Multi-Dimensional Markov Model for Performance Evaluation of an Ethernet Switch

Anatoli Manita, François Simonot, Ye-Qiong Song

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Multi-Dimensional Markov Model for Performance Evaluation of an Ethernet Switch

Anatoli Manita — François Simonot — YeQiong Song

N° 4813

Mai 2003

THÈME 1

ISSN 0249-6399 ISRN INRIA/RR--4813--FR+ENG

Multi-Dimensional Markov Model for Performance Evaluation of an Ethernet Switch

Anatoli Manita [*] , François Simonot [†] , YeQiong Song [‡]

Thème 1 —Réseaux et systèmes

Projet TRIO

Rapport de recherche n° 4813 —Mai 2003 — 34 pages

Abstract: With continuously increasing use of switched Ethernet for supporting various applications, it is important to precisely characterize the performance of Ethernet switches to provide the required QoS (Quality of Service). In this paper we propose an exact mathematical model of an Ethernet switch based on a multi-dimensional Markov process. We discuss both continuous and discrete time variants of the proposed model. We put forward the problem of asymptotic analysis of steady-state distribution of the workload to evaluate the performance of the system. This approach is deeply related with the recent results on large deviations in Markovian systems. We provide a comprehensive study of a particular model called symmetrical geometric. For the special case of Binomial input flow, the asymptotic approach is compared with the exact recurrent formulae highlighting the numerical efficiency of the asymptotic approach.

Key-words: Ethernet switch, Multi-dimensional Markov process, Asymptotic approach, Large deviation, Performance evaluation

[*] Department of Probability, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, 119992 Moscow, Russia. E-mail: manita@mech.math.msu.su . The work of this author was supported by Russian Foundation of Basic Research (project 02-01-00945) and Lyapunov French-Russian Institute (project 00-02).

[†] IECN, Université Henri Poincaré Nancy I, Esstin, 2, Rue J. Lamour, 54500 Vandoeuvre, France.
E-mail: francois.simonot@esstin.uhp-nancy.fr

[‡] LORIA - INRIA Lorraine, 2, avenue de la Forêt de Haye, 54516 Vandoeuvre-Lès-Nancy, France.
E-mail: song@loria.fr

Unité de recherche INRIA Lorraine
LORIA, Technopôle de Nancy-Brabois, Campus scientifique,

Modèle Markovien multidimensionnel pour l'évaluation de performances d'un commutateur Ethernet

Résumé : Ethernet commuté est de plus en plus utilisé comme réseau de support pour des applications demandant la garantie de la qualité de service. Il est donc important de caractériser précisément les performances des commutateurs Ethernet. Dans ce papier nous proposons un modèle Markovien multidimensionnel pour caractériser exactement le flux d'entrée dans un commutateur Ethernet, ceci à la fois en modèle continu et modèle discret. Nous mettons en avant l'approche de l'analyse asymptotique de la distribution stationnaire de la charge pour l'évaluation de performances du système. Cette approche est basée sur les résultats récents en larges déviations des systèmes Markoviens. Nous fournissons une étude exhaustive d'un modèle particulier appelé modèle géométrique symétrique. Pour le cas d'un flux d'entrée Binomial, cette approche asymptotique est comparée avec la formule exacte récurrentielle montre l'excellente qualité de cette approche asymptotique.

Mots-clés : Commutateur Ethernet, Processus de Markov multidimensionnel, Approche asymptotique, Large déviation, Evaluation de performances

1 Introduction

Ethernet switches are without any doubt the most widely spread communication switches. They are also more and more used for supporting real-time applications for which some performance guarantees must be provided. Two classes of real-time constraints: HRT (Hard Real-Time) and SRT (Soft Real-Time) should be met by a switched Ethernet. SRT constrained applications often require a probabilistic guarantee on the message response time.

By using switched Ethernet with full-duplex links instead of shared Ethernet, collisions can be completely eliminated. But the message response time can still be random because of intra-switch message buffering delay. The evaluation of this queueing delay distribution is thus important for the designer of a real-time application distributed around such a network.

The work presented in this paper is motivated by the analysis of the output queues of an $N \times N$ wire-speed non-blocking Ethernet switch (in fact almost all recent Ethernet switches are announced operating with wire-speed and non-blocking). Details on the switch functioning and the main switch fabrics can be found in [1]. Moreover the application context considered is factory communication in which per-stream oriented real-time guarantee is required. In such a context, a message stream generates a flow of constant size packets. Of course, different streams can have different packet lengths [2]. This led us to only consider the fixed length packets in our modeling and we assume that a message is always transmitted in one single packet without fragmentation.

Wire-speed means that all ports of a switch can simultaneously transmit or receive at their full bit rates. For any input port, that means the processing time of traffic classification and table lookup should not exceed the minimum packet inter-arrival interval (64 bytes minimum packet length + 8 bytes physical layer overhead + 96 bits IFG = 672 bit time). This requires that the switch fabric can operate at a bit rate equaling to the aggregate speeds of all the ports. A switch is non-blocking when it can forward a packet to the destination port as long as that port is free, while a blocking one may be not able to forward a packet although the destination port is free because of internal conflict in the switch fabric (one example is the HOL blocking in input buffering switches) [3]. Switches with output buffering are internally non-blocking. Switches with wire-speed and output buffering are non-blocking.

Buffering and thus buffering delay exists in a switch whatever the switch is with full wire-speed or not. In fact packet buffering occurs whenever the output port cannot forward all input packets at time. This corresponds to the burst traffic arrival. In a fully switched Ethernet there is only one equipment (station or switch) per switch port. The total delay introduced by a switch is composed of :

- The switching latency (traffic classification according to IEEE802.1p mapping table, destination port look-up and switch fabric set-up time),
- The packet forwarding time which depends on the forwarding mode (cut-through or store and forward) and eventually on the packet length if the store & forward mode is running,

- The buffering delay when the packet is queued.

The switching latency is a fixed value, which depends on the switch performance and often provided by the switch vendor (e.g. 11μs between 100Mbps ports and 70μs between 10Mbps ports for Cisco Catalyst 1900 and 2820). The packet forwarding time can be obtained knowing in which forwarding mode the switch is running. The technique for analyzing the buffering delay depends on the knowledge on the input traffic pattern. For periodic input traffic (or (σ,ρ)-bounded or sporadic majoring by a periodic one by taking the minimum inter arrival time as the data emission period), classic scheduling analysis [4] or (σ,ρ)-related analysis [5] can give the worst-case buffering delay, providing thus the HRT guarantee [6]. But for aperiodic input, since only few on traffic characteristics is known, a stochastic analysis is needed.

We focus our attention on the buffering delay distribution evaluation for providing SRT guarantee. Unlike for packet switches and ATM switches where exhaustive performance analysis has achieved, we can surprisingly find only measurement based performance analysis for Ethernet switches. Many measurement- based performance tests according to IETF benchmarks (RFC2285) can be found at switch suppliers web sites or for example in
(`www.NetworkComputing.com/815/815f1.html`).
The main performance parameters are RFC1944 throughput, frame/packet loss probability, many to one congestion handling capability and RFC2285 HOL (Head Of Line) blocking, X-stream performance, address-handling tests and illegal frame filtering tests. To our best knowledge, buffering delay in an Ethernet switch is never analytically addressed.

As Ethernet switch is a particular case of the packet switches. So the first question one can ask is which results on general packet switches can be used for the performance analysis of an Ethernet switch? In our previous work, the results presented in [3] and [7] have been analysed. The Markov-modulated model presented in [7] can not be directly used for our analysis as the packet lengths are assumed to be exponentially distributed whereas an Ethernet packet length is bounded between 64 bytes and 1500 bytes. Moreover the renewal process of the flow entering to the output buffer is difficult to justify as this approximation is based on their simulations which are difficult to reproduce in our context. The Binomial process used in [3] is more suitable although it can not model the exact behaviour of our system. In fact, because of the 96-bits inter-frame gap and the 64-bytes minimum packet size, there is a minimum inter-arrival interval for each input port. The resulting flow entering to an output buffer of an Ethernet switch is no longer an independent process. This makes difficult to accurately characterize the switch behaviours. In [6], we proposed to upper and lower bound the actual behaviours of such a system by considering two cases: worse case with Poisson input (for upper bound) and better case with the Binomial input (for lower bound). However, the Poisson arrival does not describe the exact behaviour of our switch as there can be more than N simultaneous arrivals. The binomial input limits effectively the simultaneous arrivals to N packet but it does not take into the eventual dependence between arrivals issued from the same input port (i.e. they should be spaced by at least the minimum inter-arrival interval). So an exact modelling is needed. Moreover a subsidiary problem is the lack of an numerically accurate method for the

tail of the distribution of the workload (from which we deduce the delay distribution). In fact, the direct use of the Markov chain balance equation [3] still meet the problem of numerical accuracy.

This paper presents the following two main contributions :

- An exact mathematical model of an Ethernet switch based on a multi-dimensional Markov is described
- The large-deviation technique is used to accurately compute the tail of the workload distribution

We notice that the packet buffering delay distribution is not explicitly addressed in this paper as it can be directly deduced from the workloads distributions of the output queues and the routing matrix. The rest of the paper is organized as follows. Section 2 firstly describes how an Ethernet switch works and then provides an exact mathematical model based on a multi-dimensional Markov description (in both continuous and discrete time). Section 3 presents the large-deviation technique that we will use to efficiently compute the steady-state probability distribution of the work load in an output buffer. Section 4 deals with a numerical case study and compares the numerical accuracy of our large-deviation method to that based on the balance equations [3]. Section 5 summarizes the contributions of the paper and outlines some extensions.

Acknowledgements. The first author is very grateful to the team TRIO (INRIA-Lorraine) for the hospitality during his stay at Nancy in summer 2002 when the main results of this paper were obtained.

2 Markovian description of the model

In this section we propose a probabilistic model of an Ethernet switch switch. We start from some basic notation, then we explain how the switch under consideration works and give the exact mathematical description of the model both in continuous and discrete time.

All random variables and stochastic processes considered below are supposed to be defined on the same probability space which we denote by $(\Omega, \mathcal{F}, \mathsf{P})$. Notation ω is used for elements of Ω.

2.1 General description of the model

2.1.1 Ethernet switch

We consider the following model of an Ethernet switch (see Figure 1). There are N input and N output ports. We denote the input ports by I_1, ..., I_N and the output ports by O_1, ..., O_N. Denote

by F_j an input flow to the input port j. A packet entering the input port j is directed to the output port m with probability R_{jm} independently of the other customers.

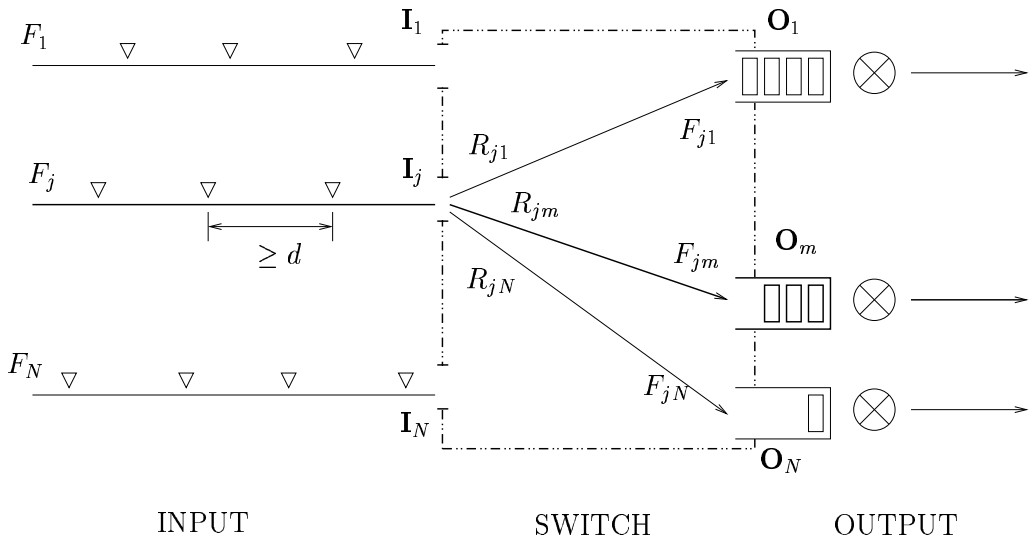


Figure 1: Communication switch

Each output port can be considered as a server node which transmits packets. Service time of a given packet is *proportional* to its length [1]. Each output port m has its own infinite transmission buffer in which packets sent to this output port are queued.

2.1.2 Input flows and subflows

In each input flow F_j packets arrive at epochs of some renewal process

$$\sigma_1^{(j)}, \sigma_2^{(j)}, \ldots, \sigma_k^{(j)}, \ldots \tag{1}$$

more precisely we assume that $\sigma_2^{(j)} - \sigma_1^{(j)}, \ldots, \sigma_{k+1}^{(j)} - \sigma_k^{(j)}, \ldots$ are independent identically distributed random variables. We assume that any packet from the flow F_j has the same (deterministic) length D_j, in other words, it requires the time D_j to pass an output port. We are mainly interested in input flows which have the following property: there exists $d > 0$ such that for all j, k

$$\mathsf{P}\left\{\sigma_{k+1}^{(j)} - \sigma_k^{(j)} \geq d\right\} = 1. \tag{2}$$

[1] More precisely, service_time = length / output_link_bit_rate. For convenience, we assume that output_link_bit_rate = 1.

We assume that all input flows $F_1, \ldots, F_j, \ldots, F_N$ are independent. Denote by F_{jm} the subflow

$$\sigma_1^{(j,m)}, \sigma_2^{(j,m)}, \ldots, \sigma_l^{(j,m)}, \ldots$$

of the flow F_j consisting of packets directed to some fixed output port m. Since each packet chooses an output port independently of the other packets, the subflow F_{jm} is still a renewal process.

2.1.3 Output port as a queueing system

Select and fix some output port m_0. Consider the composite input flow to the output port m_0:

$$s_1 \leq \ldots \leq s_n \leq \ldots . \tag{3}$$

It is the superposition of the flows F_{jm_0}, $j = 1, \ldots, N$, in the sense that

$$\{s_1(\omega), \ldots, s_n(\omega), \ldots\} = \bigcup_{j=1}^{N} \left\{\sigma_1^{(j,m_0)}(\omega), \sigma_2^{(j,m_0)}(\omega), \ldots\right\}.$$

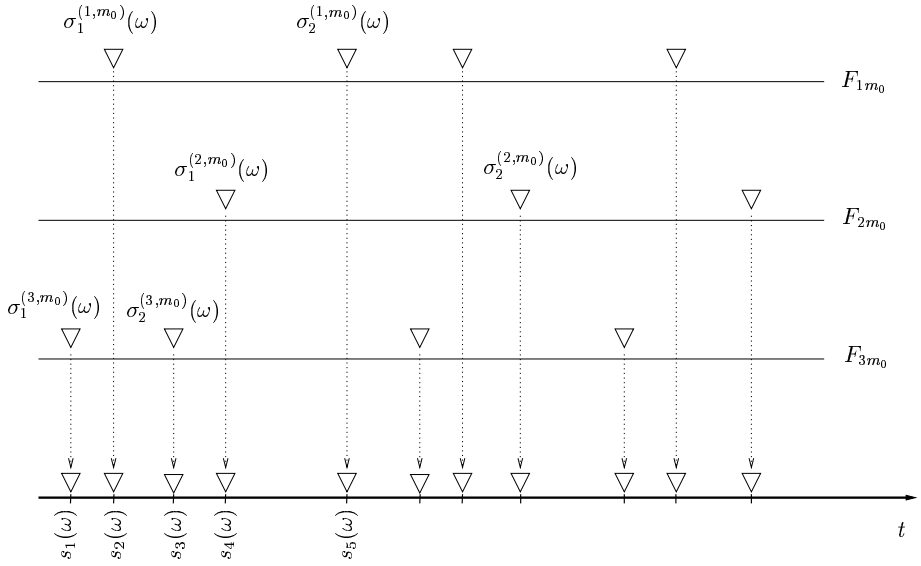


Figure 2: Composite flow $s_1(\omega), \ldots, s_n(\omega), \ldots$ as a superposition of subflows F_{jm_0}: case $N = 3$

It appears that this composite flow cannot be interpreted as a renewal process because, in general, intervals between arrivals

$$s_2 - s_1, \ldots, s_n - s_{n-1}, \ldots$$

are not independent.

So the output port m_0 can be considered as a queueing system with the input flow (3) and one server node.

We are going to study performance of this queueing system (Figure 3).

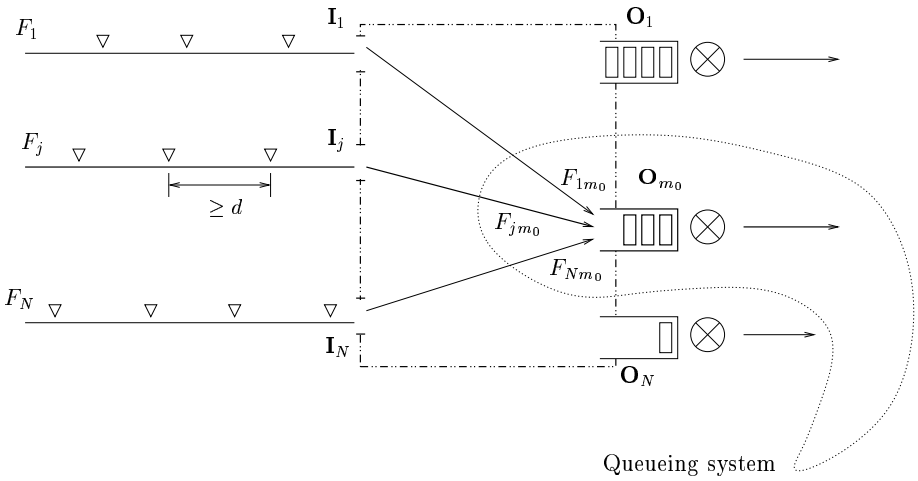


Figure 3: Queueing system associated with the output port O_{m_0}

We start from Markovian description of the model. The model can be considered in both continuous and discrete time. Let us consider these two descriptions separately.

2.2 Continuous time Markovian model

Main assumption here is that for any j the distribution of time intervals (1) is continuous, i.e., the probability distribution function

$$G_j(t) = \mathsf{P}\left\{\sigma_{k+1}^{(j)} - \sigma_k^{(j)} \leq t\right\} \tag{4}$$

is continuous in t.

2.2.1 Markovian description of the flows

First of all let us fix some j and give a Markovian description of the flow F_{jm_0}. Let $\zeta^{\circ,j}$ be a positive (real-valued) random variable with the same distribution as the intervals (1) and let $\zeta_1^{\circ,j}, \ldots, \zeta_i^{\circ,j}, \ldots$ be a sequence of independent copies of $\zeta^{\circ,j}$. Let $\nu^{(j)}$ be a geometric random variable,

$$\mathsf{P}\left\{\nu^{(j)} = l\right\} = R_{jm_0}\left(1 - R_{jm_0}\right)^l, \qquad l = 0, 1, \ldots \quad ,$$

independent of the sequence $\zeta_1^{\circ,j}, \ldots, \zeta_i^{\circ,j}, \ldots$. Define a random variable $\zeta^{(j)}$ as follows

$$\zeta^{(j)}(\omega) = \sum_{i=1}^{1+\nu^{(j)}(\omega)} \zeta_i^{\circ,j}(\omega).$$

It is clear that the probability distribution of $\zeta^{(j)}$ is the same as distribution of an interval between arrivals in the flow F_{jm_0}. Let the set $\left\{\zeta_t^{(j)}, t > 0\right\}$ consists of independent copies of $\zeta^{(j)}$.

Now we construct a continuous time stochastic process $r_j(t) = (r_j(t,\omega), t \geq 0)$ with the following trajectories:

- put $r_j(0,\omega) \equiv x_j$ for some $x_j > 0$,
- if $t > 0$ and $r_j(t,\omega) > 0$ then $\frac{d}{dt} r_j(t,\omega) = -1$,
- if $t > 0$ and $r_j(t,\omega) = 0$ then

$$r_j(t+0,\omega) - r_j(t,\omega) = \zeta_t^{(j)}(\omega).$$

It is clear that $r_j(t)$ is a Markov process with values in $\mathbf{R}_+$.

Moreover, there is a correspondence between the process $r_j(t)$ and the flow F_{jm_0} with fixed first arrival epoch: $\sigma_1^{(j,m_0)}(\omega) \equiv x_1$. To see this we should interpret the random variable $r_j(t)$ as the time remaining from the moment t to an arrival of the next packet from the flow F_{jm_0}. So an arrival in the flow F_{jm_0} occurs at time t iff $r_j(t) = 0$.

By assumption flows $F_1, \ldots, F_N$ are independent, hence the subflows $F_{1m_0}, \ldots, F_{Nm_0}$ are also independent and can be described by the Markov process

$$\mathbf{r}(t) := (r_1(t), \ldots, r_N(t)), \tag{5}$$

where the components $r_1(t), \ldots, r_N(t)$ are supposed to be independent and defined as above. Therefore the stochastic process $\mathbf{r}(t)$ with values in $\mathbf{R}_+^N$ contains all necessary information about the composite input flow to the output port m_0: an arrival in the composite flow occurs at time t iff there exists $j \in \{1, \ldots, N\}$ such that $r_j(t) = 0$.

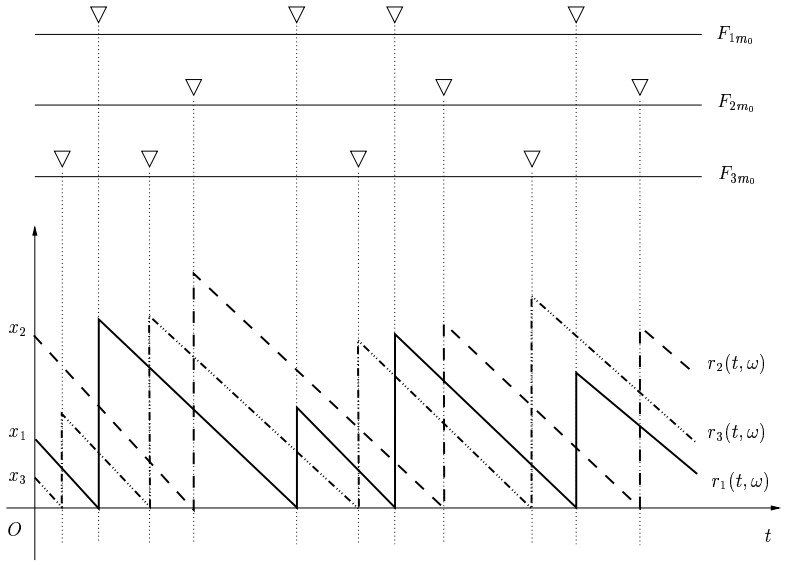


Figure 4: Correspondence between flows F_{1m_0}, F_{2m_0}, F_{3m_0} and stochastic processes $r_1(t), r_2(t), r_3(t)$

2.2.2 Markovian queueing system

When we observe the output port m_0 we are mainly interested in the total work time presented in it, i.e., the total length of packets in the queue + residual time of the current packet serving by the server node. Recall that the server node works with the constant speed (assumed to be equal to 1), so a service time of a given packet is equal to the length of this packet. Denote by $J(t)$ the total workload presented at time t [2]. Markovian description of the system is given by the multi-dimensional process

$$(r_1(t), \dots, r_N(t), J(t)) \in \mathbf{R}_+^{N+1},$$

where $r_1(t), \dots, r_N(t)$ are the same as in (5) and $J(t)$ is defined as follows

- $J(0,\omega) \equiv 0$,
- if $t > 0$ and $r_j(t,\omega) > 0\ \forall j$, then

$$\frac{d}{dt}J(t,\omega) = \begin{cases} -1, & \text{if } J(t,\omega) > 0, \\ 0, & \text{if } J(t,\omega) = 0, \end{cases}$$

[2] If $D_j = D$ for all j, then $J(t) = D\times$queue_length + residual_time_of_current_packet

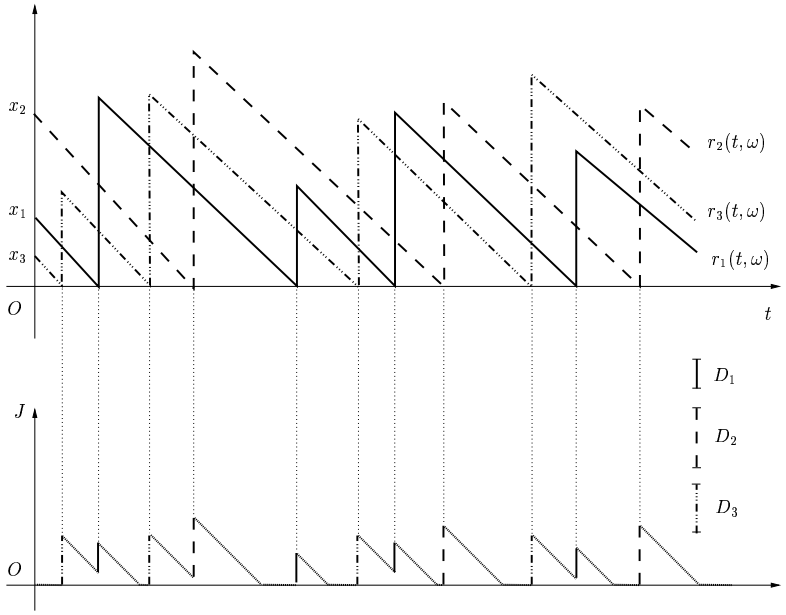


Figure 5: Trajectory of stochastic process $J(t) = J(t,\omega)$ coupled with samples $r_1(t,\omega),\ldots,r_N(t,\omega)$ in case $N=3$

- if $t>0$ and $\exists j_0:\ r_{j_0}(t,\omega)=0$, then

$$J(t+0,\omega)-J(t,\omega)=D_{j_0}.$$

This definition reflects the fact that at the moment when a packet coming from the input source j joins the queue at output source m_0 the total work to be done by the server m_0 increases by D_j.

2.2.3 Results expected from the Markovian description

1) If $\mathsf{E}\zeta^{\circ,j}=\int t\,dG_j(t)<\infty$ for all j then the Markov process $\mathbf{r}(t)$ is ergodic. Hence there exists a stationary version of the composite input flow to the selected output port m_0.

2) If the intensity of the composite input flow is sufficiently smallthen the Markov process $(r_1(t),\ldots,r_N(t),J(t))$ is ergodic and, its distribution at time t converges to the stationary distribution as $t\to\infty$. In particular, this gives an existence of the steady-state of $J(t)$ and provides a key to the problem of performance evaluation.

2.3 Discrete time Markovian model

In this section, the system is considered only at discrete epochs: $0, \Delta, 2\Delta, 3\Delta, \dots$. For the ease of exposition it will be assumed that the time unite Δ (called a "time-slot") is one. Since the packets of the input flows arrive at discrete epochs, several arrivals may happen during the same time slot. The lengths of packets D_j are integers and, of course, assumption (4) does not hold any longer.

2.3.1 Markovian description

Construction of the flow F_{jm_0} is similar to the case with continuous time. We start from a positive integer-valued random variable $\zeta^{\circ,j}$ distributed as intervals in (1) and in the same way define $\zeta^{(j)}$. Now the set $\left\{\zeta_t^{(j)},\, t=0,1,2,\dots\right\}$ of independent copies of $\zeta^{(j)}$ is countable. We introduce a discrete time Markov chain $r_j(t) = (r_j(t,\omega), t \in \mathbf{Z}_+)$ with state space $\mathbf{Z}_+$ as follows:

- put $r_j(0,\omega) \equiv x_j$ for some $x_j \in \mathbf{Z}_+$,
- if $r_j(t,\omega) > 0$ then $r_j(t+1,\omega) = r_j(t,\omega) - 1$,
- if $r_j(t,\omega) = 0$ then $r_j(t+1,\omega) = r_j(t,\omega) + \zeta_t^{(j)}(\omega)$.

This gives a Markovian description of a single flow F_{jm_0}.
Obviously, the process

$$\mathbf{r}(t) := (r_1(t), \dots, r_N(t)) \tag{6}$$

with independent components corresponds to the set of subflows $F_{1m_0}, \dots, F_{Nm_0}$. As before arrivals in the composite flow to the output port m_0 correspond to the event $\{r_j(t) = 0\}$. The only difference is that at fixed time t several arrivals from different sources may happen. It is clear, however, that the number of arrivals at given moment cannot exceed N.

Let us define a random sequence $\{J(t),\ t = 0, 1, \dots\}$:

$J(0,\omega) \equiv 0$,

$$J(t+1,\omega) = \max\Big(0, J(t,\omega) + \sum_{j=1}^{N} D_j I_{\{r_j(t)=0\}}(\omega) - 1\Big),$$

where I_A is the indicator function

$$I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases} .$$

Again the multi-dimensional process $(r_1(t), \dots, r_N(t), J(t))$ is a discrete time Markov chain with state space $\mathbf{Z}_+^{N+1}$. This Markov chain describes the complete system. As before, $J(t)$ is the total workload of the system at time t.

2.3.2 The limiting distribution of $\mathbf{r}(t)$

Since the multi-dimensional process (6) has independent components, its analysis is reduced to the consideration of one-dimensional processes $r_j(t)$. Let us fix some j. Denote by $q_1^{(j)}, q_2^{(j)}, \ldots$ a probability distribution of the integer-valued random variable $\zeta^{(j)}$:

$$\mathsf{P}\left(\zeta^{(j)} = k\right) = \mathsf{P}\left(\sigma_{l+1}^{(j,m_0)} - \sigma_l^{(j,m_0)} = k\right) = q_k^{(j)}.$$

$(r_j(t),\ t = 0, 1, \ldots)$ is a discrete time Markov chain on the state space $\mathbf{Z}_+$ with transition probabilities

$$p(i, i-1) = 1 \quad i \geq 1, \qquad p(0, k) = q_k^{(j)} \quad k \geq 1. \tag{7}$$

Our aim is to show that under very mild condition the Markov chain $(r_j(t),\ t = 0, 1, \ldots)$ is *ergodic*. Following [8, Ch. 3, § 3.1] we say that a Markov chain $(X(t),\ t = 0, 1, \ldots)$ is ergodic (or *positive recurrent*) if the next two conditions hold

1) n-step transition probabilities have limits which do not dependent on initial states:

$$p^{(n)}(i, k) \longrightarrow p_k \qquad (n \to \infty),$$

2) $\sum_k p_k = 1.$

The probability distribution $p_1, p_2, \ldots$ is called a *stationary distribution* of the Markov chain.

Theorem 1. *Assume that*

$$\mathsf{E}\zeta^{(j)} = \sum_k k q_k^{(j)} < \infty$$

and

$$\mathrm{GCD}\left\{i :\ q_i^{(j)} > 0\right\} = 1. \tag{8}$$

Then

i) the Markov chain $(r_j(t),\ t = 0, 1, \ldots)$ admits a limiting distribution $\pi^{(j)} = \left(\pi_k^{(j)},\ k \geq 0\right)$ and

$$\lim_{t \to +\infty} \mathsf{P}\left(r_j(t) = k\right) = \pi_k^{(j)}, \quad k \geq 0,$$

ii) for any bounded function f

$$\lim_{t \to +\infty} \frac{1}{t} \sum_{k=1}^{n} f(r_j(t)) = \sum_{i \geq 0} \pi_i^{(j)} f(i) \qquad (a.s.),$$

iii) $\pi^{(j)}$ *is given by*

$$\begin{aligned} \pi_0^{(j)} &= \frac{1}{\mathsf{E}\zeta^{(j)}}, \\ \pi_k^{(j)} &= \bar{Q}_k^{(j)}\pi_0^{(j)} \quad k \geq 1, \qquad \bar{Q}_k^{(j)} = \sum_{i \geq k} q_k^{(j)}. \end{aligned} \tag{9}$$

From *i)* it turns out that

$$\mathsf{P}\left\{r_1(t) = k_1, \ldots, r_N(t) = k_N\right\} \quad \stackrel{t\to\infty}{\longrightarrow} \quad \prod_{j=1}^{N} \pi_{k_j}^{(j)}.$$

The proof relies upon the following Foster-Liapounov criteria (see [9, p.116, Corollary 3], [10, 11]). Let $(X(t),\ t = 0, 1, \ldots)$ be irreducible aperiodic Markov chain on countable state space $\mathcal{X}$. Assume that there exists a non-negative function $f(\alpha)$, $\alpha \in \mathcal{X}$, a finite set A and a positive number $\epsilon > 0$ such that

$$\begin{aligned} \mathsf{E}\left(f\left(X(t+1)\right) \mid X(t) = \alpha\right) &\leq f(\alpha) - \epsilon \qquad \alpha \notin A, \\ \mathsf{E}\left(f\left(X(t+1)\right) \mid X(t) = \alpha\right) &< \infty, \qquad \alpha \in A. \end{aligned}$$

Then the Markov chain $(X(t),\ t = 0, 1, \ldots)$ is ergodic.

Proof. From (7) it follows that $(r_j(t),\ t = 0, 1, \ldots)$ is irreducible. Assumption (8) shows that the state 0 is aperiodic, therefore the Markov chain is irreducible and aperiodic. The Foster-Liapounov criteria with $f(i) = i$, $\epsilon = 1$ and $A = \{0\}$ entails that $(r_j(t),\ t \geq 0)$ is positive recurrent, hence *i)* and *ii)* hold true. A direct calculation leads to *iii)* since $\pi^{(j)}$ is also the unique stationary distribution.

□

Remark 1. It is readily seen that

$$\mathsf{E}\zeta^{(j)} = \mathsf{E}\zeta^{\circ,j}/R_{jm_0}. \tag{10}$$

2.3.3 The limiting distribution of the workload $J(t)$

In this section we are interested in the asymptotic behavior of $(J(t),\ t \geq 0)$. From subsection 2.2.2 it turns out that

$$J(t+1) = \max\left(0, J(t) + \xi(t)\right) \tag{11}$$

with

$$\xi(t) := \sum_{j=1}^{N} D_j I_{\{r_j(t)=0\}} - 1. \tag{12}$$

It is assumed here that the Markov chain

$$\mathbf{r}(t) := (r_1(t), \ldots, r_N(t))$$

is strictly stationary, therefore $(\xi(t),\, t \geq 0)$ is a strictly stationary process too. For the sake of convenience, the process $(\xi(t),\, t \geq 0)$ is extended to negative t and from now on we consider $(\xi(t),\, t \in \mathbf{Z})$.

Theorem 2. *If* $\mathsf{E}\,(\xi(t)) < 0$ *then*

$$\lim_{t \to +\infty} \mathsf{P}\,\{J(t) = j\} = \mathsf{P}\,\{Y = j\} = \nu_j,$$

where $\nu_j \geq 0, \forall j \geq 0,$ $\sum_j \nu_j = 1$ *and*

$$Y = \sup\left(0, \xi(-1), \ldots, \xi(-1) + \xi(-2) + \cdots + \xi(-n), \ldots\right).$$

Remark 2. Note that assumption $\mathsf{E}\,(\xi(t)) < 0$ can be rewritten as

$$\sum_{j=1}^{N} \pi_0^{(j)} D_j < 1. \tag{13}$$

Proof. The proof is based on the classical result [12, Ch. 1, § 3, Theorems 2]:

Let $S(t) := \sum_{m=1}^{t} \xi(m)$. *If for any fixed bounded interval* Δ

$$\lim_{t \to \infty} \mathsf{P}\,(S(t) \in \Delta) = 0 \tag{14}$$

then

$$\lim_{t \to \infty} \mathsf{P}\,(J(t) = j) = \mathsf{P}(Y = j) \tag{15}$$

for any initial state $J(0)$ *and any* $j \geq 0$.

From *ii)* of Theorem 1 it is seen that

$$\frac{1}{t} S(t) \quad = \quad \frac{1}{t} \sum_{m=1}^{t} \xi(m)$$

$$
\begin{aligned}
&= \frac{1}{t}\sum_{m=1}^{t}\left(\sum_{j=1}^{N} D_j I_{\{r_j(m)=0\}} - 1\right) \\
&= \sum_{j=1}^{N} D_j \left(\frac{1}{t}\sum_{m=1}^{t} I_{\{r_j(m)=0\}}\right) - 1 \\
&\xrightarrow{a.s.} \sum_{j=1}^{N} D_j \pi_0^{(j)} - 1 \qquad (t \to \infty).
\end{aligned}
$$

Hence $\dfrac{S(t)}{t} \xrightarrow{a.s.} \mathsf{E}\xi(1) = a < 0$ in view of (13) and finaly (14) holds true.

The Markov chain $\mathbf{r}(t)$ is positive recurrent, hence (see [13, Ch. 15, § 2]) $\mathbf{r}(t)$ is an *ergodic stationary sequence*[3]. Now condition $\mathsf{E}\xi(1) = a < 0$ together with Theorem 7 in [12, Ch. 1, § 3] entails $Y < +\infty$ a.s., therefore $\sum_j \nu_j = 1$.

□

Remark 3. The assumption of stationarity of $\mathbf{r}(t)$ in Theorem 2 is not necessary and may be omitted (see [12, Ch. 1, § 3, Theorem 6]).

2.3.4 Main examples

We describe briefly several subclasses of discrete time models interesting from practical point of view.

Symmetrical model. The model will be called symmetrical if

$$G_j(t) \equiv G(t), \qquad R_{jm} \equiv \frac{1}{N}, \qquad D_j \equiv D.$$

This means that input flows are of identical nature. It follows from (9) and (10) that

$$\pi_0^{(j)} = \frac{1}{N\,\mathsf{E}\zeta^{\circ,j}} = \frac{1}{N\,\mathsf{E}\zeta^{\circ,1}} = \pi_0^{(1)} \qquad \forall j.$$

[3] A stationary stochastic sequence $(Z(t) = Z(t,\omega),\ t \in \mathbf{Z}_+)$ is called *ergodic* if corresponding shift transformation T, defined by $Z(t+1,\omega) = Z(t, T\omega)$, is *metrically transitive*, i.e. any invariant set of T has probability 0 or 1 (see [12, 14, 11]).

Nonsymmetrical geometrical model. The key assumption here is that the intervals between arriving packets in the flows are geometrically distributed:

$$\mathsf{P}\left\{\zeta^{\circ,j} = k\right\} = p_j(1-p_j)^{k-1},\ k = 1, 2, \dots ,$$

No particular assumptions are made about D_j and the routing matrix $(R_{jm})_{j,m=1}^{N}$.

The main advantage of this model is that the process $(J(t),\ t = 0, 1, \dots)$ is now markovian. Indeed, if the $G_j(t)$ are geometrical distribution functions then the $\xi(t)$, $t = 0, 1, \dots$, defined in (12) are independent identically distributed random variables. Due to this fact we are able to study the model under consideration more deeply.

By (9) and (10) we get here $\pi_0^{(j)} = p_j R_{jm_0}$.

Geometrical model with symmetric routing. This is subcase of non-symmetrical geometric model with the following set of assumptions:

- $R_{jm} = \dfrac{1}{N}$ for all j and m,
- p_j and D_j are arbitrary.

For this model $\pi_0^{(j)} = \dfrac{p_j}{N}$.

Symmetrical geometrical model. This is a subcase of the symmetrical model with additional assumption that the distribution function $G(t)$ corresponds to the geometrical law:

$$\mathsf{P}\left\{\zeta^{\circ,j} = k\right\} = p(1-p)^{k-1},\ k \in \mathbf{N}.$$

Equivalently, we can say that this model is a totally symmetric variant of the geometrical model: $p_j \equiv p$, $R_{jm} \equiv \dfrac{1}{N}$, $D_j \equiv D$. As it was mentioned above the one-dimensional process $J(t)$ appears to be Markovian.

In this case, following (9) and (10), we have

$$\pi_0^{(j)} = \frac{1}{N\,\mathsf{E}\zeta^{\circ,j}} = \frac{p}{N} = \pi_0^{(1)} \qquad \forall j.$$

3 Performance evaluation via large-deviation technique

We consider here the discrete time geometrical models. As it was explained in Subsect. 2.3.4 for all geometrical models the stochastic process $J(t)$ is Markovian. From Subsect. 2.3.3 we recall that

$$J(t+1) = \max\left(0, J(t) + \xi(t)\right) \tag{16}$$

where $\xi(t)$, $t \in \mathbf{Z}_+$, are given by (12) and in the case of geometrical model are independent and identically distributed random variables.

Note that $J(t)$ can be nicely interpreted in the framework of queueing theory: $J(t)$ is a *virtual waiting time*. For the purpose of performance analysis it is important to estimate tail probabilities of the steady state:

$$\bar{\pi}_x := \pi\left(J \geq x\right) = \sum_{n \geq x} \pi\left\{J = n\right\}. \tag{17}$$

In the next subsection we recall some general results which will be very useful for our subsequent analysis of asymptotics of waiting time distribution.

3.1 General result on the asymptotics of stationary distribution

Denote

$$\xi := \sum_{j=1}^{N} D_j \eta_j - 1, \tag{18}$$

where $\eta_1, \ldots \eta_N$ are independent random variables having Bernoulli distribution:

$$\eta_j = \begin{cases} 1 & \text{with probability } \pi_0^{(j)} \\ 0 & \text{with probability } 1 - \pi_0^{(j)}. \end{cases}$$

So the random variable ξ has the same distribution as a distribution of any $\xi(t)$ in equation (16).

Let us introduce notation

$$\rho(\mu) := \mathsf{E} e^{\mu\xi}, \qquad \mu_+ = \sup\left\{\mu \geq 0 : \ \rho(\mu) < \infty\right\}.$$

We need the following result which can be found in [15]:
Assume that $\mu_+ > 0$, $\rho(\mu_+) > 1$ *and there exist* $\beta > 0$ *such that*

$$\rho(\beta) = 1, \qquad \rho(\mu) < 1 \quad \forall \mu \in (0, \beta),$$

$$\rho'(\beta) = \mathsf{E}\xi e^{\beta\xi} < \infty.$$

Then

$$\bar{\pi}_x = e^{-\beta x}\left(c_1 + o(e^{-\varepsilon x})\right), \qquad x \to \infty, \tag{19}$$

where $c_1 > 0$ depends on the distribution of ξ (see [12, Ch. 4, § 22, Theorem 11]) and is known in explicit form, ε is some positive number.

If ξ is defined by (18) then the constant c_1 in (19) has the following form:

$$c_1 = -\frac{a}{\rho'(\beta)} \tag{20}$$

where $a = \mathsf{E}\xi$.

3.2 Evaluation of the asymptotics

It follows from (18) that

$$a = \sum_{j=1}^{N} \pi_0^{(j)} D_j - 1. \tag{21}$$

In the case of *general (non-symmetrical) geometrical model* the function ρ takes the following form:

$$\begin{aligned}
\rho(\mu) = \mathsf{E}e^{\mu\xi} &= \mathsf{E}\exp\left(\mu\Big(\sum_{j=1}^{N} D_j\eta_j - 1\Big)\right) \\
&= e^{-\mu}\prod_{j=1}^{N}\mathsf{E}\Big\{\exp(\mu D_j\eta_j)\Big\} \\
&= e^{-\mu}\prod_{j=1}^{N}\left(\pi_0^{(j)}\exp(\mu D_j) + 1 - \pi_0^{(j)}\right) \\
&= e^{-\mu}\prod_{j=1}^{N}\left(1 + \pi_0^{(j)}\left(\exp\left(\mu D_j\right) - 1\right)\right).
\end{aligned}$$

Note that $\mu_+ = +\infty$. In a similar way we get ρ' :

$$\begin{aligned}
&\rho'(\mu) = \mathsf{E}\left\{\xi e^{\mu\xi}\right\} = \\
&= \mathsf{E}\left\{\left(\sum_{j=1}^{N} D_j\eta_j - 1\right)\exp\left(\mu\Big(\sum_{j=1}^{N} D_j\eta_j - 1\Big)\right)\right\} \\
&= e^{-\mu}\sum_{k=1}^{N}\pi_0^{(k)} D_k\, e^{\mu D_k}\prod_{j\neq k}\left(1 + \pi_0^{(j)}\left(\exp\left(\mu D_j\right) - 1\right)\right) - \rho(\mu)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{N} \frac{\pi_0^{(k)} D_k \exp(\mu D_k)}{1+\pi_0^{(k)}(\exp(\mu D_k)-1)} \rho(\mu) - \rho(\mu) \\
&= \rho(\mu) \left(\sum_{k=1}^{N} \frac{\pi_0^{(k)} D_k \exp(\mu D_k)}{1+\pi_0^{(k)}(\exp(\mu D_k)-1)} - 1 \right).
\end{aligned}
$$

Note that $\rho'(\mu) = \mathsf{E}\xi e^{\beta\mu} < \infty$ for all $\mu > 0$.

In the case of *geometric model with symmetric routing* (Sect. 2.3.4) the above formulae take the form

$$
\begin{aligned}
\rho(\mu) &= e^{-\mu} \prod_{j=1}^{N} \left(1 + \frac{p_j(\exp(\mu D_j)-1)}{N} \right), \qquad (22) \\
\rho'(\mu) &= \rho(\mu) \left(\sum_{k=1}^{N} \frac{p_k N^{-1} D_k \exp(\mu D_k)}{1+p_k N^{-1}(\exp(\mu D_k)-1)} - 1 \right)
\end{aligned}
$$

In the case of *symmetrical geometrical model* we get

$$
\rho(\mu) = e^{-\mu} \left(1 + \frac{\exp(\mu D)-1}{p^{-1}N} \right)^N \qquad (23)
$$

and, similarly,

$$
\rho'(\mu) = \rho(\mu) \left(\frac{D\exp(\mu D)}{p^{-1} + \frac{1}{N}(\exp(\mu D)-1)} - 1 \right). \qquad (24)
$$

Let us consider now the nonlinear equation $\rho(\mu) = 1$. Evidently, $\mu = 0$ always solves it, but we are interested in a positive solution. Given all parameters it is not hard to solve this equation numerically for any of models (22) and (23) (provided the positive solution β exists).

But we are going to continue below with some theoretical estimation. We shall consider in the sequel the *symmetrical geometrical model*. Denote by β_N the positive solution of the equation

$$
e^{-\mu} \left(1 + \frac{\exp(\mu D)-1}{p^{-1}N} \right)^N = 1.
$$

This equation can be rewritten as

$$
-\mu + N \log \left(1 + \frac{\exp(\mu D)-1}{p^{-1}N} \right) = 0. \qquad (25)
$$

Of course, to have existence of β_N we need additional assumptions about the parameters.

4 Detailed study of symmetrical geometrical model

4.1 Approximation for large values of N

In this particular case, in view of (12), the *load* of the queueing system is given by :

$$\mathsf{E}\left\{\sum_{j=1}^{N} D_j I_{\{r_j=0\}}\right\} = \sum_{j=1}^{N} D_j \pi_0^{(j)} = \sum_{j=1}^{N} D_j \frac{p}{N} = Dp.$$

We denote $\lambda := Dp$ and assume that

$$\lambda < 1. \tag{26}$$

To proceed further we need a lemma.

Lemma 1. *If $\lambda < 1$ then the nonlinear equation*

$$\lambda y = \log(1+y), \qquad y \geq 0, \tag{27}$$

has exactly two solutions: $y_1 = 0$ and $y_2 = y_2(\lambda) > 0$.

Proof. See Figure 6 for a proof.

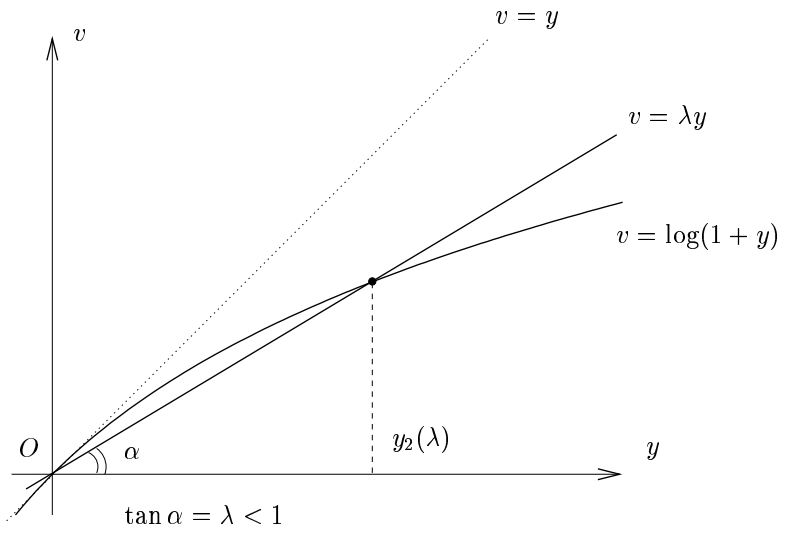


Figure 6: Proof of Lemma 1

□

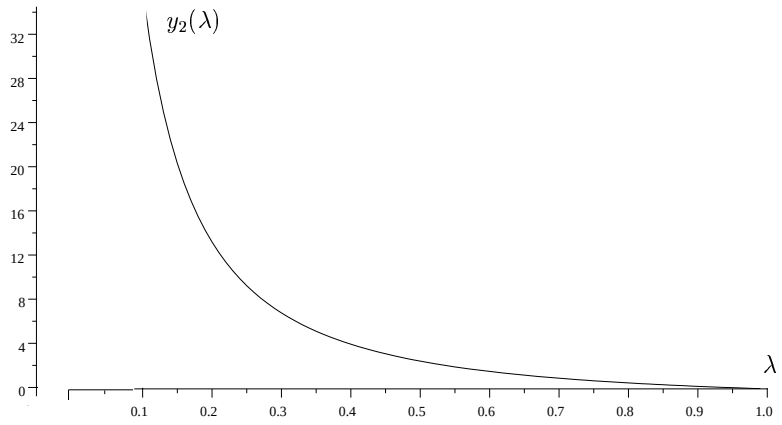


Figure 7: Graph of $y_2 = y_2(\lambda)$ obtained by a numerical calculation.

Figure 7 shows the graph of $y_2(\lambda)$ obtained by numerical calculation.

Consider $g(w) := \log(1+w) - w$ and fix w_0, $0 < w_0 < 1$, and constants $K_1 > \frac{1}{2} > K_2 > 0$ such that the following inequality

$$-K_1 w^2 < g(w) < -K_2 w^2$$

holds for $0 < w < w_0$. (In fact, K_1 and K_2 can be chosen as close to $\frac{1}{2}$ provided w_0 is sufficiently small.)

We return to equation (25). Let $M = 2y_2$ and restrict ourself to the domain $\mu \in (0, M]$. For sufficiently large N, namely,

$$N > N_0 = \frac{\lambda\left(e^{MD} - 1\right)}{Dw_0}$$

we rewrite equation (25) as follows

$$-\mu + \frac{\exp(\mu D) - 1}{p^{-1}} + Ng\left(\frac{\exp(\mu D) - 1}{p^{-1}N}\right) = 0.$$

It is convenient to take a new variable $u := \exp(\mu D) - 1$ and to consider equation

$$-\frac{1}{D}\log(1+u) + \frac{u}{p^{-1}} + Ng\left(\frac{\lambda u}{DN}\right) = 0$$

which can be evidently rewritten as

$$-\log(1+u)+\lambda u+DNg\left(\frac{\lambda u}{DN}\right)=0. \tag{28}$$

Note that for $\mu\in(0,M]$ and $N>N_0$

$$-\frac{K_1'u^2}{N}<DNg\left(\frac{\lambda u}{DN}\right)<-\frac{K_2'u^2}{N} \tag{29}$$

with $K_i'=K_i\lambda^2/D$. So the equation (28) can be considered as a *small perturbation* of the equation (27) (see Figure 8).

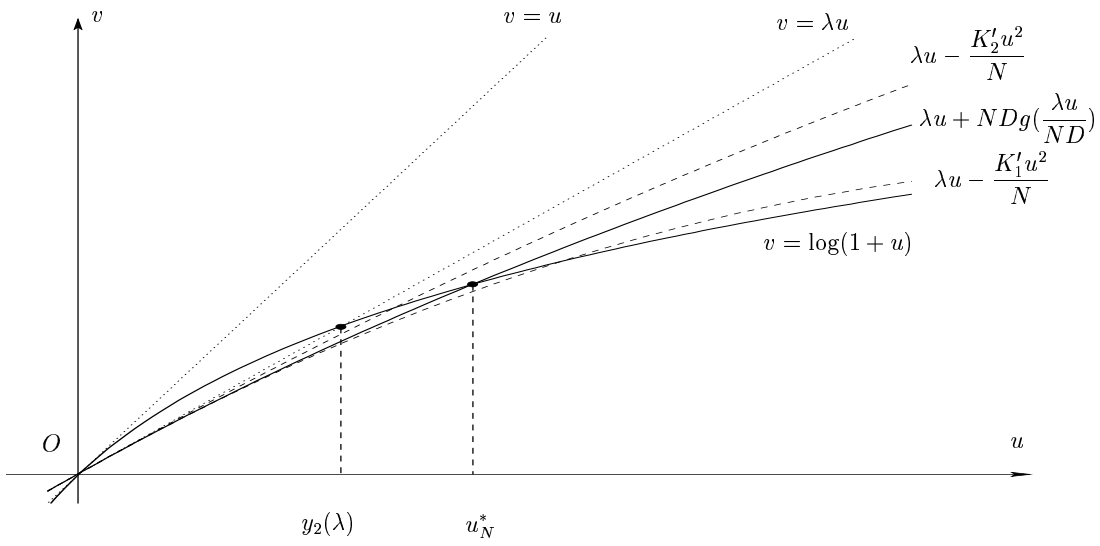


Figure 8: Perturbation of the canonical problem

Lemma 2. *Let assumption (26) holds. Then for sufficiently large $N>N_1=N_1(\lambda,D)$*

- *there exists a positive solution u_N^* of the equation (28) and there is no other solution to (28) in the interval $u\in(0,u_N^*)$;*
- *the solution u_N^* depends only on λ and ND and satisfies the following inequality*

$$y_2<u_N^*<y_2+\frac{k}{N}$$

where $y_2=y_2(\lambda)$ is the positive solution of (27) and the constant $k=k(N_1)>0$;

- *for fixed λ and D the sequence u_N^* decreases in N:*

$$u_{N+1}^* \le u_N^* \qquad \forall N \ge N_1.$$

Proof. The first two statements of the Lemma follow from perturbation arguments (28)–(29). To prove third statement let us analyze the equation (28) more carefully. It can be rewritten as

$$-\frac{1}{\lambda}\log(1+u) + u \cdot \frac{DN}{\lambda u}\log\left(1+\frac{\lambda u}{DN}\right) = 0.$$

Let us consider a function $h(x) := x\log\left(1+x^{-1}\right)$, $x>0$. It is easy to check that

$$\begin{aligned} h'(x) &= \log(1+x^{-1}) - \frac{1}{1+x}, \\ h''(x) &= -\frac{1}{x(1+x)} + \frac{1}{(1+x)^2}. \end{aligned}$$

Note that $h''(x) < 0$ if $0 < x < +\infty$ and $h'(x) \to 0$ when $x \to +\infty$. Hence $h'(x) > 0$ for $0 < x < +\infty$ and the function $h(x)$ is increasing. So if λ, D and u are fixed then a function given by the expression

$$\frac{DN}{\lambda u}\log\left(1+\frac{\lambda u}{DN}\right)$$

is increasing in N. Now the third statement of the Lemma easily follows.

□

Remark 4. Recall that $u = \exp(\mu D) - 1$. So β_N which solves (25) can be represented as $\beta_N = D^{-1}\log(1+u_N^*)$.

4.2 Some conclusions

Applying the result (19), we are able now to give sharp bounds for logarithmic asymptotics of the tail probabilities (17) for the symmetrical geometrical model (Sect. 2.3.4). Assume that the parameters p and D are given and the stability condition (26) holds: $\lambda \equiv Dp < 1$. We add the subscript N to the β, c_1 and ε in (19) to underline that the model depends on N

$$\pi(J \ge x) = e^{-\beta_N x}\left(c_{1,N} + o(e^{-\varepsilon_N x})\right), \qquad (x \to \infty). \tag{30}$$

The coefficient β_N is the most important in this asymptotics, so let us explain how it depends on N. The next theorem follows from Lemma 2 and Remark 4.

Theorem 3. *If $\lambda < 1$, then for sufficiently large $N > N_1 = N_1(\lambda, D)$*

i) there exists a positive solution β_N of (25) and there is no other solution of (25) in the interval $(0, \beta_N)$;

ii) β_N depends only on λ and ND and satisfies the following inequality

$$\frac{1}{D}\log(1+y_2) < \beta_N < \frac{1}{D}\log\left(1+\left(y_2+\frac{k}{N}\right)\right)$$

where $y_2 = y_2(\lambda)$ is the positive solution of (27) and $k = k(N_1) > 0$ is a constant;

iii) for λ and D fixed, the sequence β_N is decreasing:

$$\beta_{N+1} \leq \beta_N \qquad \forall N \geq N_1.$$

Substituting this into (30) we get

$$\left(1+(y_2+\frac{k}{N})\right)^{-x/D} < \frac{\pi(J \geq x)}{c_{1,N} + o(e^{-\varepsilon_N x})} < (1+y_2)^{-x/D}.$$

Let us study now the dependence of the constant $c_{1,N}$ on N.

Corollary. $$c_{1,N} = \frac{1-\lambda}{\lambda(1+y_2(\lambda))-1} + O\left(\frac{1}{N}\right).$$

Proof. Taking into account (20), (21) and (24) we obtain

$$\begin{aligned} c_{1,N} &= \frac{1-Dp}{\rho'(\beta_N)} = \frac{1-Dp}{\dfrac{D\exp(\beta_N D)}{p^{-1}+N^{-1}(\exp(\beta_N D)-1)} - 1} \\ &= \frac{1-\lambda}{\dfrac{D(1+u_N^*)}{p^{-1}+N^{-1}u_N^*} - 1} \\ &= \frac{1-\lambda}{\lambda(1+y_2(\lambda))-1} + O\left(\frac{1}{N}\right). \end{aligned}$$

since $\rho(\beta_N) = 1$. □

The behavior of the function

$$c_{1,\infty}(\lambda) := \frac{1-\lambda}{\lambda(1+y_2(\lambda))-1}$$

is presented on Figure 9.

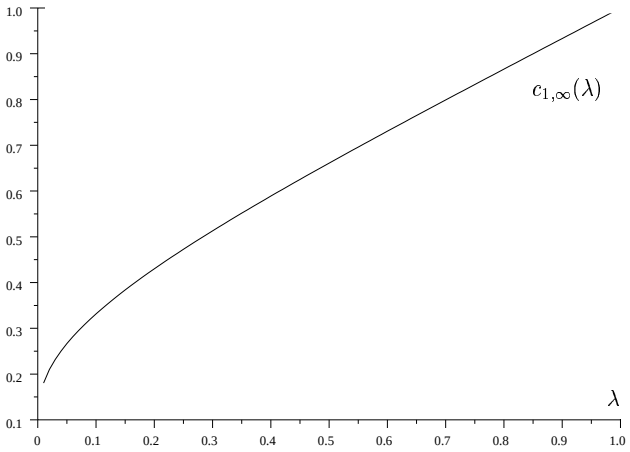


Figure 9: Graph of the function $c_{1,\infty}(\lambda)$ obtained by numerical calculations

4.3 Numerical experiments, comparison

In this section consider the case of symmetrical geometric model with the following specific choice of parameters:

$$D = 1, \quad p < 1.$$

This choice of parameters corresponds to the so-called Binomial model in terminology of [3, 6]. We propose a comparison between values of stationary probabilities $\pi_n = \pi(J = n)$ obtained numerically by two different methods.

The first method is based on the exact recurrent formulae (see [3, 6]), the corresponding numerical values of π_n will be denoted in this section by p_n. We recall exact formulae from [3]:

$$\begin{aligned} a_0 p_0 &= 1 - p, \qquad a_0 p_1 = (1 - a_0 - a_1)\, p_0, \\ a_0 p_n &= (1 - a_1)\, p_{n-1} - \sum_{i=2}^{n} a_i p_{n-i} \qquad \text{for} \quad n \geq 2, \end{aligned}$$

where a_k are defined as $\mathsf{P}\{A = k\} = a_k$ with $A = \sum_{j=1}^{N} D_j \eta_j$,

$$D_j = D = 1, \qquad \mathsf{P}\{\eta_j = 1\} = 1 - \mathsf{P}\{\eta_j = 0\} = \frac{p}{N}\,.$$

To get the values of π_n Matlab[4] was used. The values of p_n were computed until they become of order 10^{-10}.

The second method is based on the asymptotic approach presented in this paper: we compute numerically two numbers β_N and $c_{1,N}$ and take

$$\tilde{\pi}_n = c_{1,N} e^{-\beta_N n} - c_{1,N} e^{-\beta_N (n+1)}$$

as an approximation for $\pi_n = \bar{\pi}_n - \bar{\pi}_{n+1}$ (see (17) and (30)). To compute $\tilde{\pi}_n$ the package Scilab[5] was used. Below we give the results of the numerical tests for different values of N and p.

$N = 24$, $p = 0.1$. In this case we find that $\beta_N \approx$3.71910437 and $c_{1,N} \approx$0.355602842 .

n	p_n	$\tilde{\pi}_n$	$\lvert p_n - \tilde{\pi}_n \rvert$	$\lvert p_n - \tilde{\pi}_n \rvert / p_n$
1	4.95781×10^{-3}	.00841617483	0.00345836483	0.697558969
2	1.75498×10^{-4}	.000204140062	2.86420623×10^{-5}	0.163204494
3	4.92095×10^{-6}	4.95155648×10^{-6}	3.06064838×10^{-8}	0.00621962909
4	1.22495×10^{-7}	1.20103381×10^{-7}	2.39161935×10^{-9}	0.0195242202
5	2.93766×10^{-9}	2.91318944×10^{-9}	2.44705634×10^{-11}	0.00832995085
6	7.06159×10^{-11}	7.06613973×10^{-11}	4.54972697×10^{-14}	0.000644292146

$N = 24$, $p = 0.3$. In this case we find that $\beta_N \approx$2.13203465 and $c_{1,N} \approx$0.532490569 .

n	p_n	$\tilde{\pi}_n$	$\lvert p_n - \tilde{\pi}_n \rvert$	$\lvert p_n - \tilde{\pi}_n \rvert / p_n$
1	4.60243×10^{-2}	.0556616631	0.00963736308	0.209397277
2	6.39571×10^{-3}	.00660123653	0.000205526533	0.0321350612
3	7.84974×10^{-4}	.00078287858	2.09541987×10^{-6}	0.00266941309
4	9.30910×10^{-5}	9.28460703×10^{-5}	2.44929692×10^{-7}	0.00263107811
5	1.10142×10^{-5}	1.10111491×10^{-5}	3.05090174×10^{-9}	0.000276997125
6	1.30567×10^{-6}	1.30587546×10^{-6}	2.05456679×10^{-10}	0.000157357279
7	1.54864×10^{-7}	1.54871276×10^{-7}	7.27575314×10^{-12}	4.69815654×10^{-5}
8	1.83672×10^{-8}	1.83670747×10^{-8}	1.25306363×10^{-13}	6.82228992×10^{-6}
9	2.17827×10^{-9}	2.17825695×10^{-9}	1.30497612×10^{-14}	5.99088321×10^{-6}
10	2.58331×10^{-10}	2.58332011×10^{-10}	1.01097111×10^{-15}	3.91347191×10^{-6}

[4] http://www.mathworks.com/
[5] http://www-rocq.inria.fr/scilab/

$N = 24$, $p = 0.5$. In this case we find that β_N ≈1.30154338 and $c_{1,N}$ ≈0.675230878 .

n	p_n	$\tilde{\pi}_n$	$\lvert p_n - \tilde{\pi}_n \rvert$	$\lvert p_n - \tilde{\pi}_n \rvert / p_n$
1	1.21669×10^{-1}	.13374084	0.0120718398	0.099218698
2	3.59877×10^{-2}	.0363924202	0.000404720163	0.0112460692
3	9.91367×10^{-3}	.00990279594	1.08740578×10^{-5}	0.0010968751
4	2.69618×10^{-3}	.00269466463	1.51536834×10^{-6}	0.000562042721
5	7.33264×10^{-4}	.000733249228	1.47722596×10^{-8}	2.01458951×10^{-5}
6	1.99521×10^{-4}	.000199525545	4.54528141×10^{-9}	2.27809675×10^{-5}
7	5.42930×10^{-5}	5.42931949×10^{-5}	1.94883446×10^{-10}	3.58947648×10^{-6}
8	1.47738×10^{-5}	1.47738026×10^{-5}	2.55492762×10^{-12}	1.72936389×10^{-7}
9	4.02012×10^{-6}	4.02012153×10^{-6}	1.53457806×10^{-12}	3.8172444×10^{-7}
10	1.09392×10^{-6}	1.09392129×10^{-6}	1.29024953×10^{-12}	1.17947339×10^{-6}
11	2.97669×10^{-7}	2.9766856×10^{-7}	4.3974171×10^{-13}	1.4772842×10^{-6}
12	8.09990×10^{-8}	8.09990376×10^{-8}	3.76419418×10^{-14}	4.64721068×10^{-7}
13	2.20408×10^{-8}	2.20407694×10^{-8}	3.05844198×10^{-14}	1.38762748×10^{-6}
14	5.99755×10^{-9}	5.99754677×10^{-9}	3.23187363×10^{-15}	5.38865641×10^{-7}
15	1.63200×10^{-9}	1.63200143×10^{-9}	1.43142173×10^{-15}	8.77096646×10^{-7}
16	4.44088×10^{-10}	4.44086353×10^{-10}	1.6468402×10^{-15}	3.70836456×10^{-6}

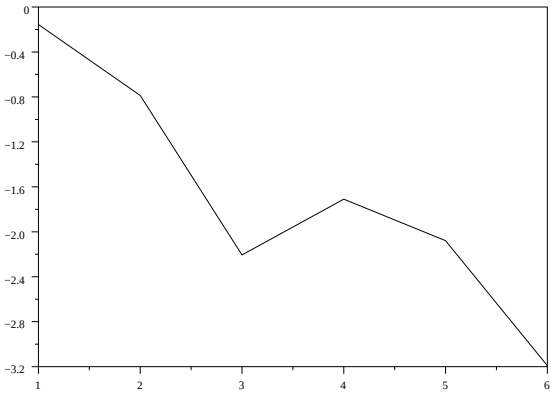
$N = 24$, $p = 0.9$. In this case we find that β_N ≈0.215840371 and $c_{1,N}$ ≈0.936198246 .

n	p_n	$\tilde{\pi}_n$	$\lvert p_n - \tilde{\pi}_n \rvert$	$\lvert p_n - \tilde{\pi}_n \rvert / p_n$
1	1.42015×10^{-1}	.146465646	0.00445064647	0.0313392703
2	1.17760×10^{-1}	.118031382	0.000271381601	0.00230453125
3	9.51324×10^{-2}	.0951172331	1.51668585×10^{-5}	0.000159428949
4	7.66551×10^{-2}	.0766515474	3.55260807×10^{-6}	4.63453582×10^{-5}
5	6.17707×10^{-2}	.0617707173	1.72877755×10^{-8}	2.79870157×10^{-7}
6	4.97788×10^{-2}	.049778793	6.95442116×10^{-9}	1.39706485×10^{-7}
7	4.01149×10^{-2}	.0401149338	3.38371849×10^{-8}	8.43506651×10^{-7}
8	3.23272×10^{-2}	.0323271783	2.17317706×10^{-8}	6.72244135×10^{-7}
9	2.60513×10^{-2}	.026051307	6.95285272×10^{-9}	2.66890816×10^{-7}
10	2.09938×10^{-2}	.0209938086	8.62507072×10^{-9}	4.10838948×10^{-7}
11	1.69182×10^{-2}	.0169181531	4.68576744×10^{-8}	2.76966075×10^{-6}
12	1.36337×10^{-2}	.0136337294	2.93941688×10^{-8}	2.15599351×10^{-6}
13	1.09869×10^{-2}	.0109869308	3.07618687×10^{-8}	2.7998679×10^{-6}

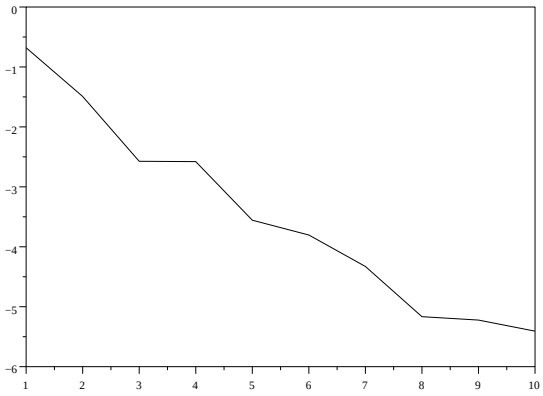
14	8.85397×10^{-3}	.00885397121	1.21184798×10^{-9}	1.36870577×10^{-7}
15	7.13510×10^{-3}	.00713509604	3.95549743×10^{-9}	5.54371687×10^{-7}
16	5.74992×10^{-3}	.00574991655	3.45451232×10^{-9}	6.00793111×10^{-7}
17	4.63365×10^{-3}	.00463365035	3.49464613×10^{-10}	7.54188628×10^{-8}
18	3.73409×10^{-3}	.00373409168	1.68485114×10^{-9}	4.51207963×10^{-7}
19	3.00917×10^{-3}	.0030091698	1.96653593×10^{-10}	6.535144×10^{-8}
20	2.42498×10^{-3}	.00242498141	1.40635045×10^{-9}	5.79943114×10^{-7}
21	1.95421×10^{-3}	.00195420505	4.95161561×10^{-9}	2.5338196×10^{-6}
22	1.57482×10^{-3}	.00157482336	3.36199783×10^{-9}	2.13484578×10^{-6}
23	1.26909×10^{-3}	.00126909334	3.34491×10^{-9}	2.63567595×10^{-6}
24	1.02272×10^{-3}	.00102271655	3.44844235×10^{-9}	3.37183428×10^{-6}
25	8.24170×10^{-4}	.00082417038	3.80394002×10^{-10}	4.61547984×10^{-7}
26	6.64169×10^{-4}	.000664169182	1.8244283×10^{-10}	2.74693383×10^{-7}
27	5.35230×10^{-4}	.000535229988	1.18804374×10^{-11}	2.21968824×10^{-8}
28	4.31323×10^{-4}	.000431322542	4.57859728×10^{-10}	1.06152403×10^{-6}
29	3.47587×10^{-4}	.000347587279	2.79277759×10^{-10}	8.03475848×10^{-7}
30	2.80108×10^{-4}	.000280108051	5.11957662×10^{-11}	1.82771524×10^{-7}
31	2.25729×10^{-4}	.000225728975	2.47977989×10^{-11}	1.09856505×10^{-7}
32	1.81907×10^{-4}	.000181906839	1.60734517×10^{-10}	8.836082×10^{-7}
33	1.46592×10^{-4}	.000146592161	1.60540831×10^{-10}	1.09515411×10^{-6}
34	1.18133×10^{-4}	.000118133335	3.34726719×10^{-10}	2.83347345×10^{-6}
35	9.51994×10^{-5}	9.51993935×10^{-5}	6.51150646×10^{-12}	6.83986082×10^{-8}
36	7.67178×10^{-5}	7.67177575×10^{-5}	4.25478588×10^{-11}	5.54602176×10^{-7}
37	6.18241×10^{-5}	6.18240736×10^{-5}	2.64128406×10^{-11}	4.27225638×10^{-7}
38	4.98218×10^{-5}	4.9821791×10^{-5}	9.03726346×10^{-12}	1.81391749×10^{-7}
39	4.01496×10^{-5}	4.01495843×10^{-5}	1.56922755×10^{-11}	3.90845127×10^{-7}
40	3.23551×10^{-5}	3.23551018×10^{-5}	1.83262453×10^{-12}	5.66409788×10^{-8}
41	2.60738×10^{-5}	2.60738095×10^{-5}	9.5462198×10^{-12}	3.66123074×10^{-7}
42	2.10119×10^{-5}	2.10119427×10^{-5}	4.2653416×10^{-11}	2.02996474×10^{-6}
43	1.69328×10^{-5}	1.69327667×10^{-5}	3.32983097×10^{-11}	1.96649755×10^{-6}
44	1.36455×10^{-5}	1.36455059×10^{-5}	5.93456228×10^{-12}	4.34909845×10^{-7}
45	1.09964×10^{-5}	1.09964211×10^{-5}	2.10510021×10^{-11}	1.91435398×10^{-6}
46	8.86162×10^{-6}	8.86161909×10^{-6}	9.06445942×10^{-13}	1.02288965×10^{-7}
47	7.14126×10^{-6}	7.1412592×10^{-6}	8.03483015×10^{-13}	1.1251278×10^{-7}
48	5.75488×10^{-6}	5.75488321×10^{-6}	3.2073965×10^{-12}	5.57335078×10^{-7}
49	4.63765×10^{-6}	4.6376528×10^{-6}	2.80315371×10^{-12}	6.04434081×10^{-7}
50	3.73732×10^{-6}	3.73731712×10^{-6}	2.88208587×10^{-12}	7.71163795×10^{-7}
51	3.01177×10^{-6}	3.01176906×10^{-6}	9.36486732×10^{-13}	3.10942314×10^{-7}
52	2.42708×10^{-6}	2.42707606×10^{-6}	3.94363415×10^{-12}	1.6248472×10^{-6}
53	1.95589×10^{-6}	1.95589305×10^{-6}	3.05127958×10^{-12}	1.56004662×10^{-6}

54	1.57618×10^{-6}	1.57618366×10^{-6}	3.66264625×10^{-12}	2.32374871×10^{-6}
55	1.27019×10^{-6}	1.27018956×10^{-6}	4.37862321×10^{-13}	3.44721908×10^{-7}
56	1.02360×10^{-6}	1.02359995×10^{-6}	4.65181734×10^{-14}	4.54456559×10^{-8}
57	8.24882×10^{-7}	8.24882282×10^{-7}	2.82141149×10^{-13}	3.42038193×10^{-7}
58	6.64743×10^{-7}	6.64742878×10^{-7}	1.21616828×10^{-13}	1.82953154×10^{-7}
59	5.35692×10^{-7}	5.35692309×10^{-7}	3.08985166×10^{-13}	5.76796304×10^{-7}
60	4.31695×10^{-7}	4.3169511×10^{-7}	1.09850345×10^{-13}	2.54462861×10^{-7}
61	3.47888×10^{-7}	3.47887518×10^{-7}	4.81842889×10^{-13}	1.38505177×10^{-6}
62	2.80350×10^{-7}	2.80350003×10^{-7}	2.87927696×10^{-15}	1.02702941×10^{-8}
63	2.25924×10^{-7}	2.25923955×10^{-7}	4.46630463×10^{-14}	1.97690579×10^{-7}
64	1.82064×10^{-7}	1.82063967×10^{-7}	3.32766857×10^{-14}	1.8277466×10^{-7}
65	1.46719×10^{-7}	1.46718784×10^{-7}	2.16108355×10^{-13}	1.47294049×10^{-6}
66	1.18235×10^{-7}	1.18235376×10^{-7}	3.75917944×10^{-13}	3.17941341×10^{-6}
67	9.52816×10^{-8}	9.52816248×10^{-8}	2.48039642×10^{-14}	2.60322709×10^{-7}
68	7.67840×10^{-8}	7.67840247×10^{-8}	2.47032662×10^{-14}	3.21724138×10^{-7}
69	6.18775×10^{-8}	6.1877476×10^{-8}	2.40253617×10^{-14}	3.88272987×10^{-7}
70	4.98648×10^{-8}	4.9864826×10^{-8}	2.60206276×10^{-14}	5.21823564×10^{-7}
71	4.01843×10^{-8}	4.01842647×10^{-8}	3.52913587×10^{-14}	8.78237487×10^{-7}
72	3.23830×10^{-8}	3.23830495×10^{-8}	4.95168312×10^{-14}	1.52909956×10^{-6}
73	2.60963×10^{-8}	2.60963316×10^{-8}	3.15768731×10^{-14}	1.21001341×10^{-6}
74	2.10301×10^{-8}	2.10300923×10^{-8}	7.65443076×10^{-15}	3.63975005×10^{-7}
75	1.69474×10^{-8}	1.69473929×10^{-8}	7.11604724×10^{-15}	4.19890204×10^{-7}
76	1.36573×10^{-8}	1.36572926×10^{-8}	7.35270623×10^{-15}	5.38371876×10^{-7}
77	1.10059×10^{-8}	1.10059195×10^{-8}	1.95376558×10^{-14}	1.77519838×10^{-6}
78	8.86927×10^{-9}	8.86927358×10^{-9}	3.58134606×10^{-15}	4.03792653×10^{-7}
79	7.14742×10^{-9}	7.14742767×10^{-9}	7.67213776×10^{-15}	1.07341359×10^{-6}
80	5.75985×10^{-9}	5.75985416×10^{-9}	4.15940768×10^{-15}	7.22138194×10^{-7}
81	4.64165×10^{-9}	4.64165871×10^{-9}	8.71408156×10^{-15}	1.87736722×10^{-6}
82	3.74054×10^{-9}	3.74054534×10^{-9}	5.33704111×10^{-15}	1.42681033×10^{-6}
83	3.01437×10^{-9}	3.01437057×10^{-9}	5.68868624×10^{-16}	1.88718911×10^{-7}
84	2.42917×10^{-9}	2.42917252×10^{-9}	2.51569766×10^{-15}	1.03562026×10^{-6}
85	1.95758×10^{-9}	1.95758251×10^{-9}	2.5122376×10^{-15}	1.28333841×10^{-6}
86	1.57754×10^{-9}	1.57754514×10^{-9}	5.13829492×10^{-15}	3.25715666×10^{-6}
87	1.27128×10^{-9}	1.27128673×10^{-9}	6.72625569×10^{-15}	5.29093173×10^{-6}
88	1.02448×10^{-9}	1.02448412×10^{-9}	4.11847076×10^{-15}	4.0200597×10^{-6}
89	8.25589×10^{-10}	8.25594799×10^{-10}	5.79881466×10^{-15}	7.02385165×10^{-6}
90	6.65312×10^{-10}	6.6531707×10^{-10}	5.06987049×10^{-15}	7.62029017×10^{-6}

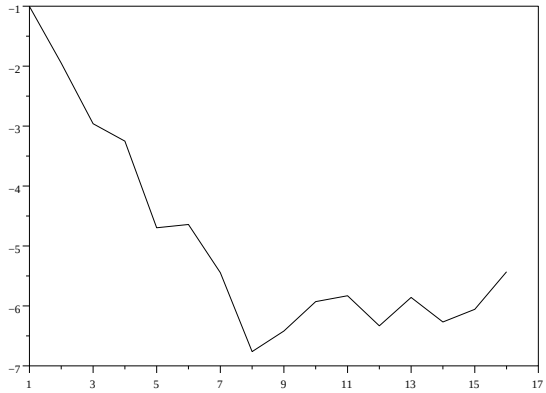
The behavior of the function $\log_{10}\left(\left|p_n - \tilde{\pi}_n\right| / p_n\right)$ for different values of parameters (N, p) is presented on the following figures.



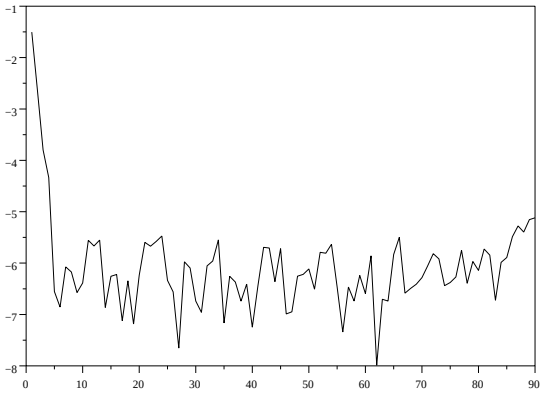
$N = 24$, $p = 0.1$



$N = 24$, $p = 0.3$



$N = 24$, $p = 0.5$



$N = 24$, $p = 0.9$

As readily seen from the above tables the approximations based on the asymptotic method give very good agreement with the values computed by using the exact recurrent formulae.

5 Conclusion

For complementing the measurement based performance tests and the approximate modelling in [6] for Ethernet switches, we presented in this paper an exact mathematical model of an Ethernet

switch based on a multi-dimensional Markov process. The main difficulty to model the flow entering to an output queue lies in the fact that the superposition of the flows from the input ports is no longer a renewal process even if the flows arriving to the input ports are renewal processes. The multi-dimensional Markov process provides a complete description of the flow entering to an output queue. Moreover this model has advantage to be general in the sense that the input flow is just to be a renewal process and whatever switching matrix may be (i.e., probability for an input packet to choose an output port). To overcome the subsidiary difficulty for accurately computing the tail of the workload distribution with classic methods (e.g. from generating function or the direct use of the Markov chain balance equation), an asymptotic approach is developed based on the recent results on large deviations for Markovian systems. A case study with Binomial input flow and symmetric routing matrix (i.e., each input packet chooses an output port with probability of $1/N$) has been performed and numerical results using both the asymptotic approach and the exact recurrent formulae [3] are compared showing the numerical efficiency of our method.

The on going extension of this work consists in developing the detailed study of non-symmetrical geometric model (which has less restrictive hypotheses than the symmetrical geometric model in section 4).

References

[1] Seifert, R., The Switch Book: The Complete Guide to LAN Switching Technology, John-Wiley, 2000.

[2] Jasperneite, J. and P. Neumann, Performance Evaluation of Switched Ethernet in real-time Applications, Proc. of 4th IFAC FeT2001, pp. 169–176, Nancy (France), Nov. 15–16, 2001.

[3] M.J. Karol, M.G. Hluchyj, S.P. Morgan, Input versus output queueing on a space division packet switch. IEEE Transactions on communications, Vol. Com-35, N. 12, 1987, P. 1347–1356.

[4] Lehoczky, J.P., Fixed priority scheduling of periodic task sets with arbitrary deadlines, Proc. of IEEE Real-time systems symposium, IEEE Computer Press, pp. 201–209, Los Alamitos, CA (USA), 1990.

[5] Cruz, R. L., A calculus for network delay, Part I, IEEE Trans. on Information Theory, 37(1): 114–131, Jan. 1991.

[6] Y. Song, A. Koubaa, F. Simonot, Switched Ethernet For Real-Time Industrial Communication: Modelling And Message Buffering Delay Evaluation, in 4th IEEE International Workshop on Factory Communication Systems (WFCS 2002), Vasteras (Sweden), pp. 27–35, August 28–30, 2002.

[7] Chen, J. S.-C., R. Guérin, T. E. Stern, Markov-modulated flow model for the output queues of a packet switch, IEEE Trans. on Commun., Vol. COM-40, No. 6, pp. 1098–1110, 1992.

[8] B.V. Gnedenko, I.N. Kovalenko, Introduction to the Queueing Theory. Nauka Moscow, 1987.

[9] V.V. Kalashnikov, Mathematical Methods in Queueing Theory, Kluwer Academic Publishers, 1994.

[10] G. Fayolle, V. Malyshev, M. Menshikov, Constructive Theory of countable Markov chains, Cambridge University Press, 1995.

[11] A.A. Borovkov, Ergodicity and stability of stochastic processes, URSS Publ. House, Moscow, 1999.

[12] A.A. Borovkov, Stochastic Processes in Queueing Theory, Springer-Verlag, 1976.

[13] A.A. Borovkov, Probability Theory, URSS Publ. House, Moscow, 1999.

[14] A.N. Shiryayev, Probability, Nauka Moscow, 1989

[15] A.A. Borovkov, D.A. Korshunov, Large deviation probabilities for one dimensional Markov chains. Part I: stationary distributions. Theory Probab. Appl. Vol. 41, No. 1, P. 1-24, 1995.

Contents

Unité de recherche INRIA Lorraine
LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399