Queueing Syst (2007) 57: 169–178
DOI 10.1007/s11134-007-9055-4

Analysis of a time-limited service priority queueing system with exponential timer and server vacations

Tsuyoshi Katayama

Received: 22 June 2006 / Revised: 10 July 2007 / Published online: 16 January 2008

Abstract We consider a multi-class priority queueing system with a non-preemptive time-limited service controlled by an exponential timer and multiple (or single) vacations. By reducing the service discipline to the Bernoulli schedule, we obtain an expression for the Laplace-Stieltjes transform (LST) of the waiting time distribution via an iteration procedure, and a recursive scheme to calculate the first two moments. It is noted that we have to select embedded Markov points based on the service beginning epochs instead of the service completion epochs adopted for most of $M/G/1$ queueing analyses. Through the queue-length analysis, we obtain a decomposition form for the LST of the waiting time in each queue having the exhaustive service.

Keywords Flexible priority system · Time-limited service · Stochastic decomposition · Iterative functional equations

Mathematics Subject Classification (2000) 60K25 · 39B12

1 Introduction

Priority queueing systems have been extensively studied by many researchers, and various priority disciplines have been proposed for computer, communications systems, production systems and other kinds of systems. However, most of classic priority disciplines, such as the head-of-the-line and the shortest-job-first priority disciplines, have no controllable parameters. An N (≥ 2)-class priority queueing system with time-limited services is defined by controllable parameters $(T_1, T_2, \ldots, T_N)$. A service period for class-n messages is controlled by a timer with limited time $T_n, n \in \{1, 2, \ldots, N\}$. The server serves messages until either the timer expires, or the queue becomes empty, whichever occurs first. The time-limited service with T_n covers a wide service range, as it is equivalent to the 1-limited service if $T_n = 0$ (nonpreemptive service), and to the exhaustive service if $T_n = \infty$. Such flexible service disciplines with controllable parameter set as $\{T_n\}$ are effective for the performance optimization, e.g. cf. [1, 2, 13] for the polling systems, and have a potential applicability to processing systems with multiple grades of service requirements, e.g. packet multiplexer systems with multiple priority classes in broadband multimedia networks and the routing scheme used in the Internet. Communication systems of two buffers ($N = 2$) with flexible service disciplines have been analyzed for various new service disciplines such as the K-limited service, the D-policy, the Bernoulli schedule, the time-limited service period scheme, the randomly-timed gated regime and so on [5, 7, 8, 15, 21] (cf. Sect. 5.3.6 in [22] for an overview).

Despite the effectiveness of flexible service disciplines, few analytical results in the literature have been obtained for flexible priority systems with multiple ($N \geq 3$) message classes. This may be due to the difficulty of queueing analysis for the above flexible priority disciplines. Kleinrock and Finkelstein [15] studied a multi-class time-dependent $M/M/1$ priority system with a controllable parameter set $\{b_n\}$, in which a single server serves a message with the highest instantaneous priority at time t defined by $q_n^{(r)} := b_n(t - t_0)^r$ for any non-negative number r, where $0 \leq b_1 \leq b_2 \leq \cdots \leq b_N$ and t_0 denotes the arrival

T. Katayama (✉)
Department of Electronics and Informatics, Toyama Prefectural University, Kosugi, Toyama 939-0398, Japan
e-mail: katayama@pu-toyama.ac.jp

time of the message. Katayama [11] analyzed a multi-class $M/G/1$ priority system with the Bernoulli schedule with a controllable parameter set $\{p_n\}$. Recently, Katayama and Kobayashi [12] analyzed a multi-class $M/D/1$ priority system with constant service times and exponential timer with a limited time set $\{T_n\}$. We review the previous literature on the time-limited service: Coffman, Fayolle and Mitrani [5] considered an $M/M/1$, alternating service model with preemptive exponentially distributed time-limits (T_1, T_2), and determined the generating function of a joint queue-length distribution by using the boundary value technique for the first time. Leung and Eisenberg [17] analyzed an $M/G/1$ vacation model with a non-gated time-limited service, and obtained an approximate distribution function of the unfinished work at the polling instant by using a weighted sum of Laguerre functions (cf. Sect. 2.6 in Takagi [21]). Leung and Lucantoni [18] analyzed $M/G/1$ vacation models with constant time-limited service and vacation-dependent, time-limited service by using exponential time stages to approximate the constant time limits. Eliazar, Fibich and Yechiali [7] studied an $M/G/1$ two-queue model with the exhaustive and randomly timed gated services. Most of papers related to polling models with time-limited services can be found in Sect. 5.3.4 in Takagi [22].

In this paper, we shall analyze multi-class $M/G/1$ priority queues with nonpreemptive time-limited services and server vacations, which gives a generalization of the previous results [12]. The rest of the paper is organized as follows: Sect. 2 describes the model and notation. In Sect. 3, we prepare some relations between the nonpreemptive time-limited service and the Bernoulli schedule. In Sect. 4, we give a queue-length analysis using embedded Markov points based on the service beginning epochs. In Sect. 5, we obtain the LST of the waiting time distribution and the first two moments. In Sect. 6, we give concluding remarks.

2 The model and notation

There are N (≥ 2) classes of messages that arrive at each queue with an infinite capacity buffer. Arrivals of class-n messages form a Poisson process with rate λ_n, $n = 1, 2, \ldots, N$, where the message classes are priority classes in the order of indexes (class 1 is the highest, and class N the lowest). Messages of the same class are served according to the FIFO rule. The priority discipline of nonpreemptive time-limited service is defined by a parameter vector $\boldsymbol{T} := (T_1, T_2, \ldots, T_N)$, where T_n (≥ 0) is the maximum attendance time (MAT) and operates as follows: Once starting service of class-n messages, a single server continues to serve only class-n messages until either queue n becomes empty or the timer with MAT T_n expires, whichever occurs first. If a message is in service at the timer expiration, the server completes the current message service (called nonpreemptive service). In succession, the highest class present in the system is next served. At service completion of a message, if there are no messages in the system, the server goes away for a vacation according to the following two rules. Multiple vacation (MV): If, on return from a vacation, the server finds the system empty, it goes on another vacation and so on until it finds at least one message waiting on return from a vacation. Single vacation (SV): The server takes exactly one vacation. If it finds no message in the system, it becomes idle until a message arrives. When a message arrives, it immediately begins the time-limited service. If the server returns from a vacation to find the system not empty, the server starts serving a message of the highest class present in the system.

The service time of class-n message (H_n), the MAT for class-n messages (T_n) and vacation time (V) are i.i.d. random variables. Given a non-negative valued random variable X, we denote throughout the paper by $X(t)$ its distribution function (DF), $X^c(t) := 1 - X(t)$ and by $X^*(s)$ its Laplace-Stieltjes transform (LST). We denote its finite first and mth moment of the DF by x and $x^{(m)}$, $m = 2, 3, \ldots$, respectively. The following notation is used for $n = 1, 2, \ldots, N$:

$$\lambda_n^+ := \sum_{j=1}^{n} \lambda_j, \qquad \lambda := \lambda_N^+, \qquad r_n := \frac{\lambda_n}{\lambda},$$

$$\rho_n := \lambda_n h_n, \qquad \rho_n^+ := \sum_{j=1}^{n} \rho_j,$$

$$\rho := \rho_N^+, \qquad \rho_v := \lambda v.$$

The LST of the DF of a service time (H_n^+) averaged over messages of class 1 through n is given by

$$H_n^{+*}(s) := \frac{1}{\lambda_n^+} \sum_{j=1}^{n} \lambda_j H_j^*(s).$$

We also use $\overline{\boldsymbol{T}} := (\overline{T}_1, \overline{T}_2, \ldots, \overline{T}_N)$, where $\overline{T}_n := \boldsymbol{E}(T_n)$, $n = 1, 2, \ldots, N$. We assume that the MAT T_n has an exponential distribution with mean $\overline{T}_n = 1/\alpha_n (\alpha_n \geq 0)$, and H_n and V have general distributions. The server utilization is less than unity $(\rho < 1)$ for the system stability.

Remark 2.1 For $T_1 < \infty$, class-1 messages are served successively until queue 1 becomes empty, because class-1 messages have priority over all other classes. That is, $(T_1 < \infty, T_2, \ldots, T_N)$ is equivalent to $(T_1 = \infty, T_2, \ldots, T_N)$.

3 Preliminaries

Let us first describe an $M/G/1$ priority model with Bernoulli schedule defined by a parameter vector $\boldsymbol{p} := (p_1, p_2,$

$\ldots, p_N)$, where $p_n \in [0, 1]$ for $n \in \{1, 2, \ldots, N\}$. Suppose that a service is given to a message of class n. At service completion of that message, if there are more class-n messages, another class-n message is next served with probability p_n, or a message of the highest class present in the system is next served with probability $\overline{p}_n := 1 - p_n$. At service completion of the class-n message, if there are no messages of class n, a message of the highest class present in the system is served with probability one. At service completion of a message, if there are no messages in the system, the server starts on vacation according to the MV or SV rule. The nonpreemptive, time-limited service priority queues with $\boldsymbol{T} = (T_1, T_2, \ldots, T_N)$ are closely related to the above priority model with Bernoulli schedule of $\boldsymbol{p} = (p_1, p_2, \ldots, p_N)$. Since the exponential timer is memoryless, we have

$$p_n := \Pr\{H_n \le T_n\} = H_n^*(\alpha_n). \tag{1}$$

For a conditional DF of a service time H_n under no expiration of the timer, we define

$$F_n^*(s) := \boldsymbol{E}[e^{-sH_n} | H_n \le T_n] = \frac{H_n^*(s + \alpha_n)}{H_n^*(\alpha_n)}. \tag{2}$$

It follows similarly from the above argument that

$$\begin{aligned} \bar{F}_n^*(s) &:= \boldsymbol{E}[e^{-sH_n} | H_n > T_n] \\ &= \frac{H_n^*(s) - H_n^*(s + \alpha_n)}{1 - H_n^*(\alpha_n)}. \end{aligned} \tag{3}$$

Here, we define joint probability generating functions (PGFs) for $n = 1, 2, \ldots, N$,

$$\begin{aligned} Q_{F_n}(z_1, z_2, \ldots, z_N) &:= F_n^*[\textstyle\sum_{j=1}^N \lambda_j (1 - z_j)], \\ Q_{\bar{F}_n}(z_1, z_2, \ldots, z_N) &:= \bar{F}_n^*[\textstyle\sum_{j=1}^N \lambda_j (1 - z_j)], \\ Q_n(z_1, z_2, \ldots, z_N) &:= H_n^*[\textstyle\sum_{j=1}^N \lambda_j (1 - z_j)], \\ Q_V(z_1, z_2, \ldots, z_N) &:= V^*[\textstyle\sum_{j=1}^N \lambda_j (1 - z_j)]. \end{aligned} \tag{4}$$

Further we define the following joint PGF,

$$\begin{aligned} &Q_{\mathrm{S}}^{\mathrm{M}}(z_1, z_2, \ldots, z_N) \\ &\quad := \begin{cases} [Q_V(z_1, z_2, \ldots, z_N) - V^*(\lambda)] \\ \quad /(1 - V^*(\lambda)) & \text{(MV)} \\ Q_V(z_1, z_2, \ldots, z_N) \\ \quad -V^*(\lambda)(1 - \sum_{j=1}^N r_j z_j) & \text{(SV).} \end{cases} \end{aligned} \tag{5}$$

The $Q_{\mathrm{S}}^{\mathrm{M}}(z_1, z_2, \ldots, z_N)$ represents the PGF for the number of messages in each queue at the beginning of the first service after multiple or single vacations.

We obtain the following lemma on the roots from the lemma of Takács [20, Sect. 1.3].

Lemma 1 *For $k = 1, 2, \ldots, N$, the following equation of z_k,*

$$\begin{aligned} z_k &= H_k^*[y_k + \lambda_k(1 - z_k)], \\ y_k &:= \frac{1}{\overline{T}_k} + \sum_{\substack{j=1 \\ (j \ne k)}}^N \lambda_j (1 - z_j) \end{aligned} \tag{6}$$

has a unique root $z_k = \beta_k(y_k)$ in the unit circle $|z_k| = 1$: For $Re(y) \ge 0$,

$$\begin{aligned} \beta_k(y) &:= \sum_{i=1}^\infty \frac{1}{i!} \\ &\quad \times \int_0^\infty \exp\{-(y + \lambda_k)t\}(\lambda_k t)^{i-1} dH_k^{(i)}(t), \end{aligned} \tag{7}$$

where $H_k^{(i)}(t)$ denotes the i-fold convolution of $H_k(t)$ with itself. The following equation

$$z_k = H_{k-1}^{+}{}^*[\lambda_k(1 - z) + \lambda_{k-1}^+(1 - z_k)] \tag{8}$$

has a unique root $z_1 = \gamma_1(z) \equiv 0$ and $z_k = \gamma_k(z)/\lambda_{k-1}^+$ ($k \ge 2$) in the unit circle $|z_k| = 1$: For $|z| \le 1$,

$$\begin{aligned} \gamma_k(z) &:= \sum_{i=1}^\infty \frac{\lambda_{k-1}^+}{i!} \\ &\quad \times \int_0^\infty \exp\{-(\lambda_k^+ - \lambda_k z)t\}(\lambda_{k-1}^+ t)^{i-1} dH_{k-1}^{+(i)}(t), \\ &\quad k \ge 2, \end{aligned} \tag{9}$$

where $H_{k-1}^{+(i)}(t)$ denotes the i-fold convolution of $H_{k-1}^+(t)$ with itself.

Remark 3.1 We have that $F_n^*(s) = \bar{F}_n^*(s) = H_n^*(s)$ only for $H_n^*(s) = \exp\{-h_n s\}$, which implies that H_n is a deterministic service time. That is, $F_n^*(s)$ and $\bar{F}_n^*(s)$ have a bias to $H_n^*(s)$. It should be noted that the queue-length sequence at *service-completion-epochs* is not a Markov chain, except for the constant service time. That is, if the system is not empty just after the service completion epoch of a message, then the server needs to know the past information on the service time of the message ($\bar{F}_n$ or F_n, i.e. the service time with or without timer expiration) in order to choose a new service for class-n messages or a service for the highest class messages present in the system. This point is essentially different from the Bernoulli schedule using no past information. We have also that $p_n F_n^*(s) = \boldsymbol{E}[e^{-sH_n} \mathbf{1}_{(H_n \le T_n)}]$, $\overline{p}_n \bar{F}_n^*(s) = \boldsymbol{E}[e^{-sH_n} \mathbf{1}_{(H_n > T_n)}]$ and $H_n^*(s) = p_n F_n^*(s) + \overline{p}_n \bar{F}_n^*(s)$, where $\mathbf{1}_{(A)}$ is the indicator function of the event A.

4 Queue-length analysis

We need to select service-beginning-epochs as the embedded Markov points for queueing analysis as pointed out in Remark 3.1. Let $\Pi_n^+(z_1, z_2, \dots, z_N)$ for $|z_n| \le 1$, $n = 1, 2, \dots, N$ be the generating function (GF) of the steady-state probability $\pi_n^+ (i_1, i_2, \dots, i_N)$ that, just after the service beginning epoch of a class-n message, there are i_j (≥ 0) messages of class-j in the system, $j = 1, 2, \dots, N$ $(j \ne n)$ and $i_n \ge 1$, and $\Pi_n(z_1, z_2, \dots, z_N)$ for $|z_n| \le 1$, $n = 1, 2, \dots, N$ the GF of the steady-state probability $\pi_n(i_1, i_2, \dots, i_N)$ that, just after the service completion epoch of a class-n message, there are i_j (≥ 0) messages of class-j left in the system, $j = 1, 2, \dots, N$, where we have

$$\sum_{n=1}^N \Pi_n^+(1, 1, \dots, 1) = \sum_{n=1}^N \Pi_n(1, 1, \dots, 1) = 1. \tag{10}$$

For $\overline{\boldsymbol{T}} = (\overline{T}_1, \overline{T}_2, \dots, \overline{T}_N)$, $|z_n| \le 1$ and $n \in \{1, 2, \dots, N\}$, we obtain the following relationship between these two GFs:

$$\begin{aligned}
&\frac{\Pi_n^+(z_1, z_2, \dots, z_N)}{\Pi_n^+(1, 1, \dots, 1)}\{p_n Q_{F_n}(z_1, z_2, \dots, z_N) \\
&\quad + \overline{p}_n Q_{\bar{F}_n}(z_1, z_2, \dots, z_N)\}\frac{1}{z_n} \\
&\quad = \frac{\Pi_n^+(z_1, z_2, \dots, z_N)}{\Pi_n^+(1, 1, \dots, 1)} Q_n(z_1, z_2, \dots, z_N)\frac{1}{z_n} \\
&\quad = \frac{\Pi_n(z_1, z_2, \dots, z_N)}{\Pi_n(1, 1, \dots, 1)},
\end{aligned} \tag{11}$$

where $p_n = H_n^*(1/\overline{T}_n) = H_n^*(\alpha_n)$. The term $p_n Q_{F_n}(z_1, z_2, \dots, z_N)$ on the left-hand side of (11) represents the GF for the number of messages arriving in each queue during a service time without timer expiration. Similarly, $\overline{p}_n Q_{\bar{F}_n}(z_1, z_2, \dots, z_N)$ represents the GF for the case with timer expiration, see Remark 3.1. The last factor $1/z_n$ corresponds to the departure of a class-n message just after the service completion. Considering the events that occur two consecutive service beginning epochs, we obtain a functional relationship for $\Pi_n^+(z_1, z_2, \dots, z_N)$ that is our starting point of the following queue-length analysis:

$$\begin{aligned}
&\Pi_n^+(z_1, z_2, \dots, z_N) \\
&\quad = \Pi_n^+(z_1, z_2, \dots, z_N) p_n Q_{F_n}(z_1, z_2, \dots, z_N)\frac{1}{z_n} \\
&\qquad - \Pi_n^+(z_1, \dots, z_{n-1}, \tilde{0}, z_{n+1}, \dots, z_N) \\
&\qquad \times p_n Q_{F_n}(z_1, \dots, z_{n-1}, 0, z_{n+1}, \dots, z_N)\frac{1}{z_n} \\
&\qquad + \sum_{j=1}^{n-1} AD_1(j) + \sum_{j=n}^{N} AD_2(j) + \sum_{j=n+1}^{N} AD_3(j) \\
&\qquad + \pi_0\{Q_{\mathrm{S}}^{\mathrm{M}}(0, \dots, 0, z_n, \dots, z_N)\} \\
&\qquad - Q_{\mathrm{S}}^{\mathrm{M}}(0, \dots, 0, z_{n+1}, \dots, z_N),
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
&\Pi_n^+(z_1, \dots, z_{n-1}, \tilde{0}, z_{n+1}, \dots, z_N) \\
&\quad := \sum_{i_1=0}^{\infty} \cdots \sum_{i_{n-1}=0}^{\infty} \sum_{i_{n+1}=0}^{\infty} \cdots \sum_{i_N=0}^{\infty} \pi_n^+ \\
&\qquad (i_1, \dots, i_{n-1}, 1, i_{n+1}, \dots, i_N) z_1^{i_1} \\
&\qquad \times \cdots z_{n-1}^{i_{n-1}} z_n z_{n+1}^{i_{n+1}} \cdots z_N^{i_N},
\end{aligned} \tag{13}$$

$$\begin{aligned}
&AD_1(j) \\
&\quad := \Pi_j^+(0, \dots, 0, z_j = \tilde{0}, 0, \dots, 0, z_n, \dots, z_N) \\
&\qquad \times \{p_j Q_{F_j}(0, \dots, 0, z_n, \dots, z_N) \\
&\qquad + \overline{p}_j Q_{\bar{F}_j}(0, \dots, 0, z_n, \dots, z_N)\}\frac{1}{z_j} \\
&\qquad - \Pi_j^+(0, \dots, 0, z_j = \tilde{0}, 0, \dots, 0, z_{n+1}, \dots, z_N) \\
&\qquad \times \{p_j Q_{F_j}(0, \dots, 0, z_{n+1}, \dots, z_N) \\
&\qquad \overline{p}_j Q_{\bar{F}_j}(0, \dots, 0, z_{n+1}, \dots, z_N)\}\frac{1}{z_j} \\
&\quad \text{for } 1 \le j \le n-1,
\end{aligned} \tag{14}$$

$$\begin{aligned}
&AD_2(j) \\
&\quad := \Pi_j^+(0, \dots, 0, z_n, \dots, z_N) \\
&\qquad \times \overline{p}_j Q_{\bar{F}_j}(0, \dots, 0, z_n, \dots, z_N)\frac{1}{z_j} \\
&\qquad - \Pi_j^+(0, \dots, 0, z_{n+1}, \dots, z_N) \\
&\qquad \times \overline{p}_j Q_{\bar{F}_j}(0, \dots, 0, z_{n+1}, \dots, z_N)\frac{1}{z_j} \\
&\quad \text{for } n \le j \le N,
\end{aligned}$$

$$\begin{aligned}
&AD_3(j) \\
&\quad := \Pi_j^+(0, \dots, 0, z_n, \dots, z_{j-1}, \tilde{0}, z_{j+1}, \dots, z_N) \\
&\qquad \times p_j Q_{F_j}(0, \dots, 0, z_n, \dots, z_j = 0, \dots, z_N)\frac{1}{z_j} \\
&\qquad - \Pi_j^+(0, \dots, 0, z_{n+1}, \dots, z_{j-1}, \tilde{0}, z_{j+1}, \dots, z_N) \\
&\qquad \times p_j Q_{F_j}(0, \dots, 0, z_{n+1}, \dots, z_j = 0, \dots, z_N)\frac{1}{z_j} \\
&\quad \text{for } n+1 \le j \le N,
\end{aligned}$$

$$\pi_0 := \sum_{j=1}^N \Pi_j(0, 0, \dots, 0)$$

$$= \sum_{j=1}^{N} \Pi_j^+(0,\ldots,0,z_j=\tilde{0},0,\ldots,0) \times Q_j(0,0,\ldots,0)\frac{1}{z_j}. \quad (15)$$

The first two terms on the right-hand side of (12) correspond to the case that a service is given to class-n, and when the service is completed, another class-n message is next served because of the service time without timer expiration (F_n with the LST $F_n^*(s)$). The terms $AD_k(j), k=1,2,3$ correspond to the case that a service is given to class-j, and when it is completed, the highest priority class of messages present in the system is class-n, i.e. to the three cases with (i) $1 \le j \le n-1$ and p_j and $\overline{p}_j$ (without and with a timer expiration), (ii) $n \le j \le N$ and $\overline{p}_j$ (with a timer expiration) and (iii) $n+1 \le j \le N$ and p_j (without a timer expiration), respectively. The last term on the right-hand side of (12) corresponds to the case that when a service is completed, there are no messages in the system and the highest priority class of messages in the system just after a vacation period is class-n. The minus terms on the right-hand side of (12) and the minus terms in $AD_k(j), k=1,2,3$ correspond to the case that the server has not any of class-n messages at the beginning of service. Note here that $\Pi_n^+(z_1,\ldots,z_{n-1},\tilde{0},z_{n+1},\ldots,z_N)/z_n$ is independent of z_n.

We obtain the following result from the steady-state functional relationship (12).

Theorem 1 *For $\overline{\boldsymbol{T}} = (\overline{T}_1, \overline{T}_2, \ldots, \overline{T}_N)$, $|z_n| \le 1$ and $n \in \{1,2,\ldots,N\}$,*

$$\Pi_n^+(z_1,z_2,\ldots,z_N) = \frac{z_n}{z_n - p_n Q_{F_n}(z_1,z_2,\ldots,z_N)} \times [\varphi_n(z_n,z_{n+1},\ldots,z_N) - \varphi_n(\beta_n(y_n),z_{n+1},\ldots,z_N)], \quad (16)$$

where

$$\varphi_n(z_n,z_{n+1},\ldots,z_N) := \sum_{j=1}^{n-1} AD_1(j) + \sum_{j=n}^{N} AD_2(j) + \sum_{j=n+1}^{N} AD_3(j) + \pi_0\{Q_S^M(0,\ldots,0,z_n,\ldots,z_N) - Q_S^M(0,\ldots,0,z_{n+1},\ldots,z_N)\}.$$

The y_n and $\beta_n(y)$ are given by (5) and (6) in Lemma 1, respectively.

$$\sum_{n=1}^{N} \frac{\Pi_n^+(z_1,z_2,\ldots,z_N)}{z_n}\{z_n - Q_n(z_1,z_2,\ldots,z_N)\} = \pi_0[Q_S^M(z_1,z_2,\ldots,z_N) - 1], \quad (17)$$

$$\pi_0 = \begin{cases} (1-\rho)\dfrac{1/\lambda}{v/(1-V^*(\lambda))} & \text{(MV)} \\ (1-\rho)\dfrac{1/\lambda}{v+V^*(\lambda)/\lambda} & \text{(SV)} \end{cases} \quad (18)$$

$$\Pi_n^+(1,1,\ldots,1) = \Pi_n(1,1,\ldots,1) = r_n. \quad (19)$$

Note that both sides of (17) are invariant for $\overline{\boldsymbol{T}} = (\overline{T}_1, \overline{T}_2, \ldots, \overline{T}_N)$.

Proof Equation (12) can be rewritten as

$$\Pi_n^+(z_1,z_2,\ldots,z_N)\left[1 - p_n Q_{F_n}(z_1,z_2,\ldots,z_N)\frac{1}{z_n}\right] + \Pi_n^+(z_1,\ldots,z_{n-1},\tilde{0},z_{n+1},\ldots,z_N) \times p_n Q_{F_n}(z_1,\ldots,z_{n-1},0,z_{n+1},\ldots,z_N)\frac{1}{z_n}$$
$$= \sum_{j=1}^{n-1} AD_1(j) + \sum_{j=n}^{N} AD_2(j) + \sum_{j=n+1}^{N} AD_3(j) + \pi_0\{Q_S^M(0,\ldots,0,z_n,\ldots,z_N) - Q_S^M(0,\ldots,0,z_{n+1},\ldots,z_N)\}$$
$$=: \varphi_n(z_n,z_{n+1},\ldots,z_N). \quad (20)$$

The right-hand side of (20) is a function of only $z_n, z_{n+1}, \ldots,$ and z_N, which is denoted by $\varphi_n(z_n,z_{n+1},\ldots,z_N)$. Putting $z_n = \beta_n(y_n)$ defined in Lemma 1 in the both sides of (20), the first term on the left-hand side of (20) equals zero, and the second term is unchanged because of the independency of z_n. That is, we get

$$\Pi_n^+(z_1,\ldots,z_{n-1},\tilde{0},z_{n+1},\ldots,z_N) \times p_n Q_{F_n}(z_1,\ldots,z_{n-1},0,z_{n+1},\ldots,z_N)\frac{1}{z_n} = \varphi_n(\beta_n(y_n),z_{n+1},\ldots,z_N). \quad (21)$$

Hence, the above statement leads to (16). Using (11) and (19), the invariant functional relationship (17) can be rewritten as

$$\sum_{n=1}^{N} \Pi_n^+(z_1,z_2,\ldots,z_N) = \sum_{n=1}^{N}\{\Pi_n(z_1,z_2,\ldots,z_N) - \Pi_n(0,0,\ldots,0)\} + \pi_0 Q_S^M(z_1,z_2,\ldots,z_N), \quad (22)$$

the both sides of which represent the PGF for the number of messages in each queue at the service beginning epochs of any class. The probability π_0 defined in (15) is determined using the normalization condition (10) and (12), see Remark 4.1. Equation (19) follows from (10), (11) and (12), see Remark 4.2. □

Theorem 2 *For* $\overline{\boldsymbol{T}} = \overline{T}_1, \overline{T}_2, \ldots, \overline{T}_N$, $|z_n| \le 1$ *and* $n \in \{1, 2, \ldots, N\}$,

$$\Pi_n(z_1, z_2, \ldots, z_N) = \frac{Q_n(z_1, z_2, \ldots, z_N)}{z_n - p_n Q_{F_n}(z_1, z_2, \ldots, z_N)} \times [\varphi_n(z_n, z_{n+1}, \ldots, z_N) - \varphi_n(\beta_n(y_n), z_{n+1}, \ldots, z_N)], \tag{23}$$

$$\sum_{n=1}^{N} \frac{z_n - Q_n(z_1, z_2, \ldots, z_N)}{z_n - p_n Q_{F_n}(z_1, z_2, \ldots, z_N)} \times [\varphi_n(z_n, z_{n+1}, \ldots, z_N) - \varphi_n(\beta_n(y_n), z_{n+1}, \ldots, z_N)] = \pi_0[Q_{\mathrm{S}}^{\mathrm{M}}(z_1, z_2, \ldots, z_N) - 1]. \tag{24}$$

Proof Eliminating $\Pi_n^+(z_1, z_2, \ldots, z_N)$ from (16) and (17) in Theorem 1 by using (11) and (19), we obtain the statement of Theorem 2. □

We know that (23) and (24) correspond to (12) and (15) in Katayama [11] and (8) and (10) in Katayama and Kobayashi [12], respectively. We have thus reached the same point with the previous results [11, 12]. For waiting time analysis in the next section, let us define the PGF for the number of class-n messages present in the system at service completion epochs of class-n messages as follows: For $n \in \{1, 2, \ldots, N\}$ and $|z| \le 1$,

$$L_n(z) := \frac{\Pi_n(1, \ldots, 1, z_n = z, 1, \ldots, 1)}{\Pi_n(1, 1, \ldots, 1)} = \frac{1}{r_n} \frac{H_n^*\{\lambda_n(1-z)\}}{z - p_n F_n^*\{\lambda_n(1-z)\}} \times [G_n(z) + r_n \overline{p}_n - G_n(1)], \tag{25}$$

where $G_n(z) := \varphi_n(z, 1, 1, \ldots, 1)$. For derivation of (25), we have used (19), (23) and the normalization condition of $L_n(z)$. That is, it follows from $L_n(1) = 1$ that the term $G_n[\beta_n(y_n)]$ after setting $z_1 = z_2 = \cdots = z_N = 1$ equals $G_n(1) - r_n \overline{p}_n$. Our remaining problem is thus the determination of the unknown function $G_n(z)$, see Remark 4.3. We have already the same type of functional equation for $G_n(z)$ and the solution in Katayama [11], following which the rest of this section is devoted to the determination of $G_n(z)$.

Theorem 3 *For* $n \in \{1, 2, \ldots, N\}$ *and* $|z| \le 1$,

$$G_k(z) - G_k[f_k(z)] = A_k(z) + \sum_{j=k+1}^{N} B_{j,k}(z), \quad k = n, n+1, \ldots, N. \tag{26}$$

The null sum ($\sum_{j=N+1}^{N}$) *is equal to zero and* $A_k(z)$, $B_{j,k}(z)$ *and* $f_k(z)$ *are defined as*

$$\begin{aligned} A_k(z) &:= \frac{z - H_k^*\{\alpha_k(z) + 1/\overline{T}_k\}}{z - H_k^*\{\alpha_k(z)\}} C_k(z), \\ B_{j,k}(z) &:= \frac{z - H_k^*\{\alpha_k(z) + 1/\overline{T}_k\}}{z - H_k^*\{\alpha_k(z)\}} \times \frac{1 - H_j^*\{\alpha_k(z)\}}{1 - H_j^*\{\alpha_k(z) + 1/\overline{T}_j\}} [G_j[\beta_j^k(z)] - G_j(1)], \\ f_k(z) &:= \beta_k\left[\lambda_{k-1}^+ - \gamma_k(z) + \frac{1}{\overline{T}_k}\right], \end{aligned} \tag{27}$$

where

$$C_k(z) := \begin{cases} \dfrac{1-\rho}{\rho_v}[V^*\{\alpha_k(z)\} - 1] & \text{(MV)} \\ \dfrac{1-\rho}{\rho_v + V^*(\lambda)} \times [V^*\{\alpha_k(z)\} - \alpha_k(z) V^*(\lambda)/\lambda - 1] & \text{(SV)} \end{cases}$$

$$\alpha_k(z) := \lambda_k^+ - \lambda_k z - \gamma_k(z), \qquad \beta_j^k(z) := \beta_j\left[\alpha_k(z) + \frac{1}{\overline{T}_j}\right], \quad j = n+1, n+2, \ldots, N. \tag{28}$$

The $\beta_k(y)$ *and* $\gamma_k(z)$ *are given in Lemma* 1.

Proof Let us consider a set $\{x_{n,k}; k = 1, 2, \ldots, n-1\}$ which makes null the terms on the left-hand side of (24), $z_k - Q_k(z_1, z_2, \ldots, z_N)$, $k = 1, 2, \ldots, n-1$. It is proved that these $n-1$ roots constitute a unique analytic solution for the set of $n-1$ equations, $z_k - Q_k(z_1, z_2, \ldots, z_N) = 0$ for $k = 1, 2, \ldots, n-1$ (Lemmas 5.1 and 5.2 in Kesten and Runnenburg [14] and Sect. 3.8 in Takagi [21]). By setting $z_k = x_{n,k}$, $k = 1, 2, \ldots, n-1$ in all factors on the both sides of (24), we have

$$\begin{aligned} &\frac{z_n - Q_n(x_{n,1}, \ldots, x_{n,n-1}, z_n, \ldots, z_N)}{z_n - p_n Q_{F_n}(x_{n,1}, \ldots, x_{n,n-1}, z_n, \ldots, z_N)} \times [\varphi_n(z_n, \ldots, z_N) - \varphi_n(\beta_n(y_n^n), z_{n+1}, \ldots, z_N)] \\ &+ \sum_{j=n+1}^{N} \frac{z_j - Q_j(x_{n,1}, \ldots, x_{n,n-1}, z_n, \ldots, z_N)}{z_j - p_j Q_{F_j}(x_{n,1}, \ldots, x_{n,n-1}, z_n, \ldots, z_N)} \times [\varphi_j(z_j, \ldots, z_N) - \varphi_j(\beta_j(y_j^n), z_{j+1}, \ldots, z_N)] \\ &= \pi_0[Q_{\mathrm{S}}^{\mathrm{M}}(x_{n,1}, \ldots, x_{n,n-1}, z_n, \ldots, z_N) - 1], \end{aligned} \tag{29}$$

where $y_j^n := 1/\overline{T}_j - \lambda_j(1-z_j) + \sum_{k=1}^{n-1}\lambda_k(1-x_{n,k}) + \sum_{k=n}^{N}\lambda_k(1-z_k)$, which is based on y_j in Lemma 1. Again, setting $z_n = z, z_{n+1} = z_{n+2} = \cdots = z_N = 1$ in all terms on the both sides of (29), we obtain the functional equation (26), where we have used $\gamma_n(z) = \sum_{k=1}^{n-1}\lambda_k x_{n,k}$, $(= x_n(z, 1, 1, \ldots, 1))$. This equation can be obtained as follows: The $x_{n,k}, k = 1, 2, \ldots, n-1$ is a function of $z_n, z_{n+1}, \ldots, z_N$, which is denoted by $x_{n,k}(z_n, z_{n+1}, \ldots, z_N)$. We introduce

$$\begin{aligned} x_n(z_n, z_{n+1}, \ldots, z_N) &:= \sum_{k=1}^{n-1}\lambda_k x_{n,k}(z_n, z_{n+1}, \ldots, z_N) \\ &= \sum_{k=1}^{n-1}\lambda_k H_k^*\left[\lambda_{n-1}^+\left\{1 - \frac{\sum_{k=1}^{n-1}\lambda_k x_{n,k}}{\lambda_{n-1}^+}\right\} + \sum_{k=n}^{N}\lambda_k(1-z_k)\right]. \end{aligned} \tag{30}$$

It is seen that (30) is identical with (7) in Lemma 1, putting again $z_n = x_n(z, 1, \ldots, 1)/\lambda_{n-1}^+ = \gamma_n(z)/\lambda_{n-1}^+$ after setting $z_n = z, z_{n+1} = z_{n+2} = \cdots = z_N = 1$ in the both sides of (30), since we have $\sum_{k=1}^{n-1}\lambda_k x_{n,k} = x_n(z_n, z_{n+1}, \ldots, z_N)$. □

Theorem 4 *For $k = n, n+1, \ldots, N$ and $0 \le z \le 1$,*

$$G_k(z) := \begin{cases} \sum_{i_k=0}^{\infty}\left[A_k(\sigma_k^{i_k}(z)) + \sum_{j=k+1}^{N} B_{j,k}(\sigma_k^{i_k}(z))\right] & \text{for } k \ge 2 \text{ and } \overline{T}_k \ne 0, \\ A_k(z) + \sum_{j=k+1}^{N} B_{j,k}(z) & \text{for } k = 1 \text{ or } \overline{T}_k = 0 (k \ge 2), \end{cases} \tag{31}$$

where the sequence $\{\sigma_k^i(z)\}$ is defined by

$$\begin{aligned} \sigma_k^0(z) &:= z, \\ \sigma_k^i(z) &:= \sigma_k^{i-1}[f_k(z)], \quad i = 1, 2, 3, \ldots, \quad 0 \le z \le 1, \end{aligned} \tag{32}$$

and $A_k(z), B_{j,k}(z), C_k(z), \alpha_k(z), \beta_j^k(z)$ and $f_k(z)$ are given in Theorem 3. The $\beta_k(y)$ and $\gamma_k(z)$ are given in Lemma 1.

Proof The functional equation for $G_k(z)$ given in Theorem 3 can be successively solved using the iterative procedure in Kuczma et al. [16], assuming that $G_{k+1}(z), G_{k+2}(z), \ldots, G_N(z)$ are known and starting from $k = N$. It should be noted that $f_k(z) = 0$ in the case of $p_k = 0$ because of $\beta_k(z) = 0$, and $f_1(z) = \beta_1(0)$. The functional equation should be solved in the following two cases, separately: (i) $f_k(z) =$ constant, i.e. $k = 1$ or $p_k = 0$ $(k \ge 2)$ and (ii) $f_k(z) \ne$ constant, i.e. $k \ge 2$ and $p_k \ne 0$. Hence the solution of (26) is given by (31) together with (32). □

Remark 4.1 The probability that the system is empty at an arbitrary epoch is identical with π_0 in the present priority queues due to the PASTA (Poisson arrivals see time averages) property and Finch's departure theorem [9] (also called Burke's theorem [21]). The decomposition property gives a probabilistic interpretation of (18) in Theorem 1 that the first factor, $1-\rho$, on the right-hand sides represents the probability that the system without vacations is empty at an arbitrary epoch, and the second factors represent the probabilities that the system is empty at an arbitrary epoch in vacation periods of the MV or SV rule, respectively, Boxma [3].

Remark 4.2 We have

$$\begin{aligned} r_n &:= \frac{\lambda_n}{\sum_{j=1}^{N}\lambda_j} = \frac{\Pi_n(1, 1, \ldots, 1)}{\sum_{j=1}^{N}\Pi_j(1, 1, \ldots, 1)} \\ &= \frac{\Pi_n^+(1, 1, \ldots, 1)}{\sum_{j=1}^{N}\Pi_j^+(1, 1, \ldots, 1)}, \end{aligned} \tag{33}$$

which means the rate balance of arrival and departure of class-n messages.

Remark 4.3 For the GF of the joint queue-length distribution $\Pi_n(z_1, z_2, \ldots, z_N)$ (not the marginal queue-length distribution, i.e., $L_n(z)$), we have to solve the functional equation,

$$\begin{aligned} &\varphi_n(z_n, z_{n+1}, \ldots, z_N) \\ &\quad - \varphi_n[f_n(z_n, z_{n+1}, \ldots, z_N), z_{n+1}, \ldots, z_N] \\ &\quad = g_n(z_n, z_{n+1}, \ldots, z_N), \end{aligned} \tag{34}$$

where $f_n(z_n, z_{n+1}, \ldots, z_N) := y_n^n$ defined in the proof of Theorem 3. $g_n(z_n, z_{n+1}, \ldots, z_N)$ is obtained from (24), however, the explicit expressions for $x_{n,k} = x_{n,k}(z_n, z_{n+1}, \ldots, z_N)$, $k = 1, 2, \ldots, n-1$ should be given, see the proof of Theorem 3.

5 Waiting time analysis

We denote by $W_n^*(s)$ the LST of the DF of the waiting time of a class-n message, $n = 1, 2, \ldots, N$. Then, we have the following equation for the FIFO order of service:

$$L_n(z) = W_n^*(\lambda_n(1-z))H_n^*(\lambda_n(1-z)), \tag{35}$$

since the number of class-n messages left behind by a departing class-n message is equal to the number of class-n messages that arrived while that message was in the system. Finally, we obtain the following result.

Theorem 5 *In the unsaturated priority queues, assured by $\rho < 1$, with exponential timer having the mean MAT $\overline{\boldsymbol{T}} =$*

$(\overline{T}_1, \overline{T}_2, \dots, \overline{T}_N)$ and the server vacation of MV or SV rule, $W_n^*(s)$, $n \in \{1, 2, \dots, N\}$ is given by

$$W_n^*(s) = \frac{\lambda}{s - \lambda_n + \lambda_n H_n^*(s + 1/\overline{T}_n)} \times [G_n(1) - G_n(1 - s/\lambda_n) - r_n \times \{1 - H_n^*(1/\overline{T}_n)\}]. \tag{36}$$

The $G_n(z)$ is given in Theorem 4. The probability of delay is given by

$$W_n^c(0) = \begin{cases} 1 & \text{(MV)}, \\ \dfrac{\rho_v + \rho V^*(\lambda)}{\rho_v + V^*(\lambda)} & \text{(SV)}. \end{cases} \tag{37}$$

Proof The LST $W_n^*(s)$ is derived from (25) and (35) together with Theorem 4. The probability $W_n^c(0)$ is obtained using the initial value theorem $W_n(0) = \lim_{s\to\infty} W_n^*(s)$. □

Corollary 1 *The LST $W_n^*(s)$ is independent of $\overline{T}_1, \overline{T}_2, \dots, \overline{T}_{n-1}$ and depends on only $\overline{T}_n, \overline{T}_{n+1}, \dots, \overline{T}_N$.*

Proof This statement follows from that $G_n(z)$ in (36) is independent of $p_k, k = 1, 2, \dots, n-1$, because $\beta_n(y)$ and $\gamma_n(z)$ defined in Lemma 1 are independent of $p_k, k = 1, 2, \dots, n-1$, and $G_n(z)$ is determined in the descent order starting from $k = N$. □

Corollary 2 *For the exhaustive service ($\overline{T}_n = \infty$), $W_n^*(s)$ is expressed by the following decomposition form of Fuhrmann and Cooper [10], cf. Theorem 4.1 in Miyazawa [19],*

$$W_n^*(s) = \frac{(1-\rho_n)s}{s - \lambda_n + \lambda_n H_n^*(s)} \left[\frac{\lambda \psi_n(1) \boldsymbol{E}(V_n)}{1-\rho_n} R_{V_n}^*(s) + \frac{\lambda \pi_0 \psi_{0n}(1) \boldsymbol{E}(V_{0n})}{1-\rho_n} R_{V_{0n}}^*(s) \right]. \tag{38}$$

The LST $R_X^*(s)$ appearing on the right-hand side of (38) is the LST of the DF of the forward (backward) recurrence time of a non-negative random variable X defined by

$$R_X^*(s) := \frac{1 - X^*(s)}{\boldsymbol{E}(X)s}, \quad V_n^*(s) := \frac{\psi_n(1 - s/\lambda_n)}{\psi_n(1)}, \quad V_{0n}^*(s) := \frac{\psi_{0n}(1 - s/\lambda_n)}{\psi_{0n}(1)}, \tag{39}$$

where

$$\psi_{0n}(z) := Q_{\mathrm{S}}^{\mathrm{M}}(0, \dots, 0, z_n = z, 1, \dots, 1) - Q_{\mathrm{S}}^{\mathrm{M}}(0, \dots, 0, z_n = 0, 1, \dots, 1),$$

$$\psi_n(z) := G_n(z) - \pi_0 \psi_{0n}(z).$$

Proof Putting $1/\overline{T}_n = 0$ in Theorem 5, (38) follows from (36). $G_n(z)$ is given by the sum of $\psi_n(z)$ and $\pi_0\psi_{0n}(z)$, that is, $\psi_n(z)$ has not the factor $Q_{\mathrm{S}}^{\mathrm{M}}(z_1, z_2, \dots, z_N)$, see Sect. 6. For the derivation of $V_n^*(s)$ and $V_{0n}^*(s)$, we have used $X^*(0) = 1$ because of the proper DF $X(t)$. □

Corollary 3 *For $0 \le \overline{T}_n < \infty$,*

$$\boldsymbol{E}(W_n) = \frac{1}{r_n \lambda_n \{1 - H_n^*(1/\overline{T}_n)\}} \times [G_n^{(1)}(1) - r_n(1 + \lambda_n p_n^{(1)})], \tag{40}$$

$$\boldsymbol{E}(W_n^2) = \frac{1}{r_n \lambda_n^2 \{1 - H_n^*(1/\overline{T}_n)\}^2} \times [2r_n(1 + \lambda_n p_n^{(1)})^2 + \lambda_n^2 p_n^{(2)} - 2(1 + \lambda_n p_n^{(1)}) G_n^{(1)}(1) + \{1 - H_n^*(1/\overline{T}_n)\} G_n^{(2)}(1)], \tag{41}$$

$$p_n^{(m)} := \frac{\mathrm{d}^m}{\mathrm{d}\alpha_n^m} H_n^*(\alpha_n), \quad m = 1, 2, \dots, \tag{42}$$

and for $\overline{T}_n = \infty$,

$$\boldsymbol{E}(W_n) = \frac{\lambda_n h_n^{(2)}}{2(1-\rho_n)} + \frac{G_n^{(2)}(1)}{2r_n\lambda_n(1-\rho_n)}, \tag{43}$$

$$\boldsymbol{E}(W_n^2) = \frac{\lambda_n h_n^{(3)}}{3(1-\rho_n)} + \frac{\lambda_n^2 (h_n^{(2)})^2}{2(1-\rho_n)^2} + \frac{h_n^{(2)}}{2r_n(1-\rho_n)^2} G_n^{(2)}(1) + \frac{G_n^{(3)}(1)}{3r_n\lambda_n(1-\rho_n)}, \tag{44}$$

where the first three derivatives of $G_n(z)$ at $z = 1$ are denoted by $G_n^{(1)}(1)$, $G_n^{(2)}(1)$ and $G_n^{(3)}(1)$, respectively.

Corollary 4 *For the exhaustive service priority queues ($\overline{\boldsymbol{T}} = \infty$) with the MV rule, $\boldsymbol{E}(W_n)$ is obtained from (43), where a set of recursive equations to calculate $G_k^{(2)}(1)$, $k = n, n+1, \dots, N$ is given by*

$$\begin{aligned}
&(1 - \rho_k^+)(1 - \rho_k^+ + 2\rho_k \rho_{k-1}^+) G_k^{(2)}(1) \\
&\quad = r_k \lambda_k \lambda_{k-1}^+ (1 - \rho_k)^2 h_{k-1}^{+(2)} + r_k \lambda_k^2 (\rho_{k-1}^+)^2 h_k^{(2)} \\
&\qquad + \lambda_k^2 (1 - \rho_k)^2 \sum_{j=k+1}^{N} \left\{ \frac{r_j h_j^{(2)}}{(1-\rho_j)^2} + \frac{h_j^2}{(1-\rho_j)^2} G_j^{(2)}(1) \right\} \\
&\qquad + 2 r_k \lambda_k (1 - \rho)(1 - \rho_k)^2 \frac{v^{(2)}}{2v}.
\end{aligned} \tag{45}$$

(The formulas for the SV rule (or no vacation model with $V = 0$) can be obtained from the above result using the replacement given in Remark 5.1, see Corollary 3 in [12] on the second moment $\boldsymbol{E}(W_n^2)$.)

Proof Setting $1/\overline{T}_k = 0$ in Theorem 3, we have

$$G_k(z) - G_k[f_k(z)] = \sum_{j=k+1}^{N} [G_j\{\beta_j^k(z)\} - G_j(1)] + \begin{cases} \dfrac{1-\rho}{\rho_v}[V^*\{\alpha_k(z)\} - 1] & \text{(MV)}, \\ \dfrac{1-\rho}{\rho_v + V^*(\lambda)} \times [V^*\{\alpha_k(z)\} - \alpha_k(z)V^*(\lambda)/\lambda - 1] & \text{(SV)}, \end{cases} \tag{46}$$

where $f_k(z), \alpha_k(z)$ and $\beta_j^k(z)$ are given in Theorem 3. The recursive equations (45) is obtained by differentiating both sides of (46). (Vanishing the term with $v^{(2)}/2v$ on the right-hand side of (45), the recursive equation for $G_k^{(2)}(1)$ is identical with (8.25a) in Sect. 3.8 [21], see Remark 5.1.) □

In two special cases of the 1-limited service or the exhaustive service for the lowest class messages, we get the following formulas independent of the mean MATs as in Corollary 1.

Corollary 5 *For $\overline{T}_1, \overline{T}_2, \ldots, \overline{T}_{N-1} \geq 0$ and $\overline{T}_N = 0$ or $\overline{T}_N = \infty$, $\boldsymbol{E}(W_N)$ is given by: For the 1-limited service ($\overline{T}_N = 0$) and the MV rule,*

$$\boldsymbol{E}(W_N) = \frac{\lambda h^{(2)}}{2(1-\rho)(1-\rho_{N-1}^+)} + \frac{1}{(1-\rho_{N-1}^+)}\frac{v^{(2)}}{2v}, \tag{47}$$

and for the exhaustive service ($\overline{T}_N = \infty$) and the MV rule,

$$\boldsymbol{E}(W_N) = \frac{\lambda_N h_N^{(2)}}{2(1-\rho_N)} + \frac{\lambda_{N-1}^+(1-\rho_N)^2 h_{N-1}^{+(2)} + \lambda_N(\rho_{N-1}^+)^2 h_N^{(2)}}{2(1-\rho_N)(1-\rho)(1-\rho+2\rho_N\rho_{N-1}^+)} + \frac{1-\rho_N}{1-\rho+2\rho_N\rho_{N-1}^+}\frac{v^{(2)}}{2v}, \tag{48}$$

see Corollary 4 in [12] on the second moment $\boldsymbol{E}(W_N^2)$.

Remark 5.1 For the present priority queues without server vacations ($V = 0$), we can obtain the moments of the waiting time distributions $\boldsymbol{E}(W_n^m)$ by vanishing the terms with $v^{(k)}/kv, k = 2, 3, \ldots$ from the equations given in Corollaries 3–5. Further we can obtain the moments for the single vacation model, $\boldsymbol{E}(W_n^m)_{\mathrm{SV}}$, $m = 1, 2, \ldots$ and the conservation law for the single vacation model corresponding to (50), using the following replacement, cf. Remark 4.2 in [12],

$$\left[\frac{v^{(k)}}{kv}\right]_{\mathrm{MV}} \leftrightarrow \left[\frac{v^{(k)}}{k\{v + V^*(\lambda)\}/\lambda}\right]_{\mathrm{SV}}, \quad k = 2, 3, \ldots. \tag{49}$$

Remark 5.2 We have the $M/G/1$ conservation law for the present time-limited service priority queues with $\overline{\boldsymbol{T}}$ and the MV rule,

$$\sum_{n=1}^{N} \rho_n \boldsymbol{E}(W_n)/\rho = \frac{\sum_{n=1}^{N}\lambda_n h_n^{(2)}}{2(1-\rho)} + \frac{v^{(2)}}{2v} \quad \text{(MV)}, \tag{50}$$

since the service discipline is *work-conserving* and non-preemptive for all n, Chap. 3 in [21].

6 Concluding remarks

As shown in Corollary 2, we consider an $M/G/1$ vacation model with arrival rate λ_n, service time H_n and two types of vacations, i.e., a nonzero-vacation time V_n and a zero-vacation (exhaustive vacation) time V_{0n}. For $0 \leq \overline{T}_n \leq \infty$, we also obtain an expression for the LST $W_n^*(s)$ corresponding to (38), using $D_n^*(s) := \bar{F}_n^*(s)W_n^*(s)$, $H_n^*(s) = p_n F_n^*(s) + \overline{p}_n \bar{F}_n^*(s)$ and (39). Taking the inverse LST after some rearrangement of (36) together with (37), we get the following level-crossing formula,

$$\begin{aligned} \frac{dW_n(x)}{dx} &= \lambda_n \int_0^x H_n^c(x-y)dW_n(y) \\ &\quad + \lambda\psi_n(1)V_n^c(x) - \lambda_n\overline{p}_n D_n^c(x) \\ &\quad + \lambda\pi_0\psi_{0n}(1)V_{0n}^c(x) - W_n(0)\delta(x-0), \\ x > 0, \end{aligned} \tag{51}$$

where $\delta(x)$ is Dirac's delta function. The left-hand and right-hand sides of (51) represent the down- and up-crossing rates at a level x, respectively, for a sample path W_t of the workload process $\{W_t\}$ [4, 6]. It is possible to give a probabilistic interpretation of each term on the right-hand side of (51). However, it should be noted that the expression for $W_n^*(s)$ contains itself, i.e., $W_n^*(s)$, in the minus term with $D_n^*(s)$ in the case of $\overline{p}_n \neq 0$, that is, we have not such decomposition form for $W_n^*(s)$ as in Corollary 2 for the non-exhaustive service ($\overline{T}_n \neq \infty$).

A direct derivation of the LST $W_n^*(s)$ (not through the queue-length analysis in this paper) may be a future research. For the generalization of the MAT instead of the exponential distribution, it seems to be rather effective to apply numerical approaches studied by Leung et al. [17, 18], who considered the case of *constant* MAT.

Acknowledgements The author would like to thank the anonymous referees for a careful reading of the original version of the paper.

References

1. Blanc, J.P.C., van der Mei, R.D.: Optimization of polling systems with Bernoulli schedules. Perform. Eval. **22**, 139–158 (1995)

2. Borst, S.C., Boxma, O.J., Levy, H.: The use of service limits for efficient operation of multistation single-medium communication systems. IEEE Trans. Netw. **3**, 602–612 (1995)
3. Boxma, O.J.: Workloads and waiting times in single-server systems with multiple customer classes. Queueing Syst. **5**, 185–214 (1989)
4. Brill, P.H.: Level crossing methods. In: Gass, S.I., Harris, C.M. (eds.) Encyclopedia of Operations Research and Management Science, Centennial edn. pp. 448–450. Kluwer Academic, Dordrecht (2001)
5. Coffman, E.G., Fayolle, G., Mitrani, I.: Two queues with alternating service periods. In: Courtois, P.-J., Latouche, G. (eds.) Performance '87, pp. 227–239. Elsevier/North-Holland, Amsterdam (1987)
6. Cohen, J.W.: On up- and downcrossings. J. Appl. Probab. **14**, 405–410 (1977)
7. Eliazar, I., Fibich, G., Yechiali, U.: A communication multiplexer problem: two alternating queues with dependent randomly-timed gated regime. Queueing Syst. **42**, 325–353 (2002)
8. Feng, W., Kowada, M., Adachi, K.: A two-queue model with Bernoulli service schedule and switching times. Queueing Syst. **30**, 405–434 (1998)
9. Finch, P.D.: On the distribution of queue size in queueing problems. Acta Math. Sci. Acad. Hung. **10**, 327–336 (1959)
10. Fuhrmann, S.W., Cooper, R.B.: Stochastic decompositions in the $M/G/1$ queue with generalized vacations. Oper. Res. **33**, 1117–1129 (1985)
11. Katayama, T.: Priority queues with Bernoulli schedules. In: Ramaswami, V., Wirth, P.E. (eds.) Teletraffic Contributions for the Information Age, pp. 309–318. Elsevier, Amsterdam (1997)
12. Katayama, T., Kobayashi, K.: Analysis of a nonpreemptive priority queue with exponential timer and server vacations. Perform. Eval. **64**, 495–506 (2007)
13. Katayama, T., Kobayashi, K., Nakagawa, K.: System delay analysis for a nonpreemptive time-limited service queueing system with two types of vacations and its application. In: Charzinski, J., Lehnert, R., Tran-Gia, P. (eds.) Providing Quality of Service in Heterogeneous Environments, pp. 591–600. Elsevier, Amsterdam (2003)
14. Kesten, H., Runnenburg, J.Th.: Priority in waiting-line problems, I and II. Proc. Koninkl. Nederl. Akad. Wetensch. Ser. A **60**, 312–336 (1957)
15. Kleinrock, L., Finkelstein, R.P.: Time dependent priority queues. Oper. Res. **15**, 104–116 (1967)
16. Kuczma, M., Choczewski, B., Ger, R.: Iterative Functional Equations. Encyclopedia of Mathematics and its Applications, vol. 32. Cambridge University Press, Cambridge (1990)
17. Leung, K.K., Eisenberg, M.: A single-server queue with vacations and non-gated time-limited service. Perform. Eval. **12**, 115–125 (1991)
18. Leung, K.K., Lucantoni, D.M.: Two vacation models for token-ring networks where service is controlled by timers. Perform. Eval. **20**, 165–184 (1994)
19. Miyazawa, M.: Decomposition formulas for single server queues with vacations: a unified approach by the rate conservation law. Stoch. Models **10**, 389–413 (1994)
20. Takács, L.: Introduction to the Theory of Queues. Oxford University Press, New York (1962)
21. Takagi, H.: Queueing Analysis, Vol. 1: A Foundation of Performance Evaluation. North-Holland, Amsterdam (1991)
22. Takagi, H.: Queueing analysis of polling models: Progress in 1990–1994. In: Dshalalow, J.H. (eds.) Frontiers in Queueing, pp. 119–146. CRC Press, Boca Raton (1997)