RAIRO-Oper. Res. 54 (2020) 675–691
https://doi.org/10.1051/ro/2019020

RAIRO Operations Research
www.rairo-ro.org

THE ANALYSIS OF A DISCRETE TIME FINITE-BUFFER QUEUE WITH WORKING VACATIONS UNDER MARKOVIAN ARRIVAL PROCESS AND PH-SERVICE TIME

Qingqing Ye[1,*], Liwei Liu[2], Tao Jiang[3] and Baoxian Chang[4]

Abstract. In this paper, we study the discrete-time MAP/PH/1 queue with multiple working vacations and finite buffer N. Using the Matrix-Geometric Combination method, we obtain the stationary probability vectors of this model, which can be expressed as a linear combination of two matrix-geometric vectors. Furthermore, we obtain some performance measures including the loss probability and give the limit of loss probability as finite buffer N goes to infinite. Waiting time distribution is derived by using the absorbing Markov chain. Moreover, we obtain the number of customers served in the busy period. At last, some numerical examples are presented to verify the results we obtained and show the impact of parameter N on performance measures.

Mathematics Subject Classification. 60K25, 68M20.

Received August 16, 2015. Accepted February 5, 2019.

1. Introduction

Vacation queues have been investigated by many authors for their wide applicability in the performance analysis of the computer networks, production management, communication and so forth. The readers may refer to surveys of Doshi [13] (prior to 1985) and to the book by Tian and Zhang [33] for works through 2006, as well as references therein.

In the real life, we may see many situations where the server can still work when the system become empty, for example, the escalators in the large supermarkets and metros, which are always designed to operate at a lower rate when there are no customers in the system in order to save the cost, the computers and usually switch to Stand By or sleep mode when there is no work to do to save the battery life. All the above behaviors can be viewed as the working vacation, which is first introduced by Servi and Finn [32]. The working vacation is a class of semi-vacation during which the customers are served at a lower rate rather than completely stopping the service, Servi and Finn [32] studied an M/M/1 queue with working vacations and applied their results to analyze a WDM optical access network using multiple wavelengths which can be reconfigured. Their model was generalized to the case of M/G/1 in Wu and Takagi [36], Kim *et al.* [19] and Li *et al.* [24], and to the case of GI/M/1 model in Baba [6] and Li and

Keywords. Working vacation, finite buffer, matrix-geometric combination method, Sojourn time, busy period.

[1] School of Science, Nanjing University of Information Science and Technology, Nanjing, P.R. China.

[2] School of Science, Nanjing University of Science and Technology, Nanjing, P.R. China.

[3] College of Economics and Management, Shandong University of Science and Technology, Qingdao, P.R. China.

[4] School of Mathematical and Physical Sciences, Nanjing Tech University, Nanjing, P.R. China.

*Corresponding author: yeqingzero@gmail.com

Tian [22]. A survey of working vacation models with emphasis on the use of matrix methods and some applications in practice can be found in Tian *et al.* [35]. Similarly, for the discrete-time queues with working vacations, Tian *et al.* [34] studied the Geom/Geom/1 queue with working vacations using quasi-birth-death process and matrix-geometric solution method. Subsequently, extension to GI/Geom/1 type with working vacation is carried out in Li *et al.* [22]. Recently, Gao *et al.* [16], Luo *et al.* [26], Yang and Wu [37] and others considered the working vacation queueing systems with various features. More details and recent works related to working vacation, the readers can refer to the survey of Chandrasekaran *et al.* [12].

A discrete MAP/PH/1 queue is a queue having the Markovian arrival process (MAP) and a single server with phase-type (PH-type) distributed service time. The discrete Markovian arrival process (D-MAP), which is introduced by Neuts [28], is convenient to capture the properties of both burstiness and correlation arising in telecommunication networks based on the Asynchronous Transfer Mode (ATM) environment. The ATM has been applied as the transport mechanism for the implementation of Broadband Integrated Services Digital Networks (B-ISDN) and provides high flexibility of network access, dynamic bandwidth allocation on demand, and flexible bearer capacity allocation. For further details on D-MAP and their applications in stochastic modeling, we refer the readers to Bruneel and Kim [8] Lucantoni [25], Neuts [30, 31] and for a review and recent works on MAP, we can refer to Chakravarthy [9] and Artalejo and He [5]. The discrete PH distribution, which is introduced in Neuts [27], is the generalization of geometric distribution, and the detail properties and applications related to discrete PH distribution, we may refer to Neuts [29] and Latouche and Ramaswami [21].

For the MAP/PH/1 queue related to vacation policy, Alfa [1] first analyzed a discrete-time MAP/PH/1 queue with exhaustive and non-exhaustive service using the matrix-geometric procedure, where the vacation time follows a phase-type distribution. Subsequently, Alfa [2] studied the discrete MAP/PH/1 queue with gated time-limited service. In Gail *et al.* [15], the discrete time MAP/PH/1 with multiple working vacations is analyzed. Recently, Chakravarthy and Ozkar [11]. By taking advantages of matrix-analytic method, all the above papers develop computation methods for computing various performance measures.

Systems with finite buffer capacity are commonly seen in practical systems due to space or process constraints. Theoretically, all physical queues have finite-capacity buffers and the buffer sizes can be small in some circumstances. From the point of view, there is a need to study the MAP/PH/1 queue with finite buffer and working vacations. To the author's knowledge, up to the present, no special works focused on this model have appeared in open literatures. For the MAP/PH/1 queue with finite buffer capacity have been widely studied by many authors, for example, Chakravarthy analyze the MAP/PH/1/K queue with service control, Heindl and Telek [18] studied the Output models of MAP/PH/1/K queues for an efficient network decomposition, Dudin *et al.* [14] analyzed the MAP/PH/1/N queue with feedback operating in a Markovian random environment. Kim *et al.* [20] studied the MAP/PH/1/N queue with flows of customers as a model for traffic control in telecommunication networks.

This remainder is structured as follows. In Section 2, we provide a description of queueing model under study. The steady-state analysis of the model is presented in Section 3. Specifically, in Section 3.1, we formulate the model under study as a finite quasi-birth-death (QBD) process, and probability vectors in steady state are provided in Section 3.2, furthermore, we analyze some properties about the two key matrixes which can be used in obtaining the stationary probability vectors in Section 3.3, and give some decomposition results which are useful to reduce the computation workload in Section 3.4. Section 4 obtains some useful performance measures and gives the limit of loss probability as the buffer N goes to infinite. The waiting time distribution is derived in Section 5, and the number of customers served in busy period is analyzed in Section 6. At last, a few numerical examples are presented in Section 7.

2. Model description

We consider a discrete-time single-server queue with working vacation and finite buffer. The queueing model is defined explicitly as follows.

Customers arrives at the system according to a discrete time Markovian arrival process, which can be described by two n dimensional substochastic matrices D_0 and D_1. Here, D_0 records the transition probability that phase transits from i to j without an arrival, and D_1 records the transition probability that phase transits from i to j with an arrival. We assume that the matrix $D = D_0 + D_1$ is an irreducible stochastic matrix and has the stationary distribution vector θ, *i.e.*, $\theta D = \theta$ and $\theta e = 1$, where e is the column vector of ones of appropriate dimension, then the stationary arrive rate can be given by $\lambda = \theta D_1 e$.

During the normal service period, the customers are served according to a phase type distribution with representation $(\beta,\ S_1)$ of order m_1. Let $S_1^0 = e - S_1 e$ and we assume that $S_1 + S_1^0\beta$ is irreducible and let ξ_1 be the row vector satisfying $\xi_1\left(S_1 + S_1^0\beta\right) = \xi_1$ and $\xi_1 e = 1$, then the service rate during the regular service period can be expressed as $\mu_b = \xi_1\left(S_1^0\beta\right)e$. When the system becomes empty, the server takes a working vacation that follows a phase type distribution with representation (α, T) of order r. During the working vacation period, the server can still serve the arriving customers, but the service time is of a PH-distribution with representation $(\delta,\ S_2)$ of order m_1. We also let $S_2^0 = e - S_2 e$ and assume $S_2 + S_2^0\beta$ is irreducible and let ξ_2 be the row vector satisfying $\xi_2\left(S_2 + S_2^0\beta\right) = \xi_2$ and $\xi_2 e = 1$, then the service rate during the working vacation period can be expressed as $\mu_v = \xi_2\left(S_2^0\beta\right)e$. Here, we assume $\mu_v < \mu_b$. When the working vacation ends, if there are no customers in the system, the sever proceeds for another working vacation until the server finds the customers waiting in the system upon arrival, the server will then resume to the normal service period.

We assume the system has a finite buffer N, any customers who meets a full buffer is rejected and lost. And the service order is first-come first served *i.e.*, FCFS. Furthermore, the inter-arrival times, service times and working vacation times are assumed mutually independent.

3. Stationary analysis

3.1. The finite quasi-birth-and-death process

In a discrete-time queueing system, the arrivals, the departures and the ends of vacations may happen at the same time. We assume that the time axis is allotted into intervals of queue length with the length of a slot being unity, to be more specific, we can let the time axis be marked by $0, 1, 2, \ldots, n$. In this paper, we analyze the model for early arrival system (EAS), that is, the potential arrival occurs in (n, n^+), where n^+ is the moment immediately after n, and the potential departure takes place in (n^-, n), where n^- represents the moment immediately before n, in addition, the beginning and ending of the working vacations also take places at the instant n^+.

For simplicity, we denote the finite buffer MAP/PH/1 queueing system with working vacations by F. Let Q_n be the number of customers in the system at $t = n^+$ and J_n be the state of server at $t = n^+$ with

$$J_n = \begin{cases} 0, & \text{the server is in working vacation period at time } n^+, \\ 1, & \text{the server is in regular busy period at time } n^+. \end{cases}$$

Denoted by V_n the phase of working vacation at time n^+, K_n^w the phase of service in working vacation at time n^+, K_n^b the phase of service in regular busy period at time n^+ and M_n the phase of service at time n^+, then, F can be characterized by a multi-dimensional discrete-time Markov process

$$Z_n = \left\{Q_n, (J_n, M_n, (V_n, K_n^w)) \cup \left(J_n, M_n, K_n^b\right), n = 1, 2, \ldots\right\},$$

with state space $\Delta = \Delta_0 \cup \Delta_1$, where

$$\Delta_0 = \{(0, 0, j, v), j = 1, 2, \ldots, m, v = 1, 2, \ldots, r\},$$

$$\Delta_1 = \{k, (0, j, v, k_1) \cup (1, j, k_2), 1 \le k \le N, 1 \le j \le n, 1 \le v \le r, 1 \le k_1 \le m_1, 1 \le k_2 \le m_2\}.$$

Δ_0 represents that the states where there are no customers in the system and the server is on working vacation, the MAP is in phase j, and the vacation is in phase v, Δ_1 represents the case in which there are k customers in the system. The four tuple represents the case where the server is in working vacation with the MAP is in phase j, the vacation is in phase v, and the service in working vacation is in phase k_1. The three tuple represents the case where the server is in regular busy period with the MAP is in phase j and the service in regular busy period is in phase k_2.

Using the lexicographical order of the state, we get the transition probability matrix P of the process F as follows.

$$P = \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ N-1 \\ N \end{matrix} \begin{bmatrix} B_{00} & B_{01} & & & & \\ B_{10} & A_1 & A_0 & & & \\ & A_2 & A_1 & \cdots & & \\ & & & \cdots & A_0 & \\ & & & \cdots & A_1 & A_0 \\ & & & & A_2 & C \end{bmatrix}, \tag{3.1}$$

where

$$\begin{aligned} B_{00} &= D_0 \otimes \left(T + T^0\alpha\right), B_{01} = \left[D_1 \otimes T \otimes \delta,\ D_1 \otimes T^0\beta\right], \\ B_{10} &= \begin{bmatrix} D_0 \otimes \left(T + T^0\alpha\right) \otimes S_2^0 \\ D_0 \otimes S_1^0\alpha \end{bmatrix}, \\ A_0 &= \begin{bmatrix} D_1 \otimes T \otimes S_2 & D_1 \otimes T^0 \otimes S_2 e\beta \\ 0 & D_1 \otimes S_1 \end{bmatrix}, \\ A_2 &= \begin{bmatrix} D_0 \otimes T \otimes S_2^0\delta & D_0 \otimes T^0 \otimes S_2^0\beta \\ 0 & D_0 \otimes S_1^0\beta \end{bmatrix}, \\ A_1 &= \begin{bmatrix} D_0 \otimes T \otimes S_2 + D_1 \otimes T \otimes S_2^0\delta & D_0 \otimes T^0 \otimes S_2 e\beta + D_1 \otimes T^0 \otimes S_2^0\beta \\ 0 & D_0 \otimes S_1 + D_1 \otimes S_1^0\beta \end{bmatrix}, \\ C &= \begin{bmatrix} (D_0 + D_1) \otimes T \otimes S_2 + D_1 \otimes T \otimes S_2^0\delta & (D_0 + D_1) \otimes T^0 \otimes S_2 e\beta + D_1 \otimes T^0 \otimes S_2^0\beta \\ 0 & (D_0 + D_1) \otimes S_1 + D_1 \otimes S_1^0\beta \end{bmatrix}. \end{aligned}$$

The symbol $\otimes$ is the Kronecker product, the block matrices A_0, A_1, A_2 and C are square matrices of dimensions $(nm_1 + nrm_2)$, the block matrix B_{00} is the square matrix with dimension nr, the block matrix B_{01} is of dimension $nr \times (nm_1 + nrm_2)$, the block matrix B_{10} is of the dimension $(nm_1 + nrm_2) \times nr$. The structure of the transition probability matrix P is a finite quasi-birth-and-death process (QBD).

We define the traffic intensity

$$\rho = \frac{\lambda}{\mu_b} \tag{3.2}$$

and assume $\rho \neq 1$.

We now show that $\rho < 1$ makes the variant of (3.1) with infinite number of buffer positive recurrent. In fact, since $A_i, i = 0, 1, 2$, are upper triangular so that the matrix $A = A_0 + A_1 + A_2$ is reducible, stochastic matrix and

$$A = \begin{bmatrix} D \otimes T \otimes \left(S_2 + S_2^0\delta\right) & D \otimes T^0 \otimes \left(S_2 e\beta + S_2^0\beta\right) \\ 0 & D \otimes \left(S_1 + S_1^0\beta\right) \end{bmatrix}. \tag{3.3}$$

For simplicity and ease of notations, we let s represent the service stage and v the working vacation stage, and we can write $A = \left(A^{ij}\right)$ and $A_k = \left(A_k^{ij}\right)$, $i, j = s, v, k = 0, 1, 2$, *i.e.*,

$$A = \begin{bmatrix} A^{vv} & A^{vs} \\ & A^{ss} \end{bmatrix}, \quad A_k = \begin{bmatrix} A_k^{vv} & A_k^{vs} \\ 0 & A_k^{ss} \end{bmatrix}, \quad k = 0, 1, 2. \tag{3.4}$$

Based on Theorem 7.3.1 in Latouche and Ramaswami [21], we know the corresponding infinite QBD process is positive recurrent if and only if

$$\hat{\pi} A_2^{ss} e > \hat{\pi} A_0^{ss} e, \tag{3.5}$$

where $\hat{\pi} = \theta \otimes \xi_1$ such that $\hat{\pi} A^{ss} = \hat{\pi}$ and $\hat{\pi} e = 1$. From (3.5), we know

$$(\theta \otimes \xi_1) \left(D_0 \otimes S_1^0 \beta\right) e > (\theta \otimes \xi_1) \left(D_1 \otimes S_1\right) e,$$

then

$$(\theta D_0) \otimes \left(\xi_1 S_1^0 \beta\right) e > (\theta D_1) \otimes (\xi_1 S_1) e,$$

adding term $(\theta D_1) \otimes \left(\xi_1 S_1^0 \beta\right) e$ on both sides of the above inequality gives

$$(\theta D) \otimes \left(\xi_1 S_1^0 \beta\right) e > (\theta D_1) \otimes \left(\xi_1 \left(S_1 + S_1^0 \beta\right)\right) e,$$

then

$$(\theta D e) \otimes \left(\xi_1 S_1^0 \beta e\right) > (\theta D_1 e) \otimes \left(\xi_1 \left(S_1 + S_1^0 \beta\right) e\right),$$

which implies $\mu_b > \lambda$, then $\rho < 1$.

3.2. The stationary probability vectors

Since the process F is a finite QBD process, so we can analyze it by the Matrix-geometric Combination method (see Chapt. 10 of Latouche and Ramaswami [21]). To this end, we should solve the minimum non-negative solution R_1 to the matrix equation

$$A_0 + RA_1 + R^2 A_2 = R, \tag{3.6}$$

and the minimum non-negative solution R_2 to the matrix equation

$$A_2 + RA_1 + R^2 A_0 = R. \tag{3.7}$$

The computation of R_1 and R_2 can be carried out by a number of well known methods, for example, we provide the following stepwise procedure to calculate R_1 and R_2, recursively.

Algorithm 1. Linear algorithm to compute the R_1 and R_2.

Input: transition blocks $\{A_0, A_1, A_2\}$ and the tolerance ε.
Output: R_1 and R_2.
1: Let $R_1[0] = 0$, and $R_2[0] = 0$.
2: Let $R_2[1] = A_0 + R_2[0] A_1 + R_2[0]^2 A_2$, $R_2[1] = A_2 + R_2[0] A_1 + R_2[0]^2 A_0$.
3: If $||R_1[1] - R_1[0]|| \geq \varepsilon$ or $||R_2[1] - R_2[0]|| \geq \varepsilon$, go to Step 4, else $R_1[1] \leftarrow R_1[0]$, $R_2[1] \leftarrow R_2[0]$, then return Step 2.
4: Set $R_1 = R_1[0]$ and $R_2 = R_2[0]$.
5: **return** R_1, R_2.

Note: for a matrix $A = (a_{i,j})_{n \times n}$, $||A|| = \max_{1 \leq j \leq n} \left\{ \sum_{i=1}^{n} |a_{i,j}| \right\}$.

Let π be the stationary probability vector associated with transition matrix P such that

$$\pi P = \pi, \quad \pi e = 1. \tag{3.8}$$

We partition the vector π into $\pi = [\pi_0, \pi_1, \pi_2, \ldots \pi_N]$, where π_0 is a row nr-vector representing the state that there is no customers in the system, π_k $(k = 1, 2, \ldots, N)$ is a row $(nrm_1 + nm_2)$-vector representing the there are k customers in system. Based on the Theorem 10.3.2 of Latouche and Ramaswami [21], we can have that if $\rho \neq 1$, the stationary probabilities expression can be given as follows:

$$\pi_k = v_1 R_1^{k-1} + v_2 R_2^{N-k}, 1 \leq k \leq N, \tag{3.9}$$

and $[\pi_0, v_1, v_2]$ is the left invariant eigenvector of $B[R_1,\ R_2]$, *i.e.*,

$$[\pi_0, v_1, v_2]\, B\,[R_1, R_2] = [\pi_0, v_1, v_2]\,, \tag{3.10}$$

normalized so that

$$\pi_0 e + (v_1 S_1' + v_2 S_2')\, e = 1, \tag{3.11}$$

where

$$B\,[R_1, R_2] = \begin{bmatrix} B_{00} & B_{01} & 0 \\ B_{10} & A_1 + R_1 A_2 & R_1^{N-2}\,(A_0 + R_1 C - R_1) \\ R_2^{N-1} B_{10} & R_2^{N-2}\,(A_2 + R_2 A_1 - R_2) & R_2 A_0 + C \end{bmatrix}, \tag{3.12}$$

and

$$S_1' = \sum_{k=0}^{N-1} R_1^k, \quad S_2' = \sum_{k=0}^{N-1} R_2^k. \tag{3.13}$$

3.3. Spectrum properties of R_1 and R_2

In order to simplify the computation procedure for computing the stationary probability vectors of our model, we follow the analysis in Akar *et al.* [4] to discuss some spectrum properties for R_1 and R_2 under the condition $\rho < 1$ and $\rho > 1$, respectively.

3.3.1. $\rho < 1$

If $\rho < 1$, it can be shown that the eigenvalues of R_1 is a power-summable matrix and all its eigenvalues lie in the open unit disk. R_2 is a power-bounded matrix, all its eigenvalues lie in the closed unit disk, one of which is 1, all other lie in the open unit disk. Because R_1 is power-summable, we can know that $(I - R_1)$ is nonsingular, where I refers to an identity matrix, then

$$S_1' = \sum_{k=0}^{N-1} R_1^k = \left(I - R_1^N\right)(I - R_1)^{-1}. \tag{3.14}$$

However, since one of eigenvalues of R_2 is 1, so a simplification form for S_2' as (3.14) is impossible. Let

$$y_1 = (A_2 - R_2 A_0)\, e, \tag{3.15}$$

and

$$M_1 = \frac{y_1 x}{x y_1}, \tag{3.16}$$

where x is the we can show that left invariant vector of A, which is also a left invariant eigenvector of R_1, *i.e.*, $xR_2 = x$. It is also readily to show that y_1 is a right invariant eigenvector of R_2, *i.e.*, $R_2 y_1 = y_1$. in fact,

$$\begin{aligned} R_2 y_1 &= \left(R_2 A_2 - R_2^2 A_0\right) e \\ &= (R_2 A_2 - R_2 + A_2 + R_2 A_1)\, e \\ &= (R_2\,(A_2 + A_1 - I) + A_2)\, e \\ &= (A_2 - R_2 A_0)\, e \\ &= y_1, \end{aligned}$$

then we can find the rank of M_1 is one, and the matrix $\left(I - \hat{R}_2\right)$ is invertible, where $\hat{R}_2 = R_2 - M_1$. By induction, we can obtain

$$R_2^k = \begin{cases} \hat{R}_2^k + M, & k \geq 1, \\ I, & k = 0. \end{cases} \tag{3.17}$$

Then S_2' can be explicitly expressed as follows:

$$S_2' = \sum_{k=0}^{N-1} R_2^k = \sum_{k=0}^{N-1} \hat{R}_2^k + (N-1)\, M_1 = \left(I - \hat{R}_2^N\right)\left(I - \hat{R}_2\right)^{-1} + (N-1)\, M_1. \tag{3.18}$$

3.3.2. $\rho > 1$

When $\rho > 1$, the form of the stationary probability vectors are still same as in (3.9), but there is a little difference on the properties of the matrix geometric factors R_1 and R_2. Following the analysis in Akar *et al.* [4], under the condition $\rho > 1$, R_1 becomes a power-bounded matrix and has a simple eigenvalue one, and R_2 become a power-summable matrix and has all its eigenvalues inside the unit disk. Therefore the expression (3.14) is not valid for R_1 under the condition $\rho > 1$ for the reason that R_1 is not invertible. We now let

$$y_2 = (A_0 - R_1 A_2)\, e, \tag{3.19}$$

and

$$M_2 = \frac{y_2 x}{x y_2}, \tag{3.20}$$

where x is the left and right invariant eigenvector of A and we can verify that y_2 are the is a right invariant eigenvector of R_1, and

$$R_1^k = \begin{cases} \hat{R}_1^k + M_2, & k \geq 1, \\ I, & k = 0. \end{cases} \tag{3.21}$$

where $\hat{R}_1 = R_1 - M_2$. Furthermore, we can write S_1' as follows:

$$S_1' = \sum_{k=0}^{N-1} R_1^k = \sum_{k=0}^{N-1} \hat{R}_1^k + (N-1)\, M_2 = \left(I - \hat{R}_1^N\right)\left(I - \hat{R}_1\right)^{-1} + (N-1)\, M_2. \tag{3.22}$$

On the other hand, because R_1 is invertible, we can know

$$S_2' = \sum_{k=0}^{N-1} R_2^k = \left(I - R_2^N\right)(I - R_2)^{-1}. \tag{3.23}$$

3.4. Some decomposition results for R_1 and R_2

From the analysis in the Section 3.2, we know that the computation of the rate matrices R_1 and R_2 is the key to analyze the stationary probability vectors. However, we can find that computing these two matrices R_1 and R_2 of order $(nm_1 + nrm_2) \times (nm_1 + nrm_2)$ are usually the heaviest computation load associated with our model, Alfa [3] presented some decomposition results for MAP/PH/1 vacation queueing system, which can be used to reduce the computation load. Following the analysis in Alfa [3], we also develop decomposition-based results that can further assist in improving on the methods for computing R_1 and R_2. Since R_1 and R_2 have similar decomposition results, here, we just analyze the decomposition of R_1.

From the structures of $A_i, i = 0, 1, 2$, we immediately see the R_1 satisfies the form of

$$R_1 = \begin{bmatrix} R_1^{vv} & R_1^{vs} \\ 0 & R_1^{ss} \end{bmatrix} \tag{3.24}$$

which leads to the following proposition.

Proposition 3.1.

$$A_0^{vv} + R_1^{vv} A_1^{vv} + \left(R_1^{vv}\right)^2 A_2^{vv} = R_1^{vv}, \tag{3.25}$$

$$A_0^{vs} + R_1^{vv} A_1^{vs} + R_1^{vs} A_1^{ss} + \left(R_1^{vv}\right)^2 A_2^{vs} + \left(R_1^{vv} R_1^{vs} + R_1^{vs} R_1^{ss}\right) A_2^{ss} = R_1^{vs}, \tag{3.26}$$

$$A_0^{ss} + R_1^{ss} A_1^{ss} + \left(R_1^{ss}\right)^2 A_2^{ss} = R_1^{ss}. \tag{3.27}$$

Proof. From (3.24), we know that

$$R_1^2 = \begin{bmatrix} (R_1^{vv})^2 & R_1^{vv}R_1^{vs} + R_1^{vs}R_1^{ss} \\ 0 & (R_1^{ss})^2 \end{bmatrix}, \tag{3.28}$$

we rewrite equation (3.6) in blocks, we can obtain (3.25)–(3.27) directly. □

It is immediately clear that R_1^{ss} is the rate matrix associated with the MAP/PH/1 queue with no vacation, R_1^{vv} corresponds to the rate matrix during the working vacation period, and the remaining third non-zero block matrix of R_1^{vv} is R_1^{vs} which connects matrices R_1^{vv} and R_1^{vv}.

Let $U_0 = R_1^{vv}$, $U_1 = A_1^{ss} + R_1^{ss}A_2^{ss}$, $U_2 = A_0^{vs} + R_1^{vv}A_1^{vs} + (R_1^{vv})^2 A_2^{vs}$, we can write (3.26) as

$$R_1^{vs} = U_2 + R_1^{vs}U_1 + U_0R_1^{vs}A_2^{ss}. \tag{3.29}$$

Let $\tilde{U}_1 = -(I - U_1)(A_2^{ss})^{-1}$ and with the assumption that the inverse of A_2^{ss} exists, we have the following proposition.

Proposition 3.2. *If the inverse of A_2^{ss} exists, then we can get the following equation*

$$R_1^{vs}\tilde{U}_1 + U_0R_1^{vs} = \tilde{U}_2. \tag{3.30}$$

This is a Sylvester function, numerous of algorithms can be used to solve this function, for example Bartels–Stewart's algorithm (see [7]).

We should note that it is necessary that the inverse of A_2^{ss} is required to be existed when we apply Bartels–Stewarts algorithm to carried out some operations. Alfa [3] proposed an alternative approach to solve this problem. For an $n \times m$ matrix B, we let

$$\operatorname{vec} B = [B_{1,1}, \dots, B_{n,1}, B_{1,2}, \dots B_{n,2}, \dots B_{1,m}, B_{n,m}]^T \tag{3.31}$$

be termed the vectorized version of B, where $B_{i,j}$ is the $(i,\ j)$th element of matrix B, and T is reserved for transpose of a matrix. Using vectorization idea and properties of Kronecker products for (3.29), we can get the following proposition directly.

Proposition 3.3.

$$\operatorname{vec} R_1^{vs} = \left[I - (A_1^{ss} + R_1^{ss}A_2^{ss})^T \otimes I - (A_2^{ss})^T \otimes R_1^{vv}\right]^{-1} \operatorname{vec} U_2. \tag{3.32}$$

The proof is direct, so we omit it, during the proof process of this proposition, we use the following property [17]: For three matrices W, X and Y, if $Z = WXY$, then $\operatorname{vec} Z = \left(Y^T \otimes W\right) \operatorname{vec} X$.

Based on the decomposition results, we can give another algorithm to compute R_1.

Algorithm 2. Compute R_1 by decomposition results.

Input: transition blocks A_k^{vv}, A_k^{vs}, A_k^{ss}, $k = 0,\ 1,\ 2$, and the tolerance ε.
Output: R_1.

1: Let $R_1^{vv}[1] = A_0^{vv} + R_1^{vv}[0]\,A_1^{vv} + R_1^{vv}[0]^2\,A_2^{vv}$ and $R_1^{ss}[1] = A_0^{ss} + R_1^{ss}[0]\,A_1^{ss} + R_1^{ss}[0]^2\,A_2^{ss}$.
2: If $||R_1^{vv}[1] - R_1^{vv}[0]|| \geq \varepsilon$ or $||R_1^{ss}[1] - R_1^{ss}[0]|| \geq \varepsilon$, go to Step 4, else $R_1^{vv}[1] \leftarrow R_1^{vv}[0]$, $R_1^{ss}[1] \leftarrow R_1^{ss}[0]$, then return Step 2.
3: Set $R_1^{vv} = R_1^{vv}[1]$ and $R_1^{ss} = R_1^{ss}[0]$.
4: Get R_1^{vs} by solving Sylvester function (3.30) using Barttel–Stewart's algorithm or through (3.32) by using vectorization idea. Then use (3.24) to get R_1.

Clearly, we can obtain the R_2 by the similar algorithm. From the Algorithm 1, we know that the complexity of computing the R_1 and R_2 will involve a complexity of $O\left((nm_1 + nrm_2)^3\right)$, however, using the decomposition results, the heaviest workload of the computation of R_1 and R_2 is in computing R_i^{vv} or R_i^{ss}, and the complexity is $\max\left\{O\left((nm_1)^3\right), O\left((nrm_2)\right)^3\right\}$.

Based on the analysis above, we give a computational procedure that can be used for computing the stationary probability vectors of our model.

Algorithm 3. Computational procedure for the stationary probability vectors.

Input: $B_{00},\ B_{01},\ B_{10},\ A_0,\ A_1,\ A_2,\ C$.
Output: Approximate solution for $\pi_0, \pi_1, \ldots, \pi_N$.

1: Use the Algorithm 1 or Algorithm 2 to compute R_1 and R_2.
2: Compute ρ, if $\rho < 1$, compute S_1' and S_2' by (3.14) and (3.18), respectively, if $\rho > 1$, compute S_1' and S_2' by (3.22) and (3.23), respectively.
3: Obtain π_{00}, v_1 and v_2 by solving $[\pi_0, v_1, v_2]\, B\, [R_1, R_2] = [\pi_0, v_1, v_2]$ and $\pi_0 e + v_1 S_1' e + v_2 S_2' e = 1$.
4: For $1 \le k \le N$, set $\pi_k = v_1 R_1^{k-1} + v_2 R_2^{K-k}$.

4. Performance measures

In this section, some performance measures are obtained as follows.

We first provide the probability that the queue is empty and full.

$$P_{\text{empty}} = \pi_0 e, \tag{4.1}$$
$$P_{\text{full}} = \pi_N e = \left(v_2 + v_1 R_1^{N-1}\right) e. \tag{4.2}$$

Using the analysis in Section 3.3 and identities

$$\sum_{k=0}^{K} kB^k = \left(I - B^{K+1}\right)(I - B)^{-2} - \left(I + KB^{K+1}\right)(I - B)^{-1}, \tag{4.3}$$

and

$$\sum_{k=1}^{K} kB^{k-1} = \left(I - B^K\right)(I - B)^{-2} - KB^K (I - B)^{-1}, \tag{4.4}$$

which hold for an arbitrary matrix B when $(I - B)$ is invertible and, we can obtain the mean queue length $E(L)$ as follows.

(i) If $\rho < 1$, note here $(I - R_1)$ is invertible, but $(I - R_2)$ is not invertible, we can obtain the mean queue length as follows.

$$\begin{aligned}
E(L) &= \sum_{k=0}^{N} k\pi_k e = \sum_{k=1}^{N} k\left(v_1 R_1^{k-1} + v_2 R_2^{N-k}\right) e \\
&= v_1 \sum_{k=1}^{N} kR_1^{k-1} e + v_2 \sum_{k=1}^{N} kR_2^{N-k} e \\
&= v_1 \sum_{k=1}^{N} kR_1^{k-1} e + v_2 \sum_{n=0}^{N-1} (N-n)\left(\hat{R}_2^n + M_1\right) e
\end{aligned}$$

$$= v_1 \sum_{k=1}^{N} k R_1^{k-1} e + v_2 \left(N \sum_{n=0}^{N-1} \hat{R}_2^n + (N-1) M_1 - \sum_{n=0}^{N-1} n \hat{R}_2^n - \frac{N(N-1)}{2} M_1 \right) e$$

$$= v_1 \left(\left(I - R_1^N\right) (I - R_1)^{-2} - N R_1^N (I - R_1)^{-1} \right) e + v_2 \left(N \left(I - \hat{R}_2^N\right) \left(I - \hat{R}_2\right)^{-1} \right.$$

$$\left. + N^2 M_1 - \left(I - \hat{R}_2^N\right) \left(I - \hat{R}_2\right)^{-2} + N \hat{R}_2^N \left(I - \hat{R}_2\right)^{-1} - \frac{N(N-1)}{2} M_1 \right) e. \quad (4.5)$$

(ii) If $\rho > 1$, note $(I - R_1)$ is not invertible, but $(I - R_2)$ is invertible, we can obtain the mean queue length as follows.

$$E(L) = \sum_{k=0}^{N} k \pi_k = \sum_{k=1}^{N} k \left(v_1 R_1^{k-1} + v_2 R_2^{N-k}\right) e$$

$$= v_1 \sum_{k=1}^{N} k R_1^{k-1} e + v_2 \sum_{k=1}^{N} k R_2^{N-k} e$$

$$= v_1 \sum_{k=1}^{N} k \left(\hat{R}_1^{k-1} + M_2\right) e + v_2 \sum_{n=0}^{N-1} (N-n) R_2^n e$$

$$= v_1 \left(\left(I - \hat{R}_1^N\right) \left(I - \hat{R}_1\right)^{-2} - N \hat{R}_1^N \left(I - \hat{R}_1\right)^{-1} + \frac{N(N+1) M_2}{2} \right) e$$

$$+ v_2 \left(N \left(I - R_2^N\right) (I - R_2)^{-1} - \left(I - R_2^N\right) (I - R_2)^{-2} + \left(I + (N-1) R_2^N\right) (I - R_2)^{-1} \right) e. \quad (4.6)$$

Now, we study the loss probability P_{loss} representimg the probability that an arbitrary customer is lost in front of system when he/she arrives. Let partition π_i as follows: $\pi_i = [\pi_{i,0}, \pi_{i,1}], 1 \le i \le N$, where $\pi_{i,0}$ represents the probability vector that there are i customers in the system with the server is in working vacation period and $\pi_{i,1}$ represents the probability vector that there are i customers in the system with the server is in normal busy period.

Proposition 4.1. *If $\rho \neq 1$, the loss probability is given by*

$$P_{\text{loss}} = \lambda^{-1} \pi_{N,0} (D_1 \otimes T \otimes S_2) e + \lambda^{-1} \left[\pi_{N,1} (D_1 \otimes S_1) + \pi_{N,0} \left(D_1 \otimes T^0 \otimes S_2 e \beta\right)\right] e, \quad (4.7)$$

$$\lim_{N \to \infty} P_{\text{loss}} = \max \left\{0, 1 - \frac{1}{\rho}\right\}. \quad (4.8)$$

Proof. Since the system capacity is N, when the customer finds N customers staying in the system when it arrives, it will leave the system at its arrival without the service, thus, the loss probability is $P_{\text{loss}} = \lambda^{-1} \pi_N A_0 e$ from which we can obtain the expression of (4.7) directly. Intuitively, the first item of right-hand size of (4.7) is the probability that the new customer finds that the system in working vacation period and there are N customers stayed in the system when it arrives, the second item is the probability that the new customer finds that the system is in vacation period and there are N customers in the system when it arrives and the last item is the probability that the new customer finds that the system is in regular busy period and there are N customers in the system. For the limit of P_{loss}, note that if $\rho < 1$, $\pi_N \to 0$, when $N \to \infty$, by $P_{\text{loss}} = \lambda^{-1} \pi_N A_0 e$, we can directly derive that $\lim_{N \to \infty} P_{\text{loss}} = 0$. On the other hand, applying Little's law to the server, we can know

$$1 - P_{\text{loss}} = \frac{1 - \pi_0 e}{\rho} \cdot \quad (4.9)$$

The right-hand side of (4.9) represents the ratio of the amount per time unit of processed workload $1 - \pi_0$ and that of offered workload ρ. If $\rho > 1$, we know that $\pi_0 \to 0$, when $N \to \infty$, from (4.9), we can derive that $\lim_{N \to \infty} P_{\text{loss}} = 1 - 1/\rho$. □

5. Waiting time analysis

In this section, we consider the waiting time of a successful customer, *i.e.*, a customer who is not lost due to the full buffer. Define $z_i, (0 \le i \le N-1)$, as the stationary vector corresponding to state of finding i customers in the system by a successful customer when he /she arrives, and the vector z_i, for $1 \le i \le N-1$, can be partitioned into $z_i = [z_{i0}, z_{i1}]$, where z_{i0} represents the probability vector that a successful customer observe i customers in the system and the system is in working vacation period when he/she arrive, z_{i1} represents the probability vector that a successful customer observe i customers in the system and the system is in normal service period when he/she arrive, then

$$z_0 = \frac{\pi_0 D_1 \otimes \left(T + T^0\alpha\right) + \pi_{10}\left(D_1 \otimes \left(T + T^0\alpha\right) \otimes S_2^0\right) + \pi_{11}\left(D_1 \otimes S_1^0\alpha\right)}{\lambda\left(1 - P_{\text{loss}}\right)}, \tag{5.1}$$

for $1 \le i \le N-1$,

$$z_{i0} = \frac{\pi_{i0}\left(D_1 \otimes T \otimes S_2\right) + \pi_{i+1,0}\left(D_1 \otimes T \otimes S_2^0\delta\right)}{\lambda\left(1 - P_{\text{loss}}\right)}, \tag{5.2}$$

$$z_{i1} = \frac{\pi_{i1}\left(D_1 \otimes S_1\right) + \pi_{i+1,1}\left(D_1 \otimes S_1^0\beta\right) + \pi_{i,0}\left(D_1 \otimes T^0 \otimes S_2 e\beta\right) + \pi_{i+1,0}\left(D_1 \otimes T^0 \otimes S_2^0\beta\right)}{\lambda\left(1 - P_{\text{loss}}\right)}. \tag{5.3}$$

We can readily show that

$$z_0 e + \sum_{i=1}^{N-1} z_{i0} e + \sum_{i=1}^{N-1} z_{i1} e = 1. \tag{5.4}$$

In order to analyze the waiting time, we define another transition probability matrix

$$\tilde{P} = \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ N-2 \\ N-1 \end{matrix} \begin{bmatrix} \tilde{B}_{00} & & & & & \\ \tilde{B}_{10} & \tilde{A}_1 & & & & \\ & \tilde{A}_2 & \tilde{A}_1 & \cdots & & \\ & & & \cdots & & \\ & & & \cdots & \tilde{A}_1 & \\ & & & & \tilde{A}_2 & \tilde{A}_1 \end{bmatrix}, \tag{5.5}$$

where

$$\tilde{B}_{00} = \left(T + T^0\alpha\right), \qquad \tilde{B}_{10} = \left[\left(T + T^0\alpha\right) \otimes S_2^0, S_1^0\delta\right]^T,$$

$$\tilde{A}_1 = \begin{bmatrix} T \otimes S_2 & T^0 \otimes S_2 e\beta \\ 0 & S_1 \end{bmatrix}, \qquad \tilde{A}_2 = \begin{bmatrix} T \otimes S_2^0\delta & T^0 \otimes S_2^0\beta \\ 0 & S_1^0\beta \end{bmatrix}.$$

Define $\tilde{z}^0 = \left[\tilde{z}_0^0, \tilde{z}_1^0, \tilde{z}_2^0, \ldots, \tilde{z}_{N-1}^0\right]$, where $\tilde{z}_0^0 = z_0\left(e \otimes I(r)\right)$, $\tilde{z}_i^0 = \left[\tilde{z}_{i0}^0, \tilde{z}_{i1}^0\right]$ for $i = 1, 2, \ldots, N-1$, and $\tilde{z}_{i0}^0 = z_{i0}\left(e \otimes I(rm_2)\right)$, $\tilde{z}_{i1}^0 = z_{i0}\left(e \otimes I(rm_1)\right)$, where $I(r)$ refers to an identity matrix I of dimension r, also let

$$\tilde{z}^{n+1} = \tilde{z}^n \tilde{P}, \ n \ge 0. \tag{5.6}$$

Further partition $\tilde{z}^n$ as $\tilde{z}^0$ is partitioned, so that $\tilde{z}^n = \left[\tilde{z}_{00}^n, \tilde{z}_1^n, \tilde{z}_2^n, \ldots, \tilde{z}_{N-1}^n\right]$, $\tilde{z}_i^n = [\tilde{z}_{i0}^n, \tilde{z}_{i1}^n]$, let W_j be the probability that a successful customers waiting time is less than or equal to j units, then

$$W_j = z_{00}^j e, \quad j \ge 0, \tag{5.7}$$

and z_{00}^j can be calculated recursively as follows

$$\begin{aligned} z_{00}^{n+1} &= z_{00}^n\left(T + T^0\alpha\right) + z_{10}^n\left(T + T^0\alpha\right) \otimes S_2^0 + z_{11}^n S_1^0\delta, & \\ z_{i0}^{n+1} &= z_{i0}^n\left(T \otimes S_2\right) + z_{i0}^n\left(T \otimes S_2^0\delta\right), & 1 \le i \le N-1, \\ z_{i1}^{n+1} &= z_{i0}^n\left(T^0 \otimes S_2 e\beta\right) + z_{i1}^n S_1 + z_{i+1,0}^n\left(T^0 \otimes S_2^0\beta\right) + z_{i+1,1}^n S_1^0\beta, & 1 \le i \le N-1. \end{aligned}$$

The mean waiting time can be obtained by Little' law

$$E\,(W) = E\,(N)\,/\lambda\,(1 - P_{\text{loss}})\,. \tag{5.8}$$

6. The number of customers served in the busy period

In this section, we concentrate on the analysis of the number of customers served in a busy period which is the consecutive duration from the server begins to serve the first arriving customer after the system becomes empty until the system becomes empty again. We define $f_{ij}\,(n) = P\left\{N = n,\ Z_{\gamma_0^{(1,i)}} = (0, j) \middle|\, Z_0 = (1, i)\,, \Delta = k\right\}$, and matrix $\tilde{F}\,(n, k) = \{f_{ij}\,(n)\}\,(1 \le i, j \le nm_1 + nrm_2)$, where Z_m represents the state of DTMC in term of a level number and a block matrix index, the stopping time $\gamma_0^{(1,i)}$ specifies the occurrence of transition that ends the busy period which starts in $Z_0 = (1, i)$, the variable Δ reckons the number of customers which counts from the current number to the greatest capacity of the system. The conditional generating function $F\,(z)$ of the number of customers that served in a busy period can be given by

$$F\,(z, k) = \sum_{n=0}^{\infty} \tilde{F}\,(n, k)\, z^k = \begin{cases} zA_2 + CF\,(z, k)\,, & k = 1, \\ zA_2 + A_1 F\,(z, k) + A_0 F\,(z, k-1)\, F\,(z, k)\,, & k > 1. \end{cases} \tag{6.1}$$

with $F_k^{(n)} = (\mathrm{d}^n/\mathrm{d}z^n)\, F\,(z, k)\,|_{z=1}$, $n \ge 0$, the derivatives are $(l \ge 0)$

$$F_k^{(l)} = \begin{cases} I_{l\in\{0,1\}} A_2 + CF_k^{(l)}, & k = 1, \\ I_{l\in\{0,1\}} A_2 + A_1 F_k^{(l)} + A_0 \sum\limits_{i=0}^{l} \begin{pmatrix} l \\ i \end{pmatrix} F_{k-1}^{(l-i)} F_k^i, & k > 1. \end{cases} \tag{6.2}$$

We focus on the conditional factorial moment vectors $\varphi_i = e^T F_k^{(i)} e\,(i = 1, 2, 3)$ for the subscript $k = N$, we may compute all matrices $F_N^{(0)}$, $F_N^{(1)}$, $F_N^{(2)}$ and $F_N^{(3)}$ explicitly from

$$F_k^{(0)} = \left(I - A_1 - A_0 F_{k-1}^{(0)}\right)^{-1} A_2, \tag{6.3}$$

$$F_k^{(l)} = \left(I - A_1 - A_0 F_{k-1}^{(0)}\right)^{-1} \left(I_{l\in\{0,1\}} A_2 + A_0 \sum_{i=0}^{l-1} \begin{pmatrix} l \\ i \end{pmatrix} F_{k-1}^{(l-i)} F_k^{(i)}\right), \quad l = 1, 2, 3, \tag{6.4}$$

starting from the initial values $F_1^{(0)} = F_1^{(1)} = (I - C)^{-1} A_2$, $F_1^{(2)} = F_1^{(3)} = 0$. Finally, $\phi_i = e^T F_N^{(i)} e\,(i = 1, 2, 3)$, where e^T is transpose of the column vector e.

7. Numerical illustration

Through the aforementioned analysis, some stationary system characteristics are derived, in this section, we explore the relationship between system parameter such as the buffer N and the system performance measures such as the loss probability, mean queue length, and the number of customers served in the busy period. We provide two examples which are under condition $\rho < 1$ and $\rho > 1$, separately.

Example 7.1. The situation under $\rho < 1$.

In this example, we give the following assumptions

– The arrival is a discrete the Markov arrival process which has representation as

$$D_0 = \begin{bmatrix} 0.25 & 0.25 \\ 0.15 & 0.55 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.2 \end{bmatrix}.$$

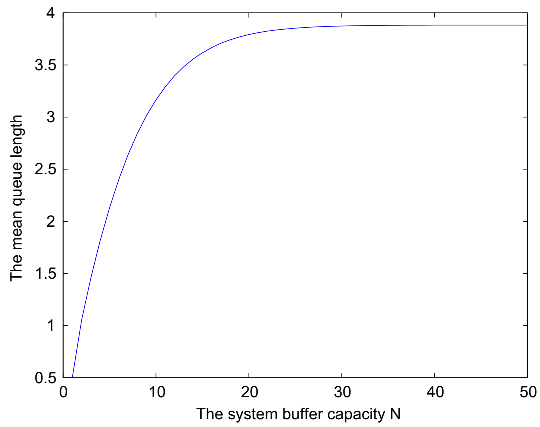

FIGURE 1. The mean queue length for changing N under $\rho < 1$.

– The working vacation time has a phase-type distribution (α, T), where

$$\alpha = [0.1,\ 0.9]\,, \quad T = \begin{bmatrix} 0.2 & 0.6 \\ 0.7 & 0.1 \end{bmatrix}.$$

– The service time during the normal service period has a phase-type (β, S_1), where

$$\beta = [0.3,\ 0.7]\,, \quad S_1 = \begin{bmatrix} 0.1 & 0.3 \\ 0.5 & 0.2 \end{bmatrix}.$$

– The service time during the working vacation period has a phase-type (δ, S_2), where

$$\delta = [0.5,\ 0.5]\,, \quad S_2 = \begin{bmatrix} 0.6 & 0.1 \\ 0.3 & 0.2 \end{bmatrix}.$$

From the assumptions above, we can calculate the arrive rate $\lambda = 0.3714$, the service rate in normal service period $\mu_b = 0.4351$, the service rate in working vacation period $\mu_v = 0.3625$, then traffic intensity $\rho = 0.8536 < 1$. From Figure 1, we can observe that the mean queue length increase and converge to a constant which is obviously the corresponding the mean queue length of MAP/PH/1 queue with working vacation (without finite buffer), and the loss probability decrease to 0. Furthermore, we plot $\log_{10}(P_{\text{loss}})$ in Figure 2 with the changing of N, from which, we can observe that $\log_{10}(P_{\text{loss}}) \sim C_1$ behaves in a straight line with N varying, therefore, P_{loss} is asymptotically exponential under the condition $\rho < 1$, that is $P_{\text{loss}} \sim we^{\sigma N}$, Here σ is a negative constant called the asymptotic decay rate and C is a positive constant called the asymptotic constant. For the tend for the loss probability with N varying, we can observe, from Figure 3, the loss probability approaches to zero when N goes to infinity, the result is consistent with that in Proposition 4.1.

For this queue, we derive the first three conditional factorial moments ϕ_1, ϕ_2 and ϕ_3, which are given in Table 1.

Example 7.2. The situation under $\rho > 1$. In this example, we give the following assumptions:

– The arrival is a discrete the Markov arrival process which has representation as

$$D_0 = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.2 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0.25 & 0.25 \\ 0.15 & 0.55 \end{bmatrix}.$$

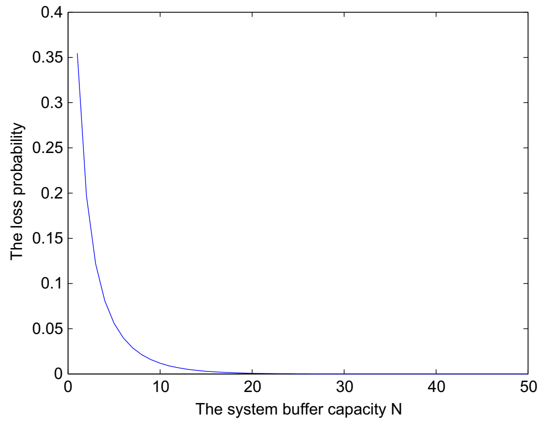

FIGURE 2. The $\log_{10}(P_{\text{loss}})$ for changing N under $\rho < 1$.

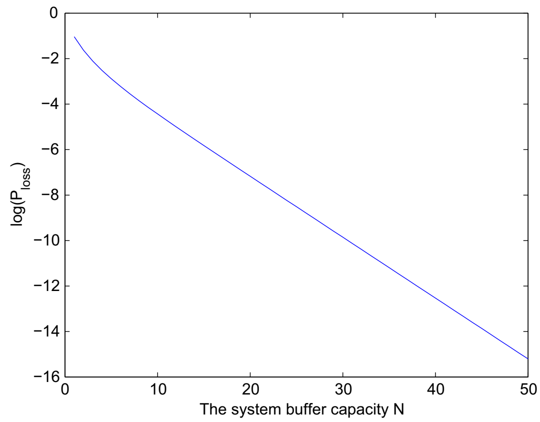

FIGURE 3. The loss probability for changing N under $\rho < 1$.

– The working vacation time has a phase-type distribution (α, T), where

$$\alpha = [0.1,\ 0.9], \quad T = \begin{bmatrix} 0.2 & 0.6 \\ 0.7 & 0.1 \end{bmatrix}.$$

– The service time during the normal service period has a phase-type (β, S_1), where

$$\beta = [0.3,\ 0.7], \quad S_1 = \begin{bmatrix} 0.1 & 0.3 \\ 0.5 & 0.2 \end{bmatrix}.$$

– The service time during the working vacation period has a phase-type (δ, S_2), where

$$\delta = [0.5,\ 0.5], \quad S_2 = \begin{bmatrix} 0.6 & 0.1 \\ 0.3 & 0.2 \end{bmatrix}.$$

For this queue, we can calculate the arrive rate $\lambda = 0.6286$, the service rate in normal service period $\mu_b = 0.4351$, the service rate in working vacation period $\mu_v = 0.3625$, then traffic intensity $\rho = 1.4446 > 1$. We plot mean queue length and loss probability as functions of N in Figures 4 and 5, from which we can observe that the

TABLE 1. The first three conditional factorial moments ϕ_1, ϕ_2 and ϕ_3.

N	2	3	4	5	6	7
ϕ_1	22.3124	30.682	37.2907	42.4318	46.4003	494 510
ϕ_2	37.0770	119.44	239.1546	383.0002	538.0982	694.1112
ϕ_3	8 803 652	777.79	2.7241e+003	6.4000e+003	1.1964e+004	1.9285e+004

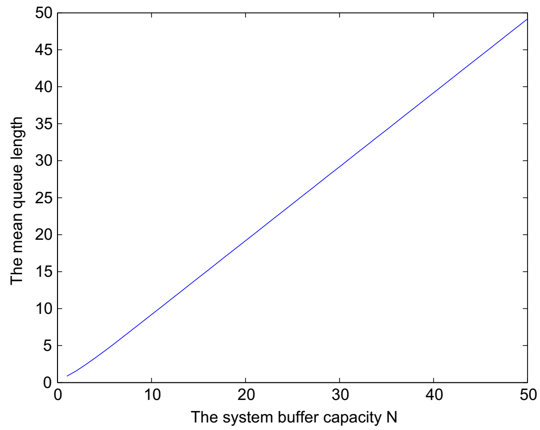

FIGURE 4. The mean queue length for changing N under $\rho > 1$.

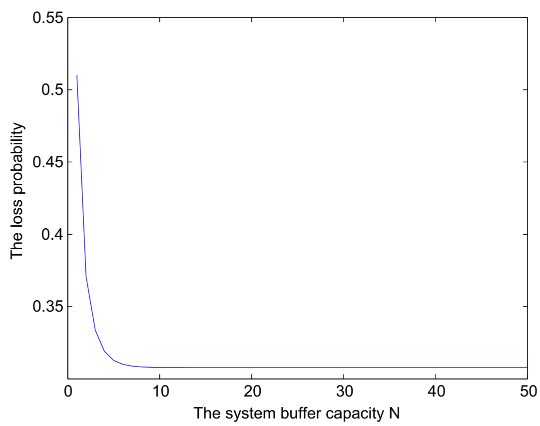

FIGURE 5. The loss probability for changing N under $\rho > 1$.

mean queue length increase to infinity as N goes to infinity, and the loss probability approaches to a constant $0.3078 = 1 - 1/\rho$, which verifies the results in Proposition 4.1.

For this queue, we derive the first three conditional factorial moments ϕ_1, ϕ_2 and ϕ_3, which are given in Table 2.

TABLE 2. The first three conditional factorial moments ϕ_1, ϕ_2 and ϕ_3.

N	2	3	4	5	6
ϕ_1	39.5122	101.3846	239.4525	546.7940	1.2304e+003
ϕ_2	181.6158	1.9017e+003	1.3155e+004	7.6946e+004	4.1498e+005
ϕ_3	1.2524e+003	5.4911e+004	1.1043e+006	1.6418e+007	2.1114e+008

8. Conclusion and future works

In this paper, we study a discrete-time MAP/PH/1 queue with multiple working vacations and finite buffer N. Using the Matrix-Geometric Combination method, we obtain the stationary probability vectors of this model, which can be expressed as a linear combination of two matrix-geometric vectors R_1 and R_2 under the condition that traffic intensity $\rho \neq 1$. Furthermore, we discuss some properties of R_1 and R_2 that are useful to obtain the stationary probability vectors and mean queue length, and give some decomposition results of R_1 and R_2 that can be used to reduce the computation load. We also obtain the loss probability and give its limit of loss probability as finite buffer N goes to infinite. Waiting time distribution is derived by using the absorbing Markov chain. Moreover, we obtain the number of customers served in the busy period.

It is clearly that if $\rho = 1$, the stationary probability vectors of this model can be obtained by Matrix-Geometric Combination method, so it would be interesting to find an effective method to deal with MAP/PH/1 queue with finite buffer under the condition.

Acknowledgements. The authors would like to thank the anonymous reference for their constructive comments that help us to improve the present paper. The research was supported by Natural Science Foundation of China (grant 11901307), and Natural Science Foundation of Jiangsu Province (grants BK20180783, 18KJB110021), and The Startup Foundation for Introducing Talent of Nanjing University of Information Science and Technology (grant 2017r082).

References

[1] A.S. Alfa, A discrete MAP/PH/1 queue with vacations and exhaustive time-limited Service. *Oper. Res. Lett.* **18** (1995) 31–44.
[2] A.S. Alfa, Discrete time analysis of MAP/PH/1 vacation queue with gated time service. *Queueing. Syst.* **29** (1998) 35–54.
[3] A.S. Alfa, Some decomposition results for a class of vacation queue. *Oper. Res. Lett.* **42** (2014) 140–144.
[4] N. Akar, N.C. Oğuz and K. Sohraby, A novel computational method for solving finite QBD processes. *Stoch. Models.* **16** (2000) 273–311.
[5] J.R. Artalejo, A. Gómez-Corral and Q.M. He, Markovian arrivals in stochastic modelling: a survey and some new results. *Sort* **34** (2000) 101–156.
[6] Y. Baba, Analysis of a GI/M/1 queue with multiple working vacations. *Oper. Res. Lett.* **33** (2005) 654–681.
[7] R.H. Bartel and G.W. Stewart, Solution of the equation AX+XB=C. *Commun. ACM.* **15** (1972) 820–826.
[8] H. Bruneel and B.G. Kim, Discrete-Time Models for Communication Systems Including ATM. Kluwer Academic Publishers, Boston (1993).
[9] S.R. Chakravarthy, Markovian Arrival Processes. Wiley Encyclopedia of Operations Research and Management Science (2010).
[10] S. Chakravarthy, Analysis of the MAP/PH/1/K queue with service control. *Appl. Stochastic. Models. Data. Anal.* **12** (1996) 179–191.
[11] S. Chakravarthy and S. Ozkar, MAP/PH/1 queueing model with working vacation and crowdsourcing. *Math. Appl.* **44** (2016) 263–294.
[12] V. Chandrasekaran, K Indhira, M. Saravanarajan and P. Rajadurai, A survey on working vacation queueing models. *Int. J. Pure Appl. Math.* **106** (2016) 33–41.
[13] B.T. Doshi, Queueing systems with vacations—A survey. *Queueing. Syst.* **1** (1986) 29–66.
[14] A.N. Dudin, A.V. Kazimirsky, V.I. Klimenok, L. Breuer and U. Krieger, The queueing model MAP/PH/1//N with feedback operating in a Markovian random environment. *Aust. J. Stat.* **34** (2005) 101–110.
[15] H.R. Gail, S.L. Hantler and B.A. Taylor, Solutions of the basic matrix equation for M/G/1 and G/M/1 type Markov chain. *Stoch. Models.* **10** (1994) 1–43.
[16] S. Gao, J. Wang and W. Li, An M/G/1 retrial queue with general retrial times, working vacations and vacation interruption. *Asia-Pac. J. Oper. Res.* **31** (2014) 6–31.

[17] A. Graham, Kronecker Products and Matrix Calculus: With Applications. John-Wiley, New York (1981).
[18] A. Heindl, M. Telek, Output models of MAP/PH/1(/K) queues for an efficient network decomposition. *Perform. Eval.* **49** (2002) 321–339.
[19] J.D. Kim, D.W. Choi and K.C. Chae, Analysis of queue-length distribution of the M/G/1 queue with Working Vacation. *Hawaii Int. Conf. Stat. Related Fields* **2003** (2003) 1191–1200.
[20] C. Kim, S. Dudin and V. Klimenok, The MAP/PH/1/N queue with flows of customers as a model for traffic control in telecommunication networks. *Perform. Eval.* **66** (2009) 564–579.
[21] G. Latouche and G. Ramaswami, Introduction to matrix analytic methods in stochastic modeling. In: Vol. 5 of *ASA-SIAM Series on Statistics and Applied Probability.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1999).
[22] J. Li and N. Tian, Performance analysis of a G/M/1 Queue with single working vacation. *Appl. Math. Comput.* **217** (2009) 4960–4971.
[23] J. Li, N. Tian and W. Liu, Discrete-time GI/Geom/1 queue with multiple working vacations. *Queueing. Syst.* **56** (2007) 53–63.
[24] J. Li, N. Tian, Z.G. Zhang and H.P. Luh, Analysis of the M/G/1 Queue with exponentially working vacations—a matrix analytic approach. *Queueing. Syst.* **61** (2011) 139–166.
[25] D.M. Lucantoni, New results on the single server queue with a batch Markovian arrival process. *Stoch. Models.* **7** (1991) 1–46.
[26] C. Luo, W. Li, K. Yu and C. Ding, The matrix-form solution for GeoX/G/1/N working vacation queue and its application to state-dependent cost control. *Comput. Oper. Res.* **67** (2016) 63–74.
[27] M.F. Neuts, Probability distributions of phase type. In: Liber Amicorum Prof. Emeritus H. Florin, University of Louvain, Belgium (1975) 173–206.
[28] M.F. Neuts, A versatile Markovian point process. *J. Appl. Probab.* **16** (1979) 764–779.
[29] M.F. Neuts, Matrix-geometric Solution in Stochastic Model. Johns Hopkins University Press, Baltimore, MD (1981).
[30] M.F. Neuts, Structured Stochastic Matrices of M/G/1 Type and Their Applications. Marcel Dekker, NY (1989).
[31] M.F. Neuts, Models based on the Markovian arrival process. *IEICE. T. Commun.* **E75B** (1992) 1255–1265.
[32] L.D. Servi and S.G. Finn, M/M/1 queues with working vacation (M/M/1/WV). *Perform. Eval.* **50** (2002) 41–52.
[33] N. Tian and Z.G. Zhang, Vacation Queueing Models-Theory and Application. Springer-Verlag, New York (2006).
[34] N. Tian, Z. Ma and M. Liu, The discrete time Geom/Geom/1 queue with multiple working vacations. *Appl. Math. Model.* **32** (2007) 2941–2953.
[35] N. Tian, J. Li and Z.G. Zhang, Matrix-analytic method and working vacation queues-survey. *Int. J. Inform. Manage. Sci.* **20** (2009) 603–633.
[36] D. Wu and H. Takagi, M/G/1 queue with multiple working vacations. *Perform. Eval.* **63** (2006) 654–681.
[37] D. Yang and D. Wu, Cost-minimization analysis of a working vacation queue with N-policy and server breakdowns. *Comput. Ind. Eng.* **82** (2015) 151–158.