

ANALYSIS OF A MULTI-ITEM QUEUEING INVENTORY SYSTEM OPERATING WITH AN EXCHANGE FACILITY

KARUMBATHIL RASMI AND MACHUVEETTIL JACOB

ABSTRACT. This paper explores a stochastic queuing-inventory system designed to efficiently manage both new and returned items, offering practical solutions for diverse customer needs. The system differentiates services for various customer classes through dedicated channels, encompassing the sale of new and returned items, as well as the purchase of used items from customers. This model incorporates four parallel queues for distinct customer classes, each serviced by a dedicated server. Customer arrivals are modeled using a Markovian Arrival Process (MMAP), with service times being exponentially distributed and independent. An (s, S) policy is implemented for replenishing fresh items. The primary objective is to enhance service accessibility for various customer types within a single facility, promoting operational efficiency and customer satisfaction. Additionally, the system's ability to purchase used items underscores its role in fostering sustainability in an evolving society. By applying the Neuts matrix geometric technique, the paper analyzes the system to derive the long-term probability distribution and significant performance metrics. Numerical methods are utilized to investigate key system parameters and performance measures, and a cost function is introduced and optimized concerning the reorder level. This comprehensive analysis offers valuable insights for optimizing inventory and queuing systems in practical, real-world applications.

1. INTRODUCTION

Second-hand markets involve the buying and selling of used or pre-owned items. Such markets that combine buying and selling of both new and used items, as well as facilitating item exchanges, offers a wide range of options for consumers. This type of market provides a versatile platform where individuals can not only purchase and sell both new and pre-owned goods but also engage in barter-style exchanges. In such a system, when a customer returns their old items, the system does not demand that these items were originally purchased from the same system. The system operates with the assumption that the returned products always meet the terms and conditions set by the system.

We come across several platforms where objects/commodities/goods could be exchanged with new ones. Customers might be interested in either new items or used items, or they wish to exchange a used item with a new one, or they just want to sell their used item. The customer approaching the system with a used item has two options: either to sell the old item and buy a new item, or the customer can choose to leave the system without buying a new item. This creates two possible classes of customers in such a system with multiple service options. Benefits of buying used items include cost savings, reduced waste, and the potential to find unique or rare items. There will be a class of customers who only wish to purchase a used item. However several customers have a preference for purchasing brand-new, unused products instead of opting for used or secondhand items. They form another class of customers in a

Received by the editors 16 June 2024; accepted 5 August 2025; published online 12 August 2025.

2020 *Mathematics Subject Classification.* 60G99.

Key words and phrases. Heterogeneous Customers, Queueing inventory, Multiple stocks, Exchange facility.

stochastic service system with new and used inventory. Thus, customers may be categorized into four distinct classes based on their characteristics.

A study by Zhao [23] focused on a queueing inventory system with two customer classes, where each service completion reduced the inventory by one unit. The authors modeled the system as a level-dependent quasi-birth-death process and provided a detailed analysis. Krishnamoorthy, Manikandan, and Lakshmy [15] presented a queueing inventory model that incorporates different inventory control policies and addresses several related optimization problems. More recent contributions, including those by Almaqbal et al. [1] and Sinu et al. [20], examine queueing inventory systems with multiple customer classes and employ a Batch Markovian Arrival Process (BMAP) to represent customer arrivals.

The field of queueing-inventory theory, which falls under the broader umbrella of operations research, plays a crucial role in aiding businesses in making decisions on how to design more efficient and cost-effective service systems. The domain of multi-type queueing inventory appears to have attained limited attention, despite the availability of numerous studies on multi-type inventories. As an illustration, [5] conducted an in-depth examination of a queueing-inventory system utilizing a single server. This system analyzed two distinct commodities, with each type having limited storage space allocated to it using the matrix analytic method. In contrast, [22] investigated a queueing inventory system that dealt with three distinct customer types and two types of inventory. [19] conducted an investigation into a two-commodity queueing inventory with two customer classes. A manufacturing system comprised of multiple products, employing a continuous review inventory control policy for each commodity type was explored by [2]. [14] also investigated an inventory model with two types of commodities. Multi-item queueing inventories were also investigated in [4]. [16] studied a Markov inventory system that incorporates flexible service policies and caters to two classes of customers. Furthermore, a production inventory model, involving job shop routings and random arrivals, was the subject of a study by [21].

We assume that the customer's arrival to the system is according to a Marked Markovian Arrival Process (MMAP). Often, the presence of prioritization coincides with the existence of diverse customer groups, such as those differing in their service time distribution, whether in single-server setups [3, 12] or multi-server scenarios [13, 7, 8]. Consequently, the arrival process exhibits variability, and one potential solution for modeling this variability is the Marked Markovian Arrival Process (MMAP), as applied in works like [3, 7, 9]. MMAP[K], an extension of Markov arrival processes (MAP), is a well-established concept in queueing theory. The original MAP, as outlined in Reference [18], was created to model non-Markovian point processes. While it serves as a valuable tool for representing point processes involving a single customer class, MMAP[K], as introduced by Neuts (as referenced in Reference [11]), offers a practical analytical framework for scenarios where multiple customer types coexist (Refer [10]). Key attributes of MMAP, including some characteristics of the arrival process, the time it takes for a specific item type to arrive, and the associated MMAP behavior during that initial arrival, are thoroughly examined in Reference [11].

In their work, Melikov [17] presented queueing inventory models that support both finite and infinite buffer capacities to address customer impatience. Yadavalli et al. [22], in a separate study, examined a system featuring three types of customers, where arrivals are governed by a Markovian Arrival Process (MAP) and two distinct types of inventory items are managed. Chakravarthy [6] explored models in which customers demand a random, but bounded, number of items, with provisions made for potential customer loss.

Although much research has focused on queueing inventory systems with multiple customer classes, no studies have addressed the concept of an inventory exchange facility while considering the presence of four distinct customer classes. This manuscript aims to fill this gap by introducing an inventory exchange mechanism within a multi-server queueing framework, offering a more comprehensive and

realistic model for modern service and production systems. This paper considers a system that treats four classes of customers, creating four different queues each with dedicated servers. On the service completion of a customer from the class 1, one item is subtracted from the stock of new items. The customer from class 2 leaves the system by exchanging of used item with a new one, which adds one item to the stock of used items, and at the same time, one is deducted from the stock of new items. On the other hand, when the service of the class 3 is completed, one item is deducted from the stock of used items. The service completion of the fourth class of customers adds one item to the stock of used inventory. This kind of system is observable all around us.

This paper is organized as follows: An introduction to the model and a review of related work are presented in Section 1. Section 2 explains the mathematical formulation of the model and the process of the system states is explained. Further, it investigates the stability analysis of the model in section 3. Following that, Section 4 derives the steady-state probability vector and the characteristic metrics of the model, and Section 5 gives the numerical illustration. Finally, section 6 concludes the analysis.

2. MODEL DESCRIPTION

In this study, customers from four different classes arrive at the system according to a marked Markovian arrival process(*MMAP*[4]). Two different stocks are incorporated into the system. Stock 1 holds inventory with new items and stock 2 consists of inventory with used items. Customers who demand service with new items belong to class 1. They are allowed to wait in an infinite waiting space. The class 2 customers are those who wish to get service with an exchange of a used item for a new one. Class 3 customers want to buy used items only. Class 4 are interested in selling their used item. The customers from class 2, 3 and 4 can wait in finite buffers with capacities, say, $l_2, l_3,$ and l_4 respectively. There are 4 service nodes, each with a single server, in the system. Server k attends the customers from class $k, k = \overline{1,4}$. The service time of the respective servers follows exponential with parameters $\mu_k, k = \overline{1,4}$. As a class 1 customer departs from the system, one inventory from stock 1 is also removed. On service completion of a class 2 customer, one item is added to the stock 2, and one item is reduced from stock 1. When a Class 3 customer leaves the system after getting service, the stock 2 level is decreased by 1. As a class 4 customer leaves the system after service completion one item is added to the stock 2. Stock 1 is running according to (S_1, s_1) policy with an exponential replenishment rate γ and Stock 2 has a maximum limit, say, S_2 . We assume the customers from class 2 and class 4 arriving in the system when the stock 2 level is S_2 are lost.

The arrival process, *MMAP*[4], is driven by an irreducible continuous time Markov chain (CTMC) $\{Z(t)\}_{t \geq 0}$ with finite state space \mathcal{W} . Let $|\mathcal{W}| = W$. The sojourn time in each state $z \in \mathcal{W}$ is distributed exponentially with rate β_z and d.f.

$$F_{z\hat{z}}(x) = 1 - e^{-\beta_z x}. \tag{2.1}$$

At sojourn time expiration epoch, with probability $P^{(k)}(z, \hat{z})$ the process $Z(t)$ moves from the state z to \hat{z} , generating a class k customer arrival, $z, \hat{z} \in \mathcal{W}, k = 0, 1, 2, 3, 4$. (conventionally, no customer arrival is generated if $k = 0$). Note that, for any $z \in \mathcal{W}$,

$$\sum_{k=0}^4 \sum_{\hat{z} \in \mathcal{W}} P^{(k)}(z, \hat{z}) = 1. \tag{2.2}$$

Thus, the *MMAP* is characterized by a set of 5 *transition rate* square matrices D_0, \dots, D_4 of order W defined as follows:

$$(D_k)_{z\hat{z}} = \begin{cases} \beta_z P^{(k)}(z, \hat{z}), & z \neq \hat{z}, \\ -\beta_z 1_{k=0}, & z = \hat{z}, \end{cases} \quad z, \hat{z} \in \mathcal{W}, \quad k = 0, 1, 2, 3, 4., \tag{2.3}$$

where $1_{k=0}$ is the indicator of a non-random condition $k = 0$. The matrices D_k constitute a *generator* matrix D of the Markov process $\{Z(t)\}_{t \geq 0}$ such that

$$D = \sum_{j=0}^4 D_j. \quad (2.4)$$

Recall that $D\mathbf{e} = \mathbf{0}$.

Let $N_k(t)$ be the number of class k customers in the system, $k = 1, 2, 3, 4$, $I_j(t)$ be the inventory level at stock j , $0 \leq I_j(t) \leq S_j$, $j = 1, 2, \dots$, and $Z(t) \in \mathcal{W}$ be the phase of the MMAP, at time $t \geq 0$. Then, the considered system can be modeled by the regular irreducible CTMC

$$\zeta(t) = \{N_1(t), N_2(t), N_3(t), N_4(t), I_1(t), I_2(t), Z(t)\}, t \geq 0,$$

with state space $(n_1, n_2, n_3, n_4, i_1, i_2, z)$, $n_1 \geq 0, 0 \leq n_k \leq l_k, k = 1, 2, 3, 4, 0 \leq i_j \leq S_j, j = 1, 2$ and $z \in \mathcal{W}$.

Fix the process $\zeta(t)$ at some state $(n_1, n_2, n_3, n_4, i_1, i_2, z)$. Due to independence of the components of $\zeta(t)$, the possible transitions are listed below:

- $(n_1 + 1, n_2, n_3, n_4, i_1, i_2, \hat{z})$ with rate $(D_1)_{z\hat{z}}$, (arrival of class 1 customer);
- $(n_1, n_2 + 1_{n_2 < l_2}, n_3, i_1, i_2, \hat{z})$ or $(n_1, n_2, n_3 + 1_{n_3 < l_3}, i_1, i_2, \hat{z})$ or $(n_1, n_2, n_3, n_4 + 1_{n_4 < l_4}, i_1, i_2, \hat{z})$ with respective rate $(D_k)_{z\hat{z}}$, (arrival of class k customer $k = 2, 3, 4$);
- $(n_1, n_2, n_3, n_4, i_1, i_2, \hat{z})$ with rate $(D_0)_{z\hat{z}}$ (MMAP phase switch);
- $(n_1 - 1, n_2, n_3, n_4, i_1 - 1, i_2, z)$ with rate μ_1 , if $n_1 > 0$ and $i_1 > 0$, (departure of class 1 customer);
- $(n_1, n_2 - 1, n_3, n_4, i_1 - 1, i_2 + 1, z)$ with rate μ_2 , if $n_2 > 0$, $i_1 > 0$, and $i_2 < S_2$ (departure of class 2 customer);
- $(n_1, n_2, n_3 - 1, n_4, i_1, i_2 - 1, z)$ with rate μ_3 , if $n_3 > 0$ and $i_2 > 0$, (departure of class 3 customer);
- $(n_1, n_2, n_3, n_4 - 1, i_1, i_2 + 1, z)$ with rate μ_4 , if $n_4 > 0$ and $i_2 < S_2$, (departure of class 4 customer);
- $(n_1, n_2, n_3, n_4, S_1, i_2, \hat{z})$ with rate γ , if $i_1 \leq s_1$ (inventory replenishment in Stock 1).

Lexicographically ordered state space allows one to write the generator matrix in block-tridiagonal form

$$Q = \begin{bmatrix} A_1^{(0)} & A_0 & \mathbb{O} & \mathbb{O} & \dots \\ A_2 & A_1 & A_0 & \mathbb{O} & \dots \\ \mathbb{O} & A_2 & A_1 & A_0 & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}, \quad (2.5)$$

where the matrix \mathbb{O} is a zero block of the corresponding dimension (we give the dimension explicitly if and when necessary).

Corresponding to each level n_1 , then are L states where,

$$\begin{aligned} L &= (l_2 + 1)(l_3 + 1)(l_4 + 1)(S_1 + 1)(S_2 + 1)W \\ W^0 &= (l_2 + 1)(l_3 + 1)(l_4 + 1)(S_1 + 1)(S_2 + 1) \\ k_2 &= (l_2 + 1), k_3 = (l_3 + 1), k_4 = (l_4 + 1) \\ s_1 &= (S_1 + 1), s_2 = (S_2 + 1) \end{aligned}$$

Then, the blocks $A_i, i = 0, 1, 2$, and $A_1^{(0)}$ are square matrices having size L .

It is quite straightforward to define A_0 , since, upon arrival of a class 1 customer, only the MMAP phase is switched; thus,

$$A_0 = I_{W^0} \otimes D_1, \quad (2.6)$$

where \otimes is the Kronecker product and I is the identity matrix of corresponding dimension. We will need the following notation: hereafter $\mathbf{e}_{i:j}^k$ is the (column) vector of dimension $k \geq 0$ with i th to j th components equal 1, and zero otherwise, $i \leq j \leq k$ (conventionally we take $\mathbf{e}_{i:j}^k \equiv \mathbf{1}$ if $k = 0$). To shorten the notation, we use for any $i \leq k$

- $\mathbf{e}_i^k \equiv \mathbf{e}_{i:i}^k$ (a single non-zero component at i th row, dimension k);
- $\mathbf{e}_k \equiv \mathbf{e}_{k:k}^k$ (last non-zero component, dimension k);
- $\mathbf{e}^k \equiv \mathbf{e}_{1:k}^k$ (all non-zero components of dimension k).

Upon departure of class 1 customer, the *level*, as well as the inventory(stock 1), are decremented by one, the matrix A_2 has the following form:

$$A_2 = I_{k_2 k_3 k_4} \otimes \mu_1 L_1^- \otimes I_{s_2 W}. \quad (2.7)$$

Where the matrix L_j^- is of order $S_j + 1$ and has only the lower diagonals 1 and all other entries are zero, $j = 1, 2$.

We need auxiliary matrices of an increment/decrement in the corresponding (phase) component:

$$N_k^{(+)} = \begin{bmatrix} \mathbf{0} & I_{l_k} \\ 0 & \mathbf{e}'_{l_k} \end{bmatrix}, \quad N_k^{(-)} = \begin{bmatrix} \mathbf{0}' & 0 \\ I_{l_k} & \mathbf{0} \end{bmatrix}.$$

Note that $N_k^{(-)}$ is a square matrix of order $l_k + 1$ having lower diagonal of ones, while $N_k^{(+)}$ is semi-upper diagonal. It is worth noting that asymmetry in $N_k^{(+)}$ (non-zero last row) is used to indicate the arrival of a class k customer that is lost, while the MMAP phase is changed according to D_k .

$$\begin{aligned} A_1 = & (N_2^+ \otimes I_{k_3 k_4 s_1 s_2} \otimes D_2) + (I_{k_2} \otimes N_3^+ \otimes I_{k_4 s_1 s_2} \otimes D_3) \\ & + (I_{k_2 k_3} \otimes N_4^+ \otimes I_{s_1 s_2} \otimes D_4) + (N_2^- \otimes I_{k_3 k_4} \otimes \mu_2 L_1^- \otimes L_2^+ \otimes I_W) \\ & + (I_{k_2} \otimes N_3^- \otimes I_{k_4 s_1} \otimes \mu_3 L_2^- \otimes I_W) + (I_{k_2 k_3} \otimes N_4^- \otimes I_{s_1} \otimes \mu_4 L_2^+ \otimes I_W) \\ & + (I_{k_2 k_3 k_4} \otimes \gamma \mathbf{e}_{1:s_1+1}^{s_1} \mathbf{e}'_{s_1} \otimes I_{s_2 W}) + (I_{W^0} \otimes D_0) - \Delta \end{aligned} \quad (2.8)$$

The matrix L_j^+ is of order $S_j + 1$ and all except upper diagonal of units are zero, $j = 1, 2$ and Δ assures $(A_0 + A_1 + A_2)\mathbf{e} = 0$.

$$\Delta = \text{diag} [\delta^{(a)} + \delta^{(b)} + \delta^{(c)}], \quad (2.9)$$

where $\delta^{(a)}$ is the vector of transition rates due to the departure of customers of classes $2, \dots, K$, $\delta^{(c)}$ contains transition rates due to a class 1 customer departure, and $\delta^{(b)}$ is the vector of transition rates due to replenishment given as follows:

$$\begin{aligned} \delta^{(a)} = & \left(\mu_2 \mathbf{e}_{2:k_2}^{k_2} \otimes \mathbf{e}^{k_3 k_4} \otimes \mathbf{e}_{2:s_1}^{s_1} \otimes \mathbf{e}_{1:s_2-1}^{s_2} \otimes \mathbf{e}^W \right) + \left(\mu_3 \mathbf{e}^{k_2} \otimes \mathbf{e}_{2:k_3}^{k_3} \otimes \mathbf{e}^{k_4 s_1} \otimes \mathbf{e}_{2:s_2}^{s_2} \otimes \mathbf{e}^W \right) \\ & + \left(\mu_4 \mathbf{e}^{k_2 k_3} \otimes \mathbf{e}_{2:k_4}^{k_4} \otimes \mathbf{e}^{s_1} \otimes \mathbf{e}_{1:s_2-1}^{s_2} \otimes \mathbf{e}^W \right), \end{aligned} \quad (2.10)$$

$$\delta^{(b)} = \gamma \mathbf{e}^{k_2 k_3 k_4} \otimes \mathbf{e}_{1:s_1+1}^{s_1} \otimes \mathbf{e}^{s_2 W}, \quad (2.11)$$

$$\delta^{(c)} = \mu_1 \mathbf{e}^{k_2 k_3 k_4} \otimes \mathbf{e}_{2:s_1}^{s_1} \otimes \mathbf{e}^{s_2 W}. \quad (2.12)$$

It remains to define the block $A_1^{(0)}$ and the only difference is in the diagonal balancing matrix. Indeed, from the condition $(A_1^{(0)} + A_0)\mathbf{e} = \mathbf{0}$, it follows that $\Delta_0 = \text{diag} [\delta^{(a)} + \delta^{(b)}]$.

3. STABILITY CONDITION

The necessary and sufficient condition for the existence of the non-zero steady-state probability is the specific version of the Foster ergodicity condition known as the Neuts ergodicity criterion

$$\pi A_0 \mathbf{e} < \pi A_2 \mathbf{e}, \quad (3.1)$$

where the stochastic vector π is the solution of the system

$$\pi A = \mathbf{0}, \quad (3.2)$$

and

$$A = A_0 + A_1 + A_2. \quad (3.3)$$

Note that, due to the properties of A_i , $i = 0, 1, 2$, the matrix A is a generator matrix of a finite state space CTMC giving the projection of the phase transition at high levels; thus, the vector π may be considered as the steady-state probability of the phase at high levels. It now follows that A has a block-tridiagonal structure indexed by the number of class-2 customers as follows:

$$A = \begin{bmatrix} U_1^{(0)} & U_0 & 0 & 0 & \dots & 0 \\ U_2 & U_1 & U_0 & 0 & \dots & 0 \\ 0 & U_2 & U_1 & U_0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & U_2 & U_1 & U_0 \\ 0 & 0 & \dots & 0 & U_2 & \widehat{U}_1 \end{bmatrix}. \quad (3.4)$$

The blocks $U_1^{(0)}$, \widehat{U}_1 , and U_i , $i = 0, 1, 2$, are square matrices. Indeed,

$$\begin{aligned} U_0 &= I_{k_3 k_4 s_1 s_2} \otimes D_2, \\ U_2 &= I_{k_3 k_4} \otimes \mu_2 L_1^- \otimes L_2^+ \otimes I_W \\ U_1 &= I_{k_3 k_4} \otimes ((\gamma \mathbf{e}_{1:s_1+1}^{s_1} \mathbf{e}'_{s_1} + \mu_1 L_1^-) \oplus (I_{s_2} \otimes (D_0 + D_1))) \\ &\quad + (N_3^+ \otimes I_{k_4 s_1 s_2} \otimes D_3) + (I_{k_3} \otimes N_4^+ \otimes I_{s_1 s_2} \otimes D_4) + \\ &\quad + (N_3^- \otimes I_{k_4 s_1} \otimes \mu_3 L_2^- \otimes I_W) + (I_{k_3} \otimes N_4^- \otimes I_{s_1} \otimes \mu_4 L_2^+ \otimes I_W) - \widehat{\Delta}, \end{aligned}$$

where \oplus is the Kronecker sum defined as $A \oplus B = I \otimes B + A \otimes I$; $\widehat{\Delta}$ being a diagonal matrix that guarantees $(U_0 + U_1 + U_2)\mathbf{e} = \mathbf{0}$, and obvious convention $\sum_{j=3}^K = 0$ for $K < 3$ is used. It remains to note that, since $U_1^{(0)}$ corresponds to states with no class-2 customers, $U_1^{(0)} = U_1 + \text{diag}[U_2 \mathbf{e}]$, while the matrix $\widehat{U}_1 = U_1 + U_0$.

Let the row vector π be presented in the form $\pi = (\pi_0, \pi_1, \dots, \pi_{l_2})$. Finally, after some algebra, the ergodicity condition becomes

$$\rho := \lambda_1 \left[\sum_{m=0}^{l_2} \pi_m \mathbf{e}^{k_3 k_4} \otimes \mu_1 \mathbf{e}_{2:s_1}^{s_1} \otimes \mathbf{e}^{s_2 W} \right]^{-1} < 1. \quad (3.5)$$

Note that the stability condition has a nice interpretation, since λ_1 is the upward drift, while the sum in brackets is the mean downward drift of class 1 customers obtained by aggregation of the components phase probabilities vector π corresponding to the states with positive departure probability of class 1 customers.

4. STEADY STATE PERFORMANCE

If the stability condition (3.5) holds, the stochastic vector of stationary probabilities, q , exists and is the solution of the steady-state equation involving the infinitesimal generator Q given in (2.5):

$$qQ = \mathbf{0}. \tag{4.1}$$

Note that the components of the vector q are ordered lexicographically, that is, for $n_k \in \{0, \dots, l_k\}$, $k = 1, \dots, K$, $i \in \{0, \dots, S\}$ and $z \in \mathcal{W}$,

$$q(n_1, n_2, n_3, n_4, i_1, i_2, z) = \lim_{t \rightarrow \infty} P \{N_1(t) = n_1, N_2(t) = n_2, N_3(t) = n_3, N_4(t) = n_4, I_1(t) = i_1, I_2(t) = i_2, Z(t) = z\}.$$

Since the matrix Q defines a level-independent QBD, the solution can be obtained in the matrix-geometric form. Indeed, splitting the vector q into finite vectors q_0, q_1, \dots by the (value of the) first coordinate, we assume that

$$q_i = q_0 R^i. \tag{4.2}$$

The matrix R is the so-called rate matrix, being the minimal non-negative solution of the matrix quadratic equation (which follows from (4.1) using the block-tridiagonal structure of Q and (4.2)),

$$R^2 A_2 + R A_1 + A_0 = \mathbf{0}. \tag{4.3}$$

The boundary conditions for obtaining q_0 follow from the first block-column of Q and are

$$q_0(A_1^{(0)} + R A_2) = \mathbf{0}, \tag{4.4}$$

$$q_0(I - R)^{-1} \mathbf{e} = 1. \tag{4.5}$$

However, in order to avoid numerical difficulties, it is recommended to use the alternative matrix quadratic equation for the sub stochastic matrix G being the minimal non-negative solution of the system

$$A_0 G^2 + A_1 G + A_2 = \mathbf{0}, \tag{4.6}$$

obtain the R matrix by the known relation

$$R = A_0(-A_1 - A_0 G)^{-1}, \tag{4.7}$$

and, finally, use (4.2) to obtain the vector q .

After obtaining the steady-state probability vector, it is straightforward to define the steady-state performance measures of interest. We use an auxiliary vector $j^{(n)}$ of size $n + 1$ containing the sequence $(0, \dots, n)$, i.e.,

$$j^{(n)} = (0, \dots, n).$$

- Average number of class 1 customers in the system:

$$\mathbf{E}_1 = q_0 R(I - R)^{-1} \mathbf{e}^L. \tag{4.8}$$

- Average number of class $k = 2, 3, 4$ customers in the system:

$$\mathbf{E}_k = q_0(I - R)^{-1} \mathbf{e}^{k_2 k_3 k_4} \otimes j^{(l_k)} \otimes \mathbf{e}^{k_{k+1} k_4} \otimes \mathbf{e}^{s_1 s_2 W} \tag{4.9}$$

- Average inventory size in the stock 1:

$$\bar{\mathbf{I}}_1 = q_0(I - R)^{-1} \mathbf{e}^{k_2 k_3 k_4} \otimes j^{(s_1)} \otimes \mathbf{e}^{s_2 W} \tag{4.10}$$

- Average inventory size in the stock 2:

$$\bar{\mathbf{I}}_2 = q_0(I - R)^{-1} \mathbf{e}^{k_2 k_3 k_4 s_1} \otimes j^{(s_2)} \otimes \mathbf{e}^W \tag{4.11}$$

- Replenishment rate,

$$\bar{\mathbf{R}} = q_0(I - R)^{-1}\delta^{(b)}. \tag{4.12}$$

- Class k loss rate, $k = 2, 3, 4$ due to filled buffer

$$\mathbf{L}_k = \sum_{i=0}^{\infty} q_i (\mathbf{e}^{k_2 k_{(k-1)}} \otimes \mathbf{e}_{l_{k+1}} \otimes \mathbf{e}^{k_{(k+1)} k_4 s_1 s_2} \otimes D_k \mathbf{e}_W) \tag{4.13}$$

- Class k loss rate, $k = 2, 4$ due to maximum inventory level

$$\mathbf{L}'_k = \sum_{i=0}^{\infty} q_i (\mathbf{e}^{k_2 k_3 k_4 s_1} \otimes \mathbf{e}_{S_2} \otimes D_k \mathbf{e}_W) \tag{4.14}$$

5. NUMERICAL ILLUSTRATIONS

In this section, a numerical study has been presented to illustrate the behavior of key performance measures under various system parameters. The results help to demonstrate the practical implications of the analytical findings and offer insights into the system’s operational dynamics. The numerical computations were carried out using MATLAB software.

We consider a fixed set of parameter values: $l_2 = 2; l_3 = 2; l_4 = 2; S_1 = 6; S_2 = 4; W = 2; K = 4; \mu_1 = 2; \mu_2 = 2; \mu_3 = 2; \mu_4 = 2$; while varying the replenishment rate γ within the range $[1, 5]$. This allows us to investigate the impact of γ on several key performance measures namely the system load ρ , and the expectations $\mathbf{E}_1, \mathbf{E}_2$, and \mathbf{L}_2 .

Figure 1 illustrates the variation of the system load ρ , as defined in equation (3.5), plotted against the replenishment rate γ for different values of $s_1 = 3, 4, 5$. The curves clearly demonstrate a non-linear inverse relationship between γ and ρ . As the replenishment rate increases, the load on the system decreases, suggesting improved inventory availability and reduced congestion. This behavior is intuitive—higher replenishment rates mitigate the risk of inventory depletion, thereby lowering customer waiting times.

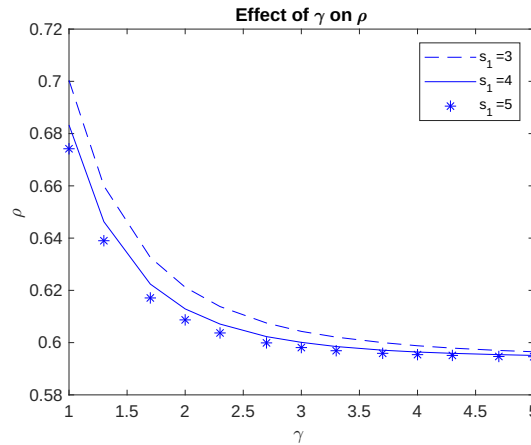


FIGURE 1. Dependency of load of the system ρ on Replenishment rate γ

We now analyze the behavior of the performance measures $\mathbf{E}_1, \mathbf{E}_2$, and \mathbf{L}_2 , focusing on their variation with respect to changes in the replenishment rate γ . By varying γ in the interval $[1, 5]$, we obtain a set of non-linear curves for different values of $s_1 = 3, 4, 5$, as illustrated in the following figures.

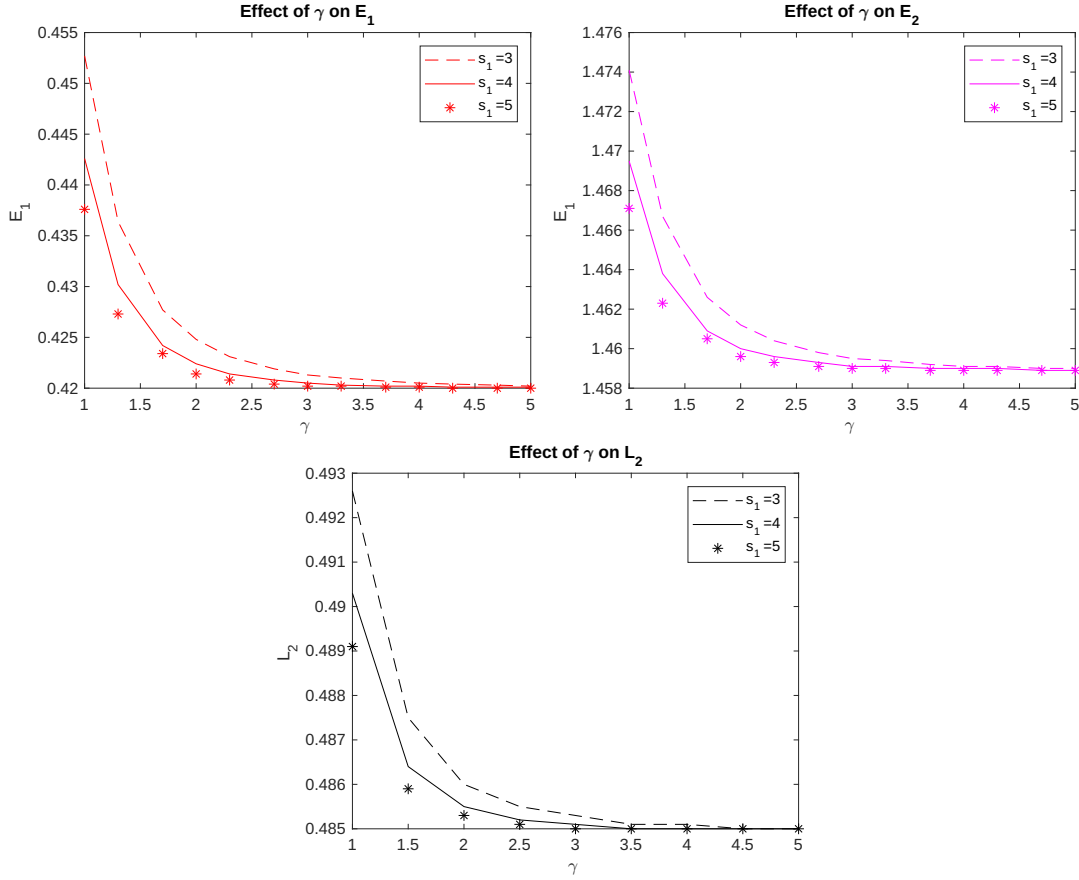


FIGURE 2. Dependency of E_1, E_2, L_2 on Replenishment rate γ

The results reveal that as the replenishment rate of stock 1 increases, the expected number of class 1 customers both those purchasing new inventory (E_1) and those returning for exchange (E_2), decreases significantly. This is attributed to improved inventory availability, which reduces delays and system congestion. Additionally, the loss rate of class 2 customers (L_2), those turned away due to a full buffer, also diminishes with higher replenishment rates.

These trends highlight the crucial role of the replenishment rate in enhancing customer service levels and minimizing the likelihood of service denial. Furthermore, the decrease in all three measures with increasing γ underscores the effectiveness of replenishment policies in managing customer flow and resource utilization.

Cost Function. One primary objective of this study is to balance the costs associated with maintaining inventory and the costs related to waiting for items to arrive. A comprehensive cost function considering various components can be introduced as,

$$E_{TC} = c_{i_1} \bar{I}_1 + c_{i_2} \bar{I}_2 + \sum_{k=1}^4 c_k E_k + c_R \hat{R} + \sum_{k=2}^4 c_{L_k} L_k + c_{L'_2} L'_2 + c_{L'_4} L'_4 \quad (5.1)$$

- c_{i_j} is the holding cost of inventory in stock j (per item per unit time), $j = 1, 2$.
- c_k is the waiting cost of customers at node k (per customer per unit time) $k = 1, 2, 3, 4$.
- c_R is the reorder cost (per system per unit time).

- c_{L_k} is the cost due to a class k customer loss due to a fully occupied buffer (per system per unit time), $k = 2, 3, 4$.
- $c_{L'_k}$ is the cost due to a class k customer loss due to maximum inventory level S_2 in stock 2 (per system per unit time), $k = 2, 4$.

The service rate plays a critical role in managing the trade-off between holding costs and waiting costs in a queueing inventory model. We may fix all the unit costs at 1. The following graphs depict the dependence of cost function on the service rates of different classes of customers. We fix the parameters as $l_2 = 3; l_3 = 3; l_4 = 3; s_1 = 2; S_1 = 6; S_2 = 4; W = 2; K = 4; \mu_3 = 2; \mu_4 = 2; \gamma = 2$; and vary μ_1 fixing $\mu_2 = 2$ and vice versa to obtain the following two graphs.

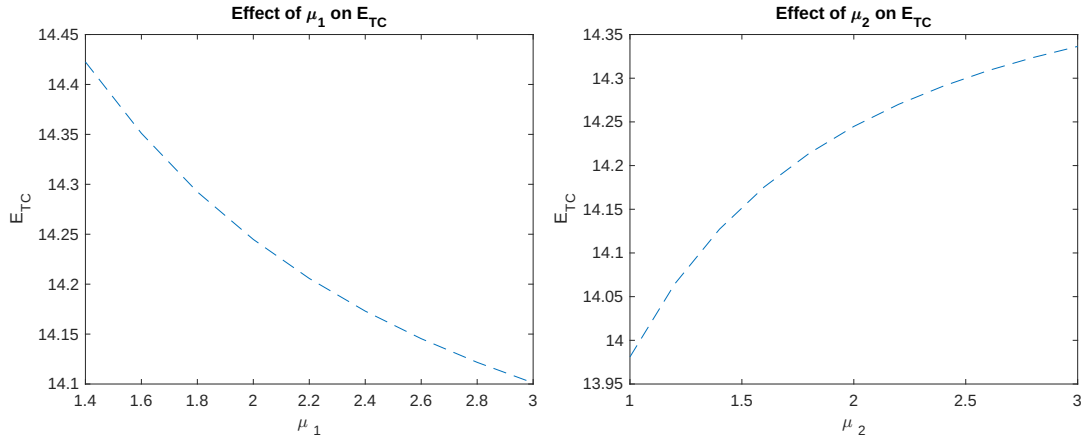


FIGURE 3. Service rates vs. Expected total cost

We observe that as the service rates of class 1 customers increase, the total cost decreases. This reduction is primarily due to the shorter waiting times and reduced number of class 1 customers in the buffer, which leads to lower holding and delay-related costs. On the other hand, an increase in the service rates of class 2 customers results in a higher total cost. This may be attributed to increased operational or service-related costs associated with handling class 2 customers more rapidly, possibly without proportional benefits in terms of system efficiency or customer retention.

The replenishment rate γ significantly influences the holding costs, ordering costs, and the risk of stockouts, which in turn impact customer waiting times. Figure 4 illustrates the effect of varying γ on the expected total cost.

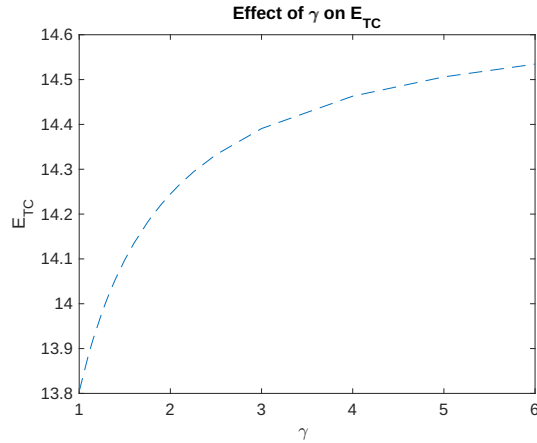


FIGURE 4. Replenishment rate vs. Expected total cost

As the figure indicates, an increase in the replenishment rate may lead to a rise in the expected total cost. While a higher replenishment rate reduces the average number of waiting customers (both class 1 and class 2) and lowers the customer loss rate due to improved inventory availability, these benefits are offset by the increased frequency of reorders. The elevated ordering activity results in higher replenishment-related costs, which ultimately outweigh the gains from reduced congestion and customer loss, leading to an overall increase in total cost.

The cost function can be optimized through strategic decisions regarding the reorder point s_1 . By optimizing this function, an effective inventory management policy can be established that minimizes the overall cost while ensuring timely service to customer demand.

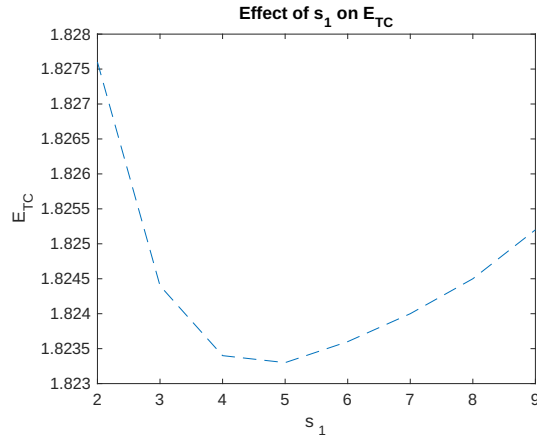


FIGURE 5. Reorder level vs. Expected total cost

Fixing the parameters as $l_2 = 2; l_3 = 2; l_4 = 2; S_1 = 10; S_2 = 4; W = 2; K = 4; \mu_1 = 2; \mu_2 = 2; \mu_3 = 2; \mu_4 = 2; \gamma = 2$; and varying s_1 from 1 – 9, the expected total cost is calculated. The results are plotted in Figure 5. As evident from the figure, the total cost initially decreases with increasing s_1 , reaches a minimum, and then begins to rise again. This indicates a trade-off between inventory holding costs and service efficiency. The minimum cost is attained when the reorder level is set at $s_1 = 5$,

suggesting that this value offers the most balanced and cost-effective inventory policy under the given system parameters.

6. CONCLUSION

In this work, a stochastic queueing inventory system has been analyzed, accommodating different types of inventory and catering to multiple classes of customers with distinct service requirements. The system handles both fresh and used items, where new inventory is replenished following an (s, S) -type policy. The model is formulated as a level-independent QBD process, and the matrix-geometric method is applied for steady-state analysis. The steady-state probability vector is obtained, which forms a basis for computing several important performance measures, such as the expected number of customers in various classes, customer loss rates, and overall system load.

A detailed numerical study is carried out to explore the behavior of the system under different parameter settings. The effect of the replenishment rate γ on key performance indicators such as customer waiting time, inventory levels, and loss rates is carefully examined through graphical analysis. It is observed that while higher replenishment rates improve service availability and reduce customer wait times, they may also lead to increased ordering costs, underscoring the need for balance.

Furthermore, a total cost function incorporating holding, ordering, and loss-related costs is introduced. Optimization of this cost function, with respect to the reorder point s_1 , reveals the optimal inventory control policy that minimizes the expected total cost. The results demonstrate that careful tuning of replenishment and service parameters can significantly enhance system performance and operational efficiency.

Overall, the study provides a comprehensive framework for analyzing and optimizing complex queueing inventory systems that serve multiple customer classes with distinct demands, including exchange facilities and service processes.

REFERENCES

- [1] K. A. K. ALMaqbali and V. C. Joshua and A. Krishnamoorthy, *Multi-Class, Multi-Server Queueing Inventory System with Batch Service*, Mathematics **11(4)**(2023): 830.
- [2] T. Ahtiok and G. A. Shiue, *Pull-type manufacturing systems with multiple product types*, IIE Transactions **32(2)**(2000), 115–124.
- [3] K. Avrachenkov, A. Dudin and V. Klimenok, *Retrial queueing model MMAP/M 2/1 with two orbits*, In International Workshop on Multiple Access Communications, 2010, pp. 107–118.
- [4] J. L. Balintfy, *On a basic class of multi-item inventory problems*, Management Science **10(2)**(1964), 287–297.
- [5] B. Benny, S. R. Chakravarthy, A. Krishnamoorthy, *Queueing-inventory system with two commodities*, Journal of the Indian Society for Probability and Statistics **19**(2018), 437–454.
- [6] S. Chakravarthy, *Queueing-inventory models with batch demands and positive service times*, Matematicheskaya Teoriya Igr i Ee Prilozheniya **11(3)**(2019), 95–120.
- [7] S. Dudin, C. Kim and O. Dudina, *MMAP/ M/ N queueing system with impatient heterogeneous customers as a model of a contact center*, Computers & Operations Research **40(7)**(2013), 1790–1803.
- [8] Q.-M. He, *Quasi-birth-and-death Markov processes with a tree structure and the MMAP [K]/PH [K]/N/LCFS non-preemptive queue*, European Journal of Operational Research **120(3)**(2000), 641–656.
- [9] Q.-M. He, *Queues with marked customers*, Advances in Applied Probability **28(2)**(1996), 567–587.
- [10] Q.-M. He, *The Versatility of MMAP[K] and the MMAP[K]/G[K]/1 Queue*, Queueing Systems **38(4)**(2001): 397–418.
- [11] Q.-M. He and M. F. Neuts, *Markov chains with marked transitions*, Stochastic Processes and Their Applications **74(1)**(1998), 37–52.
- [12] Q.-M. He and D. A. Stanford, *Distributions of the interdeparture times in FCFS and nonpreemptive priority MMAP [2]/G [2]/1 queues*, Performance Evaluation **38(2)**(1999), 85–103.
- [13] V. Klimenok, A. Dudin and V. Vishnevsky, *Priority Multi-Server Queueing System with Heterogeneous Customers*, Mathematics **8(9)**(2020), pp. 1501.

- [14] A. Krishnamoorthy, R. Iqbal Basha and B. Lakshmy, *Analysis of a two commodity inventory problem*, International Journal of Information and Management Sciences **5(1)**(1994), 127–1326.
- [15] A. Krishnamoorthy, R. Manikandan and B. Lakshmy, *A revisit to queueing-inventory system with positive service time*, Annals of Operations Research **233**(2015), 221–236.
- [16] M. Liu, M. Feng and C. Y. Wong, *Flexible service policies for a Markov inventory system with two demand classes*, International Journal of Production Economics **151**(2014), 180–185.
- [17] A. Z. Melikov, L. A. Ponomarenko and S. A. Bagirova, *Analysis of queueing-inventory systems with impatient consume customers*, J. Autom. Inf. Sci. **48(1)**(2016), 53–68.
- [18] M. F. Neuts, *A versatile Markovian point process*, Journal of Applied Probability **16(4)**(1979), 764–779.
- [19] S. Ozkar and U. Uzunoglu Kocer, *Two-commodity queueing-inventory system with two classes of customers*, Opsearch **58**(2021), 234–256.
- [20] L. Sinu, Thulaseedharan Salini and Joshua, Varghese Chaukayil and Vishnevsky, Vladimir and Kozyrev, Dmitry and Krishnamoorthy, Achyutha, *A Multi-Type Queueing Inventory System—A Model for Selection and Allocation of Spectra*, Mathematics **10(5)**(2022): 714.
- [21] N. Van, L. M. Pieter, J. W. M. Bertrand, H. P. G. Van Ooijen and N. J. Vandaele, *A heuristic to control integrated multi-product multi-machine production-inventory systems with job shop routings and stochastic arrival, set-up and processing times*, Stochastic Modeling of Manufacturing Systems: Advances in Design, Performance Evaluation, and Control Issues, 2006, pp. 253–288.
- [22] V. S. S. Yadavalli, O. Adetunji, B. Sivakumar and G. Arivarignan, G, *Two-commodity perishable inventory system with bulk demand for one commodity articles*, South African Journal of Industrial Engineering **21(1)**(2010), 137–155.
- [23] N. Zhao and Z. Lian., *A queueing-inventory system with two classes of customers*, International Journal of Production Economics **129(1)**(2011), 225–231.

K. RASMI, CORRESPONDING AUTHOR, DEPARTMENT OF MATHEMATICS, NATIONAL INSTITUTE OF TECHNOLOGY CALICUT, KOZHIKODE, 673601, KERALA, INDIA
Email address: rasmikcnair@gmail.com

M. J. JACOB, DEPARTMENT OF MATHEMATICS, NATIONAL INSTITUTE OF TECHNOLOGY CALICUT, KOZHIKODE, 673601, KERALA, INDIA
Email address: jacobnitc@gmail.com