BACHELOR

# ASIP : asymmetric simple inclusion process

Suijkerbuijk, C.A.W.M.

*Award date:*
2015

# ASIP:
# Asymmetric Simple Inclusion Process

Corné Suijkerbuijk
Student number: 0808814
c.a.w.m.suijkerbuijk@student.tue.nl

05-02-2015

Bachelor final project (2WH40)
Industrial and Applied Mathematics
University of Technology Eindhoven

Supervisor: Prof. dr. ir. O.J. Boxma

# 1 Abstract

An ASIP is an queueing system which will be introduced in the introduction.
In the literature study some properties of an homogeneous ASIP will be investigated using Monte Carlo simulation. The properties are: number of particles, load, draining time, interexit time and coalescence time. Then the traversal time will be discussed. The number of particles in an ASIP is a Markov process which gives the opportunity to investigate the expected number of particles in steady state. Using the Markovian dynamics the probability generating function for number of particles in the system will be determined. In the next section some asymptotic analysis about the traversal time, overall load, busy period, first occupied queue and draining time will be done. In the final section of the literature study some generalizations on the ASIP model will be discussed. The first generalization is that the order in which gates open is a Markov process instead of gates that open independently of each other. The second generalization is that the time between two consecutive gate openings is according to a general random variable instead of being exponentially distributed. The third generalization is that particles may enter the system at all queues instead of only entering the system via the first queue.
In my own research part the ASIP with queue capacities will be introduced. The purpose is to allocate a total capacity C over the queues of the ASIP such that the throughput of particles is maximized. In order to determine the throughput the limiting distribution of number of particles in the queues will be determined. Then we will take a look at the number of possibilities of allocating $C$ capacity over $n$ queues. Then a lemma follows which says that a small subset of all solutions contains an optimal solution. By this reduction it will be possible to solve the optimization problem by trying all solutions in this small subset. For allocating a total capacity of around 100 or less this can be done in short time. In the next section a general formula for the throughput will be derived in the case when the capacity of each queue is one. My own research part will be concluded by examples of allocating total capacity over an ASIP with 20 queues for certain gate opening intensities.

# Contents

## 2 Introduction

The ASIP (= Asymmetric Simple Inclusion Process) is a system with a number of sequential queues, say n queues. The queues have unbounded capacity. Particles will enter the system according to a Poisson process with parameter $\lambda$. After each queue there is a gate which is closed most of the time. Gate i will open according to a Poisson process with parameter $\mu_i$. At the moment gate i opens, all the particles which are located in queue i will move to queue i+1 and after that gate i will close immediately. At the moment gate n opens, all the particles which are located in queue n leave the system. If $\mu_1 = \mu_2 = ... = \mu_n$, then we call it an homogeneous ASIP. Figure 1 [2] shows a visual representation of the ASIP model.

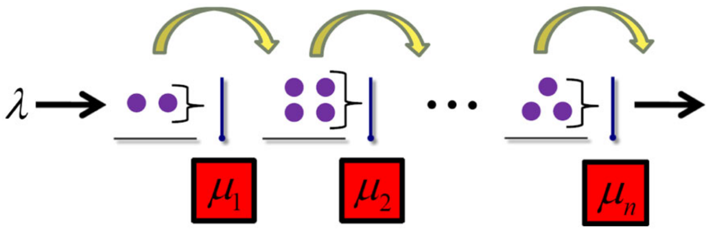


Figure 1: Visual representation of ASIP model

The ASIP is Asymmetric because the direction of movement of the particles is uniform. Particles can only move to queues with a higher index and not the other way around.
The ASIP is Simple because the queues are in a sequence; there is only one way to go through the system.
The ASIP is an Inclusion Process because all particles go to the next queue simultaneously when the gate opens. In an Exclusion Process one particle goes to the next queue if the gate opens and the next queue is empty. So in the Exclusion Process there is only one or zero particle in each queue. In an Inclusion Process it doesn't matter whether the next queue is empty.

# 3 Literature Study

This chapter is organized as follows. In Section 3.1 more insights will be given about a homogeneous ASIP in steady state. In Section 3.2 the traversal time of a particle in the system will be analyzed. In Section 3.3 the Markovian dynamics of the number of particles in the system will be given. This Markovian dynamics will be used in Section 3.4 in order to analyze the expected number of particles in the system. In Section 3.5 the Probability Generating Function of the number of particles will be formulated and analyzed in steady state. In Section 3.6 some asymptotic analysis about ASIP's observations will be done. Finally, in Section 3.7 the ASIP model will be generalized in the sense of having a general random variable for the time between two gate openings and having the possibility that particles may arrive at all queues.

## 3.1 Monte Carlo simulation homogeneous ASIP

This section is a based on the journal article "Asymmetric Inclusion Process as a Showcase of Complexity" [2]. In this section the homogeneous ASIP with $\lambda = \mu_1 = ... = \mu_n = 1$ in steady state will be considered where n is large. First a definition of the number of particles and some asymptotic analysis will be given by using Monte Carlo simulation. Then four characterizations of the ASIP will be defined: load, draining time, interexit time and coalescence time. For each of these characterizations a stochastic approximation will be done using Monte Carlo Simulation.

### 3.1.1 Number of particles

Define $X_k(t)$ as the number of particles in queue k at time t. Define $X(t) = (X_1(t), ..., X_n(t))$ as the vector of these random variables. Also define $X_k = \lim_{t\to\infty} X_k(t)$ as the number of particles in queue k at steady state. Finally define $X = (X_1, ..., X_n)$ as the vector of these random variables at steady state.
Monte Carlo Simulation of this homogeneous ASIP gives the following result:

$$\mathbb{P}(X_k > 0) \approx k^{-\frac{1}{2}}. \tag{1}$$

This sounds logic because each time a gate opens a cluster of particles will merge with a cluster of particles in the next queue, so clusters of particles will be greater in the last queues than in the first ones and particles enter the system only via the first queue. So larger clusters means less queues occupied so the probability that some particles wait in queue k decreases as k becomes larger. This also clarifies that the number of particles in queue k increases as k becomes larger which is also a result of the Monte Carlo simulation:

$$\mathbb{E}[X_k | X_k > 0] \approx k^{\frac{1}{2}} \tag{2}$$

The final result of the Monte Carlo simulation is about the standard deviation of $X_k$:

$$\frac{\sigma(X_k)}{\mathbb{E}[X_k]} \approx k^{\frac{1}{4}} \tag{3}$$

When k is large, most of the time there is no particle in the $k^{th}$ queue and if there are particles in the $k^{th}$ queue it is a relative large cluster of particles as mentioned above. So the standard deviation on the number of particles increases as k becomes larger.

### 3.1.2 Load

Load is defined as the total number of particles in the system (=$|X|$). The load can be approximated by:

$$|X| \sim \sqrt{2n} \cdot \Theta + n. \tag{4}$$

Here $\Theta$ is a Gaussian random variable with mean zero and variance one.

#### 3.1.3 Draining time

Draining time is defined as time that is takes until all particles in steady state have left the system where no particles enter the system. More formally: If $\lambda = 0$ and $X(0) = X$, then the draining time is defined as $\inf\{t \geq 0 || X(t)| = 0\}$. The draining time can be approximated by:

$$\inf\{t \geq 0 || X(t)| = 0\} \sim \sqrt{n} \cdot \Theta + n. \tag{5}$$

Here $\Theta$ is again a Gaussian random variable with mean zero and variance one.

#### 3.1.4 Interexit time

The interexit time $IT$ is defined as the time between two consecutive moments where particles leave the system. The Interexit time can be approximated by:

$$IT \sim \sqrt{\pi n} \cdot \Theta. \tag{6}$$

Here $\Theta$ is a Rayleigh distribution with mean one and tail probability:

$$\mathbb{P}(\Theta > t) = e^{-\frac{\pi t^2}{4}} \qquad \text{for} \quad t > 0. \tag{7}$$

#### 3.1.5 Coalescence time

Assume that the ASIP is circular. This means that leaving particles of the last queue will enter the first queue. Further assume that all queues are occupied at time $t = 0$. During this process particles will cluster with each other and there comes a moment that all particles in the system have been merged to one single cluster. The time till this happens is defined as the coalescence time $CT$. The coalescence time can be approximated by:

$$CT \sim \frac{n^2}{6} \cdot \Theta. \tag{8}$$

Here $\Theta$ is an inverse Gaussian random variable with mean one and density:

$$\frac{d}{dt}\mathbb{P}(\Theta \leq t) = \frac{1}{\sqrt{\frac{4\pi}{5}}} \cdot t^{-\frac{3}{2}} \cdot e^{-\frac{(t-1)^2}{\frac{4t}{5}}}. \tag{9}$$

### 3.2 Traversal time

This section is based on page 3 of the journal article "Asymmetric Inclusion Process" [3].

The traversal time T is defined as the time between the moment a particle enters the system and the moment the particle leaves the system by leaving the last queue. During the progress a particle has to visit each queue once. By the fact that gate k opens according to a Poisson Process with parameter $\mu_k$ the waiting time between two consecutive openings of gate k is exponentially distributed with parameter $\mu_k$. Because of the memoryless property of the exponential distribution the time for a particle staying in queue k is exponentially distributed with parameter $\mu_k$. Because the gates open independently of each other the traversal time is the sum of the waiting times in each queue:

$$T = E_1 + ... + E_n \qquad \text{with} \quad E_i \sim Exp(\mu_i). \tag{10}$$

The expected traversal time and his variance are then:

$$\mathbb{E}[T] = \frac{1}{\mu_1} + ... + \frac{1}{\mu_n}. \tag{11}$$

$$Var(T) = \frac{1}{\mu_1^2} + ... + \frac{1}{\mu_n^2}. \tag{12}$$

Note that the traversal time of an homogeneous ASIP is an Erlang distribution: $T_{hom} \sim Erlang(\mu_1, n)$.

### 3.3 Markovian dynamics

This section is based on page 4 of the journal article "Asymmetric Inclusion Process" [3].

The openings of the gates follow a Markov Process. The number of particles in the system at time t is $X(t) = (X_1(t), ..., X_n(t))$. In order to investigate the dynamics of the number of particles it is interesting to look at $X(t+\Delta)$ where $\Delta$ is very small. According to the Markovian dynamics $X(t+\Delta)$ is:

$$X(t+\Delta) = \begin{cases} X(t) & \text{w.p.} \quad 1-(\lambda+\mu)\Delta + o(\Delta) \\ (X_1(t)+1, X_2(t), X_3(t), ..., X_n(t)) & \text{w.p.} \quad \lambda\Delta + o(\Delta) \\ (0, X_1(t)+X_2(t), X_3(t), X_4(t), ..., X_n(t)) & \text{w.p.} \quad \mu_1\Delta + o(\Delta) \\ (X_1(t), 0, X_2(t)+X_3(t), X_4(t), ..., X_n(t)) & \text{w.p.} \quad \mu_2\Delta + o(\Delta) \\ : & : \\ (X_1(t), X_2(t), ..., X_{n-2}(t), 0, X_{n-1}(t)+X_n(t)) & \text{w.p.} \quad \mu_{n-1}\Delta + o(\Delta) \\ (X_1(t), X_2(t), ..., X_{n-1}(t), 0) & \text{w.p.} \quad \mu_n\Delta + o(\Delta). \end{cases} \tag{13}$$

In this equation $\mu = \mu_1 + ... + \mu_n$ and w.p. means with probability. The first line is the case when nothing happens. The second line is the case when one particle enters the system. The third line is the case when gate 1 opens. etc.

The number of particles and their positions in the system only change when a particle enters the system or a gate opens. Call these proceedings Poissonian events. Define $Y_k(s)$ as the number of particles in queue k after the $s^{th}$ Poissonian event took place. Also define $Y(s) = (Y_1(s), ..., Y_n(s))$. In order to investigate the dynamics of $Y(s)$ it is interesting to look at the number of particles after the $(s+1)^{th}$ Poissonian event. According to the Markovian dynamics you get the following:

$$Y(s+1) = \begin{cases} (Y_1(s)+1, Y_2(s), Y_3(s), ..., Y_n(s)) & \text{w.p.} \quad \frac{\lambda}{\lambda+\mu} \\ (0, Y_1(s)+Y_2(s), Y_3(s), Y_4(s), ..., Y_n(s)) & \text{w.p.} \quad \frac{\mu_1}{\lambda+\mu} \\ (Y_1(s), 0, Y_2(s)+Y_3(s), Y_4(s), ..., Y_n(s)) & \text{w.p.} \quad \frac{\mu_2}{\lambda+\mu} \\ : & : \\ (Y_1(s), Y_2(s), ..., Y_{n-2}(s), 0, Y_{n-1}(s)+Y_n(s)) & \text{w.p.} \quad \frac{\mu_{n-1}}{\lambda+\mu} \\ (Y_1(s), Y_2(s), ..., Y_{n-1}(s), 0) & \text{w.p.} \quad \frac{\mu_n}{\lambda+\mu}. \end{cases} \tag{14}$$

The first line is the case when a particle enters the system. The second line is the case when the first gate opens. The third line is the case when the second gate opens. etc.

### 3.4 Dynamics of expected number of particles

This section is based on page 5 and 6 of the journal article "Asymmetric Inclusion Process" [3].

The expectation of the number of particles in the system will be investigated in this section. First by finding a differential equation for $X(t)$ and a difference equation for $Y(s)$ and finally investigating the expected number of particles in steady state.

#### 3.4.1 Dynamics expectation of X(t)

Denote $\mathbb{E}[X(t)] = (\mathbb{E}[X_1(t)], ..., \mathbb{E}[X_n(t)])^\top$. Now we are going to look at $\mathbb{E}[X(t+\Delta)]$ by conditioning on $X(t)$. According to the Markovian dynamics in (13) you get the following:

$$\mathbb{E}[X(t+\Delta)] = \mathbb{E}[\mathbb{E}[X(t+\Delta)|X(t)]] = \begin{cases} (1-(\lambda+\mu)\Delta)\mathbb{E}[X(t)] \\ +(\lambda\Delta)\mathbb{E}[X(t) + (1,0,...,0)^\top] \\ +(\mu_1\Delta)\mathbb{E}[X(t) + (-X_1(t), X_1(t), 0, ..., 0)^\top] \\ +(\mu_2\Delta)\mathbb{E}[X(t) + (0, -X_2(t), X_2(t), 0, ..., 0)^\top] \\ +...+ \\ +(\mu_{n-1}\Delta)\mathbb{E}[X(t) + (0, ..., 0, -X_{n-1}(t), X_{n-1}(t))^\top] \\ +(\mu_n\Delta)\mathbb{E}[X(t) + (0, ..., 0, -X_n(t))^\top] \\ +o(\Delta). \end{cases} \tag{15}$$

Note that $1-(\lambda+\mu)\Delta + \lambda\Delta + \mu_1\Delta + \mu_2\Delta + ... + \mu_n\Delta = 1$. So we can take $1 \cdot \mathbb{E}[X(t)]$ to the other side and divide both sides by $\Delta$:

$$\frac{\mathbb{E}[X(t+\Delta)] - \mathbb{E}[X(t)]}{\Delta} = \begin{cases} \lambda \cdot (1,0,...,0)^\top \\ +\mu_1 \cdot (-\mathbb{E}[X_1(t)], \mathbb{E}[X_1(t)], 0, ..., 0)^\top \\ +\mu_2 \cdot (0, -\mathbb{E}[X_2(t)], \mathbb{E}[X_2(t)], 0, ..., 0)^\top \\ +...+ \\ +\mu_{n-1} \cdot (0, ..., 0, -\mathbb{E}[X_{n-1}(t)], \mathbb{E}[X_{n-1}(t)])^\top \\ +\mu_n \cdot (0, ..., 0, -\mathbb{E}[X_n(t)])^\top \\ +\frac{o(\Delta)}{\Delta}. \end{cases} \tag{16}$$

Now taking $\Delta \to 0$:

$$\frac{d\,\mathbb{E}[X(t)]}{dt} = \begin{pmatrix} -\mu_1 & & & & & \\ \mu_1 & -\mu_2 & & & & \\ & \mu_2 & \cdot & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & -\mu_{n-1} & \\ & & & & \mu_{n-1} & -\mu_n \end{pmatrix} \cdot \mathbb{E}[X(t)] + \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix} = M \cdot \mathbb{E}[X(t)] + \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{17}$$

Note that the empty elements in matrix $M$ mean zeros. So the mean dynamics of $X(t)$ can be characterized by this differential equation which has the following solution:

$$\mathbb{E}[X(t)] = M^{-1}(e^{Mt} - I) \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{18}$$

Here $I$ is an identity matrix of size n.

#### 3.4.2 Dynamics expectation of Y(s)

In the same way Y(s) can be analyzed. Denote $\mathbb{E}[Y(s)] = (\mathbb{E}[Y_1(s)], ..., \mathbb{E}[Y_n(s)])$. Now we are going to look at $\mathbb{E}[Y(s+1)]$ by conditioning on $Y(s)$. According to the Markovian dynamics in

(14) you get the following:

$$\mathbb{E}[Y(s+1)] = \mathbb{E}[\mathbb{E}[Y(s+1)|Y(s)]] = \begin{cases} \frac{\lambda}{\lambda+\mu} \cdot \mathbb{E}[Y(s) + (1, 0, ..., 0)^\top] \\ +\frac{\mu_1}{\lambda+\mu} \cdot \mathbb{E}[Y(s) + (-Y_1(s), Y_1(s), 0, ..., 0)^\top] \\ +\frac{\mu_2}{\lambda+\mu} \cdot \mathbb{E}[Y(s) + (0, -Y_2(s), Y_2(s), 0, ..., 0)^\top] \\ +...+ \\ +\frac{\mu_{n-1}}{\lambda+\mu} \cdot \mathbb{E}[Y(s) + (0, ..., 0, -Y_{n-1}(s), Y_{n-1}(s))^\top] \\ +\frac{\mu_n}{\lambda+\mu} \cdot \mathbb{E}[Y(s) + (0, ..., 0, -Y_n(s))^\top]. \end{cases} \tag{19}$$

Note that $\frac{\lambda}{\lambda+\mu} + \frac{\mu_1}{\lambda+\mu} + + \frac{\mu_2}{\lambda+\mu} + ... + + \frac{\mu_n}{\lambda+\mu} = 1$. So we can take $1 \cdot \mathbb{E}[Y(s)]$ to the other side and multiply both sides with $(\lambda + \mu)$:

$$(\lambda + \mu) \cdot (\mathbb{E}[Y(s+1)] - \mathbb{E}[Y(s)]) = M \cdot \mathbb{E}[Y(s)] + \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{20}$$

In this equation M is the same matrix as in Section 3.4.1. So the mean dynamics of $Y(s)$ can be characterized by this difference equation which has the following solution:

$$\mathbb{E}[Y(s)] = M^{-1}[(I + \frac{M}{\lambda+\mu})^s - I] \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{21}$$

Here $I$ is again an identity matrix of size n.

### 3.4.3 Expected number of particles in steady state

The expected number of particles in steady state is $\mathbb{E}[X]$. In steady state the expected number of particles is a constant, so $\frac{d\,\mathbb{E}[X(t)]}{dt} = 0$. Substituting this in (17) gives:

$$0 = M \cdot \mathbb{E}[X] + \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{22}$$

The expected number of particles in steady state is also $\mathbb{E}[Y]$. In steady state the expected number of particles is a constant, so $\mathbb{E}[Y(s+1)] - \mathbb{E}[Y(s)] = 0$. Substituting this in (20) gives:

$$0 = M \cdot \mathbb{E}[Y] + \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{23}$$

This is the same result as in (22) which is not surprising, because $\mathbb{E}[X]$ is the number of particles at time $t$ with $t \to \infty$ and infinitely many Poissonian events took place between time zero and time infinity.
Solving (22) and (23) leads to the following result:

$$\mathbb{E}[X_i] = \mathbb{E}[Y_i] = \frac{\lambda}{\mu_i} \qquad \forall i \in \{1, ..., n\}. \tag{24}$$

Note that this is the same result as when you have n independent ASIP's where ASIP i has one queue where particles come into ASIP i with intensity $\lambda$ and the gate opens with intensity $\mu_i$. So in steady state the expected number of particles in queue i is independent of the gate opening intensities of the other queues.

If we combine (11) and (24) we can find a formula for the number of particles in the system in steady state:

$$\mathbb{E}[\sum_{k=1}^{n} X_k(t)] = \mathbb{E}[\sum_{k=1}^{n} Y_k(s)] = \sum_{k=1}^{n} \frac{\lambda}{\mu_k} = \lambda\mathbb{E}[T]. \tag{25}$$

This formula says that the mean number of particles in steady state is the intensity of incoming particles $\lambda$ multiplied by the mean traversal time of one particle $\mathbb{E}[T]$. This can be recognized as a version of Little's law in queueing theory.

## 3.5 Probability Generating Function

This section is based on pages 7 to 10 of the journal article "Asymmetric Inclusion Process" [3].

In order to analyze the dynamics of the number of particles in the system in more detail, we want to find the probability generating functions of $X(t)$ and $Y(s)$. When we have an explicit formula for the probability generating functions, then we are able to calculate the expected number of particles in the system.

### 3.5.1 Probability Generating Function of X(t)

The probability generating functions of $X(t)$ is defined as follows:

$$G_X(t, z_1, z_2, ..., z_n) = \mathbb{E}[z_1^{X_1(t)} \cdot z_2^{X_2(t)} \cdot ... \cdot z_n^{X_n(t)}]. \tag{26}$$

Now we are going to look at the probability generating function at time $t + \Delta$ by conditioning on the number of particles at time t. According to the Markovian dynamics in (13) you get the following:

$$\mathbb{E}[\prod_{k=1}^{n} z_k^{X_k(t+\Delta)}] = \mathbb{E}[\mathbb{E}[\prod_{k=1}^{n} z_k^{X_k(t+\Delta)} | X(t)]] = \begin{cases} (1-(\lambda+\mu)\Delta)\mathbb{E}[\prod_{k=1}^{n} z_k^{X_k(t)}] \\ +(\lambda\Delta)\mathbb{E}[z_1 \prod_{k=1}^{n} z_k^{X_k(t)}] \\ +(\mu_1\Delta)\mathbb{E}[z_2^{X_1(t)} \prod_{k\neq 1}^{n} z_k^{X_k(t)}] \\ +(\mu_2\Delta)\mathbb{E}[z_3^{X_2(t)} \prod_{k\neq 2}^{n} z_k^{X_k(t)}] \\ +...+ \\ +(\mu_{n-1}\Delta)\mathbb{E}[z_n^{X_{n-1}(t)} \prod_{k\neq(n-1)}^{n} z_k^{X_k(t)}] \\ +(\mu_n\Delta)\mathbb{E}[\prod_{k\neq n}^{n} z_k^{X_k(t)}] \\ +o(\Delta). \end{cases} \tag{27}$$

Using the notation of probability generating function in (26) gives:

$$G_X(t+\Delta, z_1, z_2, ..., z_n) = \begin{cases} (1-(\lambda+\mu)\Delta)G_X(t, z_1, z_2, ..., z_n) \\ +(\lambda\Delta)z_1 G_X(t, z_1, z_2, ..., z_n) \\ +(\mu_1\Delta)G_X(t, z_2, z_2, ..., z_n) \\ +(\mu_2\Delta)G_X(t, z_1, z_3, z_3, z_4, ..., z_n) \\ +...+ \\ +(\mu_{n-1}\Delta)G_X(t, z_1, z_2, ..., z_{n-2}, z_n, z_n) \\ +(\mu_n\Delta)G_X(t, z_1, z_2, ..., z_{n-1}, 1) \\ +o(\Delta). \end{cases} \tag{28}$$

By taking $1 \cdot G_X(t, z_1, z_2, ..., z_n)$ from the right side to the left, dividing both sides by $\Delta$ and taking $\Delta \to 0$ you get the following result:

$$\frac{d\,G_X}{dt}(t, z_1, z_2, ..., z_n) = \begin{cases} (\lambda(z_1 - 1) - \mu)G_X(t, z_1, z_2, ..., z_n) \\ +\mu_1 G_X(t, z_2, z_2, ..., z_n) \\ +\mu_2 G_X(t, z_1, z_3, z_3, z_4, ..., z_n) \\ +... + +\mu_{n-1} G_X(t, z_1, z_2, ..., z_{n-2}, z_n, z_n) \\ +\mu_n G_X(t, z_1, z_2, ..., z_{n-1}, 1). \end{cases} \tag{29}$$

Note that in the last line of this equation $G_X(t, z_1, z_2, ..., z_{n-1}, 1)$ appears which is equal to $G_X(t, z_1, z_2, ..., z_{n-1})$. So we have an equivalent generating function without the last variable $z_n$. In steady state the left side of this equation is zero, so in steady state we see a way to determine probability generating functions recursively.

### 3.5.2 Probability Generating Function of Y(s)

The probability generating function $Y(s)$ is defined as follows:

$$G_Y(s, z_1, z_2, ..., z_n) = \mathbb{E}[z_1^{Y_1(s)} \cdot z_2^{Y_2(s)} \cdot ... \cdot z_n^{Y_n(s)}]. \tag{30}$$

Now we are going to look at the probability generating function after the $(s+1)^{th}$ Poissonian event by conditioning on the number of particles after the $s^{th}$ Poissonian event. According to the Markovian dynamics in (14) you get the following:

$$\mathbb{E}[\prod_{k=1}^{n} z_k^{Y_k(s+1)}] = \mathbb{E}[\mathbb{E}[\prod_{k=1}^{n} z_k^{Y_k(s+1)} | Y(s)]] = \begin{cases} \frac{\lambda}{\lambda+\mu}\mathbb{E}[z_1 \prod_{k=1}^{n} z_k^{Y_k(s)}] \\ +\frac{\mu_1}{\lambda+\mu}\mathbb{E}[z_2^{Y_1(s)} \prod_{k\neq 1}^{n} z_k^{Y_k(s)}] \\ +\frac{\mu_2}{\lambda+\mu}\mathbb{E}[z_3^{Y_2(s)} \prod_{k\neq 2}^{n} z_k^{Y_k(s)}] \\ +...+ \\ +\frac{\mu_{n-1}}{\lambda+\mu}\mathbb{E}[z_n^{Y_{n-1}(s)} \prod_{k\neq(n-1)}^{n} z_k^{Y_k(s)}] \\ +\frac{\mu_n}{\lambda+\mu}\mathbb{E}[\prod_{k\neq n}^{n} z_k^{Y_k(s)}]. \end{cases} \tag{31}$$

Using the notation of probability generating function in (30) gives:

$$G_Y(s+1, z_1, z_2, ..., z_n) - G_Y(s, z_1, z_2, ..., z_n) = \begin{cases} \frac{\lambda(z_1-1)-\mu}{\lambda+\mu} G_Y(s, z_1, z_2, ..., z_n) \\ +\frac{\mu_1}{\lambda+\mu} G_Y(s, z_2, z_2, z_3, ..., z_n) \\ +\frac{\mu_2}{\lambda+\mu} G_Y(s, z_1, z_3, z_3, z_4, ..., z_n) \\ +...+ \\ +\frac{\mu_{n-1}}{\lambda+\mu} G_Y(s, z_1, ..., z_{n-2}, z_n, z_n) \\ +\frac{\mu_n}{\lambda+\mu} G_Y(s, z_1, ..., z_{n-1}, 1). \end{cases} \tag{32}$$

### 3.5.3 Probability Generating Function in steady state

In steady state the probability generating function is time-homogeneous: $G_X(t, z_1, ..., z_n) \equiv G_X(z_1, ..., z_n)$ and $\frac{d\,G_X}{dt}(t, z_1, ..., z_n) = 0$. Substituting this in (29) gives:

$$(\lambda(1 - z_1) + \mu)G_X(z_1, z_2, ..., z_n) = \begin{cases} \mu_1 G_X(z_2, z_2, z_3, ..., z_n) \\ +\mu_2 G_X(z_1, z_3, z_3, z_4, ..., z_n) \\ +...+ \\ +\mu_{n-1} G_X(z_1, ..., z_{n-2}, z_n, z_n) \\ +\mu_n G_X(z_1, ..., z_{n-1}, 1). \end{cases} \tag{33}$$

Also holds in steady state: $G_Y(s, z_1, ..., z_n) \equiv G_Y(z_1, ..., z_n)$ and $G_Y(s+1, z_1, z_2, ..., z_n) - G_Y(s, z_1, z_2, ..., z_n) = 0$. Substituting this in (32) gives:

$$(\lambda(1-z_1)+\mu)G_Y(z_1, z_2, ..., z_n) = \begin{cases} \mu_1 G_Y(z_2, z_2, z_3, ..., z_n) \\ +\mu_2 G_Y(z_1, z_3, z_3, z_4, ..., z_n) \\ +...+ \\ +\mu_{n-1} G_Y(z_1, ..., z_{n-2}, z_n, z_n) \\ +\mu_n G_Y(z_1, ..., z_{n-1}, 1). \end{cases} \tag{34}$$

Note that (33) and (34) have the same structure. So in steady state the distributions of X(t) and Y(s) are the same. This is called the PASTA phenomenon. "The PASTA phenomenon is a central concept in queueing theory, which implies that arriving customers find, on average, the same workload in the queueing system as an outside observer looking at the system at an arbitrary point in time." [3] The PASTA phenomenon holds for queueing systems with Poisson arrivals, but not for queueing systems in general.
Another thing what you can observe from (33) and (34) is that an observation of the first m queues in an ASIP with n queues is the same as an observation of an ASIP with m queues. This can be seen in (33) and (34) by taking $z_{m+1}, ..., z_n$ equal to 1. Then at both sides of the equations you have the sum $\sum_{i=m+1}^{n} \mu_i G(z_1, ..., z_m, 1, ..., 1)$. Subtract this term on both sides of the equations and you get the same equations (33) and (34) where n is replaced by m.

### 3.5.4 Steady state analysis for ASIP with one queue

Consider the ASIP consisting of only one queue. According to (33) you get in steady state the following result:

$$(\lambda(1-z_1)+\mu)G_X(z_1) = \mu_1 \cdot G_X(1). \tag{35}$$

Note that $G_X(1) = \mathbb{E}[1^X] = 1$. Then:

$$G_X(z_1) = \frac{\mu_1}{\lambda(1-z_1)+\mu_1} = \frac{\frac{\mu_1}{\lambda+\mu_1}}{1-(1-\frac{\mu_1}{\lambda+\mu_1})z_1}. \tag{36}$$

This can be recognized as a geometric distribution with parameter $\frac{\mu_1}{\lambda+\mu_1}$ which means that:

$$\mathbb{P}(X_1 = j) = \mathbb{P}(Y_1 = j) = (1-\frac{\mu_1}{\lambda+\mu_1})^j \cdot \frac{\mu_1}{\lambda+\mu_1}. \tag{37}$$

That the number of particles in the queue in steady state is geometric distributed sounds logical. Namely the probability that the gate (=gate 1) opens as next Poissonian event is $\frac{\mu_1}{\lambda+\mu_1}$ and the probability that a particle arrives as next Poissonian event is $\frac{\lambda}{\lambda+\mu_1} = 1 - \frac{\mu_1}{\lambda+\mu_1}$. So the probability of having j particles in the queue is the probability that j particles arrive first ($= (1-\frac{\mu_1}{\lambda+\mu_1})^j$) and then multiplied by the probability that the gate opens ($= \frac{\mu_1}{\lambda+\mu_1}$).

### 3.5.5 Steady state analysis for ASIP with two queues

By using (33) you get the following formula for $G_X(z_1, z_2)$:

$$G_X(z_1, z_2) = \frac{\mu_1^2\mu_2}{(\lambda(1-z_2)+\mu_2)(\lambda(1-z_2)+\mu_1)(\lambda(1-z_1)+\mu_1+\mu_2)} + \frac{\mu_1\mu_2}{(\lambda(1-z_1)+\mu_1)(\lambda(1-z_1)+\mu_1+\mu_2)}. \tag{38}$$

Even the probability generating function of an ASIP with only two queues becomes rather complex. The solution's complexity increases drastically when an ASIP has more queues, which makes finding probability generating functions for numbers of particles in ASIP's with a lot of queues practically impossible.

## 3.6 Asymptotic analysis of ASIP

This section is based on the journal article "Limit laws for the asymmetric inclusion process" [4]

In this section we do asymptotic analysis of the following observables of the ASIP: Traversal time, Overall load, Busy period, First occupied queue and Draining time. Also the homogeneous system will be analyzed.

### 3.6.1 Traversal time

As we saw in Section 3.2 the traversal time T is sum of independent random variables $T_i$ with $i = 1, ..., n$. $T_i$ is the waiting time in the $i^{th}$ queue which is exponentially distributed with parameter $\mu_i$. In order the analyze the mean traversal time we want to find the Laplace Transform of T:

$$\mathbb{E}[e^{-\Theta T}] = \mathbb{E}[e^{-\Theta T_1} \cdot e^{-\Theta T_2} \cdot ... \cdot e^{-\Theta T_n}] = \mathbb{E}[\prod_{k=1}^{n} e^{-\Theta T_k}] = \prod_{k=1}^{n} \mathbb{E}[e^{-\Theta T_k}]. \quad (39)$$

The expectation of the product is equal to the product of expectations because the waiting times in the queues are independent of each other. Waiting time $T_k$ is exponentially distributed with parameter $\mu_k$ which has the following Laplace Transform:

$$\mathbb{E}[e^{-\Theta T_k}] = \int_0^\infty e^{-\Theta t} \cdot \mu_k e^{-\mu_k t}\, dt = \int_0^\infty \mu_k e^{-(\mu_k + \Theta)t}\, dt = \frac{\mu_k}{\mu_k + \Theta}. \quad (40)$$

The Laplace Transform of the traversal time is then:

$$\mathbb{E}[e^{-\Theta T}] = \prod_{k=1}^{n} \frac{\mu_k}{\mu_k + \Theta}. \quad (41)$$

### 3.6.2 Overall load

The Overall load is defined as the total number of particles in the system. In section 3.4.3 we derived the ASIP version of Little's law:

$$\mathbb{E}[L] = \sum_{k=1}^{n} \frac{\lambda}{\mu_k} = \lambda \sum_{k=1}^{n} \frac{1}{\mu_k} = \lambda \mathbb{E}[T]. \quad (42)$$

In order to analyze the overall load further, denote $X_{(k)}(t)$ as the number of particles in the first k queues in steady state at time t. Now we want to find the number of particles in the first k queues in steady state at time $t + \Delta$ where $\Delta$ is very small. According to the Markovian dynamics in (13) you get the following:

$$X_{(k)}(t+\Delta) = \begin{cases} X_{(k)}(t) & \text{w.p.} \quad 1 - (\lambda + \mu_k)\Delta + o(\Delta) \\ X_{(k)}(t) + 1 & \text{w.p.} \quad \lambda\Delta + o(\Delta) \\ X_{(k-1)}(t) & \text{w.p.} \quad \mu_k \Delta + o(\Delta) \end{cases}. \quad (43)$$

Here the first line is the situation when nothing happens, the second line is the situation when a particle enters the first gate and the last line is the situation when gate k opens such that all particles in the $k^{th}$ leave the number of particles in the first k queues.
Now define $G_{X_{(k)}}(t, z)$ as the probability generating functions of the number of particles in the first k queues at time t:

$$G_{X_{(k)}}(t, z) = \mathbb{E}[z^{X_{(k)}(t)}] \quad (44)$$

Note that $G_{X_{(k)}}(t, z) = G_X(t, z, ..., z, 1, ..., 1)$ where $z_1 = ... = z_k = z$ and $z_{k+1} = ... = z_n = 1$.
Now we are going to look at the probability generating function at time $t + \Delta$ by conditioning

on the number of particles at time t. According to the Markovian dynamics in (13) you get the following:

$$G_{X_{(k)}}(t+\Delta,z)=\begin{cases}(1-(\lambda+\mu)\Delta)G_{X_{(k)}}(t,z)\\+(\lambda\Delta)zG_{X_{(k)}}(t,z)\\+(\mu_1\Delta)G_{X_{(k)}}(t,z)\\+...+\\+(\mu_{k-1}\Delta)G_{X_{(k)}}(t,z)\\+(\mu_k\Delta)G_{X_{(k-1)}}(t,z)\\+(\mu_{k+1}\Delta)G_{X_{(k)}}(t,z)\\+...+\\+(\mu_n\Delta)G_{X_{(k)}}(t,z)\end{cases}. \tag{45}$$

By taking $G_{X_{(k)}}(t,z)$ to the left and dividing both sides by $\Delta$ you get the following:

$$\frac{G_{X_{(k)}}(t+\Delta,z)-G_{X_{(k)}}(t,z)}{\Delta}=(\lambda(z-1)-\mu_k)G_{X_{(k)}}(t,z)+\mu_k G_{X_{(k-1)}}(t,z)+o(\Delta). \tag{46}$$

Taking $\Delta\to 0$ gives the following:

$$\frac{d\,G_{X_{(k)}}}{dt}(t,z)=(\lambda(z-1)-\mu_k)G_{X_{(k)}}(t,z)+\mu_k G_{X_{(k-1)}}(t,z). \tag{47}$$

In steady state the probability density function is time homogeneous. This yields that $G_{X_{(k)}}(t,z)=G_{X_{(k)}}(z)$ and $\frac{d\,G_{X_{(k)}}}{dt}(t,z)=0$. Equation (3.6.2) results then in:

$$G_{X_{(k)}}(z)=\frac{\mu_k}{\mu_k+\lambda(1-z)}G_{X_{(k-1)}}(z). \tag{48}$$

By doing this step recursively gives:

$$G_{X_{(k)}}(z)=\frac{\mu_k}{\mu_k+\lambda(1-z)}\cdot...\cdot\frac{\mu_2}{\mu_2+\lambda(1-z)}\cdot G_{X_{(1)}}(z). \tag{49}$$

Note that $G_{X_{(1)}}(z)$ is equal to $G_X(z_1)$ in the situation of one ASIP with only one queue. In equation (36) we derived that $G_X(z_1)=\frac{\mu_1}{\lambda(1-z_1)+\mu_1}$. This gives the following result:

$$G_{X_{(k)}}(z)=\frac{\mu_k}{\mu_k+\lambda(1-z)}\cdot...\cdot\frac{\mu_1}{\mu_1+\lambda(1-z)}=\prod_{i=1}^{k}\frac{\mu_i}{\mu_i+\lambda(1-z)}=\prod_{i=1}^{k}\frac{\frac{\mu_i}{\lambda+\mu_i}}{1-(1-\frac{\mu_i}{\lambda+\mu_i})z}. \tag{50}$$

Note that the probability generating function of the number of particles in the first k queues is a product of k geometric distributions on non-negative integers with parameter $\frac{\mu_i}{\lambda+\mu_i}$. This results that the total number of particles in the first k queues is a sum of k geometric distributions with parameter $\frac{\mu_i}{\lambda+\mu_i}$ and the total number of particles in the system is then:

$$L=\sum_{i=1}^{n}G_i \qquad \text{with} \quad G_i\sim Geo(\frac{\mu_i}{\lambda+\mu_i}). \tag{51}$$

So the the number of particles in the first k queues behaves the same as the total number of particles in k independent ASIP's with one queue where the arrival intensity is $\lambda$ and the gate opening of the $i^{th}$ ASIP is exponentially distributed with parameter $\mu_i$.
A geometric distribution with parameter p has mean $\frac{1-p}{p}$ and variance $\frac{1-p}{p^2}$. The mean number of particles in the first k queues is:

$$\mathbb{E}[X_{(k)}(t)]=\sum_{i=1}^{k}\frac{1-\frac{\mu_i}{\lambda+\mu_i}}{\frac{\mu_i}{\lambda+\mu_i}}=\sum_{i=1}^{k}\frac{\frac{\lambda}{\lambda+\mu_i}}{\frac{\mu_i}{\lambda+\mu_i}}=\sum_{i=1}^{k}\frac{\lambda}{\mu_i}=\lambda\cdot(\frac{1}{\mu_1}+...+\frac{1}{\mu_k}). \tag{52}$$

Note that the ASIP's version of Little's law holds for the first k queues of an ASIP. This sounds logic because what happens in queue $k+1$ up to and including queue n does not influence the number of particles in the first k queues because the ASIP is asymmetric. So you might feel in advance that Little's law still holds for the first k queues.
The variance of the number of particles in the first k queues is:

$$Var(X_{(k)}(t)) = \sum_{i=1}^{k} \frac{1-\frac{\mu_i}{\lambda+\mu_i}}{(\frac{\mu_i}{\lambda+\mu_i})^2} = \sum_{i=1}^{k} \frac{\frac{\lambda}{\lambda+\mu_i}}{\frac{\mu_i^2}{(\lambda+\mu_i)^2}} = \sum_{i=1}^{k} \lambda \cdot \frac{\lambda+\mu_i}{\mu_i^2} = \lambda \cdot (\frac{\lambda+\mu_1}{\mu_1^2} + ... + \frac{\lambda+\mu_k}{\mu_k^2}). \quad (53)$$

Finally note that changing the order of the elements in the product of the probability generating function of the number of particles in the first k queues won't change the result. This means that changing the order of the first k queues in the ASIP won't change the dynamics of the number of particles in the first k queues. So the ASIP's load is invariant to gate permutations.

Now we can find with equation (51) the probability that the system is empty:

$$\mathbb{P}(L=0) = \mathbb{P}(G_1 = 0, ..., G_n = 0) = \prod_{i=1}^{n} \mathbb{P}(G_i = 0) = \prod_{i=1}^{n} \frac{\mu_i}{\lambda+\mu_i} \quad (54)$$

Note that for this probability the ratio between the $\mu_i$'s and $\lambda$ is important. If $\lambda$ is proportional larger than the $\mu_i$'s then the probability of having an empty system is quite small. On the other hand if $\lambda$ is proportional smaller than the $\mu_i$'s then the probability of having an empty system is larger. Note that if the number of queues in an ASIP increases the probability of having an empty system will decrease.

### 3.6.3 Busy period

Busy period $B$ is defined as the random variable of the duration where the system is continuously non-empty. So this is the time between an instant that a particle enters an empty system and the first time the system becomes empty again.
Suppose a particle enters an empty system and has traversal time T. Then two cases can happen: (i) this particle leaves the system before another particles enters the system; (ii) another particle enters the system before this particle leaves the system.
Denote the time between the arrivals of the two particles by $\Delta_0$. Because the arrivals of the system is Poisson distributed with parameter $\lambda$, then $\Delta_0$ is exponentially distributed with parameter $\lambda$. Note that the traversal time T and the time between two arrivals $\Delta_0$ are independent of each other.
In the first case where $T < \Delta_0$ the busy period is equal to the traversal time of the first particle: $B = T$. In the second case where $T \geq \Delta_0$ the busy period is equal to the time between the two arrivals plus busy period $B'$ where $B'$ is exactly an independent identically distributed copy of busy period $B$. This leads to the following formula for $B$:

$$B = \begin{cases} T & \text{if} \quad T < \Delta_0 \\ \Delta_0 + B' & \text{if} \quad T \geq \Delta_0 \end{cases}. \quad (55)$$

Note that $B'$ can be regarded as beginning with a particle that enters an empty system again and has traversal time $T'$. This busy period $B'$ does not depend on traversal time T because $\Delta_0 + T' \geq T$, namely all particles in a queue go in a batch to the next queue, so a particle that enters the system later than the first particle will not leave the system earlier than the first particle.
The time period during the ASIP's process can be seen as an alternating sequence of empty and non-empty moments. The empty moments are independent identically distributed copies of $\Delta_0$. The non-empty moments are independent identically distributed copies of busy period B. When the time period during ASIP's process tend to infinity the fraction of time that the system is empty

is $\frac{\mathbb{E}[\Delta_0]}{\mathbb{E}[\Delta_0]+\mathbb{E}[B]}$. Because the time period tends to infinity this fraction is equal to the probability that the system is empty in steady state given in equation (54):

$$\mathbb{P}(L=0) = \prod_{i=1}^{n} \frac{\mu_i}{\lambda+\mu_i} = \frac{\mathbb{E}[\Delta_0]}{\mathbb{E}[\Delta_0]+\mathbb{E}[B]} \tag{56}$$

The expectation of the empty time periods is $\frac{1}{\lambda}$. Filling this in and rearranging the terms gives the following result for the expected busy period:

$$\mathbb{E}[B] = \frac{1}{\lambda}(\prod_{k=1}^{n}[1+\frac{\lambda}{\mu_k}]-1) \tag{57}$$

### 3.6.4 First occupied queue

Let $I$ be the index of the first occupied queue. Consider the ASIP in steady state. Then:

$$I = \begin{cases} \min\{k|X_k > 0\} & \text{if system is non-empty} \\ \infty & \text{if system is empty} \end{cases} \tag{58}$$

The probability that the first queue is occupied is $\mathbb{P}(X_1 > 0)$. When the $k^{th}$ queue is the first occupied queue then the first $k-1$ queues are empty and $k^{th}$ queue is non-empty. Then:

$$\mathbb{P}(I=k) = \begin{cases} \mathbb{P}(X_1 > 0) & \text{if } k=1 \\ \mathbb{P}(X_1=0,...,X_{k-1}=0) - \mathbb{P}(X_1=0,...,X_k=0) & \text{if } 1 < k \leq n \\ \mathbb{P}(X_1=0,...,X_n=0) = \mathbb{P}(L=0) & \text{if } k=\infty \end{cases} \tag{59}$$

Filling in the results of equations (37) and (54) gives:

$$\mathbb{P}(I=k) = \begin{cases} 1-\mathbb{P}(X_1=0) = 1-\frac{\mu_1}{\lambda+\mu_1} = \frac{\lambda}{\lambda+\mu_1} & \text{if } k=1 \\ \prod_{i=1}^{k-1}\frac{\mu_i}{\mu_i+\lambda} - \prod_{i=1}^{k}\frac{\mu_i}{\mu_i+\lambda} = \frac{\lambda}{\mu_k}\prod_{i=1}^{k}\frac{\mu_i}{\mu_i+\lambda} & \text{if } 1 < k \leq n \\ \prod_{i=1}^{n}\frac{\mu_i}{\lambda+\mu_i} & \text{if } k=\infty \end{cases} \tag{60}$$

Note that the probability that the first occupied queue is k decreases when k increases.

### 3.6.5 Draining time

Consider the ASIP in steady state. At a time instance particles that enter the system will be blocked. The time between the instance of blocking the incoming particles and the instance that all the particles have left the system is called draining time $D$. At the moment where entering particles are blocked the draining time is zero when the system is already empty. The draining time is equal to the traversal time through queues $k$ to $n$ when the first occupied queue is $k$ ($I=k$). This traversal time can be derived from (10) is $E_k + ... + E_n$ where $E_i \sim Exp(\mu_i)$. This gives the following result for the draining time:

$$D = \begin{cases} \sum_{i=k}^{n} E_i & \text{if } I=k \text{ for } 1 \leq k \leq n \\ 0 & \text{if } I=\infty \end{cases} \tag{61}$$

The expected draining time according to the law of total chance is then:

$$\mathbb{E}[D] = \sum_{k=1}^{n}\mathbb{P}(I=k)\mathbb{E}[\text{traversal time}|I=k] = \sum_{k=1}^{n}[\frac{\lambda}{\mu_k}(\prod_{i=1}^{k}\frac{\mu_i}{\mu_i+\lambda})(\sum_{i=k}^{n}\frac{1}{\mu_i})] \tag{62}$$

Note that the first part is the probability that the first occupied queue is the $k^{th}$ queue as given in equation (60) and the second part is the expected traversal time through queues $k$ to $n$.

## 3.7 ASIP where particles may arrive at all queues

This section is based on the article "An ASIP model with general gate opening intervals." [5]

### 3.7.1 Model description

The model description is the same as the ASIP model description in Section 2 with some extensions. If gate i opens at time t, then the following gate that opens is gate j with probability $p_{i,j}$ according to a Markov Process. The time between gate i and gate j opens is a random variable $O_{i,j}$. Assuming that the Markov chain of gate openings is irreducible and aperiodic, the steady state distribution can be determined and will be denoted by $\pi_i$ with $i = 1, ..., n$. In the time period $O_{i,j}$ particles may arrive at all queues. The number of arrivals in each queue is independent of each other but may depend on which gate is gate i and which gate is gate j. The number of arrivals in each queue is captured by the following probability generating function:

$$A_{i,j}(z_1, ..., z_n) = \mathbb{E}[z_1^{X_1(i,j)} \cdot z_2^{X_2(i,j)} \cdot ... \cdot z_n^{X_n(i,j)}]. \tag{63}$$

Here $X_k(i,j)$ is the number of particles that arrive in the $k^{th}$ queue in period $O_{i,j}$.
We also define the probability generating function of the cumulative number of arrivals in the first k queues during period $O_{i,j}$ by:

$$A_{i,j,k}(z) := A_{i,j}(z, .., z, 1, ..., 1) = \mathbb{E}[z^{X_1(i,j)} \cdot ... \cdot z^{X_k(i,j)} \cdot 1^{X_{k+1}(i,j)} \cdot ... \cdot 1^{X_n(i,j)}]. \tag{64}$$

### 3.7.2 Analysis

Define $X_{(k)}(m) := X_1 + ... + X_k$ as the number of particles in the first k queues right after the $m^{th}$ gate opening and define M as the gate that has just opened. By conditioning on which gate has just opened the probability generating function of the number of particles in the first k queues is:

$$G_{k,i}(z,m) := \mathbb{E}[z^{X_{(k)}(m)} \cdot I(M = i)] = \mathbb{P}(M = i) \cdot \mathbb{E}[z^{X_{(k)}(m)} | M = i]. \tag{65}$$

Here $I$ means the indicator function. The simplest case $G_{1,1}(z,m)$ gives the following result:

$$G_{1,1}(z,m) = \mathbb{P}(M = 1) \cdot \mathbb{E}[z^{X_{(1)}(m)} | M = 1] = \mathbb{P}(M = 1) \cdot \mathbb{E}[z^0 | M = 1] = \mathbb{P}(M = 1) = \pi_1. \tag{66}$$

This is right because the first queue is immediately empty right after the first gate has been opened.
By the fact that particles in the $(k-1)^{th}$ queue only move to the $k^{th}$ queue, there is a possibility to write $G_{k,i}(z,m)$ in terms of $G_{k-1,j}(z,m)$ and after repeating this trick $G_{k,i}(z,m)$ can be written in terms of $G_{1,j}(z,m)$. By conditioning which was the previous gate that opened we obtain for $G_{1,j}(z,m)$:

$$G_{1,j}(z,m) = \sum_{i=1}^{n} p_{i,j} \cdot G_{1,i}(z, m-1) \cdot A_{i,j,1}(z) \qquad \text{for } j \neq 1. \tag{67}$$

Here $A_{i,j,1}(z)$ is the generating function of number of particles arriving at the first queue. Rewriting equation (67) gives:

$$G_{1,j}(z,m) = \sum_{i=2}^{n} p_{i,j} \cdot G_{1,i}(z, m-1) \cdot A_{i,j,1}(z) + p_{1,j} \cdot G_{1,1}(z, m-1) \cdot A_{1,j,1}(z) \qquad \text{for } j \neq 1. \tag{68}$$

In steady state this equation reduces to:

$$G_{1,j}(z) = \sum_{i=2}^{n} p_{i,j} \cdot G_{1,i}(z) \cdot A_{i,j,1}(z) + p_{1,j} \cdot G_{1,1}(z) \cdot A_{1,j,1}(z) \qquad \text{for } j \neq 1. \tag{69}$$

Now introduce the vectors:

$$\bar{G}_1(z) := (G_{1,2}(z), ..., G_{1,n}(z)). \tag{70}$$

$$R_1(z) := (p_{1,2}A_{1,2,1}(z), ..., p_{1,n}A_{1,n,1}(z)). \tag{71}$$

Now we can rewrite equation (69) into this:

$$\bar{G}_1(z) = \bar{G}_1(z)P_1(z) + G_{1,1}(z)R_1(z). \tag{72}$$

Here $P_1(z)$ is a squared matrix of size $n-1$ which has on place $(i,j)$ in the matrix the value $p_{i-1,j-1}A_{i-1,j-1,1}(z)$. Rewriting this equation gives:

$$\bar{G}_1(z) = \frac{G_{1,1}(z)R_1(z)}{I - P_1(z)}. \tag{73}$$

Here $I$ is the identity matrix of size $n-1$. We already know from equation (66) that $G_{1,1}(z) = \pi_1$. So all the terms on the right side of equation (73) are known. So we have determined the functions of $G_{1,1}(z), G_{1,2}, ..., G_{1,n}(z)$.

Now we are going to determine the terms $G_{k,j}(z)$ for $j = 1,..,n$ and $k = 2,...,n$ by expressing $G_{k,j}(z)$ into terms of $G_{k-1,i}(z)$ where $i = 1,...,n$. By doing this step recursively we can determine $G_{k,j}(z)$ in terms of $G_{1,i}$ where $i = 1,..,n$ which we have already determined.

In order to express $G_{k,j}(z)$ in terms of $G_{k-1,i}(z)$ where $i = 1,...,n$ consider two consecutive gate openings in steady state. Gate i opens first and gate j opens second. Now we can determine $G_{k,j}(z)$ by conditioning on all possible gate openings i. Then:

$$G_{k,j}(z) = \begin{cases} \sum_{i=1}^{n} G_{k,i}(z)p_{i,j}A_{i,j,k}(z) & \text{for } j \neq k \\ \sum_{i=1}^{n} G_{k-1,i}(z)p_{i,k}A_{i,k,k-1}(z) & \text{for } j = k. \end{cases} \tag{74}$$

In the case that gate k opens as the $j^{th}$ opening then right thereafter there are no particles in the $k^{th}$ site. So the total number of particles in the first k sites is equal to the total number of particles in the first k-1 sites right after the previous gate opening plus the number of new arrivals in the first k-1 queues in the time period $O_{i,j}$. Rewriting equation (74) gives:

$$G_{k,j}(z) = \sum_{i \neq k} G_{k,i}(z)p_{i,j}A_{i,j,k}(z) + G_{k,k}(z)p_{k,j}A_{k,j,k}(z). \tag{75}$$

Now introduce the vectors:

$$\bar{G}_k(z) := (G_{k,1}(z), ..., G_{k,k-1}(z), G_{k,k+1}(z), ..., G_{k,n}(z)). \tag{76}$$

$$R_k(z) := (p_{k,1}A_{k,1,k}(z), ..., p_{k,k-1}A_{k,k-1,k}(z), p_{k,k+1}A_{k,k+1,k}(z), ..., p_{k,n}A_{k,n,k}(z)). \tag{77}$$

Now we can rewrite equation (75) into this:

$$\bar{G}_k(z) = \bar{G}_k(z)P_k(z) + G_{k,k}(z)R_k(z). \tag{78}$$

Here $P_k(z)$ is a squared matrix of size $n-1$ of elements $p_{i,j}A_{i,j,k}(z)$. Rewriting this equation gives:

$$\bar{G}_k(z) = G_{k,k}(z)R_k(z)(I - P_k(z))^{-1}. \tag{79}$$

Here $I$ is the identity matrix of size $n-1$. With this equation we can determine $G_{k,j}(z)$ for $j = 1,...,n$ in terms of $G_{k,k}(z)$ and in equation (74) we have seen how to determine $G_{k,k}(z)$ in terms of $G_{k-1,i}$ with $i = 1,...,n$. By doing these two steps we can determine $G_{k,j}(z)$ in terms of $G_{1,i}(z)$ where $i = 1,...,n$ which are determined in equation (73). So we can theoretically determine the probability generating function $G_{k,j}$ for all $k,j \in \{1,...,n\}$.

### 3.8 Summary of literature study

In this literature study the number of particles in each queue of the ASIP in steady state has been analyzed. First by determining the Markovian dynamics of the number of particles in each queue of the ASIP and then using these dynamics for determining the probability generating function of the number of particles in the ASIP. Probability generating functions are very useful because summation of merging batches translates into a multiplication of probability generating functions. This probability generating function is useful for doing asymptotic analysis, because the in absolute value smallest pole in the probability generating function determines the behavior of the probability that there are n particles in the ASIP when n becomes large according to Abel's theorem. Other observables of the ASIP which are asymptotically analyzed in this literature study are the traversal time, overall load, busy period, first occupied queue and the draining time.
The ASIP can be generalized by the possibility that a particle entering the system can enter all queues instead of only the first queue. Another generalization is that the period between two consecutive gate openings is generally distributed. For both generalizations it is possible to determine the probability generating function of number of particles in the system and this is again useful for doing asymptotic analysis when the number of particles in the system becomes large.

# 4 Own research

In every queue of the ASIP model there may stay infinitely many particles. But what if you have a limited capacity available? How do you distribute the limited capacity over the queues of an ASIP in order to maximize the number of particles that complete the system? First I describe the ASIP model with capacities.

## 4.1 ASIP model with capacities

The ASIP model with capacities is quite the same as the ASIP model without capacities described in section 2. So the arrival process is a Poisson process with parameter $\lambda$ and the time between two consecutive openings of gate $i$ is exponentially distributed with parameter $\mu_i$. The only difference is that there is a limited capacity $C$. This capacity has to be distributed over the queues in an ASIP. The capacity of the $i^{th}$ queue in the ASIP has capacity $c_i$. Thus, $\sum_{i=1}^{n} c_i = C$. Figure 2 show a visual representation of the ASIP model with capacities.

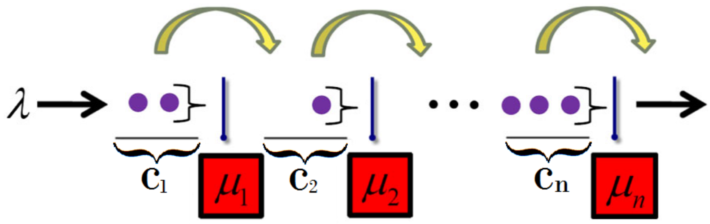


Figure 2: Visual representation of ASIP model with capacities

When a gate opens and a batch of particles merges with the batch of particles in the next queue it may happen that the number of particles of this merged batch is larger than the capacity of that specific queue. In this case I assume that the particles above the capacity will leave the system directly and are considered to be lost. A particle has completed the ASIP when it has passed all the queues and gates of the ASIP. The goal is to distribute capacity C over the queues in such a way that the number of particles that complete the system is maximized. Logically when a capacity of one of the queues is zero no particle will complete the system at all, so $c_i \geq 1$ for $i = \{1, ..., n\}$. This means directly that $C \geq n$ in order to let some particles complete the ASIP.

Define $P_{i,j}$ as the fraction of time in steady state that there are j particles in the $i^{th}$ queue where $i \in \{1, ..., n\}$ and $j \in \{0, ..., C - n + 1\}$. Because $c_i \geq 1$ the maximum capacity a queue can be assigned is $C - n + 1$. The number of particles that complete the system is equal to the number of particles in the $n^{th}$ queue at the moment the $n^{th}$ gate opens. The PASTA property says that for queueing systems with Poisson arrivals the number of particles in a queue at a time of a gate opening is on average equal to the number of particles in this queue at an arbitrary point in time. So according to the PASTA property the number of particles in the $n^{th}$ queue at the moment the $n^{th}$ gate opens is equal to the number of particles in the $n^{th}$ queue at steady state. So the number of particles that complete the ASIP per time unit can be expressed in terms of $P_{n,j}$ where $j = \{0, ..., c_n\}$. This will give the following objective function $F_{objective}$:

$$F_{objective} = \mu_n \cdot (0 \cdot P_{n,0} + 1 \cdot P_{n,1} + ... + c_n \cdot P_{n,c_n}) = \mu_n \cdot \sum_{j=0}^{c_n} j \cdot P_{n,j} \tag{80}$$

$F_{objective}$ is also called the throughput (= the number of particles that complete the ASIP per time unit).

## 4.2 Optimization problem

In the previous section I set up an optimization problem. An optimization problem has four aspects, namely parameters, variables, constraints and an objective function. In this section these four aspects are summarized below.

### 4.2.1 Parameters

The following parameters are used in the ASIP model with capacities:

- $n$, this is the number of queues in the ASIP
- $C$, the total capacity which has to be distributed over the queues.
- $\lambda$, the arrival process of the particles in the ASIP is a Poisson process with intensity $\lambda$.
- $\mu_i$, the time between two consecutive gate openings of the $i^{th}$ gate is exponentially distributed with intensity $\mu_i$.

This problem becomes interesting when total capacity $C$ is greater than number of queues $n$.

### 4.2.2 Variables

The variables in this optimization problem is $c_i$ where $i = \{1, ..., n\}$. Here $c_i$ is the capacity of the $i^{th}$ queue in the ASIP.

### 4.2.3 Constraints

This optimization problem has the following constraints:

- $\lambda, \mu_i \in \mathbb{R}^+$, the intensities of arrivals of particles and gate openings are positive.
- $n, C \in \mathbb{N}^+$, the number of queues and the total capacity are positive integers. The optimization problem becomes interesting when $C \geq n$.
- $\sum_{i=1}^{n} c_i = C$, the sum of the assigned capacities of each queue has to be equal to total capacity C.
- $c_i \in \mathbb{N}$, the assigned capacities to the queues have to be nonnegative integers. The optimization problem becomes interesting when $c_i \geq 1$ for all queues.

### 4.2.4 Objective function

The goal of this optimization problem is maximizing the throughput. So the objective function is:

$$F_{objective} = \mu_n \cdot (0 \cdot P_{n,0} + 1 \cdot P_{n,1} + ... + c_n \cdot P_{n,c_n}) = \mu_n \cdot \sum_{j=0}^{c_n} j \cdot P_{n,j} \tag{81}$$

## 4.3 Determining number of particles in steady state

First I investigate the number of particles in the first queue of the ASIP at steady state. Then I use this result for determining the number of particles in the $i^{th}$ queue for $i = \{2, ..., n\}$ in steady state.

#### 4.3.1 Number of particles in steady state in first queue

The number of particles in the first queue is dependent on two events, namely an arrival of a particle in the system and an opening of the first gate. In the case of an arrival one of two things may happen: the particle enters the first queue so the number of particles in the queue is increased by one or the queue is already full so the arriving particle is considered to be lost. In case the first gate opens all the particles in the first queue move towards the second queue so the first queue becomes empty again. The event of an arrival of a particle happens with intensity $\lambda$ and the event of an opening of the first gate happens with intensity $\mu_1$. The flow diagram of number of particles in the first queue is illustrated in figure 3.

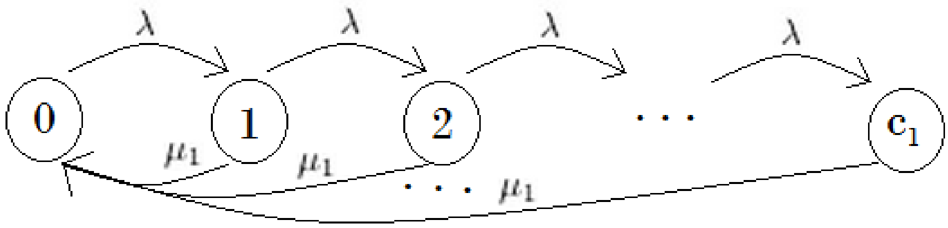


Figure 3: Flow diagram of number of particles in the first queue of the ASIP.

This leads to the following balance equations:

$$\begin{cases} \lambda P_{1,0} = \mu_1 P_{1,1} + \mu_1 P_{1,2} + ... + \mu_1 P_{1,c_1} \\ (\lambda + \mu_1) P_{1,j} = \lambda P_{1,j-1} & \text{for } j \in \{1, ..., c_1 - 1\} \\ \mu_1 P_{1,c_1} = \lambda P_{1,c_1 - 1}. \end{cases} \tag{82}$$

These balance equations have $(c_1 + 1)$ unknown variables, namely $P_{1,j}$ for $j \in \{0, ..., c_1\}$. Together with the constraint that $P_{1,0} + ... + P_{1,c_1} = 1$ we can calculate the distribution of number of particles in the first queue in steady state. Rewriting the first equation of (82) gives:

$$\lambda P_{1,0} = \mu_1 \cdot (P_{1,0} + P_{1,1} + ... + P_{1,c_1}) - \mu_1 P_{1,0}. \tag{83}$$

Using that $P_{1,0} + P_{1,1} + ... + P_{1,c_1} = 1$ gives the following expression for $P_{1,0}$:

$$P_{1,0} = \frac{\mu_1}{\lambda + \mu_1}. \tag{84}$$

For $P_{1,j}$ where $j \in \{1, ..., c_1 - 1\}$ you get the following:

$$P_{1,j} = \frac{\lambda}{\lambda + \mu_1} P_{1,j-1} = ... = (\frac{\lambda}{\lambda + \mu_1})^j \cdot P_{1,0} = (\frac{\lambda}{\lambda + \mu_1})^j \cdot \frac{\mu_1}{\lambda + \mu_1} = \frac{\lambda^j \mu_1}{(\lambda + \mu_1)^{j+1}}. \tag{85}$$

This gives the following expression for $P_{1,c_1}$:

$$P_{1,c_1} = \frac{\lambda}{\mu_1} P_{1,c_1 - 1} = \frac{\lambda}{\mu_1} \cdot \frac{\lambda^{c_1 - 1} \mu_1}{(\lambda + \mu_1)^{c_1}} = (\frac{\lambda}{\lambda + \mu_1})^{c_1}. \tag{86}$$

Summarizing this gives the following result for the distribution of number of particles in the first queue in steady state:

$$P_{1,j} = \begin{cases} \frac{\mu_1}{\lambda + \mu_1} \cdot (\frac{\lambda}{\lambda + \mu_1})^j & \text{for } j = 0, ..., c_1 - 1 \\ (\frac{\lambda}{\lambda + \mu_1})^{c_1} & \text{for } j = c_1 \\ 0 & \text{for } j > c_1. \end{cases} \tag{87}$$

Note that $P_{1,c_1}$ is the fraction of time in steady state that the queue is full. This fraction is relatively high when $\lambda$ is high compared to $\mu_1$ and relatively low when $\lambda$ is low compared to $\mu_1$. An incoming particle is considered to be lost when the first queue is full, so $P_{1,c_1}$ is also the fraction of particles that don't fit in the queue and are considered to be lost. When $c_1$ becomes higher, the probability $P_{1,c_1}$ decreases because $\frac{\lambda}{\lambda+\mu_1} < 1$, so more incoming particles fit in the first queue.
Also note that the result in (87) is a geometric distribution with parameter $\frac{\lambda}{\lambda+\mu_1}$. Suppose that there is no restriction on capacity. Then the probability of having j particles in the queue is $\frac{\mu_1}{\lambda+\mu_1} \cdot (\frac{\lambda}{\lambda+\mu_1})^j$. However the queue has capacity $c_1$. Then the limiting probability of having $c_1$ particles in the queue is equal to the probability of having $c_1$ or more particles in the queue in the situation where there is no restriction on the capacity. Then:

$$\mathbb{P}(X_1 \geq c_1) = 1-\mathbb{P}(X_1 < c_1) = 1-\sum_{n=0}^{c_1-1} \frac{\mu_1}{\lambda+\mu_1}(\frac{\lambda}{\lambda+\mu_1})^n = 1-\frac{\mu_1}{\lambda+\mu_1}\frac{1-(\frac{\lambda}{\lambda+\mu_1})^{c_1}}{1-\frac{\lambda}{\lambda+\mu_1}} = (\frac{\lambda}{\lambda+\mu_1})^{c_1}. \quad (88)$$

So (87) is a geometric distribution with parameter $\frac{\lambda}{\lambda+\mu_1}$ where the probability of having $c_1$ particles in the queue is equal to the probability of having $c_1$ or more particles in the situation where there is no restriction on capacity of queues.

At the moment the first gate opens a batch of particles moves to the second queue. The size of the batch can range from 0 to $c_1$. The intensity of moving batches from the first queue to the second is then $\mu_1$. Thanks to the PASTA property the number of particles in the first queue at the moment the first gate opens is in distribution equal to the number of particles in the first queue at an arbitrary point in time. So in steady state a batch moving to the second queue has size $j$ with probability $P_{1,j}$.

#### 4.3.2 Number of particles in rest of the queues

Knowing the steady state distribution of number of particles in the first queue we know that with intensity $\mu_1$ a batch of size $j$ with probability $P_{1,j}$ arrives at the second queue. In the rest of the queues batches of particles arrive at the queue and batches of particles leave the queue. Consider queue $i$ where $i \in \{2, ..., n\}$. A batch arrives at queue $i$ when gate $i-1$ opens, so batches arrive with intensity $\mu_{i-1}$. The distribution of the size of an arriving batch at queue $i$ is equal to the steady state distribution of number of particles in the previous queue, queue $i-1$. So an arriving batch at queue $i$ has size $j$ with probability $P_{i-1,j}$. The number of particles in queue $i$ can range from 0 to capacity $c_i$. The flow diagram of number of particles in the $i^{th}$ queue is illustrated in figure 4.

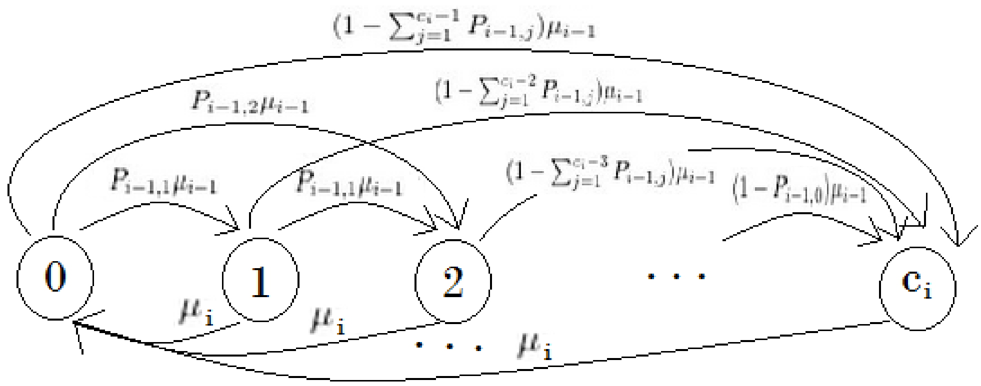


Figure 4: Flow diagram of number of particles in queue $i$ for $i \in \{2, ..., n\}$.

The number of particles in the $i^{th}$ queue increases with the size of a batch with intensity $\mu_{i-1}$. When the old number of particles in the $i^{th}$ queue plus the size of the incoming batch is greater than capacity $c_i$ then some particles are considered to be lost and the new number of particles in the $i^{th}$ queue is equal to capacity $c_i$. With intensity $\mu_i$ gate $i$ opens and queue $i$ becomes empty again.

This leads to the following balance equations:

$$\begin{cases} \mu_{i-1}(P_{i-1,1}+P_{i-1,2}+...+P_{i-1,c_{i-1}})P_{i,0} = \mu_i(P_{i,1}+P_{i,2}+...+P_{i,c_i}) \\ (\mu_{i-1}(P_{i-1,1}+P_{i-1,2}+...+P_{i-1,c_{i-1}})+\mu_i)P_{i,1} = \mu_{i-1}P_{i-1,1}P_{i,0} \\ (\mu_{i-1}(P_{i-1,1}+P_{i-1,2}+...+P_{i-1,c_{i-1}})+\mu_i)P_{i,2} = \mu_{i-1}(P_{i-1,2}P_{i,0}+P_{i-1,1}P_{i,1}) \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ (\mu_{i-1}(P_{i-1,1}+P_{i-1,2}+...+P_{i-1,c_{i-1}})+\mu_i)P_{i,c_i-1} = \mu_{i-1}(P_{i-1,c_i-1}P_{i,0}+...+P_{i-1,1}P_{i,c_i-2}) \\ \mu_iP_{i,c_i} = \mu_{i-1}((1-P_{i-1,0})P_{i,c_i-1}+(1-(P_{i-1,0}+P_{i-1,1}))P_{i,c_i-2}+...+(1-(P_{i-1,0}+...+P_{i-1,c_i-1}))P_{i,0}). \end{cases} \tag{89}$$

By using that $P_{i-1,0}+...+P_{i-1,c_{i-1}}=1$ and $P_{i,0}+...+P_{i,c_i}=1$ these balance equations reduce to:

$$\begin{cases} (\mu_{i-1}(1-P_{i-1,0})+\mu_i)P_{i,0} = \mu_i \\ (\mu_{i-1}(1-P_{i-1,0})+\mu_i)P_{i,j} = \mu_{i-1}\sum_{k=0}^{j-1}P_{i-1,j-k}P_{i,k} \quad \text{for } j\in\{1,...,c_i-1\} \\ \mu_iP_{i,c_i} = \mu_{i-1}\sum_{k=0}^{c_i-1}(1-\sum_{m=0}^{k}P_{i-1,m})P_{i,c_i-1-k}. \end{cases} \tag{90}$$

Rewriting this gives:

$$P_{i,j} = \begin{cases} \frac{\mu_i}{(1-P_{i-1,0})\mu_{i-1}+\mu_i} & \text{for } j=0 \\ \frac{\mu_{i-1}}{(1-P_{i-1,0})\mu_{i-1}+\mu_i}\sum_{k=0}^{j-1}P_{i,k}P_{i-1,j-k} & \text{for } 0<j<c_i \\ \frac{\mu_{i-1}}{\mu_i}\sum_{k=0}^{c_i-1}(1-\sum_{m=0}^{k}P_{i-1,m})P_{i,c_i-1-k} & \text{for } j=c_i \\ 0 & \text{for } j>c_i. \end{cases} \tag{91}$$

By the convolution part in the case $j=c_i$ it is very difficult to write these limiting probabilities explicitly. What we can do is putting these recursive equations of (87) and (91) in a computer and together with certain values of parameters $\lambda$, $\mu_i$ and certain allocation of total capacity C the limiting probabilities can be calculated for each queue. Knowing the limiting probabilities the value of the objective function given in (81) can be calculated.

Suppose now that parameters $\lambda$ and $\mu_i$ for $i\in\{1,...,n\}$ are known. If you calculate for each possible allocation of total capacity C over n queues the limiting probabilities and then the value of the objective function, then an allocation of capacity with the highest value of the objective function is an optimal solution.
Now the questions arises: how many possibilities are there of allocating C capacity over n queues? Here we require that $C\geq n$ and $c_i\geq 1$ for all $i\in\{1,...,n\}$ otherwise the objective function has value zero. This question will be answered in section 2.4.

### 4.3.3 Determining limiting probability of zero particles in each queue

It is possible to write $P_{i,0}$ for $i\in\{1,...,n\}$ explicitly. If we say that $\mu_0=\lambda$, then after some trying I suspect the following:

$$P_{i,0} = \frac{\mu_i}{\frac{1}{\sum_{j=0}^{i-1}\frac{1}{\mu_j}}+\mu_i} \qquad \text{for } i\in\{1,...,n\}. \tag{92}$$

For $i=1$ we have the following:

$$P_{1,0} = \frac{\mu_1}{\frac{1}{\frac{1}{\mu_0}}+\mu_1} = \frac{\mu_1}{\mu_0+\mu_1} = \frac{\mu_1}{\lambda+\mu_1}. \tag{93}$$

This corresponds to equation (84) which means that (92) holds for $i = 1$. Suppose now that (92) holds for some $i$ with $1 \leq i < n$. Now to proof:

$$P_{i+1,0} = \frac{\mu_{i+1}}{\frac{1}{\sum_{j=0}^{i} \frac{1}{\mu_j}} + \mu_{i+1}}. \tag{94}$$

We know from (91) that $P_{i+1,0} = \frac{\mu_{i+1}}{(1-P_{i,0})\mu_i + \mu_{i+1}}$, then:

$$P_{i+1,0} = \frac{\mu_{i+1}}{(1 - \frac{\mu_i}{\frac{1}{\sum_{j=0}^{i-1} \frac{1}{\mu_j}} + \mu_i})\mu_i + \mu_{i+1}}. \tag{95}$$

It remains to show that $(1 - \frac{\mu_i}{\frac{1}{\sum_{j=0}^{i-1} \frac{1}{\mu_j}} + \mu_i})\mu_i$ is equal to $\frac{1}{\sum_{j=0}^{i} \frac{1}{\mu_j}}$:

$$(1 - \frac{\mu_i}{\frac{1}{\sum_{j=0}^{i-1} \frac{1}{\mu_j}} + \mu_i})\mu_i = \frac{\frac{1}{\sum_{j=0}^{i-1} \frac{1}{\mu_j}} \cdot \mu_i}{\frac{1}{\sum_{j=0}^{i-1} \frac{1}{\mu_j}} + \mu_i} = \frac{\mu_i}{1 + \mu_i \sum_{j=0}^{i-1} \frac{1}{\mu_j}} = \frac{1}{\frac{1}{\mu_i} + \sum_{j=0}^{i-1} \frac{1}{\mu_j}} = \frac{1}{\sum_{j=0}^{i} \frac{1}{\mu_j}}. \tag{96}$$

So we have proved by induction that (92) holds for all $i \in \{1, ..., n\}$.

## 4.4 Number of possibilities of allocating C capacity over n queues

The number of possibilities of allocating a capacity $C$ over $n$ queues with the requirement that $c_i \geq 1$ for all $i \in \{1, ..., n\}$ is equal to the number of possibilities of allocating a capacity $C - n$ over $n$ queues without a requirement. You can see the ASIP with $n$ queues as $n$ bins divided by $n - 1$ gates. Now see total capacity C as C balls and you have to throw each ball in a bin. Suppose a ball is displayed as a 0 and a gate is displayed by a 1. Then the number of possibilities of allocating capacity $C$ over $n$ queues with the requirement that $c_i \geq 1$ for all $i \in \{1, ..., n\}$ is equal to the number of different sequences containing $C - n$ zeros and $n - 1$ ones. This is equal to $\binom{(C-n)+(n-1)}{n-1} = \binom{C-1}{n-1}$.

We are interested in finding an allocation of capacity C which maximizes the objective function. In order to save calculation time I want to proof that for most of the allocations of capacity C there is another solution which has minimally the same value of the objective function. This will be done by the following lemma.

**Lemma 2.1** *In the set of allocations of capacity $C$ where $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$ there is an optimal solution.*

*Proof.* Suppose you have an allocation of capacity $C$ over $n$ queues where $c_i > c_j$ for some $i, j$ where $1 \leq i < j \leq n$. Define $\Delta = c_i - c_j$ which is the extra capacity that has queue $i$ more than queue $j$. Then there are two situations: this extra capacity will be used or will not be used.

In the situation where the extra capacity will not be used then this solution has the same throughput as the solution of allocating a total capacity $C - \Delta$ over $n$ queues where queue $i$ has capacity $c_j$ and keeping the rest of the capacities the same.

In the situation where the extra capacity will be used then at some time there are $k$ particles in queue $i$ where $k > c_j$. We call these $k$ particles the initial batch. At the moment gate $i$ opens the initial batch moves to the next queue and eventually merge with another batch of particles. Every time that the initial batch moves to the next queue some particles may considered to be lost in case the next queue is or becomes full. There comes a moment when the remaining particles of the initial batch moves to queue $j$. Because $k > c_j$ minimally $k - c_j$ particles of the initial batch are considered to be lost during the movement of the initial batch from queue $i$ to queue $j$. Lost particles don't complete the ASIP completely so they do not have a positive contribution to the objective function. So again this solution has the same throughput as the solution of allocating a total capacity $C - \Delta$ over $n$ queues where queue $i$ has capacity $c_j$ and keeping the rest of the capacities the same.

So for every solution of allocating capacity C over n queues where not for all $i \in \{2, ..., n\}$ holds that $c_i \geq c_{i-1}$ the throughput is the same as the throughput of a solution of allocating $C - \Delta$ capacity over $n$ queues for some $\Delta \in \mathbb{N}^+$ where $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$. The throughput of this solution is smaller than the throughput of an optimal solution of allocating $C$ capacity over $n$ queues where $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$. This finishes the proof of the lemma.

□

Now we know that in the set of allocations of capacity $C$ where $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$ there is an optimal solution. How many of these solutions are there with the requirement that $c_i \geq 1$ for all $i \in \{1, ..., n\}$? This number of possibilities is equal to the number of possibilities of allocating $C - n$ capacity over $n$ queues where $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$ and $c_i \geq 0$ for all $i \in \{1, ..., n\}$. Determining the possibilities of allocating $C - n$ capacity over $n$ queues can be seen in rounds where in each round you assign 1 extra capacity to the last k queues. In the second round you have to allocate $C - n - k$ capacity over $k$ queues. $k$ can take values between 1 and the minimum of the remaining capacity and the remaining number of queues. Making this more formal gives:

Define $f(a, b)$ as the number of possibilities of allocating $b$ capacity over $a$ queues such that $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$, then:

$$f(a, b) = \sum_{k=1}^{\min\{a,b\}} f(k, b - k). \tag{97}$$

There are some easy situations: allocating $b$ capacity over 1 queue or 0 capacity over $a$ queues or 1 capacity over $a$ queues. There is only one possibility to allocate capacity in all these easy situations. The number of possibilities of allocating $b$ capacity over $a$ queues can be reduced to a sum of number of possibilities of easy situations for all $a, b \geq 1$. So by dynamic programming you can determine f(a,b) for any $a, b \geq 1$.

$f(n, C - n)$ gives the answer on the number of possibilities of allocating $C$ capacity over $n$ queues with the requirement that $c_i \geq c_{i-1}$ for $i \in \{2, ..., n\}$ and $c_i \geq 1$ for $i \in \{1, ..., n\}$. Take $C = 30$ and $n = 10$, then the number of possibilities is $f(10, 20) = 530$. The number of possibilities of allocating 30 capacity over 10 queues without the requirement that $c_i \geq c_{i-1}$ for $i \in \{2, ..., n\}$ is 10015005. This gives an impression that this massively reduces the number of solutions for which we have to determine the value of the objective function in order to find an optimal solution. This will also massively decrease the running time for determining an optimal solution.

## 4.5 Determining throughput of homogeneous ASIP where every queue has one capacity

Suppose you have an ASIP where the total capacity $C$ is equal to $n$ and the capacity is allocated in such a way that every queue has one capacity. In this case I'm going to determine the throughput for a general ASIP first and then look what is means for an homogeneous ASIP.

### 4.5.1 Throughput of general ASIP where every queue has one capacity

Now it holds that $P_{i,1} = 1 - P_{i,0}$ for all $i \in \{1, ..., n\}$ because $c_i = 1$ for all $i \in \{1, ..., n\}$. We already found an expression for $P_{i,0}$ in equation (92). Now the objective function (= throughput) can be rewritten in the following way:

$$F_{objective} = \mu_n \cdot \sum_{j=0}^{c_n} j \cdot P_{n,j} = \mu_n \cdot P_{n,1} = \mu_n \cdot (1 - P_{n,0}) = \mu_n \cdot (1 - \frac{\mu_n}{\frac{1}{\sum_{j=0}^{n-1} \frac{1}{\mu_j}} + \mu_n}). \tag{98}$$

Using that $\mu_0 = \lambda$, the throughput can be further rewritten as follows:

$$F_{objective} = \mu_n \cdot (1 - \frac{\mu_n}{\frac{1}{\frac{1}{\lambda} + \sum_{j=1}^{n-1} \frac{1}{\mu_j}} + \mu_n}) = \mu_n \cdot (1 - \frac{\mu_n \cdot (\frac{1}{\lambda} + \sum_{j=1}^{n-1} \frac{1}{\mu_j})}{1 + \mu_n \cdot (\frac{1}{\lambda} + \sum_{j=1}^{n-1} \frac{1}{\mu_j})}) = \frac{\mu_n}{1 + \mu_n \cdot (\frac{1}{\lambda} + \sum_{j=1}^{n-1} \frac{1}{\mu_j})}. \tag{99}$$

When we divide (99) by $\mu_n$ the throughput can be further rewritten as follows:

$$F_{objective} = \frac{1}{\frac{1}{\mu_n} + \frac{1}{\lambda} + \sum_{j-1}^{n-1} \frac{1}{\mu_j}} = \frac{1}{\frac{1}{\lambda} + \sum_{j=1}^{n} \frac{1}{\mu_j}} = \frac{1}{\sum_{j=0}^{n} \frac{1}{\mu_j}}. \tag{100}$$

From (100) you can see that the throughput only depends on $\lambda$, $n$ and $\sum_{j=1}^{n} \frac{1}{\mu_j}$. This implies that in an ASIP with $n$ queues with $c_i = 1$ for all $i \in \{1, ..., n\}$ every permutation of the gate opening intensities has the same throughput.

The result in (100) can also be explained intuitively by determining the throughput using an integrated circuit. The integrated circuit of an ASIP is as follows:

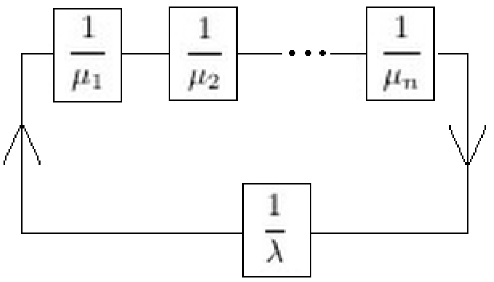


Figure 5: Integrated circuit of ASIP with $n$ queues.

Particles move on this circuit. Define C as the time that a particle does exactly one cycle in the integrated circuit. Then:

$$\mathbb{E}[C] = \frac{1}{\lambda} + \frac{1}{\mu_1} + ... + \frac{1}{\mu_n}. \tag{101}$$

Then the throughput is the average number of particles that come along a certain point in the circuit during the average time a particle does one cycle. So:

$$\text{Throughput} = \frac{\text{number of particles in circuit}}{\mathbb{E}[C]}. \tag{102}$$

We are interested in the throughput in steady state. In order to become in steady state particles are moving through the integrated circuit for an infinite amount of time. Because every queue has one capacity an arriving particle will be lost when it arrives at a queue that is already occupied by another particle. So there comes a moment that there is only one particle left in the integrated circuit. So the throughput in steady state is equal to $\frac{1}{\mathbb{E}[C]} = \frac{1}{\frac{1}{\lambda} + \frac{1}{\mu_1} + ... + \frac{1}{\mu_n}}$ which is exactly the same as in (100).

### 4.5.2 Throughput of homogeneous ASIP where every queue has one capacity

In an homogeneous ASIP the gate opening intensities are the same for all gates, say $\mu_i = \mu$ for all $i \in \{1, ..., n\}$. Then we can rewrite (100) as follows:

$$F_{objective} = \frac{1}{\sum_{j=0}^{n} \frac{1}{\mu_j}} = \frac{1}{\frac{1}{\lambda} + \sum_{j=1}^{n} \frac{1}{\mu}} = \frac{1}{\frac{1}{\lambda} + n \cdot \frac{1}{\mu}}. \tag{103}$$

In this result of the throughput we see that the throughput becomes smaller when $n$ becomes larger. In fact, the throughput will go to zero when the number of queues tends to infinity. This sounds logic because of the following: When a particle want to complete the ASIP it has to gone through all queues. When a particle arrives at queue $i$ the probability that queue $i$ is not occupied is thanks to the PASTA property equal to $P_{i,0}$. So the probability that a particle completes the ASIP is equal to the probability that all queues are not occupied when a particle arrives at each queue which is $\prod_{i=1}^{n} P_{i,0}$. This probability will go to zero when n tends to infinity. Now we have seen another way of deriving (103), namely from equation (92) we can derive that in the homogeneous case the following holds:

$$P_{i,0} = \frac{\mu_i}{\frac{1}{\sum_{j=0}^{i-1} \frac{1}{\mu_j}} + \mu_i} = \frac{\mu}{\frac{1}{\frac{1}{\lambda} + (i-1)\frac{1}{\mu}} + \mu} = \frac{\frac{\mu}{\lambda} + (i-1)}{1 + \frac{\mu}{\lambda} + (i-1)} = 1 - \frac{1}{\frac{\mu}{\lambda} + i}. \tag{104}$$

The probability of completing the ASIP is then: $\prod_{i=1}^{n} P_{i,0} = \prod_{i=1}^{n} 1 - \frac{1}{\frac{\mu}{\lambda}+i} = \frac{\frac{\mu}{\lambda}}{n+\frac{\mu}{\lambda}}$. So the probability that a particle completes the ASIP is $\frac{\frac{\mu}{\lambda}}{n+\frac{\mu}{\lambda}}$. Do we multiply this probability with arrival intensity $\lambda$ you get the expression for the throughput which is exactly the same as (103).

From (103) we can also see the relation between the arrival intensity of particles in the system $\lambda$ and gate opening intensity $\mu$. When $\lambda = \mu$ then the throughput will be $\frac{\mu}{1+n}$. When $\lambda > \mu$ then the throughput will increase compared to the situation when $\lambda = \mu$, but it will have an upper-bound, because the throughput goes to $\frac{\mu}{n}$ when $\lambda$ tends to infinity, so in this case $\frac{\mu}{n+1} < F_{objective} < \frac{\mu}{n}$. When $\lambda < \mu$ then the throughput will be smaller than $\frac{\mu}{n+1}$. When $\lambda$ tends to zero the throughput will go to zero. This sounds logic because it can't be that particles complete the ASIP successfully when the intensity of incoming particles is zero, what makes the throughput zero.
So it can be concluded that in an homogeneous ASIP with n queues where every queues has one capacity the throughput will always be smaller than $\frac{\mu}{n}$ no matter what the intensity of incoming particles is.

## 4.6 The optimal solution of homogeneous ASIP with 20 queues

Putting the equations (87) and (91) in a java program, computing the value of the objective function for every possible allocation of capacity C where $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$ and $c_i > 0$ for all $i \in \{1, ..., n\}$ and returning the solution where the objective function is maximized is a method for finding an optimal solution. In order to give an impression of the optimal solution of allocating capacity over queues I studied the homogeneous ASIP with 20 queues where $\mu = \mu_i$ for all $i \in \{1, ..., n\}$. Total capacity $C$ will range from 20 to 100.

### 4.6.1 homogeneous ASIP with 20 queues with $\lambda = \mu = 1$

Optimal solutions with corresponding throughput for an homogeneous ASIP with $\lambda = \mu = 1$ are shown in tables 1, 2 and 3.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 20 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 0.0476 |
| 21 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 | 0.0499 |
| 22 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 | 0.0522 |
| 23 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 | 0.0544 |
| 24 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 | 0.0567 |
| 25 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 | 0.0590 |
| 26 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 | 0.0612 |
| 27 | 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 | 0.0635 |
| 28 | 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 | 0.0658 |
| 29 | 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 | 0.0680 |
| 30 | 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 | 0.0703 |
| 31 | 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 | 0.0726 |
| 32 | 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0748 |
| 33 | 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0771 |
| 34 | 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0794 |
| 35 | 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0816 |
| 36 | 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0839 |
| 37 | 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0862 |
| 38 | 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0884 |
| 39 | 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0907 |
| 40 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0930 |
| 41 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 | 0.0952 |
| 42 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 | 0.0975 |
| 43 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 | 0.0998 |
| 44 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 | 0.1020 |
| 45 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 | 0.1043 |
| 46 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 | 0.1065 |
| 47 | 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 | 0.1088 |
| 48 | 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 | 0.1110 |
| 49 | 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 | 0.1133 |
| 50 | 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 | 0.1155 |
| 51 | 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 | 0.1178 |
| 52 | 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 | 0.1200 |
| 53 | 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 | 0.1223 |
| 54 | 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 | 0.1245 |
| 55 | 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 | 0.1267 |

Table 1: The optimal solution of allocating a total capacity C over a homogeneous ASIP with 20 queues where arrivals of particles and gate openings have capacity 1 and the total capacity ranges from 20 to 55.

In these tables we can see that allocating total capacity of 55 or less over the queues is quite deterministic, namely distribute $C - (C \bmod n)$ capacity equally over the queues and then assign one extra capacity to each of the $C \bmod n$ rightmost queues (= queues $j$ where $j \in \{n - (C \bmod n) + 1, ..., n\}$).

In the situation where total capacity $C$ is larger than 55 the optimal solution of allocating total capacity $C$ is not such deterministic anymore.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 56 | 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 | 0.1289 |
| 57 | 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 | 0.1311 |
| 58 | 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 | 0.1333 |
| 59 | 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 | 0.1355 |
| 60 | 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 | 0.1377 |
| 61 | 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 | 0.1399 |
| 62 | 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 | 0.1421 |
| 63 | 2 2 2 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 | 0.1443 |
| 64 | 2 2 2 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 | 0.1465 |
| 65 | 2 2 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 | 0.1487 |
| 66 | 2 2 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 | 0.1509 |
| 67 | 2 2 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 | 0.1531 |
| 68 | 2 2 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 | 0.1553 |
| 69 | 2 2 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 | 0.1574 |
| 70 | 2 2 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 5 | 0.1596 |
| 71 | 2 2 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 5 | 0.1617 |
| 72 | 2 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 5 | 0.1638 |
| 73 | 2 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 5 | 0.1660 |
| 74 | 2 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 5 5 | 0.1682 |
| 75 | 2 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 5 5 5 | 0.1703 |
| 76 | 2 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 5 5 5 | 0.1725 |
| 77 | 2 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 5 5 5 5 | 0.1746 |
| 78 | 2 3 3 3 3 3 4 4 4 4 4 4 4 4 4 5 5 5 5 5 | 0.1768 |
| 79 | 2 3 3 3 3 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 | 0.1789 |
| 80 | 2 3 3 3 3 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 | 0.1810 |
| 81 | 2 3 3 3 3 4 4 4 4 4 4 4 4 5 5 5 5 5 5 5 | 0.1831 |
| 82 | 2 3 3 3 3 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 | 0.1852 |
| 83 | 2 3 3 3 4 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 | 0.1873 |
| 84 | 2 3 3 3 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 | 0.1894 |
| 85 | 2 3 3 3 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 6 | 0.1915 |
| 86 | 2 3 3 3 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 6 | 0.1936 |
| 87 | 2 3 3 3 4 4 4 4 4 4 5 5 5 5 5 5 5 5 6 6 | 0.1957 |
| 88 | 2 3 3 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 6 6 | 0.1977 |
| 89 | 2 3 3 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 6 6 | 0.1998 |
| 90 | 2 3 3 4 4 4 4 4 4 5 5 5 5 5 5 5 5 6 6 6 | 0.2019 |

Table 2: The optimal solution of allocating a total capacity C over a homogeneous ASIP with 20 queues where arrivals of particles and gate openings have capacity 1 and the total capacity ranges from 56 to 100.

Looking at table 2 and 3 you can observe that in an optimal solution not only $c_i \geq c_{i-1}$ for $i \in \{2, ..., n\}$ holds but more strongly: $c_{i-1} + 1 \geq c_i \geq c_{i-1}$ for $i \in \{2, ..., n\}$. By this observation we see in the case of an homogeneous ASIP another way for determining an optimal solution for allocating total capacity $C$ over $n$ queues. Suppose you have an optimal solution $c_1, ..., c_n$ for allocating total capacity $C - 1$ over $n$ queues. By looking at the pattern in tables 1, 2 and 3 we only have to add 1 extra capacity to a certain queue in order to gain an optimal solution of allocating a total capacity $C$. The set of indexes of possible queues for adding 1 extra capacity is $\{i | c_i < c_{i+1}, 1 \leq i < n\} \cup \{n\}$.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 91 | 2 3 3 4 4 4 4 4 4 5 5 5 5 5 5 5 6 6 6 6 | 0.2039 |
| 92 | 2 3 3 4 4 4 4 4 5 5 5 5 5 5 5 5 6 6 6 6 | 0.2060 |
| 93 | 2 3 3 4 4 4 4 4 5 5 5 5 5 5 5 6 6 6 6 6 | 0.2080 |
| 94 | 2 3 3 4 4 4 4 4 5 5 5 5 5 5 6 6 6 6 6 6 | 0.2101 |
| 95 | 2 3 3 4 4 4 4 5 5 5 5 5 5 5 6 6 6 6 6 6 | 0.2121 |
| 96 | 3 3 4 4 4 4 4 5 5 5 5 5 5 5 5 6 6 6 6 6 | 0.2142 |
| 97 | 3 3 4 4 4 4 4 5 5 5 5 5 5 5 6 6 6 6 6 6 | 0.2162 |
| 98 | 3 3 4 4 4 4 5 5 5 5 5 5 5 5 6 6 6 6 6 6 | 0.2183 |
| 99 | 3 3 4 4 4 4 5 5 5 5 5 5 5 6 6 6 6 6 6 6 | 0.2204 |
| 100 | 3 3 4 4 4 4 5 5 5 5 5 5 6 6 6 6 6 6 6 6 | 0.2224 |
| : | : | : |
| $\infty$ | all queues have infinite capacity | 1 |

Table 3: Continuation of the optimal solution of allocating a total capacity C over a homogeneous ASIP with 20 queues where arrivals of particles and gate openings have capacity 1 and the total capacity ranges from 56 to 100.

The other way around also holds. When you have an optimal solution $c_1, ..., c_n$ for allocating total capacity $C$ over $n$ queues, then we only have to remove 1 capacity from a certain queue in order to gain an optimal solution of allocating a total capacity of $C - 1$. The set of indexes of possible queues for removing 1 capacity is $\{i | c_i > c_{i-1}, 1 < i \leq n\} \cup \{1\}$.

Suppose you have an optimal solution $c_1, ..., c_n$ for allocating total capacity $C - 1$. We have seen that an optimal solution for allocating total capacity $C$ one queue has capacity $c_i + 1$ and the rest of the assigned capacities is the same. It remains a suspicion that this always holds for all $n \in \mathbb{N}^+$, for all $C \geq n$ and for all $\lambda, \mu \in \mathbb{R}^+$, not a fact.

The relation between total capacity $C$ and the throughput of the optimal solution of allocating total capacity $C$ is shown in figure 6.

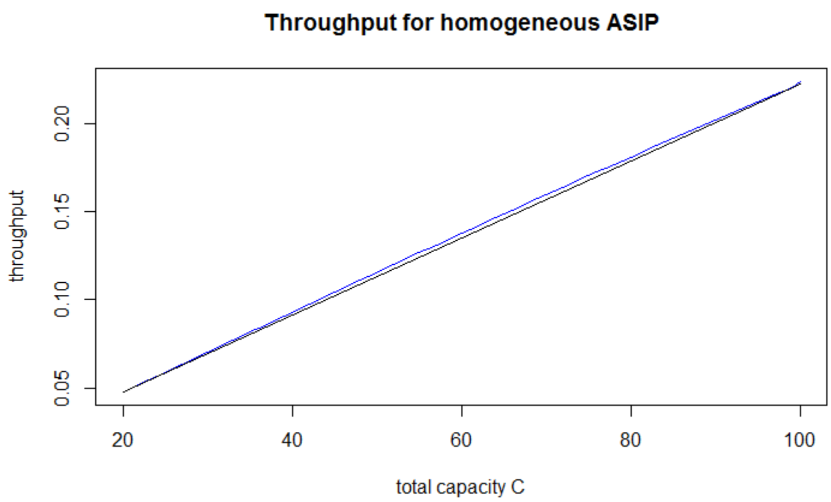


Figure 6: The throughput in particles per time unit for the optimal solution of allocating total capacity C (blue line); the black line is a linear line for showing that the throughput is asymptotic

In this figure can be seen that the throughput grows between $C = 20$ and $C = 100$ almost linearly. The black line is the linear line between the throughput for $C = 20$ and the throughput for $C = 100$. Using this black line we can see that the slope of the throughput slowly decreases. This is what we expect because the throughput converges to 1 when total capacity $C$ tends to infinity.

### 4.6.2 homogeneous ASIP with 20 queues with $\lambda = 10$ and $\mu = 1$

When we study the homogeneous ASIP with 20 queues with $\lambda = 10$ and $\mu_i = 1$ for all $i \in \{1, ..., n\}$ the optimal solution for allocating capacity $C$ where $C$ varies from 20 to 100 is more equally spread than in the case of an homogeneous ASIP with $\lambda = 1$ and $\mu_i = 1$ for all $i \in \{1, ..., n\}$. The optimal solution for some values of total capacity C can be seen below in table 4. The exact value for the throughput when $C = 20$ is according to (103) equal to $\frac{1}{\frac{1}{10}+20\cdot\frac{1}{1}} = \frac{10}{201}$ particles per time unit.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 20 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 0.0498 |
| 30 | 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 | 0.0745 |
| 40 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.0993 |
| 50 | 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 | 0.1240 |
| 60 | 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 | 0.1487 |
| 70 | 2 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 | 0.1735 |
| 80 | 2 3 3 3 4 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 | 0.1981 |
| 90 | 3 4 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5 5 | 0.2227 |
| 100 | 3 4 4 4 4 5 5 5 5 5 5 5 5 5 6 6 6 6 6 6 | 0.2473 |
| : | : | : |
| $\infty$ | all queues have infinite capacity | 10 |

Table 4: Optimal solution of allocating a total capacity $C$ over the homogeneous ASIP with 20 queues where $\lambda = 10$ and $\mu_i = 1$ for $i \in \{1, ..., n\}$.

### 4.6.3 homogeneous ASIP with 20 queues with $\lambda = 1$ and $\mu = 10$

When we study the homogeneous ASIP with 20 queues with $\lambda = 1$ and $\mu_i = 10$ for all $i \in \{1, ..., n\}$ the optimal solution of allocating capacity C where C varies from 20 to 100 is less equally spread than in the case of an homogeneous ASIP with $\lambda = 1$ and $\mu_i = 1$ for all $i \in \{1, ..., n\}$. The exact value for the throughput when $C = 20$ is according to (103) equal to $\frac{1}{\frac{1}{1}+20\cdot\frac{1}{10}} = \frac{1}{3}$ particles per time unit. The optimal solution from some values for total capacity C can be seen below in table 5.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 20 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 0.3333 |
| 30 | 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 | 0.4444 |
| 40 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.5556 |
| 50 | 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 | 0.6481 |
| 60 | 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 | 0.7215 |
| 70 | 2 2 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 5 5 | 0.7810 |
| 80 | 2 3 3 3 3 3 4 4 4 4 4 4 4 5 5 5 5 5 5 5 | 0.8291 |
| 90 | 2 3 3 3 4 4 4 4 4 4 5 5 5 5 5 6 6 6 6 6 | 0.8673 |
| 100 | 3 3 3 4 4 4 4 5 5 5 5 5 6 6 6 6 6 6 7 7 | 0.8970 |
| : | : | : |
| $\infty$ | all queues have infinite capacity | 1 |

Table 5: Optimal solution of allocating a total capacity $C$ over the homogeneous ASIP with 20 queues where $\lambda = 1$ and $\mu_i = 10$ for $i \in \{1, ..., n\}$.

So in an homogeneous ASIP the following can be concluded: the higher the arrival intensity of particles in comparison with the gate opening intensities, the more capacity has to be assigned to the first queues, the more the total capacity is equally allocated over the system.

## 4.7 The optimal solution of inhomogeneous ASIP with 20 queues

An ASIP is inhomogeneous when there is at least some $i, j \in \{1, ..., n\}$ with $\mu_i \neq \mu_j$. There are two special subsets of the set of inhomogeneous ASIP's, namely monotone increasing and decreasing ASIP's. A monotone increasing ASIP has the property that $\mu_i \geq \mu_{i-1}$ for all $i \in \{2, ..., n\}$ and analogously a monotone decreasing ASIP has the property that $\mu_i \leq \mu_{i-1}$ for all $i \in \{2, ..., n\}$. For both subsets I give an impression for the optimal solution of allocating capacity C over an ASIP with 20 queues.

### 4.7.1 Optimal solution of monotone decreasing ASIP with 20 queues

An example of a monotone decreasing ASIP with 20 queues is an ASIP with intensities: $\lambda = 1$ and $\mu_i = i^{-1}$ for $i \in \{1, ..., 20\}$.The exact value for the throughput when $C = 20$ is according to (100) equal to $\frac{1}{\frac{1}{1}+\sum_{j=1}^{20} i} = \frac{1}{211}$ particles per time unit. The optimal solutions of allocating a total capacity of $20, 30, ..., 100$ are shown table 6.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 20 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 0.0047 |
| 30 | 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 3 3 | 0.0084 |
| 40 | 1 1 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 4 4 | 0.0116 |
| 50 | 1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 4 4 4 5 5 | 0.0146 |
| 60 | 1 1 1 1 1 2 2 2 2 3 3 3 3 4 4 5 5 5 6 6 | 0.0176 |
| 70 | 1 1 1 1 2 2 2 3 3 3 3 4 4 4 5 5 6 6 7 7 | 0.0206 |
| 80 | 1 1 1 2 2 2 3 3 3 3 4 4 5 5 5 6 7 7 8 8 | 0.0235 |
| 90 | 1 1 2 2 2 2 3 3 4 4 4 5 5 6 6 7 7 8 9 9 | 0.0265 |
| 100 | 1 1 2 2 2 3 3 4 4 4 5 5 6 6 7 7 8 9 10 11 | 0.0294 |
| : | : | : |
| $\infty$ | all queues have infinite capacity | 1 |

Table 6: Optimal solution of allocating a total capacity $C$ over the monotone decreasing ASIP with 20 queues where $\lambda = 1$ and $\mu_i = i^{-1}$ for $i \in \{1, ..., n\}$.

From this table you can see that the allocation of capacity C is less equally distributed than in the homogeneous case. This sounds logic because gate $i$ opens less frequently when $i$ becomes larger. Still it holds that $c_{i-1} + 1 \geq c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$, but this will not hold anymore if we choose the intensities more extreme as we see in the next example.

In the next example of a monotone decreasing ASIP with 20 queues is the arrival intensity $\lambda = 1$ and gate opening intensity $\mu_i = 2^{-(i-1)}$ for $i \in \{1, ..., n\}$. So the first gate has a gate opening intensity of 1 opening per time unit and every following queue opens on average twice as much as the previous one. The exact value for the throughput when $C = 20$ is according to (100) equal to $\frac{1}{\frac{1}{1}+\sum_{j=1}^{20} 2^{j-1}} = \frac{1}{1048576}$ particles per time unit. The optimal solutions of allocating a total capacity of 20, 30, ... , 100 are shown table 7.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 20 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | $9.54_{\times 10^{-7}}$ |
| 30 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 4 6 | $3.95_{\times 10^{-6}}$ |
| 40 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 4 6 11 | $6.46_{\times 10^{-6}}$ |
| 50 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 6 9 15 | $8.83_{\times 10^{-6}}$ |
| 60 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 4 7 11 19 | $1.12_{\times 10^{-5}}$ |
| 70 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 5 8 14 24 | $1.35_{\times 10^{-5}}$ |
| 80 | 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 4 6 10 16 26 | $1.59_{\times 10^{-5}}$ |
| 90 | 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 4 7 11 18 32 | $1.82_{\times 10^{-5}}$ |
| 100 | 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 5 8 13 21 35 | $2.05_{\times 10^{-5}}$ |
| : | : | : |
| $\infty$ | all queues have infinite capacity | 1 |

Table 7: Optimal solution of allocating a total capacity $C$ over the monotone decreasing ASIP with 20 queues where $\lambda = 1$ and $\mu_i = 2^{-(i-1)}$ for $i \in \{1, ..., n\}$.

In this example you can see that the allocation of capacity $C$ is even less equally allocated than in the previous example. So in the case of a monotone decreasing ASIP you have to assign more and more capacity to the rightmost queues when the gate opening intensities drop.

### 4.7.2 Optimal solution of monotone increasing ASIP with 20 queues

An example of a monotone increasing ASIP with 20 queues is an ASIP with intensities $\lambda = 1$ and $\mu_i = i$ for $i \in \{1, ..., n\}$. The exact value for the throughput when $C = 20$ is according to (100) equal to $\frac{1}{\frac{1}{1}+\sum_{j=1}^{20}\frac{1}{j}} = \frac{15519504}{71354639}$ particles per time unit. The optimal solutions of allocating a total capacity of $20, 30, ..., 100$ are shown in table 8.

| total capacity C | Optimal allocation of capacities | throughput |
|---|---|---|
| 20 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 0.2175 |
| 30 | 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 | 0.2491 |
| 40 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 | 0.3877 |
| 50 | 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 | 0.4176 |
| 60 | 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 | 0.5209 |
| 70 | 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 | 0.5481 |
| 80 | 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 | 0.6251 |
| 90 | 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5 | 0.6492 |
| 100 | 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 | 0.7066 |
| : | : | : |
| $\infty$ | all queues have infinite capacity | 1 |

Table 8: Optimal solution of allocating a total capacity $C$ over the monotone increasing ASIP with 20 queues where $\lambda = 1$ and $\mu_i = i$ for $i \in \{1, ..., n\}$.

In this table you can see that the allocation of capacity is more equally spread than in the homogeneous case. This sounds logic because gate $i$ opens more frequent when $i$ becomes larger. In the homogeneous case there was till a total capacity of 55 a deterministic way to get the optimal solution, namely allocate $C - (C \bmod n)$ capacity equally over the queues and then assign one extra capacity to each of the $C \bmod n$ rightmost queues (= queues $j$ where $j \in \{n-(C \bmod n)+1, ..., n\}$). In this example this deterministic way holds even for a total capacity of 100.

When we take the gate opening intensities more extreme, namely $\mu_i = 2^{i-1}$ for $i \in \{1, ..., n\}$ we can predict that the optimal solutions of allocating a total capacity of $20, 30, ..., 100$ are the same as in table 8, because gates which open faster need less capacity and lemma 2.1 tells us that in the set of distributions of capacity $C$ where $c_i \geq c_{i-1}$ for all $i \in \{2, ..., n\}$ there is an optimal solution. After computing the optimal solutions for this case the optimal solutions are actually what we expected to be. The exact value for the throughput when $C = 20$ is according to (100) equal to $\frac{1}{\frac{1}{1}+\sum_{j=1}^{20}\frac{1}{2^{i-1}}} = \frac{524288}{1572863} \approx 0.3333$ particles per time unit.

So in the case of a monotone increasing ASIP you have to spread the total capacity more and more evenly over the queues when the gate opening intensities increase.

## References

[1] S. Reuveni, "Tandem Stochastic Systems: The Asymmetric Simple Inclusion Process," PhD Thesis, Tel Aviv University, 2014.

[2] S. Reuveni, I. Eliazar, and U. Yechiali, "Asymmetric Inclusion Process as a Showcase of Complexity," Physical Review Letters, vol.109, no. July, pp. 1-4, 2012.

[3] S.Reuveni, I. Eliazar, and U. Yechiali, "Asymmetric inclusion process," Physical Review E, vol.84, no. 4, pp. 1-16, 2011.

[4] S.Reuveni, I. Eliazar, and U. Yechiali, "Limit laws for the asymmetric inclusion process," Physical Review E, vol. 86, no. 6, pp. 1-17, 2012.

[5] O.J.Boxma, O. Kella, and U. Yechiali, "An ASIP model with general gate opening intervals," 2014, Paper in preparation.