

# Analysis of a Two-Class Priority Queue with Correlated Arrivals from Another Node

Abdulfetah Khalid<sup>(✉)</sup>, Sofian De Clercq, Bart Steyaert, and Joris Walraevens

Department of Telecommunications and Information Processing (EA07),  
Ghent University - UGent, Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium  
{[abdulfetahhadi.khalid](mailto:abdulfetahhadi.khalid@UGent.be),[sofian.declercq](mailto:sofian.declercq@UGent.be),  
[bart.steyaert](mailto:bart.steyaert@UGent.be),[joris.walraevens](mailto:joris.walraevens@UGent.be)}@UGent.be

**Abstract.** Exact analysis of tandem priority queues is a difficult problem. In this paper, we model the output process of the first stage as a three-state Markov chain and analyze the second stage. The arrival process of this second stage is the superposition of this output process and an uncorrelated arrival process. We calculate the joint probability generating function of the number of high- and low-priority packets in the second stage and show that two implicit functions appear in this expression. We demonstrate how to deal with these implicitly defined functions in the calculation of moments.

**Keywords:** Priority queue · Generating functions · Markov chain · Correlated arrivals

## 1 Introduction

HOL (Head-of-the-Line) priority scheduling is one of the main scheduling types in network buffers to diversify the delays of traffic streams with different delay requirements [5]. When delay-sensitive high-priority packets (packets of voice and video streams, interactive gaming ...) are present in the buffer, they are transmitted. Best-effort low-priority packets can thus only be transmitted when no high-priority traffic is present.

Isolated priority queues with uncorrelated arrival processes have been studied abundantly in the past [6–8, 11]. Analysis of networks of priority queues is much more complicated. Even the simplest network, a tandem queue, combined with priority has not been solved yet. Therefore, we attempt another approach for the analysis of such a tandem priority queue. We consider one of the simplest settings, namely a discrete-time tandem priority queueing system with constant (single-slot) service times, uncorrelated arrivals of both classes to the two stages, and single servers in both stages. The idea is to approximate the output process of the first stage of the tandem queue, use this output process as input (arrival) process of the second stage, and analyze the second stage.

The first step was already taken in a previous paper [9]. We modeled the output process of a discrete-time priority queue with single-slot service times as

a three-state discrete-time Markov chain, with the three states representing ‘no output’, ‘departure of a high-priority packet’ and ‘departure of a low-priority packet’ respectively. Transition probabilities were calculated such that the mean residence times in these states correspond to mean idle and busy periods of both classes.

In this paper, we take the next step. We analyze the number of high- and low-priority packets in a priority queue with an arrival process which is the superposition of an uncorrelated arrival process (coined *external arrivals* in the remainder) and a correlated arrival process as discussed in the previous paragraph (*internal arrivals*). We show that two implicitly defined functions appear in the final expression for the joint probability generating function of the stationary number of packets of both priorities. We also demonstrate that we need to calculate one constant numerically for the calculation of the moments of the number of packets of both priorities in the system. For reference, in the same queue without the second arrival process, only one implicitly defined function is encountered and expressions of the moments of the number of packets in the system can be found completely explicitly in terms of the parameters of the arrival process [10].

Other studies of discrete-time queues with correlated arrivals have been published. Khamisy and Sidi [4] study a priority queue with a correlated arrival process according to a two-state Markovian batch arrival process. Our model is in a way simpler, since the time-correlation in our model only applies at maximum one arrival. On the other hand, we will need a three-state Markov chain instead of two states in [4]. In [2], authors analysed a queue with an arrival process that is a superposition of on-off sources. The numbers of class-1 and class-2 arrivals however are independent, which is not the case in our model. An entirely different approach to correlated arrivals are train arrivals [3, 12].

In the next section, we specify the model and notations. In Sect. 3, we calculate the joint probability generating function of the number of low- and high-priority packets in the system and discuss some numerical aspects. Finally, we demonstrate the formulas by means of some numerical examples in Sect. 4 and conclude the paper.

## 2 Mathematical Model

We consider a single-server queue with infinite buffer space. Time is assumed to be slotted. Packets of two classes arrive to the system, namely packets of class 1 and packets of class 2. The packet stream is furthermore a superposition of two arrival streams.

We denote the number of arrivals of type  $j$  of the first arrival stream (external arrivals) during slot  $k$  by  $a_{j,k}$  ( $j = 1, 2$ ). The numbers of arrivals in this stream are assumed to be i.i.d. from slot to slot and are characterized by the (common) joint probability generating function (PGF)

$$A(z_1, z_2) \triangleq \mathbb{E} [z_1^{a_{1,k}} z_2^{a_{2,k}}].$$

The marginal PGFs of the numbers of per-slot arrivals of class 1 and class 2 are denoted and given by  $A_1(z) = A(z, 1)$  and  $A_2(z) = A(1, z)$  respectively. The arrival rate  $\lambda_{j,\text{ext}}$  of class- $j$  packets ( $j = 1, 2$ ) is given by  $A'_j(1)$ .

A second arrival process (internal arrivals) is governed by a three-state discrete-time Markov chain with state space  $\{0, 1, 2\}$ . We define  $t_k$  as the state of the Markov chain in slot  $k$ . When the Markov chain is in state 0 no packets arrive in that slot. When the Markov chain is in state  $j$  ( $1$  or  $2$ ), exactly one class- $j$  packet arrives. We define the transition probabilities as

$$e_i(j) \triangleq \Pr[t_{k+1} = j | t_k = i],$$

$i, j = 0, 1, 2$ . We summarize the total arrival process in the generating function matrix  $B(z_1, z_2)$ :

$$\begin{aligned} B(z_1, z_2) &\triangleq \left[ \mathbb{E} \left[ z_1^{a_{1,k} + \mathbb{1}_{\{t_k=1\}}} z_2^{a_{2,k} + \mathbb{1}_{\{t_k=2\}}} \mathbb{1}_{\{t_k=j\}} | t_{k-1} = i \right] \right]_{i,j} \\ &= A(z_1, z_2) \begin{pmatrix} e_0(0) & e_0(1)z_1 & e_0(2)z_2 \\ e_1(0) & e_1(1)z_1 & e_1(2)z_2 \\ e_2(0) & e_2(1)z_1 & e_2(2)z_2 \end{pmatrix}, \end{aligned} \quad (1)$$

with  $\mathbb{1}_{\{X\}}$  the indicator function of  $X$ . The mean internal arrival rates of class  $j$  equals  $\lambda_{j,\text{int}} = e(j)$ , with  $e(j)$  the stationary probability that the Markov chain is in state  $j$ .

We note here in passing that the case  $e_0(1) = e_2(1)$  is an important special case, as it is the case for our motivating example (cf. [9]) and it will lead to some (numerical) simplifications.

We define  $\lambda_1$  and  $\lambda_2$  as the total arrival rate of packets of class 1 and 2 respectively, i.e.,  $\lambda_j \triangleq \lambda_{j,\text{int}} + \lambda_{j,\text{ext}}$ ,  $j = 1, 2$ . Similarly we define  $\lambda_T$  as the overall arrival rate, i.e.,  $\lambda_T \triangleq \lambda_1 + \lambda_2$ .

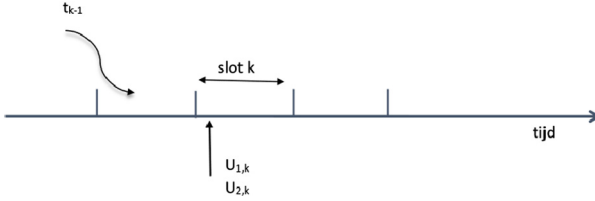
The server serves the packets at the rate of one packet per slot, i.e., all service times equal one slot. It is assumed that class-1 packets have service priority over class-2 packets and we adopt a FCFS discipline within a class.

### 3 Analysis

We denote the number of packets of class  $j$  in the system at the beginning of slot  $k$  by  $u_{j,k}$  ( $j = 1, 2$ ). The triplet  $(t_{k-1}, u_{1,k}, u_{2,k})$  is a first-order Markov chain, and is therefore suitable to analyze the queueing system. We assume that  $\lambda_T < 1$ , such that this Markov chain reaches a steady state.

#### 3.1 System Equations

Figure 1 shows the time axis for the buffer in stage 2, with  $t_{k-1}$  the class of internal packet arriving during slot  $k-1$  and  $u_{j,k}$  the buffer occupancy of class



**Fig. 1.** Time axis for the buffer analysis in stage 2

$j$  during slot  $k$  Note that  $t_{k-1} = 0$  if there is no internal arrival. The buffer occupancy of both classes are characterized by the following system of equations.

$$\begin{aligned} u_{1,k+1} &= [u_{1,k} - 1]^+ + a_{1,k} + \mathbb{1}_{\{t_k=1\}}; \\ u_{2,k+1} &= [u_{2,k} - \mathbb{1}_{\{u_{1,k}=0\}}]^+ + a_{2,k} + \mathbb{1}_{\{t_k=2\}}. \end{aligned} \tag{2}$$

These equations are understood as follow: if  $t_k = j$  ( $j = 1, 2$ ) an extra packet of class  $j$  arrives. If  $u_{1,k} > 0$ , a class-1 packet is served because of the priority scheduling; otherwise a class-2 packet is served, if any.

### 3.2 Functional Equations

We define the partial joint generating functions

$$P_j(z_1, z_2) \triangleq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p(j, m, n) z_1^m z_2^n, \quad j = 0, 1, 2,$$

with  $p(j, m, n) = \lim_{k \rightarrow \infty} \Pr[t_{k-1} = j, u_{1,k} = m, u_{2,k} = n]$ , and the row vector  $\mathbf{P}(z_1, z_2) = [P_j(z_1, z_2)]_{j=0,1,2}$ .

Transforming the system equations (2) to generating functions and letting  $k \rightarrow \infty$ , we find

$$\mathbf{P}(z_1, z_2) = \left[ \frac{\mathbf{P}(z_1, z_2) - \mathbf{P}(0, z_2)}{z_1} + \frac{\mathbf{P}(0, z_2) - \mathbf{P}(0, 0)}{z_2} + \mathbf{P}(0, 0) \right] B(z_1, z_2).$$

Solving in  $\mathbf{P}(z_1, z_2)$  yields

$$\mathbf{P}(z_1, z_2) = \left[ \frac{z_1 - z_2}{z_2} \mathbf{P}(0, z_2) + \frac{z_1(z_2 - 1)}{z_2} \mathbf{P}(0, 0) \right] B(z_1, z_2) (z_1 I - B(z_1, z_2))^{-1}, \tag{3}$$

with  $I$  the  $3 \times 3$  identity matrix.

### 3.3 Calculation of Unknowns

Expression (3) contains three unknown functions  $P_0(0, z_2)$ ,  $P_1(0, z_2)$  and  $P_2(0, z_2)$  and three unknown constants  $P_0(0, 0)$ ,  $P_1(0, 0)$  and  $P_2(0, 0)$ . However,

$P_1(0, z_2) = P_2(z_1, 0) = 0$  since at least one class- $j$  packet is present in the system if the state of the Markov chain was equal to  $j$  in the previous slot. This eliminates the unknown function  $P_1(0, z_2)$  and the unknown parameters  $P_1(0, 0)$  and  $P_2(0, 0)$ .

We now show how to calculate the other unknowns. It can be proved that  $f(z_1) \triangleq \det(z_1 I - B(z_1, z_2))$  has three zeroes inside the complex unit disk for each  $z_2$  inside the complex unit disk, see the appendix. One of them is 0. We denote the two other zeroes by  $Y_1(z_2)$  and  $Y_2(z_2)$ . Since PGFs are analytic inside the complex unit disk,  $Y_1(z_2)$  and  $Y_2(z_2)$  are zeroes of the numerators of  $P_j(z_1, z_2)$ . Using the notation  $C(z_1, z_2)$  for the cofactor matrix of  $z_1 I - B(z_1, z_2)$ , i.e.,  $C^T(z_1, z_2) = f(z_1)(z_1 I - B(z_1, z_2))^{-1}$ , this leads to the vector equations

$$[(Y_i(z_2) - z_2)\mathbf{P}(0, z_2) + Y_i(z_2)(z_2 - 1)\mathbf{P}(0, 0)]B(Y_i(z_2), z_2)C^T(Y_i(z_2), z_2) = \mathbf{0}, \tag{4}$$

for  $i = 1, 2$ .

Since  $P_i(0, 0) = 0, i \neq 0$ , we have that  $P_0(0, 0)$  equals the probability that the system is empty and thus  $P_0(0, 0) = 1 - \lambda_T$ . This leaves the calculation of the unknown functions  $P_0(0, z_2)$  and  $P_2(0, z_2)$  only. It can be shown that the three scalar equations for each  $i$  in (4) are dependent. Hence we only use the first for our system of two equations ( $i = 1, 2$ ):

$$Y_i(z_2)P_0(0, z_2) + X_i(z_2) P_2(0, z_2) = Y_i^2(z_2) \frac{1 - z_2}{Y_i(z_2) - z_2} P_0(0, 0),$$

where we used the notations

$$X_i(z_2) \triangleq A(Y_i(z_2), z_2) \sum_{j=0}^2 \frac{C_{0j}(Y_i(z_2), z_2)}{C_{00}(Y_i(z_2), z_2)} e_2(j) z_j \Big|_{z_0=1, z_1=Y_i(z_2)},$$

$$C_{ij}(z_1, z_2) \triangleq (C(z_1, z_2))_{ij},$$

and that  $f(Y_i(z_2), z_2) = 0$  can be written as

$$A(Y_i(z_2), z_2) \sum_{j=0}^2 C_{0j}(Y_i(z_2), z_2) e_0(j) z_j \Big|_{z_0=1, z_1=Y_i(z_2)} = Y_i(z_2)C_{00}(Y_i(z_2), z_2).$$

Thus we can solve for the two unknown functions, yielding

$$P_0(0, z_2) = (1 - z_2) \frac{\frac{Y_1^2(z_2)X_2(z_2)}{Y_1(z_2) - z_2} - \frac{Y_2^2(z_2)X_1(z_2)}{Y_2(z_2) - z_2}}{Y_1(z_2)X_2(z_2) - Y_2(z_2)X_1(z_2)} P_0(0, 0) \tag{5}$$

$$P_2(0, z_2) = (1 - z_2) \frac{\frac{Y_2^2(z_2)Y_1(z_2)}{Y_2(z_2) - z_2} - \frac{Y_1^2(z_2)Y_2(z_2)}{Y_1(z_2) - z_2}}{Y_1(z_2)X_2(z_2) - Y_2(z_2)X_1(z_2)} P_0(0, 0). \tag{6}$$

Not surprisingly these equations are symmetric in  $Y_1(z_2)$  and  $Y_2(z_2)$ .

### 3.4 Results

All unknown partial PGFs and constants have been calculated and this results in expressions for  $P_i(z_1, z_2)$ ,  $i = 0, 1, 2$  that depend on  $A(z_1, z_2)$ , the transition probabilities  $e_i(j)$  and the implicitly defined functions  $Y_i(z_2)$ ,  $i = 1, 2$ . In general, the latter functions have to be calculated numerically for each  $z_2$ .

From the expressions of the  $P_i(z_1, z_2)$ , the joint generating function of the stationary buffer occupancies  $u_1$  and  $u_2$  of class 1 and class 2 respectively can be calculated as

$$\begin{aligned} U(z_1, z_2) &\triangleq \mathbb{E}[z_1^{u_1} z_2^{u_2}] \\ &= P_0(z_1, z_2) + P_1(z_1, z_2) + P_2(z_1, z_2). \end{aligned}$$

From this joint PGF, marginal PGFs of the total, class-1 and class-2 buffer occupancies can be calculated as

$$\begin{aligned} U_T(z) &\triangleq \mathbb{E}[z^{u_1+u_2}] = U(z, z), \\ U_1(z) &\triangleq \mathbb{E}[z^{u_1}] = U(z, 1), \\ U_2(z) &\triangleq \mathbb{E}[z^{u_2}] = U(1, z). \end{aligned}$$

Although the resulting expressions are too large to show, we can provide some insight.

The expression for  $U_T(z)$  does not depend on the implicitly defined functions  $Y_i(z)$ , and is therefore explicitly known in terms of the arrival processes and parameters. This can also be understood from expression (3). By putting  $z_1 = z_2$  in this expression, the unknown functions  $P_i(0, z_2)$  disappear and it is these functions that introduced the implicitly defined functions in the final expressions. In fact,  $U_T(z)$  is the PGF of the buffer occupancy in a single-class buffer with a superposition of an independent arrival process and a three-state Markov chain with 1 off-state and 2 on-states.

The expression for  $U_1(z)$  does depend on the implicitly defined functions through the constants  $Y_1(1)$  and  $Y_2(1)$  only. This is logically sound as  $Y_i(z_2)$  appears in the expression of  $U(z_1, z_2)$  and we put  $z_2 = 1$  to obtain an expression for  $U_1(z)$ . We can prove that one of these two constants equals 1 and the other does not (see Appendix). Assume for instance that  $Y_1(1) = 1$ . Then,  $U_1(z)$  can be calculated completely in terms of the arrival processes and parameters apart from one constant that has to be calculated numerically ( $Y_2(1)$ ). We note that  $U_1(z)$  is, in fact, the PGF of the buffer occupancy in a single-class buffer with a superposition of an independent arrival process and a three-state Markov chain with 2 off-states and 1 on-state.

Finally, the expression for  $U_2(z)$ , which is of primary interest to us, contains both implicitly defined functions  $Y_1(z)$  and  $Y_2(z)$ , as expected.

Moments of the total buffer occupancy and the buffer occupancies of class 1 and class 2 can be calculated by means of the moment generating property of PGFs, i.e., by taking derivatives in 1. The mean buffer occupancy, for instance, is calculated by taking the first derivative of the corresponding marginal PGF in 1.

As for the PGFs, the moments of the total buffer occupancy (the class-1 buffer occupancy respectively) are found in terms of the arrival parameters (the arrival parameters and  $Y_2(1)$  respectively). The moments of the class-2 buffer occupancy are also found in terms of the arrival distributions and parameters, and  $Y_2(1)$ . The reason that no other constants but  $Y_2(1)$  appear in this expression is that  $Y_1(1) = 1$  and that the derivatives of  $Y_i(z)$  in 1 can be expressed in terms of  $Y_i(1)$ . With some abuse of notation, we obtain, for instance,

$$Y_i'(1) = - \left. \frac{\frac{\partial f(z_1, z_2)}{\partial z_2}}{\frac{\partial f(z_1, z_2)}{\partial z_1}} \right|_{z_1=Y_i(1), z_2=1},$$

by considering  $f$  as a function of  $z_1$  and  $z_2$  (instead of  $z_1$  only). We conclude that all moments of the buffer occupancies can be calculated explicitly in terms of the arrival distributions apart from one constant that has to be calculated numerically.

### 4 Numerical Results

We now demonstrate our results for some numerical examples.

We assume that the external (independent) arrival process is a two-class Bernoulli process, i.e.,

$$A(z_1, z_2) = (1 - \lambda_{1,ext} - \lambda_{2,ext}) + \lambda_{1,ext} z_1 + \lambda_{2,ext} z_2 .$$

In choosing this easy arrival process we end up with two similar Bernoulli arrival processes: the external one is a time-independent process and the internal one a time-dependent one. This will therefore demonstrate the effect of this time-dependence fairly.

The internal arrival process is characterized by the 9 transition probabilities  $e_i(j)$ ,  $i, j = 0, 1, 2$ , of which 6 can be independently chosen ( $\sum_{j=0}^2 e_i(j) = 1$ ). We first assume  $e_0(1) = e_2(1)$ , cf. the discussion in Sect. 2. The advantage is that  $Y_2(1)$  disappears from the expression of  $E[u_1]$  since the 2 off-states can be merged to 1 off-state in this particular case. Therefore, we end up – for the class-1 buffer occupancy – with a single-class system with a superposition of an independent arrival process and a two-state Markov chain with 1 off-state and 1 on-state. Therefore, the mean total buffer occupancy and the mean class-1 buffer occupancy can be calculated explicitly. Since the mean class-2 buffer occupancy is the difference of these two, this measure can also be calculated explicitly in the case  $e_0(1) = e_2(1)$ . Again, we stress that this is an important special case.

We are then left with 5 independent parameters. To demonstrate our results more intuitively, we map these to 5 other parameters, namely to the traffic intensities of both classes  $\lambda_{1,int}$  and  $\lambda_{2,int}$  and to  $L_i$ ,  $i = 0, 1, 2$  defined as

$$L_i \triangleq \frac{1}{1 - e_i(i)} .$$

The  $L_i$  are in fact the mean residence times in state  $i$  (the residence time in state  $i$  is shifted geometrically distributed with parameter  $e_i(i)$ ) and serve therefore as measures for the time-correlation: the larger  $L_i$ , the more slots on average the Markov chain resides in state  $i$ . The  $\lambda_{i,int}$ , on the other hand, are measures for the frequency of being in the different states.

A difficulty with applying this mapping is that not every ‘logical’ combination of  $(\lambda_{1,int}, \lambda_{2,int}, L_0, L_1, L_2)$  leads to values in  $[0, 1]$  for each of the 9 transition probabilities. In the remainder, we filtered out the ones that do, but this restricts the range of values for the parameters somewhat. We focus on the influence of correlation in the arrival process.

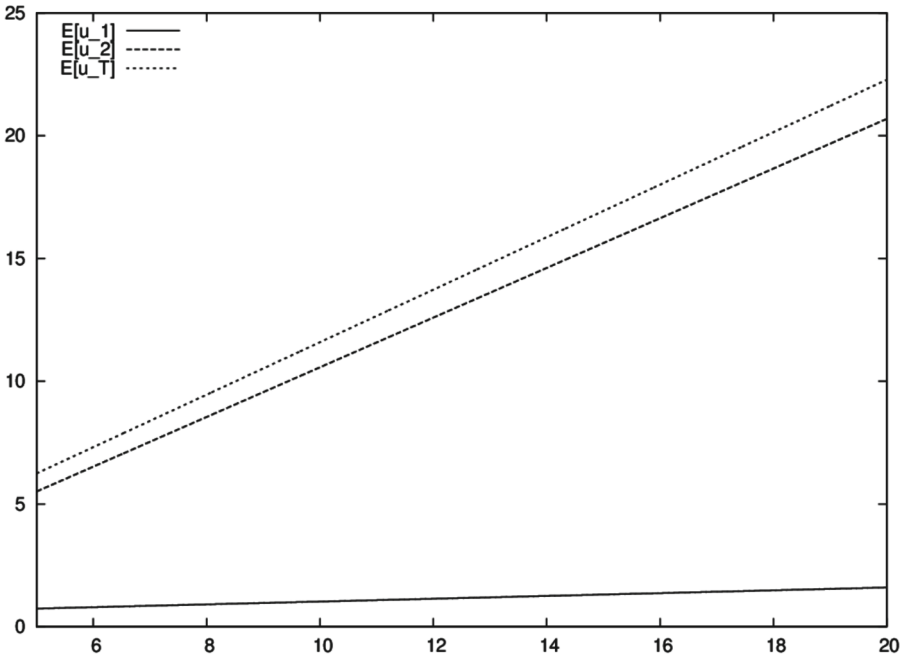


Fig. 2. Mean values of system contents versus the mean residence time  $L$

In Fig. 2, we show the influence of the lengths  $L_i$  on the mean buffer occupancy. We depict the mean total class-1 and class-2 buffer occupancies as a function of  $L$ , with  $L_i = L$ ,  $i = 0, 1, 2$ . We assume a perfect balance of class-1 and class-2 arrival rates, with  $\lambda_{1,int} = \lambda_{2,int} = 0.3, \lambda_{1,ext} = \lambda_{2,ext} = 0.15$ . We firstly observe that a larger mean residence time  $L$  in the states leads to higher mean buffer occupancies of both classes. The reason is that for long periods at least 1 packet per slot arrives (internal arrivals) while occasionally a second packet arrives in that slot (external arrivals). Therefore, the buffer occupancy



can not decrease during these (long) periods. Secondly, it is clear that the low-priority class is affected much more. In fact, its slope is about the same as the slope of the mean total buffer occupancy. This is again explained by these long consecutive periods of arrivals: class-1 is only ‘punished’ by long periods of its own class while class 2 suffers from long periods of both (which can even directly follow each other).

Figure 3 depicts the influence of the fraction of correlated and independent arrivals. The mean buffer occupancies are shown as function of  $\lambda_{T,int}$  for  $\lambda_T = 0.9$ ,  $L_i = 5$  and perfect balance between class-1 and class-2 arrivals. We notice that time-correlation is not necessarily the dominant factor: time-correlation increases when more internal packets arrive, while the mean buffer occupancies decrease. The main reason is that the variance of the number of arrivals in a slot changes as well: when  $\lambda_{T,int} = 0$  or  $\lambda_{T,ext} = 0$  (the two extremes of the plot; the left extreme cannot be shown because of the remark above), at most one packet can arrive in the system per slot and no queueing happens. When  $\lambda_{T,int}$  and  $\lambda_{T,ext}$  are more balanced, the probability of two arrivals in a slot is significantly higher. This leads to a maximum of the buffer occupancies somewhere in the middle.

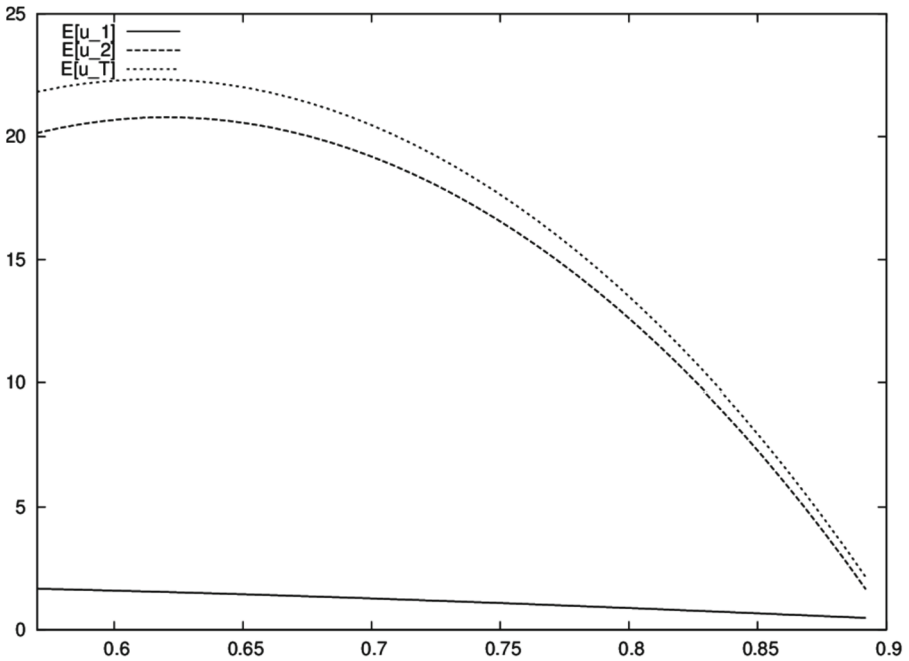


Fig. 3. Mean values of system contents versus  $\lambda_{T,int}$

## 5 Conclusions and Future Work

We analysed a discrete-time two-class priority queue with an arrival process that is a superposition of an independent process and a time-dependent process. The latter is modeled by a discrete-time Markov chain with three states, no arrivals when in state 0 and one class-1 or class-2 arrivals when in state 1 or state 2 respectively. This model is motivated by a priority tandem queue, where our system can be regarded as the second stage of the tandem and the two arrival processes as models for external and internal (output from the first stage) arrivals from the network. We analysed the buffer occupancy in this system and demonstrated some results. We also identified an important special case that simplifies the analysis and numerical work.

In future work, we will analyse sojourn times. Note that we can define several different sojourn times, not only from different classes but also either of a packet of the external or the internal arrival stream. Furthermore, since the three-state Markov chain is only a very coarse model for the output process of the first stage, we need to test whether the results of our model are accurate approximations of the second stage performance measures of a priority tandem queue. If not, we will need to come up with a more complex model for the internal arrival process which will evidently complicate analysis.

**Acknowledgement.** The authors acknowledge support of the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office and of BOF-UGent. Sofian De Clercq is a BOF postdoctoral fellow of UGent.

## Appendix

In this appendix, we first prove that  $f(z_1) = \det(z_1 I - B(z_1, z_2))$  has three zeroes inside the complex unit disk for each  $z_2$  with  $|z_2| < 1$  and that one of them is 0. The function  $f(z_1)$  can be written as  $z_1(g(z_1) - h(z_1))$  with

$$\begin{aligned} g(z_1) &= (z_1 - e_0(0)A(z_1, z_2))(1 - e_1(1)A(z_1, z_2))(z_1 - e_2(2)A(z_1, z_2)), \quad (7) \\ h(z_1) &= - (z_1 - e_0(0)A(z_1, z_2))e_1(2)e_2(1)z_2A(z_1, z_2)^2 \\ &\quad - (1 - e_1(1)A(z_1, z_2))e_0(2)e_2(0)z_2A(z_1, z_2)^2 \\ &\quad - (z_1 - e_2(2)A(z_1, z_2))e_0(1)e_1(0)A(z_1, z_2)^2 \\ &\quad - e_0(1)e_1(2)e_2(0)z_2A(z_1, z_2)^3 - e_0(2)e_2(1)z_2e_1(0)A(z_1, z_2)^3. \end{aligned}$$

The factor  $z_1$  contributes for the zero 0. We will prove that  $|g(z_1)| > |h(z_1)|$  on the complex unit circle and hence, due to Rouché's theorem, that the number of zeroes of  $g(z_1) - h(z_1)$  inside the complex unit circle equals the number of zeroes of  $g(z_1)$ . Finally, we prove that the latter has two zeroes, again by means of Rouché's theorem.

We have that

$$\begin{aligned} \frac{|h(z_1)|}{|g(z_1)|} &< \frac{e_1(2)e_2(1)|A(z_1, z_2)|^2}{|1 - e_1(1)A(z_1, z_2)||z_1 - e_2(2)A(z_1, z_2)|} \\ &+ \frac{e_0(2)e_2(0)|A(z_1, z_2)|^2}{|z_1 - e_0(0)A(z_1, z_2)||z_1 - e_2(2)A(z_1, z_2)|} \\ &+ \frac{e_0(1)e_1(0)|A(z_1, z_2)|^2}{|z_1 - e_0(0)A(z_1, z_2)||1 - e_1(1)A(z_1, z_2)|} \\ &+ \frac{e_0(1)e_1(2)e_2(0)|A(z_1, z_2)|^3}{|z_1 - e_0(0)A(z_1, z_2)||1 - e_1(1)A(z_1, z_2)||z_1 - e_2(2)A(z_1, z_2)|} \\ &+ \frac{e_0(2)e_2(1)e_1(0)|A(z_1, z_2)|^3}{|z_1 - e_0(0)A(z_1, z_2)||1 - e_1(1)A(z_1, z_2)||z_1 - e_2(2)A(z_1, z_2)|}. \end{aligned} \quad (8)$$

If we further use that, for  $|z_1| = 1$ ,  $|A(z_1, z_2)| \leq 1$ , that  $|z_1 - e_j(j)A(z_1, z_2)| \geq 1 - e_j(j)$  ( $j = 0, 2$ ), that  $1 - e_1(1)A(z_1, z_2) \geq 1 - e_1(1)$  and that  $\sum_{j=1}^2 e_i(j) = 1$ , we find that the RHS of (8) is less than or equal to 1. Therefore,  $|h(z_1)| < |g(z_1)|$  for  $|z_1| = 1$  and the number of zeroes of  $g(z_1) - h(z_1)$  in the complex unit circle equals the number of zeroes of  $g(z_1)$ . The latter function exists of three factors, see (7). Using  $|z_1 - e_j(j)A(z_1, z_2)| \geq 1 - e_j(j)$  ( $j = 0, 2$ ) and  $1 - e_1(1)A(z_1, z_2) \geq 1 - e_1(1)$  once more, we find that each of the factors has as many zeroes inside the complex unit disk as its first terms. Therefore,  $g(z_1)$  - and  $g(z_1) - h(z_1)$  - has two zeroes.

Next, we prove that the limit for  $z_2 \rightarrow 1$  of one of the two zeroes equals 1 while the other one does not. If we substitute  $z_2$  by 1 in  $f(z_1)$ , Rouché’s theorem can no longer be used directly, because  $f(z_1)$  is not necessarily analytic for  $z_2 = 1$  and  $|z_1| = 1$ . The result is however still valid, cf. [1], i.e.,  $f(z_1)$  has two zeroes inside and on the complex unit disk. One of them is equal to 1, since  $f(1) = 0$  for  $z_2 = 1$ . Finally, it is easily proved that  $f'(1) \neq 0$  for  $z_2 = 1$  which means  $z_1 = 1$  is a single zero of  $f$  for  $z_2 = 1$ . The other root is therefore different from 1 for  $z_2 \rightarrow 1$ . It will however lie in  $] - 1, 1[$  which simplifies a numerical search.

## References

1. Adan, I., Van Leeuwen, J.S.H., Winands, E.M.M.: On the application of Rouché’s theorem in queueing theory. *Oper. Res. Lett.* **34**, 355–360 (2006)
2. Ali, M.M., Song, X.: A performance analysis of a discrete-time priority queueing system with correlated arrivals. *Perform. Eval.* **57**, 307–339 (2004)
3. Choi, B.D., Choi, D.I., Lee, Y., Sung, D.K.: Priority queueing system with fixed-length packet-train arrivals. *IEE Proc.-Commun.* **145**, 331–336 (1998)
4. Khamisy, A., Sidi, M.: Discrete-time priority queues with two-state Markov modulated arrivals. *Stoch. Models* **8**, 337–357 (1992)
5. Roberts, J.: Internet traffic, QoS, and pricing. *Proc. IEEE* **92**, 1389–1399 (2005)
6. Sleptchenko, A., Selen, J., Adan, I., van Houtum, G.J.: Joint queue length distribution of multi-class, single-server queues with preemptive priorities. *QUESTA* **81**, 379–395 (2015)
7. Takagi, H.: *Queueing analysis: a foundation of performance evaluation.* (1991–1993)

8. Tarabia, A.: Two-class priority queueing system with restricted number of priority customers. *AEÜ* **61**, 534–539 (2007)
9. Walraevens, J., Fiems, D., Wittevrongel, S., Bruneel, H.: Calculation of output characteristics of a priority queue through a busy period analysis. *EJOR* **198**, 891–898 (2009)
10. Walraevens, J., Steyaert, B., Bruneel, H.: Performance analysis of a single-server ATM queue with a priority scheduling. *C&OR* **30**, 1807–1829 (2003)
11. Walraevens, J., Steyaert, B., Bruneel, H.: Analysis of a discrete-time preemptive resume priority buffer. *EJOR* **186**, 182–201 (2008)
12. Walraevens, J., Wittevrongel, S., Bruneel, H.: A discrete-time priority queue with train arrivals. *Stoch. Models* **23**, 489–512 (2007)