A *BMAP/BMSP/1* QUEUE WITH MARKOV DEPENDENT
ARRIVAL AND MARKOV DEPENDENT SERVICE BATCHES

ACHYUTHA KRISHNAMOORTHY
Centre for Research in Mathematics
C.M.S. College
Kottayam-686001, India

ANU NUTHAN JOSHUA
Department of Mathematics
Union Christian College
Aluva-683102, India

(Communicated by Bara Kim)

Abstract. Batch arrival and batch service queueing systems are of importance in the context of telecommunication networks. None of the work reported so far consider the dependence of consecutive arrival and service batches. Batch Markovian Arrival Process (*BMAP*) and Batch Markovian Service Process (*BMSP*) take care of the dependence between successive inter-arrival and service times, respectively. However in real life situations dependence between consecutive arrival and service batch sizes also play an important role. This is to regulate the workload of the server in the context of service and to restrict the arrival batch size when the flow is from the same source. In this paper we study a queueing system with Markov dependent arrival and service batch sizes. The arrival and service batch sizes are assumed to be finite. Further, successive inter-arrival and service time durations are also assumed to be correlated. Specifically, we consider a *BMAP/BMSP/1* queue with Markov dependent arrival and Markov dependent service batch sizes. The stability of the system is investigated. The steady state probability vectors of the system state and some important performance measures are computed. The Laplace-Stieltjes transform of waiting time and idle time of the server are obtained. Some numerical examples are provided.

1. Introduction. Batch arrival and batch service queueing systems have wide application in telecommunication networks. Most networks have a non stationary(bursty), self similar input flow wherein the inter-arrival times are highly correlated. Batch Markovian Arrival Process (*BMAP*) can accurately model such input flows. Also *BMAP* form a generalization of Markovian processes and hence are analytically tractable. Like *BMAP*’s, to model those service processes in which the customers are served in batches of random size with service times exhibiting high degree of correlation, Batch Markovian Service Process (*BMSP*) is used. The existing literature on queueing systems is abundant with cases in which both the

2010 Mathematics Subject Classification. Primary: 58F15, 58F17; Secondary: 53C35.

Key words and phrases. Queueing system, batch Markovian arrival process, batch Markovian service process, Markov dependent arrival/service batch sizes.

The first author is supported by UGC, Govt. of India, Emeritus Fellow(EMERITUS 2017-18 GEN 10822(SA-II)) and DST, Indo-Russian project: INT/RUS/RSF/P-15.
arrival and service processes are correlated. However, to the best of our knowledge, there is another kind of dependence that may arise while modelling batch arrival and batch service queueing systems - i.e., the dependence between successive arrival and service batch sizes, which has not received attention so far. In this paper we try to address this deficit by considering a BMAP/BMSP/1 queue with Markov dependent arrival and Markov dependent service batch sizes.

**Highlights of this paper are:**

- The arrival and service batch sizes are governed by independent Batch Markovian Processes.
- The successive arrival batch sizes are determined by a finite state Markov chain. This is the case with successive service batch sizes as well.
- As the successive service batch sizes are determined by Markov chain rule, the idle time of the server could increase for want of quorum for the next batch service.
- The conditional Laplace Steiltjes transforms of idle time in different states and conditional Laplace Steiltjes transforms of waiting times of a tagged customer in queue are derived.

Markov dependence amongst successive arrival and service batch sizes could be effectively used to restrict arrival batch sizes and to regulate the workload of the server in batch arrival and batch service queueing systems. Consider an organisation with a large fleet of vehicles. These vehicles need to be sent periodically for maintenance to optimise their performance. To adjust the requirement of available vehicles, the organisation sends vehicles in batches of varying sizes for service. The best way this could be done is to have Markov dependence between successive batches dispatched - if a large batch of vehicles is now dispatched, the next batch to be dispatched is taken sufficiently small. Similarly, at the service station (which could also be assumed to be part of the organisation), service is given in batches of varying sizes. To adjust the workload between successive batches of service, a Markov chain rule for determination of the number of vehicles in each batch is adopted. As another example consider a manufacturing plant, wherein jobs come in batches of varying sizes from various sources. Over time, by studying the previous demands a Markov chain rule could be established for selection of successive arrival batch sizes. Quite often, goods are produced in batches of varying sizes. By introducing a Markov chain rule amongst successive production batch sizes, the workload of machines in the plant could be regulated. The motivation for considering arrival and service batch sizes that are Markov dependent could be extended to telecommunication networks. Messages arrive in batches of varying sizes at the telecommunication centre. These are then transmitted in batches of varying sizes. A Markov chain rule for selection of successive arrival and service batch sizes could greatly improve the efficiency of the system.

Neuts (1979) introduced versatile Markovian point process (VMPP) [15] as a generalization of Poisson process. The name was changed to Neuts process and a detailed analysis of N/G/1 queue using matrix analytic methods was done by Ramaswami [18]. The MAP in the current formulation was introduced by Lucantoni, Hellstern and Neuts [14]. Lucantoni generalized MAP to BMAP in [13]. Since then numerous research papers with BMAP as arrival process has appeared. A detailed survey of literature devoted to studying queueing systems with Markovian and Batch Markovian arrival processes and their applications to modelling telecommunication networks can be found in [21].
Some of the recent papers with BMAP as arrival process are as follows. Kim et al [8] analyze a BMAP/PH/N type queue where the servers are subject to breakdown (which occur according to MAP) and are repaired immediately with repair period following PH distribution. The customer whose service is interrupted occupies an idle server or moves to the buffer with some fixed probability and leaves the system with complementary probability. In [12], Klimenok et al consider a multi-server queueing system with infinite buffer, N independent main servers, R independent backup servers where the arrival process is BMAP, service times of main as well as backup servers follow PH distribution. Here the backup server helps the main server whenever the service time of some customer exceeds a certain random bound. The behaviour of this queueing system is studied as a continuous time Markov chain. An infinite-buffer single server queue with BMAP arrival process and exhaustive multiple working vacation policy is analysed in [3] by Banik. The duration of service time during a working vacation is generally distributed, while during a normal busy period, it is having R-type distribution (distribution possessing rational Laplace-Stieltjes transform). In [7], Ghosh et al analyze a finite-buffer single server vacation queue with BMAP arrival process. The server serves customers according to a non exhaustive type gated-limited service discipline. It is assumed that the service and vacation durations possess rational Laplace-Stieltjes transform. A partial batch acceptance strategy is used and the queue length distributions at arbitrary, pre - arrival and post - departure epochs are obtained. A net profit optimization problem is solved in this paper. Retrial tandem queues with single server Station 1 and multiple server Station 2 are analyzed by Klimenok et al in papers [11] and [10]. The input flow to station 1 is BMAP and if the arriving customer meets a busy server it goes to an orbit and repeat attempts after an exponentially distributed time and in random time interval, respectively. The service time at Station 1 in [11] is Semi-Markovian, while in [10], it is generally distributed. The second station is a multi-server queue and the customer requires a combination of servers to complete service, with service times at each server exponentially distributed. The main difference however is that in [10], the system has two operation modes which are distinguished by a strategy of repeated attempts. The problem of computation of moments of queue length in BMAP/SM/1 queue is considered in Klimenok et al [9]. Rama et al [17] consider two BMAP arrival and bulk service C - server queues with k varying environments.

There has been very few research papers with BMSP as service process. Banik et al [4] analyze a finite-buffer single-server queue with renewal input and service provided in batches of random size according to BMSP using a combination of embedded Markov chain and supplementary variable methods. A finite-buffer single server queue where arrivals occur according to BMAP and where the server serves customers in batches of maximum size b with a minimum threshold size a (or general bulk service rule) is studied in Banik[2]. Here the service time of each batch follows general distribution independent of each other as well as the arrival process. Also, he considers a BMAP/MSP/1 queue with the same bulk service rule in this paper. Further in [1], he studies GI/BMSP/1/∞ with state dependent arrivals. In state-dependent arrival to a queueing system, the rate of arrival of jobs at the server may depend on the amount of work present in the system. Through state-dependent arrival, one may control the arrival of jobs to optimize the server performance.
Queueing systems with BMAP arrivals and correlated service processes are studied in following papers. Samanta et al [19] analyze BMAP/MSP/1 queueing system based on roots of the associated characteristic equation of the vector generating function of system-length distribution at random epoch. The steady-state system-length distributions at various epochs as well as of the actual sojourn-time distribution of an arbitrary customer in an arriving batch is derived in this paper. Ghosh et al [5] studies a BMAP/MSP/1 – GPS using RG-factorization technique applied to the Markov chain of the associated quasi-birth and death process. In a GPS (Generalized Processor Sharing) queue, if at any instant there are \(n\) number of customers present in the system then each customer experiences a service rate \(f(n)\) at that instant, where \(f(.)\) is a positive function which satisfies \(0 < nf(n) \leq C\) for all \(n(\geq 1)\) and \(C\) is a positive constant. The conditional sojourn time of a randomly chosen tagged customer in a BMAP/MSP/1 – ROS (Random order Service) queueing system is studied in Ghosh et al [6]. Sandhya et al [20] analyze a BMAP/BMSP/1 queue with randomly varying environment where the environment changes as per changes in a continuous time Markov chain and derive stationary queue length probabilities and other performance measures for the same. They consider 2 models: a) the maximum of arrival sizes is greater than maximum of service sizes b) the maximum of arrival sizes is less than maximum of service sizes. It may be noted that block circulant structure is noticed in the basic system infinitesimal generators.

In this paper, we consider a queueing system in which the arrival process is BMAP with consecutive arrival batch sizes forming a finite first order Markov chain. Service is provided according to BMSP with successive batch service sizes determined by a finite first order Markov chain. Whenever there are not enough customers to initiate the next service at a service completion epoch (as determined by Markov chain rule) the service process is freezed and the server remains idle. Also, it is assumed that the maximum arrival and service batch sizes are finite.

This paper is organised as follows. The mathematical model is described in Section 2. Section 3 deals with the steady state analysis of the queueing system. The Laplace - Stieltjes transforms for the idle time and waiting time are provided in Sections 4 and 5, respectively. Other performance measures computed are presented in Section 6 and numerical results are presented in Section 7. Section 8 provides the conclusions of the present study.

Notations and abbreviations used in the sequel are:
- \(e\) : column vector of 1’s of appropriate order
- \(e(j)\): column vector of 1’s of order \(j\)
- \(e_i\): column vector of appropriate order with 1 in \(i^{th}\) position and 0 elsewhere
- \(e'(j)\): transpose of \(e(j)\)
- \(e_i'(j)\): column vector of order \(j\) with 1 in \(i^{th}\) position and 0 elsewhere
- \(e'(j)\): transpose of \(e_i'(j)\)
- \(0\): zero matrix of appropriate order
- \(I\): identity matrix of appropriate order
- \(I_r\): identity matrix of order \(r\)
- CTMC: Continuous time Markov Chain
- LIQBD: Level independent quasi birth and death process
- tpm: Transition probability matrix
- LST: Laplace - Stieltjes transform
- TC: Tagged customer
2. The mathematical model. We assume that customers arrive at a single server queueing system according to a BMAP with maximum arrival batch size \( a \). The successive arrival batch sizes form a first order Markov chain \( \{X_n; n \geq 1\} \) with one step \( tpm P = [p_{ij}] \) on state space \( \{1, 2, 3 \ldots a\} \). Service is provided according to BMSP with maximum service batch size \( b \). The successive service batch sizes form a first order Markov chain \( \{Y_n; n \geq 1\} \) with one step \( tpm Q = [q_{ij}] \) on state space \( \{1, 2, 3 \ldots b\} \). If there are not sufficient number of customers in the system to initiate service as per the service batch size Markov chain rule then the service process is freeze. i.e., if according to the Markov chain rule, \( k \) customers are required to initiate the next service and these many customers are not available at the service completion epoch, the server waits till the number of customers in the system reaches \( k \). In such cases we use \( 0(k) \) to denote the status of the service process.

The next arrival batch size is determined by the Markov chain rule with \( tpm P \). Hence if the last arrival batch is of size \( i \), the next arrival batch is of size \( j \), with probability \( p_{ij} \). The arrival process is defined using two stochastic matrices \( D_0 \) and \( D_c \) of order \( r \) where \( D_0 \) denote transition rates of the underlying Markov chain of BMAP without arrival and \( D_c \) denote transition rates of the underlying Markov chain of BMAP with batch arrivals. i.e., the transition rates of the underlying Markov chain, if the last arrival batch size is \( i \) and next arrival of batch size \( j \), is specified by the matrix \( p_{ij}D_c \). Similarly the size of the next batch to be served is determined by the Markov chain with \( tpm Q \). The service process is described using two stochastic matrices \( S_0 \) and \( S_d \) of order \( s \). \( S_0 \) denote transition rates of the underlying Markov chain of BMSP without departures. \( S_d \) denote transition rates of the underlying Markov chain of BMSP with bulk departures. i.e., the transition rates of the underlying Markov chain, if the current service batch size is \( i \) and next batch to be served is of size \( j \), is specified by the matrix \( q_{ij}S_d \).

The queueing model described above can be studied as a CTMC,

\[
\{\{(N(t), A(t), B(t), J(t), I(t)), t \geq 0\} \ |
\begin{align*}
N(t) & \text{ - number of customers in the queue at time } t \\
A(t) & \text{ - size of the batch that arrived before time } t \\
B(t) & \text{ - size of the service batch at time } t \\
J(t) & \text{ - the state of underlying MC of BMSP} \\
I(t) & \text{ - the state of underlying MC of BMAP}
\end{align*}
\]

on state space

\[
\Omega_1 = \{(0, n_1, n_2, j, i) : 1 \leq n_1 \leq a; n_2 = 0(1), 0(2) \ldots 0(b), 1, 2, 3 \ldots b; 1 \leq j \leq s; 1 \leq i \leq r \} \cup \{(n, n_1, n_2, j, i) : n \geq 1; 1 \leq n_1 \leq a; n_2 = 0(n + 1), 0(n + 2), \ldots 0(b), 1, 2, 3 \ldots b; 1 \leq j \leq s; 1 \leq i \leq r \}.
\]

As an illustration, the infinitesimal generator of the CTMC when \( a = 2, b = 3 \) is

\[
Q_1 = \\
\begin{bmatrix}
B_{00} & B_{01} & B_{02} & B_{03} & B_{04} & B_{05} & B_{06} & B_{07} \\
B_{10} & B_{11} & B_{12} & B_{13} & B_{14} & B_{15} & B_{16} & B_{17} \\
B_{20} & B_{21} & B_{22} & B_{23} & B_{24} & B_{25} & B_{26} & B_{27} \\
C_3 & C_2 & C_1 & F & G_1 & G_2 & C_3 & C_2 \\
C_3 & C_2 & C_1 & F & G_1 & G_2 & C_3 & C_2 \\
C_3 & C_2 & C_1 & F & G_1 & G_2 & C_3 & C_2 \\
C_3 & C_2 & C_1 & F & G_1 & G_2 & C_3 & C_2 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\end{bmatrix}
\]
The infinitesimal generator of the CTMC is not LIQBD. This can be made a LIQBD by redefining the level \(N(t)\) of the system at \(t\).

Let \(q = \max\{a, b\}\). Suppose \(N(t) = n\). Then, \(n = lq + p, 0 \leq p \leq q - 1\) (the remainder modulo \(q\)). So, \(n\) could be replaced by an ordered pair \((l, p); l \geq 0, 0 \leq p \leq q - 1\). The resulting CTMC is a LIQBD with state space, 

\[ \Omega_2 = \{(0, p, n_1, n_2, j, i) : 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; n_2 = 0(p+1), 1, 2, 3...b; 1 \leq j \leq s; 1 \leq i \leq r\} \cup \{(l, p, n_1, n_2, j, i) : l \geq 1; 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; 1 \leq n_2 \leq b; 1 \leq j \leq s; 1 \leq i \leq r\} \]

and the generator matrix changes to \(Q_2\), where

\[ Q_2 = \begin{bmatrix}
A_{00} & A_{01} \\
A_{10} & A_1 & A_0 \\
A_2 & A_1 & A_0 \\
\vdots & \vdots & \vdots \\
A_1 & A_0 & \vdots \\
A_0 & \vdots & \vdots \\
\end{bmatrix} \]

where \(A_2, A_1, A_0\) are matrices of order \(tq, t = r \times s \times a \times b\) of the form

\[ A_2 = \begin{bmatrix}
C_q & C_{q-1} & C_{q-2} & \cdots & C_1 \\
C_q & C_{q-1} & \cdots & C_2 \\
\vdots & \vdots & \ddots & \vdots \\
C_q & \cdots & \cdots & C_q \\
\end{bmatrix} \]

\[ C_i = I_a \otimes Q \times E_{ii} \otimes (S_d \otimes I_r) \]

\[ C_i = 0, i > b \]

\[ E_{ii} = e_i(b) \times e_i'(b) \]

\[ A_1 = \begin{bmatrix}
F & G_1 & G_2 & \cdots & G_{q-1} \\
C_1 & F & G_1 & \cdots & G_{q-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
C_{q-1} & C_{q-2} & C_{q-3} & \cdots & F \\
\end{bmatrix} \]

\[ F = I_{ab} \otimes (S_0 \oplus D_0) \]

\[ A_0 = \begin{bmatrix}
G_q \\
G_{q-1} & G_q \\
\vdots & \vdots & \ddots \\
G_1 & G_2 & G_3 & \cdots & G_q \\
\end{bmatrix} \]

\[ G_j = P \times \bar{E}_{jj} \otimes I_s \otimes (I_s \otimes D_c) \]

\[ G_j = 0, j > a, \]

\[ \bar{E}_{jj} = e_j(a) \times e_j'(a) \]

\(A_{00}\) is a matrix of order \(t(q + \frac{b+1}{2})\) whose transition sub matrices and the corresponding rates are given in Table 1.

\(A_{01}\) and \(A_{10}\) are matrices of order \(t(q + \frac{b+1}{2}) \times tq\) and \(tq \times t(q + \frac{b+1}{2})\) respectively. The first two rows of the Table 2 indicate the transition sub matrices and the corresponding rates for \(A_{01}\) and \(A_{10}\) respectively, while others give the transition sub matrices and the corresponding rates from level 1 to any other level.
To Rate

| From       | To                | Rate      |
|------------|-------------------|-----------|
| (0, p, n₁, n₂ = 0(k)) | (0, p, n₁, n₂ = 0(k)) | Iᵢ ⊗ D₀   |
| (0, p, n₁, n₂)          | (0, p, n₁, n₂)       | S₀ ⊕ D₀   |
| (0, p, m₁, n₂)          | (0, p + m₁, m₁, n₂)  | Iₓ ⊗ pₙ₁m₁Dₑ |
| (0, p, m₁, n₂)          | (0, p + m₁ - k, m₁, k) | Iₓ ⊗ pₙ₁m₁Dₑ |
| (0, p, n₁, n₂ = 0(k))   | (0, p + m₁, 0(k))    | Iₓ ⊗ pₙ₁m₁Dₑ |
| (0, p, n₁, n₂)          | (0, p, n₁, m₂ = 0(k)) | qₙ₂mₛDᵣ ⊕ Iᵣ |
| (0, p, n₁, n₂)          | (0, p - m₂, n₁, m₂)  | qₙ₂mₛDᵣ ⊕ Iᵣ |

Table 1. Transition rate submatrices within level 0

| From       | To                | Rate      |
|------------|-------------------|-----------|
| (0, p, n₁, n₂)          | (1, p + m₁ - q, m₁, n₂) | Iₛ ⊗ pₙ₁m₁Dₑ |
| (1, p, n₁, n₂)          | (0, q + m₂ - n₁, m₂)   | qₙ₂mₛDᵣ ⊕ Iᵣ |
| (l, p, n₁, n₂)          | (l - 1, p + m₂, n₁, n₂) | qₙ₂mₛDᵣ ⊕ Iᵣ |
| (l, p, n₁, n₂)          | (l, p + m₁, n₁, n₂)    | Iₛ ⊗ pₙ₁m₁Dₑ |
| (l, p, n₁, n₂)          | (l, p - m₂, n₁, n₂)    | qₙ₂mₛDᵣ ⊕ Iᵣ |
| (l, p, n₁, n₂)          | (l + 1, p + m₁ - q, m₁, n₂) | Iₛ ⊗ pₙ₁m₁Dₑ |
| (l, p, n₁, n₂)          | (l, p, n₁, n₂)         | S₀ ⊕ D₀   |

Table 2. Transition rate submatrices except those within level 0

3. Steady state analysis.

3.1. Stability condition. Let \( \boldsymbol{π} = (\boldsymbol{π₁}, \boldsymbol{π₂}, \ldots, \boldsymbol{πₚ}) \) denote the steady state probability vector of the generator matrix \( \mathbf{A} = \mathbf{A₀} + \mathbf{A₁} + \mathbf{A₂} \)

\[
\mathbf{A} = \begin{bmatrix}
F + C_q + G_q & C_{q-1} + G_{1} & C_{q-2} + G_{2} & \cdots & C_{1} + G_{q-1} \\
C_{1} + G_{q-1} & F + C_q + G_q & C_{q-1} + G_{1} & \cdots & C_{2} + G_{q-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
C_{q-1} + G_{1} & C_{q-2} + G_{2} & C_{q-3} + G_{3} & \cdots & F + C_q + G_{q}
\end{bmatrix}
\]

This matrix is block circulant and hence solution to equations \( \boldsymbol{πA} = \mathbf{0} \) and \( \boldsymbol{πe} = 1 \) is given by \( \boldsymbol{π} = \frac{1}{q}(\mathbf{e}'(q) \otimes \mathbf{v}) \) where \( \mathbf{v} \) is a vector of order \( t \) satisfying the system of equations \( \mathbf{v}[F + C_q + G_q + C_{1} + G_{q-1} + \ldots C_{q-1} + G_{1}] = \mathbf{0} \).

In the special case, when \( p_{ij} = \frac{1}{a} \) and \( q_{ij} = \frac{1}{b} \); \( \forall i, j, \boldsymbol{π} = \frac{1}{abq} (\mathbf{e}'(abq) \otimes \mathbf{v'}) \) where \( \mathbf{v'} \) is a row vector of order \( rs \) satisfying \( \mathbf{v'}(S \oplus D) = \mathbf{0} \). \( S = S₀ + S_d \) is the generator matrix of the underlying MC of BMAP and \( D = D₀ + D_c \) is the generator matrix of the underlying MC of BMSP.

The \( LIQBD \) description of the model indicates that the queueing system is stable (see Neuts[16]) if and only if the left drift exceeds that of right drift. That is, \( \boldsymbol{πA₀e} < \boldsymbol{πA₂e} \).

Lemma 3.1. The given system is stable iff \( \mathbf{v} \cdot \sum_{k=1}^{q} kG_k. e(t) < \mathbf{v} \cdot \sum_{k=1}^{q} kC_k. e(t) \).

3.2. Steady state probability vector. Let \( \mathbf{x} \) be the steady state probability vector of \( Q₂ \). We partition this vector as

\[
\mathbf{x} = (\mathbf{x₀}, \mathbf{x₁}, \mathbf{x₂} \ldots)
\]
where \( x_0 \) is a row vector of dimension \( t(q + \frac{b+1}{2}) \), \( x_1, x_2, \ldots \) are of dimension \( t \). Under the stability condition (see Neuts[16]), we have

\[
x_i = x_1 R^{i-1}, i \geq 2
\]

where the matrix \( R \) is the minimal nonnegative solution to the matrix quadratic equation,

\[
R^2 A_2 + RA_1 + A_0 = 0.
\]

The vectors \( x_0 \) and \( x_1 \) are obtained by solving the equations

\[
x_0 A_{00} + x_1 A_{10} = 0,
\]

\[
x_0 A_{01} + x_1 (A_1 + RA_2) = 0.
\]

subject to the normalizing condition \( x_0 e + x_1 (I - R)^{-1} e = 1 \).

4. **Idle time analysis.** In the queueing system under consideration, the server can go idle even when there are customers in queue, as in queues with bulk service rule. This situation arises when the next service batch size, as determined by the Markov chain rule, exceeds the number of customers waiting at the departure epoch. The service process is freezeed till the requisite number is reached. Specifically, if there are \( p \) customers in queue and if the Markov chain rule specifies that the service is initiated only when \( n_2 \) customers accumulate with \( p < n_2 \), then server stays idle. The Laplace - Stieltjes transform of idle time when the system is in different states is computed here.

Suppose the process is in one of the states given by \( E_1 = \{ (0, p, n_1, n_2, j, i) : 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; n_2 = 0(p + 1); 1 \leq j \leq s; 1 \leq i \leq r \} \). The idle time is the time till one more customer arrives to the system. Thus the idle time in the given state is interrupted by the arrival of at least 1 customer to the system in a batch.

The probability that the arrival of a customer happens in \( (x, x + dx) \) with the arrival phase changing to \( i' \), is given by the \( (i, i') \) entry of the matrix \( e^{D_0 x} p_n D_c dx \). The \( LST \) of idle time interrupted by the arrival of a batch of size 1 (which results in the arrival phase changing to \( i' \) ) is given by the \( (i, i') \) entry of the matrix

\[
\int_0^\infty e^{-s x} e^{D_0 x} p_n D_c dx = (s I - D_0)^{-1} p_n D_c.
\]

The idle time can be interrupted by the arrival of a batch of size 1 or 2 or 3 or...a which results in phase changing to \( i' \). Hence the \( LST \) of idle time, given the system is in \( E_1 \) and the arrival phase changes to \( i' \) at the end of the idle time, denoted as \( I^*(s/E_1) \), is the \( (i, i') \) entry of the matrix, \( H_1 = (s I - D_0)^{-1} (p_{n_1} + p_{n_2} + \ldots D_c) = (s I - D_0)^{-1} D_c \).

Suppose the process is in one of the states given by \( E_2 = \{ (0, p, n_1, n_2, j, i) : 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; n_2 = 0(p + 2); 1 \leq j \leq s; 1 \leq i \leq r \} \). The idle time is the time till two more customers arrive to the system. Thus the idle time in the given state is interrupted by the arrival of at least 2 customers to the system in a batch or the arrival of a customer batch of size exactly 1, followed by the arrival of a customer batch of size at least 1. In the latter case, with the arrival of a customer batch of size 1, the system moves to state \( E_1 \) and the idle time is time till a batch of size 1 arrives plus idle time in \( E_1 \). Hence the \( LST \) of idle time, given the system is in \( E_2 \) and the arrival phase changes to \( i' \) at the end of the idle time, denoted as \( I^*(s/E_2) \) is the \( (i, i') \) entry of the matrix \( H_2 = (s I - D_0)^{-1} (p_{n_1} + p_{n_2} + \ldots D_c) = (s I - D_0)^{-1} D_c H_1 \).

Thus in general **suppose the system is in state** \( E_h = \{ (0, p, n_1, n_2, j, i) : 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; n_2 = 0(p + h); 1 \leq j \leq s; 1 \leq i \leq r \} **and the arrival phase changes to \( i' \) at the end of the idle time in that state.** The idle time is the time till \( h \) more customers arrive to the system.
Case 1: $1 \leq h \leq a$

The idle time is time till at least $h$ customers arrive to the system in a batch or time till exactly 1 customer arrives to the system in a batch plus idle time in $E_{h-1}$ ... or time till exactly $h-1$ customers arrive to the system in a batch plus idle time in $E_1$. Hence the LST of idle time, given the system is in $E_h$ and the arrival phase changes to $i'$ at the end of the idle time in that state, denoted as $I^*{(s/E_h)}$ is the $(i, i')$ entry of the matrix,

$$H_h = \frac{(p_{n1} + \cdots + p_{n1})D_c}{sI - D_0} + \frac{(p_{n2})D_c}{sI - D_0}H_{h-1} + \cdots + \frac{(p_{n1-h})D_c}{sI - D_0}H_1.$$

Case 2: $a + 1 \leq h \leq b$

The idle time is time till exactly 1 customer arrives to the system in a batch plus idle time in $E_{h-1}$ ... or time till exactly $a$ customers arrive to the system in a batch, plus idle time in $E_{h-a}$. Hence the LST of idle time, given the system is in $E_h$ and the arrival phase changes to $i'$ at the end of the idle time in that state, denoted as $I^*{(s/E_h)}$ is the $(i, i')$ entry of the matrix,

$$H_h = \frac{(p_{n2})D_c}{sI - D_0}H_{h-1} + \cdots + \frac{(p_{n1})D_c}{sI - D_0}H_{h-a}.$$

5. Analysis of waiting time. The waiting time of the tagged customer in queue, upon his arrival to the system in a batch of $n_1$ customers, is analysed and Laplace-Stieltjes transforms for the same is computed in this section. When the service process is freezed and there are customers in the queue while the $TC$ arrives, his waiting time is the idle time in that state. When there are more than 1 customer in queue, two cases may arise, assuming that the customer gets served in $f^{th}$ batch, depending on whether there are sufficiently many customers to initiate the $f^{th}$ batch service or not.

Suppose the $TC$ on his arrival to the system in a batch of $n_1$ customers, finds himself in state $E_h = \{(0, p, n_1, n_2, j, i) : 0 \leq p = q-1; 1 \leq n_1 \leq a; n_2 = 0(p+h); 1 \leq j \leq s; 1 \leq i \leq r\}$. Then, he has to wait till $h$ customers arrive to this system. His waiting time is the same as the idle time of the server in $E_h$. The LST of waiting time, given the system is in $E_h$ and the arrival phase changes to $i'$ is $W^*{(s/E_h)} = I^*{(s/E_h)}$.

Suppose on the arrival of the $TC$ to the system in a batch of $n_1$ customers the state of the system becomes $F = \{(l, p, n_1, n_2, j, i) : l \geq 0; 0 \leq p = q-1; 1 \leq n_1 \leq a; 1 \leq n_2 \leq b; 1 \leq j \leq s; 1 \leq i \leq r\}$. We assume that the $TC$ is served in $f^{th}$ batch and his position in the batch is $r'$ and position in queue is $t' = n - (n_1 - r')$, $[n = lq + p, 1 \leq r' \leq n_1]$. Let $n^1, n^2, \ldots, n^{f-1}, n^f$ be the successive batch sizes determined by Markov chain rule at each service completion epoch after the service of $n_2$ customers currently in service. Two cases arise:

Case 1: At the previous service completion epoch of $n^{f-1}$ customers (wherein the service phase changes to $j'$), there are sufficiently many customers to initiate the next service (the event denoted as $F_i$). The waiting time of $TC$ in queue is residual service time of $n_2$ customers + service time of all customer batches of sizes $n^1, n^2, \ldots, n^{f-1}$.

The probability that the residual service of a customer batch of size $n_2$ currently in service completes in $(x, x + dx]$ within time $t$, with the service phase changing to $j'$ is given by the $(j, j')$ entry of the matrix $\int_0^t xe^{s_0x}q_{n_2}S_dSdx$. LST of residual
service time (which results in the service phase changing to \( j' \)) is given by the \((j, j')\) entry of the matrix \( \int_0^\infty e^{-st}e^{S q_n_{n+1}S_d}dt = (sI - S_0)^{-1}q_n_{n+1}S_d \).

The probability that the service completion of a customer batch of size \( n^1 \) completes in \([x, x + dx]\), with the service phase changing to \( j' \) and the next batch taken for service is of size \( n^2 \), is given by the \((j, j')\) entry of the matrix \( e^{S_0xq_n_{n+1}S_d}dx \). \( LST \) of service time, given the current service batch is of size \( n^1 \) and next service batch is of size \( n^2 \) is the \((j, j')\) entry of the matrix (provided at the service completion epoch the service phase is \( j' \)), is \( \int_0^\infty e^{-sx}e^{S_0xq_n_{n+1}S_d}dx = (sI - S_0)^{-1}(q_n_{n+1}S_d) \).

The \( LST \) of waiting time of \( TC \) (who is the \( l^{th} \) customer in queue) given \( F_1 \) is the \((j, j')\) entry of the matrix,

\[
\sum_{f=\left(\frac{j-1}{t}\right)+1, \frac{j'}{t}}^{t'} \sum_{(n_1, n_2, \ldots, n_{l-1})} \frac{q_{n_1n_2S_d}q_{n_2n_2S_d}q_{n_{l-2}n_{l-1}S_d}}{(sI - S_0)^2 sI - S_0 \cdots sI - S_0} M_h,
\]

\( n^1 + n^2 + \cdots + n^f \leq n. \)

**Case 2:** At the previous service completion epoch of \( n^{l-1} \) customers (wherein the service phase changes to \( j' \)), there are not sufficiently many customers to initiate the service for a batch of size \( n^l \). The system reaches a state \( E_h = \{(0, p, n_1, n_2, j, i) : 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; n_2 = 0(p + h); 1 \leq j \leq s; 1 \leq i \leq r\} \) at the end of the \((f - 1)^{th}\) batch service completion. The server waits for the arrival of \( h \) customers to initiate the next batch service. Or, the server is idle and upon the arrival of customers the arrival phase changes to \( j' \). The event is denoted as \( F_2 \). The waiting time of \( TC \) in queue is residual service time of \( n_2 \) customers + service time of all customer batches of sizes \( n^1, n^2, \ldots, n^{l-1} \), \( n^f \) + an expected idle time in \( E_h(M_h) \).

The \( LST \) of waiting time of \( TC \) (who is the \( l^{th} \) customer in queue) given \( F_2 \) is the \((j, j')\) entry of the matrix,

\[
\sum_{f=\left(\frac{j-1}{t}\right)+1, \frac{j'}{t}}^{t'} \sum_{(n_1, n_2, \ldots, n_{l-1})} \frac{q_{n_1n_2S_d}q_{n_2n_2S_d}q_{n_{l-2}n_{l-1}S_d}}{(sI - S_0)^2 sI - S_0 \cdots sI - S_0} M_h,
\]

\( n^1 + n^2 + \cdots + n^f > n. \)

6. **Other performance measures.**

- **Expected queue length,**

\[
E_{ql} = \sum_{p=0}^{q-1} \sum_{n_1=1}^{a} \sum_{n_2=0(p+1)}^{b} \sum_{j=1}^{s} \sum_{i=1}^{r} p x_{0p,n_1n_2ji} + \sum_{l=1}^{\infty} \sum_{p=0}^{q-1} \sum_{n_1=1}^{a} \sum_{n_2=1}^{b} \sum_{i=1}^{s} (lq + p) x_{lp,n_2ji},
\]

- **The fraction of time server is providing service to a customer batch of size \( k \),**

\[
T_k = \sum_{p=0}^{q-1} \sum_{n_1=1}^{a} \sum_{j=1}^{s} \sum_{i=1}^{r} x_{0p,n_1jki} + \sum_{l=1}^{\infty} \sum_{p=0}^{q-1} \sum_{n_1=1}^{a} \sum_{j=1}^{s} \sum_{i=1}^{r} x_{lp,n_2jki}.
\]

- **The fraction of time server is providing service to a customer batch,**

\[
T_{(busy)} = \sum_{k=1}^{b} T_k.
\]
• Probability that the server is idle with no customers in the system,
\[ P_{(idle,0)} = \sum_{n_1=1}^{a} \sum_{n_2=1}^{b} \sum_{j=1}^{s} \sum_{i=1}^{r} x_{00n_1,0(n_2)ji}. \]

• Probability that the server is idle with \( p \) customers in the system,
\[ P_{(idle,p)} = \sum_{n_1=1}^{a} \sum_{n_2=p+1}^{b} \sum_{j=1}^{s} \sum_{i=1}^{r} x_{0pn_1,0(n_2)ji}. \]

• Probability that the server is idle,
\[ P_{(idle)} = \sum_{p=0}^{b-1} P_{(idle,p)}. \]

7. **Numerical results.** For the purpose of numerical illustration, we consider a queueing system where the arrival is according to BMAP with representation \((D_0, D_c)\) and the maximum arrival batch size 2. The successive arrival batch sizes form a first order Markov chain with \( tpm, P = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \). The service is provided according to BMSP with representation \((S_0, S_d)\) and the maximum service batch size is 3. The successive service batch sizes form a first order Markov chain with \( tpm, Q = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix} \). We study the effect of correlation in inter-arrival times and service times on the following performance measures for the queueing system under consideration:

- Expected queue length
- Expected idle time when an empty system waits for 1, 2, 3 customers respectively to initiate the next service
- Conditional expected waiting time, when the TC is expected to be served in \( f^{th} \) batch and there are enough customers to initiate the next batch service or not.

For the arrival process we consider the following set of matrices for \( D_0 \) and \( D_c \). These five BMAP processes are normalized so as to have an arrival rate of 1.

1. **BMAP with** \( D_0 \) and \( D_c \) **having zero correlation (ZCA)**
   \[ D_0 = \begin{bmatrix} -2 & 2 & 0 \\ 0 & -3 & 3 \\ 0 & 0 & -5 \end{bmatrix}, \quad D_c = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 5 & 0 & 0 \end{bmatrix} \]

2a. **BMAP with** \( D_0 \) and \( D_c \) **having negative correlation (−0.4889)(NCA)**
   \[ D_0 = \begin{bmatrix} -1.0022 & 1.0022 & 0 \\ 0 & -1.0022 & 0 \\ 0 & 0 & -225.7466 \end{bmatrix}, \quad D_c = \begin{bmatrix} 0 & 0 & 0 \\ 0.0100 & 0 & 0.9922 \\ 223.4892 & 0 & 2.2575 \end{bmatrix} \]

2b. **BMAP with** \( D_0 \) and \( D_c \) **having negative correlation (−0.2452)(NCA)**
   \[ D_0 = \begin{bmatrix} -0.5423 & 0.0603 & 0.0603 \\ 0.0603 & -1.6270 & 0.0603 \\ 0.0603 & 0.1205 & -19.8861 \end{bmatrix}, \quad D_c = \begin{bmatrix} 0 & 0.0603 & 0.3616 \\ 0.0603 & 0.0603 & 1.3860 \\ 18.6809 & 0.9039 & 0.1205 \end{bmatrix} \]

3a. **BMAP with** \( D_0 \) and \( D_c \) **having positive correlation (0.2782)(PCA)**
7.2. Expected idle time when an empty system waits for 1, 2, 3 customers respectively to initiate the next service. In Section 4, we derived LST’s of idle time given the process is in state $E_h = \{(0, p, n_1, n_2, j, i) : 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; n_2 = 0(p + h); 1 \leq j \leq s; 1 \leq i \leq r\}$ and the arrival phase changes to $i'$. 

\[
D_0 = \begin{bmatrix}
-0.1957 & 0.0391 & 0.0391 \\
0.0391 & -0.2349 & 0.0391 \\
0.0391 & 0.0783 & -12.9188
\end{bmatrix},
D_c = \begin{bmatrix}
0.0391 & 0.0391 & 0.0391 \\
0.0783 & 0.0783 & 0 \\
0.0783 & 0.5872 & 12.1358
\end{bmatrix}
\]

3b. BMAP with $D_0$ and $D_c$ having positive correlation (0.4889)(PCA)

\[
D_0 = \begin{bmatrix}
-1.0022 & 1.0022 & 0 \\
0 & -1.0022 & 0 \\
0 & 0 & -225.7466
\end{bmatrix},
D_c = \begin{bmatrix}
0.9922 & 0 & 0.0100 \\
0 & 2.2575 & 0 \\
0 & 223.4892 & 0
\end{bmatrix}
\]

For the service process we consider the following set of matrices for $S_0$ and $S_d$.

1. BMSP with $S_0$ and $S_d$ having zero correlation (ZCS)

\[
S_0 = \begin{bmatrix}
-2 & 2 \\
0 & -2
\end{bmatrix},
S_d = \begin{bmatrix}
0 & 0 \\
2 & 0
\end{bmatrix}
\]

2a. BMSP with $S_0$ and $S_d$ having negative correlation—0.0025 (NCS)

\[
S_0 = \begin{bmatrix}
-1.3889 & 1.1111 \\
1.6667 & -5.5556
\end{bmatrix},
S_d = \begin{bmatrix}
0 & 0.2778 \\
3.8889 & 0
\end{bmatrix}
\]

2b. BMSP with $S_0$ and $S_d$ having negative correlation—0.3010 (NCS)

\[
S_0 = \begin{bmatrix}
-9.3844 & 0.0938 \\
0.0001 & -0.5285
\end{bmatrix},
S_d = \begin{bmatrix}
0.0938 & 9.1967 \\
0.5283 & 0.0001
\end{bmatrix}
\]

3. BMSP with $S_0$ and $S_d$ having positive correlation 0.2738 (PCS)

\[
S_0 = \begin{bmatrix}
-2.1739 & 0.0072 \\
0.0072 & -0.4348
\end{bmatrix},
S_d = \begin{bmatrix}
2.1449 & 0.0217 \\
0.0072 & 0.4203
\end{bmatrix}
\]

7.1. Expected queue length. The expected queue length under various arrival and service processes is tabulated in Table 3.

|          | NCA (2a) | NCA (2b) | ZCA   | PCA (3a) | PCA (3b) |
|----------|----------|----------|-------|----------|----------|
| NCS(2b)  | 4.5898   | 5.7659   | 3.1824| 42.5934  | 179.1428 |
| NCS(2a)  | 3.3006   | 4.3856   | 2.1585| 41.1375  | 177.7154 |
| ZCS      | 1.1070   | 1.1905   | 0.9351| 8.1323   | 40.2739  |
| PCS      | 40.5835  | 42.1713  | 32.6996| 82.5983  | 214.3443 |

From this table it can be drawn that:

- As the correlation in arrival process increases from 0 to higher values, the expected queue length increases while, the same cannot be said about negatively correlated arrival process. The expected queue length is least for arrival processes with 0 correlation.
- As the correlation in service process increases to 0 from negative values, the expected queue length decreases and as the correlation in service process increases from 0 to higher values, the expected queue length increases. The expected queue length is least for service processes with 0 correlation.
- The expected queue length is least for a queue with arrival process having 0 correlation and service process having 0 correlation.
at the end of the idle time in that state. Using this result, the expected idle time, given the process is in states (0, 0, 1, 0(1), 1, 1), (0, 0, 1, 0(2), 1, 1), (0, 0, 1, 0(3), 1, 1) respectively and arrival phase changes to 1, 2, 3 respectively at the end of idle time under various arrival and service processes, are computed and presented in Table 4.

### Table 4. Expected idle time in (0, 0, 1, 0(1), 1, 1), (0, 0, 1, 0(2), 1, 1), (0, 0, 1, 0(3), 1, 1) respectively and arrival phase changes to 1, 2, 3 respectively at the end of idle time under various arrival and service processes

| L(1,1) | M(1,1) | N(1,1) | L(1,2) | M(1,2) | N(1,2) | L(1,3) | M(1,3) | N(1,3) |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0200 | 1.5725 | 0.7685 | 0      | 0      | 0      | 1.9756 | 0.4126 | 2.5034 |
| 0.2199 | 1.2456 | 1.1507 | 0.2274 | 0.1551 | 0.2979 | 1.4797 | 0.8040 | 1.7725 |
| 1.0333 | 1.8600 | 2.7280 | 0      | 0      | 0      | 0      | 0      | 0      |
| 1.8098 | 2.2052 | 2.3602 | 1.8682 | 2.3471 | 2.5799 | 2.5189 | 4.2899 | 6.0942 |
| 1.9756 | 1.5247 | 5.1267 | 0      | 0      | 0      | 0      | 0      | 0      |
| 0.0200 | 1.5725 | 0.7685 | 0      | 0      | 0      | 1.9756 | 0.4126 | 2.5034 |
| 0.2199 | 1.2456 | 1.1507 | 0.2274 | 0.1551 | 0.2979 | 1.4797 | 0.8040 | 1.7725 |
| 1.0333 | 1.8600 | 2.7280 | 0      | 0      | 0      | 0      | 0      | 0      |
| 1.8098 | 2.2052 | 2.3602 | 1.8682 | 2.3471 | 2.5799 | 2.5189 | 4.2899 | 6.0942 |

### Table 5. Expected idle time in (0, 0, 1, 0(1), 1, 2), (0, 0, 1, 0(2), 1, 2), (0, 0, 1, 0(3), 1, 2) respectively and arrival phase changes to 1, 2, 3 respectively at the end of idle time under various arrival and service processes

| L(2,1) | M(2,1) | N(2,1) | L(2,2) | M(2,2) | N(2,2) | L(2,3) | M(2,3) | N(2,3) |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0100 | 0.1881 | 0.3977 | 0      | 0      | 0      | 0.9878 | 0.2292 | 1.8764 |
| 0.0567 | 0.5166 | 0.5367 | 0.0351 | 0.0530 | 0.2016 | 0.5964 | 0.3021 | 1.2421 |
| 0.5333 | 1.3600 | 2.2280 | 0      | 0      | 0      | 0      | 0      | 0      |
| 1.9276 | 2.5820 | 2.8420 | 1.9727 | 2.6957 | 3.0442 | 1.4113 | 3.6018 | 5.8677 |
| 0.9878 | 2.5447 | 4.1548 | 0      | 0      | 0      | 0      | 0      | 0      |
| 0.0100 | 0.7881 | 0.3977 | 0      | 0      | 0      | 0      | 0      | 0      |
| 0.0567 | 0.5166 | 0.5367 | 0.0351 | 0.0530 | 0.2016 | 0.5964 | 0.3021 | 1.2421 |
| 0.5333 | 1.3600 | 2.2280 | 0      | 0      | 0      | 0      | 0      | 0      |
| 1.9276 | 2.5820 | 2.8420 | 1.9727 | 2.6957 | 3.0442 | 1.4113 | 3.6018 | 5.8677 |
| 0.9878 | 2.5447 | 4.1548 | 0      | 0      | 0      | 0      | 0      | 0      |
| 0.0100 | 0.7881 | 0.3977 | 0      | 0      | 0      | 0      | 0      | 0      |
| 0.0567 | 0.5166 | 0.5367 | 0.0351 | 0.0530 | 0.2016 | 0.5964 | 0.3021 | 1.2421 |
| 0.5333 | 1.3600 | 2.2280 | 0      | 0      | 0      | 0      | 0      | 0      |
| 1.9276 | 2.5820 | 2.8420 | 1.9727 | 2.6957 | 3.0442 | 1.4113 | 3.6018 | 5.8677 |
| 0.9878 | 2.5447 | 4.1548 | 0      | 0      | 0      | 0      | 0      | 0      |
From Table 4, 5, 6 one can observe that:

- The expected idle time in a particular state depends only on the arrival process and is independent of the service process.
- For positively correlated arrival process, the expected idle time is maximum when 3 customers are required to initiate the service process.
- For negatively correlated arrival process, the expected idle time shows no such pattern.
- The sparse structure of $D_0$ and $D_c$ results in values 0 for expected idle time. A zero in $D_c$ indicates the impossibility of transitions amongst the underlying Markov chain with arrivals.

### 7.3. Conditional expected waiting time, when the TC is expected to be served in $f^{th}$ batch and there are enough customers to initiate the next batch service or not.

In Section 5, we derived $LST$’s of waiting time given the $TC$ finds himself in state $F = \{(l, p, n_1, n_2, j, i) : l \geq 1; 0 \leq p \leq q - 1; 1 \leq n_1 \leq a; 1 \leq n_2 \leq b; 1 \leq j \leq s; 1 \leq i \leq r\}$ given two events which we denoted as $F_1$ and $F_2$ respectively. Here the conditional expected waiting time of a tagged customer who find himself in states $A(0, 1, 1, 2, j, 1)$ and $B(0, 2, 1, 2, j, 1); 1 \leq j \leq 2$ given $F_1$ are computed and presented in Table 7. In the event $F_1$, the customer is served in the $f^{th}$ batch and there are enough customers to initiate this batch service and hence there is a phase change from $j$ to either 1 or 2 in service process alone, at the end of the waiting time. If the service phase change from $j$ to $j'$ at the end of waiting time, then the conditional expected waiting time in $A$ and $B$ given $F_1$, are respectively denoted as $A_1(j, j')$ and $B_1(j, j')$.

From Table 7 one can observe that:

- The conditional expected waiting time given $F_1$, depends only on the service process and is independent of the arrival process.
- When there are more number of customers in queue, the conditional expected waiting time given $F_1$ increases.
- The sparse structure of $S_0$ and $S_d$ results in values 0 for expected waiting time. A zero in $S_d$, indicates the impossibility of transitions amongst the underlying Markov chain with service completions.

#### Table 6. Expected idle time in $(0, 0, 1, 0(1), 1, 3), (0, 0, 1, 0(2), 1, 3), (0, 0, 1, 0(3), 1, 3)$ respectively and arrival phase changes to 1, 2, 3 respectively at the end of idle time under various arrival and service processes

|             | L(3,1) | M(3,1) | N(3,1) | L(3,2) | M(3,2) | N(3,2) | L(3,3) | M(3,3) | N(3,3) |
|-------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| ZCS NCA(2a) | 0.0044 | 0.0165 | 1.2524 | 0.0032 | 0.0088 | 0.1684 | 0.1684 | 0.1684 | 0.1684 |
| ZCS NCA(2b) | 0.0038 | 0.0088 | 1.0924 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| ZCS ZCA    | 0.0090 | 0.0100 | 1.0924 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| ZCS PCA(3a) | 0.0176 | 0.1166 | 0.2426 | 0.0244 | 0.1252 | 0.2598 | 0.2598 | 0.2598 | 0.2598 |
| ZCS PCA(3b) | 0.0250 | 0.0100 | 0.4168 | 0.0100 | 0.1252 | 0.2598 | 0.2598 | 0.2598 | 0.2598 |
| PCS NCA(2a) | 0.0041 | 0.0165 | 1.2524 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| PCS NCA(2b) | 0.0041 | 0.0165 | 1.2524 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| PCS ZCA    | 0.0090 | 0.0100 | 1.0924 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| PCS PCA(3a) | 0.0176 | 0.1166 | 0.2426 | 0.0244 | 0.1252 | 0.2598 | 0.2598 | 0.2598 | 0.2598 |
| PCS PCA(3b) | 0.0250 | 0.0100 | 0.4168 | 0.0100 | 0.1252 | 0.2598 | 0.2598 | 0.2598 | 0.2598 |
| NCS(2a) NCA(2a) | 0.0040 | 0.0165 | 1.2524 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| NCS(2a) NCA(2b) | 0.0040 | 0.0165 | 1.2524 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| NCS(2a) ZCA    | 0.0090 | 0.0100 | 1.0924 | 0.0124 | 0.1684 | 0.0088 | 0.0088 | 0.0088 | 0.0088 |
| NCS(2a) PCA(3a) | 0.0176 | 0.1166 | 0.2426 | 0.0244 | 0.1252 | 0.2598 | 0.2598 | 0.2598 | 0.2598 |
| NCS(2a) PCA(3b) | 0.0250 | 0.0100 | 0.4168 | 0.0100 | 0.1252 | 0.2598 | 0.2598 | 0.2598 | 0.2598 |
Table 7. $A_1(j, j')$ and $B_1(j, j')$.

|       | $A(1, 1)$ | $A(1, 2)$ | $A(2, 1)$ | $A(2, 2)$ | $B(1, 1)$ | $B(1, 2)$ | $B(2, 1)$ | $B(2, 2)$ |
|-------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ZCS NCA(2a) | 0.3000    | 0.3000    | 0.1000    | 0.1000    | 0.3400    | 0.3400    | 0.3400    | 0.3400    |
| ZCS NCA(2b) | 0.3000    | 0.3000    | 0.1000    | 0.1000    | 0.3400    | 0.3400    | 0.3400    | 0.3400    |
| ZCS ZCA     | 0.3000    | 0.3000    | 0.1000    | 0.1000    | 0.3400    | 0.3400    | 0.3400    | 0.3400    |
| ZCS PCA(3a) | 0.3000    | 0.3000    | 0.1000    | 0.1000    | 0.3400    | 0.3400    | 0.3400    | 0.3400    |
| ZCS PCA(3b) | 0.3000    | 0.3000    | 0.1000    | 0.1000    | 0.3400    | 0.3400    | 0.3400    | 0.3400    |
| PCS NCA(2a) | 0.0838    | 0.0993    | 0.0784    | 0.0784    | 0.4063    | 0.0963    | 0.0110    | 0.0974    |
| PCS NCA(2b) | 0.0838    | 0.0993    | 0.0784    | 0.0784    | 0.4063    | 0.0963    | 0.0110    | 0.0974    |
| PCS ZCA     | 0.0838    | 0.0993    | 0.0784    | 0.0784    | 0.4063    | 0.0963    | 0.0110    | 0.0974    |
| PCS PCA(3a) | 0.0838    | 0.0993    | 0.0784    | 0.0784    | 0.4063    | 0.0963    | 0.0110    | 0.0974    |
| PCS PCA(3b) | 0.0838    | 0.0993    | 0.0784    | 0.0784    | 0.4063    | 0.0963    | 0.0110    | 0.0974    |

The conditional expected waiting time of a tagged customer who find himself in states $A(0, 1, 2, j, 1)$ and $B(0, 2, 1, 2, j, 1)$: $1 \leq j \leq 2$ given $F_2$ are computed and presented in Table 8. In the event $F_2$, the customer is served in the $i^{th}$ batch and there are not enough customers to initiate this batch service. Hence, the waiting time includes an idle time whence the system waits for customers to initiate the batch service. Hence there is a phase change from $j$ to $j'$ in service process, at the end of the waiting time and at the end of idle time, the arrival phase changes to $i'$ (Here for convenience we take $i' = 1$). If the service phase change from $j$ to $j'$ at the end of waiting time then the conditional expected waiting time in A and B given $F_2$ are respectively denoted as $A_2(j, j')$ and $B_2(j, j')$.

Table 8. $A_2(j, j')$ and $B_2(j, j')$.

|       | $A_2(1, 1)$ | $A_2(1, 2)$ | $A_2(2, 1)$ | $A_2(2, 2)$ | $B_2(1, 1)$ | $B_2(1, 2)$ | $B_2(2, 1)$ | $B_2(2, 2)$ |
|-------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| ZCS NCA(2a) | 0.4212      | 0.4737      | 0.4599      | 0.4599      | 0.1717      | 0.1717      | 0.1717      | 0.1717      |
| ZCS NCA(2b) | 0.1870      | 0.3957      | 0.5642      | 0.5642      | 0.2033      | 0.2033      | 0.2033      | 0.2033      |
| ZCS ZCA     | 0.9840      | 0.6613      | 1.5335      | 1.5335      | 0.5363      | 0.5363      | 0.5363      | 0.5363      |
| ZCS PCA(3a) | 0.5277      | 0.8426      | 2.3911      | 2.3911      | 0.8288      | 0.8288      | 0.8288      | 0.8288      |
| ZCS PCA(3b) | 3.7649      | 1.2550      | 2.9230      | 2.9230      | 1.0221      | 1.0221      | 1.0221      | 1.0221      |
| PCS NCA(2a) | 0.3972      | 0.3716      | 0.6942      | 0.6942      | 0.3162      | 0.2146      | 0.2146      | 0.2146      |
| PCS NCA(2b) | 0.3317      | 0.3103      | 0.8964      | 0.8964      | 0.1693      | 0.2257      | 0.2257      | 0.2257      |
| PCS ZCA     | 0.5544      | 0.5187      | 13.5328     | 13.5328     | 0.6255      | 0.5298      | 0.5298      | 0.5298      |
| PCS PCA(3a) | 0.7063      | 0.6608      | 17.2411     | 17.2411     | 0.6924      | 0.7879      | 0.7879      | 0.7879      |
| PCS PCA(3b) | 1.0521      | 0.9843      | 25.6805     | 25.6805     | 0.8533      | 1.0087      | 1.0087      | 1.0087      |
| NCS(2a) NCA(2a) | 1.7200 | 0.5080 | 0.6618 | 0.1843 | 0.6196 | 0.1777 | 0.2529 | 0.0732 |
| NCS(2a) NCA(2b) | 1.4365 | 0.5243 | 0.5528 | 0.1539 | 0.7351 | 0.2128 | 0.2949 | 0.0844 |
| NCS(2a) ZCA | 2.4011 | 0.7091 | 0.9239 | 0.2573 | 1.9419 | 0.5664 | 0.7671 | 0.2174 |
| NCS(2a) PCA(3a) | 3.0590 | 0.9035 | 1.1771 | 0.3278 | 3.0025 | 0.8777 | 1.1804 | 0.3345 |
| NCS(2a) PCA(3b) | 4.5564 | 1.3457 | 1.7533 | 0.4882 | 3.7009 | 1.0794 | 1.4617 | 0.4143 |

From Table 8 one can observe that:

- The conditional expected waiting time given $F_2$, depends on service process and arrival process.
- When there are more number of customers in queue the conditional expected waiting time given $F_2$ need not always increase.
- For a particular service process, the conditional expected waiting time given $F_2$ increases when the correlation in arrival process increase from 0 to higher values.
• The conditional expected waiting time given $F_2$, is maximum for positively correlated arrival and service process.
• The sparse structure of $S_0$ and $S_d$ results in values 0 for expected waiting time. A zero in $S_d$ indicates the impossibility of transitions amongst the underlying Markov chain with service completions.

8. Conclusion. In this paper we considered a $BMAP/BMSP/1$ queue with arrival and service batch sizes forming a first order Markov chain. The conditional LST’s of idle time in different states and conditional LST’s of waiting times of a tagged customer in queue are derived. Numerical experiments were performed to evaluate performance measures. Several extensions to the model presented are proposed to be carried out in future.

Acknowledgments. The authors thank the reviewers for helpful suggestions that improved the presentation of the paper. Anu Nuthan Joshua thanks Department of Mathematics, Cochin University of Science and Technology, Cochin -22 for providing facilities for conducting research.

REFERENCES
[1] A. D. Banik, Analyzing state-dependent arrival in $GI/BMSP/1/\infty$ queues, Math. Comput. Modelling, 53 (2011), 1229–1246.
[2] A. D. Banik, Queueing analysis and optimal control of $BMAP/G^{(a,b)}/1/N$ and $BMAP/MSP^{(a,b)}/1/N$ systems, Comput. Industrial Engineering, 57 (2009), 748–761.
[3] A. D. Banik, Stationary analysis of a $BMAP/R/1$ queue with $R$-type multiple working vacations, Comm. Statist. Simulation Comput., 46 (2017), 1033–1061.
[4] A. D. Banik, M. L. Chaudhry and U. C. Gupta, On the finite buffer queue with renewal input and batch Markovian service process: $GI/BMSP/1/N$, Methodol. Comput. Appl. Probab., 10 (2008), 559–575.
[5] S. Ghosh and A. D. Banik, An algorithmic analysis of the $BMAP/MSP/1$ generalized processor-sharing queue, Comput. Oper. Res., 79 (2017), 1–11.
[6] S. Ghosh and A. D. Banik, Computing conditional sojourn time of a randomly chosen tagged customer in a $BMAP/MSP/1$ queue under random order service discipline, Ann. Oper. Res., 261 (2018), 185–206.
[7] S. Ghosh and A. D. Banik, Efficient computational analysis of non-exhaustive service vacation queues: $BMAP/R/1/N(\infty)$ under gated-limited discipline, Appl. Math. Model., 68 (2019), 540–562.
[8] C. Kim, V. I. Klimenok and A. N. Dudin, Analysis of unreliable $BMAP/P\bar{H}/N$ type queue with Markovian flow of breakdowns, Appl. Math. Comput., 314 (2017), 154–172.
[9] V. Klimenok, A. N. Dudin and K. Samouylov, Computation of moments of queue length in the $BMAP/SM/1$ queue, Oper. Res. Lett., 45 (2017), 467–470.
[10] V. Klimenok and O. Dudina, Retrial tandem queue with controllable strategy of repeated attempts, Quality Tech. Quantitative Manag., 14 (2017), 74–93.
[11] V. Klimenok, O. Dudina, V. Vishnevsky and K. Samouylov, Retrial tandem queue with $BMAP$ input and semi-Markovian service process, in International Conference on Distributed Computer and Communication Networks, Communications in Computer and Information Science, 700, Springer, Cham, 2017, 159–173.
[12] V. I. Klimenok, A. N. Dudin and K. E. Samouylov, Analysis of the $BMAP/P\bar{H}/N$ queueing systems with backup servers, Appl. Math. Model., 57 (2018), 64–84.
[13] D. M. Lucantoni, New results on the single server queue with a batch Markovian arrival process, Comm. Statist. Stochastic Models, 7 (1991), 1–46.
[14] D. M. Lucantoni, K. S. Meier-Hellstern and M. F. Neuts, A single-server queue with server vacations and a class of non-renewal arrival processes, Adv. in Appl. Probab., 22 (1990), 676–705.
[15] M. F. Neuts, Versatile Markovian point process, J. Appl. Probab., 16 (1979), 764–779.
[16] M. F. Neuts, Matrix-Geometric Solutions in Stochastic Models. An Algorithmic Approach, Johns Hopkins Series in the Mathematical Sciences, 2, John Hopkins University Press, Baltimore, MD, 1981.

[17] G. Rama, R. Ramshankar, R. Sandhya, V. Sundar and R. Ramanarayanan, BMAP/M/C bulk service queue with randomly varying environment, IOSR J. Engineering, 5 (2015), 33–47.

[18] V. Ramaswami, The N/G/1 queue and its detailed analysis, Adv. in Appl. Probab., 12 (1980), 222–261.

[19] S. K. Samanta, M. L. Chaudhry and A. Pacheco, Analysis of BMAP/MSP/1 queue, Methodol. Comput. Appl. Probab., 18 (2016), 419–440.

[20] R. Sandhya, V. Sundar, G. Rama, R. Ramshankar and R. Ramanarayanan, BMAP/BMSP/1 queue with randomly varying environment, IOSR J. Engineering, 5 (2015), 01–12.

[21] V. M. Vishnevskii and A. N. Dudin, Queueing systems with correlated arrival flows and their applications to modeling telecommunication networks, Autom. Remote Control, 78 (2017), 1361–1403.

Received for publication November 2019.

E-mail address: achyuthacusat@gmail.com
E-mail address: anunuthanjoshua@gmail.com