On a doubly dynamically controlled supermarket model with impatient customers

Quan-Lin Li *, Ye Du, Guirong Dai, Meng Wang

School of Economics and Management Sciences, Yanshan University, Qinhuangdao 066004, PR China

A R T I C L E   I N F O
Available online 16 October 2014

Keywords:
Supermarket model
Impatient customer
Operator semigroup
Mean-field limit
Fixed point
Performance analysis

A B S T R A C T

In this paper, we provide a key generalization of the supermarket model both from the impatient customers and from a doubly dynamic control, which may also be related to the size-based scheduling through the centered management of the customer resource as well as the total service ability. We first use an infinite-dimensional Markov process to express the states of this supermarket model, and set up an infinite-dimensional system of differential equations satisfied by the expected fraction vector. Then we use the operator semigroup to provide a mean-field limit for the sequence of infinite-dimensional Markov processes, which asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of limiting differential equations. Finally, we provide an effective and efficient algorithm for computing the fixed point of the infinite-dimensional system of limiting differential equations, and use the fixed point to give performance analysis of this supermarket model. Also, some numerical examples are given to demonstrate how the performance measures depend on some crucial parameters of this supermarket model. We believe that the mean-field method developed in this paper will be useful and effective for analyzing more complicated supermarket models in many practical areas.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Dynamic randomized load balancing is often referred to as the supermarket model. The supermarket models have a wide range of practically applicable areas such as call centers, health care, computer networks, production/inventory systems, and transportation networks. Recently, the supermarket models have been analyzed by means of queuing methods as well as Markov processes. For a simple supermarket model, Vvedenskaya et al. [50] applied the operator semigroup and mean-field limit to compute the stationary queue length distribution of any queuing process, and obtained its doubly exponential decay tail which is a substantial improvement of system performance over that in the ordinary M/M/1 queue. At nearly the same time, Mitzenmacher [37] also analyzed the same supermarket model in terms of the density-dependent jump Markov processes. Turner [48] provided a martingale approach to further discuss the supermarket model. The path space evolution of the supermarket model was studied by Graham [23,24] who showed that starting from independent initial states, as $N \to \infty$ the queues of the limiting process evolve independently. Luczak and McDiarmid [33] showed that the length of the longest queue scales as ($\log \log N$)/$\log d$ plus $O(1)$. Certain generalization of the supermarket model has been explored in studying various variations, for example, modeling more crucial factors by Vvedenskaya and Suhov [51], Mitzenmacher [38], Mitzenmacher et al. [39], Bramson et al. [10–12], Li and Lui [30,31], Li et al. [32,29] and Li [27,28]; fast Jackson networks by Martin and Suhov [35], Martin [34] and Suhov and Vvedenskaya [46].

Queues with impatient customers represent a wide range of service systems in which customers may become impatient when they do not receive service fast enough, and they are always useful in modeling many practical situations such as banks, hospitals, supermarkets, supply chains and transportation systems. Readers may refer to some important publications for more details, among which are the M/M/1 queue with impatient customers by Choi et al. [16], Brandt and Brandt [14] and Altman and Yechiali [1]; the M/M/c queue with impatient customers by Stanford [44], Bhatachtarya and Ephremides [5], Boxma and de Waal [7], Movaghar [40], Brandt and Brandt [13], Boots and Tijms [6], Yechiali [52] and Perel and Yechiali [43]; the GI/M/s queue with impatient customers by Teghem [47], Swensen [45] and Perry et al. [42]; the M/G/1 queue with impatient customers by de Kok and Tijms [21], Bae et al. [4], Perry and Asmussen [41], Martin and Artalejo [36], Boxma et al. [8], Iravani and Balcociu [25] and Boxma et al. [9];
and the GI/G/1 queue with impatient customers by Daley [19], Baccelli and Hébuterne [3] and Baccelli et al. [2]. The matrix-analytic method is applied to study queues with impatient customers, e.g., see Combé [18], Choi et al. [17], Van Velthoven et al. [49] and Chakravarthy [15]. In applications of queues with impatient customers, examples include Kaspi and Perry [26] for inventory systems, Zohar et al. [54] for call centers, and Zeltyn and Mandelbaum [53] for telecommunications networks.

The purpose of this paper is to provide a key generalization of the supermarket model both from the impatient customers and from a doubly dynamic control. The generalized supermarket model is described as a queueing network which consists of one server and \( N \) waiting lines with impatient customers under two centered management modes: Each arriving customer joins the shortest one among the \( d_1 \) randomly selected waiting lines (JSQ \((d_1)\)), and the server provides its service in the longest one among the \( d_2 \) randomly selected waiting lines (SLQ \((d_2)\)). Firstly, such two centered management modes may be related to the size-based scheduling both for a suitable allocation of customer resource and for a total use of multiple service units. Then our numerical examples indicate that for improving system performance, the SLQ\((d_2)\) controlling the service processes is more effective than the JSQ\((d_1)\) controlling the arrival processes. Finally, it is necessary to analyze the useful relations among some special cases. When \( d_2 = 1 \) and the impatience rate \( \theta = 0 \), this supermarket model is equivalently degraded to an ordinary supermarket model, that is, each waiting line has a server with exponential service times of rate \( \mu \). When \( d_1 = 1 \) and \( d_2 = 1 \), this supermarket model is equivalent to a system of \( N \) independent M/M/1 queues with impatient customers.

The main contributions of this paper are twofold. The first one is to provide a key generalization of the supermarket model both from the impatient customers and from a doubly dynamic control. We use the operator semigroup to provide the mean-field limit for the sequence of infinite-dimensional Markov processes, and give a computational framework for organizing a Lipschitzian condition, which is always a key for proving existence and uniqueness of solution to the infinite-dimensional system of differential equations by means of the Picard approximation. In Section 5, we provide a computational framework for organizing a Lipschitzian condition, which is always a key for proving existence and uniqueness of solution to the infinite-dimensional system of differential equations. In Section 5, we provide an effective and efficient algorithm for computing the fixed point. In Section 7, we use the fixed point give performance analysis of this supermarket model, and provide some numerical examples to analyze how the performance measures depend on some crucial parameters of this supermarket model. Some concluding remarks are given in the final section.

2. A supermarket model with impatient customers

In this section, we describe a supermarket model with impatient customers, which consists of one server and \( N \) waiting lines under a doubly dynamic randomized load balancing control. We use the fraction vector of waiting lines with at least \( k \) customers for \( k \geq 0 \) to organize an infinite-dimensional Markov process which expresses the state of the supermarket model at time \( t \geq 0 \).

2.1. Model description

Let us describe the supermarket model with impatient customers, which consists of one server and \( N \) waiting lines under a doubly dynamic randomized load balancing control, where the random factors of this supermarket model are listed as follows:

\textbf{The arrival process and its dynamically control:} Customers arrive at the supermarket model as a Poisson process with arrival rate \( \lambda \) for \( \lambda > 0 \). Upon arrival, each customer chooses \( d_1 \geq 1 \) waiting lines from the \( N \) waiting lines independently and uniformly at random, and joins the one whose queue length is the shortest among the \( d_1 \) selected waiting lines. If there is a tie, waiting lines with the shortest queue are chosen randomly.

\textbf{The service process and its dynamical control:} There is one server working among the \( N \) waiting lines in this system, and the service times are i.i.d. and is exponential with service rate \( N \mu \) for \( \mu > 0 \). The server always chooses \( d_2 \geq 1 \) waiting lines independently and uniformly at random from the \( N \) waiting lines, and provides its service in the one whose queue length is the longest among the \( d_2 \) selected waiting lines.

\textbf{The impatient process:} The waiting customers excluding the one in service are impatient, and each of them has an exponential sojourn time with impatience rate \( N \theta \) for \( \theta > 0 \). Obviously, if there is a customer at the server and \( k = 1 \) customers in the waiting room (note that there are \( k \) customers in the system), then \((k−1)N\theta \) is the total impatience rate due to the independently impatient behavior of the \( k−1 \) customers.

We assume that all the random variables defined above are independent of each other. Fig. 1 provides a physical illustration of the supermarket model with impatient customers.

![Fig. 1. A physical illustration of the supermarket model with impatient customers.](image-url)
In this supermarket model depicted in Fig. 1, the choice number \( d_1 \geq 1 \) expresses the route scheduling of arriving customers and the choice number \( d_2 \geq 1 \) shows the effective use of total service ability. When \( d_2 = 1 \), this supermarket model is equivalently degraded to an ordinary supermarket model, where each waiting line has a server with exponential service times of rate \( \mu \).

It is easy to see from the above doubly dynamically control for the arrival and service processes that the generalized supermarket model may be related to the size-based scheduling through the centered management of the customer resource as well as the total service ability.

The following lemma provides a sufficient condition under which the supermarket model with impatient customers is stable.

**Lemma 1.** For the supermarket model with impatient customers which consists of one server and \( N \) waiting lines under a doubly dynamic randomized load balancing control, it is stable if \( \rho = \lambda / \mu < 1 \).

**Proof.** If \( d_1 = d_2 = 1 \), then this supermarket model is equivalent to a system of \( N \) independent \( M/M/1 \) queues with impatient customers. It is well-known that the \( M/M/1 \) queue with impatient customers is stable if \( \rho < 1 \). Using a coupling method, as given in Martin and Suhov [35, Theorems 4 and 5], this supermarket model with impatient customers is stable if \( \rho < 1 \). This completes the proof. \( \square \)

**Remark 1.** In a practical supermarket model, the impatient behavior of customers may have different classes, such as,

- **Class one:** The impatient behavior only exists for customers in the waiting room. Once a customer enters a server, it must finish its complete service and then leave this system. Such an impatient example can be seen in, such as, bank, supermarket, and hospital.

- **Class two:** The impatient behavior exists for customers with a time-length constraint, thus the impatient customers may leave this system both from the waiting room and from the server. Such an impatient example can be seen in either meeting, interview or perishable product.

### 2.2. An infinite-dimensional Markov process

For \( k \geq 0 \), we denote by \( n_k(t) \) the numbers of waiting lines with at least \( k \) customers at time \( t \geq 0 \). Clearly, \( n_0(t) = N \) and \( 0 \leq n_k(t) \leq N \) for \( k \geq 1 \). We write that for \( k \geq 0 \)

\[
U_{k}^{(N)}(t) = n_k(t) / N,
\]

which is the fraction of waiting lines with at least \( k \) customers at time \( t \geq 0 \). Let

\[
U_{k}^{(N)}(t) = (U_{0}^{(N)}(t), U_{1}^{(N)}(t), U_{2}^{(N)}(t), \ldots).
\]

Then the state of this supermarket model with impatient customers is described as an infinite-dimensional stochastic process \( \{U_{k}^{(N)}(t) : t \geq 0\} \). Since the arrival process is Poisson and the service and impatient times are exponential, \( \{U_{k}^{(N)}(t) : t \geq 0\} \) is an infinite-dimensional Markov process whose state space is given by

\[
E_N = \{u^{(N)} : 1 \geq U_{0}^{(N)} \geq U_{1}^{(N)} \geq U_{2}^{(N)} \cdots \geq 0, N u_{k}^{(N)} \text{ is a nonnegative integer for } k \geq 0\}.
\]

For a fixed pair \( (t, N) \) with \( t \geq 0 \) and \( N = 1, 2, 3, \ldots \), it is easy to see from the stochastic order that \( U_{k}^{(N)}(t) \geq U_{k+1}^{(N)}(t) \) for \( k \geq 0 \). This gives that

\[
1 = U_{0}^{(N)}(t) \geq U_{1}^{(N)}(t) \geq U_{2}^{(N)}(t) \cdots \geq 0.
\]

Since for a fixed pair \( (t, N) \), the sample value of \( U_{k}^{(N)}(t) \) is in the set

\[
\{0, 1 \ldots, N-1, N\} / N \ldots / N\} = \{0, 1 \ldots, N-1, 0\},
\]

there exists a larger positive integers \( K \) such that \( U_{K}^{(N)}(t) \geq U_{K+1}^{(N)}(t) \geq \cdots \geq U_{K+1}^{(N)}(t) > 0 \) and \( U_{K}^{(N)}(t) = 0 \) for \( k \geq K+1 \).

To study the infinite-dimensional Markov process \( \{U_{k}^{(N)}(t) : t \geq 0\} \), we write the expected fractions as follows:

\[
u_{k}^{(N)}(t) = E[U_{k}^{(N)}(t)]
\]

and

\[
u^{(N)}(t) = (u_{0}^{(N)}(t), u_{1}^{(N)}(t), u_{2}^{(N)}(t), \ldots).
\]

It is clear that

\[1 = u_{0}^{(N)}(t) \geq u_{1}^{(N)}(t) \geq u_{2}^{(N)}(t) \geq \cdots.\]

### 3. The system of differential equations

In this section, for this supermarket model with impatient customers, we set up an infinite-dimensional system of differential equations satisfied by the expected fraction vector \( \nu^{(N)}(t) \) for \( t \geq 0 \) by means of the technique of tail probabilities, e.g., see Li [28].

In the supermarket model with impatient customers, to set up an infinite-dimensional system of differential equations satisfied by the expected fraction vector \( \nu^{(N)}(t) \), we need to determine the expected change in the number of waiting lines with at least \( k \) customers over a small time period \([0, dt)\). To that end, our computation contains the following three parts:

(a) **The arrival process:** The rate that any arriving customer selects \( d_1 \) waiting lines from the \( N \) waiting lines independently and uniformly at random, joins the one whose queue length is the shortest, and the queue lengths of the other selected \( d_1 - 1 \) waiting line are not shorter than \( k-1 \) is given by

\[
N \lambda \sum_{m=1}^{d_1} \binom{d_1}{m} \nu_{m}^{(N)}(t) - \nu_{m+1}^{(N)}(t) \nu_{m}^{(N)}(t)^m - \nu_{m}^{(N)}(t)^m dt = N \lambda \nu_{m}^{(N)}(t) - \nu_{m}^{(N)}(t)^m dt,
\]

where \( \binom{d_1}{m} dt = [m!(d_1-m)!] \) is a binomial coefficient. Note that in the \( d_1 \) selected waiting lines and \( 1 \leq m \leq d_1 \), \( \nu_{m}^{(N)}(t) - \nu_{m+1}^{(N)}(t) \nu_{m}^{(N)}(t)^m \) is the probability that any arriving customer who can only choose one server makes \( m \) independent selections from the \( m \) servers with the shortest queue length \( k-1 \), while \( \nu_{m}^{(N)}(t)^m \) is the probability that there are \( d_1 - m \) selected waiting lines whose queue length are not shorter than \( k \).

(b) **The service process:** The server always chooses \( d_2 \geq 1 \) waiting lines independently and uniformly at random from the \( N \) waiting lines, and provides its service in the one whose queue length is the longest among the \( d_2 \) selected waiting lines. Based on this, the rate that the server goes to one waiting line with the longest queue length \( k \) among the \( d_2 \) selected waiting lines, and the queue lengths of the other selected \( d_2 - 1 \) waiting lines are not longer than \( k \), is given by

\[
N \mu \sum_{m=1}^{d_2} \binom{d_2}{m} \nu_{m}^{(N)}(t) - \nu_{m+1}^{(N)}(t) \nu_{m}^{(N)}(t)^m - \nu_{m}^{(N)}(t)^m dt = N \mu ([1 - \nu_{m}^{(N)}(t)^m] dt.
\]

Note that in the \( d_2 \) selected waiting lines and \( 1 \leq m \leq d_2 \), \( \nu_{m}^{(N)}(t) - \nu_{m+1}^{(N)}(t) \nu_{m}^{(N)}(t)^m \) is the probability that the server makes \( m \) independent selections from the \( m \) waiting lines with the longest queue length \( k \), while \( [1 - \nu_{m}^{(N)}(t)^m] dt \) is the probability that there are \( d_2 - m \) selected waiting lines whose queue length are not longer than \( k - 1 \).

(c) **The impatient process:** The rate that one impatient customer cannot wait for his service and leave this system from its waiting
line queued by \( k \) customers is given by
\[
N(k-1)\partial t[u_1^{(N)}(t) - u_{k+1}^{(N)}(t)] dt. \tag{3}
\]

Based on the above analysis from (1) to (3), we obtain
\[
d[u_k^{(N)}(t)] = N\lambda [u_{k-1}^{(N)}(t) - [u_k^{(N)}(t)]^k] dt - N\mu ([1 - u_{k+1}^{(N)}(t)]^k - [1 - u_k^{(N)}(t)]^k) dt - N(k-1)\partial [u_k^{(N)}(t) - u_{k+1}^{(N)}(t)] dt.
\]

This gives that for \( k \geq 1 \)
\[
\frac{d}{dt}u_k^{(N)}(t) = \lambda \left( [u_{k-1}^{(N)}(t)]^k - [u_k^{(N)}(t)]^k \right) - \mu \left( [1 - u_{k+1}^{(N)}(t)]^k - [1 - u_k^{(N)}(t)]^k \right) - (k-1)\partial [u_k^{(N)}(t) - u_{k+1}^{(N)}(t)], \tag{4}
\]
which is due to \( u_k^{(N)}(t) = E[u_k^{(N)}(t)]/N \).

Using some similar analysis to Eqs. (4), we can obtain an infinite-dimensional system of differential equations satisfied by the expected fraction vector \( u^{(N)}(t) = (u_0^{(N)}(t), u_1^{(N)}(t), u_2^{(N)}(t), \ldots) \) as follows:
\[
\frac{d}{dt}u_0^{(N)}(t) = \lambda \left( u_0^{(N)}(t) \right) - \mu \left( u_0^{(N)}(t) \right), \tag{5}
\]
and for \( k \geq 2 \)
\[
\frac{d}{dt}u_k^{(N)}(t) = \lambda \left( [u_{k-1}^{(N)}(t)]^k - [u_k^{(N)}(t)]^k \right) - \mu \left( [1 - u_{k+1}^{(N)}(t)]^k - [1 - u_k^{(N)}(t)]^k \right) - (k-1)\partial [u_k^{(N)}(t) - u_{k+1}^{(N)}(t)] \tag{6}
\]
with the boundary condition
\[
u_0^{(N)}(t) = 1, \quad t \geq 0,
\]
and the initial conditions
\[
u_k^{(N)}(0) = g_k, \quad k \geq 0,
\]
where
\[1 = g_0 \geq g_1 \geq g_2 \geq \cdots \geq 0.
\]

In the remainder of this section, we consider a special supermarket model with \( d_1 = d_2 = 1 \), that is, the \( M/M/1 \) queue with impatient customers. We denote by \( N(t) \) the number of customers in the \( M/M/1 \) queue with impatient customers. Then \( N(t) = 0, 1, 2, \ldots \) It is easy to see that \( N(t) : t \geq 0 \) is a Markov chain whose state transition relation is depicted in Fig. 2.

For \( k \geq 0 \), we write
\[
p_k(t) = P(N(t) = k).
\]

Using Fig. 2, we obtain
\[
\frac{d}{dt}p_0(t) = -\lambda p_0(t) + \mu p_1(t), \tag{9}
\]
\[
\frac{d}{dt}p_1(t) = -\lambda p_1(t) + \mu p_0(t) + (\mu + \theta) p_2(t), \tag{10}
\]
for \( k \geq 2 \)
\[
\frac{d}{dt}p_k(t) = -[\lambda + \mu + (k-1)\theta] p_k(t) + \lambda p_{k-1}(t) + (\mu + k\theta) p_{k+1}(t). \tag{11}
\]

Let
\[
Q_k(t) = \sum_{i=k}^{\infty} p_i(t), \quad k \geq 0,
\]
Then it follows from (9) to (11) that
\[
\frac{d}{dt}Q_k(t) = \lambda [Q_{k-1}(t) - Q_k(t)] - \mu [Q_k(t) - Q_{k+1}(t)], \tag{12}
\]
for \( k \geq 2 \)
\[
\frac{d}{dt}Q_k(t) = \lambda [Q_{k-1}(t) - Q_k(t)] - [\lambda + (k-1)\theta] [Q_k(t) - Q_{k+1}(t)]. \tag{13}
\]
with the boundary condition
\[
Q_0(t) = 1. \tag{14}
\]

It is easy to see that Eqs. (12) and (13) are the same as Eqs. (5) and (6) for \( d_1 = d_2 = 1 \).

### 4. A mean-field limit

In this section, we use the operator semigroup to provide a mean-field limit for the sequence \( \{u^{(N)}(t) : t \geq 0\} \) of infinite-dimensional Markov processes, and show that this sequence of Markov processes can asymptotically approach a single trajectory identified by the unique and global solution to the infinite-dimensional system of limiting differential equations.

For the vector \( u^{(N)} = (u_0^{(N)}, u_1^{(N)}, u_2^{(N)}, u_3^{(N)}, \ldots) \), we write
\[
\tilde{\Omega}_N = \{ u^{(N)} : 1 = u_0^{(N)} \geq u_1^{(N)} \geq u_2^{(N)} \geq \cdots \geq 0, \quad N u_k^{(N)} \text{ is a nonnegative integer for } k \geq 0 \}.
\]

![Fig. 2. State transition relation in the M/M/1 queue with impatient customers.](image-url)
Note that the vector space $\tilde{\Omega}$ is separable and compact under the metric $\rho((\mathbf{u}, \mathbf{u}'), (\mathbf{u}, \mathbf{u}'))$

Now, we consider the sequence $\{U(t) : t \geq 0\}$ of infinite-dimensional Markov processes on state space $\Omega_N$ (or $\tilde{\Omega}$). It indicates that the stochastic evolution of this supermarket model with impatient customers is described as the infinite-dimensional Markov process $\{U(t) : t \geq 0\}$, where

$$\frac{\mathrm{d} U(t)}{\mathrm{d} t} = A_N f(U(t)),$$

where $A_N$ acting on functions $f : \Omega_N \to \mathbb{R}$ is the generating operator of the Markov process $\{U(t) : t \geq 0\}$.

$$A_N = A_{N0} + A_{N1},$$

for $g \in \Omega_N$

$$A_{N0} f = 2N \sum_{k=1}^{\infty} (g_k^{d_k} - g_k^{-d_k}) \left\{ f\left( g + \frac{e_k}{N} \right) - f(g) \right\},$$

and

$$A_{N1} f = \mu N \sum_{k=1}^{\infty} \left\{ (1-g_k^{d_k} - (1-g_k^{-d_k}) \left\{ f\left( g + \frac{e_k}{N} \right) - f(g) \right\} \right\} + \theta N \sum_{k=2}^{\infty} \left( (k-1)(g_k^{d_k} - g_{k-1}^{-d_k}) \left\{ f\left( g + \frac{e_k}{N} \right) - f(g) \right\} \right).$$

The operator semigroup of the Markov process $\{U(t) : t \geq 0\}$ is defined as $T_N(t)$, where if $f : \Omega_N \to \mathbb{R}$, then for $g \in \Omega_N$ and $t \geq 0$

$$T_N(t)f(\mathbf{u}) = E(f(U_N(t)|\mathbf{U}_N(0) = \mathbf{u}).$$

Note that $A_N f$ is the generating operator of the operator semigroup $T_N(t)$, it is easy to see that $T_N(t)f = \exp(A_N t)f$ for $t \geq 0$.

**Definition 1.** A operator semigroup $\{S(t) : t \geq 0\}$ on the Banach space $L = C(\tilde{\Omega})$ is said to be strongly continuous if $\lim_{t \to 0} S(t) f = f$ for every $f \in L$; it is said to be a contractive semigroup if $\|S(t)f\| \leq 1$ for all $f \in L$.

Let $L = C(\tilde{\Omega})$ be the Banach space of continuous functions $f : \tilde{\Omega} \to \mathbb{R}$ with uniform metric $\|f\| = \max_{u \in \tilde{\Omega}} |f(u)|$, and similarly, let

$$L_N = C(\Omega_N).$$

The inclusion $\Omega_N \subset \tilde{\Omega}$ induces a contraction mapping $\Pi_N : L \to L_N$, $\Pi_N f(u) = f(u)$ for $f \in L$ and $u \in \Omega_N$.

Now, we consider the limiting process $\{U(t) : t \geq 0\}$ of the sequence $\{U_N(t) : t \geq 0\}$ of Markov processes for $N = 1, 2, 3, . . .$. That is, $U(t) = \lim_{N \to \infty} U_N(t)$ for $t \geq 0$, which leads to that $U(t) = \lim_{N \to \infty} U_N(t)$ for $t \geq 0$. In this case, two formal limits for the sequence $\{A_N\}$ of generating operators and for the sequence $\{T_N(t)\}$ of semigroups are expressed as $A = \lim_{N \to \infty} A_N$ and $T(t) = \lim_{N \to \infty} T_N(t)$ for $t \geq 0$, respectively. Note that as $N \to \infty$

$$N\left[ f\left( g + \frac{e_k}{N} \right) - f(g) \right] \to \frac{\partial f}{\partial g_k}(g)$$

and

$$N\left[ f\left( g - \frac{e_k}{N} \right) - f(g) \right] \to -\frac{\partial f}{\partial g_k}(g),$$

it follows from (19) that as $N \to \infty$

$$A f(g) = \lambda \sum_{k=1}^{\infty} (g_k^{d_k} - g_k^{-d_k}) \frac{\partial f}{\partial g_k}(g)$$

$$- \mu \sum_{k=1}^{\infty} (1 - g_k^{d_k}) - (1 - g_k^{-d_k}) \frac{\partial f}{\partial g_k}(g)$$

$$- \theta \sum_{k=2}^{\infty} (k-1)(g_k^{d_k} - g_{k-1}^{-d_k}) \frac{\partial f}{\partial g_k}(g).$$

Let $u(t) = \lim_{N \to \infty} U_N(t)$ for $t \geq 0$, where $u(t) = \lim_{N \to \infty} U_N(t)$ for $t \geq 0$. Based on the limiting generating operator $A$ given in (21), as $N \to \infty$ it follows from the system of differential equations (5)–(8) that $u(t)$ is a solution to the following system of differential equations

$$\frac{\mathrm{d} u_1(t)}{\mathrm{d} t} = \mu \sum_{m=1}^{\infty} \left\{ [u_1(t)]^{d_1} - [u_1(t)]^{-d_1} \right\} - \mu \sum_{m=1}^{\infty} \left\{ [1 - u_2(t)]^{d_1} - [1 - u_2(t)]^{-d_1} \right\},$$

and for $k \geq 2$

$$\frac{\mathrm{d} u_k(t)}{\mathrm{d} t} = \lambda \frac{d_k}{g_k} \sum_{m=1}^{\infty} \left\{ [u_{k-1}(t)]^{d_k} - [u_{k-1}(t)]^{-d_k} \right\} - \mu \sum_{m=1}^{\infty} \left\{ [1 - u_k(t)]^{d_k} - [1 - u_k(t)]^{-d_k} \right\} - (k-1) \theta [u_k(t) - u_{k-1}(t)],$$

with the boundary condition

$$u_0(t) = 1,$$

and the initial condition

$$u_0(0) = g_k, \quad k \geq 0,$$

where

$$1 = g_0 \geq g_1 \geq g_2 \geq g_3 \geq \ldots \geq 0.$$

It is worthwhile to note that in the next section, we shall prove the existence and uniqueness of solution to the system of differential equations (22)–(25) by means of the Picard approximation.

We define a mapping $g \to u(t, g)$, where $u(t, g)$ is a solution to the system of differential equations (22)–(25). It is clear that $u(0, g) = g$. For the operator semigroup $T(t)$ acts in the space $L$. If $f \in L$ and $g \in \tilde{\Omega}$, then

$$T(t)f(g) = f(u(t, g)).$$

It is easy to see that the operator semigroups $T_N(t)$ and $T(t)$ are strongly continuous and constructive. We denote by $\mathcal{D}(A)$ the domain of the generating operator $A$. It follows from (26) that if $f$ is a function from $L$ and has the partial derivatives $(\partial / \partial g_k) f(g)$ and $(\partial^2 / \partial g_k \partial g_m) f(g)$, and there exists a
positive constant $C = C(f) < +\infty$ such that

$$\sup_{\begin{array}{l} i \geq 0 \\ t \geq 0 \end{array}} \left\{ \frac{\partial}{\partial g_k} f(g) \right\} < C$$

and

$$\sup_{\begin{array}{l} i \geq 0 \\ t \geq 0 \end{array}} \left\{ \frac{\partial^2}{\partial g_k^2} f(g) \right\} < C.$$  \hspace{1cm} (28)

We call that $f \in D$ depends only on the first $K$ variables if for $g^{(1)}, g^{(2)} \in \Omega$, it follows from $g^{(1)} = g^{(2)}$ for $0 \leq i \leq K - 1$ that $f(g^{(1)}) = f(g^{(2)})$. Proposition 2 in Vvedenskaya et al. [50] indicates that the set of functions from $f$ that depends on the first $K$ variables is dense in $L$.

**Definition 2.** Let $A$ be a closed linear operator on the Banach space $L = C(\Omega)$. A subspace $D$ of $\text{dom}(A)$ is said to be a core for $A$ if the closure of the restriction of $A$ to $D$ is equal to $A$, i.e., $\overline{\text{dom}(A)} = A$.

The following Lemma comes from Proposition 1 in Vvedenskaya, Dobrushin and Karpelevich. We restated it here for convenience of description.

**Lemma 2.** Consider an infinite-dimensional system of differential equations: For $k \geq 0$,

$$z_k(0) = c_k$$

and

$$\frac{dz_k(t)}{dt} = \sum_{i=0}^{\infty} z_i(t) a_{ik}(t) + b_k(t).$$

and let $\sum_{i=0}^{\infty} |a_{ik}(t)| \leq a, |b_k(t)| \leq b_0 \exp(bt), |c_k| \leq q, b_0 \geq 0$ and $a < b$. Then

$$z_k(t) \leq q \exp(at) + \frac{b_0}{b-a} \exp(bt) - \exp(at).$$

The following lemma is a key to prove that the set $D$ is a core for the generating operator $A$.

**Lemma 3.** Let $u(t, g)$ be a solution to the system of differential equations (22)–(25). Then

$$\sup_{\begin{array}{l} i \geq 0 \\ t \geq 0 \end{array}} \left\{ \frac{\partial}{\partial g_k} f(g) \right\} \leq \exp(at)$$

and

$$\sup_{\begin{array}{l} i \geq 0 \\ t \geq 0 \end{array}} \left\{ \frac{\partial^2}{\partial g_k^2} f(g) \right\} \leq \frac{a \lambda d_1 + \mu d_2 + (k-1)\theta}{\lambda d_1 + \mu d_2 + (k-1)\theta} \exp(2at) - \exp(at).$$

where

$$a = \lambda d_1 + \mu d_2 + (k-1)\theta.$$  \hspace{1cm} (29)

**Proof.** We only prove Inequality (29), while Inequality (30) can be proved similarly.

It is easy to verify that the solution $u(t, g)$ to the system of differential equations (22)–(25) possesses the partial derivatives $\partial u_i(t, g)/\partial g_j$ and $\partial^2 u_i(t, g)/\partial g_j \partial g_k$ for $i, j \geq 0$.

For simplicity of description, we will omit the argument of $\partial^2 u_i(t, g)/\partial g_j \partial g_k$. We write

$$u_k = u_k(t, g), u_{ij} = \frac{\partial u_i(t, g)}{\partial g_j}.$$  \hspace{1cm} (30)

It follows from (23) that for all $k, j \geq 2$.

$$\frac{du_{ij}}{dt} = d_1 (u_{i-1,j-1} u_{i-1,j} - u_{i-1,j} u_{i-1,j-1}) + d_2 (1 - u_{i-1,j}^2 - u_{i-1,j} - (1 - u_{i-1,j})^2 - u_{i-1,j} - u_{i-1,j-1})$$

$$- \theta (k-1) u_{ij} u_{i-1,j-1}.$$  \hspace{1cm} (27)

Using Lemma 2, we obtain Inequality (29) with $a = \lambda d_1 + \mu d_2 + (k-1)\theta, b_0 = 0, q = 1$.

This completes the proof. \hspace{1cm} ⊖

**Lemma 4.** The set $D$ is a core for the generating operator $A$.

**Proof.** It is obvious that $D$ is dense in $L$ and $D \in \text{dom}(A)$. Let $D_0$ be the set of functions from $f$ which depend only on the first finite variables. It is easy to see that $D_0$ is dense in $L$. Using Proposition 3.3 in Chapter 1 of Ethier and Kurtz [22], it can be shown that for any $t > 0$, the operator semigroup $\text{exp}((t) D_0$ does not bring $D_0$ out of $D$. Select an arbitrary function $\varphi \in D_0$ and let $f(g) = \varphi(u(t, g))$ for $g \in \Omega$. It follows from Lemma 3 that $f$ has the partial derivatives $\partial u_i(t, g)/\partial g_j$ and $\partial^2 u_i(t, g)/\partial g_j \partial g_k$ and they satisfy the inequalities (29) and (30). Therefore $f \in D$. This completes the proof. \hspace{1cm} ⊖

The following theorem applies the operator semigroup to provide an mean-field limit, which shows that the sequence $(\text{U}^{(n)}(t) : t \geq 0)$ of Markov processes asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of differential equations (23)–(25).

**Theorem 1.** For any continuous function $f: \Omega \to \mathbb{R}$ and $t > 0$,

$$\lim_{N \to +\infty} \sup_{g \in \mathbb{Z}_k} \left| \text{U}^{(n)}(t) f(g) - f(u(t, g)) \right| = 0,$$

and the convergence is uniform in $t$ within any bounded interval.

**Proof.** This proof is to use the convergence of the operator semigroups as well as the convergence of their corresponding generating generators, e.g., see Ethier and Kurtz [22, Theorem 6.1 in Chapter 1]. Lemma 4 shows that the set $D$ is a core for the generating operator $A$. For any function $f \in D$, we have

$$\frac{\text{U}^{(n)}(t) f(g) - f(g)}{N} + \frac{\partial f(g)}{\partial g_k} - \frac{\partial^2 f(g)}{\partial g_k^2} \left( \frac{g - (1 - g_k)}{N} \right)$$

and

$$\frac{\text{U}^{(n)}(t) f(g) - f(g)}{N} - \frac{\partial f(g)}{\partial g_k} - \frac{\partial^2 f(g)}{\partial g_k^2} \left( \frac{g - (1 - g_k)}{N} \right)$$

where $0 < \lambda_{i,k}^1 - 1 < \lambda_{i,k}^1 < 1$ for $i = 1, 2$. Since

$$\left\{ \left| \frac{\lambda_{i,k}^1}{N} \right| \right\} \leq C \left\{ \left| \frac{\lambda_{i,k}^1}{N} \right| \right\}$$

we obtain

$$\left| A f(g) - f(g) \right| \leq C \left\{ \left| \frac{\lambda_{i,k}^1}{N} \right| \right\}$$

$$\left| A f(g) - f(g) \right| \leq C \left\{ \left| \frac{\lambda_{i,k}^1}{N} \right| \right\}$$

$$\left| A f(g) - f(g) \right| \leq C \left\{ \left| \frac{\lambda_{i,k}^1}{N} \right| \right\}$$

$$\left| A f(g) - f(g) \right| \leq C \left\{ \left| \frac{\lambda_{i,k}^1}{N} \right| \right\}$$

because $C$ and $g^{(1)} + (1 - g^{(1)}) + \sum_{k=1}^{\infty} g_k$ are all finite (note that $\sum_{k=1}^{\infty} g_k$ is the expected queue length at time 0, hence $\sum_{k=1}^{\infty} g_k < +\infty$, it is clear that as $N \to +\infty$.

$$\lim_{N \to +\infty} \sup_{g \in \mathbb{Z}_k} \left| A f(g) - f(g) \right| = 0.$$  \hspace{1cm} (31)

This completes the proof. \hspace{1cm} ⊖

Finally, we provide some interpretation on Theorem 1. If $\lim_{N \to +\infty} \text{U}^{(n)}(t) = u(t) = g \in \Omega$ in probability, then Theorem 1 shows that $u(t) = \lim_{N \to +\infty} \text{U}^{(n)}(t)$ is concentrated on the trajectory
It is easy to check that if the infinite-dimensional function $G(x)$ is Gateaux differentiable, then it is also Frechet differentiable. At the same time, we have

$$G_{C}(x) = G(x) = \mathcal{DG}(x).$$

(33)

Let $t = (t_{1}, t_{2}, t_{3}, ...)$ with $0 \leq t_{k} \leq 1$ for $k \geq 1$. We write

$$\mathcal{DG}(x + t\mathcal{O}(y - x)) =$$

$$= \begin{pmatrix}
\frac{\partial G_{1}(x + t_{1}(y - x))}{\partial x_{1}} & \frac{\partial G_{1}(x + t_{2}(y - x))}{\partial x_{2}} & \cdots & \frac{\partial G_{1}(x + t_{N}(y - x))}{\partial x_{N}} \\
\frac{\partial G_{2}(x + t_{1}(y - x))}{\partial x_{1}} & \frac{\partial G_{2}(x + t_{2}(y - x))}{\partial x_{2}} & \cdots & \frac{\partial G_{2}(x + t_{N}(y - x))}{\partial x_{N}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial G_{N}(x + t_{1}(y - x))}{\partial x_{1}} & \frac{\partial G_{N}(x + t_{2}(y - x))}{\partial x_{2}} & \cdots & \frac{\partial G_{N}(x + t_{N}(y - x))}{\partial x_{N}} \\
\end{pmatrix}.$$ 

(34)

Now, we provide two useful properties for the Gateaux derivatives of the infinite-dimensional functions. Obviously, the two useful properties also hold for the Frechet derivatives.

If the infinite-dimensional function $G : \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$ is Gateaux differentiable, then there exists a vector $t = (t_{1}, t_{2}, t_{3}, ...)$ with $0 \leq t_{k} \leq 1$ for $k \geq 1$ such that

$$\mathcal{G}(y) - G(x) = (y - x)\mathcal{DG}(x + t\mathcal{O}(y - x)).$$

(35)

Furthermore, we have

$$\|\mathcal{G}(y) - G(x)\| \leq \sup_{0 \leq t \leq 1} \|\mathcal{DG}(x + t(y - x))\| \|y - x\|.$$ 

(36)

For simplicity of description, we write that $u_{k} = u_{k}(t)$ for $k \geq 0$.

In this case, the system of differential equations (22)–(25) can be simplified as an initial value problem as follows:

$$\frac{d}{dt}u_{1} = \lambda(u_{2}^{(0)} - u_{1}^{(0)}) - \mu[(1 - u_{2}^{(0)})^{2} - (1 - u_{1}^{(0)})^{2}],$$

(37)

$$\frac{d}{dt}u_{k} = \lambda(u_{k+1}^{(0)} - u_{k}^{(0)}) - \mu[(1 - u_{k+1}^{(0)})^{2} - (1 - u_{k}^{(0)})^{2}] - \theta(k - 1)(u_{k} - u_{k+1}),$$

(38)

for $k \geq 1$.

Let $x = (x_{1}, x_{2}, x_{3}, ...)$ and $G(x) = (G_{1}(x), G_{2}(x), G_{3}(x), ...)$, where $x_{k}$ and $G_{k}(x)$ are scalar for $k \geq 1$. Then the matrix of partial derivatives of the infinite-dimensional function $G(x)$ is defined as

$$\mathcal{D}G(x) = \begin{pmatrix}
\frac{\partial G_{1}(x)}{\partial x_{1}} & \frac{\partial G_{1}(x)}{\partial x_{2}} & \cdots & \frac{\partial G_{1}(x)}{\partial x_{N}} \\
\frac{\partial G_{2}(x)}{\partial x_{1}} & \frac{\partial G_{2}(x)}{\partial x_{2}} & \cdots & \frac{\partial G_{2}(x)}{\partial x_{N}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial G_{N}(x)}{\partial x_{1}} & \frac{\partial G_{N}(x)}{\partial x_{2}} & \cdots & \frac{\partial G_{N}(x)}{\partial x_{N}} \\
\end{pmatrix}.$$ 

(39)

Now, we define two classes of derivatives for the infinite-dimensional function $G(x)$. In fact, they are a direct and minor generalization from the derivatives of finite-dimensional functions.

**Definition 3.** Let the infinite-dimensional function $G : \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$. Then we need to use the derivative of the infinite-dimensional function $G : \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$. Thus it is necessary to first provide some useful notation and definitions of derivatives. Readers may refer to Li et al. [29] for more details.

For the infinite-dimensional function $G : \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$, we write $x = (x_{1}, x_{2}, x_{3}, ...)$ and $G(x) = (G_{1}(x), G_{2}(x), G_{3}(x), ...)$, where $x_{k}$ and $G_{k}(x)$ are scalar for $k \geq 1$. Then the matrix of partial derivatives of the infinite-dimensional function $G(x)$ is defined as

$$\mathcal{D}G(x) = \begin{pmatrix}
\frac{\partial G_{1}(x)}{\partial x_{1}} & \frac{\partial G_{1}(x)}{\partial x_{2}} & \cdots & \frac{\partial G_{1}(x)}{\partial x_{N}} \\
\frac{\partial G_{2}(x)}{\partial x_{1}} & \frac{\partial G_{2}(x)}{\partial x_{2}} & \cdots & \frac{\partial G_{2}(x)}{\partial x_{N}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial G_{N}(x)}{\partial x_{1}} & \frac{\partial G_{N}(x)}{\partial x_{2}} & \cdots & \frac{\partial G_{N}(x)}{\partial x_{N}} \\
\end{pmatrix}.$$ 

(40)

For simplicity of description, we write that $u_{k} = u_{k}(t)$ for $k \geq 0$.

In this case, the system of differential equations (22)–(25) can be simplified as an initial value problem as follows:

$$\frac{d}{dt}u_{1} = \lambda(u_{2}^{(0)} - u_{1}^{(0)}) - \mu[(1 - u_{2}^{(0)})^{2} - (1 - u_{1}^{(0)})^{2}],$$

(41)

$$\frac{d}{dt}u_{k} = \lambda(u_{k+1}^{(0)} - u_{k}^{(0)}) - \mu[(1 - u_{k+1}^{(0)})^{2} - (1 - u_{k}^{(0)})^{2}] - \theta(k - 1)(u_{k} - u_{k+1}),$$

(42)

for $k \geq 2$.

Let $x = (x_{1}, x_{2}, x_{3}, ...)$ and $G(x) = (G_{1}(x), G_{2}(x), G_{3}(x), ...)$, where $x_{k}$ and $G_{k}(x)$ are scalar for $k \geq 1$. Then the matrix of partial derivatives of the function $G(x)$ is given by

$$\mathcal{D}G(x) = \begin{pmatrix}
A_{1}(x) & B_{1}(x) \\
C_{1}(x) & A_{2}(x) & B_{2}(x) \\
C_{2}(x) & A_{3}(x) & B_{3}(x) \\
\vdots & \vdots & \ddots & \vdots \\
C_{n-1}(x) & A_{n}(x) & B_{n}(x) \\
\end{pmatrix}.$$ 

(43)

where

$$A_{1}(x) = -\lambda d_{1}x_{1}^{(0)} - \mu d_{2}(1 - x_{1})^{2} - 1,$$

(44)

$$C_{1}(x) = \mu d_{2}(1 - x_{1})^{2} - 1,$$

and for $k \geq 2$

$$A_{k}(x) = -\lambda d_{1}x_{k}^{(0)} - \mu d_{2}(1 - x_{k})^{2} - (k - 1)\theta,$$

$$B_{k}(x) = \lambda d_{1}x_{k}^{(0)} - 1,$$

$$C_{k}(x) = \mu d_{2}(1 - x_{k+1})^{2} - 1 + (k - 1)\theta.$$ 

Hence it follows from (41) that

$$\|\mathcal{D}F(x)\| = \max_{k \geq 2} \left\{ \|A_{1}(x) + C_{1}(x)\|, \sup_{k \geq 2} \|B_{k}(x) + A_{k}(x) + C_{k}(x)\| \right\}.$$ 

(45)

Note that

$$\|A_{1}(x) + C_{1}(x)\| < \mu d_{2},$$

(46)

and for $k \geq 2$

$$\|B_{k}(x) + A_{k}(x) + C_{k}(x)\| < \lambda d_{1} + \mu d_{2}.$$ 

(47)
it follows from Equations (42) and (43) that
\[ \|DF(x)\| \leq \lambda d_1 + \mu d_2 \Rightarrow M. \]

Note that \( x = u \), this gives that for \( u \in \tilde{\Omega} \)
\[ \|DF(u)\| \leq M. \tag{44} \]

It follows from (35) and (44) that for \( u, v \in \tilde{\Omega} \)
\[ F(u) - F(v) \leq M \| u - v \|. \tag{45} \]

This indicates that the function \( F(u) \) is Lipschitzian for \( u \in \tilde{\Omega} \).

Note that \( x = u \) and \( u(0) = g \), it follows from Equations (22) to (25) that for \( u \in \tilde{\Omega} \)
\[ u(t) = u(0) + \int_0^t F(u(s)) \, ds, \]
this gives
\[ u(t) = g + \int_0^t F(u(s)) \, ds. \tag{46} \]

Based on the integral equation (46), we can apply the Picard approximation to prove the existence and uniqueness of solution to the system of differential equations (22)-(25), where the Lipschitzian condition organized above is a key when using the Picard approximation. At the same time, it is worthwhile to note that once the Lipschitzian condition is determined by our above procedure, the existence and uniqueness of solution can be proved by a standard method in the Banach spaces. Readers may refer to Li et al. [29] for more details.

6. The fixed point

In this section, we study the fixed point of the infinite-dimensional system of differential equations (22)-(25), and provide an effective and efficient algorithm for computing the fixed point by means of an infinite-dimensional system of nonlinear equations. Note that the fixed point is a key for performance analysis of this supermarket model with impatient customers.

Let \( \pi = (\pi_0, \pi_1, \pi_2, \pi_3, \ldots) \). A row vector \( \pi \) is called a fixed point of the infinite-dimensional system of differential equations (22)-(25) satisfied by the limiting expected fraction vector \( \pi(t) \), if \( \pi = \lim_{t \to +\infty} \pi(t) \), or \( \pi_k = \lim_{t \to +\infty} \pi_k(t) \) for \( k \geq 0 \). It is well-known that if \( \pi \) is a fixed point of the expected fraction vector \( \pi(t) \), then
\[ \lim_{t \to +\infty} \left[ \frac{d}{dt} \pi(t) \right] = 0. \]

Taking \( t \to +\infty \) in both sides of the system of differential equations (22)-(25), we obtain
\[ \lambda (\pi_0 - \pi_1^d) - \mu (1 - \pi_2^d) (1 - \pi_1^d) = 0, \tag{47} \]
for \( k \geq 2, \)
\[ \lambda (\pi_{k-1}^d - \pi_k^d) - \mu (1 - \pi_{k+1}^d) (1 - \pi_k^d) - \theta (k-1)(\pi_k - \pi_{k+1}) = 0, \tag{48} \]
with the boundary condition \( \pi_0 = 1. \tag{49} \)

In what follows we show how to solve the system of nonlinear equations (47)-(49).

From (47) and (48), we get
\[ \lambda \pi_k^d - \mu (1 - \pi_k^d) (1 - \pi_k^d) - \theta \sum_{k=2}^{\infty} \pi_k = 0, \tag{50} \]
which follows
\[ \pi_1 = 1 - \frac{\sqrt{\lambda - \mu - \theta}}{1 - \rho} \sum_{k=2}^{\infty} \pi_k. \tag{51} \]

It follows from (48) that
\[ \lambda \pi_k^d - \mu (1 - \pi_k^d) (1 - \pi_k^d) - \theta \sum_{k=2}^{\infty} \pi_k = 0, \]
this gives that for \( k \geq 2 \)
\[ \pi_k = \frac{\sqrt{\lambda - \mu - \theta}}{1 - \rho} \sum_{k=2}^{\infty} \pi_k. \tag{52} \]

Remark 2. If \( \theta = 0 \), then \( \pi_1 = 1 - \sqrt{1 - \rho} \) and \( \pi_{k+1} = 1 - \sqrt{1 - \rho \pi_k^d} \) for \( k \geq 1 \). In this case, it is easy to compute the fixed point \( \pi = (\pi_0, \pi_1, \pi_2, \ldots) \) for various values of \( d_1, d_2 \geq 1 \). At the same time, we have
\[ \pi_{k+1} = \frac{\sum_{i=0}^{d_2-1} \left( \rho^{(1/2) \left( 1 - \pi_k^d \right)} \right)}{\pi_k} \frac{\sum_{i=0}^{d_1-1} \left( \rho^{(1/2) \left( 1 - \pi_k^d \right)} \right)}{\pi_k}. \tag{53} \]

This product decomposition shows the associated role played by the two choice numbers \( d_1 \) and \( d_2 \), respectively.

It is observed from (51) and (52) that the parameter \( \theta \) related to the impatient customers makes computation of the fixed point more difficult. In this case, it follows from (47) that
\[ \pi_2 = 1 - \sqrt{\rho (1 - \pi_1^d) + (1 - \pi_1^d)^2} \tag{54} \]
and from (48) that
\[ \lambda \pi_2^d - \mu (1 - \pi_2^d) (1 - \pi_2^d) - \theta \sum_{i=2}^{\infty} \pi_i = 0, \]
they give
\[ \lambda (\pi_2^d - \pi_1^d) - \mu (1 - \pi_2^d) (1 - \pi_2^d) - \theta (1 - \pi_2^d) = 0. \]

Hence we obtain that for \( k \geq 1 \)
\[ \pi_{k+1} = 1 - \frac{\sqrt{\rho (\pi_k^d - \pi_{k-1}) + (1 - \pi_k^d)}}{\pi_k} \tag{55} \]
This gives some useful iterative relations for computing the fixed point once \( \pi_1 \) can be first determined.

Now, we further discuss the computational role played by the iterative relations (53) and (54).

Let \( \pi_1 = g \in (0,1) \). Then it is easy to see from the iterative relations (53) and (54) that \( \pi_1 \) is a function of the underdetermined number \( g \), denoted as \( \pi_1(g) \). From (50), it is shown that \( g \) is the maximal positive solution in \((0,1)\) to the nonlinear equation
\[ \lambda \pi_0^d - \mu (1 - \pi_1^d) (1 - \pi_1^d) - \theta \sum_{k=2}^{\infty} \pi_k = 0, \]
this gives
\[ \lambda - \mu (1 - g^d) - \theta \sum_{k=2}^{\infty} \pi_k(g) = 0. \tag{55} \]

In general, \( \pi_k \) always decreases very fast for \( d_1, d_2 \geq 2 \), as \( k \) increases. In this case, we take \( N \) as a suitable positive integer such that \( \pi_{N+j} = 0 \) for \( j > 1 \). Therefore, \( g \) can be approximated solved from the nonlinear equation
\[ \lambda - \mu (1 - g^d) - \theta \sum_{k=2}^{N} \pi_k(g) = 0. \]
Specifically, if \( \theta = 0 \), then \( g = 1 - \sqrt{1 - \rho} \).
Summarizing the above analysis, the following theorem provides an iterative algorithm for computing the fixed point $\pi = (\pi_0, \pi_1, \pi_2, \ldots)$, while its proof is easy and is omitted here.

**Theorem 2.** If $g$ is approximately solved from the nonlinear equation

$$\lambda - \mu [1 - (1 - g)^d] - \theta \sum_{k=2}^{N} \pi_k(g) = 0, \quad (56)$$

then

$$\pi_0 = 1,$$

$$\pi_1 = g$$

and for $k \geq 1$

$$\pi_{k+1} = 1 - \frac{\lambda}{\theta} \pi_k + \rho(\pi_{k-1} - \pi_k) + (1 - \pi_k) (1 - g)\pi_k.$$

Using Theorem 2, we consider some special cases as follows:

1. If $\theta = 0$ and $d_2 = 1$, then

$$\pi_k = \rho^k (1 - \pi_{k+1}^d), \quad k \geq 0,$$

where the fixed point $\{\pi_k\}$ is called to be doubly exponential.

2. If $\theta > 0$ and $d_2 = 1$, then

$$\pi_k = \pi_1 - (1 - \pi_{k+1}^d)$$

and for $k \geq 2$

$$\pi_{k+1} = \pi_1 + \theta \pi_{k+1} - \pi_{k+1}^d.$$

This gives

$$\sum_{k=2}^{\infty} \pi_k = \frac{\theta - \pi_1}{\theta} = \frac{\theta - g}{\theta}.$$

Using (56), we have

$$\lambda - \mu g - \theta \sum_{k=2}^{N} \pi_k = 0,$$

giving $\pi_1 = g = \rho$.

3. If $\theta = 0$ and $d_1, d_2 \geq 2$, then

$$\pi_1 = 1 - \sqrt[1-d_1]{1 - \rho}$$

and

$$\pi_k = \pi_1 - \frac{\rho^d}{1-d_1 - 1} \pi_{k-1}^d$$

for $k \geq 2$. At the same time, it is clear that

$$\lim_{k \to \infty} \pi_k = 0.$$

or

$$\lim_{k \to \infty} \frac{\ln \pi_k}{d_1 - 1} = 0,$$

where the fixed point $\{\pi_k\}$ is called to be super-exponential.

4. If $\theta > 0$ and $d_1, d_2 \geq 2$, then

$$\lim_{k \to \infty} \pi_k(\theta > 0) = 0$$

and

$$\lim_{k \to \infty} \pi_k(\theta = 0) = 0,$$

where the fixed point $\{\pi_k\}$ is also called to be super-exponential.

In the remainder of this section, we discuss some useful convergent properties of $u(t, g)$ as $N \to \infty$ and/or $t \to 0$.

Under the condition $\rho < 1$, the following theorem describes two important limiting processes which are related to the fixed point. The two limiting processes are crucial for understanding the convergence of $u(t, g)$ as $N \to \infty$ and/or $t \to 0$, and specifically, for performance approximation of this supermarket model with a bigger number $N$. Here, we omit its proof, while the proof can be completed by a similar discussion to those of Theorems 1 (iii) and Theorem 4 in Martin and Suhov [35].

**Theorem 3.** (1) If $\rho < 1$, then for any $g \in \Omega$ we have

$$\lim_{t \to \infty} u(t, g) = \pi.$$

Furthermore, there exists a unique probability measure $\phi$ on $\Omega$, which is invariant under the map $g \to u(t, g)$, that is, for any continuous function $f : \Omega \to \mathbb{R}$ and $t > 0$

$$\int f(g) \, d\phi(g) = \int f(u(t, g)) \, d\phi(g).$$

Also, $\phi = \delta_\pi$ is the probability measure concentrated at the fixed point $\pi$.

(2) If $\rho < 1$, then for a fixed number $N = 1, 2, 3, \ldots$, the Markov process $(U^{(N)}(t), t \geq 0)$ is positive recurrent, and hence it has a unique invariant distribution $\phi_N$. Furthermore, $(\phi_N)$ weakly converges to $\delta_\pi$, that is, for any continuous function $f : \Omega \to \mathbb{R}$

$$\lim_{N \to \infty} E_{\phi_N} f(g) = f(\pi).$$

Based on Theorems 3, we obtain a useful relation as follows:

$$\lim_{t \to \infty} \lim_{N \to \infty} u^{(N)}(t, g) = \lim_{N \to \infty} \lim_{t \to \infty} u^{(N)}(t, g) = \pi.$$

Therefore, we have

$$\lim_{N \to \infty} u^{(N)}(t, g) = \pi.$$

**7. Performance analysis**

In this section, based on the fixed point given in Theorem 2, we provide performance analysis of this supermarket model with impatient customers. Also, we use some numerical examples to discuss how the performance measures depend on some crucial factors of this supermarket model. Specifically, we observe that for improving system performance, the SLQ($d_1$) controlling the service processes is more effective than the JSQ($d_2$) controlling the arrival processes.

Now, we provide two useful performance measures of this supermarket model with impatient customers: The mean of stationary queue length in any waiting line, and the expected sojourn time of any arriving customer spending in this system.

Let $Q$ be the stationary queue length of any waiting line (including the customer being served in the server, if any). Then

$$E[Q] = \sum_{k=1}^{\infty} P(Q \geq k) = \sum_{k=1}^{\infty} \pi_k,$$

(57)

If $\rho < 1$, then this supermarket model is stable. In this case, we denote by $S$ the sojourn time of any arriving customer spending in this system. Note that a customer may leave the system because either it is impatient or after it is served, we need to introduce some events

$I = \{\text{The arriving customer shall leave the system because it is impatient}\}$

for $k \geq 1$

$I_k = \{\text{The arriving customer finds } k \text{ customers in its selected waiting line, and it shall leave the system because it is served}\}$

for $l \geq 0$
\( J_i = \) (The arriving customer finds \( k \) customers in its selected waiting line, and it shall leave the system because it is served).

It is easy to see that when computing \( E[S] \), the sample space is given by

\[ \Lambda = \bigcup_{i=0}^{\infty} J_i, \]

where

\[ \bigcup_{k=1}^{\infty} I_k, \quad J = \bigcup_{i=0}^{\infty} J_i. \]

Based on the total probability theorem, we obtain

\[ E[S] = E[S, I] + E[S, J], \]

where

\[ E[S, I] = \sum_{k=1}^{\infty} E[S | I_k] P[I_k] \]

and

\[ E[S, J] = \sum_{i=0}^{\infty} E[S | J_i] P[J_i]. \]

Since the probability that an arriving customer finds \( k \) customers in the selected waiting line is given by \( \pi_k^i - \pi_{k+1}^i \) for \( k \geq 0 \), we obtain

\[ E[S | I_k] = \frac{1}{\theta} \]

\[ P[I_k] = (\pi_k^i - \pi_{k+1}^i) P\left\{ X \leq \sum_{i=1}^{k} Y_i \right\}, \]

\[ E[S | J_i] = \frac{l + 1}{\mu} \]

and

\[ P[J_i] = (\pi_k^i - \pi_{k+1}^i) P\left\{ X > \sum_{i=1}^{k} Y_i \right\}, \]

where \( X \) is the exponential impatient time of impatience rate \( \theta \), and \( Y_i \) is the exponential service time of the \( i \)th customer. This gives

\[ E[S, I] = \frac{1}{\theta} \sum_{k=1}^{\infty} (\pi_k^i - \pi_{k+1}^i) P\left\{ X \leq \sum_{i=1}^{k} Y_i \right\} \] (58)

and

\[ E[S, J] = \frac{1}{\mu} (\pi_0^i - \pi_1^i) + \sum_{k=1}^{\infty} \frac{k+1}{\mu} (\pi_k^i - \pi_{k+1}^i) P\left\{ X > \sum_{i=1}^{k} Y_i \right\} \] (59)

It is easy to see that \( \sum_{i=1}^{k} Y_i \) is the sum of \( k \) exponential service times, and it is the Erlang distribution of order \( k \), given by

\[ P\left\{ \sum_{i=1}^{k} Y_i \leq y \right\} = 1 - \frac{1}{\Gamma(k) \lambda^k} \]

Hence we obtain

\[ P\left\{ X > \sum_{i=1}^{k} Y_i \right\} = \int_{0}^{+\infty} P\left\{ \sum_{i=1}^{k} Y_i < y \right\} dP[X \leq y] \]

\[ = \theta \int_{0}^{+\infty} e^{-\theta y} \left[ 1 - \frac{1}{\theta^k} \sum_{j=0}^{k-1} (\theta y)^j e^{-\theta y} \right] dy. \]

Let

\[ \omega_k = \theta \int_{0}^{+\infty} e^{-\theta y} \left[ 1 - \frac{1}{\theta^k} \sum_{j=0}^{k-1} (\theta y)^j e^{-\theta y} \right] dy. \]

Then it follows from (58) and (59) that

\[ E[S] = \frac{1}{\mu} \left[ (\pi_0^i - \pi_1^i) + \sum_{k=1}^{\infty} \omega_k (k+1)(\pi_k^i - \pi_{k+1}^i) \right] \]

\[ + \frac{1}{\mu} \sum_{k=1}^{\infty} (1 - \omega_k)(\pi_k^i - \pi_{k+1}^i). \] (60)

In the remainder of this section, using (57) and (60) we provide some numerical examples to analyze how the two performance measures \( E[Q] \) and \( E[S] \) depend on some crucial parameters of this supermarket model, for example, \( d_1, d_2, \lambda, \mu \) and \( \theta \).

Example one: the role of the arrival process: We consider the supermarket model with impatient customers, where the Poisson arrival rate \( \lambda \in [0.1, 0.9] \), the exponential service rate \( \mu = 1 \) and the exponential impatience rate \( \theta = 2 \).

Figs. 3 and 4 respectively indicate how \( E[Q] \) and \( E[S] \) depend on the arrival rate \( \lambda \in [0.1, 0.9] \) under the different values: \( d_1, d_2 = 1, 2, 3, 5, 6 \). It is seen from Figs. 3 and 4 that \( E[Q] \) and \( E[S] \) increase as \( \lambda \) increases, while each of them decreases as \( d_1 \) and \( d_2 \) increase. At the same time, it is easy to see that for improving system performance, the choice number \( d_2 \) for controlling the service process is more effective than the choice number \( d_1 \) for controlling the arrival process.

![Fig. 3. The mean of stationary queue length vs. \( \lambda \) for the different values of \( d_1, d_2 \).](image)

![Fig. 4. The expected sojourn time vs. \( \lambda \) for the different values of \( d_1, d_2 \).](image)
doubly dynamic control, which may also be related to the size-based scheduling with respect to the arrival and service processes. Firstly, we provide a probability method to set up the infinite-dimensional system of differential equations. Then we use the operator semigroup to provide the mean-field limit, which shows that the sequence of infinite-dimensional Markov processes asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of limiting differential equations. Finally, we provide an effective and efficient algorithm for computing the fixed point of the infinite-dimensional system of limiting differential equations, and use the fixed point to give performance analysis of this supermarket model. Also, some numerical examples are given to demonstrate how the performance measures depend on some crucial parameters of this supermarket model. From many practical applications, analysis of such a supermarket model is an important parallel queueing network to analyze the relation between the system performance and the job routing rule, and it can also help to design reasonable architecture to improve the performance and to balance the load.

Note that this paper provide a clear picture for how to use the mean-field theory as well as the numerical computation to analyze performance measures of more general supermarket models. We show that this picture is organized as three key parts: (1) setting up system of differential equations, (2) giving system of nonlinear equations satisfied by the fixed point through the mean-field limit, and (3) performance numerical computation of this supermarket model. Therefore, the results of this paper give new highlight on understanding influence of resource scheduling and information utilization on performance measures of more general supermarket models. Along such a line, there are a number of interesting directions for potential future research, for example:

- analyzing non-Poisson inputs such as renewal processes;
- studying non-exponential service time distributions, for example, general distributions, matrix-exponential distributions and heavy-tailed distributions; and
- discussing the bulk arrival processes, and/or the bulk service processes, where effective algorithms both for the fixed point and for the performance measures are necessary and interesting.

Acknowledgments

The authors thank the referees for their valuable comments to sufficiently improve the presentation of this paper. At the same time, Q.L. Li was supported by the National Natural Science Foundation of China under Grant nos. 71271187 and 71471160, the Hebei Natural Science Foundation of China under Grant no. A2012203125, and the Fostering Plan of Innovation Team and Leading Talent in Hebei Universities under Grant no. IJRC027.

References
