Approximating multiple arrival streams by using aggregation

TU/e technische universiteit eindhoven

SPOR-Report 2004-16

Approximating multiple arrival streams by using aggregation

M. van Vuuren
I.J.B.F. Adan

SPOR-Report
Reports in Statistics, Probability and Operations Research

Eindhoven, December 2004
The Netherlands

/department of mathematics and computing science

SPOR-Report
Reports in Statistics, Probability and Operations Research

Eindhoven University of Technology
Department of Mathematics and Computing Science
Probability theory, Statistics and Operations research
P.O. Box 513
5600 MB Eindhoven - The Netherlands

Secretariat: Main Building 9.10
Telephone: + 31 40 247 3130
E-mail: wscosor@win.tue.nl
Internet: http://www.win.tue.nl/math/bs/cosor.html

ISSN 1567-5211

Approximating multiple arrival streams by using aggregation

Marcel van Vuuren, Ivo J.B.F. Adan

Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands
E-mail: m.v.vuuren@tue.nl, i.j.b.f.adan@tue.nl

Abstract: *In this paper we consider the superposition of independent Coxian arrival streams. We propose a method to efficiently approximate the superposition process, based on state space aggregation. The method is applied to the $\Sigma C_{k_i}/G/1$ queue and to an inventory control model. Simulation results show that the method accurately estimates performance characteristics and that it significantly outperforms two-moment approximations.*

Keywords: *superposition, Markovian arrival process, aggregation, queueing system, inventory control.*

1 Introduction

Systems fed by multiple arrival streams are very common in practice. For example, in a production line, the input to a work station is the output of the machines in the upstream station. Another example is a wholesale house where many retailers place replenishment orders. In this paper we consider the superposition of independent arrival streams, each with independent Coxian distributed inter-arrival times. A complicating feature of the superposition is that its inter-arrival times may be no longer independent. Further, an exact description of the superposition is computationally not feasible, because the state space explodes when the number of arrival streams and the number of phases of the Coxian distributions increase. Therefore accurate and efficient approximations are needed. However, not much work has been done in approximating multiple arrival streams. Usually multiple arrival streams are approximated by a renewal process, the inter-arrival times of which are determined by a two-moment fit. Thus dependencies between successive inter-arrival times are completely ignored. For example, this approach was employed by Van Vuuren et al. [10] in a production environment and by Smits et al. [8] in an inventory environment. In general, this approximation can lead to severe errors. A more sophisticated method has been developed by Albin [1] and Whitt [11]. They also approximate the superposition by a renewal process, but the second moment of the inter-arrival time is determined differently: the squared coefficient of variation of the inter-arrival time is determined as a convex combination of the squared coefficients of variation obtained from the so-called asymptotic and stationary-interval approximations. This method gives reasonable results for an $\Sigma G_i/G/1$ queue, but Van Nyen et al. [7] applied the method to a production-inventory system and concluded that it may result in serious errors. Mitchell [5] developed a method to fit a matrix exponential distributed process on a correlated arrival process leaving the first order properties invariant. This method works well, but it cannot handle the specific correlation structure of multiple arrival streams. The method assumes that the magnitude of the correlation coefficients decreases in the lag and that the correlation coefficients are all positive or alternating in sign. Typically, these assumptions are not satisfied by superpositions of arrival streams.

The superposition of independent Coxian arrival stream can be exactly described by a Markovian

Arrival Process (MAP), but as indicated above, its state space grows exponentially in the number of streams and in the number of phases of the Coxian distributions. We propose to aggregate the state space, i.e., to aggregate the states in which the total number of completed phases is the same. The number of aggregate states grows polynomially, instead of exponentially. The approximate arrival process is obtained by assuming the aggregate process is again a MAP. To obtain the parameters of this MAP we develop two algorithms, one for Coxian arrival streams and another (more efficient) one for the case that all streams have identical mixed-Erlang inter-arrival times.

We use the aggregation method to approximate the performance of a $\Sigma C_{k_i}/G/1$ queue and of an inventory control model. The results are compared with a two-moment fit and the method of Whitt and Albin. Simulation results show that the aggregation method is very accurate and it works much better than the other ones. The average errors in performance characteristics of a $\Sigma C_{k_i}/G/1$ queue are around 2% and for the inventory control model, they are around 0.2%. So we can conclude that the aggregate (compact) MAP provides an efficient and accurate tool to describe multiple arrival streams, and it opens the door to the use of powerful matrix-analytical techniques to approximately evaluate the performance of $\sum GI/G/s$ queues.

2 Model and Aggregation

We consider the superposition of m independent arrival processes, labeled $1, \ldots, m$. Arrival process $i = 1, \ldots, m$ has independent Coxian$_{k_i}$ distributed inter-arrival times with parameters $\nu_{i,j}$ and $p_{i,j}$ with $j = 0, \ldots, k_i - 1$; the phase-diagram is shown in Figure 1. Note that the family of Coxian$_k$ distributions is dense in the family of all distributions on $(0, \infty)$, see, e.g., [2].

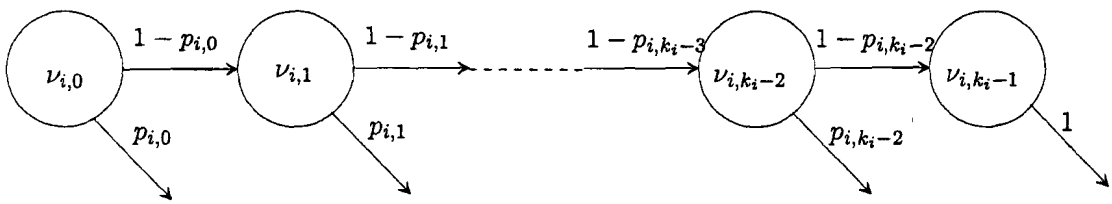

Figure 1: A phase diagram of the Coxian distribution of the inter-arrival time of the i-th arrival process.

The superposition of Coxian arrival processes can be described by a Markovian Arrival Process (MAP), a useful model for point processes (see [3]). A MAP is defined in terms of a continuous-time Markov process with finite state space, $\{0, \cdots, k-1\}$ say, and generator $A_0 + A_1$. The element $A_{0,ij}$ denotes the intensity of transitions from i to j accompanied by an arrival, whereas for $i \neq j$ element $A_{0,ij}$ denotes the intensity of the remaining transitions from i to j and the diagonal elements $A_{0,ii}$ are negative and chosen such that the row sums of $A_0 + A_1$ are zero. For example, for a single Coxian arrival stream with parameters k, ν_i, and p_i, $i = 0, \ldots, k-1$, the transition rate matrices A_0 and A_1 (of dimension $k \times k$) are given by:

$$
\begin{aligned}
A_{0,(i,i)} &= -\nu_i, && \text{for } i = 0, \ldots, k-1, \\
A_{0,(i,i+1)} &= (1-p_i)\nu_i, && \text{for } i = 0, \ldots, k-2, \\
A_{1,(i,0)} &= p_i\nu_i, && \text{for } i = 0, \ldots, k-1,
\end{aligned} \tag{1}
$$

and all other entries are zero. The long-run fraction of time π_i that this MAP is in state i is equal to

$$
\pi_i = \frac{\prod_{j=0}^{i-1}(1-p_j)}{C}, \tag{2}
$$

where C is the normalization constant. For more information on MAPs, the reader is referred to [6].

The superposition of m independent Coxian arrival processes can be described as a MAP with states $(j_1, j_2, \ldots, j_m)$, where $j_i = 0, \ldots, k_i - 1$ represents the number of completed phases of the inter-arrival time of arrival process i. The number of states is $k_1 \cdots k_m$, which grows exponentially in the number of arrival streams. Therefore, to keep the size of the state space limited, we aggregate the state space as follows. We take together all states with the same total number of completed arrival phases, i.e., aggregate state i corresponds to the set of states $(j_1, j_2, \ldots, j_m)$ with $j_1 + \cdots + j_m = i$, where i runs from 0 to $K = k_1 + \cdots + k_m - m$. Note that K grows polynomially in m. To illustrate the aggregation procedure we show in Figure 2 the phase diagram of the superposition of two Erlang$_4$ arrival processes; the aggregated states are indicated by the rings.

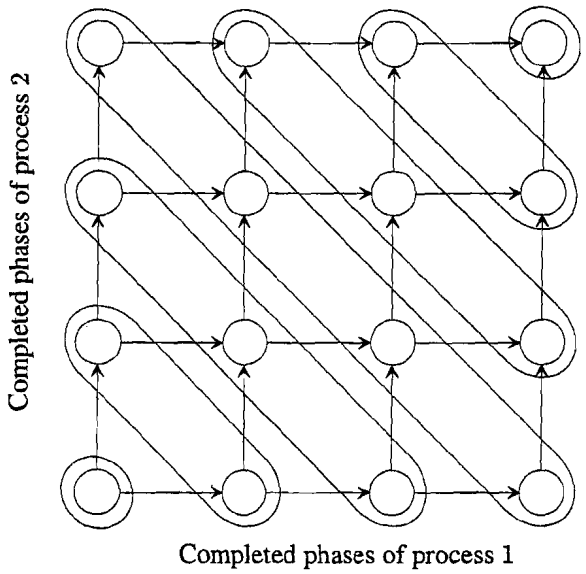

Figure 2: A phase diagram of the superposition of two Erlang$_4$ arrival processes and its aggregation.

For the aggregated process we can exactly determine the fraction of time π_i spent in state i, and the

number of transitions per time unit $r_{i,j}$ from state i to j. Then the transition rate from i to j is given by $q_{i,j} = r_{i,j}/\pi_i$. Note that rates q_{ij} with $i > j$ correspond to arrivals; the ones with $i < j$ do not. Figure 3 shows the aggregated states and their transition rates for the example in Figure 2. An efficient algorithm for computing the transition rates $q_{i,j}$ is presented in the next section. The aggregated process is, in general, *not Markovian* and the sojourn times in the states may not be exponential (except when for each stream $\nu_{i,j} = \nu_i, j = 0, \ldots, k_i - 1$; then the sojourn time in an aggregated state is exponential with parameter $\nu_1 + \cdots + \nu_m$). Now, the crucial step is that we treat the aggregated process again as a MAP with transition rates $q_{i,j}$ and thus we act as if the sojourn times are exponential and the transitions are memoryless. Further the rates $q_{i,j}$ with $i > j$ (corresponding to arrivals) are put in the matrix A_1 and the rest are put in A_0. So, for the example of two Erlang$_4$ arrival streams, we use Figure 3 as flow diagram for the MAP obtained from aggregation. This procedure yields a compact MAP approximately describing the original arrival process and that naturally preserves some of the dependencies present in the original process.

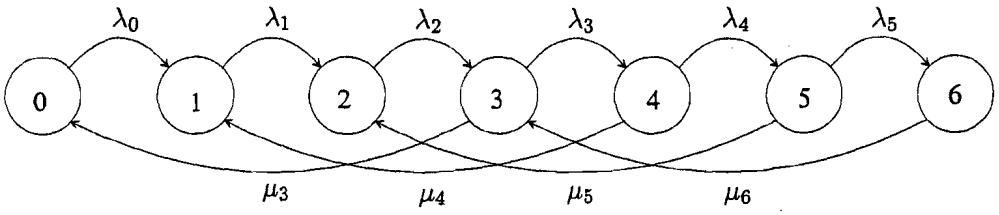

Figure 3: A diagram of the aggregated MAP of the superposition of two Erlang$_4$ arrival processes.

In Section 3 we develop algorithms to efficiently compute the transition rates of the aggregated MAP.

Remark 1 The superposition of m identical Coxian$_2$ arrival streams can be described by a MAP with states j_2 denoting the number of streams in the second phase of the Coxian inter-arrival time (the other $m - j_2$ streams are in the first phase). Thus the number of states is linear in m. Now the proposed aggregation does not help, i.e., it leads to exactly the same states as the original arrival process.

3 Transition rates of the aggregated MAP

The MAP obtained by aggregation has a special structure. The process is skip-free to the right; it can jump from state i to $i + 1$, but it is not possible to make larger jumps to the right. Transitions to the right do not correspond to arrivals. Hence, the transition rate matrix A_0 has negative elements on the diagonal and non-zero elements on the super-diagonal; the other elements are zero. Note that, in case of m arrival streams with Erlang$_{1,\ldots,k_i}$ distributed inter-arrival times with parameters ν_i and $p_{i,j}$ with $j = 0, \ldots, k_i - 1$, the time the aggregated process spends in each state is exponentially distributed with parameter $\nu = \sum_{i=1}^{m} \nu_i$. So, in this special case, the diagonal elements of A_0 are all equal to $-\nu$. The aggregated process can make large jumps to the left; these transitions correspond to an arrival. This means that the transition matrix A_1 is a lower triangular matrix. In the following section we present a method to find the transition matrices A_0 and A_1. In Section 3.2 we describe a

more efficient method for the special case of identical Erlang$_{k-1,k}$ arrival streams.

3.1 Coxian arrival streams

To determine the transition rates of the aggregated process we aggregate the Coxian arrival streams one by one, i.e., we aggregate the first two streams, then add the third one and so on. Below we describe how to add an arrival stream.

Suppose we have aggregated a number of arrival streams and now we want to add the next stream, with transition rates A_0 and A_1 and state-space $\{0,\ldots,k_1-1\}$. The long-run fraction of time spent in state i is π_i (see (1) and (2)). The transition rates of the aggregated streams are B_0 and B_1 and the state space is $\{0,\ldots,k_2-1\}$. Obviously, the number of states k_2 depends on the number of streams that is aggregated. The long-run fraction of time the aggregated process spends in state i is ξ_i. Let us refer to these processes as process 1 (the new stream) and process 2 (the ones already aggregated).

By aggregating process 1 and 2 we obtain a process with state space $\{0,\ldots,k_3\}$ where $k_3 = k_1 + k_2 - 2$. Now we want to determine its transition rates C_0 and C_1 and the long-run fraction of time η_i it spends in state i. Since process 1 and 2 are independent of each other, we immediately have

$$\eta_i = \sum_{j=\max(0,i-k_2+1)}^{\min(i,k_1-1)} \pi_j \xi_{i-j}, \qquad i = 0,\ldots,k_3. \tag{3}$$

To determine the number of transitions per time unit out of state i, note that, if process 1 is in state j and process 2 in state $i-j$, the total number of transitions out of these states is $\pi_j\xi_{j-i}(A_{0,(j,j)} + B_{0,(i-j,i-j)})$. Hence, adding over all feasible states yields

$$\eta_i C_{0,(i,i)} = \sum_{j=\max(0,i-k_2+1)}^{\min(i,k_1-1)} \pi_j \xi_{i-j} \left(A_{0,(j,j)} + B_{0,(i-j,i-j)}\right), \qquad i = 0,\ldots,k_3. \tag{4}$$

Similarly we obtain, for all $i = 0,\ldots,k_3-1$,

$$\eta_i C_{0,(i,i+1)} = \sum_{j=\max(0,i-k_2+1)}^{\min(i,k_1-2)} \pi_j \xi_{i-j} A_{0,(j,j+1)} + \sum_{j=\max(0,i-k_2+2)}^{\min(i,k_1-1)} \pi_j \xi_{i-j} B_{0,(i-j,i-j+1)}, \tag{5}$$

and for all $i = 1,\ldots,k_3$ and $l = 1,\ldots,i$,

$$\eta_i C_{1,(i,i-l)} = \sum_{j=\max(l,i-k_2+1)}^{\min(i,k_1-1)} \pi_j \xi_{i-j} A_{1,(j,j-l)} + \sum_{j=\max(0,i-k_2+1)}^{\min(i-l,k_1-1)} \pi_j \xi_{i-j} B_{1,(i-j,i-j-l)}. \tag{6}$$

The time needed to determine the transition rates C_0 and C_1 is $O(k_3^2 \min(k1,k2))$. Hence, the time complexity to aggregate m Coxian$_{k_i}$ arrival streams is $O(mK^2k_{\max})$ where $K = \sum_{i=1}^{m} k_i$ and $k_{\max} = \max(k_1,\ldots,k_m)$.

3.2 Identical mixed Erlang arrival streams

For identical arrival streams we can develop a simpler and more efficient algorithm to determine the transition rates of the aggregated arrival process. Suppose we have m arrival streams, each with $\text{Erlang}_{k-1,k}$ distributed inter-arrival times with parameters p and μ. The number of states of the aggregated process is $(k-1)m+1$. First we use (3) to determine the long-run fraction of time π_i the aggregated process spends in state i, where $i = 0, \ldots, (k-1)m$. Second we determine the transition rates A_0 and A_1 of the aggregated process in one go. Of course, the rate at which we leave a state is always equal to $\nu = m\mu$, so

$$A_{0,(i,i)} = -\nu \qquad i = 0, \ldots, (k-1)m. \tag{7}$$

To determine the transition rates we start with the last state $(k-1)m$ and then work our way back to the front. From state $(k-1)m$ we can only jump to state $(k-1)(m-1)$, thus

$$A_{1,((k-1)m,(k-1)(m-1))} = \nu. \tag{8}$$

State i with $(k-1)(m-1)+1 \leq i \leq (k-1)m$ can only be reached from state $i-1$. Hence, application of the balance principle to state i yields

$$\pi_{i-1}A_{0,(i-1,i)} = -\pi_i A_{0,(i,i)},$$

and thus, by (7),

$$A_{0,(i-1,i)} = \frac{\pi_i}{\pi_{i-1}}\nu, \qquad i = (k-1)(m-1)+1, \ldots, (k-1)m. \tag{9}$$

State i with $1 \leq i \leq (k-1)(m-1)$ can be reached from state $i-1$, but also from $i+k-2$ and $i+k-1$. So, by application of the balance principle to state i we obtain

$$A_{0,(i-1,i)} = \frac{\pi_i}{\pi_{i-1}}\nu - \frac{\pi_{i+k-1}}{\pi_{i-1}}A_{1,(i+k-1,i)} - \frac{\pi_{i+k-2}}{\pi_{i-1}}A_{1,(i+k-2,i)}, \tag{10}$$

valid for all $i = 1, \ldots, (k-1)(m-1)$. Note that a transition from state $i+k-2$ to i corresponds to the completion of the $k-1$th and last phase of an inter-arrival time (and thus to an arrival); a transition from $i+k-1$ to i corresponds to the completion of the kth phase. Since the fraction of inter-arrival times consisting of $k-1$ phases is p (the rest consists of k phases), it follows that the ratio of the number of transitions per time unit from state $i+k-2$ to i and the number of transitions per time unit from state $i+k-1$ to i is equal to $p/(1-p)$. Thus

$$\pi_{i+k-2}(1-p)A_{1,(i+k-2,i)} = \pi_{i+1}pA_{1,(i+k-1,i)},$$

which can be rewritten as

$$A_{1,(i,i-(k-2))} = \frac{\pi_{i+1}(1-p)}{\pi_i p}A_{1,(i+1,i+1-(k-1))}, \qquad i = k-2, \ldots, (k-1)m-1. \tag{11}$$

Finally, the transition rates from states i to $i-(k-1)$ for $i = k-1, \ldots, (k-1)m$ can be found by using that the rate at which we leave a state is always ν, so we get

$$A_{1,(i,i-(k-1))} = \nu - A_{1,(i,i-(k-2))} - A_{0,(i,i+1)}, \qquad i = k-1, \ldots, (k-1)m. \tag{12}$$

Using the equations (7)-(12) the transition rates can be determined recursively. The other rates in A_0 and A_1 are 0. The time complexity of this algorithm is only $O(m^2k^2)$.

4 Numerical Results

In this section we investigate how accurate the aggregated MAP describes the superposition of arrival streams. To do so we use simulation. First, in Section 4.1, we consider the $\Sigma C_{k_i}/G/1$ queue. We compare the mean waiting time (W) and the probability that the system is empty (p_e) for the real system with the corresponding characteristics for:

(i) the system where the inflow is the aggregated MAP;

(ii) the $GI/G/1$ system where the first two moments of the inter-arrival time match with the first two moments of an arbitrary inter-arrival time of the superposition of arrival streams (see [10]);

(iii) the $GI/G/1$ system where the first two moments of the inter-arrival time are determined according to the method of Whitt [11] and Albin [1].

The distribution of the inter-arrival time in (ii) and (iii) is determined by matching the first two moments, i.e., we use a mixed Erlang or Coxian$_2$ distribution, depending on whether the squared coefficient of variation is less or greater than 1 (see, e.g., [9]). More specifically, let $1/\lambda$ and c_a^2 denote the mean and squared coefficient of variation of the inter-arrival time. If $1/k \leq c_a^2 \leq 1/(k-1)$ for some $k = 2, 3, \ldots$, then the mean and squared coefficient of variation of the Erlang$_{k-1,k}$ distribution with density

$$f_a(t) = p\mu^{k-1}\frac{t^{k-2}}{(k-2)!}e^{-\mu t} + (1-p)\mu^k\frac{t^{k-1}}{(k-1)!}e^{-\mu t}, \qquad t \geq 0,$$

matches with $1/\lambda$ and c_a^2, provided the parameters p and μ are chosen as

$$p = \frac{1}{1+c_a^2}[kc_a^2 - \{k(1+c_a^2) - k^2c_a^2\}^{1/2}], \qquad \mu = (k-p)\lambda. \tag{13}$$

Second, in Section 4.2, we consider an inventory control system, consisting of a central stock point that faces demand from several retailers. The inter-arrival times of orders from a retailer are assumed assumed to be Coxian distributed. We compare the fraction of demand delivered from stock (β) and the average inventory level (I) for the real system with the corresponding characteristics for:

(i) the inventory system with an aggregated MAP demand process;

(ii) the inventory system where the demand process is a renewal process, the first two moments of the inter-arrival times of which match with the first two moments of an arbitrary inter-arrival time of the superposition of the demand processes of the retailers.

Now we can not compare it with the system with a renewal demand process, the first two moments of the inter-arrival times of which are determined by the method of Whitt [11] and Albin [1]. The reason is that their method is specifically designed for $\Sigma G_i/G/1$ queues.

In Section 4.3 we briefly discuss the results. All simulation results have 95% confidence intervals, the widths of which are smaller than 1%.

4.1 The $\Sigma C_{k_i}/G/1$ queue

We have simulated systems with identical arrival processes and systems with different arrival processes. We start with the identical arrival processes.

The inter-arrival times of the m identical arrival processes are $\text{Erlang}_{k-1,k}$ distributed with squared coefficient of variation c_a^2. We vary the number of arrival streams between 2, 5 and 10. The squared coefficient of variation of the inter-arrival times are varied between 0.4, 0.2 and 0.1. The service times are exponentially distributed, Erlang_k distributed or deterministic, depending on the squared coefficient of variation c_d. We vary c_d^2 between 1, 0.2 and 0. The parameters of the Erlang distributions are determined according to (13). Note that we only consider arrival streams, the inter-arrival times of which have a squared coefficient of variation less than 0.5; the reason is that the aggregation is exact when the squared coefficient of variation is greater or equal to 0.5 (see Remark 1). Finally, the occupation rate ρ of the queueing system is varied between 0.8 and 0.9. This results in a total of 54 test cases. A summary of the results can be found in Table 1.

Real system	Error in W					Error in p_e				
compared with	Avg.	0-5 %	5-10 %	10-15 %	> 15 %	Avg.	0-5 %	5-10 %	10-15 %	> 15 %
Aggregated MAP	1.6 %	96.3 %	3.7 %	0.0 %	0.0 %	3.3 %	74.1 %	25.9 %	0.0 %	0.0 %
Whitt and Albin	4.8 %	66.7 %	22.2 %	9.3 %	1.9 %	10.9 %	5.6 %	48.2 %	25.9 %	20.4 %
Two moment fit	34.8 %	0 %	1.8 %	11.1 %	87.0 %	10.9 %	5.6 %	48.2 %	25.9 %	20.4 %

Table 1: Overall results for the queueing system with identical arrival streams.

In case of nonidentical arrival streams we use $m = 5$ arrival streams. The arrival rates are varied between the sets $(0.6, 0.8, 1.0, 1.2, 1.4)$ and $(1, 1, 1, 1, 1)$ and the squared coefficients of variation of the inter-arrival times are varied between the sets $(0.2, 0.25, 0.3, 0.4, 0.5)$ and $(0.3, 0.3, 0.3, 0.3, 0.3)$. The case in which both the rates and the squared coefficients of variation are the same is left out. The rest of the parameters are the same as in case of identical streams. This leads to 18 test cases. A summary of the results is presented in Table 2.

Real system	Error in W					Error in p_e				
compared with	Avg.	0-5 %	5-10 %	10-15 %	> 15 %	Avg.	0-5 %	5-10 %	10-15 %	> 15 %
Aggregated MAP	2.8 %	83.3 %	16.7 %	0.0 %	0.0 %	1.2 %	100.0 %	0.0 %	0.0 %	0.0 %
Whitt and Albin	15.7 %	16.7 %	16.7 %	33.3 %	33.3 %	9.0 %	0.0 %	72.2 %	27.8 %	0.0 %
Two moment fit	31.5 %	0.0 %	0.0 %	0.0 %	100.0 %	9.0 %	0.0 %	72.2 %	27.8 %	0.0 %

Table 2: Overall results for the queueing system with different arrival streams.

4.2 Inventory control system

For the inventory control system we also make a distinction between identical and different order arrival processes. We start with the first case.

The inventory control system consists of one stock point, facing demand from m retailers. A schematic representation of the inventory system is shown in Figure 4. The stock point is controlled by an (s, nQ)-installation stock policy; for more information see Smits [8]. The batch size Q is varied between 800 and 900, the reorder level s is varied between 900 and 1000 and the replenishment lead-time L is exactly 8 time units. The number of retailers m is varied between 2, 5 and 10. The batch

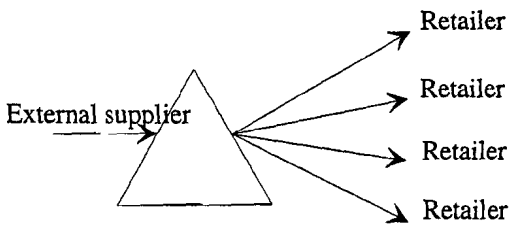

Figure 4: Schematic representation of the inventory system.

sizes of all orders have mean μ where $m\mu = 150$ and squared coefficient of variation 0.5. The number of retailers m is varied between 2, 5 and 10. The average time between two successive orders of a retailer is 1 time unit and the squared coefficient of variation c_a^2 is varied between 0.4, 0.2 and 0.1. Again, the distribution of the inter-arrival time of an order is $\text{Erlang}_{k-1,k}$, the parameters of which are determined by (13). This gives 36 test cases. Table 3 summarizes the results of these tests.

Real system compared with	Error in I					Error in β				
	Avg.	0-1 %	1-2 %	2-3 %	> 3 %	Avg.	0-1 %	1-2 %	2-3 %	> 3 %
Aggregated MAP	0.14 %	100.0 %	0.0 %	0.0 %	0.0 %	0.09 %	100.0 %	0.0 %	0.0 %	0.0 %
Two moment fit	3.53 %	0.0 %	8.3 %	33.3 %	58.4 %	1.85 %	0.0 %	61.1 %	38.9 %	0.0 %

Table 3: Overall results for the inventory system with identical order streams.

In case of different order arrival processes we use $m = 5$ retailers. The arrival rates of the orders $\bar{\mu}$ are varied between $(0.6, 0.8, 1.0, 1.2, 1.4)$ and $(1, 1, 1, 1, 1)$ and the squared coefficients of variation $\bar{c}_a^2$ are varied between $(0.2, 0.25, 0.3, 0.4, 0.5)$ and $(0.3, 0.3, 0.3, 0.3, 0.3)$. The case in which both the rates and the coefficients of variation are the same is left out. The rest of the parameters are the same as in case of identical order arrival processes. This leads to 12 cases. The results are summarized in Table 4.

Real system compared with	Error in I					Error in β				
	Avg.	0-1 %	1-2 %	2-3 %	> 3 %	Avg.	0-1 %	1-2 %	2-3 %	> 3 %
Aggregated MAP	0.26 %	100.0 %	0.0 %	0.0 %	0.0 %	0.19 %	100.0 %	0.0 %	0.0 %	0.0 %
Two moment fit	4.06 %	0.0 %	0.0 %	16.7 %	83.3 %	2.22 %	0.0 %	25.0 %	75.0 %	0.0 %

Table 4: Overall results for the inventory system with different order streams.

4.3 Discussion

We see from the results in the previous sections that use of the aggregated (compact) MAP gives much better results than the other two methods, both for the queueing system and the inventory control system. As expected, the method of Whitt and Albin predicts the mean waiting time better than the two moment method. But their performance is the same for the probability of an empty system p_e. Further we see that the results for different arrival streams are about the same as for identical streams. However, the results of Whitt and Albin's method get worse in case of different arrival

streams. Overall we can say that use of the aggregated MAP works very well for both systems. It outperforms the two moment methods, but of course, the price to be paid is that it is more complicated to analyze a model with a MAP than one with a (renewal) phase-type arrival stream.

5 Conclusions and future research

In this paper we proposed to use an aggregated MAP for the (approximate) description of the superposition of Coxian arrival streams. Simulation results suggest that the aggregated MAP captures the characteristics of the superposition of arrival streams very well. Application to a queueing model and an inventory control model leads to average errors in predicting performance characteristics around 2%. So we may conclude that aggregated MAP can be very useful to accurately approximate superpositions of arrival processes.

In future work we will try to use MAP descriptions for *both* the arrival and service process in the $\Sigma G_i/G/s$ and the $\Sigma G_i/G/s/K$ queue. The resulting (approximate) models may be analyzed efficiently by the matrix geometric method and the spectral expansion method (see, e.g. Bertsimas [4]). The ultimate goal is to use these models to evaluate the performance of queueing networks with blocking [10].

References

[1] S.L. Albin (1984) Approximating a Point Process by a Renewal Process, 2: Superposition of Arrival Processes to Queues. *Operations Research* 32, 1133-1162.

[2] S. Asmussen (2003) *Applied Probability and Queues.* Springer, New York.

[3] S. Asmussen and G. Koole (1993) Marked point processes as limits of Markovian arrival streams. *Journal of Applied Probability* 30, 365-372.

[4] D. Bertsimas (1990) An analytic approach to a general class of G/G/s queueing systems. *Operations Research*, 1, 139-155.

[5] K. Mitchell (2001) Constucting a correlated sequence of matrix exponentials with invariant first-order properties. *Operations Research Letter* 28, 27-34.

[6] M.F. Neuts (1989) *Structured Stochastic Matrices of M/G/1-type and their Applications.* Marcel Dekker, New York, NY.

[7] P.L.M. Van Nyen, H.P.G. Van Ooijen and J.W.M. Bertrand (2004) Simulation results on the performance of Albin and Whitt's estimation method for waiting times in integrated production-inventory systems. *To appear in Int. J. Production Economics*

[8] S.R. Smits, A.G. De Kok and G.P. Kiesmüller (2002) Approximation for the evaluation of multi-echelon inventory systems with order consolidation. *Submitted for publication.*

[9] H.C. Tijms (1994) *Stochastic models: an algorithmic approach.* John Wiley & Sons, Chichester.

[10] M. van Vuuren, I.J.B.F. Adan and S.A. Resing-Sassen (2003) Performance Analysis of Multi-Server Tandem Queues with Finite Buffers and blocking. *Submitted for publication.*

[11] W. Whitt (1982) Approximating a Point Process by a Renewal Process, I: Two Basic Methods. *Operations Research* 30, 125-147.