
DOI: 10.1080/0740817X.2016.1222470

©2020. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
On the Output Dynamics of Production Systems Subject to Blocking

Barış Tan
College of Administrative Sciences and Economics, Koç University, Rumeli Feneri Yolu, Istanbul, Turkey, 34450,
Email: btan@ku.edu.tr Phone: +90 212 338 1459 Fax: +90 212 338 1653

Svenja Lagershausen
Department of Production Management, Leibniz University of Hanover, Hanover, Germany,
Email: svenja.lagershausen@prod.uni-hannover.de Phone: +49 511 762 2982 Fax: +49 511 762 4863

IIE Transactions, forthcoming, 2016. Accepted Manuscript.

Analyzing the output dynamics of a production system gives valuable information for operation and performance evaluation of a production system. In this paper, we present an analytical method to determine the auto-correlation of the inter-departure times in queueing networks subject to blocking that can be represented by a Continuous Time Markov Chain. We particularly focus on production systems that are modeled as open or closed queueing networks where stations have phase-type service time distributions. We use the analytical results for the mean and the variance of the time to produce a given number of products in queueing networks to determine the correlation of inter-departure times with different lags. We present a computationally efficient recursive method to determine the correlation of the inter-departure times in open and closed queueing networks. The method also yields closed-form expressions for the correlations of a two-station production line with exponential servers and a finite buffer. We show how the correlations develop with increasing lags subject to different processing time distributions, buffer capacities, and number of stations, both in open and closed queueing networks. As a result, we propose the analytical method given in this study as a tool to study the effects of design and control parameters on the output dynamics of production systems.

Key words: Performance Evaluation of Manufacturing Systems, Stochastic methods, Queueing Theory, Production System Design

1. Introduction

A wide range of production, service, and telecommunication systems can be modeled as queueing networks subject to blocking. Our motivation for this study comes from using queueing network models to develop analytical techniques that will be used in design of production systems. Analyzing the output dynamics of a production system, i.e. the quantity and timing of flow units departing from the system, gives valuable information for operation and for performance evalua-
tion of these systems. The output of one system is usually an input to another system. Therefore analyzing the output dynamics of one stage gives information about the arrival dynamics of the following stages. Representing the arrival dynamics in a way that ignores possible correlation affects the performance of the system negatively.

The effects of the system parameters such as the service time variability, the buffer capacities, the number of stations, and the control parameters on the output dynamics can be investigated by evaluating the performance of the system. This information can then be used to design and control production systems in a more effective way.

Many studies in the manufacturing systems literature focus on the production rate that is the average number of parts produced per time unit in the long term as a first order performance measure to summarize the output process. In addition, average work-in-process inventory is also used as a first order performance measure. The number of studies that focus on statistical properties of the output dynamics is limited.

In general, statistical properties of the output process can be summarized by the distribution of the inter-departure time and by the statistical dependencies of the sequence of inter-departure times. Many performance measures related to the number of units leaving the system and the time to obtain a given number of parts can be derived from the distribution of the inter-departure time and the autocorrelation vector of the output which is a vector containing the correlation of inter-departure times with lag $i$ as its $i$th element.

Figure 1 shows the frequency distribution of the inter-departure time and the sample autocorrelation of inter-departure times for cars leaving an assembly line of a Japanese automotive manufacturer. The data we collected from this assembly line encompass the inter-departure times of 1,276 cars (the data related to cars produced immediately after breaks are excluded). This figure shows that the cars leave the assembly line with a time interval close to 8 minutes on average, and the inter-departure time exhibits a low variability (coefficient of variation of 0.12). Furthermore, the dependency of inter-departure times on the previous departures is present but it gets weaker for longer lags. Observing low variability and weak long-term dependency in the inter-departure times in this system is a result of using a predetermined time that is referred to as the takt time.
Figure 1  Frequency Distribution of the Inter-departure Time and Sample Autocorrelation of Inter-departure Times for a Car Assembly Line

![Graph showing frequency distribution and autocorrelation of inter-departure times.]

to control the flow of cars in this assembly line. Similarly, Inman (1999) shows that the autocorrelation of inter-arrival times to work centers collected from two automotive body welding lines exhibit low positive autocorrelation. These inter-arrival times are also the inter-departure of the preceding work centers that are composed of closed loops of unreliable machines.

Although the long-term dependencies of the inter-departure times are not strong in the example of the automotive manufacturing (see Figure 1), significant long-term dependencies of inter-departure times have been reported for semiconductor manufacturing (Schomig and Mittler 1995). They show that correlation of inter-departure times affects downstream stages such as packaging lines at the end of a semiconductor plant negatively in terms of cycle time performance. In previous simulation studies of semiconductor manufacturing plants, it has also been observed that closed-loop control mechanisms, such as CONWIP control, cause long-term dependencies of the output process (Wein 1988). Furthermore, statistical properties of the arrivals to a system have a significant effect on its performance. Since the output of one stage is an input to another stage, designing a system in a way to control its output dynamics is an alternative way to control the downstream stages. As these examples indicate, analyzing the output dynamics of production systems gives valuable information about the operation of a production system.

Following our recent work on determining the distribution of the inter-departure time of closed queueing networks subject to blocking (Lagershausen and Tan 2013), in this study we present an analytical method to determine the autocorrelation of inter-departure times from queueing networks that can be modeled as Continuous Time Markov Chains. We focus on production systems that are modeled as open or closed queueing networks with stations that have random
service times described by phase-type distributions. The method only uses the state-transition rate matrix of the system and an indicator matrix that identifies the states where a departure occurs as its inputs. The mean and the variance of the time to produce a given number of products derived from the inter-departure time distribution are used to determine the correlation of inter-departure times with different lags. We propose the analytical method given in this study as a tool to study the effects of design and control parameters on the output dynamics of production systems. In order to model distributions with a coefficient of variation unequal to one and additionally to carry out an exact analysis, we use phase-type distributions. These distributions can be used to approximate a wide range of distributions.

The organization of the remaining part of this paper is as follows. In Section 2, the pertinent literature is reviewed. The methodology is presented in Section 3. Section 4 analyzes a two-station production line with exponential servers and a finite buffer to explain the methodology. The methodology presented in this paper yields closed-form expressions for the expected value and variance of the inter-departure time and the correlations of lag 1 and lag 2 for this specific system. Numerical results for multi-station systems with different service time distributions are presented in Section 5. Finally, conclusions are given in Section 6.

2. Literature Review

There are numerous studies on stochastic models of production systems (Dallery and Gershwin 1992). In this section, we focus on the studies that are closely related to the output processes from queueing networks subject to blocking.

Lagershausen and Tan (2015) developed a method to determine the inter-departure, inter-start, and cycle-time distributions for closed queueing networks. This paper develops a method to determine the auto-correlation of inter-departure times for open and closed queueing networks by using the method given in (Lagershausen and Tan 2015) to determine the expectation and variance of the sums of inter-departure times to derive recursive equations for correlations. In this process, this paper also gives recursive expressions for expectation and variance that were not given in the earlier paper.
This paper is closely related to (Hendricks 1992) and (Hendricks and McClain 1993). Hendricks (1992) derived the marginal probability function for the inter-departure time of exponential, finite-buffered open production lines using Markov chains. Hendricks and McClain (1993) investigated the influence of the input parameters on the variance of the inter-departure time for general, finite-buffered open queueing systems by simulation.

This study extends the analytical results of Hendricks (1992) for exponential open serial production lines to open and closed queueing networks with phase-type distributions. As a result, the analysis conducted by simulation, such as the one in Hendricks and McClain (1993), for production lines with general service distributions can be conducted analytically for production systems with phase-type service distributions by using the analytical method presented in this paper.

The number of studies that analyze the variability of the output and timing in addition to its expected performance is quite limited, see e.g. (Duenyas et al. 1993), (Li and Meerkov 2000), (Li and Meerkov 2001), (Manitz and Tempelmeier 2011), (Tan 2013), (Assaf et al. 2014), (Angius et al. 2014), and (Shi and Gershwin 2016). In the manufacturing systems literature, although there are many analytical results for first- and second-order performance measures related to the output process from production systems, the interdependency of the output process from production lines has been studied mainly by using simulation. For example, Sabuncuoglu, Erel, and Kok (2002) studied the effect of number of stations, processing time distributions, and buffers on throughput and inter-departure time variability of assembly systems by using simulation. Kalir and Sarin (2009) proposed a method to reduce the inter-departure time variability in serial production lines via simulation.

Assaf, Colledani, and Matta (2014) proposed approximations for the correlation of the output in order to incorporate the variance of the output in the analysis of two-station subsystems. These are then used in a decomposition approach for capacititated multi-stage production lines where machines are prone to random failures. He, Wu, and Li (2007) computed the autocorrelation for the inter-arrival times.

In the queueing literature, there are several studies that focus on the inter-departure time process from single server queues. King (1974) studied the covariance structure of the departure
process for a M/G/1 queue with finite waiting space. Reynolds (1972) surveys the works carried out so far regarding the covariance structure in queues. Daley (1976) discusses the second-order properties of the inter-departure times for G/G/1/N systems where the probability function is approximated by the Palm-Khinchin equation. Zhang, Hendll, and Smirni (2005) modeled the departure process of a BMAP/MAP/1 queue using Markov chains providing results for the autocorrelation of the inter-departure times. Nazarathy and Weiss (2005) calculate the asymptotic variance rate of the output of finite birth and death queues. Lipsky, Fiorini, Hsin, and Van de Liefvoort (1995) studied the inter-departure time distribution and the correlation of departures for G/G/1, M/G/1, and G/M/1 systems using Markov chains.

For a Markov Arrival Process (MAP), Neuts (1995) presented a recursive method to determine the moments of inter-arrival times and the autocorrelation of lag k based on the generator matrices of an absorbing Markov chain describing the arrival process. Kriege and Buchholz (2014) presented a method to fit a MAP that matches the statistical properties of real traffic data and uses the Neuts’ method to evaluate the fitted MAP in terms of the autocorrelation of the observed data and the fit.

Queueing networks are also commonly used to model telecommunication networks. In the communication literature, the correlation of the output from queueing networks is studied analytically and also by using simulation. DeSimone (1991) carried out a simulation study on the correlation effects in networks of queues. Mitchell, van de Liefvoort, and Place (1998) focused on the correlation properties of the token leaky bucket based on the covariance lag-k of the inter-departure process. These studies show that determining the inter-departure time correlation is necessary to determine buffer size requirements at a downstream node or switch in an ATM network. Ignoring autocorrelation of inter-arrival times and representing the input stream as a renewal process leads to setting buffer levels incorrectly, and causes higher blocking compared to the intended design levels.

This is the first study that gives an analytical method to determine the autocorrelation function for the inter-departure time of open and closed queueing networks with phase-type servers subject to blocking. The main contribution of this study is to present a general method that uses only
the state space model of a queueing network as its input to analytically determine the correlation of the inter-departure time with a given lag. This method allows to analyze the effects of system parameters such as the variability of service times, buffer capacities, and number of stations on the output correlation structure analytically, as an alternative to the existing simulation studies for production systems.

3. Methodology

We consider a queueing network subject to blocking that can be modeled as a Continuous-Time Markov Chain. There are $K$ stations. Each station has a finite input buffer. The processing time on each station follows a phase-type distribution. The state of the system at time $t$ is $X(t) \in S$. There are $|S|$ states in the state space $S$. $X(t)$ is a $K$-tuple where each element shows the number of parts on a given station including its input buffer, its phase, and whether it is blocked or not. Since the processing times have phase-type distributions, the process $\{X(t), t \geq 0\}$ is a Continuous-Time Markov Chain. The corresponding infinitesimal generator or the rate matrix of the process $\{X(t), t \geq 0\}$ is denoted with $Q = \{q_{i,j}\}$.

To compute the correlation of the inter-departure time for a given number of lags, we use the mean and variance of the sum of $k$ sequential inter-departure times. A short review of the methodology to determine the inter-departure time distribution presented in Lagershausen and Tan (2015) is given in Sections 3 for completeness. We then discuss how the original state space is extended for the time to produce $k$ parts (Section 3.2.1) and show how to compute the expected value and the variance of this sequential sum (Section 3.2.2). We present the iterative procedure using this variance to determine the correlations in Section 3.3.

3.1. Computation of the Inter-departure Time Distribution

We analyze the output process of the departure of a part from the last station by focusing on the inter-departure time. The inter-departure time is the time from the instant a part departs from a station to the instant the next part departs from the same station. In other words, the inter-departure time $T_n$ is the first-passage time from the state at time $t_{n-1}$ to the state at time $t_n$ where $t_n$ is the departure time of the $n$th part. Then the steady-state distribution of the inter-departure time is the distribution of $T = \lim_{n \to \infty} \{t_{n+1} - t_n\}$. 
In order to determine the inter-departure time distribution by using a first passage time analysis, a new process is defined based on \( \{X(t), t \geq 0\} \) to identify transitions that lead to a departure from the last station. Accordingly, the states in the state space \( S \) are duplicated in the state space of a new process \( \{X^{(1)}(t), t \geq 0\} \). The rate matrix of the process \( \{X^{(1)}(t), t \geq 0\} \) is denoted by a \( 2|S| \times 2|S| \) matrix \( Q^{(1)} \). The rate matrix \( Q^{(1)} \) is

\[
Q^{(1)} = \begin{bmatrix}
Q_d & R_d \\
0 & 0
\end{bmatrix}
\]

(1)

where \( Q_d \) is the \( |S| \times |S| \) infinitesimal generator submatrix for the transitions from the transient states to other transient states, \( R_d \) is the \( |S| \times |S| \) submatrix with elements that are the transition rates from the transient states where the part on the last station is being processed to the absorbing states where this part departs the last station, and 0s are \( |S| \times |S| \) matrices of zeroes.

The matrices \( Q_d \) and \( R_d \) in Equation (1) are determined directly from the state transition matrix of the original process \( Q \) and an indicator matrix \( G_d \) that identifies and represents the transitions that lead to a departure. The methodology we present uses \( Q \) and \( G_d \) as the main inputs.

### 3.1.1. Determining the Indicator Matrix \( G_d \) and the Transition-rate Matrix \( Q_d \)

In order to identify a departure event, let \( G_d = \{g_{i,j}\} \) be a \( |S| \times |S| \) indicator matrix where \( g_{i,j} \) is 1 if a transition from state \( i \) to state \( j \) increases the number of departures by one and it is zero otherwise. The elements of \( G_d = \{g_{i,j}\} \) are determined by checking if the event of a departure occurs by a transition from state \( i \) to state \( j \).

In general, since the output of one system usually feeds into a downstream system, the departure event can be determined by considering the service completions at the last station in the system. For example, the matrix \( G_d \) to determine the correlation of the inter-departure time from the last station is given explicitly for a two-station system in Section 4. However, with this setting, the inter-departure times from any station in a production system can be analyzed by defining the departure event accordingly. The statistical properties such as the variance of the inter-departure time, its distribution, and the auto-correlation structure depend on the station used to define \( G_d \).
Figure 9 in Section 5 depicts the correlations of inter-departure times for a three-station line when the correlations are determined for the first, second, and the third stations.

The matrices $Q_d$ and $R_d$ are determined using $G_d$ and the original generator matrix of the queueing network $Q$ as given below:

$$R_d = Q \circ G_d$$

(2)

where $(Q \circ G_d)_{i,j} = q_{i,j}g_{d_{i,j}}$ is the element-wise product (Schur product) of matrices $Q$ and $G_d$ and

$$Q_d = Q - R_d.$$ 

(3)

3.1.2. Distribution, Mean, and Variance of the Inter-departure Time. By using the state transition matrix of the original process $Q$ and the indicator matrix $G_d$, the inter-departure time distribution can be determined from the steady-state first passage time distribution. The steady-state inter-departure time probability distribution $F_T(t)$, the mean and the variance of the inter-departure time, $E[T]$ and $Var[T]$ are given as

$$F_T(t) = 1 - \pi^{\text{entry}} e^{Q_d t} u,$$

(4)

$$E(T) = -\pi^{\text{entry}} Q_d^{-1} u,$$

(5)

$$Var(T) = 2 \pi^{\text{entry}} Q_d^{-2} u - (\pi^{\text{entry}} Q_d^{-1} u)^2$$

(6)

where $\pi^{\text{entry}}$ is determined from the solution of the following equations

$$\pi^{\text{entry}} (I + Q_d^{-1} R_d) = 0,$$

(7)

$$\pi^{\text{entry}} u = 1$$

(8)

where $u$ is a $|S| \times 1$ column vector of ones (Lagershausen and Tan 2013).

3.2. Analysis of Sequence of Inter-departure Times

In this study, we are interested in characterizing the output process of a production system in terms of the correlation of the inter-departure times with a given lag. Namely, we are interested in the dependency of one inter-departure time duration $T_i$ with the next inter-departure time $T_{i+1}$,
with the subsequent one $T_{i+2}$, etc. The steady-state covariance of $T_i$ and $T_{i+k}$, $\text{Cov}(T_i, T_{i+k})$, and the lag $(k)$ correlations, $\text{Corr}(T)_k$, are defined as

$$\text{Cov}(T_i, T_{i+k}) = E(T_i \cdot T_{i+k}) - E(T_i) \cdot E(T_{i+k})$$

(9)

and

$$\text{Corr}(T)_k = \frac{\text{Cov}(T_i, T_{i+k})}{\text{Var}(T)}$$

(10)

for $i = 1, 2, \ldots$ and $k = 1, 2, \ldots$

where $E(T)$ and $\text{Var}(T)$ are determined from Equations (2) to (4). In steady state, the covariance between two inter-departure times depends only on the lag between these times but not on the time itself. That means

$$\text{Corr}(T)_k = \frac{\text{Cov}(T_i, T_{i+k})}{\text{Var}(T)}$$

(11)

for $i = 1, 2, \ldots$ and $k = 1, 2, \ldots$

### 3.2.1. Extending the State Space to Analyze the Sequential Sum of Inter-departure Times

In order to analyze the sequence of inter-departure times $T_i, \ldots, T_k$, we first extend the state space. We define a new process based on $\{X(t), t \geq 0\}$ to identify transitions that lead to $k$ subsequent departures from the last station. The new process is denoted with $\{X^{(k)}(t), t \geq 0\}$.

In order to form the extended state space, all the states in $S$ are duplicated in the state space of the new process $\{X^{(k)}(t), t \geq 0\}$. The same state $i$ in $S$ appears $(k+1)$ times for the states that correspond to the production of the first part, second part, ..., $k$th part. In order to differentiate the duplicated states, we refer to them as states in $S^{(1)}, S^{(2)}, \ldots, S^{(k)}$. As a result, the state space of $\{X^{(k)}(t), t \geq 0\}$ is $S \cup S^{(1)} \cup S^{(2)} \cup \cdots \cup S^{(k)}$ and has $(k+1)|S|$ states.

Accordingly, the state transition rate matrix that describes the transitions to produce $k$ parts, i.e., $k$ parts departing from the last station, is given as

$$Q^{(k)} =$$

$$
\begin{bmatrix}
Q_d & R_d & 0 & 0 & 0 & \cdots & 0 \\
0 & Q_d & R_d & 0 & 0 & \cdots & 0 \\
0 & 0 & Q_d & R_d & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \cdots & Q_d & R_d \\
0 & 0 & 0 & \cdots & \cdots & 0 & 0
\end{bmatrix}.
$$

(13)

The rate matrix $Q^{(k)}$ can be put in the form
where $Q_d^{(k)}$ is the $k|S| \times k|S|$ infinitesimal generator submatrix for the transitions from the transient states to other transient states, $R_d^{(k)}$ is a submatrix with elements that are the transition rates from the transient states to the absorbing states, and the 0 entries are zero matrices of appropriate sizes.

### 3.2.2. Mean and Variance of the Time to Produce $k$ Parts.

Using the above defined matrices, the mean and the variance of the time to produce $k$ parts, i.e. the time until $k$ parts depart, $E(\Gamma_k)$ and $Var(\Gamma_k)$ where $\Gamma_k = \sum_{i=1}^{k} T_i$ are obtained directly from the matrices $Q_d^{(k)}$ and $R_d^{(k)}$ following the expressions for the mean and variance of the inter-departure time of a single part given in Equations (5) and (6):

$$E(\Gamma_k) = -\pi^{\text{entry}(k)}(Q_d^{(k)})^{-1} u^{(k)},$$

$$Var(\Gamma_k) = 2\pi^{\text{entry}(k)}(Q_d^{(k)})^{-2} u^{(k)} - (E(\Gamma_k))^2,$$

where $\pi^{\text{entry}(k)}$ is a $1 \times k|S|$ row vector of the steady-state probability of entering the states in the extended state space and $u^{(k)}$ is a $k|S| \times 1$ column vector of ones.

The above equations require the inversion of the matrix $Q_d^{(k)}$ that is of the dimension $k|S| \times k|S|$. As shown in the following, the mean and the variance of the time to produce $k$ parts can be determined by only using the submatrices $Q_d$ and $R_d$ that are of the size $|S| \times |S|$ without determining the inverse of $Q_d^{(k)}$.

**Proposition 1** Let $\Phi = Q_d^{-1}$ and $\Omega = -R_d \Phi$. By exploiting the special structure of the matrix $Q^{(k)}$, the mean and the variance of the time for $k$ parts to depart, $\Gamma_k = \sum_{i=1}^{k} T_i$, can be determined as

$$E(\Gamma_k) = -\pi^{\text{entry}} \sum_{i=1}^{k} \Phi \Omega^{i-1} u = kE(T) = -k\pi^{\text{entry}} \Phi u,$$

$$Var(\Gamma_k) = 2\pi^{\text{entry}} \sum_{i=1}^{k} \Phi \Omega^{i-1} \sum_{j=1}^{i} \Phi \Omega^{j-1} u - (E(\Gamma_k))^2.$$
3.3. Iterative Procedure to Compute the Autocorrelations between $k$ successive Inter-departure Times

In general, the variance of the sum of $k$ dependent variables $\Gamma_k = \sum_{i=1}^{k} T_i$ is given as

\[
Var(\Gamma_k) = \sum_{i=1}^{k} \sum_{j=1}^{k} Cov(T_i, T_j) = \sum_{i=1}^{k} Var(T_i) + 2 \sum_{i=1}^{k} \sum_{j=i+1}^{k} Cov(T_i, T_j). \tag{19}
\]

By using the stationarity of the covariance of the inter-departure time in queueing networks, Equation (19) is simplified in the following way:

\[
Var(\Gamma_k) = kVar(T) + 2 \sum_{j=1}^{k-1} (k-j)Cov(T_1, T_{j+1}). \tag{20}
\]

Using the variance of the inter-departure time $Var(T)$ and all variances of the time to produce $i$ parts, $Var(\Gamma_i)$, with $i = 2, 3, ..., k$ in Equation (20), the covariances can be computed iteratively. The covariance of the first and the second inter-departure time is

\[
Cov(T_1, T_2) = \frac{1}{2} (Var(\Gamma_2) - 2Var(T)). \tag{21}
\]

Continuing recursively for $k = 3, 4, ..., \text{the covariance for lag of } (k-1), k \geq 3 \text{ can be written as}

\[
Cov(T_1, T_k) = \frac{1}{2} \left( Var(\Gamma_k) - kVar(T) - 2 \sum_{j=1}^{k-2} (k-j)Cov(T_1, T_{j+1}) \right). \tag{22}
\]

Namely calculating the expected value and the variance of the time to produce $k$ parts and the covariance of inter-departure times with 1 to $k-1$ lags yield the correlation of inter-departure times with lag $k$.

The expected value and the variance of the sum of a given number of inter-departure times can also be written in a recursive way. The following result gives a set of equations that can be solved recursively to determine the mean and variance of the time until $k$ units depart and also the correlation of inter-departure times with lags 1 to $k$.

**Proposition 2** Starting with $Var[\Gamma_1] = Var[T]$, $\Upsilon_1 = \Psi_1 = \Phi = Q_d^{-1}$ and $\Omega = -R_d \Phi$, the following equations are solved for $k = 2, 3, ..., m$ iteratively to determine the covariance between the departure times with a lag of $1, 2, ..., m-1$ periods, $Corr(T)_1, ..., Corr(T)_{m-1}$:

\[
\Upsilon_k = \Upsilon_{k-1} \Omega \tag{23}
\]
\[
\Psi_k = \Psi_{k-1} + \Upsilon_k
\]  

(24)

\[
\text{Var}(\Gamma_k) = \text{Var}(\Gamma_{k-1}) + (1 - 2k)(E(T))^2 + 2\pi^{\text{entry}} \Psi_k \Upsilon_k u
\]  

(25)

\[
\text{Cov}(T_1, T_k) = \begin{cases} 
\frac{1}{2} (\text{Var}(\Gamma_k) - k\text{Var}(T)) & \text{for } k = 2 \\
\frac{1}{2} \left( \text{Var}(\Gamma_k) - k\text{Var}(T) - 2\sum_{j=1}^{k-2} (k-j)\text{Cov}(T_1, T_{j+1}) \right) & \text{for } k \geq 3 
\end{cases}
\]  

(26)

\[
\text{Corr}(T)_k = \frac{\text{Cov}(T_1, T_k)}{\text{Var}(T)}.
\]  

(27)

The proof is given in Appendix B.

Note that starting with matrices \(Q^{-1}_d\) and \(-R_d\Phi\), the above recursions have only two matrix multiplications, \(\Upsilon_{k-1}\Omega\) and \(\Psi_k \Upsilon_k\) at each iteration and one matrix addition \(\Psi_{k-1} + \Upsilon_k\) and do not include any matrix inversion. As a result, the recursion is much more efficient compared to the direct calculations given in Equations (15) and (16) that require determining the square of the inverse of a \(k|S| \times k|S|\) matrix, especially for calculation of autocorrelation for longer lags.

4. Two-Station Exponential Production Line Model with a Finite Buffer

In this section, we focus on a production line with two stations with exponential processing times with rates \(\mu_i, i = 1, 2\) and a finite inter-station buffer of capacity of \(M\). By using this model, we demonstrate how the methodology presented in the preceding sections is applied to a specific example. In addition, we show that the methodology presented in this paper yields closed-form expressions for the correlations of the inter-departure times with lags of 1 and 2, \(\text{Corr}(T)_1\) and \(\text{Corr}(T)_2\) for a homogenous line with \(\mu = \mu_1 = \mu_2\).

4.1. State Space Model

The state of the system for this system is a 2-tuple \(X(t) = (X_1, X_2)\) where \(X_k\) is the state of Station \(k, k = 1, 2\). \(X_1\) gives the number of parts at Station 1 and also indicates whether it is blocked. \(X_2\) gives the number of parts including the part in the buffer. For example \(X_1 = 1\) indicates that there is one part being processed at station 1, \(X_1 = B\) indicates that station 1 is blocked. \(X_2 = 0\) corresponds to the event that the second station is starved. When the buffer capacity is \(M\), there are \(M + 3\) states in the state space:

\[
S = \{(1,0), (1,1), (1,2), \cdots, (1,M-1), (1,M), (1,M+1), (B,M+1)\}.
\]  

(28)
Table 1 shows the corresponding transition rate matrix $Q$ for the case of $M = 2$.

<table>
<thead>
<tr>
<th></th>
<th>(1,0)</th>
<th>(1,1)</th>
<th>(1,2)</th>
<th>(1,3)</th>
<th>(B,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,0)</td>
<td>$-\mu_1$</td>
<td>$\mu_1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(1,1)</td>
<td>$\mu_2$</td>
<td>$-\mu_1 - \mu_2$</td>
<td>$\mu_1$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(1,2)</td>
<td>0</td>
<td>$\mu_2$</td>
<td>$-\mu_1 - \mu_2$</td>
<td>$\mu_1$</td>
<td>0</td>
</tr>
<tr>
<td>(1,3)</td>
<td>0</td>
<td>0</td>
<td>$\mu_2$</td>
<td>$-\mu_1 - \mu_2$</td>
<td>$\mu_1$</td>
</tr>
<tr>
<td>(B,3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\mu_2$</td>
<td>$-\mu_2$</td>
</tr>
</tbody>
</table>

### 4.2. Inter-departure Time Distribution

In order to determine the inter-departure time distribution, we define a new process by duplicating
the state space $S$. Since $Cov(T)_1$ and $Cov(T)_2$ are of interest, the state space is extended to
include three departures from the last station. Figure 2 shows the state transition diagram of the
process that captures the production of parts on Station 1 for three succeeding inter-departure
time cycles. The time until 3 parts produced is the first passage time from the states on the first
column to the states on the fourth column. Similarly, the time until 1 and 2 parts are the first
passage time from the states on the first column to the states on the second and third columns
respectively. Furthermore, the steady-state probabilities of starting in one of the entry states in
the first column is equal to the steady-state probabilities of absorption in one of the exit states in
the fourth column.

The transitions that correspond to an increase in the number of parts produced on Station 2
are captured by the indicator matrix $G_d$. Table 2 gives the indicator matrix $G_d$ for the process
depicted in Figure 2. Applying $Q$ and $G_d$ (given in Tables 1 and 2) to Equations (2) and (3),
the transition rate matrices $Q_d$ and $R_d$ are obtained. Once $Q_d$ and $G_d$ are available, Equations
(4) to (6)) yield the distribution function of the inter-departure time, its density, expectation and
variance.

### 4.3. Correlation of Inter-departure Times with Lags 1 and 2

The matrices $\Phi = Q_d^{-1}$ and $\Omega = -R_d\Phi$ can be given in closed form for a given $M$. For example,
Table 4 gives the matrix $\Omega$ for $M = 2$. 

The state transition diagram of the process $\{X_n^{(3)}(t), t \geq 0\}$ associated with three departures from a two station production line with exponential servers and a finite buffer with a capacity of 2

Table 2  The Indicator Matrix $G_d$ for the inter-departure time of a two station production line with exponential servers and a finite buffer

<table>
<thead>
<tr>
<th></th>
<th>(1,0)</th>
<th>(1,1)</th>
<th>(1,2)</th>
<th>(1,3)</th>
<th>(B,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(1,1)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(1,2)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(1,3)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(B,3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3  Matrix $\Omega$ for inter-departure time of a two station production line with exponential servers and a finite buffer

<table>
<thead>
<tr>
<th></th>
<th>(1,0)</th>
<th>(1,1)</th>
<th>(1,2)</th>
<th>(1,3)</th>
<th>(B,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,0)</td>
<td>$\mu_2$</td>
<td>$\mu_2$</td>
<td>$\mu_2$</td>
<td>$\mu_2$</td>
<td>$\mu_2$</td>
</tr>
<tr>
<td>(1,1)</td>
<td>$\mu_1 + \mu_2$</td>
<td>$(\mu_1 + \mu_2)^2$</td>
<td>$(\mu_1 + \mu_2)^2$</td>
<td>$(\mu_1 + \mu_2)^2$</td>
<td></td>
</tr>
<tr>
<td>(1,2)</td>
<td>$\mu_1 + \mu_2$</td>
<td>$\mu_1 + \mu_2$</td>
<td>$(\mu_1 + \mu_2)^2$</td>
<td>$(\mu_1 + \mu_2)^2$</td>
<td></td>
</tr>
<tr>
<td>(1,3)</td>
<td>$\mu_1 + \mu_2$</td>
<td>$\mu_1 + \mu_2$</td>
<td>$\mu_1 + \mu_2$</td>
<td>$(\mu_1 + \mu_2)^2$</td>
<td></td>
</tr>
<tr>
<td>(B,3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\mu_1 + \mu_2$</td>
<td>$\mu_1 + \mu_2$</td>
</tr>
</tbody>
</table>

The mean and the variance of the inter-departure time and the covariance of the inter-departure times can be written in closed form by using the expressions for matrices $\Phi$ and $\Omega$. The closed-form expressions for the special case $\mu_1 = \mu_2 = \mu$ are given below:

$$E(T) = \frac{(M + 3)}{\mu(M + 2)},$$  \hspace{1cm} (29)

$$Var(T) = \frac{M^2 + 6M + 7}{\mu^2(M + 2)^2}.$$  \hspace{1cm} (30)
Following Equation (25), the variance of the time to produce two parts, \( \text{Var}(\Gamma_2) = \text{Var}(T_1 + T_2) \) is given as

\[
\text{Var}(\Gamma_2) = \text{Var}(T) + (E(T))^2 - (E(T_1 + T_2))^2 + 2\pi^\text{entry} X_2 \Psi_2 \psi u = \frac{2(M^2 + 6M + 6)}{\mu^2(M + 2)^2}.
\]  

(31)

The covariance of the inter-departure times with a lag of one is determined from the variance of the sum of two inter-departure times as given in Equation (21):

\[
\text{Cov}(T_1, T_2) = \frac{1}{2} (\text{Var}(\Gamma_2) - 2\text{Var}(T)) = -\frac{1}{\mu^2(M + 2)^2}.
\]  

(32)

Therefore, the correlation coefficient of inter-departure times with lag 1 is

\[
\text{Corr}(T_1) = -\frac{1}{(M^2 + 6M + 7)}.
\]  

(33)

Similarly, the covariance and the correlation of inter-departure times with a lag of 2 can be determined from Equation (21) as

\[
\text{Cov}(T_1, T_3) = -\frac{1}{\mu^2(M + 2)^2}.
\]  

(34)

As a result, the correlation coefficient of inter-departure times with lag 2 is

\[
\text{Corr}(T_2) = \frac{\text{Cov}(T_1, T_3)}{\text{Var}(T)} = -\frac{1}{(M^2 + 6M + 7)}.
\]  

(35)

Note that \( \text{Corr}(T_1) \) and \( \text{Corr}(T_2) \) are equal to each other and negative for this system. Furthermore, the correlation decreases very rapidly with the buffer capacity. The general approach for the derivations is described in Appendix C.

5. Numerical Experiments

We present two sets of numerical experiments. We first study the effects of the number of stations, the buffer capacities, the coefficient of variation of the service time and the limiting of a two-station production line with a finite buffer where the processing times have balanced Coxian \((C_{2:b})\) distribution. In the second set of experiments, we consider open and closed multi-station production lines with exponential and phase-type service time distributions. The analysis of exponential production lines allows us to validate our results with those given in the literature (Hendricks 1992) directly.
5.1. Open Two-Station Production Lines with Finite Buffers

In this part, we analyze a production system with two stations and a finite buffer with a capacity of $M$ separating these stations. The service time on each station has a balanced Coxian-2 distribution. The processing rate of station $i$ is $\mu_i$ and the coefficient of variation of the service time on station $i$ is $cv_i$.

In order to analyze the effects of the processing time variability and the buffer capacity on the autocorrelation structure, an experiment is designed with 3000 different cases corresponding to the system parameters $\mu_1 = 1$, $\mu_2 = \{0.5, 1, 1.5\}$, $cv_i = \{0.5, 0.6, .., 1.4\}$, $i = 1, 2$, and $M = \{1, 2, ..., 10\}$ using all parameter combinations. The auto-correlation function is determined analytically by using the methodology presented in this study for each case.

Figure 3 depicts the effects of the coefficient of variation of the processing times and the buffer capacity on the correlation of inter-departure times with lags 1 to 19 for the case the upstream and the downstream stations have the same processing rate ($\mu_1 = 1$, $\mu_2 = 1$). From left to right, the pictures show the correlation for an increasing coefficient of variation of the processing time at the downstream station, $cv_2$, for the values 0.5, 1 and 1.4. From top to bottom, the pictures show the correlation for an increasing coefficient of variation of the processing time at the upstream station, $cv_1$, again for the values of 0.5 over 1 to 1.4. For very low buffer capacity ($M = 1$), the correlation function subject to the number of lags converges quickly to zero whereas this takes more lags when the buffer capacity is higher.

The figures for the cases where the upstream station has a higher processing rate ($\mu_1 = 1$, $\mu_2 = 0.5$) and the downstream station has a higher processing rate ($\mu_1 = 1$, $\mu_2 = 1.5$) are given in Appendix D. The setting with the different coefficients of variations is the same here. With the downstream station being the bottleneck in Figure 3, the correlation of the inter-departure time is a lot smaller in all cases compared to those of Figure 3. In Figure 4, where the upstream station is the bottleneck, the correlations are higher for the first lags.

However, in all cases the correlation approaches zero as the number of lags $k$ increases as expected. In the case where the downstream station is slower, the correlations are quite small. Absolute values of the correlations are higher when the buffer is larger when $\mu_1 \leq \mu_2$. 
Figure 3  Correlation of Inter-departure Times with Different Lags for the Case $\mu_1 = 1$, $\mu_2 = 1$ for a Two-Station Line with Coxian Service Time Distribution and a Finite Buffer
Figure 4 presents the minimum and the maximum of the correlation of inter-departure times with different lags ranging from 1 to 19. The maximum and minimum values correspond to different cases of the downstream coefficient of variation, processing rates, and buffer capacities are considered to have an overall understanding about how the range behaves as the lag increases. As can be seen from this figure, the range of values decreases with an increasing number of lags.

In a similar way, Figure 5 shows the minimum and maximum values of the correlation of the inter-departure time for different coefficients of variation of the upstream station \( cv_1 \). The maximum and minimum values correspond to different cases of the downstream coefficient of variation, processing rates, and buffer capacities for the given \( cv_1 \). Finally, Figure 6 shows the effect of the coefficient of variation of the downstream station processing time on the correlation of inter-departure times for all cases, and Figure 7 shows the effect of the inter-station buffer capacity on the correlation of the inter-departure times including the correlations at different lags for all different system parameters. These figures show that for the given range of system parameters, the inter-departure time autocorrelation values are more likely to increase with an increasing coefficient of variation of the upstream station and increasing buffer capacity and are more likely to decrease with an increasing coefficient of variation of the downstream station. Note that the dependency of the autocorrelation function on the system parameters is not always monotonic, and are affected by the specific values of all the system parameters, see Figure 6.
5.2. Open and Closed Multi-Station Exponential Production Lines with Finite Buffers

In this section, we investigate the output processes from multi-station production lines with stations that have exponential service time distributions and finite buffers, both open and closed lines.

Figure 8 shows the effect of the number of stations and the buffer capacity on the correlation of the inter-departure time with different lags for an open multi-station line with stations that have exponential service times with the same processing rate, $\mu_i = 1$, $i = 1, \cdots, N$ and inter-station
buffers with identical capacities, $B_i = M$, $i = 1, \ldots, N - 1$. This model is analyzed by Hendricks (1992) analytically. Figure 7 depicts the effects of buffer capacity on the correlation of inter-departure times for an exponential serial line with three stations, calculated for the departing products from the first, second, and the last station, for the instances reported in Hendricks (1992). Our method yields the same results given by Hendricks (1992) for these systems. The largest model solved by Hendricks (1992) was a model for an unbuffered line of six machines that had 144 states. For buffered lines, the numerical results for three-station lines with capacities $M \leq 10$ were given. In Figure 8, the largest model reported is the one for the four-station line with a capacity of $M = 10$ which has 1338 states.

For closed queueing networks subject to blocking, there are no analytical results available. Figure 9 shows the effect of the number of stations and the buffer capacity on the correlation of the inter-departure time with different lags for a closed multi-station line. The stations all have exponential service times with the same processing rate $\mu_i = 1$, $i = 1, \ldots, k$, the inter-station buffers have identical capacities, $B_i = M$, $i = 1, \ldots, k$ and there are 2, 6, or 11 customers in the two, three, and four station systems respectively. Figure 9 indicates that having a closed queueing network introduces higher dependency for the inter-departure times especially when the buffer capacities are small. This observation is similar to the empirical studies for the autocorrelation of inter-departure times in semiconductor manufacturing (Schomig and Mittler 1995).
Figure 8  Correlation of Inter-departure Times with Different Lags for Open Multi-Station Production Lines with Exponential Servers and Finite Buffers

![Graphs showing correlation of inter-departure times with different lags for open multi-station production lines with exponential servers and finite buffers.](image)

Figure 9  Effects of the Buffer Capacity on the Correlation of Inter-departure Times with Different Lags for an Open Three-Station Production Line with Exponential Servers and Finite Buffers

![Graphs showing the effect of buffer capacity on the correlation of inter-departure times for an open three-station production line.](image)
5.3. Effect of the Number of Pallets $N$ on the Open and Closed Multi-Station Production Lines with Phase-Type Service Time Distributions and Finite Buffers

There are no analytical results available for the autocorrelation of inter-departure times for open and closed multi-station production lines with non-exponential servers. Figure 11 and Figure 12 show the correlation of inter-departure times for three- and four-station closed queueing networks with phase-type servers with different coefficient of variations and a given number of pallets $N$.

For the case with a single pallet $N = 1$, the inter-departure time distribution is equal to the cycle time distribution. The figure shows that when the number of pallets is limited compared to the total buffer capacity of the system such as the case when $N = 4$ for the four-station line with buffer capacities of $M_i = 1$, as the variability of the service time increases, i.e. the coefficient of variation increases, the correlation decreases.

6. Conclusion

In this study, we presented an analytical method to determine the correlation of inter-departure times from a queueing network subject to blocking that can be modeled as a Continuous Time Markov Chain. Our primary focus was to analyze production systems that are modeled as open or closed queueing networks with stations that have service times following phase-type distributions. Our method uses the state-transition rate matrix of the system and an indicator matrix that indicates the transitions where a departure occurs. Furthermore it utilizes the mean and variance of the time to produce a given number of products in a queueing network subject to blocking to determine the correlation of inter-departure times with different lags. We developed
The main contribution of this study is to present a general method that uses only the state-space model of a queueing network as its input to determine the correlation of inter-departure times with a given lag analytically. This method allows us to analyze the effects of system parameters such as the variability of service times, buffer capacities, number of stations on the output correlation structure analytically, as an alternative to the existing simulation studies.

The results of this study helps us to understand the effects of the system parameters such as the service time variability, buffer capacity, and the number of stations on the output process. This information can be used to design production systems in a more effective way.

Many approximate methods are based on decomposing the original production system into smaller subsystems. In these approaches, the interdependency of input and output streams is
not captured properly. By using the methodology we presented, we can investigate whether the statistical properties of a stream of products leaving a system can be captured effectively by a simpler system. This analysis can be used to develop more accurate approximation methods to determine performance measures related to the variability of the output.

For a given system with a given number of states, the number of computations in the proposed method increases polynomially with the number of states in the state space and linearly with the number of lags to be calculated. However, as it is known in the literature, the number of states in the state space increase very rapidly with the number of stations and buffers in the system. This is the limitation of all state-space based methods and it is not specific to our method. Therefore, exact solutions obtained by this method can be used to compare approximate methods with the exact results for computationally feasible systems and also to analyze the effects of system parameters.

Furthermore, this approach can be extended to build stochastic models of production systems by using observed data from the departing units from a production system. Namely, the parameters of a simple system, such as a two-station closed production line with Coxian service time distributions and a finite buffer, can be set in such a way that the statistical properties of the output process from this model, e.g. inter-departure time distribution and the autocorrelation function, is close to the observed statistical properties of the output process of a real system. If the output process of production systems can be represented satisfactorily by this model, many performance measures can be evaluated using such a small system.

Similarly, this approach can be used as a building block that exhibits similar dynamics as the statistical properties of a demand sequence of interest. For example, a correlated demand sequence can be used as an input to the models of the production systems by embedding this subsystem as a demand generator block. By using these blocks, the effects of modelling arrival and demand streams with including or ignoring the auto-correlation structure on the performance of a system can be analyzed for a production-inventory system. These extensions are left for future research.

References


Appendix. Proofs of Statements and Additional Numerical Results

A. Proof of Proposition I

The infinitesimal generator submatrix for the transitions from the transient states to other transient states \( Q_d^{(k)} \) is given as

\[
Q_d^{(k)} = \begin{bmatrix}
Q_d & R_d & 0 & 0 & 0 & \ldots & 0 \\
0 & Q_d & R_d & 0 & 0 & \ldots & 0 \\
0 & 0 & Q_d & R_d & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \ldots & Q_d
\end{bmatrix}.
\]

By using the upper triangular structure of the matrix \( Q_d^{(k)} \), its inverse can be written as follows

\[
(Q^{(k)})^{-1} = \begin{bmatrix}
\Phi & \Phi\Omega & \Phi\Omega^2 & \Phi\Omega^3 & \Phi\Omega^4 & \ldots & \Phi\Omega^{k-1} \\
0 & \Phi & \Phi\Omega & \Phi\Omega^2 & \Phi\Omega^3 & \ldots & \Phi\Omega^{k-2} \\
0 & 0 & \Phi & \Phi\Omega & \Phi\Omega^2 & \ldots & \Phi\Omega^{k-3} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \ldots & \Phi
\end{bmatrix}
\]

(37)

where \( \Phi = Q_d^{-1} \) and \( \Omega = -R_d\Phi \). Since the process enters in one of the states where the first part departs, the steady-state probability vector of entering into the process \( \pi_{\text{entry}}^{(k)} \) has the following structure

\[
\pi_{\text{entry}}^{(k)} = \begin{bmatrix}
\pi_{\text{entry}}^{(k)} 0 0 0 \ldots 0
\end{bmatrix}.
\]

(38)

where in the above equation \( 0 \) is a \( 1 \times |S| \) row vector of zeros.

Therefore the expected value of the sum of \( k \) inter-departure times \( E(\Gamma_k) \) can be written directly as

\[
E(\Gamma_k) = -\pi_{\text{entry}}^{(k)}(Q_d^{(k)})^{-1}u^{(k)} = \pi_{\text{entry}}^{(k)} \sum_{i=1}^{k} \Phi\Omega^{i-1}u.
\]

(39)

Since the expected value of the sum of \( k \) inter-departure times is the sum of the expected inter-departure times,

\[
E(\Gamma_k) = kE(T) = -k\pi_{\text{entry}}^{(k)}\Phi u.
\]

(40)

Similarly, multiplying \((Q^{(k)})^{-1}\) given above with itself and calculating \( \pi_{\text{entry}}^{(k)}(Q_d^{(k)})^{-2}u^{(k)} \) in a similar way yields the variance of the sum of \( k \) inter-departure times \( Var(\Gamma_k) \):

\[
\pi_{\text{entry}}^{(k)}(Q_d^{(k)})^{-2}u^{(k)} = \pi_{\text{entry}}^{(k)} \sum_{i=1}^{k} \sum_{j=1}^{i} \Phi\Omega^{i-1}\Phi\Omega^{j-1}u.
\]

(41)

Therefore,

\[
Var(\Gamma_k) = 2\pi_{\text{entry}}^{(k)}(Q_d^{(k)})^{-2}u^{(k)} - (E(\Gamma_k))^2
\]

(42)
Finally, since 

\[ \sum_{i=1}^{k} \Phi \Omega^{-1} \Phi \Omega^{i-j} u = (E(\Gamma_k))^2 \]  

\( k \geq 2, \ldots \), and adding the above equations in the order of calculations, i.e. calculating first Equation (46), then Equation (47), followed by Equation (43) for \( k = 2, \ldots, m \) yield \( \text{Corr}(T)_1, \ldots, \text{Corr}(T)_{(m-1)}. \quad \square \)
C. Derivation of the Closed-Form Performance Measures for Two Station Production Line with Exponential Servers and a Finite Buffer

When \( \mu_1 = \mu_2 = \mu \), the elements of matrices \( \Phi \) and \( \Omega \) can be written as a function of \( M \):

\[
\Phi = -\frac{1}{\mu} \begin{bmatrix}
1 & 1 & \frac{1}{2} & \cdots & \frac{1}{2^{M+1}} \\
0 & 1 & \frac{1}{2} & \cdots & \frac{1}{2^{M+1}} \\
0 & 0 & 1 & \cdots & \frac{1}{2^{M+1}} \\
0 & 0 & 0 & \cdots & \frac{1}{2} \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}, \tag{52}
\]

\[
\Omega = \frac{1}{\mu} \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
1 & 1 & \frac{1}{2} & \cdots & \frac{1}{2^{M+1}} \\
0 & 1 & \frac{1}{2} & \cdots & \frac{1}{2^{M+1}} \\
0 & 0 & 1 & \cdots & \frac{1}{2^{M+1}} \\
0 & 0 & 0 & \cdots & \frac{1}{2} \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}, \tag{53}
\]

Furthermore, since \( \mu_1 = \mu_2 \), it is equally likely to enter and exit to all the states excluding the state \((B, M+1)\) where the first station is blocked. It is not possible to enter a departure cycle in this state since all the departures from the second station eliminates the blocking of the first station. Therefore the steady-state entry probability vector \( \pi_{entry} \) can also be expressed as a function of \( M \). Since there are \( M+3 \) states and it is equally likely to enter in one of the \( M+2 \) states, the steady-state entry probability vector \( \pi_{entry} \) is

\[
\pi_{entry} = \frac{1}{M+2} \begin{bmatrix}1 & 1 & 1 & \cdots & \cdots & 1 & 1 \end{bmatrix}. \tag{54}
\]

From these equations, we can show that

\[
\Phi u = -\frac{1}{\mu} \begin{bmatrix}2 & 1 & 1 & \cdots & \cdots & 1 & 1 \end{bmatrix}^T. \tag{55}
\]

The direct product of the vectors \( \pi_{entry} \) and \( \Phi u \) yields

\[
E[T] = -\pi_{entry} \Phi u = \frac{(M+3)}{\mu(M+2)}. \tag{56}
\]

For the variance of the inter-departure time, we first determine the term \( 2\pi_{entry} \Psi^2 u = 2\pi_{entry} \Psi_1 \Psi_1 u \) as the product of the vector \( 2\pi_{entry} \Psi_1 \) and the vector \( \Psi_1 u \).

We can show that

\[
\pi_{entry} \Psi_1 = \frac{1}{M+2} \begin{bmatrix}1 & 1 & 1 & \cdots & \cdots & 1 \end{bmatrix}. \tag{57}
\]

Since \( \Psi_1 u = \Phi u \) is given in Equation (55), the direct vector multiplication and algebraic simplifications yield

\[
Var(T) = 2\pi_{entry} \Psi_1 \Psi_1 u - (E(T))^2 = \frac{(M^2 + 6M + 7)}{\mu^2(M+2)^2}. \tag{58}
\]

In order to determine \( Var(\Gamma_k) \) for \( k = 2, 3 \) we use the recursive equation

\[
Var(\Gamma_k) = Var(\Gamma_{k-1}) + (1-2k)(E(T))^2 + 2\pi_{entry} \Psi_k u. \tag{59}
\]
We show that \( \pi_{\text{entry}} \Upsilon_2 = \pi_{\text{entry}} \Upsilon_3 = \pi_{\text{entry}} \Upsilon_1 \) where \( \pi_{\text{entry}} \Upsilon_1 \) is given in Equation (57). Furthermore
\[
\Phi\Omega u = -\frac{1}{\mu} \left[ 1.5 \ 1.5 \ 1 \ \cdots \ 1 \ 1 \right]^T
\]
that yields
\[
\Psi_{2u} = \Phi u + \Phi\Omega u = -\frac{1}{\mu} \left[ 3.5 \ 2.5 \ 2 \ \cdots \ 2 \ 2 \right]^T.
\]
(61)

As a result, the vector product of \( \pi_{\text{entry}} \Upsilon_2 \) and \( \Psi_{2u} \) yields
\[
2\pi_{\text{entry}} \Upsilon_2 \Psi_{2u} = \frac{4(M + 4)}{\mu^2(M + 2)}.
\]
(62)

Equation (61) yields \( Var(T_1 + T_2) \) for \( k = 2 \). Similarly, we show that
\[
2\pi_{\text{entry}} \Upsilon_3 \Psi_{3u} = \frac{6(M + 4)}{\mu^2(M + 2)}
\]
(63)

that gives \( Var(T_1 + T_2 + T_3) \) through Equation (61) for \( k = 3 \).

Once \( Var(T) \), \( Var(T_1 + T_2) \), and \( Var(T_1 + T_2 + T_3) \) are determined \( Cov(T_1, T_2) \), \( Cov(T_1, T_3) \), \( Corr(T_1) \), and \( Corr(T_2) \) are determined directly.

D. Correlation of Inter-departure Times with Different Lags for a Two-Station Line with Coxian Service Time Distribution and a Finite Buffer
Figure 13 Correlation of Inter-departure Times with Different Lags for the Case $\mu_1 = 1$, $\mu_2 = 0.5$ for a Two-Station Line with Coxian Service Time Distribution and a Finite Buffer.
Figure 14  Correlation of Inter-departure Times with Different Lags for the Case $\mu_1 = 1$, $\mu_2 = 1.5$ for a Two-Station Line with Coxian Service Time Distribution and a Finite Buffer