Discrete Event Dyn Syst
DOI 10.1007/s10626-013-0171-5

The mean-field computation in a supermarket model with server multiple vacations

Quan-Lin Li · Guirong Dai ·
John C. S. Lui · Yang Wang

Received: 21 November 2012 / Accepted: 8 October 2013

Abstract While vacation processes are considered to be ordinary behavior for servers, the study of queueing networks with server vacations is limited, interesting, and challenging. In this paper, we provide a unified and effective method of functional analysis for the study of a supermarket model with server multiple vacations. Firstly, we analyze a supermarket model of N identical servers with server multiple vacations, and set up an infinite-dimensional system of differential (or mean-field) equations, which is satisfied by the expected fraction vector, in terms of a technique of tailed equations. Secondly, as $N \to \infty$ we use the operator semigroup to provide a mean-field limit for the sequence of Markov processes, which asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of limiting differential equations. Thirdly, we provide an effective 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, including the mean of stationary queue length in any server and the expected sojourn time that any arriving customer spends in this system. Finally, we use some numerical examples to analyze how the performance measures depend on some crucial factors of this supermarket model. Note that the method of this paper will be useful and effective for performance analysis of

Q.-L. Li (✉) · G. Dai
School of Economics and Management Sciences, Yanshan University,
Qinhuangdao 066004, People's Republic of China
e-mail: liquanlin@mail.tsinghua.edu.cn

J. C. S. Lui
Department of Computer Science & Engineering,
The Chinese University of Hong Kong, Shatin, NT, Hong Kong

Y. Wang
Institute of Network Computing & Information Systems, Peking University,
Beijing 100871, People's Republic of China

Published online: 08 November 2013

complicated supermarket models with respect to resource management in practical areas such as computer networks, manufacturing systems and transportation networks.

Keywords Supermarket model · Randomized load balancing · Server vacation · Join the shortest queue · Expected fraction vector · Operator semigroup · Mean-field limit · Fixed point · Performance analysis

1 Introduction

During the last three decades considerable attention has been paid to studying queueing systems with server vacations. Queues with server vacations are always useful in modeling many real life situations such as digital communication, computer networks, production/inventory systems, transportation networks and business systems. Various queueing models with server vacations have been extensively reported by a number of authors, for example, basic vacation policies include server multiple vacations, server single vacations, server working vacations, N-policy, D-policy and T-policy. Reader may refer to Takagi (1991), Dshalalow (1995, 1997) and Tian and Zhang (2006) for more details. In the study of queueing systems with server vacations, an important result is stochastic decompositions of stationary queue length and of stationary waiting time. For single-server queues, the stochastic decompositions in the M/G/1 queue with server vacations were first established by Fuhrmann and Cooper (1985); while in multiple-server queues, the conditional stochastic decompositions for the M/M/c queues with server vacations were first analyzed in Tian et al. (1999). Up to now, extensive research on the single-server (or multiple-server) queueing systems with server vacations has been well-documented, such as, by three survey papers of Doshi (1986, 1990) and Alfa (2003), and by two books of Takagi (1991) and Tian and Zhang (2006).

Until now, the available results of queueing networks with server vacations has been very limited. Note that the supermarket models are an important class of queueing networks and play a key role in the area of networking resource management, thus the supermarket model with server vacations is very interesting in the study of queueing networks with server vacations, and it can also provide some new understanding and valuable highlight for the ordinary queues with server vacations which are described in Takagi (1991) and Tian and Zhang (2006). For queueing networks with server vacations, Vvedenskaya and Suhov (2005) first discussed a supermarket model with server On/Off vacations, and analyzed the stationary queue length distribution by means of the fixed point. However, the On/Off vacation discipline is not accurate for understanding the vacation processes, because it is not clear why to begin a vacation and how to end this vacation. This motivates us in this paper to further consider a supermarket model with server multiple vacations, while for other cases such as server single vacations and server working vacations, we can similarly give performance analysis. Note that the results given in Vvedenskaya and Suhov (2005) is very interesting, it also inspires us to further provide an effective algorithm to compute the fixed point with respect to the choice number $d \geq 3$, which have not been given a complete solution in the literature up to now. Note that the choice constant $d \leq N$, where N is the number of servers in the supermarket model.

Dynamic randomized load balancing is often referred to as the supermarket model. Recently, some supermarket models have been analyzed by means of queueing methods as well as Markov processes. For the simplest supermarket model (that is, Poisson arrivals and exponential service times), Vvedenskaya et al. (1996) applied the operator semigroups of Markov processes to analyze the stationary distribution and obtained an important result: Super-exponential decay tail. The super-exponential solution is a substantial improvement of system performance over that in the ordinary M/M/1 queue. At nearly the same time, Mitzenmacher (1996) also analyzed the same supermarket model in terms of the density-dependent jump Markov processes, e.g., see Kurtz (1981). Later, Turner (1998) provided a martingale approach to further discuss this supermarket model. The path space evolution of the supermarket model was studied by Graham (2000a, b, 2004) who showed that starting from independent initial states, as $N \to \infty$ the queues of the limiting process evolve independently. Luczak and Norris (2005) provided a strong approximation for the supermarket model, and Luczak and McDiarmid (2006, 2007) showed that the length of the longest queue scales as $(\log \log N) / \log d + O(1)$. The positive Harris recurrence of the Markov processes underlying some supermarket models was discussed in Foss and Chernova (1998) and Bramson (2008, 2011). Certain generalization of the supermarket model has been explored in studying various variations, for example, modeling more crucial factors by Mitzenmacher (1999), Jacquet and Vvedenskaya (1998), Jacquet et al. (1999) and Vvedenskaya and Suhov (2005); analyzing non-exponential server times or non-Poisson input by Bramson et al. (2010, 2012, 2011), Vvedenskaya and Suhov (1997), Mitzenmacher et al. (2001), Li et al. (2011, 2012), Li and Lui (2010) and Li (2011); fast Jackson networks by Martin and Suhov (1999), Martin (2001) and Suhov and Vvedenskaya (2002). Up to now, there have been three excellent survey papers by Turner (1996), Vvedenskaya and Suhov (1997) and Mitzenmacher et al. (2001), and one book by Mitzenmacher and Upfal (2005).

The mean-field equations and mean-field limits play an important role in the study of supermarket models. Readers may refer to recent publications for the mean-field models, among which are Sznitman (1989), Vvedenskaya and Suhov (1997), Le Boudec et al. (2007), Benaim and Le Boudec (2008), Bordenave et al. (2009), Gast and Gaujal (2009, 2012), Gast et al. (2011) and Tsitsiklis and Xu (2012). This paper provides a clear picture for illustrating how to use mean-field models to numerically analyze performance measures of complicated supermarket models, and is organized into three key parts: (Part one) setting up system of differential equations, see Section 2. (Part two) theoretical support, see Sections 3 and 4. In Section 3, we use the operator semigroup to give some strict proofs for the mean-field limit (or propagation of chaos), which shows the asymptotic independence of queues in the supermarket model with server vacations. Section 4 is a necessary supplementary part of the mean-field limit, in which the Lipschitzian condition is established for guaranteeing the existence and uniqueness of solution to the system of limiting differential equations. (Part three) performance analysis, Sections 5 and 6 provide a novel mean-field method for being able to numerically analyze performance measures of this supermarket model after the basic preparation given in Sections 3 and 4. Although analysis of the supermarket model with a finitely big N is very difficult, we use the mean-field limit to be able to numerically analyze performance measures of one queue, the information of which will help us to understand the total behavior of this supermarket model as $N \to \infty$. It is worthwhile to note that some simulations in Bramson et al. (2010, 2012, 2011) indicated that the asymptotic independence of queues can be formed well when $N \geq 100$. Therefore, the method of this paper is effective for performance analysis of complicated supermarket models.

The main contributions of this paper are threefold. The first one is to provide a unified and effective method for setting up an infinite-dimensional system of differential (or mean-field) equations, which is satisfied by the expected fraction vector in terms of a technique of tailed equations. Specifically, we derive an important relation: *the invariance of environment factor*. Note that the invariance of environment factor plays a key role in our later study with respect to this supermarket model. The second contribution is the development of a useful technique for establishing the Lipschitzian condition for the infinite-dimensional fraction vector function $f : \mathbf{R}_+^\infty \to \mathbf{C}^1(\mathbf{R}_+^\infty)$ for the general choice number $d \geq 1$. Note that the choice number $d = 2$ was always assumed in several important references, e.g., see Vvedenskaya and Suhov (1997, 2005) and Mitzenmacher et al. (2001). As seen in this paper, the case with $d = 2$ has a special structure in the system of nonlinear equations satisfied by the fixed point, which is easily dealt with from some simple computation; while for the case with $d \geq 3$, this paper gives some new and interesting results when establishing the the Lipschitzian condition, which leads to the strict proofs for the mean-field limit. The third contribution of this paper is to provide an effective algorithm for computing the fixed point, and also to provide performance analysis of this supermarket model. Note that our algorithm has a key which has the ability to determine the boundary probabilities in the system of nonlinear equations satisfied by the fixed point.

The remainder of this paper is organized as follows. In Section 2, we describe a supermarket model of N identical servers with server multiple vacations, introduce the sequence of fraction vectors which express the supermarket model as infinite-dimensional Markov processes, and set up an infinite-dimensional system of differential equations satisfied by the expected fraction vector in terms of a technique of tailed equations. In Section 3, we use the operator semigroup to provide a mean-field limit for the sequence of Markov processes, which asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of limiting differential equations. In Section 4, we provide a unified and effective method for organizing the Lipschitzian condition for the infinite-dimensional fraction vector function $f : \mathbf{R}_+^\infty \to \mathbf{C}^1(\mathbf{R}_+^\infty)$. Then we apply the Lipschitzian condition and the Picard approximation to show that the limiting expected fraction vector is the unique and global solution to the system of limiting differential equations. In Section 5, we provide an effective algorithm to compute the fixed point of the infinite-dimensional system of limiting differential equations. In Section 6, we use the fixed point to give performance analysis of this supermarket model, including the mean of the stationary queue length in any server and the expected sojourn time that any arriving customer spends in this system. Furthermore, we use some numerical examples to analyze how the performance measures depend on some crucial factors of this supermarket model. Some concluding remarks are given in the final section.

2 A supermarket model with server multiple vacations

In this section, we first describe a supermarket model of $N \geq 1$ identical servers with server multiple vacations. Then we introduce the sequence of fraction vectors, which are used to express the supermarket model as infinite-dimensional Markov processes. Finally, we provide a unified and effective method to set up an infinite-dimensional system of differential equations satisfied by the expected fraction vector of the supermarket model in terms of a technique of tailed equations.

The supermarket model consists of N identical servers, where each server has an infinite buffer. The service times of each server are i.i.d. with an exponential distribution of service rate $\mu = 1$. The vacation process of each server is based on the multiple vacation policy: When there is not any customer at one server and its buffer, it immediately takes a vacation and keeps taking vacations until it finds at least one customer waiting in the server or its buffer at the vacation completion instant. The vacation time distribution of each server is exponential with vacation rate $\theta > 0$. The common input flow is Poisson with arrival rate $N\lambda$ for $\lambda > 0$. Upon arrival, each customer chooses $d \geq 1$ servers from the N servers independently and uniformly at random, and joins the one whose queue length is the shortest. If there is a tie, servers with the shortest queue length are chosen randomly. All customers in any server will be served in the first-come-first-served (FCFS) manner, and the arrival, service and vacation processes are independent of each other. Figure 1 provides a physical illustration for the supermarket model of N identical servers with server multiple vacations.

Lemma 1 *The supermarket model of N identical servers with server multiple vacations is stable if* $0 < \lambda < 1$.

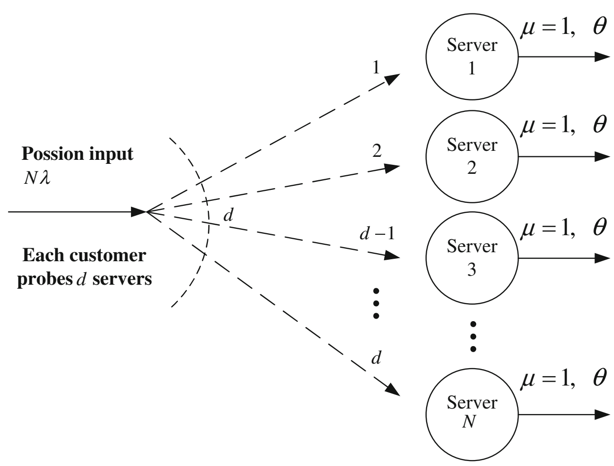


Fig. 1 A supermarket model with each customer choosing the loading of d servers

Proof If $d = 1$, then this supermarket model of N identical servers with server multiple vacations is equivalent to a system of N independent M/M/1 queues with server multiple vacations. From Chapter 1 of Tian and Zhang (2006), it is seen that the M/M/1 queue with server multiple vacations is stable if $\rho = \lambda/\mu = \lambda < 1$. Using a coupling method, as given in Theorems 4 and 5 of Martin and Suhov (1999), it is easy to see that for a fixed number $N = 1, 2, 3, \ldots$, this supermarket model of N identical servers is stable if $\rho = \lambda < 1$. This completes the proof. □

2.1 An infinite-dimensional Markov process

For this supermarket model, let $L_k^{(N)}(t)$ be the number of working servers with at least $k \geq 1$ customers (the serving customer is also taken into account) at time t, and $M_l^{(N)}(t)$ the number of vacation servers with at least $l \geq 0$ customers at time t. We write

$$U_k^{(N)}(t) = \frac{L_k^{(N)}(t)}{N}, \quad k \geq 1,$$

and

$$V_l^{(N)}(t) = \frac{M_l^{(N)}(t)}{N}, \quad l \geq 0.$$

Clearly, $U_k^{(N)}(t)$ for $k \geq 1$ and $V_l^{(N)}(t)$ for $l \geq 0$ are the fractions of these working servers with at least k customers at time t and the fractions of these vacation servers with at least l customers at time t, respectively. Set

$$\mathbf{U}^{(N)}(t) = (U_1^{(N)}(t), U_2^{(N)}(t), U_3^{(N)}(t), \ldots)$$

and

$$\mathbf{V}^{(N)}(t) = (V_0^{(N)}(t), V_1^{(N)}(t), V_2^{(N)}(t), \ldots).$$

It is easy to see that for any given $N \geq 1$, $\mathbf{U}^{(N)}(t)$ and $\mathbf{V}^{(N)}(t)$ are all random vectors. Based on the exponential or Poisson assumptions of the arrival, service and vacation processes, $\left\{\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right), t \geq 0\right\}$ is an infinite-dimensional Markov process whose state space $\mathbf{E}_N$ is given by

$$\begin{aligned}\mathbf{E}_N = \Big\{ & \left(u_1^{(N)}, u_2^{(N)}, u_3^{(N)}, \ldots; v_0^{(N)}, v_1^{(N)}, v_2^{(N)}, \ldots\right) : 1 \geq u_1^{(N)} \geq u_2^{(N)} \\ & \geq u_3^{(N)} \geq \cdots \geq 0, 1 \geq v_0^{(N)} \geq v_1^{(N)} \geq v_2^{(N)} \geq v_3^{(N)} \geq \cdots \geq 0, \\ & Nu_k^{(N)} \text{ and } Nv_l^{(N)} \text{ are nonnegative integers for } k \geq 1 \text{ and } l \geq 0 \Big\}.\end{aligned}$$

Note that $M_l^{(N)}(t) \geq M_{l+1}^{(N)}(t) \geq 0$ for $l \geq 0$ and $t \geq 0$, it is obvious that $1 \geq V_0^{(N)}(t) \geq V_1^{(N)}(t) \geq V_2^{(N)}(t) \geq \cdots \geq 0$. Similarly, the fact that $L_k^{(N)}(t) \geq L_{k+1}^{(N)}(t) \geq 0$ for $k \geq 1$ and $t \geq 0$ can yield that $1 \geq U_1^{(N)}(t) \geq U_2^{(N)}(t) \geq U_2^{(N)}(t) \geq \cdots \geq 0$. Furthermore, since the two random variables $U_k^{(N)}(t)$ and $V_l^{(N)}(t)$ take values in the set $\{0, 1/N, 2/N, \ldots, (N-1)/N, 1\}$ for $k \geq 1$, $l \geq 0$ and $t \geq 0$, this gives that for $t \geq 0$, there exist two positive integers K and L such that

$$1 \geq U_1^{(N)}(t) \geq U_2^{(N)}(t) \geq \cdots \geq U_K^{(N)}(t) > 0, \quad U_k^{(N)}(t) = 0 \text{ for } k \geq K+1;$$

and

$$1 \geq V_0^{(N)}(t) \geq V_1^{(N)}(t) \geq \cdots \geq V_L^{(N)}(t) > 0, \quad V_l^{(N)}(t) = 0 \text{ for } l \geq L+1.$$

To analyze the infinite-dimensional Markov process $\left\{\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right), t \geq 0\right\}$ on state space $\mathbf{E}_N$, we write

$$u_k^{(N)}(t) = E\left[U_k^{(N)}(t)\right], \quad k \geq 1,$$

and

$$v_l^{(N)}(t) = E\left[V_l^{(N)}(t)\right], \quad l \geq 0.$$

Let

$$\mathbf{u}^{(N)}(t) = \left(u_1^{(N)}(t), u_2^{(N)}(t), u_3^{(N)}(t), \ldots\right)$$

and

$$\mathbf{v}^{(N)}(t) = \left(v_0^{(N)}(t), v_1^{(N)}(t), v_2^{(N)}(t), \ldots\right).$$

It is easy to see that

$$1 \geq u_1^{(N)}(t) \geq u_2^{(N)}(t) \geq u_3^{(N)}(t) \geq \cdots \geq 0$$

and

$$1 \geq v_0^{(N)}(t) \geq v_1^{(N)}(t) \geq v_2^{(N)}(t) \geq \cdots \geq 0$$

with

$$v_0^{(N)}(t) + u_1^{(N)}(t) = 1.$$

In the remainder of this section, we set up an infinite-dimensional system of differential equations whose purpose is to be able to determine the expected fraction vector $(\mathbf{u}^{(N)}(t), \mathbf{v}^{(N)}(t))$.

2.2 The system of differential equations

To determine the expected fraction vector $\left(\mathbf{u}^{(N)}(t), \mathbf{v}^{(N)}(t)\right)$, this subsection provides a unified and effective method to set up an infinite-dimensional system of differential equations satisfied by the expected fraction vector in terms of a technique of tailed equations. To that end, we first provide an example with $k \geq 2$ to indicate how to derive these differential equations.

In the supermarket model of N identical servers, we need to determine the expected change in the number of servers with at least k customers over a small time period $\left[0, \mathrm{d}t\right)$, that is, we shall compute the rate that any arriving customer selects d servers from the N servers independently and uniformly at random and joins the one whose queue length is the shortest. From Figs. 2 and 3, it is seen that any arriving customer joins either server works or server vacations among the selected d servers, thus we need to consider the following two cases:

Case one: **Entering one working server**. In this case, the rate that any arriving customer joins a working server with the queue length $k-1$ and the

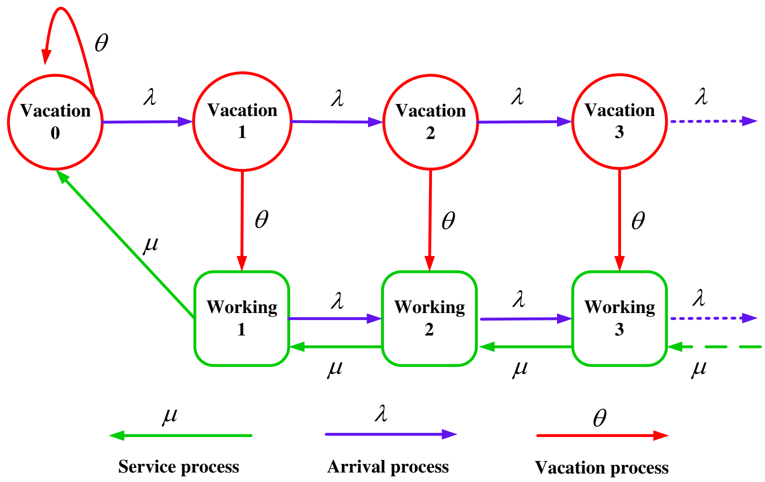


Fig. 2 The state transition relation in the M/M/1 queue with server vacations

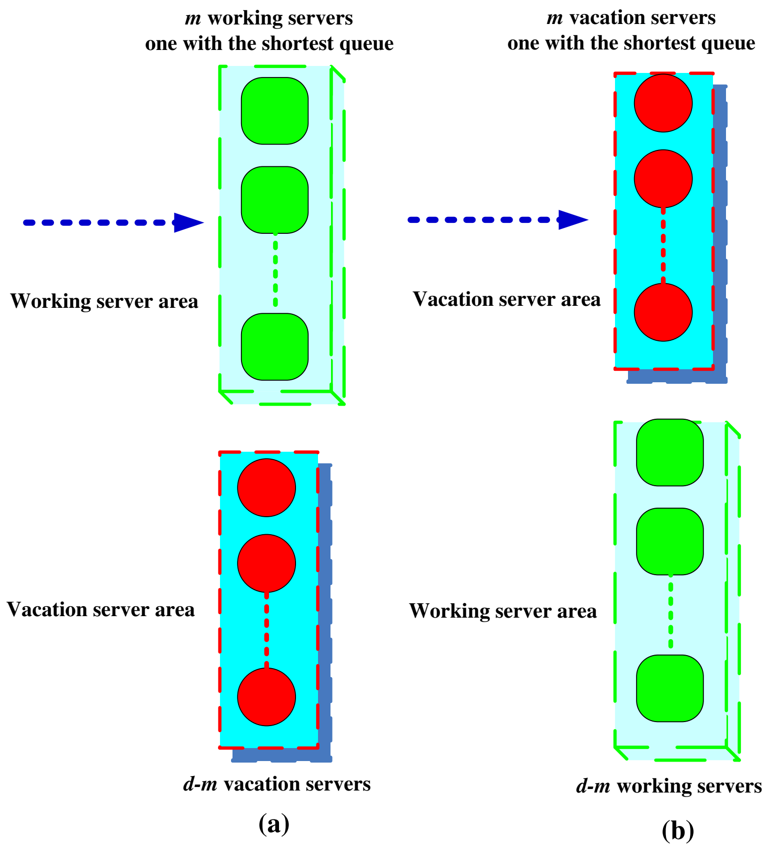


Fig. 3 Two different cases when joining a working server or a vacation server

queue lengths of the other selected $d-1$ servers are not shorter than $k-1$ is given by

$$N\lambda \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t) \right] W_k \left(u_{k-1}, u_k; v_{k-1}, v_k; t \right) \mathrm{d}t, \tag{1}$$

where

$$\begin{aligned} W_k^{(N)}\left(u_{k-1}, u_k; v_{k-1}, v_k; t\right) = & \sum_{m=1}^{d} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t) \right]^{m-1} \left[u_k^{(N)}(t) \right]^{d-m} \\ & + \sum_{m=1}^{d-1} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t) \right]^{m-1} \\ & \times \sum_{j=1}^{d-m} C_{d-m}^j \left[u_k^{(N)}(t) \right]^{d-m-j} \left[v_k^{(N)}(t) \right]^j \\ & + \sum_{m=2}^{d} C_d^m \sum_{m_1=1}^{m-1} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t) \right]^{m_1-1} \\ & \times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t) \right]^{m-m_1} \\ & \times \sum_{r=0}^{d-m} C_{d-m}^r \left[u_k^{(N)}(t) \right]^r \left[v_k^{(N)}(t) \right]^{d-m-r}. \end{aligned}$$

It is necessary to provide a detailed interpretation for how to derive Eq. 1. From the joining process expressed by (a) in Fig. 3 and from the set decomposition of all possible events indicated in Fig. 4, it is seen that the probability $W_k^{(N)}\left(u_{k-1}, u_k; v_{k-1}, v_k; t\right)$ given in Eq. 1 contain the following three parts.

Part I: **Neither of the selected d servers is taking a vacation**, that is, each of the selected d servers is working for service. In this case, the probability that any arriving customer joins a working server with the queue length $k-1$ and the queue

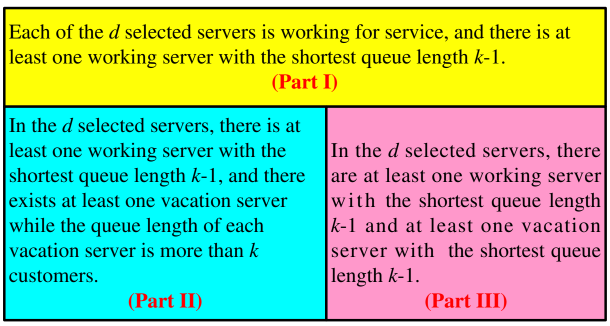


Fig. 4 Set decomposition of all possible events when joining a working server

lengths of the other selected $d-1$ working servers are not shorter than $k-1$ is given by

$$\sum_{m=1}^{d} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^m \left[u_k^{(N)}(t)\right]^{d-m}$$
$$= \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]$$
$$\times \sum_{m=1}^{d} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1} \left[u_k^{(N)}(t)\right]^{d-m},$$

where $C_d^m = d!/\left[m!\,(d-m)!\right]$ is a binomial coefficient, and $\left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^m$ is the probability that any arriving customer who can only choose one queue makes m independent selections during the m selected working servers with the queue length $k-1$ at time t.

Part II: **For the selected d servers, there is at least one working server with $k-1$ customers, and there exist at least one vacation server while the queue length of each vacation server is more than k customers**. In this case, the probability that any arriving customer joins a working server with the shortest queue length $k-1$; and for the other selected $d-1$ servers, the queue lengths of the selected working servers are not shorter than $k-1$, and there exist at least one vacation server while the queue length of each vacation server is more than k customers, is given by

$$\sum_{m=1}^{d-1} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^m$$
$$\times \sum_{j=1}^{d-m} C_{d-m}^j \left[u_k^{(N)}(t)\right]^{d-m-j} \left[v_k^{(N)}(t)\right]^j$$
$$= \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right] \sum_{m=1}^{d-1} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1}$$
$$\times \sum_{j=1}^{d-m} C_{d-m}^j \left[u_k^{(N)}(t)\right]^{d-m-j} \left[v_k^{(N)}(t)\right]^j.$$

Part III: **For the selected d servers, there are at least one working server with $k-1$ customers and at least one vacation server with $k-1$ customers**. In this case, if there are the selected m servers with the shortest queue length $k-1$ where there are $m_1 \geq 1$ working servers and $m-m_1$ vacation servers, then the probability that any arriving customer joins a working server is equal to m_1/m. Therefore, the probability that any arriving customer joins a working server with the queue length $k-1$, the queue lengths of the other selected $d-1$ servers are not shorter than $k-1$, and there are at least one working server with $k-1$ customers and at least one vacation server with $k-1$ customers, is given by

$$\sum_{m=2}^{d} C_d^m \sum_{m_1=1}^{m-1} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1} \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1}$$
$$\times \sum_{r=0}^{d-m} C_{d-m}^r \left[u_k^{(N)}(t)\right]^r \left[v_k^{(N)}(t)\right]^{d-m-r}$$
$$= \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right] \sum_{m=2}^{d} C_d^m$$
$$\times \sum_{m_1=1}^{m-1} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1}$$
$$\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1}$$
$$\times \sum_{r=0}^{d-m} C_{d-m}^r \left[u_k^{(N)}(t)\right]^r \left[v_k^{(N)}(t)\right]^{d-m-r}.$$

Using the above three parts, Eq. 1 can be obtained immediately.
Besides the above analysis for the arrival process, in what follows we consider the service and vacation processes. The rate that a customer leaves a server queued by k customers is given by

$$N\left[u_k^{(N)}(t) - u_{k+1}^{(N)}(t)\right] \mathrm{d}t. \tag{2}$$

The rate that a server queued by at least k customers completes its vacation is given by

$$N\theta v_k^{(N)}(t) dt. \tag{3}$$

Using Eqs. 1, 2 and 3, we obtain

$$\mathrm{d}E\left[L_k^{(N)}(t)\right] = N\lambda \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right] W_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t)\, \mathrm{d}t$$
$$-N\left[u_k^{(N)}(t) - u_{k+1}^{(N)}(t)\right] \mathrm{d}t + N\theta v_k^{(N)}(t) \mathrm{d}t,$$

this gives

$$\frac{\mathrm{d}}{\mathrm{d}t} u_k^{(N)}(t) = \lambda \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right] W_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t)$$
$$-\left[u_k^{(N)}(t) - u_{k+1}^{(N)}(t)\right] + \theta v_k^{(N)}(t) \tag{4}$$

by means of $u_k^{(N)}(t) = E\left[L_k^{(N)}(t)/N\right]$.

Case two: **Entering one vacation server**. In this case, the rate that any arriving customer joins a vacation server with the queue length $k-1$ and the queue lengths of the other selected $d-1$ servers are not shorter than $k-1$ is given by

$$N\lambda\left[v_{k-1}^{(N)}(t)-v_{k}^{(N)}(t)\right]V_{k}^{(N)}\left(u_{k-1},u_{k};v_{k-1},v_{k};t\right)\mathrm{d}t, \tag{5}$$

where

$$\begin{aligned}
V_{k}^{(N)}\left(u_{k-1},u_{k};v_{k-1},v_{k};t\right)=&\sum_{m=1}^{d}C_{d}^{m}\left[v_{k-1}^{(N)}(t)-v_{k}^{(N)}(t)\right]^{m-1}\left[v_{k}^{(N)}(t)\right]^{d-m}\\
&+\sum_{m=1}^{d-1}C_{d}^{m}\left[v_{k-1}^{(N)}(t)-v_{k}^{(N)}(t)\right]^{m-1}\\
&\times\sum_{j=1}^{d-m}C_{d-m}^{j}\left[v_{k}^{(N)}(t)\right]^{d-m-j}\left[u_{k}^{(N)}(t)\right]^{j}\\
&+\sum_{m=2}^{d}C_{d}^{m}\sum_{m_1=1}^{m-1}\frac{m_1}{m}C_{m}^{m_1}\left[v_{k-1}^{(N)}(t)-v_{k}^{(N)}(t)\right]^{m_1-1}\\
&\times\left[u_{k-1}^{(N)}(t)-u_{k}^{(N)}(t)\right]^{m-m_1}\\
&\times\sum_{r=0}^{d-m}C_{d-m}^{r}\left[v_{k}^{(N)}(t)\right]^{r}\left[u_{k}^{(N)}(t)\right]^{d-m-r}.
\end{aligned}$$

Note that Eq. 5 can be derived similarly to that in Case one by means of (b) in Figs. 3 and 5. Using a similar analysis to Eq. 4, it follows from Eq. 5 that

$$\frac{\mathrm{d}}{\mathrm{d}t}v_{k}^{(N)}(t)=\lambda\left[v_{k-1}^{(N)}(t)-v_{k}^{(N)}(t)\right]V_{k}^{(N)}\left(u_{k-1},u_{k};v_{k-1},v_{k};t\right)-\theta v_{k}^{(N)}(t). \tag{6}$$

The following theorem simplifies expressions for $V_{1}^{(N)}\left(u_{1};v_{0},v_{1};t\right)$, $V_{k}^{(N)}\left(u_{k-1},u_{k};v_{k-1},v_{k};t\right)$ and $W_{k}^{(N)}\left(u_{k-1},u_{k};v_{k-1},v_{k};t\right)$ for $k\geq 2$. Note that the simplified expressions will be a key in our later study.

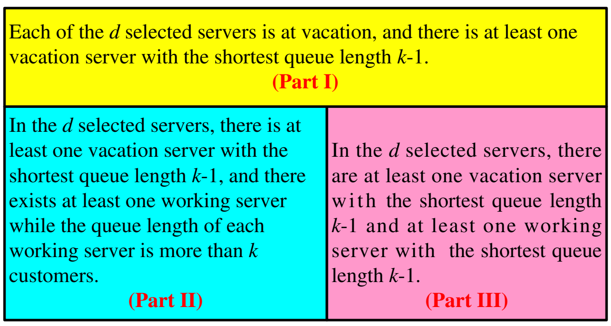


Fig. 5 Set decomposition of all possible events when joining a vacation server

Theorem 1

$$V_1^{(N)}(u_1; v_0, v_1; t) = \sum_{m=1}^{d} C_d^m \left[v_0^{(N)}(t) - v_1^{(N)}(t)\right]^{m-1} \left[v_1^{(N)}(t) + u_1^{(N)}(t)\right]^{d-m},$$

for $k \geq 2$

$$W_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = \sum_{m=1}^{d} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t) + v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-1} \times \left[u_k^{(N)}(t) + v_k^{(N)}(t)\right]^{d-m}$$

and

$$V_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = \sum_{m=1}^{d} C_d^m \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t) + u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1} \times \left[u_k^{(N)}(t) + v_k^{(N)}(t)\right]^{d-m}.$$

Hence, for $k \geq 2$ *we have*

$$W_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = V_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t).$$

Proof It is easy to see that

$$\begin{aligned} V_1^{(N)}(u_1; v_0, v_1; t) &= \sum_{m=1}^{d} C_d^m \left[v_0^{(N)}(t) - v_1^{(N)}(t)\right]^{m-1} \sum_{r=0}^{d-m} C_{d-m}^r \left[v_1^{(N)}(t)\right]^{d-m-j} \left[u_1^{(N)}(t)\right]^j \\ &= \sum_{m=1}^{d} C_d^m \left[v_0^{(N)}(t) - v_1^{(N)}(t)\right]^{m-1} \left[v_1^{(N)}(t) + u_1^{(N)}(t)\right]^{d-m}. \end{aligned}$$

For $k \geq 2$, we obtain

$$\begin{aligned} W_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = C_d^d &\left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{d-1} \\ &+ \sum_{m=1}^{d-1} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1} \left[u_k^{(N)}(t)\right]^{d-m} \\ &+ \sum_{m=1}^{d-1} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1} \end{aligned}$$

$$
\begin{aligned}
&\times \sum_{j=1}^{d-m} C_{d-m}^{j} \left[u_k^{(N)}(t)\right]^{d-m-j} \left[v_k^{(N)}(t)\right]^{j} \\
&+ \sum_{m=2}^{d} C_d^m \sum_{m_1=1}^{m-1} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1} \\
&\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1} \\
&\times \sum_{r=0}^{d-m} C_{d-m}^{r} \left[u_k^{(N)}(t)\right]^{r} \left[v_k^{(N)}(t)\right]^{d-m-r} \\
= {}& C_d^d \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{d-1} \\
&+ \sum_{m=1}^{d-1} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1} \\
&\times \sum_{j=0}^{d-m} C_{d-m}^{j} \left[u_k^{(N)}(t)\right]^{d-m-j} \left[v_k^{(N)}(t)\right]^{j} \\
&+ \sum_{m=2}^{d} C_d^m \sum_{m_1=1}^{m-1} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1} \\
&\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1} \sum_{r=0}^{d-m} C_{d-m}^{r} \left[u_k^{(N)}(t)\right]^{r} \left[v_k^{(N)}(t)\right] \\
= {}& C_d^1 \sum_{j=0}^{d-1} C_{d-1}^{j} \left[u_k^{(N)}(t)\right]^{d-1-j} \left[v_k^{(N)}(t)\right]^{j} \\
&+ \sum_{m=2}^{d} C_d^m \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1} \\
&\times \sum_{j=0}^{d-m} C_{d-m}^{j} \left[u_k^{(N)}(t)\right]^{d-m-j} \left[v_k^{(N)}(t)\right]^{j} \\
&+ \sum_{m=2}^{d} C_d^m \sum_{m_1=1}^{m-1} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1} \\
&\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1} \sum_{r=0}^{d-m} C_{d-m}^{r} \left[u_k^{(N)}(t)\right]^{r} \left[v_k^{(N)}(t)\right] \\
= {}& C_d^1 \sum_{j=0}^{d-1} C_{d-1}^{j} \left[u_k^{(N)}(t)\right]^{d-1-j} \left[v_k^{(N)}(t)\right]^{j}
\end{aligned}
$$

$$+\sum_{m=2}^{d} C_d^m \sum_{m_1=1}^{m} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1}$$

$$\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1} \sum_{r=0}^{d-m} C_{d-m}^r \left[u_k^{(N)}(t)\right]^r \left[v_k^{(N)}(t)\right]$$

$$= \sum_{m=1}^{d} C_d^m \sum_{m_1=1}^{m} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1}$$

$$\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1} \sum_{r=0}^{d-m} C_{d-m}^r \left[u_k^{(N)}(t)\right]^r \left[v_k^{(N)}(t)\right]$$

$$= \sum_{m=1}^{d} C_d^m \sum_{m_1=1}^{m} \frac{m_1}{m} C_m^{m_1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1}$$

$$\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1} \left[u_k^{(N)}(t) + v_k^{(N)}(t)\right]^{d-m}$$

$$= \sum_{m=1}^{d} C_d^m \left[u_k^{(N)}(t) + v_k^{(N)}(t)\right]^{d-m} \sum_{m_1=1}^{m} C_{m-1}^{m_1-1} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m_1-1}$$

$$\times \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-m_1}$$

$$= \sum_{m=1}^{d} C_d^m \left[u_k^{(N)}(t) + v_k^{(N)}(t)\right]^{d-m} \left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t) + v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]^{m-1},$$

similarly, we have

$$V_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = \sum_{m=1}^{d} C_d^m \left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t) + u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right]^{m-1}$$

$$\times \left[u_k^{(N)}(t) + v_k^{(N)}(t)\right]^{d-m}.$$

This completes the proof. □

Set

$$L_1^{(N)}(u_1; v_0, v_1; t) = V_1^{(N)}(u_1; v_0, v_1; t)$$

and for $k \geq 2$

$$L_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = W_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = V_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t).$$

The sequence: $L_1^{(N)}(u_1; v_0, v_1; t)$ and $L_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t)$ for $k \geq 2$, is called the invariance of environment factor, which will play a key role in our later study with respect to how to set up the system of differential equations.

Using some similar analysis to Eqs. 4 and 6, we obtain an infinite-dimensional system of differential equations satisfied by the expected fraction vector $\left(\mathbf{u}^{(N)}(t), \mathbf{v}^{(N)}(t)\right)$ as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}v_0^{(N)}(t) = \left[u_1^{(N)}(t) - u_2^{(N)}(t)\right] - \theta v_1^{(N)}(t), \tag{7}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}v_1^{(N)}(t) = \lambda\left[v_0^{(N)}(t) - v_1^{(N)}(t)\right] L_1^{(N)}(u_1; v_0, v_1; t) - \theta v_1^{(N)}(t), \tag{8}$$

for $k \geq 2$

$$\frac{\mathrm{d}}{\mathrm{d}t}v_k^{(N)}(t) = \lambda\left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right] L_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) - \theta v_k^{(N)}(t) \tag{9}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}u_k^{(N)}(t) = \lambda\left[u_{k-1}^{(N)}(t) - u_k^{(N)}(t)\right] L_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) - \left[u_k^{(N)}(t) - u_{k+1}^{(N)}(t)\right] + \theta v_k^{(N)}(t) \tag{10}$$

with the boundary condition

$$v_0^{(N)}(t) + u_1^{(N)}(t) = 1, \quad t \geq 0, \tag{11}$$

and with the initial conditions

$$\begin{cases} u_k^{(N)}(0) = g_k, & k \geq 1, \\ v_l^{(N)}(0) = h_l, & l \geq 0. \end{cases} \tag{12}$$

where

$$\begin{cases} 1 \geq g_1 \geq g_2 \geq g_3 \geq \cdots \geq 0, \\ 1 \geq h_0 \geq h_1 \geq h_2 \geq \cdots \geq 0, \\ h_0 + g_1 = 1. \end{cases}$$

Remark 1 If $d = 2$, then $W_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t) = u_{k-1}^{(N)}(t) + u_k^{(N)}(t) + v_{k-1}^{(N)}(t) + v_k^{(N)}(t)$ for $k \geq 2$ and $V_l^{(N)}(u_{l-1}, u_l; v_{l-1}, v_l; t) = u_{l-1}^{(N)}(t) + u_l^{(N)}(t) + v_{l-1}^{(N)}(t) + v_l^{(N)}(t)$ for $l \geq 1$. In this case, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}v_k^{(N)}(t) &= \lambda\left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]\left[u_{k-1}^{(N)}(t) + u_k^{(N)}(t) + v_{k-1}^{(N)}(t) + v_k^{(N)}(t)\right] - \theta v_k^{(N)}(t) \\ &= \lambda\left\{\left[v_{k-1}^{(N)}(t)\right]^2 - \left[v_k^{(N)}(t)\right]^2\right\} + \lambda\left[v_{k-1}^{(N)}(t) - v_k^{(N)}(t)\right]\left[u_{k-1}^{(N)}(t) + u_k^{(N)}(t)\right] \\ &\quad -\theta v_k^{(N)}(t). \end{aligned}$$

Therefore, the system of differential equations (7) to (12) is the same as those in Vvedenskaya and Suhov (2005).

2.3 A useful probabilistic interpretation

In this subsection, we provide a useful probabilistic interpretation for the invariance of environment factor $L_1^{(N)}(u_1; v_0, v_1; t)$ and $L_k^{(N)}(u_{k-1}, u_k; v_{k-1}, v_k; t)$ for $k \geq 2$, this will help us to further understand the system of differential equations (7) to (12).

Using Theorem 1, it is easy to check that

$$\left[v_0^{(N)}(t)-v_1^{(N)}(t)\right]V_1^{(N)}(u_1;v_0,v_1;t)=\left[v_0^{(N)}(t)+u_1^{(N)}(t)\right]^d-\left[v_1^{(N)}(t)+u_1^{(N)}(t)\right]^d,$$

for $k\geq 2$

$$\left[v_{k-1}^{(N)}(t)-v_k^{(N)}(t)\right]L_k^{(N)}(u_{k-1},u_k;v_{k-1},v_k;t)=\frac{v_{k-1}^{(N)}(t)-v_k^{(N)}(t)}{v_{k-1}^{(N)}(t)-v_k^{(N)}(t)+u_{k-1}^{(N)}(t)-u_k^{(N)}(t)}$$
$$\times\left\{\left[u_{k-1}^{(N)}(t)+v_{k-1}^{(N)}(t)\right]^d-\left[u_k^{(N)}(t)+v_k^{(N)}(t)\right]^d\right\}$$

and

$$\left[u_{k-1}^{(N)}(t)-u_k^{(N)}(t)\right]L_k^{(N)}(u_{k-1},u_k;v_{k-1},v_k;t)=\frac{u_{k-1}^{(N)}(t)-u_k^{(N)}(t)}{v_{k-1}^{(N)}(t)-v_k^{(N)}(t)+u_{k-1}^{(N)}(t)-u_k^{(N)}(t)}$$
$$\times\left\{\left[u_{k-1}^{(N)}(t)+v_{k-1}^{(N)}(t)\right]^d-\left[u_k^{(N)}(t)+v_k^{(N)}(t)\right]^d\right\}.$$

To give the probabilistic interpretation for $\left[v_0^{(N)}(t)-v_1^{(N)}(t)\right]V_1^{(N)}(u_1;v_0,v_1;t)$, $\left[v_{k-1}^{(N)}(t)-v_k^{(N)}(t)\right]L_k^{(N)}(u_{k-1},u_k;v_{k-1},v_k;t)$ and $\left[u_{k-1}^{(N)}(t)-u_k^{(N)}(t)\right]L_k^{(N)}(u_{k-1},u_k;v_{k-1},v_k;t)$ for $k\geq 2$, we introduce some notation

- $\mathbf{W}_{k-1}$ and $\mathbf{V}_{k-1}$ are the events in which any arriving customer is redirected to a working server or a vacation server with the queue length $k-1$, respectively.
- $\mathbf{X}_{k-1}$ denotes the number of times in the randomized load balancing policy we choose a server with the exactly queue length $k-1$, where we do not distinguish working and vacation servers.
- $\mathbf{Y}_{k-1}$ denotes the number of times in the randomized load balancing policy we choose a server with whose queue length is not shorter than $k-1$.

Now, we compute the two probabilities $P\{\mathbf{W}_{k-1}\}$ and $P\{\mathbf{V}_{k-1}\}$ for $k\geq 2$. Using the law of total probability, we obtain

$$P\{\mathbf{W}_{k-1}\}=\sum_{m=1}^{d}P\{\mathbf{W}_{k-1}|\mathbf{X}_{k-1}=m,\mathbf{Y}_{k-1}=d-m\}\,P\{\mathbf{X}_{k-1}=m,\mathbf{Y}_{k-1}=d-m\}.$$

We can compute the conditional probability

$$P\{\mathbf{W}_{k-1}|\mathbf{X}_{k-1}=m,\mathbf{Y}_{k-1}=d-m\}=\frac{u_{k-1}^{(N)}(t)-u_k^{(N)}(t)}{v_{k-1}^{(N)}(t)-v_k^{(N)}(t)+u_{k-1}^{(N)}(t)-u_k^{(N)}(t)},\quad(13)$$

which is independent of the number m. In fact, the conditional probability is easy to compute, e.g., by thinking in terms of an urn model with black and white balls, from which one draws m balls, black ones with probability

$$q=\frac{u_{k-1}^{(N)}(t)-u_k^{(N)}(t)}{v_{k-1}^{(N)}(t)-v_k^{(N)}(t)+u_{k-1}^{(N)}(t)-u_k^{(N)}(t)}$$

and white ones with probability $1-q$. Then, once m balls are extraced, one draws at random one ball from the m ones. The probability of having chosen a black ball is equal to q.

By means of Mitzenmacher (1996), we obtain

$$\sum_{m=1}^{d} P\left\{\mathbf{X}_{k-1}=m, \mathbf{Y}_{k-1}=d-m\right\}=\left[u_{k-1}^{(N)}(t)+v_{k-1}^{(N)}(t)\right]^{d}-\left[u_{k}^{(N)}(t)+v_{k}^{(N)}(t)\right]^{d}. \tag{14}$$

Based on Eqs. 13 and 14, we have the following probabilistic setting

$$\left[u_{k-1}^{(N)}(t)-u_{k}^{(N)}(t)\right] L_{k}^{(N)}\left(u_{k-1}, u_{k}; v_{k-1}, v_{k}; t\right)=P\left\{\mathbf{W}_{k-1}\right\}.$$

Similarly, we have

$$\left[v_{k-1}^{(N)}(t)-v_{k}^{(N)}(t)\right] L_{k}^{(N)}\left(u_{k-1}, u_{k}; v_{k-1}, v_{k}; t\right)=P\left\{\mathbf{V}_{k-1}\right\}.$$

Thus we obtain

$$P\left\{\mathbf{W}_{k-1}\right\}=P\left\{\mathbf{W}_{k-1} | \mathbf{X}_{k-1}=m, \mathbf{Y}_{k-1}=d-m\right\} \cdot \sum_{m=1}^{d} P\left\{\mathbf{X}_{k-1}=m, \mathbf{Y}_{k-1}=d-m\right\}$$

and

$$P\left\{\mathbf{V}_{k-1}\right\}=P\left\{\mathbf{V}_{k-1} | \mathbf{X}_{k-1}=m, \mathbf{Y}_{k-1}=d-m\right\} \cdot \sum_{m=1}^{d} P\left\{\mathbf{X}_{k-1}=m, \mathbf{Y}_{k-1}=d-m\right\}.$$

3 A mean-field limit

In this section, we use the operator semigroup to provide a mean-field limit for the sequence $\left\{\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right), t \geq 0\right\}$ of infinite-dimensional Markov processes for $N=1,2,3,\ldots$, and show that this sequence of Markov processes asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of limiting differential equations.

For the two vectors $\mathbf{u}^{(N)}=\left(u_{1}^{(N)}, u_{2}^{(N)}, u_{3}^{(N)}, \ldots\right)$ and $\mathbf{v}^{(N)}=\left(v_{0}^{(N)}, v_{1}^{(N)}, v_{2}^{(N)}, \ldots\right)$, we write

$$\begin{aligned}\widetilde{\Omega}_{N}=\Big\{\left(\mathbf{u}^{(N)}, \mathbf{v}^{(N)}\right): & \ 1 \geq u_{1}^{(N)} \geq u_{2}^{(N)} \geq u_{3}^{(N)} \geq \cdots \geq 0, \\ & 1 \geq v_{0}^{(N)} \geq v_{1}^{(N)} \geq v_{2}^{(N)} \geq \cdots \geq 0, \\ & N u_{k}^{(N)} \text{ and } N v_{l}^{(N)} \text{ are nonnegative integers for } k \geq 1 \text{ and } l \geq 0\Big\}\end{aligned}$$

and

$$\Omega_{N}=\left\{\left(\mathbf{u}^{(N)}, \mathbf{v}^{(N)}\right):\left(\mathbf{u}^{(N)}, \mathbf{v}^{(N)}\right) \in \widetilde{\Omega}_{N} \text{ and } \mathbf{u}^{(N)} e+\mathbf{v}^{(N)} e<+\infty\right\},$$

where e is a column vector of ones with a suitable dimension in the context.

For the two vectors $\mathbf{u}=\left(u_{1}, u_{2}, u_{3}, \ldots\right)$ and $\mathbf{v}=\left(v_{0}, v_{1}, v_{2}, v_{3}, \ldots\right)$, set

$$\widetilde{\Omega}=\left\{(\mathbf{u}, \mathbf{v}): 1 \geq u_{1} \geq u_{2} \geq u_{3} \geq \cdots \geq 0,1 \geq v_{0} \geq v_{1} \geq v_{2} \geq v_{3} \geq \cdots \geq 0\right\}$$

and

$$\Omega = \left\{ (\mathbf{u}, \mathbf{v}) : (\mathbf{u}, \mathbf{v}) \in \widetilde{\Omega} \text{ and } \mathbf{u}e + \mathbf{v}e < +\infty \right\}.$$

Obviously, $\Omega_N \subsetneqq \Omega \subsetneqq \widetilde{\Omega}$ and $\Omega_N \subsetneqq \widetilde{\Omega}_N \subsetneqq \widetilde{\Omega}$.

In the vector space $\widetilde{\Omega}$, we take a metric

$$\rho\left((\mathbf{u}, \mathbf{v}), (\mathbf{u}', \mathbf{v}')\right) = \sup_{k \geq 1} \left\{ \max \left\{ \frac{|u_k - u_k'|}{k}, \frac{|v_{k-1} - v_{k-1}'|}{k} \right\} \right\} \tag{15}$$

for $(\mathbf{u}, \mathbf{v}), (\mathbf{u}', \mathbf{v}') \in \widetilde{\Omega}$. Note that under the metric $\rho\left((\mathbf{u}, \mathbf{v}), (\mathbf{u}', \mathbf{v}')\right)$, the vector space $\widetilde{\Omega}$ is separable and compact.

For $(\mathbf{g}, \mathbf{h}) \in \Omega_N$, we write

$$L_1\left(g_1; h_0, h_1\right) = \sum_{m=1}^{d} C_d^m \left(h_0 - h_1\right)^{m-1} \left(h_1 + g_1\right)^{d-m},$$

for $k \geq 2$

$$L_k\left(g_{k-1}, g_k; h_{k-1}, h_k\right) = \sum_{m=1}^{d} C_d^m \left(g_{k-1} - g_k + h_{k-1} - h_k\right)^{m-1} \left(g_k + h_k\right)^{d-m}.$$

Now, we consider the infinite-dimensional Markov process $\{\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right), t \geq 0\}$ on state space Ω_N (or $\widetilde{\Omega}_N$ in a similar analysis) for $N = 1, 2, 3, \ldots$.

Note that the stochastic evolution of this supermarket model of N identical servers is described as the Markov process $\left\{\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right), t \geq 0\right\}$, where

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right) = \mathbf{A}_N f\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right),$$

where $\mathbf{A}_N$ acting on functions $f : \Omega_N \to \mathbf{R}$ is the generating operator of the Markov process $\left\{\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right), t \geq 0\right\}$,

$$\mathbf{A}_N = \mathbf{A}_N^{\text{In}} + \mathbf{A}_N^{\text{Out}}, \tag{16}$$

for $(\mathbf{g}, \mathbf{h}) \in \Omega_N$

$$\begin{aligned}
\mathbf{A}_N^{\text{In}} f(\mathbf{g}, \mathbf{h}) = {} & \lambda N \sum_{k=2}^{\infty} \left[\left(g_{k-1} - g_k\right) L_k\left(g_{k-1}, g_k; h_{k-1}, h_k\right)\right] \left[f(\mathbf{g} + \frac{\mathbf{e}_k}{N}, \mathbf{h}) - f(\mathbf{g}, \mathbf{h}) \right] \\
& + \lambda N \left[\left(h_0 - h_1\right) L_1\left(g_1; h_0, h_1\right)\right] \left[f(\mathbf{g}, \mathbf{h} + \frac{\mathbf{e}_1}{N}) - f(\mathbf{g}, \mathbf{h}) \right] \\
& + \lambda N \sum_{k=2}^{\infty} \left[\left(h_{k-1} - h_k\right) L_k\left(g_{k-1}, g_k; h_{k-1}, h_k\right)\right] \left[f(\mathbf{g}, \mathbf{h} + \frac{\mathbf{e}_k}{N}) - f(\mathbf{g}, \mathbf{h}) \right]
\end{aligned} \tag{17}$$

and

$$\mathbf{A}_N^{\text{Out}} = N\sum_{k=1}^{\infty}(g_k - g_{k+1})\left[f(\mathbf{g} - \frac{\mathbf{e}_k}{N}, \mathbf{h}) - f(\mathbf{g}, \mathbf{h})\right]$$
$$+\theta N\sum_{k=1}^{\infty}h_k\left[f(\mathbf{g} + \frac{\mathbf{e}_k}{N}, \mathbf{h} - \frac{\mathbf{e}_k}{N}) - f(\mathbf{g}, \mathbf{h})\right], \tag{18}$$

where $\mathbf{e}_k$ stands for a row vector with the kth entry 1 and all others 0.

For $(\mathbf{g}, \mathbf{h}) \in \Omega_N$, it follows from Eqs. 16 to 18 that

$$\mathbf{A}_N f(\mathbf{g}, \mathbf{h}) = \lambda N\sum_{k=2}^{\infty}\left[(g_{k-1} - g_k)\, L_k\,(g_{k-1}, g_k; h_{k-1}, h_k)\right]\left[f(\mathbf{g} + \frac{\mathbf{e}_k}{N}, \mathbf{h}) - f(\mathbf{g}, \mathbf{h})\right]$$
$$+\lambda N\left[(h_0 - h_1)\, L_1\,(g_1; h_0, h_1)\right]\left[f(\mathbf{g}, \mathbf{h} + \frac{\mathbf{e}_1}{N}) - f(\mathbf{g}, \mathbf{h})\right]$$
$$+\lambda N\sum_{k=2}^{\infty}\left[(h_{k-1} - h_k)\, L_k\,(g_{k-1}, g_k; h_{k-1}, h_k)\right]\left[f(\mathbf{g}, \mathbf{h} + \frac{\mathbf{e}_k}{N}) - f(\mathbf{g}, \mathbf{h})\right]$$
$$+N\sum_{k=1}^{\infty}(g_k - g_{k+1})\left[f(\mathbf{g} - \frac{\mathbf{e}_k}{N}, \mathbf{h}) - f(\mathbf{g}, \mathbf{h})\right]$$
$$+\theta N\sum_{k=1}^{\infty}h_k\left[f(\mathbf{g} + \frac{\mathbf{e}_k}{N}, \mathbf{h} - \frac{\mathbf{e}_k}{N}) - f(\mathbf{g}, \mathbf{h})\right] \tag{19}$$

The operator semigroup of the Markov process $\left\{\left(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)\right), t \geq 0\right\}$ is defined as $\mathbf{T}_N(t)$, where if $f : \Omega_N \to \mathbf{C}^1$, then for $\left(\mathbf{g}, \mathbf{h}\right) \in \Omega_N$ and $t \geq 0$

$$\mathbf{T}_N(t) f(\mathbf{g}, \mathbf{h}) = E\left[f(\mathbf{U}_N(t), \mathbf{V}_N(t)) | \mathbf{U}_N(0) = \mathbf{g}, \mathbf{V}_N(0) = \mathbf{h}\right]. \tag{20}$$

Note that $\mathbf{A}_N$ is the generating operator of the operator semigroup $\mathbf{T}_N(t)$, it is easy to see that $\mathbf{T}_N(t) = \exp\{\mathbf{A}_N t\}$ for $t \geq 0$.

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

Let $L = C(\widetilde{\Omega})$ be the Banach space of continuous functions $f : \widetilde{\Omega} \to \mathbf{R}$ with uniform metric $\|f\| = \max_{u\in\widetilde{\Omega}} |f(u)|$, and similarly, let $L_N = C(\Omega_N)$. The inclusion $\Omega_N \subset \widetilde{\Omega}$ 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 behavior of the sequence $\{(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)), t \geq 0\}$ of Markov processes for $N = 1, 2, 3, \ldots$. Two formal limits for the sequence $\{\mathbf{A}_N\}$ of generating operators and for the sequence $\{\mathbf{T}_N(t)\}$ of semigroups are expressed as $\mathbf{A} = \lim_{N\to\infty} \mathbf{A}_N$ and $\mathbf{T}\,(t) = \lim_{N\to\infty} \mathbf{T}_N(t)$ for $t \geq 0$, respectively. It follows from

Eq. 19 that as $N \to \infty$

$$\begin{aligned}\mathbf{A} f(\mathbf{g},\mathbf{h}) = & \lambda \sum_{k=2}^{\infty} \left[(g_{k-1} - g_k) L_k (g_{k-1}, g_k; h_{k-1}, h_k) \right] \frac{\partial}{\partial g_k} f(\mathbf{g},\mathbf{h}) \\ & + \lambda \left[(h_0 - h_1) L_1 (g_1; h_0, h_1) \right] \frac{\partial}{\partial h_1} f(\mathbf{g},\mathbf{h}) \\ & + \lambda N \sum_{k=2}^{\infty} \left[(h_{k-1} - h_k) L_k (g_{k-1}, g_k; h_{k-1}, h_k) \right] \frac{\partial}{\partial h_k} f(\mathbf{g},\mathbf{h}) \\ & - \sum_{k=1}^{\infty} (g_k - g_{k+1}) \frac{\partial}{\partial g_k} f(\mathbf{g},\mathbf{h}) + \theta \sum_{k=1}^{\infty} h_k \left[\frac{\partial f(\mathbf{g},\mathbf{h})}{\partial g_k} - \frac{\partial f(\mathbf{g},\mathbf{h})}{\partial h_k} \right]. \quad (21)\end{aligned}$$

We write

$$L_1 (u_1; v_0, v_1; t) = \sum_{m=1}^{d} C_d^m \left[v_0 (t) - v_1 (t) \right]^{m-1} \left[v_1 (t) + u_1 (t) \right]^{d-m},$$

for $k \geq 2$

$$\begin{aligned}L_k (u_{k-1}, u_k; v_{k-1}, v_k; t) = & \sum_{m=1}^{d} C_d^m \left[u_{k-1} (t) - u_k (t) + v_{k-1} (t) - v_k (t) \right]^{m-1} \\ & \times \left[u_k (t) + v_k (t) \right]^{d-m}.\end{aligned}$$

Let $\mathbf{u}(t) = \lim_{N\to\infty} \mathbf{u}^{(N)}(t)$ and $\mathbf{v}(t) = \lim_{N\to\infty} \mathbf{v}^{(N)}(t)$ for $t \geq 0$, where $u_k(t) = \lim_{N\to\infty} u_k^{(N)}(t)$ for $k \geq 1$ and $v_l(t) = \lim_{N\to\infty} v_l^{(N)}(t)$ for $l \geq 0$. Based on the limiting generating operator $\mathbf{A}$ given in Eq. 21, as $N \to \infty$ it follows from the system of differential equations (7) to (12) that $(\mathbf{u}(t), \mathbf{v}(t))$ is a solution to the following system of differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t} v_0 (t) = u_1 (t) - u_2 (t) - \theta v_1 (t), \quad (22)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} v_1 (t) = \lambda \left[v_0 (t) - v_1 (t) \right] L_1 (u_1; v_0, v_1; t) - \theta v_1 (t), \quad (23)$$

for $k \geq 2$

$$\frac{\mathrm{d}}{\mathrm{d}t} v_k (t) = \lambda \left[v_{k-1} (t) - v_k (t) \right] L_k (u_{k-1}, u_k; v_{k-1}, v_k; t) - \theta v_k (t), \quad (24)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} u_k (t) = \lambda \left[u_{k-1} (t) - u_k (t) \right] L_k (u_{k-1}, u_k; v_{k-1}, v_k; t) - \left[u_k (t) - u_{k+1} (t) \right] + \theta v_k (t), \quad (25)$$

with the boundary condition

$$v_0 (t) + u_1 (t) = 1, \quad t \geq 0, \quad (26)$$

and with the initial conditions

$$\begin{cases} u_k(0) = g_k, & k \geq 1, \\ v_l(0) = h_l, & l \geq 0. \end{cases} \tag{27}$$

Remark 2 In the next section, we shall prove that the vector $(\mathbf{u}(t,\mathbf{g},\mathbf{h}), \mathbf{v}(t,\mathbf{g},\mathbf{h}))$ is the unique and global solution to the system of differential equations (22) to (27) for $t \geq 0$, where $\mathbf{u}(0,\mathbf{g},\mathbf{h}) = \mathbf{g}$ and $\mathbf{v}(0,\mathbf{g},\mathbf{h}) = \mathbf{h}$.

We define a mapping: $(\mathbf{g},\mathbf{h}) \to (\mathbf{u}(t,\mathbf{g},\mathbf{h}), \mathbf{v}(t,\mathbf{g},\mathbf{h}))$, where $(\mathbf{u}(t,\mathbf{g},\mathbf{h}), \mathbf{v}(t,\mathbf{g},\mathbf{h}))$ is a solution to the system of differential equations (22) to (27). For the operator semigroup $\mathbf{T}(t)$ acts in the space L. If $f \in L$ and $(\mathbf{g},\mathbf{h}) \in \widetilde{\Omega}$, then

$$\mathbf{T}(t) f(\mathbf{g},\mathbf{h}) = f(\mathbf{u}(t,\mathbf{g},\mathbf{h}), \mathbf{v}(t,\mathbf{g},\mathbf{h})). \tag{28}$$

From Eqs. 19 and 21, it is easy to see that the operator semigroups $\mathbf{T}_N(t)$ and $\mathbf{T}(t)$ are strongly continuous and contractive, see, for example, Section 1.1 in Chapter one of Ethier and Kurtz (1986). We denote by $\mathcal{D}(\mathbf{A})$ the domain of the generating operator $\mathbf{A}$. It follows from Eq. 28 that if f is a function from L and has the partial derivatives $\frac{\partial}{\partial g_i} f(\mathbf{g},\mathbf{h})$ for $i \geq 1$ and $\frac{\partial}{\partial h_j} f(\mathbf{g},\mathbf{h}) \in L$ for $j \geq 0$, and $\sup\limits_{i\geq 1, j\geq 0} \left\{ \left| \frac{\partial}{\partial g_i} f(\mathbf{g},\mathbf{h}) \right|, \left| \frac{\partial}{\partial h_j} f(\mathbf{g},\mathbf{h}) \right| \right\} < \infty$, then $f \in \mathcal{D}(\mathbf{A})$.

Let D be the set of all functions $f \in L$ that have the partial derivatives $\frac{\partial}{\partial g_i} f(\mathbf{g},\mathbf{h})$, $\frac{\partial}{\partial h_j} f(\mathbf{g},\mathbf{h})$, $\frac{\partial^2}{\partial g_i \partial g_j} f(\mathbf{g},\mathbf{h})$, $\frac{\partial^2}{\partial g_i \partial h_j} f(\mathbf{g},\mathbf{h})$ and $\frac{\partial^2}{\partial h_i \partial h_j} f(\mathbf{g},\mathbf{h})$, and there exists $C = C(f) < +\infty$ such that

$$\sup_{i\geq 1, j\geq 0} \left\{ \left| \frac{\partial}{\partial g_i} f(\mathbf{g},\mathbf{h}) \right|, \left| \frac{\partial}{\partial h_j} f(\mathbf{g},\mathbf{h}) \right| \right\} < C \tag{29}$$

and

$$\sup_{\substack{i,i'\geq 1 \\ j,j'\geq 0 \\ (\mathbf{g},\mathbf{h})\in\widetilde{\Omega}}} \left\{ \left| \frac{\partial^2}{\partial g_i \partial g_{i'}} f(\mathbf{g},\mathbf{h}) \right|, \left| \frac{\partial^2}{\partial g_i \partial h_j} f(\mathbf{g},\mathbf{h}) \right|, \left| \frac{\partial^2}{\partial h_j \partial h_{j'}} f(\mathbf{g},\mathbf{h}) \right| \right\} < C. \tag{30}$$

We call that $f \in L$ depends only on the first K two dimensional variables if for $(\mathbf{g}^{(1)},\mathbf{h}^{(1)})$, $(\mathbf{g}^{(2)},\mathbf{h}^{(2)}) \in \widetilde{\Omega}$, it follows from $g_i^{(1)} = g_i^{(2)}$ for $1 \leq i \leq K$ and $h_j^{(1)} = h_j^{(2)}$ for $0 \leq j \leq K$ that $f(\mathbf{g}^{(1)},\mathbf{h}^{(1)}) = f(\mathbf{g}^{(2)},\mathbf{h}^{(2)})$. A similar and simple proof to that in Proposition 2 in Vvedenskaya et al. (1996) can show that the set of functions from L that depends on the first finite two dimensional variables is dense in L.

The following lemma comes from Proposition 1 in Vvedenskaya et al. (1996). We restated it here for convenience of description.

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

$$z_k(0) = c_k$$

and

$$\frac{\mathrm{d}z_k(t)}{\mathrm{d}t} = \sum_{i=0}^{\infty} z_i(t) a_{i,k}(t) + b_k(t),$$

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

$$z_k(t) \leq \varrho \exp\{at\} + \frac{b_0}{b-a}\left[\exp\{bt\} - \exp\{at\}\right].$$

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

We introduce some notations

$$M_1 = \sum_{m=1}^{d} C_d^m 2^{m-1} 2^{d-m} = 2^{d-1}\left(2^d - 1\right),$$

$$M_2 = \sum_{m=1}^{d} C_d^m (d-m)\, 2^{m-1} 2^{d-m} + \sum_{m=1}^{d} C_d^m (m-1)\, 2^{m-1} 2^{d-m}$$
$$= (d-1)\, 2^{d-1}\left(2^d - 1\right),$$

$$M_3 = 4\sum_{m=1}^{d} C_d^m (d-m)(d-m-1)\, 2^{m-1} 2^{d-m-2} + 16\sum_{m=1}^{d} C_d^m (m-1)(m-2)\, 2^{m-3} 2^{d-m}$$
$$+16\sum_{m=1}^{d} C_d^m (d-m)(m-1)\, 2^{m-2} 2^{d-m-1}$$
$$= 2^{d+1}\sum_{m=1}^{d} C_d^m (m-1)(d-2) + 2^{d-1}\sum_{m=1}^{d} C_d^m (d-m)(d-m-1),$$

$$a = 2M_1 + \theta + 2 + 2M_2$$

and

$$a' = 2M_1 + \theta + 2M_2.$$

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

Lemma 3 *Let* $(\mathbf{u}(t), \mathbf{v}(t))$ *be a solution to the system of differential equations* (22) *to* (27). *Then*

$$\sup_{\substack{k\geq 1\\ i\geq 1, j\geq 0}} \left\{\left|\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial g_i}\right|, \left|\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial h_j}\right|\right\} \leq \exp\{(2M_1 + \theta + 2 + 2M_2)t\}, \tag{31}$$

$$\sup_{\substack{k\geq 1\\ i\geq 1, j\geq 0}} \left\{\left|\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial g_i}\right|, \left|\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial h_j}\right|\right\} \leq \exp\{(2M_1 + \theta + 2M_2)t\}, \tag{32}$$

$$\sup_{\substack{k\geq 1\\ i,i'\geq 1\\ j,j'\geq 0}}\left\{\left|\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial g_{i'}}\right|,\left|\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial h_j}\right|,\left|\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial h_j \partial h_{j'}}\right|\right\}$$

$$\leq \frac{2M_2+2M_3}{2M_1+\theta+2+2M_2}\left[\exp\{2at\}-\exp\{at\}\right] \tag{33}$$

and

$$\sup_{\substack{k\geq 1\\ i,i'\geq 1\\ j,j'\geq 0}}\left\{\left|\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial g_{i'}}\right|,\left|\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial h_j}\right|,\left|\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial h_j \partial h_{j'}}\right|\right\}$$

$$\leq \frac{2M_2+2M_3}{2M_1+\theta+2M_2}\left[\exp\left\{2a't\right\}-\exp\left\{a't\right\}\right]. \tag{34}$$

Proof We only prove (31), while (32) to (34) can be proved similarly.

It is easy to verify that the solution $(\mathbf{u}(t),\mathbf{v}(t))$ to the system of differential equations (22) to (27) possesses the partial derivatives

$$\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial g_j},\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial h_j},\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial g_j},\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial h_j},$$

$$\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial g_j},\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial h_j},\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial h_i \partial h_j},$$

and

$$\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial g_j},\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial g_i \partial h_j},\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial h_i \partial h_j}.$$

In what follows we only compute the two derivatives $\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial g_j}$ and $\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial h_j}$, while the other derivatives can be computed similarly.

For simplicity of description, we write that $u_k=u_k(t,\mathbf{g},\mathbf{h})$, $u'_{k,j}=\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial g_j}$ or $v'_{k,j}=\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial h_j}$. It follows from Eqs. 22, 25 and 26 that

$$\frac{\mathrm{d}}{\mathrm{d}t}u'_{1,j}(t)=u'_{2,j}(t)-u'_{1,j}(t)+\theta v'_{1,j}(t)$$

and for all $k,j\geq 2$,

$$\begin{aligned}\frac{\mathrm{d}u'_{k,j}}{\mathrm{d}t}&=\lambda\left(u'_{k-1,j}-u'_{k,j}\right)L_k\left(u_{k-1},u_k;v_{k-1},v_k\right)-\left(u'_{k,j}-u'_{k+1,j}\right)+\theta v'_{k,j}\\&+\lambda\left(u_{k-1}-u_k\right)L'_{k,j}\left(u_{k-1},u_k;v_{k-1},v_k\right),\end{aligned}$$

and

$$L'_{k,j}(u_{k-1},u_k;v_{k-1},v_k;t)=\sum_{m=1}^{d}C_d^m(d-m)(u_{k-1}-u_k+v_{k-1}-v_k)^{m-1}$$
$$\times(v_k+u_k)^{d-m-1}\left(v'_{k,j}+u'_{k,j}\right)$$
$$+\sum_{m=1}^{d}C_d^m(m-1)(u_{k-1}-u_k+v_{k-1}-v_k)^{m-2}$$
$$\times(v_k+u_k)^{d-m}\left(u'_{k-1,j}-u'_{k,j}+v'_{k-1,j}-v'_{k,j}\right).$$

Using Lemma 2, we obtain Inequalities (31) with

$$a=2M_1+\theta+2+2M_2,\ a'=2M_1+\theta+2M_2,\ b=1,\ b_0=0,\ \varrho=1.$$

This completes the proof. □

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

Proof It is obvious that D is dense in L and $D\in\mathcal{D}(\mathbf{A})$. Let D_0 be the set of functions from D, which depend only on the first finite two dimensional variables. It is easy to see that D_0 is dense in L. Using Proposition 3.3 in Chapter 1 of Ethier and Kurtz (1986), it can show that for any $t\geq 0$, the operator semigroup $\mathbf{T}(t)$ does not bring D_0 out of D. Select an arbitrary function $\varphi\in D_0$ and let $f(\mathbf{g},\mathbf{h})=\varphi(\mathbf{u}(t,\mathbf{g},\mathbf{h}),\mathbf{v}(t,\mathbf{g},\mathbf{h}))$ for $(\mathbf{g},\mathbf{h})\in\widetilde{\Omega}$. It follows from Lemma 3 that f has the partial derivatives

$$\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial g_j},\frac{\partial u_k(t,\mathbf{g},\mathbf{h})}{\partial h_j},\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial g_j},\frac{\partial v_k(t,\mathbf{g},\mathbf{h})}{\partial h_j},$$

$$\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial g_i\partial g_j},\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial g_i\partial h_j},\frac{\partial^2 u_k(t,\mathbf{g},\mathbf{h})}{\partial h_i\partial h_j},$$

and

$$\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial g_i\partial g_j},\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial g_i\partial h_j},\frac{\partial^2 v_k(t,\mathbf{g},\mathbf{h})}{\partial h_i\partial h_j}.$$

and they satisfy the inequalities (31) to (34). Therefore $f\in D$. This completes the proof. □

The following theorem applies the operator semigroup to provide an mean-field limit, which shows that the sequence $\left\{\left(\mathbf{U}^{(N)}(t),\mathbf{V}^{(N)}(t)\right),t\geq 0\right\}$ of Markov processes asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of differential equations (22) to (27).

Theorem 2 *For any continuous function $f:\widetilde{\Omega}\to\mathbf{R}$ and $t>0$,*

$$\lim_{N\to\infty}\sup_{(\mathbf{g},\mathbf{h})\in\Omega_N}\left|\mathbf{T}_N(t)f(\mathbf{g},\mathbf{h})-f(\mathbf{u}(t,\mathbf{g},\mathbf{h}),\mathbf{v}(t;\mathbf{g},\mathbf{h}))\right|=0,$$

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 Theorem 6.1 in Chapter 1 of Ethier and Kurtz (1986). Lemma 4 shows that the set D is a core for the generating operator $\mathbf{A}$. For any function $f \in D$, we have

$$N\left[f(\mathbf{g}-\frac{\mathbf{e}_k}{N},\mathbf{h})-f(\mathbf{g},\mathbf{h})\right]+\frac{\partial f(\mathbf{g},\mathbf{h})}{\partial g_k}=-\frac{\gamma_{1,k}(g)}{N}\frac{\partial^2 f(\mathbf{g}-\gamma_{2,k}(g)\frac{\mathbf{e}_k}{N},\mathbf{h})}{\partial g_k^2},$$

$$N\left[f(\mathbf{g},\mathbf{h}-\frac{\mathbf{e}_k}{N})-f(\mathbf{g},\mathbf{h})\right]+\frac{\partial f(\mathbf{g},\mathbf{h})}{\partial h_k}=-\frac{\gamma_{1,k}(h)}{N}\frac{\partial^2 f(\mathbf{g},\mathbf{h}-\gamma_{2,k}(h)\frac{\mathbf{e}_k}{N})}{\partial g_k^2},$$

where $0<\gamma_{i,k}(g),\gamma_{i,k}(h)<1$ for $i=1,2$. Since

$$\left|\frac{\gamma_{1,k}(g)}{N}\frac{\partial^2 f(\mathbf{g}-\gamma_{2,k}(g)\frac{\mathbf{e}_k}{N},\mathbf{h})}{\partial g_k^2}\right|\le\frac{C}{N}$$

and

$$\left|\frac{\gamma_{1,k}(h)}{N}\frac{\partial^2 f(\mathbf{g},\mathbf{h}-\gamma_{2,k}(h)\frac{\mathbf{e}_k}{N})}{\partial g_k^2}\right|\le\frac{C}{N},$$

we obtain

$$\begin{aligned}|\mathbf{A}_N f(\mathbf{g},\mathbf{h})-f(\mathbf{g},\mathbf{h})|\le\frac{C}{N}&\left[\sum_{k=2}^{\infty}(g_{k-1}-g_k)L_k(g_{k-1},g_k;h_{k-1},h_k)\right.\\&+\sum_{k=1}^{\infty}(h_{k-1}-h_k)L_k(g_{k-1},g_k;h_{k-1},h_k)\\&\left.+(h_0-h_1)L_1(g_1;h_0,h_1)+\sum_{k=1}^{\infty}(g_k-g_{k+1})+2\theta\sum_{k=1}^{\infty}h_k\right]\\&\le\frac{C}{N}\left\{M_1\left[\sum_{l=2}^{\infty}(g_{l-1}-g_l)+\sum_{k=1}^{\infty}(h_{k-1}-h_k)\right]+g_1+2\theta\sum_{k=1}^{\infty}h_k\right\}\\&\le\frac{C}{N}\left[M_1(g_1+h_0)+g_1+2\theta\sum_{k=1}^{\infty}h_k\right].\end{aligned}$$

Note that C, M_1 and $\sum_{k=1}^{\infty}h_k$ are all finite, it is clear that as $N\to\infty$,

$$\lim_{N\to\infty}\sup_{(\mathbf{g},\mathbf{h})\in\Omega_N}|\mathbf{A}_N f(\mathbf{g},\mathbf{h})-\mathbf{A}f(\mathbf{g},\mathbf{h})|=0.$$

This completes the proof. □

Remark 3

(1) As discussed in Ethier and Kurtz (1986), there have been at least three basic techniques: operator semigroup, martingale and density dependent population process, for analyzing the weak approximation of the sequences of Markov processes. In fact, the three techniques have been applied to the mean-field limit

of supermarket models up to now, e.g., see the operator semigroup by Vvedenskaya et al. (1996), the density dependent population process by Mitzenmacher (1996) and the martingale by Turner (1996, 1998). In this section, we use the operator semigroup to provide a mean-field limit for the supermarket model with server multiple vacations, and show that the sequence of corresponding Markov processes asymptotically approaches a single trajectory identified by the unique and global solution to the infinite-dimensional system of mean-field limit equations.

(2) The mean-field limit over the finite time intervals has been generalized in the PhD thesis of Mitzenmacher (1996) to the infinite-dimensional case when the right-hand side of the system of differential equations is Lipschitz. Therefore, using the generalized results by Mitzenmacher (1996), we may significantly simplify some results in this section, and also we can possibly obtain a stronger form of convergence and some error bounds. Furthermore, readers may refer to Graham (2000a, b, 2004) and Luczak and McDiarmid (2006, 2007) for the longest queue length, the asymptotic independence and chaoticity on path space.

4 Existence and uniqueness

In this section, we first provide an effective technique to organize the Lipschitzian condition of the infinite-dimensional fraction vector function $F:\mathbf{R}_{+}^{\infty}\rightarrow\mathbf{C}^{1}(\mathbf{R}_{+}^{\infty})$. Note that the Lipschitzian condition for $d\geq 3$ is always difficult in the literature. Then we apply the Lipschitzian condition, together with the Picard approximation, to show that the limiting expected fraction vector is the unique and global solution to the system of differential equations.

For convenience of description, we write that $u_k=u_k(t,\mathbf{g},\mathbf{h})$ for $k\geq 1$ and $v_l=v_l(t,\mathbf{g},\mathbf{h})$ for $l\geq 0$. Using $v_0(t,\mathbf{g},\mathbf{h})+u_1(t,\mathbf{g},\mathbf{h})=1$, the system of differential equations (22) to (27) can be simplified as an initial value problem as follows

$$\frac{\mathrm{d}}{\mathrm{d}t}v_0=u_1-u_2-\theta v_1, \tag{35}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}v_1=\lambda\left(v_0-v_1\right)L_1\left(u_1;v_0,v_1;t\right)-\theta v_1, \tag{36}$$

for $k\geq 2$,

$$\frac{\mathrm{d}}{\mathrm{d}t}v_k=\lambda\left(v_{k-1}-v_k\right)L_k\left(u_{k-1},u_k;v_{k-1},v_k;t\right)-\theta v_k \tag{37}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}u_k=\lambda\left(u_{k-1}-u_k\right)L_k\left(u_{k-1},u_k;v_{k-1},v_k;t\right)+\left(u_{k+1}-u_k\right)+\theta v_k, \tag{38}$$

with the boundary condition

$$v_0+u_1=1 \tag{39}$$

and with the initial condition

$$\begin{cases} u_k(0)=g_k, & k\geq 1,\\ v_l(0)=h_l, & l\geq 0.\end{cases} \tag{40}$$

4.1 A Lipschitzian condition

To establish the Lipschitzian condition, we need to use the derivative of the infinite-dimensional function $G : \mathbf{R}_+^\infty \to \mathbf{C}^1(\mathbf{R}_+^\infty)$. Thus it is necessary to provide some useful notation and definitions of derivatives as follows.

For the infinite-dimensional function $G : \mathbf{R}_+^\infty \to \mathbf{C}^1(\mathbf{R}_+^\infty)$, we write $x = (x_1, x_2, x_3, \ldots)$ and $G(x) = (G_1(x), G_2(x), G_3(x), \ldots)$, 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

$$DG(x) = \begin{pmatrix} \frac{\partial G_1(x)}{\partial x_1} & \frac{\partial G_2(x)}{\partial x_1} & \frac{\partial G_3(x)}{\partial x_1} & \cdots \\ \frac{\partial G_1(x)}{\partial x_2} & \frac{\partial G_2(x)}{\partial x_2} & \frac{\partial G_3(x)}{\partial x_2} & \cdots \\ \frac{\partial G_1(x)}{\partial x_3} & \frac{\partial G_2(x)}{\partial x_3} & \frac{\partial G_3(x)}{\partial x_3} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}. \tag{41}$$

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, e.g., see Chapter 1 of Taylor (1996) and Chapter 3 of Fleming (1977) for more details.

Definition 3 Let the infinite-dimensional function $G : \mathbf{R}_+^\infty \to \mathbf{C}^1(\mathbf{R}_+^\infty)$.

(1) If there exists a linear operator $A : \mathbf{R}_+^\infty \to \mathbf{R}_+^\infty$ such that for any vector $h \in \mathbf{R}^\infty$ and a scalar $t \in \mathbf{R}$

$$\lim_{t \to 0} \frac{||G(x + ht) - G(x) - hAt||}{t} = 0,$$

then the function $G(x)$ is called to be Gateaux differentiable at $x \in \mathbf{R}_+^\infty$. In this case, we write the Gateaux derivative $G'_G(x) = A$.

(2) If there exists a linear operator $B : \mathbf{R}_+^\infty \to \mathbf{R}_+^\infty$ such that for any vector $h \in \mathbf{R}^\infty$

$$\lim_{||h|| \to 0} \frac{||G(x + h) - G(x) - hB||}{||h||} = 0,$$

then the function $G(x)$ is called to be Frechet differentiable at $x \in \mathbf{R}_+^\infty$. In this case, we write the Frechet derivative $G'_F(x) = B$.

It is easy to check that if the infinite-dimensional function $G(x)$ is Frechet differentiable, then it is also Gateaux differentiable. At the same time, we have

$$G'_G(x) = G'_F(x) = DG(x). \tag{42}$$

Let $\boldsymbol{t} = (t_1, t_2, t_3, \ldots)$ with $0 \le t_k \le 1$ for $k \ge 1$. Then we write

$$DG(x + \boldsymbol{t} \oslash (y - x))$$
$$= \begin{pmatrix} \frac{\partial G_1(x + t_1 (y-x))}{\partial x_1} & \frac{\partial G_2(x + t_2 (y-x))}{\partial x_1} & \frac{\partial G_3(x + t_3 (y-x))}{\partial x_1} & \cdots \\ \frac{\partial G_1(x + t_1 (y-x))}{\partial x_2} & \frac{\partial G_2(x + t_2 (y-x))}{\partial x_2} & \frac{\partial G_3(x + t_3 (y-x))}{\partial x_2} & \cdots \\ \frac{\partial G_1(x + t_1 (y-x))}{\partial x_3} & \frac{\partial G_2(x + t_2 (y-x))}{\partial x_3} & \frac{\partial G_3(x + t_3 (y-x))}{\partial x_3} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

The following lemma provides two useful properties for the Gateaux derivatives of the infinite-dimensional functions. Obviously, the two useful properties also hold for the Frechet derivatives.

Lemma 5 *If the infinite-dimensional function* $G : \mathbf{R}_+^\infty \to \mathbf{C}^1(\mathbf{R}_+^\infty)$ *is Gateaux differentiable, then there exists a vector* $\boldsymbol{t} = (t_1, t_2, t_3, \ldots)$ *with* $0 \le t_k \le 1$ *for* $k \ge 1$ *such that*

$$G(y) - G(x) = (y - x)\, DG(x + \boldsymbol{t} \oslash (y - x)). \tag{43}$$

Furthermore, we have

$$||G(y) - G(x)|| \le \sup_{0 \le t \le 1} ||DG(x + t(y - x))|| \, ||y - x||. \tag{44}$$

Proof For the function $G_k(x)$, it is easy to check that there exists a number $t_k \in [0, 1]$ such that

$$\begin{aligned} G_k(y) - G_k(x) &= \sum_{i=1}^{\infty} (y_i - x_i) \frac{\partial G_k(x + t_k (y - x))}{\partial x_i} \\ &= (y - x) \left(\frac{\partial G_k(x + t_k (y - x))}{\partial x_1}, \frac{\partial G_k(x + t_k (y - x))}{\partial x_2}, \ldots \right)^T. \end{aligned}$$

Note that

$$G(y) - G(x) = (G_1(y) - G_1(x), G_2(y) - G_2(x), G_3(y) - G_3(x), \ldots),$$

we obtain

$$G(y) - G(x) = (y - x)\, DG(x + \boldsymbol{t} \oslash (y - x)).$$

Since

$$||DG(x + \boldsymbol{t} \oslash (y - x))|| \le \sup_{0 \le t \le 1} ||DG(x + t(y - x))||,$$

it follows

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

This completes the proof. □

Note that $v_0 + u_1 = 1$, we set that $x = (u_1, v_1; u_2, v_2; u_3, v_3; u_4, v_4; \ldots)$ and $F(x) = (F_1(x), F_2(x), F_3(x), \ldots)$, where

$$F_1(x) = (u_2 - u_1 + \theta v_1, \lambda(v_0 - v_1) L_1(u_1; v_0, v_1; t) - \theta v_1) \tag{45}$$

and $k \geq 2$

$$\begin{aligned} F_k(x) = (&\lambda(u_{k-1} - u_k) L_k(u_{k-1}, u_k; v_{k-1}, v_k; t) + (u_{k+1} - u_k) + \theta v_k, \\ &\lambda(v_{k-1} - v_k) L_k(u_{k-1}, u_k; v_{k-1}, v_k; t) - \theta v_k). \end{aligned} \tag{46}$$

Then $F(x)$ is in $\mathbf{C}^2(\mathbf{R}_+^\infty)$, and the system of differential vector equations (35) to (40) is rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t} x = F(x) \tag{47}$$

with initial condition

$$x(0) = (g_1, h_1; g_2, h_2; g_3, h_3; g_4, h_4; \ldots). \tag{48}$$

In what follows we show that the infinite-dimensional function $F(x)$ is Lipschitzian for $t \geq 0$. From (1) of Definition 3 and Eq. 42, the matrix of partial derivatives of the function $F(x)$ is given by

$$DF(x) = \begin{pmatrix} A_1(x) & B_2(x) & & & \\ C_1(x) & A_2(x) & B_3(x) & & \\ & C_2(x) & A_3(x) & B_4(x) & \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{49}$$

where $A_k(x)$ is a matrix of size 2 for $k \geq 1$, and the sizes of all the others can be determined accordingly. Let

$$L_1(t) = L_1(u_1; v_0, v_1; t) = V_1(u_1; v_0, v_1; t)$$

and for $k \geq 2$

$$L_k(t) = W_k(u_{k-1}, u_k; v_{k-1}, v_k; t) = V_k(u_{k-1}, u_k; v_{k-1}, v_k; t).$$

Then

$$A_1(x) = \begin{pmatrix} -1 & \lambda(v_0 - v_1) \dfrac{\partial L_1(t)}{\partial u_1} \\ \theta & -\lambda L_1(t) + \lambda(v_0 - v_1) \dfrac{\partial L_1(t)}{\partial v_1} - \theta \end{pmatrix},$$

for $k \geq 2$

$$A_k(x) = \begin{pmatrix} -\lambda L_k(t) + \lambda(u_{k-1} - u_k) \dfrac{\partial L_k(t)}{\partial u_k} - 1 & \lambda(v_{k-1} - v_k) \dfrac{\partial L_k(t)}{\partial u_k} \\ \lambda(u_{k-1} - u_k) \dfrac{\partial L_k(t)}{\partial v_k} + \theta & -\lambda V_k(t) + \lambda(v_{k-1} - v_k) \dfrac{\partial L_k(t)}{\partial v_k} - \theta \end{pmatrix},$$

$$B_k(x) = \begin{pmatrix} \lambda L_k(t) + \lambda(u_{k-1} - u_k) \dfrac{\partial L_k(t)}{\partial u_{k-1}} & \lambda(v_{k-1} - v_k) \dfrac{\partial L_k(t)}{\partial u_{k-1}} \\ \lambda(u_{k-1} - u_k) \dfrac{\partial L_k(t)}{\partial v_{k-1}} & \lambda V_k(t) + \lambda(v_{k-1} - v_k) \dfrac{\partial L_k(t)}{\partial v_{k-1}} \end{pmatrix},$$

$$C_j(x) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad j \geq 1,$$

where

$$\frac{\partial L_1(t)}{\partial u_1} = \sum_{m=1}^{d} C_d^m (d-m) [v_0(t) - v_1(t)]^{m-1} [v_1(t) + u_1(t)]^{d-m-1},$$

$$\frac{\partial L_1(t)}{\partial v_1} = \sum_{m=1}^{d} C_d^m (1-m) [v_0(t) - v_1(t)]^{m-2} [v_1(t) + u_1(t)]^{d-m} + \sum_{m=1}^{d} C_d^m (d-m) [v_0(t) - v_1(t)]^{m-1} [v_1(t) + u_1(t)]^{d-m-1},$$

and for $k \geq 2$

$$\frac{\partial L_k(t)}{\partial u_k} = \sum_{m=1}^{d} C_d^m (1-m) \left[u_{k-1}(t) - u_k(t) + v_{k-1}(t) - v_k(t)\right]^{m-2} [u_k(t) + v_k(t)]^{d-m} + \sum_{m=1}^{d} C_d^m (d-m) \left[u_{k-1}(t) - u_k(t) + v_{k-1}(t) - v_k(t)\right]^{m-1} [u_k(t) + v_k(t)]^{d-m-1},$$

$$\frac{\partial L_k(t)}{\partial v_k} = \sum_{m=1}^{d} C_d^m (1-m) \left[u_{k-1}(t) - u_k(t) + v_{k-1}(t) - v_k(t)\right]^{m-2} [u_k(t) + v_k(t)]^{d-m} + \sum_{m=1}^{d} C_d^m (d-m) \left[u_{k-1}(t) - u_k(t) + v_{k-1}(t) - v_k(t)\right]^{m-1} [u_k(t) + v_k(t)]^{d-m-1},$$

$$\frac{\partial L_k(t)}{\partial u_{k-1}} = \sum_{m=1}^{d} C_d^m (m-1) \left[u_{k-1}(t) - u_k(t) + v_{k-1}(t) - v_k(t)\right]^{m-2} [u_k(t) + v_k(t)]^{d-m},$$

$$\frac{\partial L_k(t)}{\partial v_{k-1}} = \sum_{m=1}^{d} C_d^m (m-1) \left[u_{k-1}(t) - u_k(t) + v_{k-1}(t) - v_k(t)\right]^{m-2} [u_k(t) + v_k(t)]^{d-m}.$$

Hence it follows from Eq. 49 that

$$\| DF(x) \| = \max \left\{ \left\| e^{\mathrm{T}} \left[A_1(x) + B_2(x) \right] \right\|, \sup_{k \geq 2} \left\| e^{\mathrm{T}} \left[C_{k-1}(x) + A_k(x) + B_{k+1}(x) \right] \right\| \right\}. \tag{50}$$

Let

$$M_5 = \sum_{m=1}^{d} C_d^m (d-m).$$

Based on the expression $L_1(u_1; v_0, v_1; t)$ and $L_k(u_{k-1}, u_k; v_{k-1}, v_k; t)$ for $k \geq 2$, it is easy to check that

$$L_1(u_1; v_0, v_1; t) \leq M_1$$

and for $k \geq 2$

$$L_k(u_{k-1}, u_k; v_{k-1}, v_k; t) \leq M_1.$$

At the same time, we have

$$\left| \frac{\partial L_1(t)}{\partial u_1} + \frac{\partial L_1(t)}{\partial v_1} \right| \leq M_5,$$

and for $k \geq 2$

$$\left| \frac{\partial L_k(t)}{\partial u_k} + \frac{\partial L_k(t)}{\partial u_{k-1}} \right| \leq M_5,$$
$$\left| \frac{\partial L_k(t)}{\partial v_k} + \frac{\partial L_k(t)}{\partial v_{k-1}} \right| \leq M_5.$$

Note that

$$e^{\mathrm{T}}\left[A_1(x) + C_1(x)\right] = \left(\theta, -\lambda V_1(t) + \lambda(v_0 - v_1)\left[\frac{\partial L_1(t)}{\partial u_1} + \frac{\partial L_1(t)}{\partial v_1}\right] - \theta\right),$$

we obtain

$$\left\| e^{\mathrm{T}}\left[A_1(x) + B_2(x)\right] \right\| \leq \lambda M_1 + 2\lambda M_5 + \theta. \tag{51}$$

Since for $k \geq 2$

$$\begin{aligned} & e^{\mathrm{T}}\left[A_k(x) + B_k(x) + C_k(x)\right] \\ & = \left(\lambda(u_{k-1} - u_k)\left[\frac{\partial L_k(t)}{\partial u_{k-1}} + \frac{\partial L_k(t)}{\partial u_k} + \frac{\partial L_k(t)}{\partial v_{k-1}} + \frac{\partial L_k(t)}{\partial v_k}\right] + \theta, \right. \\ & \qquad \left. \lambda(v_{k-1} - v_k)\left[\frac{\partial L_k(t)}{\partial u_{k-1}} + \frac{\partial L_k(t)}{\partial u_k} + \frac{\partial L_k(t)}{\partial v_{k-1}} + \frac{\partial L_k(t)}{\partial v_k}\right] - \theta\right), \end{aligned}$$

we obtain

$$\left\| e^{\mathrm{T}}\left[A_k(x) + B_k(x) + C_k(x)\right] \right\| \leq 4\lambda M_5 + \theta. \tag{52}$$

Then it follows form Eqs. 50, 51 and 52 that

$$\|DF(x)\| \leq M,$$

where

$$M = \max\left(\lambda M_1 + 2\lambda M_5 + \theta, 4\lambda M_5 + \theta\right).$$

Note that for $(\mathbf{u}, \mathbf{v}), (\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}) \in \widetilde{\Omega}$, it follows from Eq. 43 that

$$F(\mathbf{u}, \mathbf{v}) - F(\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}) = [(\mathbf{u}, \mathbf{v}) - (\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}})]\, DF((\mathbf{u}, \mathbf{v}) + \boldsymbol{t} \oslash [(\mathbf{u}, \mathbf{v}) - (\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}})])$$

and from Eq. 44 that

$$\begin{aligned} \| F(\mathbf{u}, \mathbf{v}) - F(\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}) \| &\leq \sup_{0 \leq t \leq 1} || DF((\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}) + t [(\mathbf{u}, \mathbf{v}) - (\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}})]) |||| (\mathbf{u}, \mathbf{v}) - (\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}) || \\ &\leq M \| (\mathbf{u}, \mathbf{v}) - (\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}) \| . \end{aligned} \tag{53}$$

Therefore, the function $F(\mathbf{u}, \mathbf{v})$ is Lipschitzian for $(\mathbf{u}, \mathbf{v}) \in \widetilde{\Omega}$.

Remark 4 Let $G : \mathbf{R}^m \to \mathbf{R}^n$ be continuously differentiable. Then Proposition 4.5 or Proposition 4.4 in Fleming (1977) proved that the function $G(\mathbf{t})$ is locally Lipschitzian for $\mathbf{t} \in \Lambda$. Note that Eq. 53 extends the Lipschitzian condition to the infinite-dimensional continuously differentiable function $F : \mathbf{R}^\infty \to \mathbf{R}^\infty$, and such a generalization is always necessary in the study of supermarket models.

Remark 5 The Lipschitzian condition for $d \geq 3$ is always difficult in the literature, while few available results were organized in the supermarket models with $d = 2$, e.g., see Vvedenskaya and Suhov (1997, 2005) and Mitzenmacher et al. (2001). Therefore, here we provide a unique and effective method to compute the Lipschitzian condition for more complicated supermarket models with $d \geq 1$.

4.2 The Picard approximation

In this subsection, we apply the Lipschitzian condition, together with the Picard approximation, to show that the limiting expected fraction vector is the unique and global solution to the system of differential equations.

Let $v_0 + u_1 = 1$ and $x = (u_1, v_1; u_2, v_2; u_3, v_3; u_4, v_4; \ldots)$. We write

$$\widetilde{\Omega}_0 = \{x : 1 \geq u_1 \geq u_2 \geq \cdots \geq 0, 1 \geq v_1 \geq v_2 \geq v_3 \geq \cdots \geq 0\}$$

It follows from Eqs. 47 to 48 that for $x \in \widetilde{\Omega}_0$

$$x(t) = x(0) + \int_0^t F(x(s)) \mathrm{d}s,$$

this gives

$$x(t) = \widetilde{(\mathbf{g}, \mathbf{h})} + \int_0^t F(x(s)) \mathrm{d}s, \tag{54}$$

where

$$\widetilde{(\mathbf{g}, \mathbf{h})} = (g_1, h_1; g_2, h_2; g_3, h_3; g_4, h_4; \ldots).$$

Based on the integral equation (54), the following theorem indicates that there exists the unique and global solution to the system of differential equations (35) to (40) for $t \geq 0$.

Theorem 3 *For $\widetilde{(\mathbf{g}, \mathbf{h})} \in \widetilde{\Omega}_0$, there exists the unique and global solution to the Eq. 54 for $t \geq 0$.*

Proof We take the Picard sequence as follows

$$x^{\langle 0\rangle}(t) \equiv 0,$$

and for $n \geq 1$

$$x^{\langle n\rangle}(t) = \widetilde{(\mathbf{g}, \mathbf{h})} + \int_0^t F\left(x^{\langle n-1\rangle}(s)\right) \mathrm{d}s.$$

It follows from Eq. 53 that

$$\begin{aligned}\left\|x^{\langle n+1\rangle}(t) - x^{\langle n\rangle}(t)\right\| &\leq \int_0^t \left\|F\left(x^{\langle n\rangle}(s)\right) - F\left(x^{\langle n-1\rangle}(s)\right)\right\| \mathrm{d}s \\ &\leq Mt\left\|x^{\langle n\rangle}(t) - x^{\langle n-1\rangle}(t)\right\| \leq \cdots \\ &\leq \frac{(Mt)^{n-1}}{(n-1)!}\left\|x^{\langle 2\rangle}(t) - x^{\langle 1\rangle}(t)\right\|.\end{aligned}$$

From the boundary condition: $\left\|x^{\langle 2\rangle}(t) - x^{\langle 1\rangle}(t)\right\| \leq 1$, it is clear that if $0 \leq t \leq 1/M$, then $\lim_{n\to\infty}(Mt)^{n-1}/(n-1)! = 0$, which leads to that as $n \to \infty$, the Picard sequence $\{x^{\langle n\rangle}(t)\}$ is uniformly convergent for $t \in [0, 1/M]$.

Let $x(t) = \lim_{n\to\infty} x^{\langle n\rangle}(t)$ for $t \in [0, 1/M]$. Then $x(t)$ is a solution to Eq. 54 for $t \in [0, 1/M]$.

Let $y(t)$ is another solution to Eq. 54 for $t \in [0, 1/M]$. Then it is easy to check that

$$\|x(t) - y(t)\| \leq \frac{(Mt)^{n-1}}{(n-1)!},$$

this gives that for $t \in [0, 1/M]$,

$$y(t) = \lim_{n\to\infty} x^{\langle n\rangle}(t) = x(t).$$

This shows that $x(t)$ is the unique solution to Eq. 54 for $t \in [0, 1/M]$.

We consider the following equation

$$x(t) = x\left(\frac{1}{M}\right) + \int_{\frac{1}{M}}^t F(x(s))\mathrm{d}s.$$

Take the Picard sequence

$$x^{\langle 0\rangle}(t) \equiv 0,$$

and for $n \geq 1$

$$x^{\langle n\rangle}(t) = x\left(\frac{1}{M}\right) + \int_{\frac{1}{M}}^t F\left(x^{\langle n-1\rangle}(s)\right) \mathrm{d}s.$$

It is easy to show that $x(t) = \lim_{n\to\infty} x^{\langle n\rangle}(t)$ is the unique solution to Eq. 54 for $t \in [1/M, 2/M]$.

We assume that for $l = k$, $x(t)$ is the unique solution to Eq. 54 for $t \in [k/M, (k+1)/M]$. Then for $l = k+1$, we consider the following equation

$$x(t) = x\left(\frac{k+1}{M}\right) + \int_{\frac{k+1}{M}}^t F(x(s))\mathrm{d}s.$$

Take the Picard sequence

$$x^{\langle 0\rangle}(t) \equiv 0$$

and for $n \geq 1$

$$x^{\langle n\rangle}(t) = x(\frac{k+1}{M}) + \int_{\frac{k+1}{M}}^{t} F(x^{\langle n-1\rangle}(s))\mathrm{d}s.$$

It is easy to indicate that $x(t) = \lim_{n\to\infty} x^{\langle n\rangle}(t)$ is the unique solution to Eq. 54 for $t \in [(k+1)/M, (k+2)/M]$.

By induction, it is easy to show that $x(t) = \lim_{n\to\infty} x^{\langle n\rangle}(t)$ is the unique solution to Eq. 54 for $t \in [l/M, (l+1)/M]$ for $l = 0, 1, 2, \ldots$. Note that

$$[0, +\infty) = \left[0, \frac{1}{M}\right] \cup \left[\frac{1}{M}, \frac{2}{M}\right] \cup \left[\frac{2}{M}, \frac{3}{M}\right] \cup \cdots,$$

thus, $x(t)$ is the unique and global solution to Eq. 54 for $t \geq 0$. This completes the proof. □

Remark 6 Comparing with the finite-dimensional system of integral equations (e.g., see Chapter 1 of Hale (1980)), Theorem 3 makes some necessary generalization of the Picard approximation in order to deal with existence and uniqueness of solution to the infinite-dimensional system of integral equations. Note that such a generalization is always necessary in the study of supermarket models.

Remark 7 Note that the infinite-dimensional system of differential equations for the supermarket model with server multiple vacations is defined on a Banach space, the existence and uniqueness of solution is immediately obtained due to the existing results on Banach spaces given that the right-hand side of the infinite-dimensional system of differential equations is Lipschitzian.

5 The fixed point

In this section, we analyze the fixed point of the infinite-dimensional system of differential equations (22) to (27), and set up an infinite-dimensional system of nonlinear equations satisfied by the fixed point. Based on this, we provide an effective algorithm for computing the fixed point. Note that the fixed point is a key in performance analysis of this supermarket model.

Let $\pi = (\pi_1^{(W)}, \pi_2^{(W)}, \pi_3^{(W)}, \ldots; \pi_0^{(V)}, \pi_1^{(V)}, \pi_2^{(V)}, \pi_3^{(V)}, \ldots)$. The row vector π is called a fixed point of the infinite-dimensional system of differential equations (22) to (27) if $\pi = \lim_{t\to+\infty} (\mathbf{u}(t), \mathbf{v}(t))$, where $\pi_k^{(W)} = \lim_{t\to+\infty} u_k(t)$ for $k \geq 1$ and $\pi_l^{(V)} = \lim_{t\to+\infty} v_l(t)$ for $l \geq 0$. It is well-known that if π is a fixed point of the limiting expected fraction vector $(\mathbf{u}(t), \mathbf{v}(t))$, then

$$\lim_{t\to+\infty} \left[\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}(t)\right] = 0, \quad \lim_{t\to+\infty} \left[\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v}(t)\right] = 0. \qquad (55)$$

For two sequences of positive numbers $\{x_k : k \geq 1\}$ and $\{y_l : l \geq 0\}$ with $x_1 + y_0 = 1$, we introduce two function notations

$$L_1 (x_1; y_0, y_1) = \sum_{m=1}^{d} C_d^m (y_0 - y_1)^{m-1} (x_1 + y_1)^{d-m},$$

for $k \geq 2$

$$L_k (x_{k-1}, x_k; y_{k-1}, y_k) = \sum_{m=1}^{d} C_d^m (x_{k-1} - x_k + y_{k-1} - y_k)^{m-1} (x_k + y_k)^{d-m}.$$

Taking $t \to +\infty$ in both sides of the system of differential equations (22) to (27), we obtain

$$\pi_2^{(W)} - \pi_1^{(W)} + \theta \pi_1^{(V)} = 0, \tag{56}$$

$$\lambda \left[\pi_0^{(V)} - \pi_1^{(V)}\right] L_1 \left(\pi_1^{(W)}; \pi_0^{(V)}, \pi_1^{(V)}\right) - \theta \pi_1^{(V)} = 0, \tag{57}$$

for $k \geq 2$,

$$\lambda \left[\pi_{k-1}^{(W)} - \pi_k^{(W)}\right] L_k \left(\pi_{k-1}^{(W)}, \pi_k^{(W)}; \pi_{k-1}^{(V)}, \pi_k^{(V)}\right) + \pi_{k+1}^{(W)} - \pi_k^{(W)} + \theta \pi_k^{(V)} = 0 \tag{58}$$

and

$$\lambda \left[\pi_{k-1}^{(V)} - \pi_k^{(V)}\right] L_k \left(\pi_{k-1}^{(W)}, \pi_k^{(W)}; \pi_{k-1}^{(V)}, \pi_k^{(V)}\right) - \theta \pi_k^{(V)} = 0, \tag{59}$$

with the boundary condition

$$\pi_0^{(V)} + \pi_1^{(W)} = 1. \tag{60}$$

To solve the system of nonlinear equations, it is a key to first compute the boundary probabilities $\pi_0^{(V)}$ and $\pi_1^{(W)}$ in the fixed point. The following theorem provide an effective method to determine the boundary probabilities.

Theorem 4 *If $\lambda < 1$, then the boundary probabilities $\pi_0^{(V)}$ and $\pi_1^{(W)}$ are given by*

$$\pi_0^{(V)} = 1 - \lambda$$

and

$$\pi_1^{(W)} = \lambda.$$

Proof For $d \geq 1$, we have

$$\begin{aligned}\left[\pi_0^{(V)} - \pi_1^{(V)}\right] L_1 \left(\pi_1^{(W)}; \pi_0^{(V)}, \pi_1^{(V)}\right) &= \sum_{m=1}^{d} C_d^m \left[\pi_0^{(V)} - \pi_1^{(V)}\right]^m \left[\pi_1^{(V)} + \pi_1^{(W)}\right]^{d-m} \\ &= \left[\pi_0^{(V)} + \pi_1^{(W)}\right]^d - \left[\pi_1^{(V)} + \pi_1^{(W)}\right]^d,\end{aligned}$$

$$\left[\pi_{k-1}^{(V)}-\pi_k^{(V)}\right] L_k\left(\pi_{k-1}^{(W)},\pi_k^{(W)};\pi_{k-1}^{(V)},\pi_k^{(V)}\right) = \left[\pi_{k-1}^{(V)}-\pi_k^{(V)}\right]\sum_{m=1}^{d} C_d^m\left[\pi_k^{(V)}+\pi_k^{(W)}\right]^{d-m}$$
$$\times\left[\pi_{k-1}^{(V)}-\pi_k^{(V)}+\pi_{k-1}^{(W)}-\pi_k^{(W)}\right]^{m-1}$$
$$=\frac{\pi_{k-1}^{(V)}-\pi_k^{(V)}}{\pi_{k-1}^{(V)}-\pi_k^{(V)}+\pi_{k-1}^{(W)}-\pi_k^{(W)}}$$
$$\times\left\{\left[\pi_{k-1}^{(V)}+\pi_{k-1}^{(W)}\right]^d-\left[\pi_k^{(V)}+\pi_k^{(W)}\right]^d\right\}$$

and

$$\left[\pi_{k-1}^{(W)}-\pi_k^{(W)}\right] L_k\left(\pi_{k-1}^{(W)},\pi_k^{(W)};\pi_{k-1}^{(V)},\pi_k^{(V)}\right)$$
$$=\frac{\pi_{k-1}^{(W)}-\pi_k^{(W)}}{\pi_{k-1}^{(V)}-\pi_k^{(V)}+\pi_{k-1}^{(W)}-\pi_k^{(W)}}\left\{\left[\pi_{k-1}^{(V)}+\pi_{k-1}^{(W)}\right]^d-\left[\pi_k^{(V)}+\pi_k^{(W)}\right]^d\right\},$$

this gives

$$\left[\pi_0^{(V)}-\pi_1^{(V)}\right] L_1\left(\pi_1^{(W)};\pi_0^{(V)},\pi_1^{(V)}\right)+\sum_{k=2}^{\infty}\left[\pi_{k-1}^{(V)}-\pi_k^{(V)}\right] L_k\left(\pi_{k-1}^{(W)},\pi_k^{(W)};\pi_{k-1}^{(V)},\pi_k^{(V)}\right)$$
$$+\sum_{k=2}^{\infty}\left[\pi_{k-1}^{(W)}-\pi_k^{(W)}\right] L_k\left(\pi_{k-1}^{(W)},\pi_k^{(W)};\pi_{k-1}^{(V)},\pi_k^{(V)}\right)$$
$$=\left[\pi_0^{(V)}+\pi_1^{(W)}\right]^d-\left[\pi_1^{(V)}+\pi_1^{(W)}\right]^d+\sum_{k=2}^{\infty}\left\{\left[\pi_{k-1}^{(V)}+\pi_{k-1}^{(W)}\right]^d-\left[\pi_k^{(V)}+\pi_k^{(W)}\right]^d\right\}$$
$$=\left[\pi_0^{(V)}+\pi_1^{(W)}\right]^d=1,$$

Together with $\pi_0^{(V)}+\pi_1^{(W)}=1$, it follows from Eqs. 56 to 60 that

$$\pi_1^{(W)}=\lambda\left\{\sum_{k=2}^{\infty}\left[\pi_{k-1}^{(W)}-\pi_k^{(W)}\right] L_k\left(\pi_{k-1}^{(W)},\pi_k^{(W)};\pi_{k-1}^{(V)},\pi_k^{(V)}\right)\right.$$
$$+\left[\pi_0^{(V)}-\pi_1^{(V)}\right] L_1\left(\pi_1^{(W)};\pi_0^{(V)},\pi_1^{(V)}\right)$$
$$\left.+\sum_{k=2}^{\infty}\left[\pi_{k-1}^{(V)}-\pi_k^{(V)}\right] L_k\left(\pi_{k-1}^{(W)},\pi_k^{(W)};\pi_{k-1}^{(V)},\pi_k^{(V)}\right)\right\}=\lambda. \tag{61}$$

Thus, using $\pi_0^{(V)}+\pi_1^{(W)}=1$ we obtain

$$\pi_0^{(V)}=1-\lambda.$$

This completes the proof. □

Remark 8 Note that

$$\sum_{k=2}^{\infty}\left[\pi_{k-1}^{(W)}-\pi_{k}^{(W)}\right]L_k\left(\pi_{k-1}^{(W)},\pi_{k}^{(W)};\pi_{k-1}^{(V)},\pi_{k}^{(V)}\right)+\left[\pi_{0}^{(V)}-\pi_{1}^{(V)}\right]L_1\left(\pi_{1}^{(W)};\pi_{0}^{(V)},\pi_{1}^{(V)}\right)$$
$$+\sum_{k=2}^{\infty}\left[\pi_{k-1}^{(V)}-\pi_{k}^{(V)}\right]L_k\left(\pi_{k-1}^{(W)},\pi_{k}^{(W)};\pi_{k-1}^{(V)},\pi_{k}^{(V)}\right)=1,$$

the sequence

$$\left\{\left[\pi_{k-1}^{(W)}-\pi_{k}^{(W)}\right]L_k\left(\pi_{k-1}^{(W)},\pi_{k}^{(W)};\pi_{k-1}^{(V)},\pi_{k}^{(V)}\right);\left[\pi_{0}^{(V)}-\pi_{1}^{(V)}\right]L_1\left(\pi_{1}^{(W)};\pi_{0}^{(V)},\pi_{1}^{(V)}\right),\right.$$
$$\left.\left[\pi_{k-1}^{(V)}-\pi_{k}^{(V)}\right]L_k\left(\pi_{k-1}^{(W)},\pi_{k}^{(W)};\pi_{k-1}^{(V)},\pi_{k}^{(V)}\right):k\geq 2\right\}$$

is a probability vector.

Now, we provide an iterative algorithm for computing the fixed point of the infinite-dimensional system of differential equations (56) to (60).

We define

$$F_k(x)=\theta x-\lambda\left(\xi_{k-1}-x\right)L_k\left(\delta_{k-1},\delta_k;\xi_{k-1},x\right),\quad k\geq 2. \tag{62}$$

We assume that $\xi_0=1-\lambda$ and $\delta_1=\lambda$. Using Eq. 60, we take that $\pi_0^{(V)}=\xi_0$ and $\pi_1^{(W)}=\delta_1$. We denote by ξ_1 a solution in $(0,\xi_0)$ to the nonlinear equation

$$F_1(x)=\theta x-\lambda\left(\xi_0-x\right)L_1\left(\delta_1;\xi_0,x\right)=0, \tag{63}$$

and set

$$\delta_2=\delta_1-\theta\xi_1. \tag{64}$$

Let ξ_2 be a solution in $(0,\xi_1)$ to the nonlinear equation

$$F_2(x)=\theta x-\lambda\left(\xi_1-x\right)L_2\left(\delta_1,\delta_2;\xi_1,x\right)=0, \tag{65}$$

and set

$$\delta_3=\delta_2-\theta\xi_2-\lambda\left(\delta_1-\delta_2\right)L_2\left(\delta_1,\delta_2;\xi_1,\xi_2\right). \tag{66}$$

We assume that the k pairs $(\xi_0,\delta_1),(\xi_1,\delta_2),\ldots,(\xi_{k-1},\delta_k)$ have been obtained iteratively. Then we denote by ξ_k a solution in $(0,\xi_{k-1})$ to the nonlinear equation $F_k(x)=0$, and set

$$\delta_{k+1}=\delta_k-\theta\xi_k-\lambda\left(\delta_{k-1}-\delta_k\right)L_k\left(\delta_{k-1},\delta_k;\xi_{k-1},\xi_k\right). \tag{67}$$

It is clear that $0<\xi_k<\xi_{k-1}<\cdots<\xi_1<\xi_0=1-\lambda$ and $0<\delta_{k+1}<\delta_k<\cdots<\delta_2<\delta_1=\lambda$.

The following theorem expresses the fixed point of the infinite-dimensional system of differential equations (56) to (60) by means of the iterative algorithm. Note that it is a key in the proof that we need to indicate the uniqueness of the sequences $\{(\xi_{k-1},\delta_k):k\geq 1\}$.

Theorem 5 *If* $\lambda < 1$, *then the fixed point* $\pi = (\pi_1^{(W)}, \pi_2^{(W)}, \pi_3^{(W)}, \ldots; \pi_0^{(V)}, \pi_1^{(V)}, \pi_2^{(V)}, \ldots)$ *is uniquely given by*

$$\pi_l^{(V)} = \xi_l, \quad l \geq 0,$$

and

$$\pi_k^{(W)} = \delta_k, \quad k \geq 1.$$

Proof It is obvious that $\pi_0^{(V)} = \xi_0 = 1 - \lambda$ and $\pi_1^{(W)} = \delta_1 = \lambda$.

It follows from Eq. 57 that

$$\lambda \left[\pi_0^{(V)} - \pi_1^{(V)} \right] L_1 \left(\pi_1^{(W)}; \pi_0^{(V)}, \pi_1^{(V)} \right) - \theta \pi_1^{(V)} = 0,$$

that is

$$\lambda \sum_{m=1}^{d} C_d^m \left[1 - \lambda - \pi_1^{(V)} \right]^m \left(\pi_1^{(V)} + \lambda \right)^{d-m} - \theta \pi_1^{(V)} = 0.$$

Let

$$\begin{aligned} F_1(x) &= \theta x - \lambda \sum_{m=1}^{d} C_d^m \left[1 - \lambda - x \right]^m (x + \lambda)^{d-m} \\ &= \theta x - \lambda + \lambda \left[x + \lambda \right]^d. \end{aligned}$$

Then $F_1(0) = -\lambda + \lambda^{d+1} < 0$, $F_1(\xi_0) = \theta \xi_0 > 0$ and for $x \in (0, \xi_0)$

$$\frac{\mathrm{d}}{\mathrm{d}x} F_1(x) = \theta + d\lambda \left[x + \lambda \right]^{d-1} > 0.$$

Note that $F_1(x)$ is a continuous function for $x \in (0, \xi_0)$, thus there exists a unique positive solution $x = \xi_1$ to the nonlinear equation $F_1(x) = 0$ for $x \in (0, \xi_0)$. Hence, $\pi_1^{(V)} = \xi_1$. It follows from Eq. 56 that

$$\pi_2^{(W)} = \pi_1^{(W)} - \theta \pi_1^{(V)} = \delta_1 - \theta \xi_1 = \delta_2.$$

It follows from Eq. 59 for $k = 2$ that

$$\lambda \left[\pi_1^{(V)} - \pi_2^{(V)} \right] L_2 \left(\pi_1^{(W)}, \pi_2^{(W)}; \pi_1^{(V)}, \pi_2^{(V)} \right) - \theta \pi_2^{(V)} = 0,$$

that is

$$\lambda \left[\xi_1 - \pi_2^{(V)} \right] L_2 \left(\delta_1, \delta_2; \xi_1, \pi_2^{(V)} \right) - \theta \pi_2^{(V)} = 0.$$

Let

$$\begin{aligned} F_2(x) &= \theta x - \lambda \left[\xi_1 - x \right] L_2 \left(\delta_1, \delta_2; \xi_1, x \right) \\ &= \theta x - \lambda \frac{\xi_1 - x}{\xi_1 - x + \delta_1 - \delta_2} \left[(\xi_1 + \delta_1)^d - (x + \delta_2)^d \right]. \end{aligned}$$

Then $F_2(0) = -\lambda \xi_1 L_2(\delta_1, \delta_2; \xi_1, 0) < 0$, $F_2(\xi_1) = \theta \xi_1 > 0$ and for $x \in (0, \xi_1)$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x} F_2(x) &= \theta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda \frac{(\xi_1 - x)(\xi_1 + \delta_1)^d}{\xi_1 - x + \delta_1 - \delta_2}\right) + \frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda \frac{(\xi_1 - x)(x + \delta_2)^d}{\xi_1 - x + \delta_1 - \delta_2}\right) \\ &= \theta + \lambda \frac{(\xi_1 + \delta_1)^d (\xi_1 - x + \delta_1 - \delta_2)}{(\xi_1 - x + \delta_1 - \delta_2)^2} + \lambda \frac{(\xi_1 - x)(\xi_1 + \delta_1)^d}{(\xi_1 - x + \delta_1 - \delta_2)^2} \\ &\quad - \lambda \frac{(x + \delta_2)^d (\xi_1 - x + \delta_1 - \delta_2)}{(\xi_1 - x + \delta_1 - \delta_2)^2} - \lambda \frac{(\xi_1 - x)(x + \delta_2)^d}{(\xi_1 - x + \delta_1 - \delta_2)^2} \\ &\quad + \lambda \frac{d(\xi_1 - x)(x + \delta_2)^{d-1}(\xi_1 - x + \delta_1 - \delta_2)}{(\xi_1 - x + \delta_1 - \delta_2)^2} > 0.\end{aligned}$$

Since $F_2(x)$ is a continuous function for $x \in (0, \xi_1)$, there exists a unique positive solution $x = \xi_2$ to the nonlinear equation $F_2(x) = 0$ for $x \in (0, \xi_1)$. Hence, $\pi_2^{(V)} = \xi_2$. It follows from Eqs. 58 and 67 that

$$\begin{aligned}\pi_3^{(W)} &= \pi_2^{(W)} - \theta \pi_2^{(V)} - \lambda \left[\pi_1^{(W)} - \pi_2^{(W)}\right] L_2\left(\pi_1^{(W)}, \pi_2^{(W)}; \pi_1^{(V)}, \pi_2^{(V)}\right) \\ &= \delta_2 - \theta \xi_2 - \lambda [\delta_1 - \delta_2] L_2(\delta_1, \delta_2; \xi_1, \xi_2) = \delta_3.\end{aligned}$$

We assume that for $l = k$, $\pi_k^{(V)} = \xi_k$ and $\pi_{k+1}^{(W)} = \delta_{k+1}$, where $0 < \xi_k < \xi_{k-1} < \cdots < \xi_1 < \xi_0 = 1 - \lambda$ and $0 < \delta_{k+1} < \delta_k < \cdots < \delta_2 < \delta_1 = \lambda$. Then for $l = k + 1$, it follows from Eq. 59 that

$$\lambda \left[\pi_k^{(V)} - \pi_{k+1}^{(V)}\right] L_{k+1}\left(\pi_k^{(W)}, \pi_{k+1}^{(W)}; \pi_k^{(V)}, \pi_{k+1}^{(V)}\right) - \theta \pi_{k+1}^{(V)} = 0,$$

that is

$$\lambda \left[\xi_k - \pi_{k+1}^{(V)}\right] L_{k+1}\left(\delta_k, \delta_{k+1}; \xi_k, \pi_{k+1}^{(V)}\right) - \theta \pi_{k+1}^{(V)} = 0.$$

Let

$$\begin{aligned}F_{k+1}(x) &= \theta x - \lambda [\xi_k - x] L_{k+1}(\delta_k, \delta_{k+1}; \xi_k, x) \\ &= \theta x - \lambda \frac{\xi_k - x}{\xi_k - x + \delta_k - \delta_{k+1}} \left[(\xi_k + \delta_k)^d - (x + \delta_{k+1})^d\right].\end{aligned}$$

Then $F_{k+1}(0) = -\lambda \xi_k L_{k+1}(\delta_k, \delta_{k+1}; \xi_k, 0) < 0$, $F_{k+1}(\xi_k) = \theta \xi_k > 0$ and for $x \in (0, \xi_k)$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x} F_{k+1}(x) &= \theta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda \frac{(\xi_k - x)(\xi_k + \delta_k)^d}{\xi_k - x + \delta_k - \delta_{k+1}}\right) + \frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda \frac{(\xi_k - x)(x + \delta_{k+1})^d}{\xi_k - x + \delta_k - \delta_{k+1}}\right) \\ &= \theta + \lambda \frac{(\xi_k + \delta_k)^d (\xi_k - x + \delta_k - \delta_{k+1})}{(\xi_k - x + \delta_k - \delta_{k+1})^2} + \lambda \frac{(\xi_k - x)(\xi_k + \delta_k)^d}{(\xi_k - x + \delta_k - \delta_{k+1})^2} \\ &\quad - \lambda \frac{(x + \delta_{k+1})^d (\xi_k - x + \delta_k - \delta_{k+1})}{(\xi_k - x + \delta_k - \delta_{k+1})^2} - \lambda \frac{(\xi_k - x)(x + \delta_{k+1})^d}{(\xi_k - x + \delta_k - \delta_{k+1})^2} \\ &\quad + \lambda \frac{d(\xi_k - x)(x + \delta_{k+1})^{d-1}(\xi_k - x + \delta_k - \delta_{k+1})}{(\xi_k - x + \delta_k - \delta_{k+1})^2} > 0.\end{aligned}$$

Note that $F_{k+1}(x)$ is a continuous function for $x \in (0, \xi_k)$, there exists a unique positive solution $x = \xi_{k+1}$ to the nonlinear equation $F_{k+1}(x) = 0$ for $x \in (0, \xi_k)$. Hence, $\pi_{k+1}^{(V)} = \xi_{k+1}$. It follows from Eqs. 58 and 67 that

$$\begin{aligned}\pi_{k+2}^{(W)} &= \pi_{k+1}^{(W)} - \theta\pi_{k+1}^{(V)} - \lambda\left[\pi_k^{(W)} - \pi_{k+1}^{(W)}\right] L_{k+1}\left(\pi_k^{(W)}, \pi_{k+1}^{(W)}; \pi_k^{(V)}, \pi_{k+1}^{(V)}\right)\\ &= \delta_{k+1} - \theta\xi_{k+1} - \lambda\left[\delta_k - \delta_{k+1}\right] L_{k+1}(\delta_k, \delta_{k+1}; \xi_k, \xi_{k+1}) = \delta_{k+2}.\end{aligned}$$

By induction, this completes the proof. □

Under the condition $\lambda < 1$, the following theorem describes two important limiting processes which are related to the fixed point. The two limiting processes are interesting when the convergence of fraction vector sequence is understood as $N \to \infty$ and/or $t \to 0$. Here, we omit its proof, while the proof can be completed by a similar discussion to those of Theorem 1 (iii) and Theorem 4 in Martin and Suhov (1999).

Theorem 6

(1) *If $\lambda < 1$, then for any $(\mathbf{g}, \mathbf{h}) \in \Omega$*

$$\lim_{t\to+\infty} \left(\mathbf{u}(t, \mathbf{g}, \mathbf{h}), \mathbf{v}(t, \mathbf{g}, \mathbf{h})\right) = \pi.$$

Furthermore, there exists a unique probability measure φ on Ω, which is invariant under the map $(\mathbf{g}, \mathbf{h}) \longmapsto \left(\mathbf{u}(t, \mathbf{g}, \mathbf{h}), \mathbf{v}(t, \mathbf{g}, \mathbf{h})\right)$, that is, for any continuous function $f : \Omega \to \mathbf{R}$ and $t > 0$

$$\int_\Omega f(\mathbf{g}, \mathbf{h}) \mathrm{d}\varphi(\mathbf{g}, \mathbf{h}) = \int_\Omega f(\mathbf{u}(t, \mathbf{g}, \mathbf{h}), \mathbf{v}(t, \mathbf{g}, \mathbf{h})) d\varphi(\mathbf{g}, \mathbf{h}).$$

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

(2) *If $\lambda < 1$, then for a fixed number $N = 1, 2, 3, \ldots$, the Markov process $\left\{(\mathbf{U}^{(N)}(t), \mathbf{V}^{(N)}(t)), t \geq 0\right\}$ is positive recurrent, and hence it has a unique invariant distribution φ_N. Furthermore, $\{\varphi_N\}$ weakly converges to δ_π, that is, for any continuous function $f : \Omega \to \mathbf{R}$*

$$\lim_{N\to\infty} E_{\varphi_N}\left[f(\mathbf{g}, \mathbf{h})\right] = f(\pi).$$

Based on Theorems 6, we obtain a useful relation as follows

$$\lim_{t\to+\infty}\lim_{N\to\infty} \left(\mathbf{u}^{(N)}(t, \mathbf{g}, \mathbf{h}), \mathbf{v}^{(N)}(t, \mathbf{g}, \mathbf{h})\right) = \lim_{N\to\infty}\lim_{t\to+\infty} \left(\mathbf{u}^{(N)}(t, \mathbf{g}, \mathbf{h}), \mathbf{v}^{(N)}(t, \mathbf{g}, \mathbf{h})\right) = \pi.$$

Therefore, we have

$$\lim_{\substack{N\to\infty\\ t\to+\infty}} \left(\mathbf{u}^{(N)}(t, \mathbf{g}, \mathbf{h}), \mathbf{v}^{(N)}(t, \mathbf{g}, \mathbf{h})\right) = \pi.$$

6 Performance analysis and numerical examples

In this section, we provide performance analysis of the supermarket model with server multiple vacations, including the mean of the stationary queue length in any server and the expected sojourn time that any arriving customer spends in this system. Furthermore, we use some numerical examples to analyze how the two performance measures depend on some crucial factors of this supermarket model.

Let $\mathcal{Q}$ be the stationary queue length of any server. Then

$$E\left[\mathcal{Q}\right] = \sum_{k=1}^{\infty} P\left\{\mathcal{Q} \geq k\right\} = \sum_{k=1}^{\infty} \left[\pi_k^{(W)} + \pi_k^{(V)}\right] = \sum_{k=1}^{\infty} \left(\xi_k + \delta_k\right). \tag{68}$$

Note that

$$E\left[\mathcal{Q}\right] = E\left[\mathcal{Q} \text{ this server is at vacation}\right] + E\left[\mathcal{Q}, \text{ this server is working}\right],$$

where

$$E\left[\mathcal{Q}, \text{ this server is at vacation}\right] = \sum_{k=1}^{\infty} \pi_k^{(V)} = \sum_{k=1}^{\infty} \xi_k$$

and

$$E\left[\mathcal{Q}, \text{ this server is working}\right] = \sum_{k=1}^{\infty} \pi_k^{(W)} = \sum_{k=1}^{\infty} \delta_k.$$

If $\lambda < 1$, then this supermarket model with server multiple vacations is stable. In this case, we denote by $\mathcal{S}$ the sojourn time that any arriving customer spends in this system. It is easy to see that

$$\begin{aligned} E\left[\mathcal{S}, \text{ this server is at vacation}\right] = &\left(\frac{1}{\theta} + 1\right) \left(\xi_0 - \xi_1\right) L_1\left(\delta_1; \xi_0; \xi_1\right) \\ &+ \sum_{k=2}^{\infty} \left(\frac{1}{\theta} + k\right) \left(\xi_{k-1} - \xi_k\right) L_k\left(\delta_{k-1}, \delta_k; \xi_{k-1}; \xi_k\right) \end{aligned}$$

and

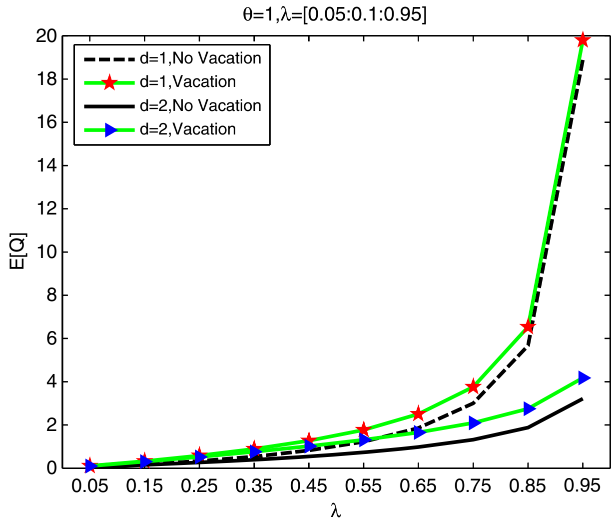
$$E\left[\mathcal{S}, \text{ this server is working}\right] = \sum_{k=2}^{\infty} k\left(\delta_{k-1} - \delta_k\right) L_k\left(\delta_{k-1}, \delta_k; \xi_{k-1}; \xi_k\right),$$

thus we obtain

$$\begin{aligned} E\left[\mathcal{S}\right] =& E\left[\mathcal{S}, \text{ this server is at vacation}\right] + E\left[\mathcal{S}, \text{ this server is working}\right] \\ =& \frac{(1+\theta)}{\theta} \left(\xi_0 - \xi_1\right) L_1\left(\delta_1; \xi_0; \xi_1\right) \\ &+ \frac{1}{\theta} \sum_{k=2}^{\infty} (1 + k\theta) \left(\xi_{k-1} - \xi_k\right) L_k\left(\delta_{k-1}, \delta_k; \xi_{k-1}; \xi_k\right) \\ &+ \sum_{k=2}^{\infty} k\left(\delta_{k-1} - \delta_k\right) L_k\left(\delta_{k-1}, \delta_k; \xi_{k-1}; \xi_k\right). \end{aligned} \tag{69}$$

In the remainder of this section, using Eqs. 68 and 69 we provide some numerical examples to analyze how the two performance measures $E\left[\mathcal{Q}\right]$ and $E\left[\mathcal{S}\right]$ depend on some crucial parameters of this supermarket model under several different values of d.

Fig. 6 The stationary expected queue length depends on vacation or no vacation



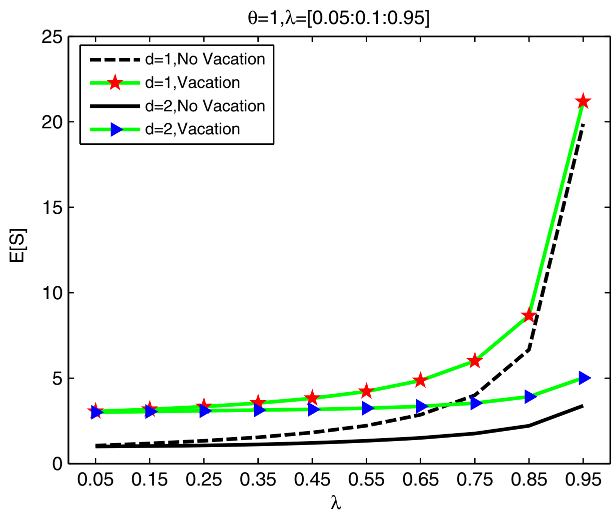
(1) The role of vacation processes

In this supermarket model, we assume that the exponential service rate $\mu = 1$, the exponential vacation rate $\theta = 1$ and the Poisson arrival rate $\lambda \in (0.05, 0.95)$.

Figures 6 and 7 indicate how the two performance measures of the supermarket model depend on the role played by the vacation processes, where $\lambda \in (0.05, 0.95)$.

Figure 6 shows that the vacation processes increase $E\left[\mathcal{Q}\right]$ under the two cases with $d = 1, 2$. At the same time, the choice number d decreases $E\left[\mathcal{Q}\right]$ when the servers have either vacations or no vacations.

Fig. 7 The expected sojourn time depends on vacation or no vacation



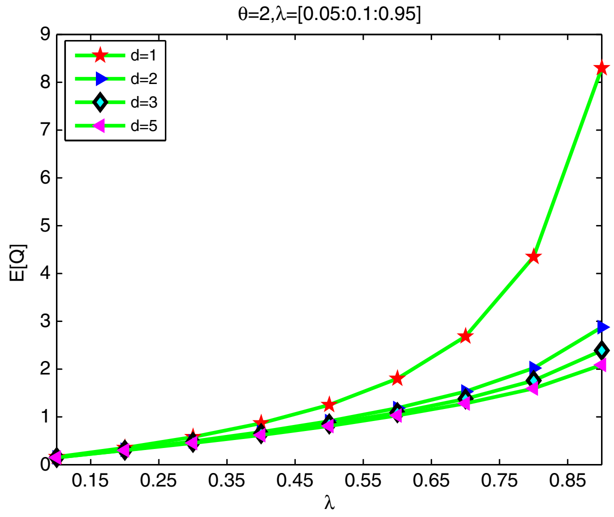


Fig. 8 The mean of the stationary queue length vs λ for the different values of d

Figure 7 illustrates that the vacation processes also increase $E\left[\mathcal{S}\right]$ under the two cases with $d = 1, 2$. When the servers have either vacations or no vacations, $E\left[\mathcal{S}\right]$ decreases as the choice number d increases.

(2) The role of Poisson arrival rates

In this supermarket model, we assume that the exponential service rate $\mu = 1$, the exponential vacation rate $\theta = 0.4$ and the Poisson arrival rate $\lambda \in (0.05, 0.95)$.

Figures 8 and 9 indicate how the two performance measures of the supermarket model depend on the Poisson arrival rate $\lambda \in (0.05, 0.95)$.

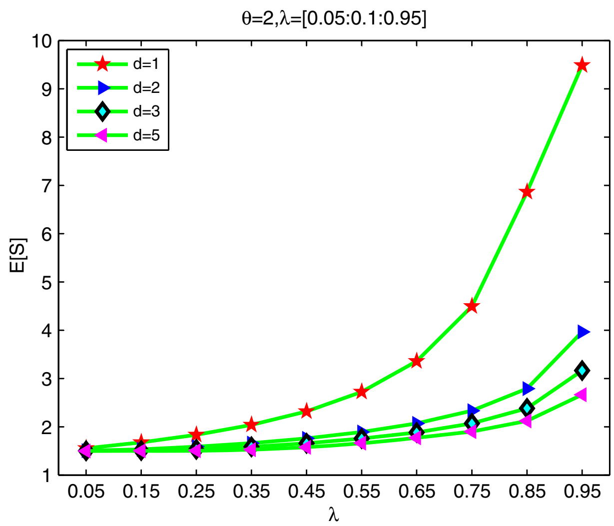


Fig. 9 The expected sojourn time vs λ for the different values of d

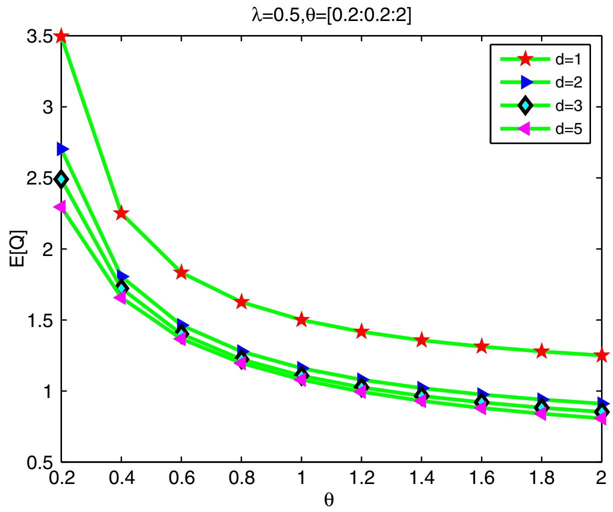


Fig. 10 The mean of the stationary queue length vs θ for the different values of d

Figure 8 shows that $E\left[\mathcal{Q}\right]$ increases as λ increases under the three cases with $d = 1, 2, 5$. However, the role of choice number d is complicated; only when λ is big, $E\left[\mathcal{Q}\right]$ decreases as d increases.

Figure 9 illustrates that $E\left[\mathcal{S}\right]$ increases as λ increases under the four cases with $d = 1, 2, 3, 4$. At the same time, $E\left[\mathcal{S}\right]$ decreases as the choice number d increases, this is different from that in $E\left[\mathcal{Q}\right]$.

(3) The role of exponential vacation rates

In this supermarket model, we assume that the exponential service rate $\mu = 1$, the Poisson arrival rate $\lambda = 0.5$ and the exponential vacation rate $\theta \in (0.2, 2)$.

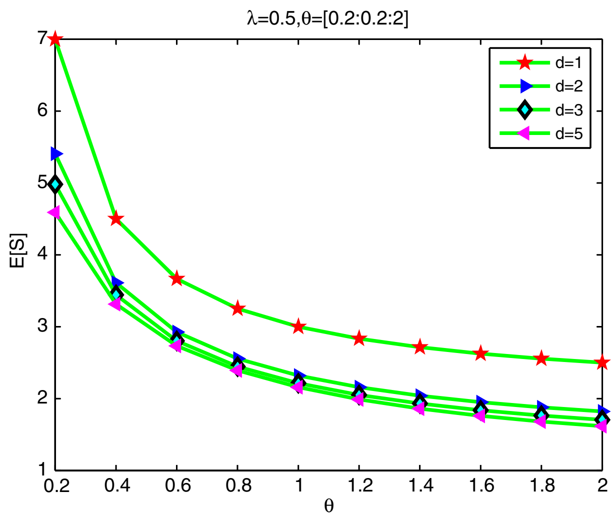


Fig. 11 The expected sojourn time vs θ for the different values of d

Figures 10 and 11 indicate how the two performance measures of the supermarket model depend on the exponential vacation rate $\theta \in (0.2, 2)$.

Figure 10 shows that $E\left[\mathcal{Q}\right]$ decreases as θ increases under the four cases with $d = 1, 2, 3, 4$. However, the role of choice number d is complicated, we can not describe how $E\left[\mathcal{Q}\right]$ depends on d yet.

Figure 11 illustrates that $E\left[\mathcal{S}\right]$ decreases as θ increases under the four cases with $d = 1, 2, 3, 4$. At the same time, $E\left[\mathcal{S}\right]$ decreases as the choice number d increases, this is different from that in $E\left[\mathcal{Q}\right]$.

7 Concluding remarks

In this paper, we first analyze a supermarket model of N identical servers with server multiple vacations, and set up an infinite-dimensional system of differential equations satisfied by the expected fraction vectors in terms of the technique of tailed equations. Then, as $N \to \infty$ we use the operator semigroup to provide a mean-field limit for the sequence of Markov processes, which weakly converges to the unique and global solution for the infinite-dimensional system of limiting differential equations. Finally, we provide an effective algorithm for computing the fixed point of the infinite-dimensional system of limiting differential equations. Using the fixed point, we provide performance analysis of this supermarket model, and also give some numerical examples to analyze how the two performance measures depend on some crucial factors of this supermarket model.

This paper provides a clear picture for how to use the mean-field models to numerically analyze performance measures of complicated supermarket models, and this picture is organized into three key parts: (1) Setting up system of differential equations, (2) strict proofs of the mean-field limit and (3) performance analysis of system. Therefore, the method given in this paper can be applied to performance analysis of complicated supermarket models with more random factors, such as cases where each server is a retrial queue or a processor-sharing queue, each server may be failure and repaired, the customers may be impatient or negative. Along these lines, there are a number of interesting directions for potential future research, for example:

1. We need to develop effective algorithms for computing the fixed point of complicated supermarket models, and specifically, to analyze the influence of crucial random factors on the design of algorithms. In fact, this paper indicates that the nonlinear dynamics of the supermarket model with server multiple vacations makes the system of limiting differential equations more complicated, making the computation of the fixed points more difficult and challenging. On the other hand, it is worthwhile to note that for a Markov process, the fixed point with nonlinear structure is very different from the stationary probability vectors with linear structure (e.g., see the RG-factorizations given in Li (2010)), therefore there are many interesting topics for computing the fixed point with the nonlinear dynamic structure.
2. How to apply the operator semigroup to provide the mean-field limit in the study of supermarket models with either non-Poisson inputs or with non-exponential service times is still an open and interesting problem. Reader may refer to, such

as, Bramson et al. (2010, 2012, 2011), Li and Lui (2010) and Li (2011). Recently, Li et al. (2012) made crucial advances in applying the operator semigroup and the mean-field limit to analyzing the supermarket model with PH service times. However, we believe that a large gap still exists for dealing with either renewal inputs or general service times, because of the fact that a more complicated nonlinear dynamic structure exists and needs to be given a detailed analysis in order to set up the infinite-dimensional system of differential equations.

Acknowledgements The authors thank the two referees for their valuable comments and remarks, and acknowledge Professor Benny van Houdt for many valuable suggestions to sufficiently improve the system of differential equations derived in Section 2 of this paper. At the same time, Q.L. Li and Y. Wang thank that this research is partly supported by the National Natural Science Foundation of China (No. 71271187, No. 61001075) and the Hebei Natural Science Foundation of China (No. A2012203125).

References

Alfa AS (2003) Vacation models in discrete time. Queueing Syst 44:5–30

Benaim M, Le Boudec JY (2008) A class of mean-field interaction models for computer and communication systems. Perform Eval 65:823–838

Bordenave C, McDonald D, Proutiere A (2009) A particle system in interaction with a rapidly varying environment: mean-field limits and applications. Available: arXiv:0701363v3

Le Boudec JY, McDonald D, Mundinger J (2007) A generic mean-field convergence result for systems of interacting objects. In: Proc. Conf. IEEE on the quantitative evaluation of systems, pp 3–18

Bramson M (2008) Stability of queueing networks. Springer-Verlag, New York

Bramson M (2011) Stability of join the shortest queue networks. Ann Appl Probab 21:1568–1625

Bramson M, Lu Y, Prabhakar B (2010) Randomized load balancing with general service time distributions. In: Proceedings of the ACM SIGMETRICS international conference on measurement and modeling of computer systems, pp 275–286

Bramson M, Lu Y, Prabhakar B (2011) Decay of tails at equilibrium for FIFO join the shortest queue networks. Available: arXiv:1106.4582

Bramson M, Lu Y, Prabhakar B (2012) Asymptotic independence of queues under randomized load balancing. Queueing Syst 71:247–292

Doshi BT (1986) Queueing system with vacation: a survey. Queueing Syst 1:29-66

Doshi BT (1990) Generalization of the stochastic decomposition results for single server queues with vacations. Stoch Model 6:307–333

Dshalalow JH (1995) Advances in queueing theory, methods and open problems. CRC Press, Boca Raton

Dshalalow JH (1997) Frontiers in queueing models and applications in science and engineering. CRC Press, Boca Raton

Ethier SN, Kurtz TG (1986) Markov processes: characterization and convergence. John Wiley & Sons, New York

Fleming W (1977) Functions of several variables, 2nd edn. Springer, New York

Foss S, Chernova N (1998) On the stability of a partially accessible multi-station queue with state-dependent routing. Queueing Syst 29:55–73

Fuhrmann SW, Cooper RB (1985) Stochastic decomposition in the M/G/1 queue with generalized vacation. Oper Res 33:1117–1129

Gast N, Gaujal B (2009) A mean-field approach for optimization in particles systems and applications. Technical Report N^0 6877, INRIA Rocquencourt, Frence

Gast N, Gaujal B (2012) Markov chains with discontinuous drift have differential inclusions limits. Technical Report N^0 7315, INRIA Rocquencourt, Frence

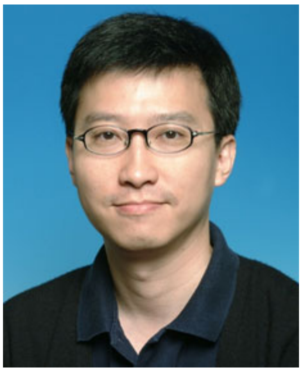
Gast N, Gaujal B, Le Boudec JY (2011) Mean-field for Markov decision processes: from discrete to continuous optimization. Technical Report N^0 7239, INRIA Rocquencourt, Frence

Graham C (2000a) Kinetic limits for large communication networks. In: Bellomo N, Pulvirenti M (eds) Modelling in applied sciences. Birkhäuser, pp 317–370

Graham C (2000b) Chaoticity on path space for a queueing network with selection of the shortest queue among several. J Appl Probab 37:198–201
Graham C (2004) Functional central limit theorems for a large network in which customers join the shortest of several queues. Probab Theory Relat Fields 131:97–120
Hale JK (1980) Ordinary differential equations. Roberte E. Krieger Publishing, Melbourne
Jacquet P, Vvedenskaya ND (1998) On/off sources in an interconnection networks: performance analysis when packets are routed to the shortest queue of two randomily selected nodes. Technical Report N^0 3570, INRIA Rocquencourt, Frence
Jacquet P, Suhov YM, Vvedenskaya ND (1999) Dynamic routing in the mean-field approximation. Technical Report N^0 3789, INRIA Rocquencourt, Frence
Kurtz TG (1981) Approximation of population processes. SIAM
Li QL (2010) Constructive computation in stochastic models with applications: the RG-factorizations. Springer and Tsinghua Press
Li QL (2011) Super-exponential solution in Markovian supermarket models: framework and challenge. Available: arXiv:1106.0787
Li QL, Dai G, Lui JCS, Wang Y (2012) Matrix-structured super-exponential tail in the supermarket model with PH service times. http://www.liquanlin.com/_d274639327.htm
Li QL, Lui JCS (2010) Doubly exponential solution for randomized load balancing models with Markovian arrival processes and PH service times. Available: arXiv:1105.4341
Li QL, Lui JCS, Wang Y (2011) A matrix-analytic solution for randomized load balancing models with PH service times. In: Performance evaluation of computer and communication systems: milestones and future challenges. Lecture notes in computer science, vol 6821, pp 240–253
Luczak MJ, McDiarmid C (2006) On the maximum queue length in the supermarket model. Ann Probab 34:493–527
Luczak MJ, McDiarmid C (2007) Asymptotic distributions and chaos for the supermarket model. Electron J Probab 12:75–99
Luczak MJ, Norris JR (2005) Strong approximation for the supermarket model. Ann Appl Probab 15:2038–2061
Martin JB (2001) Point processes in fast Jackson networks. Ann Appl Probab 11:650–663
Martin JB, Suhov YM (1999) Fast Jackson networks. Ann Appl Probab 9:854–870
Mitzenmacher MD (1996) The power of two choices in randomized load balancing. PhD thesis, Department of Computer Science, University of California at Berkeley, USA
Mitzenmacher MD (1999) On the analysis of randomized load balancing schemes. Theory Comput Syst 32:361–386
Mitzenmacher MD, Richa A, Sitaraman R (2001) The power of two random choices: a survey of techniques and results. In: Handbook of randomized computing, vol 1, pp 255–312
Mitzenmacher MD, Upfal E (2005) Probability and computing: randomized algorithms and probabilistic analysis. Cambridge University Press
Suhov YM, Vvedenskaya ND (2002) Fast Jackson networks with dynamic routing. Probl Inf Transm 38:136–153
Sznitman A (1989) Topics in propagation of chaos. In: Springer-Verlag lecture notes in mathematics 1464. École d'Été de Probabilités de Saint-Flour XI, pp 165–251
Takagi H (1991) Queueing analysis: a foundation of performance evaluation, vacation and priority systems, part 1. North-Holland, Amsterdam
Taylor ME (1996) Partial differential equations, I Basic Theory. Springer-Verlag, New York
Tian NS, Li QL, Cao J (1999) Conditional stochastic decompositions in the M/M/C queue with server vacations. Stoch Model 15:367–378
Tian NS, Zhang ZG (2006) Vacation queueing models: theory and applications. Springer, New York
Tsitsiklis JN, Xu K (2012) On the power of (even a little) resource pooling. Stochastic Systems 2:1–66
Turner SRE (1996) Resource pooling in stochastic networks. Ph.D. Thesis, Statistical Laboratory, Christ's College, University of Cambridge
Turner SRE (1998) The effect of increasing routing choice on resource pooling. Probability in the Engineering and Informational Sciences 12:109–124
Vvedenskaya ND, Dobrushin RL, Karpelevich FI (1996) Queueing system with selection of the shortest of two queues: an asymptotic approach. Probl Inf Transm 32:15–27
Vvedenskaya ND, Suhov YM (1997) Dobrushin's mean-field approximation for a queue with dynamic routing. Markov Processes and Related Fields 3:493–526
Vvedenskaya ND, Suhov YM (2005) Dynamic routing queueing systems with vacations. Information Processes, Electronic Scientific Journal. The Keldysh Institute of Applied Mathematics. The Institute for Information Transmission Problems, vol 5, pp 74–86

Quan-Lin Li is Full Professor in School of Economics and Management Sciences, Yanshan University, Qinhuangdao, China. He received the Ph.D. degree in Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, China in 1998. He has published a book (Constructive Computation in Stochastic Models with Applications: The RG-Factorizations, Springer, 2010) and over 40 research papers in a variety of journals, such as, Advances in Applied Probability, Queueing Systems, Stochastic Models, European Journal of Operational Research, Computer Networks, Performance Evaluation, Discrete Event Dynamic Systems, Computers & Operations Research, Computers & Mathematics with Applications, Annals of Operations Research, and International Journal of Production Economics. His main research interests concern with Queueing Theory, Stochastic Models, Matrix-Analytic Methods, Manufacturing Systems, Computer Networks, Network Security, and Supply Chain Risk Management.

Guirong Dai is a master student in School of Economics and Management Sciences, Yanshan University, Qinhuangdao, China. Her research interests include Queueing Theory, Stochastic Models, Computer Networks.



John C. S. Lui (M93-SM02-F10) was born in Hong Kong. He received the Ph.D. degree in computer science from the University of California, Los Angeles, 1992. He is currently a Professor with the Department of Computer Science and Engineering, The Chinese University of Hong Kong (CUHK), Hong Kong. He was the chairman of the Department from 2005 to 2011. His current research interests are in communication networks, network/system security (e.g., cloud security, mobile security, etc.), network economics, network sciences (e.g., online social networks, information spreading, etc.), cloud computing, large-scale distributed systems, and performance evaluation theory. Professor Lui is a Fellow of the Association for Computing Machinery (ACM), a Fellow of IEEE, a Croucher Senior Research Fellow, and an elected member of the IFIP WG 7.3. He serves on the Editorial Board of IEEE/ACM Transactions on Networking, IEEE Transactions on Computers, IEEE Transactions on Parallel and Distributed Systems, Journal of Performance Evaluation and International Journal of Network Security. He received various departmental teaching awards and the CUHK Vice-Chancellors Exemplary Teaching Award. He is also a co-recipient of the IFIP WG 7.3 Performance 2005 and IEEE/IFIP NOMS 2006 Best Student Paper Awards.

Yang Wang is an Assistant Professor in the Department of Computer Science and Technology, Peking University, Beijing, PR China. She received her Ph.D. degree from the Department of Computer Science and Technology, Tsinghua University in 2009. Her research interests include network security, system performance evaluation, wireless networks and communication networks etc.