Top (2007) 15: 146–160
DOI 10.1007/s11750-007-0007-2

ORIGINAL PAPER

Analyzing the finite buffer batch arrival queue under Markovian service process: $GI^X/MSP/1/N$

A.D. Banik · U.C. Gupta

Received: 20 March 2006 / Accepted: 9 January 2007 /
Published online: 10 March 2007

Abstract We consider a batch arrival finite buffer single server queue with inter-batch arrival times are generally distributed and arrivals occur in batches of random size. The service process is correlated and its structure is presented through Markovian service process (*MSP*). The model is analyzed for two possible customer rejection strategies: partial batch rejection and total batch rejection policy. We obtain steady-state distribution at pre-arrival and arbitrary epochs along with some important performance measures, like probabilities of blocking the first, an arbitrary, and the last customer of a batch, average number of customers in the system, and the mean waiting times in the system. Some numerical results have been presented graphically to show the effect of model parameters on the performance measures. The model has potential application in the area of computer networks, telecommunication systems, manufacturing system design, etc.

Keywords General independent arrival · Batch arrival · Finite buffer · Queue · Markovian service process

Mathematics Subject Classification (2000) 60K25 · 90B22

1 Introduction

During the last few decades queueing systems under various type of arrival and service processes have been investigated by researchers due to their applicability in various networking situations, production and manufacturing systems, etc. Traditional

A.D. Banik (✉) · U.C. Gupta
Department of Mathematics, Indian Institute of Technology, Kharagpur 721302, India
e-mail: banikad@gmail.com

U.C. Gupta
e-mail: umesh@iitkgp.ac.in

queueing analysis using Poisson process is not powerful enough to capture the correlated nature of arrival (service) processes. Markovian arrival process (*MAP*) is important because it can capture the correlation among the inter-arrival times. Similarly, batch Markovian arrival process (*BMAP*) is used to capture correlation among the inter-batch arrival times. Considerable amount of literature is available on the queueing systems with *MAP* arrival, see Lucantoni et al. (1990), Kasahara et al. (1996), Choi et al. (1998), etc. Also plenty of studies are available on queueing systems with *BMAP* arrival, see Lucantoni (1991), Ferrandiz (1993), Niu et al. (2003), etc. There exist quite a few examples on batch service queue with *MAP* arrival, e.g., Chakravarthy (1993), Gupta and Vijaya (2001), etc.

Like the *MAP*, Markovian service process (*MSP*) is a versatile service process and can capture the correlation among the service times. Several other service processes, e.g., Poisson process, Markov modulated Poisson process (*MMPP*), PH-type renewal process, etc., can be considered as special cases of *MSP*. For details on *MSP* readers are referred to Bocharov (1996), and Albores and Tajonar (2004). The analysis of finite as well as infinite buffer $G/MSP/1/r$ ($r \leq \infty$) queue has been performed by Bocharov et al. (2003). The multi-server $GI/MSP/c/r$ queue has been analyzed by Albores and Tajonar (2004). Alfa et al. (2000) discussed the asymptotic behaviour of the $GI/MSP/1$ queue using perturbation theory. The $G/SM/1/\infty$ queueing system with vacations was considered by Machihara (1995). However, to the best of authors' knowledge there are no studies available on *MSP* with renewal input of batches of random size.

In this paper we analyze the $GI^X/MSP/1/N$ queue where customers arrive in batches of random size. Unlike the other batch arrival queue with finite buffer, batches which upon arrival find not enough space in the buffer are either fully or partially rejected. Some queueing protocols are based on the former strategy, and it is known as the total batch rejection policy. The later situation is known as the partial batch rejection policy. Both policies are analyzed in this paper and steady-state distribution at pre-arrival and arbitrary epochs, and the LST's of actual waiting time of the first, an arbitrary, and the last customer in an accepted batch have been obtained. Several other queueing models analyzed in the past, e.g., $M^X/G/1/N$ (Baba 1984), $GI^X/E_k/1/N$ (Gupta and Vijaya 2000), $GI^X/G/1/N$ (Nobel 1989 studied $GI^X/G/1/N$ queue using approximation technique based on the exact results of $M^X/E_k/1/N$ and $GI^X/E_k/1/N$ queues), $GI/E_k/1/N$ (Ohsone 1981), $GI/MSP/1/N$ (Bocharov et al. 2003), etc., can be obtained as special cases from our model. It may be remarked here that recently Gupta and Banik (2007) have carried out the analysis of finite as well as infinite buffer $GI/MSP/1$ queue.

The paper is organized as follows. In Sect. 2 we give a description of the model and introduce the notation used to describe the model parameters. In Sects. 3 and 4 we present the analytic analysis of the model and obtain steady-state distribution at various epochs. In Sect. 5 performance measures and waiting times are discussed. Finally, Sect. 6 concludes with the numerical results and a study of the behaviour of some performance measures against the variation of critical model parameters.

2 Description of the model

Let us consider a single server queueing system wherein customers arrive in batches of random size X with $P(X=i)=g_i$, $i=1,2,\ldots,$ and mean batch size $E(X)=\sum_{i=1}^{\infty} i g_i=\bar{g}$. The inter-batch arrival times are independent identically distributed (i.i.d.) random variables (r.vs.) with distribution function (DF) $A(x)$, probability density function $a(x)$, Laplace–Stieltjes transform (LST) $A^*(\theta)$ and mean inter-batch arrival time $(1/\lambda,\ \text{say})=-A^{*(1)}(0)$, i.e., λ is the mean arrival rate of batches, where $f^{*(j)}(\zeta)$ is the jth ($j\geq 1$) derivative of $f^*(\theta)$ at $\theta=\zeta$.

The service process of the queueing system is a Markovian service process (*MSP*) and is governed by an underlying m-state Markov chain having transition rate L_{ij}, $1\leq i,\ j\leq m$, with a transition from state i to j without service completion and having transition rate M_{ij}, $1\leq i,\ j\leq m$, with a transition from state i to j with service completion. The matrix $\mathbf{L}$ has nonnegative off-diagonal and negative diagonal elements. The matrix $\mathbf{M}$ has nonnegative elements, and both have at least one positive entry. Let $N(t)$ denote the number of customers served in $(0,t]$ with state space $\{n : n\geq 0\}$, and let $J(t)$ be the state of the underlying Markov chain at time t with state space $\{i : 1\leq i\leq m\}$. Then $\{N(t),\ J(t)\}$ is a two-dimensional Markov process with state space $\{(n,i) : n\geq 0,\ \ 1\leq i\leq m\}$. The infinitesimal generator of the above Markov process, when the system is not empty, is given by

$$\mathbf{Q}^*=\begin{pmatrix} \mathbf{L} & \mathbf{M} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{L} & \mathbf{M} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{L} & \mathbf{M} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Therefore, we have $(\mathbf{L}+\mathbf{M})\mathbf{e}=\mathbf{0}$, where $\mathbf{e}$ is the $m\times 1$ column vector with all its elements equal to 1. Since $(\mathbf{L}+\mathbf{M})$ is the infinitesimal generator of the underlying Markov chain $\{J(t)\}$, there exists a stationary probability vector $\overline{\boldsymbol{\pi}}$ such that $\overline{\boldsymbol{\pi}}(\mathbf{L}+\mathbf{M})=\mathbf{0}$, $\overline{\boldsymbol{\pi}}\mathbf{e}=1$. The fundamental service rate of the stationary *MSP* is given by $\mu^*=\overline{\boldsymbol{\pi}}\mathbf{M}\mathbf{e}$. When the server remains idle for a certain time interval, and then a batch of customers enters, the service process starts with the initial phase distribution given by f_j, $j=1,2,\ldots,m$, $\sum_{j=1}^{m} f_j=1$, independently of the path followed in the previous service period. Therefore, *MSP* is characterized by the matrices $\mathbf{L}$, $\mathbf{M}$ along with the vector $\mathbf{f}=(f_1,f_2,\ldots,f_m)$. For further illustration on *MSP* readers are referred to Bocharov (1996) and Albores and Tajonar (2004).

The customers are served individually according to a *MSP* with mean service time $1/\mu^*$. The offered load ρ is as usual defined to be $\rho=\lambda\bar{g}/\mu^*$. Let N be the capacity of the system including the one in service. The state of the system at time t is described by the following random variables:

- $N_t=$ number of customers present in the system including the one in service
- $J_t=\{i\}$, where the server is in the ith ($1\leq i\leq m$) phase of the service process
- $U_t=$ remaining inter-batch arrival time for the next arrival of a batch

We define the joint probability densities of N_t, state of the server J_t, and the remaining inter-batch arrival time U_t, respectively, by

$$\pi_{n,i}(x,t)\Delta x = P\{N_t = n,\, J(t) = i, x < U_t < x + \Delta x\},$$
$$0 \le n \le N,\ 1 \le i \le m,\ x \ge 0.$$

As we shall discuss the model in the limiting case, i.e., when $t \to \infty$, the above probabilities will be denoted by $\pi_{n,i}(x) = \lim_{t\to\infty} \pi_{n,i}(x,t)$. Furthermore, let $\boldsymbol{\pi}_n(x)$ be the row vector of order $1 \times m$ whose ith component is $\pi_{n,i}(x)$ and denotes the probability that n customers are in the system and the service process is in phase i while remaining inter-batch arrival time is equal to x for the next arrival of a batch.

3 Steady-state distribution at pre-arrival epoch

Consider the system just before the arrival of a batch which is taken as embedded points. Let $t_0, t_1, t_2, \ldots$ be the time epochs at which batch arrival is about to occur, and let t_i^- denote the time epoch just before the arrival instant t_i. The state of the system at t_i^- $(i \ge 0)$ is defined as $\{N_{t_i^-}, \xi_{t_i^-}\}$ where $N_{t_i^-}$ is the number of customers in the system and $\xi_{t_i^-}$ indicates the phase of the service process. In the limiting case $\pi_{j,n}^-$ represents the probability that there are n customers in the system at pre-arrival epoch of a batch when the service process is in phase j, that is,

$$\pi_{n,j}^- = \lim_{i\to\infty} P\left(N_{t_i^-} = n,\ \xi_{t_i^-} = j\right), \quad 0 \le n \le N,\ 1 \le j \le m.$$

Let $\boldsymbol{\pi}_n^-$ $(0 \le n \le N)$ be the row vectors of order $1 \times m$ whose ith component is $\pi_{n,i}^-$.

Let $\mathbf{S}_k$, $k \ge 0$, denote an $m \times m$ matrix whose (i, j)th element represents the conditional probability that k customers have been served during an inter-batch arrival time and the underlying Markov chain of the service process is in phase j just before the arrival of a batch, given that the underlying Markov chain was in phase i at the previous pre-arrival epoch of a batch.

Observing the state of the system at two consecutive embedded points, we have an embedded Markov chain whose state space is equivalent to $\Omega = \{(i, j),\ 0 \le i \le N, 1 \le j \le m\}$. The elements $[\mathbf{P}_{ij}]_{m\times m}$ of the one step transition probability matrix (TPM) $\mathcal{P}_{(N+1)m\times(N+1)m}$ of the above mentioned Markov chain under partial and total batch rejection policies are given by

Partial batch rejection:

$$\mathbf{P}_{ij} = \begin{cases} \sum_{r=j-i}^{N-i} g_r \mathbf{S}_{i+r-j} + (1 - \sum_{r=1}^{N-i} g_r)\mathbf{S}_{N-j}, & j > i \ge 0, \\ \sum_{r=1}^{N-i} g_r \mathbf{S}_{i+r-j} + (1 - \sum_{r=1}^{N-i} g_r)\mathbf{S}_{N-j}, & 1 \le j \le i, \end{cases} \tag{1}$$

$$\text{and} \quad \mathbf{P}_{i0} = \left(\mathbf{I}_m - \sum_{k=1}^{N} \mathbf{P}_{ik}\right)\mathbf{ef} = \mathbf{B}^{00} - \sum_{k=1}^{N} \mathbf{P}_{ik}\mathbf{B}^{00}, \quad 0 \le i \le N, \tag{2}$$

where $\mathbf{B}^{00} = \mathbf{ef}$ is a stochastic matrix and has the invariant vector $\mathbf{f}$, and $\mathbf{I}_m$ is the identity matrix of the order given in the suffix.

Total batch rejection:

$$\mathbf{P}_{ij} = \begin{cases} \sum_{r=j-i}^{N-i} g_r \mathbf{S}_{i+r-j}, & j > i \geq 0, \\ \sum_{r=1}^{N-i} g_r \mathbf{S}_{i+r-j} + (1 - \sum_{r=1}^{N-i} g_r) \mathbf{S}_{i-j}, & 1 \leq j \leq i, \end{cases} \tag{3}$$

where $\mathbf{P}_{i0}$ $(0 \leq i \leq N)$ are same as for the partial rejection case. It may be noted here that if any of the $f_j = 0$ $(1 \leq j \leq m)$ then the jth column of the matrix $\mathbf{B}^{00}$ is zero, and as a result the jth column of each matrix $\mathbf{P}_{i0}$ will be equal to zero. Therefore, the jth column of the matrix $\mathcal{P}$ will also be equal to zero. In order to obtain the stochastic matrix it should be deleted from the first row and column of the matrix $\mathcal{P}$, and the corresponding component of the pre-arrival epoch probability $\pi_{0,j}^{-}$ will be equal to zero.

Remark 3.1 It may be remarked here that instead of assuming idle-restart service phase distribution it is natural to consider that a new busy period starts with the same service phase where previous busy period ends. If this is the case then TPM will remain the same as above, the only difference is that the expression of $\mathbf{P}_{i0}$ will slightly change, and it will be given by

$$\mathbf{P}_{i0} = \mathrm{diag}\left(\mathbf{I}_m - \sum_{k=1}^{N} \mathbf{P}_{ik}\mathbf{e}\mathbf{e}'\right), \quad 0 \leq i \leq N,$$

where $\mathbf{e}'$ is the $1 \times m$ row vector with all its elements equal to 1, and diag($\mathbf{A}$) is a diagonal matrix whose diagonal elements are equal to the diagonal elements of the matrix $\mathbf{A}$ and the rest of the elements of diag($\mathbf{A}$) are *zero*. Remaining part of the analysis of the paper stays same under this assumption.

The matrices $\mathbf{S}_k$ involved in the TPM, in general, for arbitrary inter-batch arrival time distribution require numerical integration which can be carried out along the lines proposed by Lucantoni and Ramaswami (1985). However, when the inter-batch arrival time distributions are of phase type (PH-distribution), these matrices can be evaluated without any numerical integration, Neuts (1981, pp. 67–70). It may be noted here that various inter-batch arrival time distributions arising in practical applications can be approximated by PH-distributions. The following theorem gives a procedure for the computation of the matrices $\mathbf{S}_k$.

Theorem 3.1 *Let $A(x)$ have a PH-distribution with irreducible representation $(\boldsymbol{\alpha}, \mathbf{T})$, where $\boldsymbol{\alpha}$ and $\mathbf{T}$ are of dimension γ. Then the matrices $\mathbf{S}_n$ are given by*

$$\mathbf{S}_n = \mathbf{B}_n\big(\mathbf{I}_m \otimes \mathbf{T}^0\big), \quad n \geq 0, \tag{4}$$

$$\text{where} \quad \mathbf{B}_0 = -(\mathbf{I}_m \otimes \boldsymbol{\alpha})[\mathbf{L} \otimes \mathbf{I}_\gamma + \mathbf{I}_m \otimes \mathbf{T}]^{-1},$$

$$\mathbf{B}_n = -\mathbf{B}_{n-1}(\mathbf{M} \otimes \mathbf{I}_\gamma)[\mathbf{L} \otimes \mathbf{I}_\gamma + \mathbf{I}_m \otimes \mathbf{T}]^{-1}, \quad n \geq 1,$$

where $\mathbf{T}^0 = -\mathbf{T}\mathbf{e}$ and the symbol $\otimes$ denotes the Kronecker product of two matrices.

Proof is similar to Neuts (1981) for PH-type service and PH-type arrival, and can be generalized to *MAP* arrival, see Machihara (1999). □

The pre-arrival epoch probabilities $\boldsymbol{\pi}_n^-$ $(0 \le n \le N)$ can be evaluated by solving the system of equations: $(\boldsymbol{\pi}_0^-, \boldsymbol{\pi}_1^-, \ldots, \boldsymbol{\pi}_N^-) = (\boldsymbol{\pi}_0^-, \boldsymbol{\pi}_1^-, \ldots, \boldsymbol{\pi}_N^-)\mathcal{P}$. We have used GTH algorithm developed by Grassmann et al. (1985) for solving the system of equations.

4 Steady-state distribution at arbitrary epoch

4.1 Partial batch rejection

To obtain steady-state distribution at an arbitrary epoch we will develop relations between distributions of the number of customers in the system at pre-arrival and arbitrary epochs. For this we use supplementary variable method and relate the states of the system at two consecutive time epochs t and $t + \Delta t$. Using probabilistic arguments and considering partial batch rejection policy, we get a set of partial differential equations. Taking limit as $t \to \infty$ and using matrices and vector notation, we have the following differential-difference equations:

$$-\frac{d}{dx}\boldsymbol{\pi}_0(x) = \boldsymbol{\pi}_1(x)\mathbf{M}, \tag{5}$$

$$-\frac{d}{dx}\boldsymbol{\pi}_n(x) = \boldsymbol{\pi}_n(x)\mathbf{L} + \boldsymbol{\pi}_{n+1}(x)\mathbf{M} + \sum_{i=0}^{n-1}\boldsymbol{\pi}_i(0)a(x)g_{n-i}, \quad 1 \le n \le N-1, \tag{6}$$

$$-\frac{d}{dx}\boldsymbol{\pi}_N(x) = \boldsymbol{\pi}_N(x)\mathbf{L} + \boldsymbol{\pi}_N(0)a(x) + \sum_{i=0}^{N-1}\boldsymbol{\pi}_i(0)a(x)\sum_{j=N-i}^{\infty} g_j, \tag{7}$$

where $\boldsymbol{\pi}_n(0)$ are the respective probabilities with remaining inter-batch arrival time equal to zero. Let us define the Laplace transform of $\boldsymbol{\pi}_n(x)$ as $\boldsymbol{\pi}_n^*(\theta)$, so that $\boldsymbol{\pi}_n \equiv \boldsymbol{\pi}_n^*(0)$, where $\boldsymbol{\pi}_n$ is the $1 \times m$ vector whose ith component $\pi_{n,i}$ denotes the probability that n customers are in the system and the service process is in phase i at an arbitrary time. Multiplying (5–7) by $e^{-\theta x}$ and integrating w.r.t. x from 0 to ∞, we get

$$-\theta\boldsymbol{\pi}_0^*(\theta) = \boldsymbol{\pi}_1^*(\theta)\mathbf{M} - \boldsymbol{\pi}_0(0), \tag{8}$$

$$-\theta\boldsymbol{\pi}_n^*(\theta) = \boldsymbol{\pi}_n^*(\theta)\mathbf{L} + \boldsymbol{\pi}_{n+1}^*(\theta)\mathbf{M} + \sum_{i=0}^{n-1}\boldsymbol{\pi}_i(0)A^*(\theta)g_{n-i} - \boldsymbol{\pi}_n(0), \quad 1 \le n \le N-1, \tag{9}$$

$$-\theta\boldsymbol{\pi}_N^*(\theta) = \boldsymbol{\pi}_N^*(\theta)\mathbf{L} + \boldsymbol{\pi}_N(0)A^*(\theta) + \sum_{i=0}^{N-1}\boldsymbol{\pi}_i(0)A^*(\theta)\sum_{j=N-i}^{\infty} g_j - \boldsymbol{\pi}_N(0). \tag{10}$$

One important result listed below in the form of a lemma uses (8–10).

Lemma 1

$$\sum_{n=0}^{N} \boldsymbol{\pi}_n(0)\mathbf{e} = \lambda. \tag{11}$$

The left hand side denotes the mean number of entrances into the system per unit time and is obviously equal to the mean arrival rate λ.

Proof Post-multiplying (8–10) by $\mathbf{e}$ and adding them, we obtain

$$\sum_{n=0}^{N} \boldsymbol{\pi}_n^*(\theta)\mathbf{e} = \frac{1 - A^*(\theta)}{\theta}\left(\sum_{n=0}^{N} \pi_n(0)\mathbf{e}\right). \tag{12}$$

Taking the limit as $\theta \to 0$ and applying normalization condition one can achieve the result of the lemma. □

4.1.1 Relation between steady-state distribution at arbitrary and pre-arrival epochs

We first relate the pre-arrival epoch probabilities $\boldsymbol{\pi}_n^-$ with $\boldsymbol{\pi}_n(0)$ which are given by

$$\boldsymbol{\pi}_n^- = \frac{1}{\lambda}\boldsymbol{\pi}_n(0), \quad 0 \le n \le N. \tag{13}$$

Now we will express arbitrary epoch probabilities in terms of pre-arrival epoch probabilities. Setting $\theta = 0$ in (10), (9), and using (13), we obtain

$$\boldsymbol{\pi}_N = \lambda \sum_{i=0}^{N-1} \boldsymbol{\pi}_i^- \sum_{j=N-i}^{\infty} g_j(-\mathbf{L})^{-1}, \tag{14}$$

$$\boldsymbol{\pi}_n = \left(\boldsymbol{\pi}_{n+1}\mathbf{M} + \lambda\left(\sum_{i=0}^{n-1} \boldsymbol{\pi}_i^- g_{n-i} - \boldsymbol{\pi}_n^-\right)\right)(-\mathbf{L})^{-1},$$
$$n = N-1, N-2, \ldots, 1. \tag{15}$$

It may be noted here that we do not have explicit expressions for $\boldsymbol{\pi}_0$. However, it can be computed using the normalization condition, i.e., $\boldsymbol{\pi}_0 = \overline{\boldsymbol{\pi}} - \sum_{n=1}^{N} \boldsymbol{\pi}_n$.

4.2 Total batch rejection

As we did in the earlier Sect. 4.1, here we also relate the states of the system at two consecutive time epochs t and $t + \Delta t$. Using probabilistic arguments and considering total batch rejection policy, we get a set of partial differential equations. Taking the limit as $t \to \infty$ and using matrices and vector notation, we obtain

$$-\frac{d}{dx}\boldsymbol{\pi}_0(x) = \boldsymbol{\pi}_1(x)\mathbf{M} + \boldsymbol{\pi}_0(0)a(x)\left(1 - \sum_{i=1}^{N} g_i\right), \tag{16}$$

$$-\frac{d}{dx}\boldsymbol{\pi}_n(x) = \boldsymbol{\pi}_n(x)\mathbf{L} + \boldsymbol{\pi}_{n+1}(x)\mathbf{M} + \sum_{i=0}^{n-1}\boldsymbol{\pi}_i(0)a(x)g_{n-i}$$
$$+\boldsymbol{\pi}_n(0)a(x)\left(1-\sum_{i=1}^{N-n}g_i\right), \quad 1 \le n \le N-1, \tag{17}$$

$$-\frac{d}{dx}\boldsymbol{\pi}_N(x) = \boldsymbol{\pi}_N(x)\mathbf{L} + \boldsymbol{\pi}_N(0)a(x) + \sum_{i=0}^{N-1}\boldsymbol{\pi}_i(0)a(x)g_{N-i}. \tag{18}$$

Multiplying (16–18) by $e^{-\theta x}$ and integrating w.r.t. x from 0 to ∞, we get

$$-\theta\boldsymbol{\pi}_0^*(\theta) = \boldsymbol{\pi}_1^*(\theta)\mathbf{M} + \boldsymbol{\pi}_0(0)A^*(\theta)\left(1-\sum_{i=1}^{N}g_i\right) - \boldsymbol{\pi}_0(0), \tag{19}$$

$$-\theta\boldsymbol{\pi}_n^*(\theta) = \boldsymbol{\pi}_n^*(\theta)\mathbf{L} + \boldsymbol{\pi}_{n+1}^*(\theta)\mathbf{M} + \sum_{i=0}^{n-1}\boldsymbol{\pi}_i(0)A^*(\theta)g_{n-i}$$
$$+\boldsymbol{\pi}_n(0)A^*(\theta)\left(1-\sum_{i=1}^{N-n}g_i\right) - \boldsymbol{\pi}_n(0), \quad 1 \le n \le N-1, \tag{20}$$

$$-\theta\boldsymbol{\pi}_N^*(\theta) = \boldsymbol{\pi}_N^*(\theta)\mathbf{L} + \boldsymbol{\pi}_N(0)A^*(\theta) + \sum_{i=0}^{N-1}\boldsymbol{\pi}_i(0)A^*(\theta)g_{N-i} - \boldsymbol{\pi}_N(0). \tag{21}$$

Adding (19–21), after some algebraic manipulation, and taking the limit as $\theta \to 0$ we get a similar result to Lemma 1.

4.2.1 *Relation between steady-state distribution at arbitrary and pre-arrival epochs*

As described in Sect. 4.1.1, arbitrary epoch probabilities are expressed in terms of pre-arrival epoch probabilities. Setting $\theta = 0$ in (21), (20), and using (13), we obtain

$$\boldsymbol{\pi}_N = \lambda\sum_{i=0}^{N-1}\boldsymbol{\pi}_i^- g_{N-i}(-\mathbf{L})^{-1}, \tag{22}$$

$$\boldsymbol{\pi}_n = \left(\boldsymbol{\pi}_{n+1}\mathbf{M} + \lambda\left(\sum_{i=0}^{n-1}\boldsymbol{\pi}_i^- g_{n-i} + \boldsymbol{\pi}_n^-\left(1-\sum_{i=1}^{N-i}g_i\right) - \boldsymbol{\pi}_n^-\right)\right)(-\mathbf{L})^{-1},$$
$$n = N-1, N-2, \ldots, 1. \tag{23}$$

$\boldsymbol{\pi}_0$ can be computed the same way as discussed before, i.e., $\boldsymbol{\pi}_0 = \overline{\boldsymbol{\pi}} - \sum_{n=1}^{N}\boldsymbol{\pi}_n$.

5 Performance measures

5.1 Partial batch rejection

As state probabilities at various epochs are known, performance measures can be easily obtained. The average number of customers in the system (queue) at an arbitrary epoch is given by

$$L=\sum_{i=0}^{N} i\boldsymbol{\pi}_i\mathbf{e}, \qquad L_q=\sum_{i=1}^{N}(i-1)\boldsymbol{\pi}_i\mathbf{e}.$$

One useful performance measure is the blocking probability. Since the pre-arrival epoch probabilities are known, the blocking probability of the first customer of an arriving batch is $(P_{BF})=\boldsymbol{\pi}_N^-\mathbf{e}$. To obtain the blocking probability of an arbitrary customer of a batch let us define the r.v. G^- that denotes the number of customers ahead of an arbitrary customer within the batch. The distribution of G^- is given by Takagi (1993, p. 419)

$$g_r^-=P[G^-=r]=\frac{1}{\bar{g}}\sum_{i=r+1}^{\infty} g_i, \quad r\geq 0.$$

Hence, the blocking probability of an arbitrary customer of a batch (P_{BA}) and the blocking probability of the last customer of a batch (P_{BL}) are given by

$$P_{BA}=\sum_{i=0}^{N}\sum_{j=N-i}^{\infty}\boldsymbol{\pi}_i^- g_j^-\mathbf{e}, \qquad P_{BL}=\sum_{i=0}^{N}\sum_{j=N-i+1}^{\infty}\boldsymbol{\pi}_i^- g_j\mathbf{e}. \tag{24}$$

Let w_A (w_{qA}) denote the average waiting time in the system (queue) of an arbitrary customer of a batch. Then, by Little's rule, $w_A=\frac{L}{\lambda'}$, $W_{qA}=\frac{L_q}{\lambda'}$, where $\lambda'=\lambda\bar{g}(1-P_{BA})$ is the effective arrival rate.

5.1.1 Waiting time analysis

In this section we obtain the LST of waiting time distribution of a customer who is accepted in the system. Let $\boldsymbol{\phi}_k(\theta)$ be the LST of the probability that k customers will be served within a time x and the service process upon completion of service passes to phase j, provided k customers were in the system and the service process was in phase i at the beginning of service. Since the probability that the service of a customer is completed in the interval $[x,x+dx]$ is given by the matrix $e^{-\mathbf{L}x}\mathbf{M}\,dx$ and the total service time of k customers is the sum of their service times,

$$\boldsymbol{\phi}_1(\theta)=\int_0^{\infty} e^{-\theta x}e^{\mathbf{L}x}\mathbf{M}\,dx=(\theta\mathbf{I}-\mathbf{L})^{-1}\mathbf{M}, \qquad \boldsymbol{\phi}_k(\theta)=\boldsymbol{\phi}_1^k(\theta), \quad k\geq 2. \tag{25}$$

Let $W_F^*(\theta)$, $W_A^*(\theta)$ and $W_L^*(\theta)$ denote the LST of the actual waiting time distribution of the first customer, an arbitrary customer and the last customer in an accepted batch, respectively. They can be derived as follows:

$$W_F^*(\theta) = \frac{1}{1-P_{BF}}\left[\sum_{n=0}^{N-1} \boldsymbol{\pi}_n^- \boldsymbol{\phi}_1^{n+1}(\theta)\mathbf{e}\right],$$

$$W_A^*(\theta) = \frac{1}{1-P_{BA}}\left[\sum_{n=0}^{N-1} \boldsymbol{\pi}_n^- \sum_{j=0}^{N-n-1} g_j^- \boldsymbol{\phi}_1^{n+j+1}(\theta)\mathbf{e}\right],$$

$$W_L^*(\theta) = \frac{1}{1-P_{BL}}\left[\sum_{n=0}^{N-1} \boldsymbol{\pi}_n^- \sum_{j=1}^{N-n} g_j \boldsymbol{\phi}_1^{n+j}(\theta)\mathbf{e}\right].$$

One can find mean waiting times in the system of the above three types (w_F, w_A, and w_L, respectively) through differentiating the above formulae and putting $\theta = 0$. They are given by

$$w_F = \frac{1}{1-P_{BF}}\left[\sum_{n=0}^{N-1} \boldsymbol{\pi}_n^- \sum_{j=0}^{n}\left(-\mathbf{L}^{-1}\mathbf{M}\right)^j\left(-\mathbf{L}^{-1}\right)\mathbf{e}\right], \tag{26}$$

$$w_A = \frac{1}{1-P_{BA}}\left[\sum_{n=0}^{N-1} \boldsymbol{\pi}_n^- \sum_{j=0}^{N-n-1}\sum_{k=0}^{n+j} g_j^-\left(-\mathbf{L}^{-1}\mathbf{M}\right)^k\left(-\mathbf{L}^{-1}\right)\mathbf{e}\right], \tag{27}$$

$$w_L = \frac{1}{1-P_{BL}}\left[\sum_{n=0}^{N-1} \boldsymbol{\pi}_n^- \sum_{j=1}^{N-n}\sum_{k=0}^{n+j-1} g_j\left(-\mathbf{L}^{-1}\mathbf{M}\right)^k\left(-\mathbf{L}^{-1}\right)\mathbf{e}\right]. \tag{28}$$

5.2 Total batch rejection

The average system (queue) length L (L_q) is the same as given in Sect. 5.1. But the blocking probabilities will differ. The blocking probability of the first customer (or, equivalently, the blocking probability of the whole batch or, equivalently, the blocking probability of the last customer), and the blocking probability of an arbitrary customer of a batch are given by

$$P_{BF} = P_{BL} = \sum_{i=0}^{N} \boldsymbol{\pi}_i^- \sum_{j=N-i+1}^{\infty} g_j\mathbf{e}, \qquad P_{BA} = \sum_{i=0}^{N} \boldsymbol{\pi}_i^- \sum_{j=N-i+1}^{\infty} \frac{jg_j}{\bar{g}}\mathbf{e}. \tag{29}$$

5.2.1 Waiting time analysis

Like in Sect. 5.1.1, one can obtain the LST of the actual waiting time distributions of the first customer and an arbitrary customer in an accepted batch, respectively. They are given as

$$W_F^*(\theta) = \frac{1}{1-P_{BF}}\left[\sum_{n=0}^{N-1} \boldsymbol{\pi}_n^- \sum_{j=1}^{N-n} g_j \boldsymbol{\phi}_1^{n+1}(\theta)\mathbf{e}\right],$$

$$W_A^*(\theta) = \frac{1}{1-P_{BA}}\left[\sum_{n=0}^{N-1}\boldsymbol{\pi}_n^- \sum_{j=1}^{N-n}\sum_{r=0}^{j-1}\frac{g_j}{\bar{g}}\boldsymbol{\phi}_1^{n+r+1}(\theta)\mathbf{e}\right].$$

Their means are given by

$$w_F = \frac{1}{1-P_{BF}}\left[\sum_{n=0}^{N-1}\boldsymbol{\pi}_n^- \sum_{j=1}^{N-n}\sum_{k=0}^{n} g_j\left(-\mathbf{L}^{-1}\mathbf{M}\right)^k\left(-\mathbf{L}^{-1}\right)\mathbf{e}\right], \tag{30}$$

$$w_A = \frac{1}{1-P_{BA}}\left[\sum_{n=0}^{N-1}\boldsymbol{\pi}_n^- \sum_{j=1}^{N-n}\sum_{r=0}^{j-1}\sum_{k=0}^{n+r}\frac{g_j}{\bar{g}}\left(-\mathbf{L}^{-1}\mathbf{M}\right)^k\left(-\mathbf{L}^{-1}\right)\mathbf{e}\right]. \tag{31}$$

6 Numerical result and discussion

To demonstrate the applicability of the results obtained in the previous sections some numerical results have been presented in self explanatory tables and graphs. In the bottom of the tables various performance measures are given. Since the results reported here were rounded, the sum of the elements of probabilities may not add to one.

In Table 1 we have presented the steady-state distributions at various epochs in a $HE_2^X/E_5/1/10$ queue with the following parameters: inter-batch arrival time is HE_2 and its PH-type representation is given by $\boldsymbol{\alpha} = [0.2 \quad 0.8]$, $\mathbf{T} = \begin{bmatrix} -0.2 & 0.0 \\ 0.0 & -0.3 \end{bmatrix}$ with $\lambda = 0.272727$, batch size distribution is taken as $g_1 = 0.2$, $g_2 = 0.3$, $g_3 = 0.5$, i.e., mean batch size $\bar{g} = 2.3$, *MSP* representation of E_5 is taken as

$$\mathbf{L} = \begin{bmatrix} -5.0 & 5.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & -5.0 & 5.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -5.0 & 5.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & -5.0 & 5.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & -5.0 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 5.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix},$$

and $\mathbf{f} = [1.0 \quad 0.0 \quad 0.0 \quad 0.0 \quad 0.0]$ with stationary mean service rate $\mu^* = 1.0$, so that $\rho = 0.627273$.

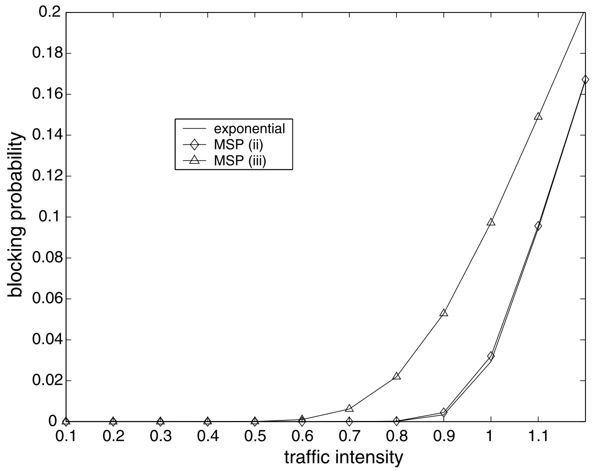
Here one can see that the arbitrary epoch probabilities presented in the third and the fifth columns of this table exactly match with the one given in Gupta and Vijaya (2000, p. 168, Table 2). Moreover in all numerical experiments, mean waiting time w_A obtained by Little's rule exactly matches with the one obtained by the direct method given in Sect. 5.

In Figs. 1 and 2 we have plotted blocking probability of an arbitrary customer in a $E_2^X/MSP/1/50$ queue against traffic intensity for partial and total batch rejection model, respectively. Batch size distribution is taken as $g_1 = 0.2$, $g_2 = 0.5$, $g_3 = 0.1$, $g_5 = 0.2$, so that $\bar{g} = 2.5$. Three types of service time distributions are assumed, viz., (i) exponential, (ii) *MSP* with lag-2 correlation equal to 0.000050, and (iii) *MSP* with lag-2 correlation equal to 0.143181. For (ii) the *MSP* representation is taken as $\mathbf{L} = \begin{pmatrix} -1.5125 & 0.750 \\ 0.875 & -1.025 \end{pmatrix}$, $\mathbf{M} = \begin{pmatrix} 0.7625 & 0.000 \\ 0.125 & 0.025 \end{pmatrix}$. *MSP* representation for (iii) is taken to

Table 1 Distribution of the number of customers in the system for $HE_2^X/E_5/1/10$ queue

	Partial rejection		Total rejection	
n	$\boldsymbol{\pi}_n^- \mathbf{e}$	$\boldsymbol{\pi}_n \mathbf{e}$	$\boldsymbol{\pi}_n^- \mathbf{e}$	$\boldsymbol{\pi}_n \mathbf{e}$
0	0.3785	0.3876	0.3827	0.3918
1	0.1164	0.1152	0.1177	0.1165
2	0.1269	0.1251	0.1283	0.1265
3	0.1200	0.1181	0.1214	0.1194
4	0.0731	0.0720	0.0739	0.0728
5	0.0590	0.0581	0.0597	0.0588
6	0.0440	0.0433	0.0446	0.0438
7	0.0319	0.0314	0.0323	0.0317
8	0.0239	0.0235	0.0228	0.0223
9	0.0177	0.0173	0.0126	0.0123
10	0.0085	0.0083	0.0041	0.0040
sum	1.0000	1.0000	1.000000	1.000000
$P_{BF} =$	0.008508		0.025543	
$P_{BA} =$	0.023722		0.030437	
$P_{BL} =$	0.034645		0.025543	
$L =$	2.204885		2.127259	
$w_A =$	3.600445		3.497743	

Fig. 1 ρ versus P_{BA} (partial rejection)



be $\mathbf{L} = \begin{pmatrix} -6.9375 & 0.9375 \\ 0.0625 & -0.1958 \end{pmatrix}$, $\mathbf{M} = \begin{pmatrix} 6.0 & 0.0 \\ 0.0 & 0.1333 \end{pmatrix}$. They all have equal mean service rate $\mu^* = 0.5$, and also the same $\mathbf{f} = (0.3\ 0.7)$ for the cases (ii) and (iii), and for the case (i) $\mathbf{f} = 1.0$. Inter-batch arrival time is E_2 and its PH representation is given by $\alpha = (1.0\ 0.0)$, $\mathbf{T} = \begin{pmatrix} -\gamma & \gamma \\ 0.0 & -\gamma \end{pmatrix}$, where $\lambda = \gamma/2.0$. We suitably vary γ and as a result achieve various values of ρ. From Fig. 1 it can be observed that up to certain extent of

Fig. 2 ρ versus P_{BA} (total rejection)

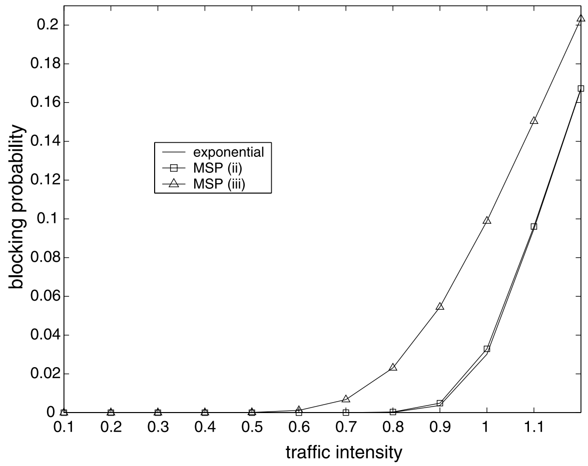
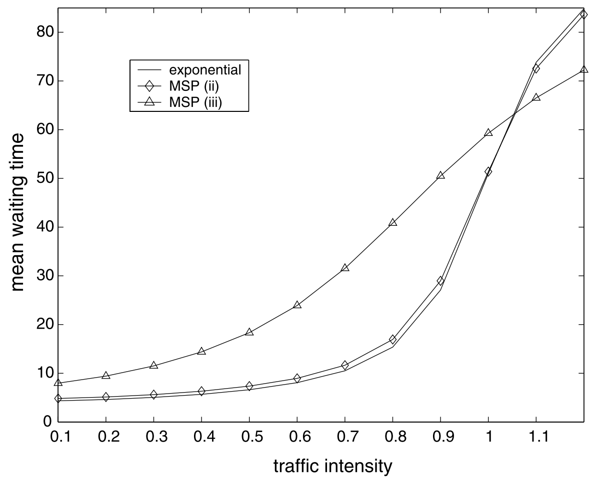


Fig. 3 ρ versus w_A (partial rejection)



an increase of traffic load (say, $\rho = 0.5$) all the blocking probabilities are low. After that we observe that the blocking probability for the case of highly correlated *MSP* is greater than for the other two. Similar behavior has been observed in Fig. 2. This is due to the fact that for large buffer size partial and total batch rejection policies act similarly.

In Figs. 3 and 4 we have plotted mean waiting time (w_A) of an arbitrary customer against traffic intensity under the above condition. From Fig. 3 it can be observed that the mean waiting time for highly correlated *MSP*, case (iii), is greater than the corresponding mean waiting time of cases (ii) and (i). But they merge at around traffic load equal to 1.1; after that we observe opposite behavior of mean waiting time, that is, w_A for highly correlated *MSP* is smaller than for the other two. Quite naturally Fig. 4 exhibits almost the same behavior.

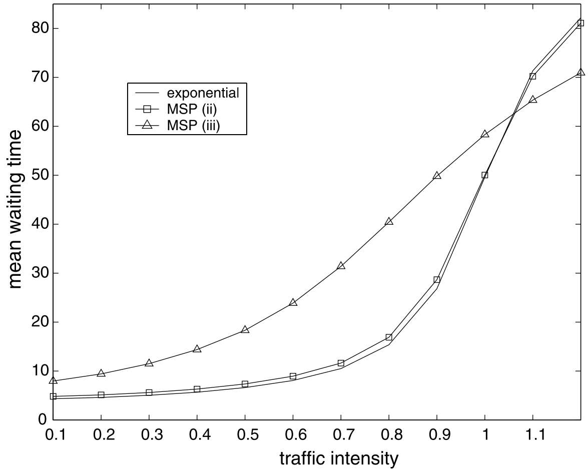


Fig. 4 ρ versus w_A (total rejection)

Acknowledgements The authors would like to thank the referees for their valuable comments and suggestions. The first author wishes to thank the Council of Scientific and Industrial Research (CSIR), New Delhi, India for its financial support.

References

Albores-Velasco FJ, Tajonar-Sanabria FS (2004) Analysis of the $GI/MSP/c/r$ queueing system. Inf Process 4:46–57

Alfa AS, Qiang JX, Ye Q (2000) Perturbation theory for the asymptotic decay rates in the queues with Markovian arrival process and/or Markovian service process. Queueing Syst 36:287–301

Baba Y (1984) The $M^X/G/1$ queue with finite waiting room. J Oper Res Soc Jpn 27:260–272

Bocharov PP (1996) Stationary distribution of a finite queue with recurrent flow and Markovian service. Autom Remote Control 57:66–78

Bocharov PP, D'Apice C, Pechinkin A, Salerno S (2003) The stationary characteristics of the $G/MSP/1/r$ queueing system. Autom Remote Control 64:288–301

Chakravarthy S (1993) Analysis of a finite $MAP/G/1$ queue with group services. Queueing Syst 13:385–407

Choi BD, Hwang GU, Han DH (1998) Supplementary variable method applied to the $MAP/G/1$ queueing system. J Aust Math Soc B 40:86–96

Ferrandiz JM (1993) The $BMAP/G/1$ queue with server set-up times and server vacations. Adv Appl Probab 25:235–254

Grassmann WK, Taksar MI, Heyman DP (1985) Regenerative analysis and steady state distributions for Markov chains. Oper Res 33:1107–1116

Gupta UC, Vijaya Laxmi P (2000) On the finite buffer bulk arrival $GI^X/E_k/1/N$ queueing system. Opsearch 37:154–169

Gupta UC, Vijaya Laxmi P (2001) Analysis of $MAP/G^{a,b}/1/N$ queue. Queueing Syst 38:109–124

Gupta UC, Banik AD (2007) Complete analysis of finite and infinite buffer $GI/MSP/1$ queue—a computational approach. Oper Res Lett 35:273–280

Kasahara S, Takine T, Takahashi Y, Hasegawa T (1996) $MAP/G/1$ queues under N-policy with and without vacations. J Oper Res Soc Jpn 39:188–212

Lucantoni DM (1991) New results on the single server queue with a batch Markovian arrival process. Stoch Model 7:1–46

Lucantoni DM, Ramaswami V (1985) Efficient algorithms for solving non-linear matrix equations arising in phase type queues. Stoch Model 1:29–52

Lucantoni DM, Meier-Hellstern KS, Neuts MF (1990) A single-server queue with server vacations and a class of non-renewal process. Adv Appl Probab 22:676–705

Machihara F (1995) A $G/SM/1$ queue with vacations depending on service times. Stoch Model 11:671–690
Machihara F (1999) A $BMAP/SM/1$ queue with service times depending on the arrival process. Queueing Syst 33:277–291
Neuts MF (1981) Matrix-geometric solutions in stochastic models: an algorithmic approach. Johns Hopkins Univ Press, Baltimore
Niu Z, Shu T, Takahashi Y (2003) A vacation queue with set up and close-down times and batch Markovian arrival processes. Perform Eval 54:225–248
Nobel RD (1989) Practical approximations for finite buffer queueing models with batch arrivals. Eur J Oper Res 38:44–55
Ohsone T (1981) The $GI/E_k/1$ queue with finite waiting room. J Oper Res Soc Jpn 24:375–391
Takagi H (1993) Queueing analysis—a foundation of performance evaluation, finite systems: vol 2. North-Holland, New York