1 - 1)] = \nu x_1, \end{cases} \tag{4.1}$$

where  $x \mapsto r(x)$ , referred to as the *rate function*, is some fixed function on  $\mathbb{N}^2$  with values in  $[0, 1]$ .

From a modeling perspective, this Markov process describes the following system. Requests for a single file arrive at rate  $\lambda$ , the first component  $X_0(t)$  is the number of requests which did not get the file, whereas the second component is the number of requests having the file and acting as servers until they leave the file-sharing network. The constants  $\mu$  and  $\nu$  can be viewed as the file transmission rate and the rate at which servers having the file leave, respectively. The term  $r(x_0, x_1)$  describes the interaction of downloaders and uploaders in the system. The term  $x_1 \vee 1$  can be interpreted so that there is one permanent server in the network, which is contacted if there are no other uploader nodes in the system. A related system where there is always one permanent server for the file can be modeled by replacing the term  $x_1 \vee 1$  by  $x_1 + 1$ . See the remark at the end of this section.

Several related examples of this class of models have been recently investigated. The case

$$r(x_0, x_1) = \frac{x_0}{x_0 + x_1}$$

was considered in [9] and [12]; in this case the downloading time of the file is neglected. Susitaival *et al.* [18] analyzed the rate function  $r(x)$ ,

$$r(x_0, x_1) = 1 \wedge \left( \alpha \frac{x_0}{x_1} \right),$$

where  $\alpha > 0$  and  $a \wedge b$  denotes  $\min(a, b)$  for  $a, b \in \mathbb{R}$ . This model allows us to take into account the fact that a request cannot be served by more than one server. See also [13].

With a slight abuse of notation, for  $0 < \delta \leq 1$ , the matrix  $\Omega_\delta$  will refer to the case when the function  $r$  is identically equal to  $\delta$ . Note that the boundary condition  $x_1 \vee 1$  for departures from the first queue prevents the second coordinate from ending up in the absorbing state 0. Other possibilities are discussed at the end of this section. In the following  $(X^r(t)) = (X_0^r(t), X_1^r(t))$  and  $(X^\delta(t)) = (X_0^\delta(t), X_1^\delta(t))$  will denote Markov processes with  $Q$ -matrices  $\Omega_r$  and  $\Omega_\delta$ , respectively.

#### 4.1. Free process

For  $\delta > 0$ ,  $Q_\delta$  denotes the following  $Q$ -matrix:

$$\begin{cases} Q_\delta[(y, z), (y + 1, z)] = \lambda, \\ Q_\delta[(y, z), (y - 1, z + 1)] = \mu \delta (z \vee 1), \\ Q_\delta[(y, z), (y, z - 1)] = \nu z. \end{cases} \tag{4.2}$$

The process  $(Y^\delta(t)) = (Y_0^\delta(t), Y_1^\delta(t))$ , referred to as the free process, will denote a Markov process with  $Q$ -matrix  $Q_\delta$ . Note that the first coordinate  $(Y_0^\delta(t))$  may become negative. The

second coordinate  $(Y_1^\delta(t))$  of the free process is the birth-and-death process with the same distribution as the process  $(Z(t))$  introduced in Section 3. It is easily checked that if  $\rho_\delta$  defined as  $\delta\mu/\nu$  is such that  $\rho_\delta < 1$  then  $(Y_1^\delta(t))$  is an ergodic Markov process converging in distribution to  $Y_1^\delta(\infty)$  and

$$\lambda^*(\delta) := \nu E(Y_1^\delta(\infty)) = \delta\mu E(Y_1^\delta(\infty) \vee 1) = \frac{\delta\mu}{(1 - \rho_\delta)(1 - \log(1 - \rho_\delta))}. \tag{4.3}$$

When  $\rho_\delta > 1$ , then the process  $(Y_1^\delta(t))$  converges almost surely and in expectation to  $\infty$ . In the sequel,  $\lambda^*(1)$  is simply denoted as  $\lambda^*$ .

In the following it will be assumed, see condition (4.6) below, that the rate function  $r$  converges to 1 as the first coordinate goes to  $\infty$ ; as will be seen, the special case  $r \equiv 1$  then plays a special role, and so before analyzing the stability properties of  $(X^r(t))$ , we begin with an informal discussion when the rate function  $r$  is identically equal to 1. Since the departure rate from the system is proportional to the number of requests/servers in the second queue, a large number of servers in the second queue gives a high departure rate, irrespectively of the state of the first queue. The input rate of new requests being constant, the real bottleneck with respect to stability is therefore when the first queue is large. The interaction of the two processes  $(X_0^1(t))$  and  $(X_1^1(t))$  is expressed through the indicator function of the set  $\{X_0^1(t) > 0\}$ . The second queue  $(X_1^1(t))$  locally behaves like the birth-and-death process  $(Y_1^1(t))$  as long as  $(X_0^1(t))$  is away from 0. The two cases  $\rho_1 > 1$  and  $\rho_1 < 1$  are considered.

If  $\rho_1 > 1$ , i.e.  $\mu > \nu$ , the process  $(X_1^1(t))$  is a transient process as long as the first coordinate is nonzero. Consequently, departures from the second queue occur faster and faster. Since, on the other hand, arrivals occur at a steady rate, departures eventually outpace arrivals. The fact that the second queue grows when  $(X_0(t))$  is away from 0 stabilizes the system independently of the value of  $\lambda$ , and so the system should be stable for any  $\lambda > 0$ .

If  $\rho_1 < 1$ , and as long as  $(X_0(t))$  is away from 0, the coordinate  $(X_1^1(t))$  locally behaves like the ergodic Markov process  $(Y_1^1(t))$ . Hence, if  $(X_0^1(t))$  is nonzero for long enough, the requests in the first queue see, on average,  $E(Y_1^1(\infty) \vee 1)$  servers which work at rate  $\mu$ . Therefore, the stability condition for the first queue should be

$$\lambda < \mu E(Y_1^1(\infty) \vee 1) = \lambda^*,$$

where  $\lambda^* = \lambda^*(1)$  is defined by (4.3). Otherwise, if  $\lambda > \lambda^*$ , the system should be unstable.

### 4.2. Transience and recurrence criteria for $(X^r(t))$

**Proposition 4.1.** (Coupling.) *If  $X^r(0) = Y^1(0) \in \mathbb{N}^2$ , there exists a coupling of the processes  $(X^r(t))$  and  $(Y^1(t))$  such that the relations*

$$X_0^r(t) \geq Y_0^1(t) \quad \text{and} \quad X_1^r(t) \leq Y_1^1(t) \tag{4.4}$$

hold for all  $t \geq 0$  and any sample path.

For any  $0 \leq \delta \leq 1$ , if

$$\tau_\delta = \inf\{t \geq 0 : r(X^r(t)) \leq \delta\} \quad \text{and} \quad \sigma = \inf\{t \geq 0 : X_0^r(t) = 0\},$$

and if  $X^r(0) = Y^\delta(0) \in \mathbb{N}^2$ , then there exists a coupling of the processes  $(X^r(t))$  and  $(Y^\delta(t))$  such that, for any sample path, the relations

$$X_0^r(t) \leq Y_0^\delta(t) \quad \text{and} \quad X_1^r(t) \geq Y_1^\delta(t) \tag{4.5}$$

hold for all  $t \leq \tau_\delta \wedge \sigma$ .

*Proof.* Let  $X^r(0) = (x_0, x_1)$  and  $Y^1(0) = (y, z)$  be such that  $x_0 \geq y$  and  $x_1 \leq z$ . We have to prove that the processes  $(X^r(t))$  and  $(Y^1(t))$  can be constructed such that (4.4) holds at the time of the next jump of one of them. See [8] for the existence of couplings using analytical techniques.

The arrival rates in the first queue are the same for both processes. If  $x_1 < z$ , a departure from the second queue for  $(Y^1(t))$  or  $(X^r(t))$  preserves the order relation (4.4), and if  $x_1 = z$ , this departure occurs at the same rate for both processes and, thus, the corresponding instant can be chosen at the same (exponential) time. For the departures from the first to the second queue, the departure rate for  $(X^r(t))$  is  $\mu r(x_0, x_1)(x_1 \vee 1) \mathbf{1}_{\{x_0 > 0\}} \leq \mu(z \vee 1)$ , which is the departure rate for  $(Y^1(t))$ ; hence, the corresponding departure instants can be taken in the reverse order so that (4.4) also holds at the next jump instant. The first part of the proposition is proved.

The rest of the proof is done in a similar way. The initial states  $X^r(0) = (x_0, x_1)$  and  $Y^\delta(0) = (y, z)$  are such that  $x_0 \leq y$  and  $x_1 \geq z$ . With the killing of the processes at time  $\tau_\delta \wedge \sigma$  we can assume additionally that  $x_0 \neq 0$  and that the relation  $r(x_0, x_1) \geq \delta$  holds. Under these assumptions, we can check by inspecting the next transition that (4.5) holds. The proposition is proved.

**Proposition 4.2.** *Under the condition  $\mu < \nu$ , the relation*

$$\liminf_{t \rightarrow +\infty} \frac{X_0^r(t)}{t} \geq \lambda - \lambda^*$$

*holds almost surely. In particular, if  $\mu < \nu$  and  $\lambda > \lambda^*$ , then the process  $(X^r(t))$  is transient.*

*Proof.* By Proposition 4.1 we can assume that there exists a version of  $(Y^1(t))$  such that  $X_0^r(0) = Y_0^1(0)$  and that the relation  $X_0^r(t) \geq Y_0^1(t)$  holds for any  $t \geq 0$ . From definition (4.2) of the  $Q$ -matrix of  $(Y^1(t))$ , we have, for  $t \geq 0$ ,

$$Y_0^1(t) = Y_0^1(0) + \mathcal{N}_\lambda(t) - A(t),$$

where  $(\mathcal{N}_\lambda(t))$  is a Poisson process with parameter  $\lambda$  and  $(A(t))$  is the number of arrivals (jumps of size 1) for the second coordinate  $(Y_1^1(t))$ : in particular,

$$E(A(t)) = \mu E\left(\int_0^t Y_1^1(s) \vee 1 \, ds\right).$$

Since  $(Y_1^1(t))$  is an ergodic Markov process under the condition  $\mu < \nu$ , the ergodic theorem in this setting gives

$$\lim_{t \rightarrow +\infty} \frac{1}{t} A(t) = \lim_{t \rightarrow +\infty} \frac{1}{t} E(A(t)) = \mu E(Y_1^1(\infty) \vee 1) = \lambda^*,$$

by (4.3); hence,  $(Y_0^1(t)/t)$  converges almost surely to  $\lambda - \lambda^*$ . The proposition is proved.

The next result establishes the ergodicity result of this section.

**Proposition 4.3.** *If the rate function  $r$  is such that, for any  $x_1 \in \mathbb{N}$ ,*

$$\lim_{x_0 \rightarrow +\infty} r(x_0, x_1) = 1, \tag{4.6}$$

*and if  $\mu \geq \nu$ , or if  $\mu < \nu$  and  $\lambda < \lambda^*$  with*

$$\lambda^* = \frac{\mu}{(1 - \rho)(1 - \log(1 - \rho))},$$

*and  $\rho = \mu/\nu$ , then  $(X^r(t))$  is an ergodic Markov process.*

Note that condition (4.6) is satisfied for the functions  $r$  considered in the models of Núñez-Queija and Prabhu [12] and Susitaival *et al.* [18]. See above.

*Proof of Proposition 4.3.* If  $x = (x_0, x_1) \in \mathbb{R}^2$ ,  $|x|$  denotes the norm of  $x$ ,  $|x| = |x_0| + |x_1|$ . The proof uses Foster’s criterion as stated in [14, Theorem 9.7]. If there exist constants  $K_0$ ,  $K_1$ ,  $t_0$ , and  $t_1$  such that

$$\sup_{x_1 \geq K_1} E_{(x_0, x_1)}(|X^r(t_1)| - |x|) < 0, \tag{4.7}$$

$$\sup_{x_1 < K_1, x_2 \geq K_2} E_{(x_0, x_1)}(|X^r(t_0)| - |x|) < 0, \tag{4.8}$$

then the Markov process  $(X^r(t))$  is ergodic.

Relation (4.7) is straightforward to establish: if  $x_1 \geq K_1$ , we obtain, by considering only  $K_1$  of the  $x_1$  initial servers in the second queue and the Poisson arrivals,

$$E_{(x_0, x_1)}(|X^r(1)| - |x|) \leq \lambda - K_1(1 - e^{-\nu});$$

hence, it is enough to take  $t_1 = 1$  and  $K_1 = (\lambda + 1)/(1 - e^{-\nu})$  to establish relation (4.7).

Now we establish inequality (4.8). Let  $\tau_\delta$  and  $\sigma$  be the stopping times introduced in Proposition 4.1. We first prove an intermediate result: for any  $t > 0$  and any  $x_1 \in \mathbb{N}$ ,

$$\lim_{x_0 \rightarrow +\infty} P_{(x_0, x_1)}(\sigma \wedge \tau_\delta \leq t) = 0. \tag{4.9}$$

Fix  $x_1 \in \mathbb{N}$  and  $t \geq 0$ . For  $\varepsilon > 0$ , there exists  $D_1$  such that

$$P_{x_1} \left( \sup_{0 \leq s \leq t} Y_1^1(s) \geq D_1 \right) \leq \varepsilon;$$

from Proposition 4.1, this gives the relation valid for all  $x_0 \geq 0$ :

$$P_{(x_0, x_1)} \left( \sup_{0 \leq s \leq t} X_1^r(s) \geq D_1 \right) \leq \varepsilon.$$

By condition (4.6), there exists  $\gamma \geq 0$  (that depends on  $x_1$ ) such that  $r(x_0, x_1) \geq \delta$  when  $x_0 \geq \gamma$ . As long as  $(X^r(t))$  stays in the subset  $\{(y_0, y_1) : y_1 \leq D_1\}$ , the transition rates of the first component  $(X_0^r(t))$  are uniformly bounded. Consequently, there exists  $K$  such that, for  $x_0 \geq K$ ,

$$P_{(x_0, x_1)} \left( \sup_{s \leq t} X_0^r(s) \leq \gamma, \sup_{s \leq t} X_1^r(s) \leq D_1 \right) \leq \varepsilon.$$

Relation (4.9) follows from the last two inequalities and the identity

$$P_{(x_0, x_1)}(\sigma \wedge \tau_\delta \leq t) \leq P_{(x_0, x_1)} \left( \sup_{s \leq t} X_0^r(s) \leq \gamma \right).$$

We return to the proof of inequality (4.8). By definition of the  $Q$ -matrix of the process  $(X^r(t))$ ,

$$E_{(x_0, x_1)}(|X^r(t)| - |x|) = \lambda t - \nu \int_0^t E_{(x_0, x_1)}(X_1^r(u)) du, \quad x \in \mathbb{N}^2, t \geq 0.$$

For any  $x \in \mathbb{N}^2$ , there exists a version of  $(Y^\delta(t))$  with initial condition  $Y^\delta(0) = X^r(0) = x$ , and such that relation (4.5) holds for  $t < \tau_\delta \wedge \sigma$ ; in particular,

$$\begin{aligned} E_x(X_1^r(t)) &\geq E_x(X_1^r(t); t < \tau_\delta \wedge \sigma) \\ &\geq E_x(Y_1^\delta(t); t < \tau_\delta \wedge \sigma) \\ &= E_x(Y_1^\delta(t)) - E_x(Y_1^\delta(t); t \geq \tau_\delta \wedge \sigma). \end{aligned}$$

The Cauchy–Schwarz inequality shows that, for any  $t \geq 0$  and  $x \in \mathbb{N}^2$ ,

$$\begin{aligned} \int_0^t E_x(Y_1^\delta(u); \tau_\delta \wedge \sigma \leq u) \, du &\leq \int_0^t \sqrt{E_x((Y_1^\delta(u))^2)} \sqrt{P_x(\tau_\delta \wedge \sigma \leq u)} \, du \\ &\leq \sqrt{P_x(\tau_\delta \wedge \sigma \leq t)} \int_0^t \sqrt{E_x((Y_1^\delta(u))^2)} \, du. \end{aligned}$$

By gathering these inequalities, and by using the fact that the process  $(Y_1^\delta(t))$  depends only on  $x_1$  and not  $x_0$ , we finally obtain the relation

$$\frac{1}{t} E_x(|X(t)| - |x|) \leq \lambda - \frac{\nu}{t} \int_0^t E_{x_1}(Y_1^\delta(u)) \, du + c(x_1, t) \sqrt{P_x(\tau_\delta \wedge \sigma \leq t)} \tag{4.10}$$

with

$$c(x_1, t) = \frac{\nu}{t} \int_0^t \sqrt{E_{x_1}((Y_1^\delta(u))^2)} \, du.$$

Two cases are considered.

(i) If  $\mu > \nu$ , and if  $\delta < 1$  is such that  $\delta\mu > \nu$ , the process  $(Y_1^\delta(t))$  is transient, so that

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t E_{x_1}(Y_1^\delta(u)) \, du = +\infty$$

for each  $x_1 \geq 0$ .

(ii) If  $\mu < \nu$ , we take  $\delta = 1$ , or if  $\mu = \nu$ , we take  $\delta < 1$  close enough to 1 so that  $\lambda < \lambda^*(\delta)$ . In both cases,  $\lambda < \lambda^*(\delta)$  and the process  $(Y_1^\delta(t))$  converges in distribution; hence,

$$\lim_{t \rightarrow +\infty} \frac{\nu}{t} \int_0^t E_{x_1}(Y_1^\delta(u)) \, du = \nu E(Y_1^\delta(\infty)) = \lambda^*(\delta) > \lambda$$

for each  $x_1 \geq 0$ .

Consequently, in both cases, there exist constants  $\eta > 0$ ,  $\delta < 1$ , and  $t_0 > 0$  such that, for any  $x_1 \leq K_1$ ,

$$\lambda - \frac{\nu}{t_0} \int_0^{t_0} E_{x_1}(Y_1^\delta(u)) \, du \leq -\eta. \tag{4.11}$$

With relation (4.10), we find that if  $x_1 \leq K_1$  then

$$\frac{1}{t_0} E_x(|X(t_0)| - |x|) \leq -\eta + c^* \sqrt{P_x(\tau_\delta \wedge \sigma \leq t_0)},$$

where  $c^* = \max(c(n, t_0), 0 \leq n \leq K_1)$ . By identity (4.9), there exists  $K_0$  such that, for all  $x_0 \geq K_0$  and  $x_1 \leq K_1$ , the relation

$$c^* \sqrt{P_{(x_0, x_1)}(\tau_\delta \wedge \sigma \leq t_0)} \leq \frac{\eta}{2}$$

holds. This relation and inequalities (4.10) and (4.11) give inequality (4.8). The proposition is proved.

### 4.3. Another boundary condition

The boundary condition  $x_1 \vee 1$  in the transition rates of  $(X(t))$ , (4.1), prevents the second coordinate from ending up in the absorbing state 0. It amounts to supposing that a permanent server gets activated when no node may offer the file. Another way to avoid this absorbing state is to suppose that a permanent node is always active, which gives transition rates with  $x_1 + 1$  instead. This choice was, for instance, made in [12]. All our results apply for this other boundary condition: the only difference is that, when  $\nu > \mu$ , the value of the threshold  $\lambda^*$  in (4.3) is given by the quantity  $\lambda^* = \mu\nu/(\nu - \mu)$ .

## 5. The two-chunk network

In this section it is assumed that a file of two chunks is distributed by the file-sharing system corresponding to Figure 1. Chunks are delivered in sequential order. The analysis of this simple file-sharing system gives a hint of the difficulties one can encounter when dealing with multiple chunks, in particular with nodes distributing *and* receiving chunks.

For  $k = 0, 1$  and  $t \geq 0$ , the variable  $X_k(t)$  denotes the number of requests downloading the  $(k+1)$ th chunk; for  $k = 2$ ,  $X_2(t)$  is the number of requests having all the chunks. When taking into account the boundaries in the transition rates described in Figure 1, we obtain the following  $Q$ -matrix for the three-dimensional Markov process  $(X(t)) = (X_k(t), k = 0, 1, 2)$ :

$$Q(f)(x) = \lambda[f(x + e_0) - f(x)] + \mu_1(x_1 \vee 1)[f(x + e_1 - e_0) - f(x)] \mathbf{1}_{\{x_0 > 0\}} + \mu_2(x_2 \vee 1)[f(x + e_2 - e_1) - f(x)] \mathbf{1}_{\{x_1 > 0\}} + \nu x_2[f(x - e_2) - f(x)],$$

where  $x \in \mathbb{N}^3$ ,  $f: \mathbb{N}^3 \rightarrow \mathbb{R}_+$  is a function and, for  $0 \leq k \leq 2$ ,  $e_k \in \mathbb{N}^3$  is the  $k$ th unit vector. Note that, as before, to avoid absorbing states, it is assumed that there is a server for the  $k$ th chunk when  $x_k = 0$ .

The stability behavior of  $(X(t))$  depends on the values of the parameters  $\mu_1$ ,  $\mu_2$ , and  $\nu$ . Three cases need to be distinguished: (i)  $\mu_1 > \mu_2 - \nu > 0$ , (ii)  $\mu_2 - \nu > \mu_1$ , and (iii)  $\mu_2 < \nu$ . In each case, the method of proof is similar to the one used in the proof of Proposition 4.3 and relies on Foster’s criterion, as stated in [14, Theorem 9.7]. We investigate the case when the size  $x_0 + x_1 + x_2$  of the initial state  $X(0) = x = (x_0, x_1, x_2)$  is large. We briefly discuss the three situations which will be analyzed in the following.

- (i) If  $x_2$  is large then the total number of customers decreases instantaneously.
- (ii) If  $x_1$  is large and  $x_2$  is small, then the last queue  $(X_2(t))$  behaves like the birth-and-death process  $(Z(t))$  of Section 3, which is transient if  $\mu_2 > \nu$  and stable otherwise.
- (iii) If  $x_0$  is large, and  $x_1$  and  $x_2$  are small, the two last queues  $(X_1(t), X_2(t))$  behave like the system  $(X^S(t)) = (X_1^S(t), X_2^S(t))$ , defined by its generator

$$Q^S(f)(z) = \mu_1(z_1 \vee 1)[f(z + e_1) - f(z)] + \mu_2(z_2 \vee 1)[f(z + e_2 - e_1) - f(z)] \mathbf{1}_{\{z_1 > 0\}} + \nu z_2[f(z - e_2) - f(z)],$$

where  $z \in \mathbb{N}^2$  and  $f: \mathbb{N}^2 \rightarrow \mathbb{R}_+$  is an arbitrary function.

Since  $(X^S(t))$  corresponds to the two last queues when the first queue is saturated, we call it the *saturated system*: it plays a crucial role in the stability of  $(X(t))$ , in particular the asymptotic behavior of  $E(X_2^S(t))$  determines the output of the system when the first queue is saturated. The properties of  $(X^S(t))$  that will be needed are gathered in the next proposition.

**Proposition 5.1.** (Saturated system.) *Let  $x \in \mathbb{N}^2$ .*

- (i) *If  $\mu_1 > \mu_2 - \nu > 0$  then  $E_x(X_2^S(t)) \rightarrow +\infty$  as  $t \rightarrow +\infty$ .*
- (ii) *If  $\mu_2 - \nu > \mu_1$  then  $(X^S(t))$  is positive recurrent.*
- (iii) *If  $\nu > \mu_2$  then  $\nu E_x(X_2^S(t)) \rightarrow \lambda^*$  as  $t \rightarrow +\infty$ , with  $\lambda^*$  given by*

$$\lambda^* = \frac{\mu_2}{(1 - \mu_2/\nu)(1 - \log(1 - \mu_2/\nu))}. \tag{5.1}$$

*Proof.* The crucial observation is that as long as  $X_1^S(t) > 0$ , the process  $(X^S(t))$  can be coupled with the process  $(W(t), Z(t))$  of the last part of Section 3, i.e.  $(X_2^S(t))$  behaves like a birth-and-death process and  $(X_1^S(t))$  is a Yule process with particles killed at the instants of birth of  $(X_2^S(t))$ .

More formally, in the sequel let  $(Z(t))$  be the process with  $Q$ -matrix defined by relation (3.1) with  $\mu_Z = \mu_2$ , and define  $(\sigma_n)$  to be its sequence of birth times. Let  $(W(t))$  be a Yule process with parameter  $\mu_1$  killed at  $(\sigma_n)$ . Then  $(X^S(t))$  and  $(W(t), Z(t))$  can be coupled so that

$$(X_1^S(t), X_2^S(t), 0 \leq t \leq T) = (W(t), Z(t), 0 \leq t \leq T)$$

with  $T = \inf\{t \geq 0: X_1^S(t) = 0\}$ . We give a full proof of the two first cases; the last case uses similar techniques and so we leave the details to the reader.

*Proof of case (i).* Since  $\mu_2 > \nu$ , the process  $(e^{-(\mu_2-\nu)t} Z(t))$  converges almost surely to a finite and positive random variable  $M_Z(\infty)$  by Corollary 3.2. Moreover, since  $\mu_1 > \mu_2 - \nu > 0$ , Corollary 3.1 entails that

$$\sum_{n \geq 1} e^{-\mu_1 \sigma_n} < +\infty \quad \text{almost surely.}$$

By Proposition 2.1, this shows that  $(W(t))$  survives with positive probability. In this event, say  $\mathcal{E}$ , it holds that  $W(t) \geq 1$  at all times and  $\lim_{t \rightarrow +\infty} Z(t) = +\infty$ . In the event  $\mathcal{E}$ ,  $W(t) \geq 1$  and, therefore,  $(X_1^S(t), X_2^S(t)) = (W(t), Z(t))$ ; in particular, we have exhibited an event of positive probability where  $X_2^S(t) \rightarrow +\infty$ , which proves the claim. Note that we implicitly assumed that no coordinate of the initial state is 0, these cases being dealt with easily.

*Proof of case (ii).* This is the most delicate case. By Proposition 3.4 and since  $\mu_2 - \nu > \mu_1$ ,  $(X_1^S(t))$  returns infinitely often to 0. When  $(X_1^S(t))$  is at 0, it jumps to 1 after an exponential time with parameter  $\mu_1$ . We define the sequences  $(H_{0,n})$ ,  $(E_{\mu,n})$ , and  $(S_n)$  by induction:  $S_0 = 0$  by convention, and, for  $n \geq 1$ ,

$$S_n = \sum_{k=1}^n (H_{0,k} + E_{\mu,k}),$$

$$H_{0,n} = \inf\{t \geq S_{n-1}: X_1^S(t) = 0\}, \quad \text{and} \quad E_{\mu,n} = \inf\{t \geq S_{n-1} + H_{0,n}: X_1^S(t) = 1\}.$$

The sequence  $(E_{\mu,n})$  represents the successive times at 0,  $(S_n)$  represents the times at which  $(X_1^S(t))$  hits 0, and  $(H_{0,n})$  represents the times needed for  $(X_1^S(t))$  to hit 0. On the time intervals  $[S_n, S_n + H_{0,n+1}]$ ,  $(X^S(t))$  can be coupled with a version of  $(W(t), Z(t))$ . This shows that, with the notation of Proposition 3.4,  $X_2^S(S_n)$  is equal in distribution to  $Z(H_0)$  with the initial condition  $(W(0), Z(0)) = (X_1^S(S_{n-1}), X_2^S(S_{n-1}))$ . Then while  $X_1^S$  is at 0, the dynamics of  $X_2^S$

are simple, since it just empties. These two remarks show that  $X_2(S_{n+1})$  is equal in distribution to

$$(X_2^S(S_{n+1}) \mid X^S(S_n) = x) \stackrel{D}{=} \left( \sum_{i=1}^{Z(H_0)} \mathbf{1}_{\{E_{v,i} > E_{\mu_1}\}} \mid (W(0), Z(0)) = x \right),$$

with  $(E_{v,i})$  i.i.d., exponentially distributed with parameter  $v$ , and independent of  $E_{\mu_1}$ . For  $i \geq 1$ ,  $E_{v,i}$  represents the residence time of the  $i$ th customer present in the second queue at time  $S_n + H_{0,n+1}$ .

Consequently, and since  $X_1^S(S_n) = 1$  for  $n \geq 1$ ,  $(X_2^S(S_n), n \geq 1)$  is a Markov chain whose transitions are defined by relation (3.8) with  $p = v/(v + \mu_1)$ ; note that  $(X_2^S(S_n), n \geq 0)$  has the same dynamics only when  $X_1^S(0) = 1$ . In the sequel, define, for any  $K > 0$ , the stopping  $T_K$ :

$$T_K = \inf\{n \geq 0: X_2^S(S_n) \leq K, X_1^S(S_n) = 1\}.$$

The ergodicity of  $(X^S(t))$  will follow from the finiteness of  $E_{(x_1, x_2)}(T_K)$  for some large enough  $K$  and arbitrary  $x = (x_1, x_2) \in \mathbb{N}^2$ ; indeed,  $T_K$  is greater than the time needed to return to the finite set  $\{x_1, x_2: x_1 = 1, x_2 \leq K\}$ . The strong Markov property of  $(X^S(t))$  applied at time  $S_1$  gives

$$E_{(x_1, x_2)}(T_K) \leq 2 E_{(x_1, x_2)}(S_1) + E_{(x_1, x_2)}(E_{(1, X_2^S(S_1))}(T_K)),$$

and so we need to study only  $T_K$  conditioned on  $\{X_1^S(0) = 1\}$  since  $E_{(x_1, x_2)}(S_1)$  is finite in view of Proposition 3.4. Thus, we can assume that  $(X_2^S(S_n), n \geq 0)$  is a time-homogeneous Markov chain whose transitions are defined by relation (3.8), and, with  $N_K$  defined in Proposition 3.5, the identity

$$T_K = \sum_{i=1}^{N_K} (H_{0,i} + E_{\mu_1, i}) \tag{5.2}$$

holds. For  $i \geq 0$ , the Markov property of  $(X_2^S(S_n), n \geq 0)$  gives, for any  $x_2 \geq 0$ ,

$$E_{(1, x_2)}(H_{0,i} \mathbf{1}_{\{i \leq N_K\}}) = E_{(1, x_2)}(E_{(1, X_2^S(S_i))}(H_0 \mathbf{1}_{\{i \leq N_K\}})).$$

With the same argument as in the proof of Proposition 3.5, we have

$$E_{(1, X_2^S(S_i))}(H_0) \leq E_{(1, 0)}(H_0) < +\infty;$$

with (5.2) and (3.9) of Proposition 3.5, we obtain, for some  $\gamma > 0$  and some  $K > 0$ ,

$$E_{(x_1, x_2)}(T_K) \leq 2 E_{(x_1, x_2)}(S_1) + C E_{(x_1, x_2)}[\log(1 + X_2^S(S_1))],$$

where the constant  $C = (E_{(1, 0)}(H_0) + 1/\mu_2)/\gamma$ . This last term is finite for any  $(x_1, x_2)$  in view of Proposition 3.4, which proves the proposition.

*Sketch of the proof of case (iii).* Since  $\mu_1 > 0 > \mu_2 - v$ , Corollary 3.1 shows that  $\sum_{n \geq 1} e^{-\mu_1 \sigma_n}$  is finite almost surely, and Proposition 2.1 implies that  $(X_1^S(t))$  survives with positive probability. Using the fact that the process  $(Z(t))$  is stable since  $\mu_2 < v$ , together with similar arguments as before, we can show that the process  $(X_1^S(t))$  actually survives almost surely. Thus, after some time  $(X_2^S(t))$  can be coupled forever with  $(Z(t))$ , and  $\lambda^*$  is precisely the value of  $v E(Z(\infty))$  in the stationary regime.



**Proposition 5.2.** *Let  $(X(t)) = (X_0(t), X_1(t), X_2(t))$  be the Markov process describing the linear file-sharing network with parameters  $\lambda, \mu_1, \mu_2,$  and  $\nu$ .*

- (i) *If  $\mu_1 > \mu_2 - \nu > 0$  then  $(X(t))$  is ergodic for any  $\lambda > 0$ .*
- (ii) *If  $\mu_2 - \nu > \mu_1$ , let  $\lambda^S$  be the value of the expectation of  $\nu X_2^S$  at equilibrium, which exists by Proposition 5.1. Then  $(X(t))$  is ergodic for  $\lambda < \lambda^S$  and transient for  $\lambda > \lambda^S$ .*
- (iii) *If  $\nu > \mu_2$  then  $(X(t))$  is ergodic for  $\lambda < \lambda^*$  and transient for  $\lambda > \lambda^*$ , with  $\lambda^*$  given by (5.1).*

In the second case, preliminary investigations seem to suggest that  $\lambda^S = +\infty$  is possible. In such situations, the stability region would actually be infinite. The intuitive reason is that  $(X_2^S(t))$  grows exponentially fast in periods where  $X_1^S(t) > 0$ , and for a duration  $H_0$ . Thus, when  $(X_1^S(t))$  hits 0,  $X_2^S(H_0) \approx e^{(\mu_2 - \nu)H_0}$ , and although some exponential moments of  $H_0$  are finite by (3.7), it seems plausible that  $E(e^{(\mu_2 - \nu)H_0}) = +\infty$  sometimes.

*Proof of Proposition 5.2.* In all three cases the proof relies on Foster’s criterion, as stated in Theorem 9.7 of [14]. Let  $X(0) = x = (x_0, x_1, x_2) \in \mathbb{N}^3$ , and assume that  $\|x\| = x_0 + x_1 + x_2$  is large. First we inspect the case when  $x_2$  is large, which is common to all three cases.

Since the last queue serves each request at rate  $\nu$ , for  $t \geq 0$ ,

$$E_x(\|X(t)\|) \leq \|x\| + \lambda t - x_2(1 - e^{-\nu t}).$$

Define  $t_2 = 1$ , and let  $K_2$  be such that  $\lambda - K_2(1 - e^{-\nu}) \leq -1$ . Then the relation

$$E_x(\|X(t_2)\|) - \|x\| \leq -1$$

holds when  $x_2 \geq K_2$  for all  $x_0, x_1 \geq 0$ . In the sequel, the constants  $t_2$  and  $K_2$  are fixed, and the process  $(Z(t))$  is defined similarly as in the proof of Proposition 5.1. We give a full proof of case (i); the other cases use similar techniques and so we leave the details to the reader.

*Proof of case (i).* Assume that  $\mu_1 > \mu_2 - \nu > 0$ : in particular,  $(Z(t))$  is transient; thus there exists some  $t_1$  such that, for any  $x_2 < K_2$ ,

$$\nu \int_0^{t_1} E_{x_2}(Z(u)) \, du \geq \lambda t_1 + 2.$$

The two processes  $(Z(t))$  and  $(X(t))$  can be built on the same probability space such that if they start from the same initial state then the two processes  $(Z(t))$  and  $(X_2(t))$  are identical as long as  $X_1(t)$  stays positive. Since, moreover, the hitting time  $\inf\{t \geq 0 : X_1(t) = 0\}$  goes to  $\infty$  as  $x_1$  goes to  $\infty$  for any  $x_2 < K_2$ , there exists  $K_1$  such that if  $x_1 \geq K_1$  and  $x_2 < K_2$ , then the relation

$$\begin{aligned} E_x(\|X(t_1)\|) - \|x\| &= \lambda t_1 - \nu \int_0^{t_1} E_x(X_2(u)) \, du \\ &\leq \lambda t_1 - \left( \nu \int_0^{t_1} E_{x_2}(Z(u)) \, du - 1 \right) \\ &\leq -1 \end{aligned}$$

holds.

In a similar way, by saturating the first queue and coupling the two last queues  $(X_1(t), X_2(t))$  with the saturated system  $(X^S(t))$ , we find in view of Proposition 5.1(i) that there exist constants  $t_0$  and  $K_0$  such that if  $x_2 < K_2, x_1 < K_1$ , and  $x_0 \geq K_0$ , then

$$E_x(\|X(t_0)\|) - \|x\| \leq -1,$$

which completes the proof of this case.

*Sketch of the proof of case (ii).* Since  $\mu_2 > \nu$ , the process  $(Z(t))$  is still transient and we can choose the same constants,  $t_1$  and  $K_1$ , as in the previous case to deal with the case  $x_1 \geq K_1$  and  $x_2 < K_2$ . When  $x_0 \geq K_0, x_1 < K_1$ , and  $x_2 < K_2$ , then we can justify for large enough  $K_0$  and  $t_0$  the approximation

$$E_x(\|X(t_0)\|) - \|x\| \approx \lambda t_0 - \nu \int_0^{t_0} E_{(x_1, x_2)}(X_2^S(u)) du \approx (\lambda - \lambda^S)t_0.$$

Hence, for  $\lambda < \lambda^S$ , the stability result follows. For  $\lambda > \lambda^S$ , this suggests that the first coordinated  $(X_0(t))$  builds up, and similarly as in the proof of Proposition 4.2 we can show that indeed in this case

$$\liminf_{t \rightarrow +\infty} \frac{X_0(t)}{t} \geq \lambda - \lambda^S > 0 \quad \text{almost surely.}$$

*Sketch of the proof of case (iii).* Since  $\nu > \mu_2$ , the process  $(Z(t))$  is stable, and it acts as a bottleneck on the system. When  $x_0 < K_0$  and  $x_1 \geq K_1$ , then we can, as before, justify the approximation, for large enough  $K_1$  and  $t_1$ ,

$$E_x(\|X(t_1)\|) - \|x\| \approx \lambda t_1 - \nu \int_0^{t_1} E_{x_2}(Z(u)) du \approx (\lambda - \lambda^*)t_1.$$

Similarly, the third case of Proposition 5.1 shows that when  $x_0 \geq K_0$  and  $x_1 < K_1, x_2 < K_2$ , again for large enough  $K_0$  and  $t_0$ ,

$$E_x(\|X(t_0)\|) - \|x\| \approx \lambda t_0 - \nu \int_0^{t_0} E_{(x_1, x_2)}(X_2^S(u)) du \approx (\lambda - \lambda^*)t_0.$$

This shows that the system is ergodic for  $\lambda < \lambda^*$ . Transience in the case  $\lambda > \lambda^*$  follows easily from the fact that the last queue  $(X_2(t))$  is stochastically dominated by  $(Z(t))$ ; thus,  $\lambda^*$  is the maximal output rate of the system.

### Appendix A. Proof of Proposition 3.3

Here the notation of Section 3 is used. Since the random variable  $(B_\sigma(t) \mid Z(0) = 0)$  is stochastically smaller than  $(B_\sigma(t) \mid Z(0) = z)$  for any  $z \in \mathbb{N}$ , it is enough to show that, for  $\eta < \eta^*(\nu/\mu_Z)$ ,

$$E_0\left(\sup_{t \geq \sigma_1} (e^{\eta\alpha t} B_\sigma(t)^{-\eta})\right) < +\infty,$$

where  $\alpha = \mu_Z - \nu > 0$ .

Note that the process  $(B_\sigma(t + \sigma_1), t \geq 0)$  under  $P_0$  has the same distribution as  $(B_\sigma(t) + 1, t \geq 0)$  under  $P_1$ , and by the independence of  $\sigma_1$ , an exponentially random variable with parameter  $\mu_Z$ , and  $(B_\sigma(t + \sigma_1), t \geq 0)$ , we obtain

$$E_0\left(\sup_{t \geq \sigma_1} (e^{\eta\alpha t} B_\sigma(t)^{-\eta})\right) = E_0(e^{\eta\alpha\sigma_1}) E_1\left(\sup_{t \geq 0} (e^{\eta\alpha t} (B_\sigma(t) + 1)^{-\eta})\right).$$

Since  $\alpha < \mu_Z$  and  $\eta^*(\nu/\mu_Z) < 1$ , then  $E_0(e^{\eta\alpha\sigma_1})$  is finite, and all we need to prove is that the second term is finite as well.

Define  $\tau$  as the last time  $Z(t) = 0$ :

$$\tau = \sup\{t \geq 0: Z(t) = 0\},$$

with the convention that  $\tau = +\infty$  if  $(Z(t))$  never returns to 0. Recall that, because of the assumption that  $\mu_Z > \nu$ , with probability 1, the process  $(Z(t))$  returns to 0 a finite number of times.

Conditioned on the event  $\{\tau = +\infty\}$ , the process  $(Z(t))$  is a  $(p, \lambda)$ -branching process conditioned on survival, with  $\lambda = \mu_Z + \nu$  and  $p = \mu_Z/\lambda$ . Such a branching process conditioned on survival can be decomposed as  $Z = Z_{(1)} + Y$ , where  $(Y(t))$  is a Yule process  $(Y(t))$  with parameter  $\alpha$ . See [3, Theorem 1, p. 48]. Consequently, for any  $0 < \eta < 1$ ,

$$E_1\left(\sup_{t \geq 0}(e^{\eta\alpha t}(B_\sigma(t) + 1)^{-\eta}) \mid \tau = +\infty\right) \leq E_1\left(\sup_{t \geq 0}(e^{\eta\alpha t}Y(t)^{-\eta})\right).$$

Since the  $n$ th split time  $t_n$  of  $(Y(t))$  is distributed like the maximum of  $n$  i.i.d. exponential random variables,  $Y(t)$  for  $t \geq 0$  is geometrically distributed with parameter  $1 - e^{-\alpha t}$ ; hence,

$$\begin{aligned} \sup_{t \geq 0}\left[e^{\eta\alpha t} E_1\left(\frac{1}{Y(t)^\eta}\right)\right] &= \sup_{t \geq 0}\left[e^{-(1-\eta)\alpha t} \sum_{k \geq 1} \frac{(1 - e^{-\alpha t})^{k-1}}{k^\eta}\right] \\ &\leq \sup_{0 \leq u \leq 1}\left[(1 - u)^{1-\eta} \sum_{k \geq 1} \frac{u^{k-1}}{k^\eta}\right]. \end{aligned}$$

For  $0 < u < 1$ , the relation

$$\begin{aligned} (1 - u)^{1-\eta} \sum_{k \geq 1} \frac{u^{k-1}}{k^\eta} &\leq (1 - u)^{1-\eta} \int_0^\infty \frac{u^x}{(1 + x)^\eta} dx \\ &= \left(\frac{1 - u}{-\log u}\right)^{1-\eta} \int_0^\infty \frac{e^{-x}}{(x - \log u)^\eta} dx, \end{aligned}$$

holds; hence,

$$\sup_{t \geq 0}\left[e^{\eta\alpha t} E_1\left(\frac{1}{Y(t)^\eta}\right)\right] < +\infty.$$

The process  $(e^{-\alpha t}Y(t))$  being a martingale, by convexity, the process  $(e^{\eta\alpha t}Y(t)^{-\eta})$  is a non-negative submartingale. For any  $\eta \in (0, 1)$ , Doob’s  $L_p$  inequality gives the existence of a finite  $q(\eta) > 0$  such that

$$E_1\left(\sup_{t \geq 0}(e^{\eta\alpha t}Y(t)^{-\eta})\right) \leq q(\eta) \sup_{t \geq 0}\left[e^{\eta\alpha t} E_1\left(\frac{1}{Y(t)^\eta}\right)\right] < +\infty.$$

The following result has therefore been proved.

**Lemma A.1.** *For any  $0 < \eta < 1$ ,*

$$E_1\left[\sup_{t \geq 0}(e^{\eta\alpha t}(B_\sigma(t) + 1)^{-\eta}) \mid \tau = +\infty\right] < +\infty.$$

On the event  $\{\tau < +\infty\}$ ,  $(Z(t))$  hits 0 a geometric number of times and then couples with a  $(p, \lambda)$ -branching process conditioned on survival. On this event,

$$\begin{aligned} \sup_{t \geq 0} (e^{\eta\alpha t} (B_\sigma(t) + 1)^{-\eta}) &= \max \left( \sup_{0 \leq t \leq \tau} (e^{\eta\alpha t} (B_\sigma(t) + 1)^{-\eta}), \sup_{t \geq \tau} (e^{\eta\alpha t} (B_\sigma(t) + 1)^{-\eta}) \right) \\ &\leq e^{\eta\alpha\tau} \left( 1 + \sup_{t \geq 0} (e^{\eta\alpha t} (B'_\sigma(t) + 1)^{-\eta}) \right), \end{aligned}$$

where  $B'_\sigma(t)$  for  $t \geq \tau$  is the number of births in  $(\tau, t]$  of a  $(p, \lambda)$ -branching process conditioned on survival and independent of the variable  $\tau$ ; consequently,

$$\begin{aligned} &E_1 \left( \sup_{t \geq 0} (e^{\eta\alpha t} (B_\sigma(t) + 1)^{-\eta}) \mid \tau < +\infty \right) \\ &\leq E_1 (e^{\eta\alpha\tau} \mid \tau < +\infty) \left( 1 + E_1 \left( \sup_{t \geq 0} (e^{\eta\alpha t} (B_\sigma(t) + 1)^{-\eta}) \mid \tau = +\infty \right) \right). \end{aligned}$$

In view of Lemma A.1, the proof of Proposition 3.3 will be complete if we can prove that

$$E_1 (e^{\eta\alpha\tau} \mid \tau < +\infty) < +\infty,$$

which actually comes from the following decomposition: under  $P_1(\cdot \mid \tau < +\infty)$ , the random variable  $\tau$  can be written as

$$\tau = \sum_{k=1}^{1+G} (T_k + E_{\mu_Z, k}),$$

where  $G$  is a geometric random variable with parameter  $q = \nu/\mu_Z$ ,  $(T_k)$  is an i.i.d. sequence with the same distribution as the extinction time of a  $(p, \lambda)$ -branching process starting with one particle and conditioned on extinction, and the  $(E_{\mu_Z, k})$  are i.i.d. exponential random variables with parameter  $\mu_Z$ .

Since  $q$  is the probability of extinction of a  $(p, \lambda)$ -branching process started with one particle,  $G + 1$  represents the number of times  $(Z(t))$  hits 0 before going to  $\infty$ . This representation entails

$$E_1 (e^{\eta\alpha\tau} \mid \tau < +\infty) = E(\gamma(\eta)^{G+1}), \quad \text{where } \gamma(\eta) = E(e^{\eta\alpha(T_1 + E_{\mu_Z, 1})}).$$

A  $(p, \lambda)$ -branching process conditioned on extinction is actually a  $(1 - p, \lambda)$ -branching process. See again [3, Theorem 3, p. 52]. Thus,  $T_1$  satisfies the following recursive distributional equation:

$$T_1 \stackrel{D}{=} E_\lambda + \mathbf{1}_{\{\xi=2\}}(T_1 \vee T_2),$$

where  $P(\xi = 2) = 1 - p$  and  $E_\lambda$  is an exponential random variable with parameter  $\lambda$ . This equation yields

$$P(T_1 \geq t) \leq e^{-\lambda t} + 2\lambda(1 - p) \int_0^t P(T_1 \geq t - u) e^{-\lambda u} du,$$

and Gronwall's lemma applied to the function  $t \mapsto e^{\lambda t} P(T_1 \geq t)$  gives

$$P(T_1 \geq t) \leq e^{(\lambda - 2\lambda p)t} = e^{(\nu - \mu_Z)t};$$

hence, for any  $0 < \eta < 1$ ,

$$E_1(e^{\eta\alpha T_1}) \leq \frac{1}{1-\eta}.$$

Since  $G$  is a geometric random variable with parameter  $q$ ,  $E(\gamma(\eta)^G)$  is finite if and only if  $\gamma(\eta) < q$ . Since, finally,

$$\gamma(\eta) = \frac{\mu_Z}{\mu_Z - \eta\alpha} E(e^{\eta\alpha T_1}) \leq \frac{\mu_Z}{(1-\eta)(\mu_Z - \eta\alpha)},$$

we can easily check that  $\gamma(\eta) < q$  for  $\eta < \eta^*(\nu/\mu_Z)$  as defined by (3.4), which concludes the proof of Proposition 3.3.

### Acknowledgements

This paper has benefited from various interesting discussions with S. Borst, I. Norros, R. Núñez-Queija, B. J. Prabhu, and H. Reittu.

### References

- [1] ALDOUS, D. AND PITMAN, J. (1998). Tree-valued Markov chains derived from Galton–Watson processes. *Ann. Inst. H. Poincaré Prob. Statist.* **34**, 637–686.
- [2] ALSMEYER, G. (1993). On the Galton–Watson predator-prey process. *Ann. Appl. Prob.* **3**, 198–211.
- [3] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
- [4] BRAMSON, M. (2008). *Stability of Queueing Networks* (Lecture Notes Math. **1950**). Springer, Berlin.
- [5] CHEN, H. AND YAO, D. D. (2001). *Fundamentals of Queueing Networks*. Springer, New York.
- [6] KINGMAN, J. F. C. (1975). The first birth problem for an age-dependent branching process. *Ann. Prob.* **3**, 790–801.
- [7] LAMPERTI, J. (1967). Continuous state branching processes. *Bull. Amer. Math. Soc.* **73**, 382–386.
- [8] LESKELÄ, L. (2010). Stochastic relations of random variables and processes. *J. Theoret. Prob.* **23**, 523–546.
- [9] MASSOULIÉ, L. AND VOJNOVIĆ, M. (2005). Coupon replication systems. In *Proc. ACM SIGMETRICS 2005*, ACM Press, New York, pp. 2–13.
- [10] NERMAN, O. (1981). On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitsthe.* **57**, 365–395.
- [11] NEVEU, J. (1986). Erasing a branching tree. In *Analytic and Geometric Stochastics (Papers in honour of G. E. H. Reuter; Adv. Appl. Prob. Spec. Suppl. 19)*, ed. D. G. Kendall, Applied Probability Trust, Sheffield, pp. 101–108.
- [12] NÚÑEZ-QUEIJA, R. AND PRABHU, B. J. (2008). Scaling laws for file dissemination in P2P networks with random contacts. In *Proc. 16th Internat. Workshop on Quality of Service*, IEEE, pp. 75–79.
- [13] QIU, D. AND SRIKANT, R. (2004). Modeling and performance analysis of BitTorrent-like peer-to-peer networks. In *Proc. 2004 Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communications*, ACM Press, New York, pp. 367–378.
- [14] ROBERT, P. (2003). *Stochastic Networks and Queues* (Appl. Math. **52**). Springer, Berlin.
- [15] ROGERS, L. C. G. AND WILLIAMS, D. (1987). *Diffusions, Markov Processes, and Martingales*, Vol. 2. John Wiley, New York.
- [16] SENETA, E. (1970). On the supercritical branching process with immigration. *Math. Biosci.* **7**, 9–14.
- [17] SIMATOS, F., ROBERT, P. AND GUILLEMIN, F. (2008). A queueing system for modeling a file sharing principle. In *Proc. ACM SIGMETRICS 2008*, ACM Press, New York, pp. 181–192.
- [18] SUSITAIVAL, R., AALTO, S. AND VIRTAMO, J. (2006). Analyzing the dynamics and resource usage of P2P file sharing by a spatio-temporal model. In *Computational Science – ICCS 2006* (Lecture Notes Comput. Sci. **3994**), Springer, Berlin, pp. 420–427.
- [19] TRANG, D. D., PERECZES, R. AND MOLNÁR, S. (2007). Modeling the population of file-sharing peer-to-peer networks with branching processes. In *IEEE Symp. Comput. Commun.* (Aveiro, Portugal, July 2007), pp. 809–815.
- [20] WILLIAMS, D. (1991). *Probability with Martingales*. Cambridge University Press.
- [21] YANG, X. AND DE VECIANA, G. (2004). Service capacity of peer to peer networks. In *Proc. IEEE INFOCOM*, ACM Press, New York, pp. 2242–2252.