Queueing Systems 39, 109–135, 2001


# Departure Processes of BMAP/G/1 Queues *

HUEI-WEN FERNG wen@santos.ee.ntu.edu.tw
*Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan*

JIN-FU CHANG jfchang@ncnu.edu.tw
*National Chi-Nan University, Puli, Nantou 545, Taiwan*

Received 20 April 2000; Revised 1 June 2001

**Abstract.** A unified approach is applied to analyze the departure processes of finite/infinite BMAP/G/1 queueing systems for both vacationless and vacation arrangements via characterizing the moments, the $z$-transform of the scaled autocovariance function of interdeparture times $C_P(z)$, and lag $n$ ($n \geqslant 1$) covariance of interdeparture times. From a structural point of view, knowing departure process helps one to understand the impact of service mechanisms on arrivals. Through numerical experiments, we investigate and discuss how the departure statistics are affected by service and vacation distributions as well as the system capacity. From a practical perspective, output process analysis serves to bridge the nodal performance and connectionwise performance. Our results can be then used to facilitate connection- or networkwise performance analysis in the current high-speed networks.

**Keywords:** departure process, BMAP/G/1, vacation

## 1. Introduction

Network technology has been progressing in a very rapid pace in recent years. Connection- or networkwise rather than nodal performance analysis is currently receiving a great deal of attention. Although Jacksonian and BCMP [9] queueing networks have been thoroughly studied, the Poisson arrival assumption frequently employed in the traditional teletraffic analysis is not powerful enough to capture the correlative and bursty feature of traffic streams in the contemporary/future high-speed packet/cell-based networks, e.g., the asynchronous transfer mode (ATM) [1,17] networks. More sophisticated traffic models are desired. To address this need, the Markov-modulated Poisson process (MMPP) [7] has been extensively used to model multimedia traffic. Lucantoni [15] proposed a versatile batch Markovian arrival process (BMAP) which includes the Markovian arrival process (MAP) [15,16] and the MMPP as special cases and showed that the BMAP is essentially equivalent to the $N$-process, a rich class of point processes introduced by Neuts [18]. Comparing with the $N$-process, Lucantoni [15] demonstrated

* This work was done when the authors were with Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan.

that the BMAP gives neater results and more efficient algorithms. Thus, we decide to use the BMAP as the input process in this paper.

Exact connectionwise queueing analysis involving BMAP inputs is intractable. Therefore, an approximate approach is inevitable. Bearing the efficiency concern in mind, one viable method is to recursively perform the nodal performance analysis together with the output process analysis. This justifies the need of the output process analysis. Besides this practical motivation, knowing the departure process helps one to understand the impact imposed by the service mechanisms on the arrival streams. In the literature, Burke [3] first showed that the output of an $M/M/m$ queue is again a Poisson process. Later Takács [21] derived the Laplace–Stieltjes transform (LST) of the interdeparture time distribution of the stationary $M/G/1$ queue. Daley [4] wrote a general survey on the departure process of queueing systems. All the above works concentrated on the renewal type arrival processes, specifically, the Poisson family. For non-renewal or correlated arrivals, Saito [20] has acquired the departure process of a finite $N/G/1$ queue using the first passage technique [15]. Results obtained by Saito in [20] include the mean and variance of interdeparture times of an $N/G/1$ queue and the $C_P(z)$ (to be defined in a later section) and lag 1 covariance of interdeparture times of the $N/D/1$ queue. Recently, Yeh [22] derived the mean, variance, and lag 1 covariance of interdeparture times for the (infinite) $MAP/G/1$ queue that allows server vacations to take place. Neither [20] nor [22] has acquired the lag $n$ ($n \geqslant 2$) covariance of interdeparture times. In [23], Yeh and Chang have developed a recursive/iterative formula/method for the lag $n$ ($n \geqslant 1$) covariance of (infinite) $M/G/1$ type [19] queues. In this paper, we consider a finite/infinite $BMAP/G/1$ queue and treat both the vacationless and vacation service disciplines using a unified approach similar to that employed in our previous work [6] done for discrete-time queues. The vacationless case is equivalent to the finite $N/G/1$ queue employed by Saito [20]. But we obtain more comprehensive results including the moments of any order, $C_P(z)$, and lag $n$ ($n \geqslant 1$) covariance under a general service distribution rather than restricting to the deterministic server of [20]. Moreover, we derive the results for the finite/infinite $BMAP/G/1$ queue with server vacations which includes the $MAP/G/1$ queue with server vacations used by Yeh [22] as a special case. In [22], Yeh derived results for an infinite queue, but we provide results for both finite and infinite queues. Therefore, we provide results more general than Saito [20] and Yeh [22]. Unlike [23], we provide a direct and efficient matrix representation for the lag $n$ ($n \geqslant 1$) covariance for both finite and infinite capacity queues. Hence, our results save considerable computation time than [23].

In section 2, we first review the definition of the BMAP and present the queueing models considered in this paper. In section 3, using the preliminaries prepared in section 3.1, the departure process is analyzed and validated. Section 4 discusses different perspectives of the departure statistics through numerical examples. Finally, section 5 concludes the paper.

## 2. Traffic and system models

### 2.1. The batch Markovian arrival process (BMAP)

The definition and notation in [15] related to the BMAP are briefly reviewed here for readers' convenience.

Let $N(t)$ and $J(t)$ be the number of arrivals in $(0, t]$ and the phase at time $t$, respectively. Then $\{N(t), J(t)\}$ forms a 2-dimensional Markov process on the state space $\{(i, j) \mid i \geqslant 0, 1 \leqslant j \leqslant m\}$ with an infinitesimal generator

$$Q^* = \begin{bmatrix} \boldsymbol{D}_0 & \boldsymbol{D}_1 & \boldsymbol{D}_2 & \dots \\ \mathbf{0} & \boldsymbol{D}_0 & \boldsymbol{D}_1 & \dots \\ \mathbf{0} & \mathbf{0} & \boldsymbol{D}_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{1}$$

where $\mathbf{0}$ is an $m \times m$ zero matrix and $\boldsymbol{D}_k$, $k \geqslant 0$, are $m \times m$ matrices in which $\boldsymbol{D}_0$ (governing the transitions that correspond to no arrival) is a stable matrix implying that $\boldsymbol{D}_0$ is nonsingular with nonnegative off-diagonal elements, $\boldsymbol{D}_k$, $k \geqslant 1$ (governing the transitions corresponding to arrival batches of size $k$) are nonnegative matrices; and their sum $\boldsymbol{D} = \sum_{k=0}^{\infty} \boldsymbol{D}_k$ is an irreducible infinitesimal generator of the underlying Markov process of states. The above constructed 2-dimensional Markov process then defines a batch arrival process, e.g., transitions from state $(i, j)$ to state $(i+k, l)$, $k \geqslant 1$, $1 \leqslant j, l \leqslant m$, indicate $k$-batch arrivals. Now let

$$P_{ij}(n, t) = \Pr\bigl\{N(t) = n,\ J(t) = j \mid N(0) = 0,\ J(0) = i\bigr\} \tag{2}$$

be the $(i, j)$th element of an $m \times m$ matrix $\boldsymbol{P}(n, t)$. It is shown in [15] that the matrix generating function $\boldsymbol{P}^*(z, t)$ of $\boldsymbol{P}(n, t)$, i.e., $\boldsymbol{P}^*(z, t) = \sum_{n=0}^{\infty} \boldsymbol{P}(n, t) z^n$, has the following form

$$\boldsymbol{P}^*(z, t) = \mathrm{e}^{\boldsymbol{D}(z)t}, \tag{3}$$

where $\boldsymbol{D}(z)$ is the matrix generating function of $\boldsymbol{D}_k$, i.e., $\boldsymbol{D}(z) = \sum_{k=0}^{\infty} \boldsymbol{D}_k z^k$.

As mentioned earlier, the BMAP contains many familiar arrival processes such as MAP, MMPP, and $PH$-renewal processes as special cases.

### 2.2. System model

We first consider a first-in-first-out (FIFO) single server queue receiving a BMAP. The system capacity is assumed to be $K$ ($K \geqslant 1$), i.e., the buffer size is $K - 1$. A customer finding that system size equals $K$ (or queue length $K - 1$) on its arrival gets rejected. Since batch arrivals are considered, the partial batch acceptance policy is adopted in this paper when the buffer overflows due to a batch arrival. The service times of different customers are assumed to be independent and identically distributed (i.i.d.). Let $\widetilde{H}(x)$, $\bar{h}$, and $H(s)$ represent the cumulative distribution function (c.d.f.) of the service time, the mean service time, and the corresponding LST, respectively. The model described above is a vacationless BMAP/G/1/$K$ queue. In the literature, the $\cdot$/G/1/$K$ queues are frequently used to model an ATM multiplexer or an output port of an ATM switch [13]. Note that the case of $K = \infty$ is the case of infinite BMAP/G/1.

Taking server vacations into consideration creates another class of queueing systems. The vacation models have found many applications in communication networks. For example, they can be applied to assess the performance of the media access control (MAC) protocols in the token rings and satellite networks [2]; as suggested by Lucantoni [16], they may be employed as an effective method to solve queueing systems with multiclass arrivals of different priorities. Other examples can be found in [5] which gives an excellent survey on vacation models. In this paper, we concentrate on the *exhaustive service with multiple vacations*, i.e., a vacation begins when the system becomes idle and the server may take a repeated number of vacations if on its return from a vacation an empty system is seen again. We assume that vacation periods are i.i.d. and follow the c.d.f. $\widetilde{V}(x)$ with finite mean $\bar{v}$ and the corresponding LST $V(s)$.

## 3. Departure process analysis

### *3.1. Preliminaries*

Let $\{\tau_n\colon\ n \geqslant 0\}$ denote the successive departure epochs with $\tau_0 \equiv 0$ and further define $X_n$ and $J_n$ to be the queue length and the state of the arrival process at time $\tau_n^+$. Then it can be easily seen that $\{(X_n, J_n, \tau_{n+1} - \tau_n)\colon\ n \geqslant 0\}$ (for both vacationless and vacation BMAP/G/1/$K$ queues) forms a semi-Markov sequence at the departure epochs with state-space $\{0, 1, \ldots, K-1\} \times \{1, \ldots, m\}$ and the state transition probability matrix

$$\widetilde{\boldsymbol{Q}}(x) = \begin{bmatrix} \widetilde{\boldsymbol{B}}_0(x) & \widetilde{\boldsymbol{B}}_1(x) & \ldots & \widetilde{\boldsymbol{B}}_{K-2}(x) & \sum_{n=K-1}^{\infty} \widetilde{\boldsymbol{B}}_n(x) \\ \widetilde{\boldsymbol{A}}_0(x) & \widetilde{\boldsymbol{A}}_1(x) & \ldots & \widetilde{\boldsymbol{A}}_{K-2}(x) & \sum_{n=K-1}^{\infty} \widetilde{\boldsymbol{A}}_n(x) \\ \boldsymbol{0} & \widetilde{\boldsymbol{A}}_0(x) & \ldots & \widetilde{\boldsymbol{A}}_{K-3}(x) & \sum_{n=K-2}^{\infty} \widetilde{\boldsymbol{A}}_n(x) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \widetilde{\boldsymbol{A}}_0(x) & \sum_{n=1}^{\infty} \widetilde{\boldsymbol{A}}_n(x) \end{bmatrix}, \quad k \geqslant 1, \tag{4}$$

where $\widetilde{\boldsymbol{A}}_n(x)$ and $\widetilde{\boldsymbol{B}}_n(x)$, $n \geqslant 0$, $x \geqslant 0$, are defined as follows:

$[\widetilde{\boldsymbol{A}}_n(x)]_{ij}$ = Pr{given a departure at time 0, which left at least one customer in the system and the arrival process in state $i$, the next departure occurs no later than time $x$ with the arrival process in state $j$, and during that service there were $n$ customers},

$[\widetilde{\boldsymbol{B}}_n(x)]_{ij}$ = Pr{given a departure at time 0, which left an empty system and the arrival process in state $i$, the next departure occurs no later than time $x$ with the arrival process in state $j$, and in the mean time there were $n + 1$ customers}.

where $[\cdot]_{ij}$ is used to represent the $(i, j)$th element of the corresponding matrix. For a vacationless BMAP/G/1/$K$ queue, $\widetilde{\boldsymbol{B}}_n(x)$, $n \geqslant 0$, only depend on the arrival process and $\widetilde{\boldsymbol{A}}_n(x)$, $n \geqslant 0$ (see a later section). But for a vacation permissible queue, $\widetilde{\boldsymbol{B}}_n(x)$, $n \geqslant 0$, further depend on the vacation periods. Therefore, we need to define the following additional conditional probabilities:

$\big[\widetilde{\boldsymbol{V}}_n(x)\big]_{ij}$ = Pr{given that a vacation began at time 0, with the arrival process in state $i$, the vacation ends no later than time $x$ with the arrival process in state $j$, and during the vacation period there were $n$ customers}.

According to the definition of $\boldsymbol{P}(n,t)$, it is clear that

$$\widetilde{\boldsymbol{A}}_n(x) = \int_0^x \boldsymbol{P}(n,t)\,\mathrm{d}\widetilde{H}(t), \tag{5}$$

$$\widetilde{\boldsymbol{V}}_n(x) = \int_0^x \boldsymbol{P}(n,t)\,\mathrm{d}\widetilde{V}(t). \tag{6}$$

Let $\boldsymbol{A}_n(s)$, $\boldsymbol{B}_n(s)$, $\boldsymbol{V}_n(s)$, and $\boldsymbol{Q}(s)$ represent the matrix transforms of $\widetilde{\boldsymbol{A}}_n(\cdot)$, $\widetilde{\boldsymbol{B}}_n(\cdot)$, $\widetilde{\boldsymbol{V}}_n(\cdot)$, and $\widetilde{\boldsymbol{Q}}(\cdot)$, respectively, i.e.,

$$\boldsymbol{A}_n(s) = \int_0^\infty \mathrm{e}^{-sx}\,\mathrm{d}\widetilde{\boldsymbol{A}}_n(x), \qquad \boldsymbol{B}_n(s) = \int_0^\infty \mathrm{e}^{-sx}\,\mathrm{d}\widetilde{\boldsymbol{B}}_n(x),$$
$$\boldsymbol{V}_n(s) = \int_0^\infty \mathrm{e}^{-sx}\,\mathrm{d}\widetilde{\boldsymbol{V}}_n(x), \quad \text{and} \quad \boldsymbol{Q}(s) = \int_0^\infty \mathrm{e}^{-sx}\,\mathrm{d}\widetilde{\boldsymbol{Q}}(x).$$

Also let

$$\boldsymbol{A}(z,s) = \sum_{n=0}^\infty \boldsymbol{A}_n(s)z^n, \qquad \boldsymbol{B}(z,s) = \sum_{n=0}^\infty \boldsymbol{B}_n(s)z^n, \quad \text{and} \quad \boldsymbol{V}(z,s) = \sum_{n=0}^\infty \boldsymbol{V}_n(s)z^n.$$

For convenience, we set $\boldsymbol{A}(s) = \boldsymbol{A}(1,s)$, $\boldsymbol{B}(s) = \boldsymbol{B}(1,s)$, and $\boldsymbol{V}(s) = \boldsymbol{V}(1,s)$. Applying the LST to (4) yields

$$\boldsymbol{Q}(s) = \begin{bmatrix} \boldsymbol{B}_0(s) & \boldsymbol{B}_1(s) & \dots & \boldsymbol{B}_{K-2}(s) & \sum_{n=K-1}^\infty \boldsymbol{B}_n(s) \\ \boldsymbol{A}_0(s) & \boldsymbol{A}_1(s) & \dots & \boldsymbol{A}_{K-2}(s) & \sum_{n=K-1}^\infty \boldsymbol{A}_n(s) \\ \boldsymbol{0} & \boldsymbol{A}_0(s) & \dots & \boldsymbol{A}_{K-3}(s) & \sum_{n=K-2}^\infty \boldsymbol{A}_n(s) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{A}_0(s) & \sum_{n=1}^\infty \boldsymbol{A}_n(s) \end{bmatrix}. \tag{7}$$

In the following, we set $\boldsymbol{A}_i = \boldsymbol{A}_i(0)$, $\boldsymbol{B}_i = \boldsymbol{B}_i(0)$, $\boldsymbol{V}_i = \boldsymbol{V}_i(0)$, and $\boldsymbol{Q} = \boldsymbol{Q}(0)$ for notational simplicity. The stationary vector $\boldsymbol{x}$ of the Markov chain $\boldsymbol{Q}$ represents the joint probability of the stationary queue length and the state of the arrival process. The vector $\boldsymbol{x}$ can be solved via $\boldsymbol{x}\boldsymbol{Q} = \boldsymbol{x}$ and $\boldsymbol{x}\hat{\boldsymbol{e}} = 1$ where $\hat{\boldsymbol{e}}$ is an $mK \times 1$ column vector of all ones and can be partitioned into the form $\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{K-1})$ where $\boldsymbol{x}_i = (x_{i,1}, \dots, x_{i,j}, \dots, x_{i,m})$ with $x_{i,j}$ = Pr{there are $i$ customers in the system right behind a departure epoch, and the state of the arrival process is in state $j$}. Grassmann et al. [8] proposed the (block) GTH algorithm using state-reduction technique [10] to obtain $\boldsymbol{x}$ for the case of finite $K$. For $K = \infty$, $\boldsymbol{x}$ can be solved using algorithms provided in [15,16]

([15] for vacationless while [16][1] for vacation queues). In the following lemma, we give several useful identities.

**Lemma 1.** The matrices $\widetilde{\boldsymbol{A}}_n(x)$, $\boldsymbol{A}(s)$, $\widetilde{\boldsymbol{V}}_n(x)$, and $\boldsymbol{V}(s)$ satisfy the following identities

$$\sum_{n=0}^{\infty} \widetilde{\boldsymbol{A}}_n(x)\boldsymbol{e} = \widetilde{H}(x)\boldsymbol{e}, \tag{8}$$

$$\boldsymbol{A}(s)\boldsymbol{e} = H(s)\boldsymbol{e}, \tag{9}$$

$$\sum_{n=0}^{\infty} \widetilde{\boldsymbol{V}}_n(x)\boldsymbol{e} = \widetilde{V}(x)\boldsymbol{e}, \tag{10}$$

$$\boldsymbol{V}(s)\boldsymbol{e} = V(s)\boldsymbol{e}, \tag{11}$$

where $\boldsymbol{e}$ is an $m \times 1$ column vector of all ones.

*Proof.* Using (5), (6), and $\sum_{n=0}^{\infty} \boldsymbol{P}(n,t)\boldsymbol{e} = \boldsymbol{e}$ from the law of total probability, one can easily prove the lemma. □

Define $\widetilde{D}^{[n]}(x_1, x_2, \ldots, x_n)$ and $D^{[n]}(s_1, s_2, \ldots, s_n)$ to be the joint distribution of $n$ successive interdeparture times $T_{D,i}$ $(1 \leqslant i \leqslant n)$ and the corresponding $n$-dimensional LST, respectively. And let $D^{[1]}(s_i) = D^{[n]}(0, \ldots, s_i, \ldots, 0)$ represent the LST of the marginal c.d.f. $\widetilde{D}^{[1]}(x_i) = \widetilde{D}^{[n]}(\infty, \ldots, x_i, \ldots, \infty)$ of the interdeparture time. Following the philosophy similar to our previous work [6], we derive the departure process of BMAP/G/1(/$K$) queues with or without server vacations.

Utilizing the Markovian property of $\widetilde{\boldsymbol{Q}}(\cdot)$ makes

$$\widetilde{D}^{[n]}(x_1, x_2, \ldots, x_n) = \boldsymbol{x}\, \widetilde{\boldsymbol{Q}}(x_1) * \widetilde{\boldsymbol{Q}}(x_2) * \cdots * \widetilde{\boldsymbol{Q}}(x_n)\hat{\boldsymbol{e}}, \tag{12}$$

$$D^{[n]}(s_1, s_2, \ldots, s_n) = \boldsymbol{x} \prod_{i=1}^{n} \boldsymbol{Q}(s_{n-i+1})\hat{\boldsymbol{e}}, \tag{13}$$

where the operator $*$ denotes the matrix convolution and $\prod$ follows the left multiplication here and in the sequel, that is, $\prod_{i=1}^{n} \boldsymbol{Q}(s_i) = \boldsymbol{Q}(s_n) \cdots \boldsymbol{Q}(s_1)$. Using (13), applying $\boldsymbol{x}\boldsymbol{Q} = \boldsymbol{x}$ together with $\boldsymbol{Q}\hat{\boldsymbol{e}} = \hat{\boldsymbol{e}}$, and utilizing lemma 1, one obtain

$$D^{[1]}(s) = (1 - \boldsymbol{x}_0\boldsymbol{e})H(s) + \boldsymbol{x}_0\boldsymbol{B}(s)\boldsymbol{e}. \tag{14}$$

Differentiating (13) with respect to $s_1$ and $s_n$, and setting $s_j = 0$ $(1 \leqslant j \leqslant n)$ further obtain

$$E\{T_{D,1}T_{D,n}\} = \frac{\partial^2}{\partial s_1 \partial s_n} D^{[n]}(s_1, \ldots, s_n)\Big|_{s_j=0, 1\leqslant j\leqslant n} = \boldsymbol{x}\, \boldsymbol{Q}^{(1)}\, \boldsymbol{Q}^{n-2}\, \boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}} \tag{15}$$

[1] Lucantoni et al. [16] provides results only for an MAP/G/1 queue; but the results can be easily extended to a BMAP/G/1 queue.

which yields

$$\mathrm{Cov}(T_{D,1}, T_{D,n}) = \boldsymbol{x}\,\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{n-2}\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}} - \big[\boldsymbol{x}\,\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}}\big]^2, \quad n \geqslant 2, \tag{16}$$

where $\boldsymbol{Q}^{(k)} \equiv (\mathrm{d}^k/\mathrm{d}s^k)\,\boldsymbol{Q}(s)|_{s=0}$. From (7) and lemma 1,

$$\boldsymbol{Q}^{(k)}\hat{\boldsymbol{e}} = \big[\big[\boldsymbol{B}^{(k)}\boldsymbol{e}\big]^{\mathrm{T}} \quad H^{(k)}(0)\boldsymbol{e}^{\mathrm{T}} \quad \dots \quad H^{(k)}(0)\boldsymbol{e}^{\mathrm{T}}\big]^{\mathrm{T}}, \tag{17}$$

where $\boldsymbol{B}^{(k)} \equiv (\mathrm{d}^k/\mathrm{d}s^k)\boldsymbol{B}(s)|_{s=0}$, $H^{(k)}(0) \equiv (\mathrm{d}^k/\mathrm{d}s^k)H(s)|_{s=0}$ and $^{\mathrm{T}}$ represents the matrix transpose. Now define $c_n \triangleq \mathrm{Cov}(T_{D,i}, T_{D,i+n})$, $n \geqslant 0$, to be the lag $n$ covariance of $T_{D,i}$ and $T_{D,i+n}$ or the autocovariance function (sequence) of interdeparture times. Then

$$c_0 = E\big\{T_{D,i}^2\big\} - E^2\{T_{D,i}\} = \boldsymbol{x}\,\boldsymbol{Q}^{(2)}\hat{\boldsymbol{e}} - \big[\boldsymbol{x}\,\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}}\big]^2, \tag{18}$$

$$c_n = \boldsymbol{x}\,\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{n-1}\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}} - \big[\boldsymbol{x}\,\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}}\big]^2, \quad n \geqslant 1. \tag{19}$$

Further define $C_P(z)$, the $z$-transform of the autocovariance function of interdeparture times scaled by $E^2\{T_{D,i}\}$, as follows:

$$C_P(z) \triangleq \frac{\sum_{n=0}^{\infty} c_n z^n}{E^2\{T_{D,i}\}}. \tag{20}$$

Using (18)–(20), we obtain

$$C_P(z) = \frac{\boldsymbol{x}\,\boldsymbol{Q}^{(2)}\hat{\boldsymbol{e}} + z\boldsymbol{x}\,\boldsymbol{Q}^{(1)}[\boldsymbol{I} - z\,\boldsymbol{Q}]^{-1}\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}}}{[\boldsymbol{x}\,\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}}]^2} - (1-z)^{-1}. \tag{21}$$

Note that $C_P(z)$ is the squared coefficient of variation for the renewal processes since $c_n = 0$ for $n \geqslant 1$. In particular, $C_P(z) = 1$ for the Poisson process. Therefore, $C_P(z)$ can be seen as an index to indicate the burstiness of a traffic stream [20].

### 3.2. *Departure process of a vacationless queue*

For a closer examination of the departure process, we must find $\boldsymbol{B}(z, s)$. In [15], Lucantoni has provided the following results (theorem 1 and corollary 1.1) for the BMAP/G/1 queue. For the BMAP/G/1/$K$ queue, these results still hold because the definition of $\widetilde{\boldsymbol{B}}_n(x)$ is the same for both BMAP/G/1 and BMAP/G/1/$K$ queues.

**Theorem 1.** For the vacationless BMAP/G/1 and BMAP/G/1/$K$ queues, we have

$$\boldsymbol{B}(z, s) = z^{-1}(s\boldsymbol{I} - \boldsymbol{D}_0)^{-1}\big[\boldsymbol{D}(z) - \boldsymbol{D}_0\big]\boldsymbol{A}(z, s). \tag{22}$$

**Corollary 1.1.** The matrices $\boldsymbol{B}_n(s)$, $n \geqslant 0$, satisfy

$$\boldsymbol{B}_n(s) = (s\boldsymbol{I} - \boldsymbol{D}_0)^{-1}\sum_{j=0}^{n}\boldsymbol{D}_{j+1}\boldsymbol{A}_{n-j}(s). \tag{23}$$

**Corollary 1.2.**

$$\boldsymbol{B}(s)\boldsymbol{e} = H(s)\boldsymbol{e} - sH(s)(s\boldsymbol{I} - \boldsymbol{D}_0)^{-1}\boldsymbol{e}. \tag{24}$$

We omit the proof of corollary 1.2, but note that $\boldsymbol{De} = \mathbf{0}$ must be employed in the proof where $\mathbf{0}$ represents an $m \times 1$ column vector of all zeros.

Let us first derive the $n$th moment of the interdeparture time. From (14) and (24), we obtain the LST of the interdeparture time as follows:

$$D^{[1]}(s) = H(s) - sH(s)\boldsymbol{x}_0(s\boldsymbol{I} - \boldsymbol{D}_0)^{-1}\boldsymbol{e}. \tag{25}$$

Directly differentiate (25) $n$ times, set $s = 0$, and multiply by $(-1)^n$ to obtain

$$\begin{aligned} E\{T_{D,i}^n\} &= (-1)^n \left\{ H^{(n)}(0) + \sum_{j=0}^{n} \binom{n}{j} H^{(j)}(0)\boldsymbol{x}_0 \frac{d^{n-j}}{ds^{n-j}} \left[ -s(s\boldsymbol{I} - \boldsymbol{D}_0)^{-1} \right]\Big|_{s=0} \boldsymbol{e} \right\} \\ &= (-1)^n \left[ \sum_{j=0}^{n-1} \frac{n!}{j!} H^{(j)}(0)\boldsymbol{x}_0 \boldsymbol{D}_0^{-(n-j)} \boldsymbol{e} \right] + (-1)^n H^{(n)}(0). \end{aligned} \tag{26}$$

(26) gives the moments of the interdeparture time of any order. In particular, the first three moments are given by

$$E\{T_{D,i}\} = \bar{h} - \boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}, \tag{27}$$
$$E\{T_{D,i}^2\} = H^{(2)}(0) - 2\bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} + 2\boldsymbol{x}_0\boldsymbol{D}_0^{-2}\boldsymbol{e}, \tag{28}$$
$$E\{T_{D,i}^3\} = -H^{(3)}(0) - 3H^{(2)}(0)\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} + 6\bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-2}\boldsymbol{e} - 6\boldsymbol{x}_0\boldsymbol{D}_0^{-3}\boldsymbol{e}. \tag{29}$$

Although (26)–(29) are derived based on the finite capacity assumption, they are still valid for infinite queues because the capacity condition is implicitly imposed on $\boldsymbol{x}_i$ ($i \geqslant 0$) only. From [15], we have $\boldsymbol{x}_0 = -(1-\rho)\boldsymbol{g}\boldsymbol{D}_0/\bar{\lambda}$ for an infinite queue which yields $-\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} = 1/\bar{\lambda} - \bar{h}$, where $\bar{\lambda}$ represents the effective arrival rate, $\rho = \bar{\lambda}\bar{h}$, and $\boldsymbol{ge} = 1$. Then (27) reduces to $E\{T_{D,i}\} = 1/\bar{\lambda}$ that can also be deduced from the flow conservation. Using (27) and (28), $c_0$ is obtained as follows:

$$c_0 = H^{(2)}(0) - \bar{h}^2 + 2\boldsymbol{x}_0\boldsymbol{D}_0^{-2}\boldsymbol{e} - \left[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\right]^2. \tag{30}$$

Next, we derive the lag $n$ ($n \geqslant 1$) covariance. From (24), we have

$$\boldsymbol{B}^{(k)}\boldsymbol{e} = H^{(k)}(0)\boldsymbol{e} + \sum_{j=0}^{k-1} \frac{k!}{j!} H^{(j)}(0)\boldsymbol{D}_0^{-(k-j)}\boldsymbol{e}. \tag{31}$$

Using (17) and (31), we obtain

$$\begin{aligned} \boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}} &= \left[ -\bar{h}\boldsymbol{e}^{\mathrm{T}} + \left(\boldsymbol{D}_0^{-1}\boldsymbol{e}\right)^{\mathrm{T}} \quad -\bar{h}\boldsymbol{e}^{\mathrm{T}} \quad \ldots \quad -\bar{h}\boldsymbol{e}^{\mathrm{T}} \right]^{\mathrm{T}} \\ &= -\bar{h}\hat{\boldsymbol{e}} + \left[ \left(\boldsymbol{D}_0^{-1}\boldsymbol{e}\right)^{\mathrm{T}} \quad \mathbf{0}^{\mathrm{T}} \quad \ldots \quad \mathbf{0}^{\mathrm{T}} \right]^{\mathrm{T}} \end{aligned} \tag{32}$$

which makes (19) become

$$c_n = \boldsymbol{x}\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{n-1}\big[\big(\boldsymbol{D}_0^{-1}\boldsymbol{e}\big)^{\mathrm{T}} \quad \mathbf{0}^{\mathrm{T}} \quad \ldots \quad \mathbf{0}^{\mathrm{T}}\big]^{\mathrm{T}} + \bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2, \quad n \geqslant 1. \tag{33}$$

The structure of $\boldsymbol{Q}$ allows one to further simplify (33) as follows:

$$c_n = \begin{cases} (\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n)\big[\boldsymbol{Q}^{(1)}\big]^{n+1,n}\displaystyle\prod_{i=1}^{n-1}[\boldsymbol{Q}]^{i+1,i}\boldsymbol{D}_0^{-1}\boldsymbol{e} + \bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2, \\ \quad 1 \leqslant n \leqslant K-1, \\ \boldsymbol{x}\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{n-K}\displaystyle\prod_{i=1}^{K-1}[\boldsymbol{Q}]^{i+1,i}\boldsymbol{D}_0^{-1}\boldsymbol{e} + \bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2, \quad n \geqslant K, \end{cases} \tag{34}$$

where $[\cdot]^{i,j}$ stands for the submatrix comprising of blocks $(s,t)$, $1 \leqslant s \leqslant i$, $1 \leqslant t \leqslant j$ (here a block is an $m \times m$ matrix with $m$ denoting the number of states). For an infinite queue, i.e., $K = \infty$, $c_n$ $(n \geqslant 1)$ are given by

$$c_n = (\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n)\big[\boldsymbol{Q}^{(1)}\big]^{n+1,n}\prod_{i=1}^{n-1}[\boldsymbol{Q}]^{i+1,i}\boldsymbol{D}_0^{-1}\boldsymbol{e} + \bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2. \tag{35}$$

(34) or (35) gives a neater matrix-form representation and a more direct and efficient method of calculating $c_n$, $n \geqslant 1$ (rather than a recursive/iterative method in [23]). For example,

$$\begin{aligned} c_2 &= (\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)\begin{bmatrix} \boldsymbol{B}_0^{(1)} & \boldsymbol{B}_1^{(1)} \\ \boldsymbol{A}_0^{(1)} & \boldsymbol{A}_1^{(1)} \\ \mathbf{O} & \boldsymbol{A}_0^{(1)} \end{bmatrix}\begin{bmatrix} \boldsymbol{B}_0 \\ \boldsymbol{A}_0 \end{bmatrix}\boldsymbol{D}_0^{-1}\boldsymbol{e} + \bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2, \quad \text{if } K \geqslant 3, \\ c_2 &= (\boldsymbol{x}_0, \boldsymbol{x}_1)\begin{bmatrix} \boldsymbol{B}_0^{(1)} & \sum_{n=1}^{\infty}\boldsymbol{B}_n^{(1)} \\ \boldsymbol{A}_0^{(1)} & \sum_{n=1}^{\infty}\boldsymbol{A}_n^{(1)} \end{bmatrix}\begin{bmatrix} \boldsymbol{B}_0 \\ \boldsymbol{A}_0 \end{bmatrix}\boldsymbol{D}_0^{-1}\boldsymbol{e} + \bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2, \quad \text{if } K = 2, \\ c_2 &= \boldsymbol{x}_0\boldsymbol{B}^{(1)}\boldsymbol{B}(0)\boldsymbol{D}_0^{-1}\boldsymbol{e} + \bar{h}\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2, \quad \text{if } K = 1, \end{aligned} \tag{36}$$

where $\boldsymbol{A}_n^{(k)} \equiv (\mathrm{d}^k/\mathrm{d}s^k)\boldsymbol{A}_n(s)|_{s=0}$ and $\boldsymbol{B}_n^{(k)} \equiv (\mathrm{d}^k/\mathrm{d}s^k)\boldsymbol{B}_n(s)|_{s=0}$. The values of $\boldsymbol{A}_0$, $\boldsymbol{A}_0^{(1)}$, $\boldsymbol{A}_1$, and $\boldsymbol{A}_1^{(1)}$ can be calculated using the formulae in appendix A (for some specific distributions) or the algorithm in appendix B, and the values of $\boldsymbol{B}_0$, $\boldsymbol{B}_0^{(1)}$, $\boldsymbol{B}_1$, and $\boldsymbol{B}_1^{(1)}$ can be obtained using corollary 1.1. Then, the lag $n$ covariance can be calculated from (34) (or (35)). More specifically, $c_1$ can be written in explicit form as follows:

$$c_1 = \big[\bar{h}\boldsymbol{x}_0 + \boldsymbol{x}_0\boldsymbol{B}_0^{(1)} + \boldsymbol{x}_1\boldsymbol{A}_0^{(1)}\big]\boldsymbol{D}_0^{-1}\boldsymbol{e} - \big[\boldsymbol{x}_0\boldsymbol{D}_0^{-1}\boldsymbol{e}\big]^2. \tag{37}$$

*3.3. Departure process of a vacation queue*

Again, let us first find $\boldsymbol{B}(z,s)$.

**Theorem 2.** For vacation BMAP/G/1 and BMAP/G/1/$K$ queues, we have

$$\boldsymbol{B}(z,s) = z^{-1}\big[\boldsymbol{I} - \boldsymbol{V}_0(s)\big]^{-1}\big[\boldsymbol{V}(z,s) - \boldsymbol{V}_0(s)\big]\boldsymbol{A}(z,s). \tag{38}$$

*Proof.* Using arguments analogous to [16], one can prove the theorem. □

**Corollary 2.1.** The matrices $\boldsymbol{B}_n(s)$, $n \geqslant 0$, satisfy

$$\boldsymbol{B}_n(s) = \sum_{j=0}^{n} \boldsymbol{V}_j^0(s)\boldsymbol{A}_{n-j}(s), \tag{39}$$

where $\boldsymbol{V}_j^0(s) = [\boldsymbol{I} - \boldsymbol{V}_0(s)]^{-1}\boldsymbol{V}_{j+1}(s)$.

*Proof.* By equating the corresponding powers of $z$ in theorem 2. □

Note that the results in theorem 2 and corollary 2.1 when narrowing BMAP to MAP are identical to those obtained by Lucantoni et al. in [16].

**Corollary 2.2.**

$$\boldsymbol{B}(s)\boldsymbol{e} = H(s)\big\{\big[V(s) - 1\big]\big[\boldsymbol{I} - \boldsymbol{V}_0(s)\big]^{-1} + \boldsymbol{I}\big\}\boldsymbol{e}. \tag{40}$$

Similar to corollary 1.2, the proof is omitted.

From (14) and (40), we obtain the LST of the interdeparture time as follows:

$$D^{[1]}(s) = H(s)\big\{1 + \big(V(s) - 1\big)\boldsymbol{x}_0\big[\boldsymbol{I} - \boldsymbol{V}_0(s)\big]^{-1}\boldsymbol{e}\big\}. \tag{41}$$

Directly differentiate (41) $n$ times, set $s = 0$, and multiply by $(-1)^n$ to yield

$$\begin{aligned} E\big\{T_{D,i}^n\big\} = (-1)^n \Bigg\{ & H^{(n)}(0) + \sum_{i=1}^{n}\sum_{j=0}^{n-i} \frac{n!}{(n-i-j)!i!j!} V^{(i)}(0) \\ & \times H^{(j)}(0)\boldsymbol{x}_0 \frac{\mathrm{d}^{n-i-j}}{\mathrm{d}s^{n-i-j}}\big[\boldsymbol{I} - \boldsymbol{V}_0(s)\big]^{-1}\big|_{s=0}\boldsymbol{e}\Bigg\}. \end{aligned} \tag{42}$$

Applying (42), we obtain the first three moments as follows:

$$E\big\{T_{D,i}\big\} = \bar{h} + \bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}, \tag{43}$$

$$\begin{aligned} E\big\{T_{D,i}^2\big\} = {} & H^{(2)}(0) + \big(2\bar{h}\bar{v} + V^{(2)}(0)\big)\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} \\ & - 2\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{V}_0^{(1)}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}, \end{aligned} \tag{44}$$

$$
\begin{aligned}
E\{T_{D,i}^3\} = {} & -H^{(3)}(0) + \big[3\bar{h}V^{(2)}(0) + 3\bar{v}H^{(2)}(0) - V^{(3)}(0)\big]\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} \\
& - 3\big[2\bar{h}\bar{v} + V^{(2)}(0)\big]\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{V}_0^{(1)}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} \\
& + 3\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\big\{\boldsymbol{V}_0^{(2)} + 2\boldsymbol{V}_0^{(1)}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{V}_0^{(1)}\big\}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}, \quad (45)
\end{aligned}
$$

where $\boldsymbol{V}_0 \equiv \boldsymbol{V}_0(0)$, $\boldsymbol{V}_0^{(i)} \equiv (\mathrm{d}^i/\mathrm{d}s^i)\boldsymbol{V}_0(s)|_{s=0}$, and $V^{(i)}(0) \equiv (\mathrm{d}^i/\mathrm{d}s^i)V(s)|_{s=0}$. The values of $\boldsymbol{V}_0$, $\boldsymbol{V}_0^{(1)}$, and $\boldsymbol{V}_0^{(2)}$ can be obtained using the formulae in appendix A (for some specific distributions) or the algorithm in appendix B. Since the differentiation of $[\boldsymbol{I} - \boldsymbol{V}_0(s)]^{-1}$ is more involved, we present here the first three moments only. The higher moments can still be obtained through laborious manipulations using (42). Again, (42)–(45) are also valid for the infinite queues for reasons given earlier in section 3.2. Note that $\boldsymbol{x}_0[\boldsymbol{I}-\boldsymbol{V}_0]^{-1}\boldsymbol{e} = (1-\rho)/(\bar{\lambda}\bar{v})$ for an infinite queue (This can be derived using arguments analogous to those in [16].) Then we obtain $\bar{v}\boldsymbol{x}_0[\boldsymbol{I}-\boldsymbol{V}_0]^{-1}\boldsymbol{e} = 1/\bar{\lambda} - \bar{h}$ which makes (43) reduce to $E\{T_{D,i}\} = 1/\bar{\lambda}$ that complies with the flow conservation. Applying (43) and (44), we obtain

$$
\begin{aligned}
c_0 = {} & H^{(2)}(0) - \bar{h}^2 - 2\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{V}_0^{(1)}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} \\
& + V^{(2)}(0)\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} - \big\{\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}\big\}^2. \quad (46)
\end{aligned}
$$

From (40), we obtain

$$
\begin{aligned}
\boldsymbol{B}^{(k)}\boldsymbol{e} = {} & H^{(k)}(0)\boldsymbol{e} + \sum_{i=1}^{k}\sum_{j=0}^{k-i}\frac{k!}{(k-i-j)!i!j!}V^{(i)}(0)H^{(j)}(0) \\
& \times \frac{\mathrm{d}^{k-i-j}}{\mathrm{d}s^{k-i-j}}\big[\boldsymbol{I} - \boldsymbol{V}_0(s)\big]^{-1}\big|_{s=0}\boldsymbol{e}. \quad (47)
\end{aligned}
$$

Apply (17) and (47) to yield

$$
\begin{aligned}
\boldsymbol{Q}^{(1)}\hat{\boldsymbol{e}} &= \big[-\bar{h}\boldsymbol{e}^{\mathrm{T}} - \big(\bar{v}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}\big)^{\mathrm{T}} \quad -\bar{h}\boldsymbol{e}^{\mathrm{T}} \quad \ldots \quad -\bar{h}\boldsymbol{e}^{\mathrm{T}}\big]^{\mathrm{T}} \\
&= -\bar{h}\hat{\boldsymbol{e}} + \big[\big(-\bar{v}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}\big)^{\mathrm{T}} \quad \boldsymbol{0}^{\mathrm{T}} \quad \ldots \quad \boldsymbol{0}^{\mathrm{T}}\big]^{\mathrm{T}} \quad (48)
\end{aligned}
$$

which leads (19) to become the following form after algebraic manipulations analogous to section 3.2

$$
c_n = \begin{cases}
-\bar{v}(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n)\big[\boldsymbol{Q}^{(1)}\big]^{n+1,n}\prod_{i=1}^{n-1}[\boldsymbol{Q}]^{i+1,i}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} - \bar{h}\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} \\
\quad - \big[\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}\big]^2, \quad 1 \leqslant n \leqslant K-1, \\
-\bar{v}\boldsymbol{x}\,\boldsymbol{Q}^{(1)}\boldsymbol{Q}^{n-K}\prod_{i=1}^{K-1}[\boldsymbol{Q}]^{i+1,i}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} - \bar{h}\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} \\
\quad - \big[\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}\big]^2, \quad n \geqslant K.
\end{cases} \quad (49)
$$

For an infinite queue, $c_n$ $(n \geqslant 1)$ are given as follows:

$$c_n = -\bar{v}(\boldsymbol{x}_0, \dots, \boldsymbol{x}_n)\big[\boldsymbol{Q}^{(1)}\big]^{n+1,n} \prod_{i=1}^{n-1}[\boldsymbol{Q}]^{i+1,i}[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} \\ - \bar{h}\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} - \big[\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}\big]^2. \tag{50}$$

Again, (49) or (50) provides a neater matrix-form representation and a more direct and efficient method of calculating $c_n$, $n \geqslant 1$. In particular, we obtain $c_1$ in the following explicit forms:

$$c_1 = -\bar{v}\big[\bar{h}\boldsymbol{x}_0 + \boldsymbol{x}_0\boldsymbol{B}_0^{(1)} + \boldsymbol{x}_1\boldsymbol{A}_0^{(1)}\big][\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e} - \big\{\bar{v}\boldsymbol{x}_0[\boldsymbol{I} - \boldsymbol{V}_0]^{-1}\boldsymbol{e}\big\}^2. \tag{51}$$

As mentioned earlier, the calculation of $\boldsymbol{A}_0$ and $\boldsymbol{A}_0^{(1)}$ can be done using the appendices and $\boldsymbol{B}_0$ and $\boldsymbol{B}_0^{(1)}$ can be obtained using corollary 2.1.

### 3.4. *Formulae validation*

For the vacationless queues, Saito [20] has derived the departure process of a finite $N/\mathrm{G}/1$ queue which is equivalent to the BMAP/G/1 queue. For the vacation queues, Yeh [22] has previously derived the output process of an infinite MAP/G/1 queue. Both of them are special cases of the models considered in this paper. In the following, we examine the correctness of our results through checking with those in [20,22].

From [15,20], one can easily find the following correspondence

$$\boldsymbol{D}_0 = \boldsymbol{R}(0), \tag{52}$$

$$\boldsymbol{D}_k = -\boldsymbol{R}(0)\boldsymbol{U}_k(0), \quad k \geqslant 1, \tag{53}$$

where $\boldsymbol{R}(0)$ and $\boldsymbol{U}_k(0)$ are the symbols related to the $N$-process [18,20]. Replacing $\boldsymbol{D}_0$ by $\boldsymbol{R}(0)$ makes (27) and (28) the same as those in [20]. This supports the correctness of (27) and (28), and indirectly supports (26) since (27) and (28) are derived from (26). For further checks, let us consider the deterministic service distribution with $\widetilde{H}(t) = 0$ when $t < \bar{h}$; and 1 otherwise. For the deterministic case, it can be easily shown that $\boldsymbol{A}_0(s) = \mathrm{e}^{-s\bar{h}}\boldsymbol{P}(0, \bar{h})$, thus $\boldsymbol{A}_0^{(1)} = -\bar{h}\boldsymbol{A}_0$. Also from corollary 1.1, we have $\boldsymbol{B}_0(s) = (s\boldsymbol{I} - \boldsymbol{D}_0)^{-1}\boldsymbol{D}_1\boldsymbol{A}_0(s)$ which yields $\boldsymbol{B}_0^{(1)} = -\boldsymbol{D}_0^{-2}\boldsymbol{D}_1\boldsymbol{A}_0 - \boldsymbol{D}_0^{-1}\boldsymbol{D}_1\boldsymbol{A}_0^{(1)} = -\boldsymbol{D}_0^{-2}\boldsymbol{D}_1\boldsymbol{A}_0 + \bar{h}\boldsymbol{D}_0^{-1}\boldsymbol{D}_1\boldsymbol{A}_0$. And note that $\boldsymbol{x}_1\boldsymbol{A}_0 = \boldsymbol{x}_0 - \boldsymbol{x}_0\boldsymbol{B}_0$ from (7) and $\boldsymbol{x}\boldsymbol{Q} = \boldsymbol{x}$. Using these results together with (52) and (53), (37) now reduces to the form

$$c_1 = \boldsymbol{x}_0\boldsymbol{R}^{-1}(0)\boldsymbol{U}_1(0)\boldsymbol{A}_0\boldsymbol{R}^{-1}(0) - \big[\boldsymbol{x}_0\boldsymbol{R}^{-1}(0)\boldsymbol{e}\big]^2 \tag{54}$$

which is the same as that in [20]. This indirectly checks the correctness of (34) and (35).

We also note that (43), (44), (46) and (51) indeed reduce to those in [22] when narrowing the BMAP to the MAP under the infinite-buffer assumption. This checks the correctness of (42)–(45) and (49)–(51).

## 4. Numerical examples and discussions

In this section, we investigate the phenomena exhibited by the departure statistics of BMAP/G/1 queues under different service and vacation distributions (e.g., deterministic (D), exponential (Exp), $k$-stage Erlangian ($E_k$) [11,12], and $R$-stage hyperexponential ($H_R$) [12]) through numerical experiments. The effect of system capacity is also incorporated.

In the following, we use a set of two-state (i.e., $m = 2$) BMAPs with parameters specified as follows:

$$\boldsymbol{D}_0 = \begin{bmatrix} \alpha_0 & \beta_0 \\ \gamma_0 & \delta_0 \end{bmatrix}, \qquad \boldsymbol{D}_j = p(1-p)^{j-1} \begin{bmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{bmatrix}, \quad j \geqslant 1, \tag{55}$$

where $\alpha_0 = -\beta_0 - \alpha_1 - \beta_1$, $\delta_0 = -\gamma_0 - \gamma_1 - \delta_1$, $\beta_0 > 0$, $\gamma_0 > 0$, $\alpha_1 \geqslant 0$, $\beta_1 \geqslant 0$, $\gamma_1 \geqslant 0$, $\delta_1 \geqslant 0$, and $\{p(1-p)^{j-1}, j \geqslant 1\}$ defines the batch size distribution with mean $1/p$. From (55), we have

$$\boldsymbol{D} = \begin{bmatrix} -\beta_0 - \beta_1 & \beta_0 + \beta_1 \\ \gamma_0 + \gamma_1 & -\gamma_0 - \gamma_1 \end{bmatrix} \tag{56}$$

and

$$\boldsymbol{D}(z) = \boldsymbol{D}_0 + \frac{z}{1-(1-p)z}\boldsymbol{D}_1, \quad |z| < 1. \tag{57}$$

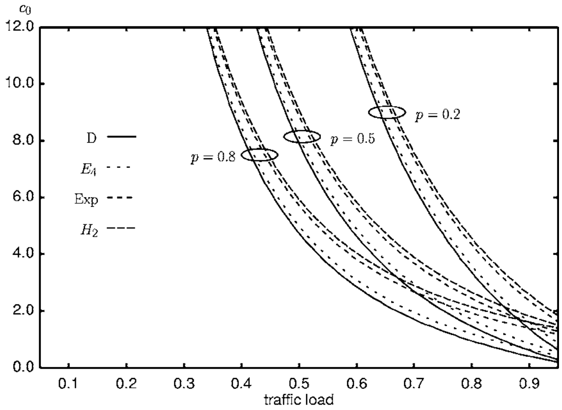
Let $\boldsymbol{\pi}$ denote the stationary probability vector of $\boldsymbol{D}$, i.e., $\boldsymbol{\pi}\boldsymbol{D} = \boldsymbol{\pi}$, $\boldsymbol{\pi}\boldsymbol{e} = 1$ (here, $\boldsymbol{e}$ is a $2 \times 1$ column vector of ones). It can be easily shown that

$$\boldsymbol{\pi} = \left( \frac{\gamma_0 + \gamma_1}{\beta_0 + \beta_1 + \gamma_0 + \gamma_1}, \frac{\beta_0 + \beta_1}{\beta_0 + \beta_1 + \gamma_0 + \gamma_1} \right). \tag{58}$$

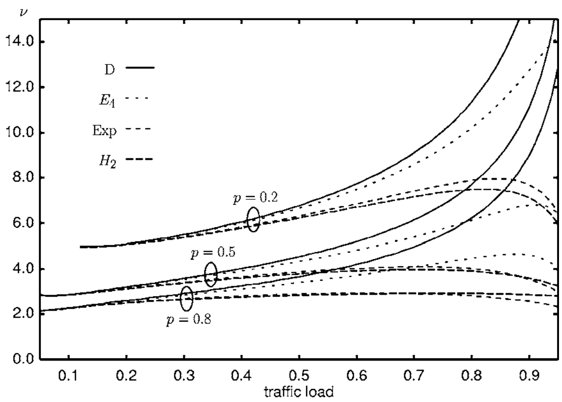
Note that the effective arrival rate $\bar{\lambda} = \boldsymbol{\pi}\boldsymbol{D}'(1)\boldsymbol{e} = \boldsymbol{\pi}\boldsymbol{D}_1\boldsymbol{e}/p^2$. In the following examples, we set $\beta_0 = 0.01$, $\gamma_0 = 0.02$, $\beta_1 = 0.02$, $\gamma_1 = 0.03$, $\alpha_1 = r\delta_1$, $\alpha_0 = -0.03 - \alpha_1$, $\delta_0 = -0.05 - \delta_1$ to yield $\bar{\lambda} = [(5r+3)\delta_1 + 0.19]/(8p)$. We notice that larger value of the ratio $r(= \alpha_1/\delta_1)$ causes more fluctuation to the arrival rate.

### 4.1. *Departure statistics of the infinite-capacity case*

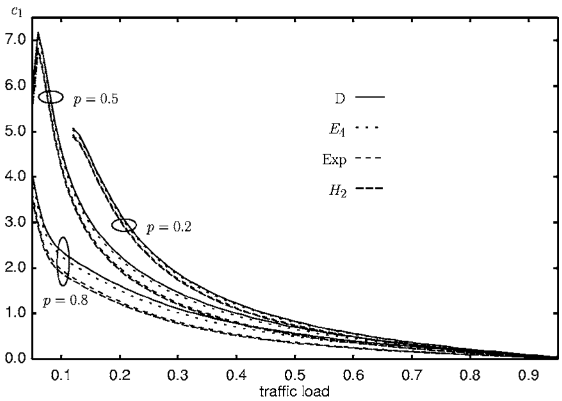
We first examine the departure statistics of the infinite vacationless queue. Consider four different types of server: D, $E_4$, Exp, and $H_2$ with the mean service time $\bar{h} = 1$ in which $H_2$ is specified by two Exps with the mean service times $1/1.2$ and $1/0.6$ weighted respectively by the probabilities 0.8 and 0.2. We note that $c_V^2 = 0$ for D, $0 < c_V^2 < 1$ for $E_4$, $c_V^2 = 1$ for Exp, and $c_V^2 > 1$ for $H_2$, where $c_V^2$ is the squared coefficient of variation (see the definition given in [12]). Varying $\rho$ from 0.05 to 0.95, we obtain the variance $c_0$, skewness $\nu$ ($\nu = E\{(T_{D,i} - E\{T_{D,i}\})^3\}/c_0^{3/2}$ is an index of measure of symmetry [14]: $\nu = 0$ indicates a symmetric distribution and a distribution is skewed to the right/left if $\nu > 0$/$\nu < 0$), and lag 1 covariance $c_1$ in figures 1(a)–(c) for $p = 0.2$,




(a) variance




(b) skewness




(c) lag 1 covariance

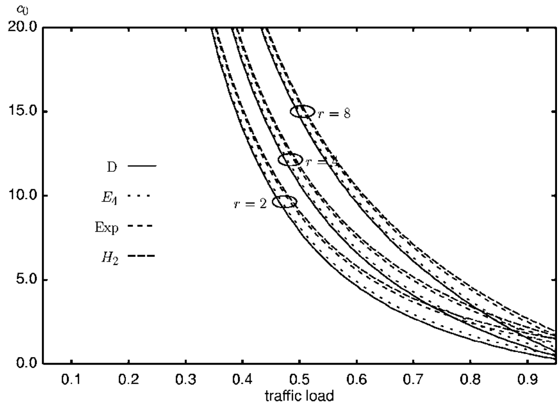
Figure 1. Departure statistics vs. traffic load of a infinite vacationless BMAP/G/1 queue for four different types of server: D, $E_4$, Exp, and $H_2$ with $\bar{h} = 1$ under $p = 0.2$, 0.5 and 0.8 with $r = 2$: (a) variance $c_0$, (b) skewness $\nu$, (c) lag 1 covariance $c_1$.

0.5 and 0.8, respectively, under $r = 2$. Before contemplating on the numerical results, let us first relate the following parameters to their physical meanings:

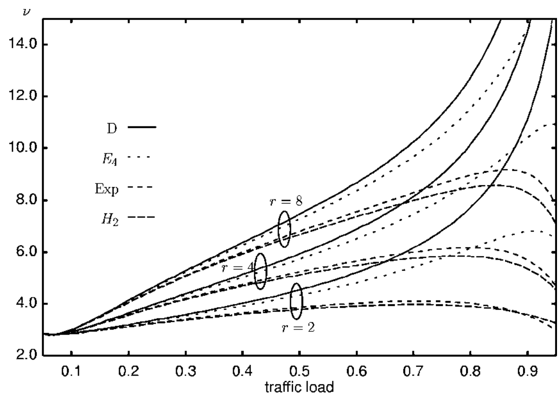
- $c_V^2$: larger $c_V^2$ of service or vacation distribution induces more variation and randomness which in turn reduces correlation between interdeparture times.
- $\rho$: for a fixed service time, larger $\rho$, i.e., smaller interarrival times, results in smaller absolute values of moments and lag $n$ covariance; here we call this phenomenon the "scaling effect".
- $p$: since the mean burst size equals the reciprocal of $p$, smaller $p$ causes more burstiness; thus more variation and higher correlation between interdeparture times.
- $r$: this parameter is the ratio $\alpha_1/\delta_1$ and has contribution to burstiness.
- $\bar{v}$: larger value of $\bar{v}$ causes $T_{D,i}$ to deviate more profoundly from its mean and induces larger or more negative terms when calculating $c_1$ in order to maintain a constant level of $T_{D,i}$ (this will be explained later).

The above correspondence helps to understand the phenomena exhibited by the departure statistics under the influence of these parameters. We summarize the followings from figures 1(a)–(c): First, due to less burstiness, larger value of $p$ yields smaller $c_0$, $\nu$, and $c_1$ under the same type of server. Second, under a fixed $p$, larger $c_V^2$ of a service distribution produces larger $c_0$ since more variation and randomness arise in interdeparture time $T_{D,i}$. Conversely, larger $c_V^2$ of a service distribution makes $c_1$ smaller because the raised randomness reduces correlation. A similar phenomenon is observed for $\nu$ except crossovers occur for Exp and $H_2$ at (extremely) high traffic load. Third, as traffic load increases, $c_0$ monotonically decreases due to the scaling effect. For $c_1$, the scaling effect dominates except at very light traffic load in which different values of $p$ may disrupt the trends, e.g., a peak for $c_1$ under $p = 0.5$. Since $E\{(T_{D,i} - E\{T_{D,i}\})^3\}$ and $c_0^{3/2}$ possess a similar trend as traffic load goes up, $\nu = E\{(T_{D,i} - E\{T_{D,i}\})^3\}/c_0^{3/2}$ is somewhat unpredictable. But we observe that $\nu$ for a D server has a monotonically increasing trend regardless of the value of $p$. Note that the observed skewness $\nu$ is positive which implies the interdeparture time distribution is skewed to the right. Shown in figures 2(a)–(c) are the above three departure statistics v.s. traffic load for a fixed $p = 0.5$ and three different values of $r = 2$, 4 and 8, respectively. Similar phenomena are observed, but larger $r$ (more fluctuation or more burstiness in the arrival rate) causes larger $c_0$, $\nu$, and $c_1$ under the same type of server (except at extremely low traffic load, where $c_1$ has crossovers to yield reverse results).

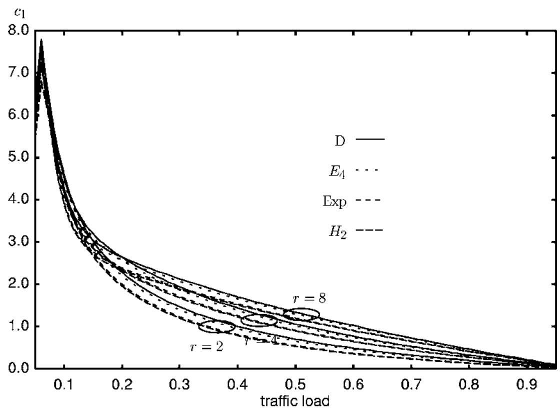
Next, we explore the phenomena exhibited by the departure statistics of the infinite and vacation BMAP/G/1 queue under four different vacation distributions: D, $E_4$, Exp, and $H_2$. Here $H_2$ is specified by two Exps with the mean vacation periods $\bar{v}/1.2$ and $\bar{v}/0.6$ weighted respectively by the probabilities 0.8 and 0.2. Under $p = 0.5$, $r = 2$, and $\bar{v} = 1$, 10 and 40, we obtain figures 3(a)–(c) for a D server with $\bar{h} = 1$ and figures 4(a)–(c) for an $H_2$ server with $\bar{h} = 1$ (the $H_2$ server is specified the same as that in the vacationless case). Based on figures 3(a)–(c), we first notice that $c_0$ under the same type of vacation distribution increases as $\bar{v}$ increases while $c_1$ decreases (even decreases




(a) variance

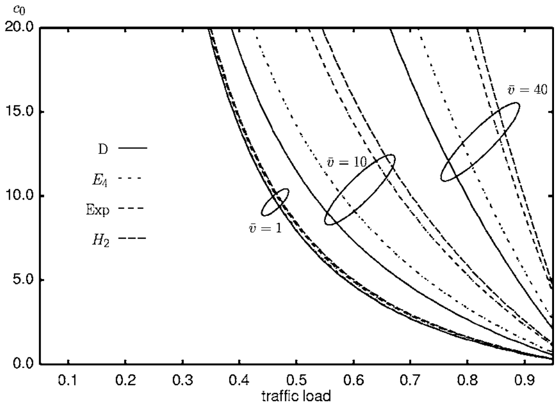



(b) skewness

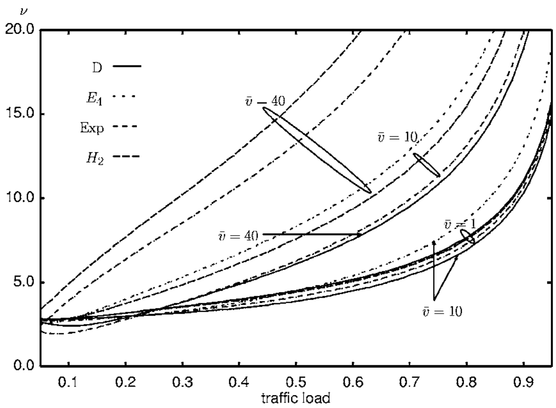



(c) lag 1 covariance

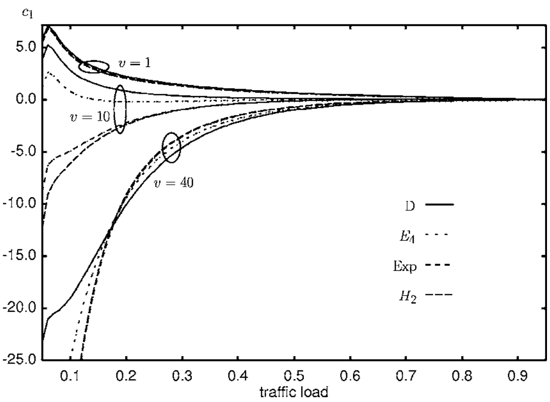
Figure 2. Departure statistics vs. traffic load of a infinite vacationless BMAP/G/1 queue for four different types of server: D, $E_4$, Exp, and $H_2$ with $\bar{h} = 1$ under $r = 2$, 4 and 8 with $p = 0.5$: (a) variance $c_0$, (b) skewness $\nu$, (c) lag 1 covariance $c_1$.




(a) variance

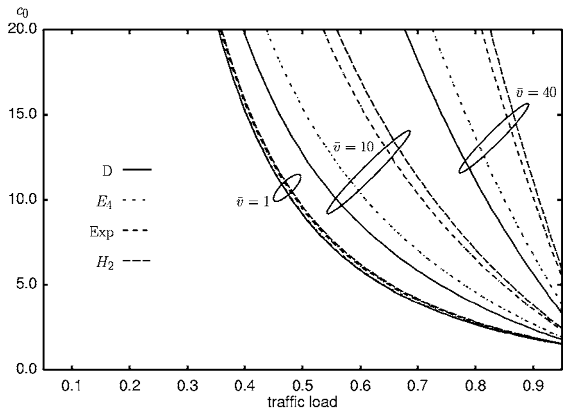



(b) skewness

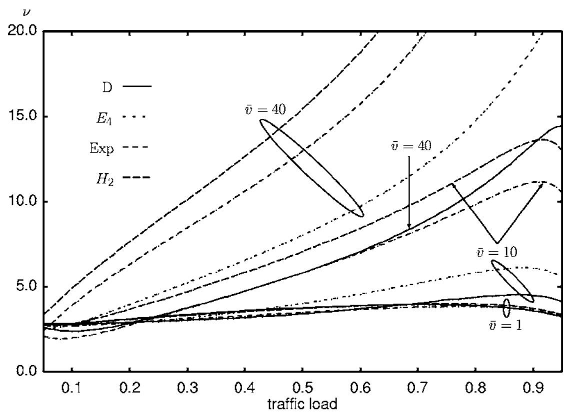



(c) lag 1 covariance

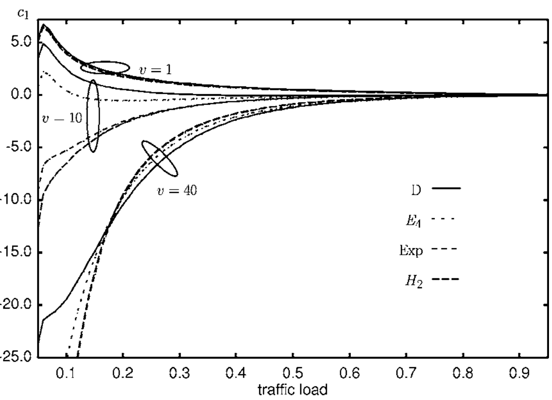
Figure 3. Departure statistics vs. traffic load of a infinite vacation BMAP/D/1 queue for four different types of vacation distribution: D, $E_4$, Exp, and $H_2$ with $\bar{h} = 1$, $\bar{v} = 1$, 10 and 40 under $r = 2$ and $p = 0.5$: (a) variance $c_0$, (b) skewness $\nu$, (c) lag 1 covariance $c_1$.




(a) variance




(b) skewness




(c) lag 1 covariance

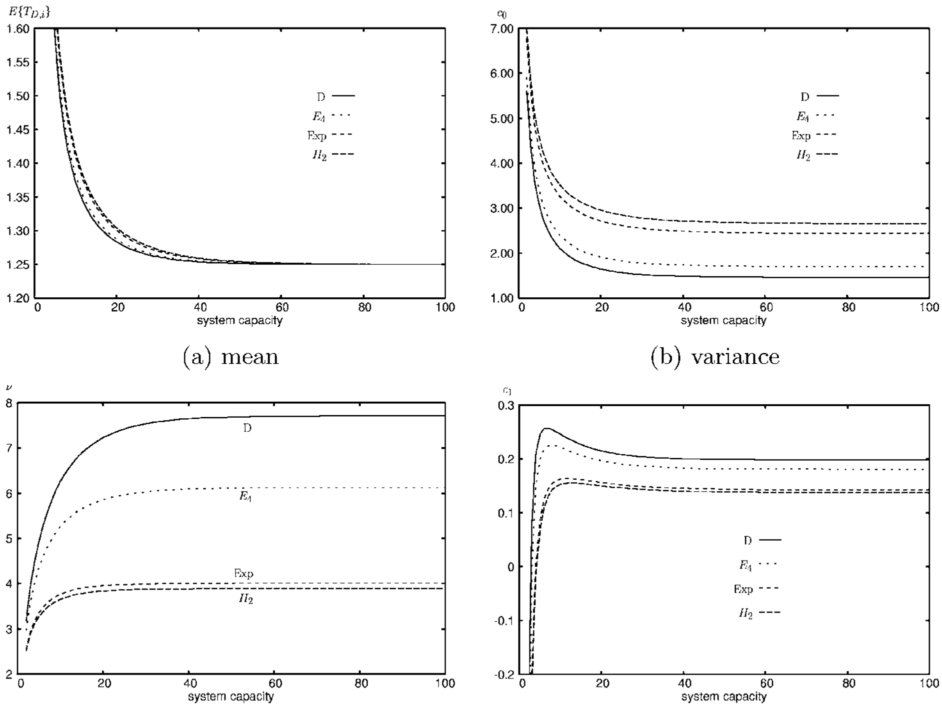
Figure 4. Departure statistics vs. traffic load of a infinite vacation BMAP/$H_2$/1 queue for four different types of vacation distribution: D, $E_4$, Exp, and $H_2$ with $\bar{h} = 1$, $\bar{v} = 1$, 10 and 40 under $r = 2$ and $p = 0.5$: (a) variance $c_0$, (b) skewness $\nu$, (c) lag 1 covariance $c_1$.

to negative values) when $\bar{v}$ becomes high. We now explain this phenomenon. The vacation policy causes interdeparture time $T_{D,i}$ to deviate more profoundly from its mean (i.e., $E\{T_{D,i}\}$) when $\bar{v}$ gets larger. Therefore, $c_0$ increases when $\bar{v}$ goes up. We further note that for a fixed arrival rate $\bar{\lambda}$, $E\{T_{D,i}\}$ is a constant (which is the reciprocal of $\bar{\lambda}$) and is independent of the vacation policy. Then, $T_{D,i}$'s deviate bidirectionally with respect to the mean in order to maintain a constant level of $E\{T_{D,i}\}$. Thus, larger (or more) negative terms arise when calculating $c_1$ and cause $c_1$ to decrease (even to negative values) as $\bar{v}$ gets large. The skewness $\nu$ increases as $\bar{v}$ gets larger for all vacation distributions except the deterministic vacation distribution. Second, for a fixed $\bar{v}$, larger $c_V^2$ of a vacation distribution induces larger $c_0$ due to variation of the vacation distribution. This phenomenon is more obvious for larger $\bar{v}$. Same as the vacationless case, $c_0$ decreases when traffic load increases. For smaller $\bar{v}$ ($\bar{v}$ varies from 1 to 10 in figure 3(c)), $c_1$ is larger if $c_V^2$ of a vacation distribution is larger. We note that a transition region in which $c_1$ may be positive and negative for different values of $c_V^2$ of a vacation distribution occurs (see figure 3(c) when $\bar{v} = 10$). As $\bar{v}$ further increases, crossovers occur for $c_1$ under different vacation distributions. This is due to that larger or more negative terms arise for $c_1$ with larger value of $c_V^2$ at low traffic load in which vacations occur more frequently. As traffic load increases, the effect of $c_V^2$ takes over that of $\bar{v}$. Therefore, crossovers occur for $c_1$ with different values of $c_V^2$ under a larger value of $\bar{v}$. Under a larger value of $\bar{v}$, larger $c_V^2$ of a vacation distribution also induces larger skewness $\nu$ due to explicit vacation effect while the effect of $c_V^2$ is insensitive for a smaller $\bar{v}$, say, $\bar{v} = 1$. Using the parameters of figure 3 (except that D server is replaced by $H_2$), figure 4 is obtained. It reveals that $c_0$, and $c_1$ are insensitive to service distributions, but $\nu$ is not. Although the effect of $\bar{v}$ and $c_V^2$ remains the same as that in figure 3, $\nu$ may first go up then down rather than continue to go up as traffic load increases since an $H_2$ server is employed.

### *4.2. Departure statistics of the finite-capacity case*

The effect of system capacity is investigated in figure 5 for vacationless queues while in figures 6 and 7 for vacation queues under $r = 2$, $p = 0.5$, $\bar{h} = 1$, and $\rho = 0.8$. For four different types of server ($H_2$ is specified analogously to those examples in section 4.1), the mean , variance, skewness, and lag 1 covariance of interdeparture times are obtained in figures 5(a)–(d). These figures show the followings. First, as the system capacity increases, $E\{T_{D,i}\}$ and $c_0$ monotonically decrease since smaller capacity results in more extra gaps which widens the interdeparture time. Second, the extra gaps caused by smaller system capacity result in negative terms in the calculation of $c_1$. Therefore, $c_1$ first sharply increases then smoothly decreases as the system capacity increases. Third, following the reason given in section 4.1, server with larger $c_V^2$ also possesses larger $E\{T_{D,i}\}$ and $c_0$ but smaller $c_1$. Fourth, the skewness $\nu$ increases as the system capacity increases and has a larger value for a server with smaller $c_V^2$.

For four different vacation distributions (with $\bar{v}$ taking three different values 1, 10 and 40), the departure statistics are obtained in figure 6 under the D server and figure 7 under the $H_2$ server. From these figures, we conclude the followings. First, for a fixed




(a) mean (b) variance (c) skewness (d) lag 1 covariance

Figure 5. Departure statistics vs. system capacity of a vacationless BMAP/G/1/$K$ queue for four different types of server, i.e., D, $E_4$, Exp, and $H_2$ with $\bar{h} = 1$ under $r = 2$, $p = 0.5$, and $\rho = 0.8$: (a) mean $E\{T_{D,i}\}$, (b) variance $c_0$, (c) skewness $\nu$, (d) lag 1 covariance $c_1$.

value of $\bar{v}$, the trends of $E\{T_{D,i}\}$ and $c_0$ caused by various $c_V^2$'s of vacation distributions are similar to those of vacationless queues (figures 5(a), (b)) as the system capacity increases. Second, as the value of $\bar{v}$ gets larger, $E\{T_{D,i}\}$ and $c_0$ also become larger because larger $\bar{v}$ results in longer interdeparture times as well as more severe deviation of the interdeparture time from its mean, while $c_1$ becomes smaller because more or larger negative terms arise in calculating $c_1$. Third, irregular trends are observed for skewness $\nu$ (figure 6(c)). For figure 7 in which the $H_2$ server is employed, results similar to figure 6 are observed.

### 4.3. *Lag n* ($n \geqslant 1$) *covariance for both infinite- and finite-capacity cases*

Long-term correlation of the arrival process may have critical impact on system performance, e.g., loss probability. Therefore, lag $n$ ($n \geqslant 1$) covariance of interdeparture times is worth exploring. Now, the lag $n$ covariance $c_n$ of infinite queues for both vacationless and vacation cases is investigated in figures 8(a)–(c). The trend of $c_n$ shown in figure 8(a) for vacationless queues is straightforward but exhibits diverse characteristics for small lags and converges to zero for large lags in figures 8(b), (c) for vacation queues. Using the same reason given in section 4.1 for $c_1$, larger $\bar{v}$ may cause $c_n$ to become negative.

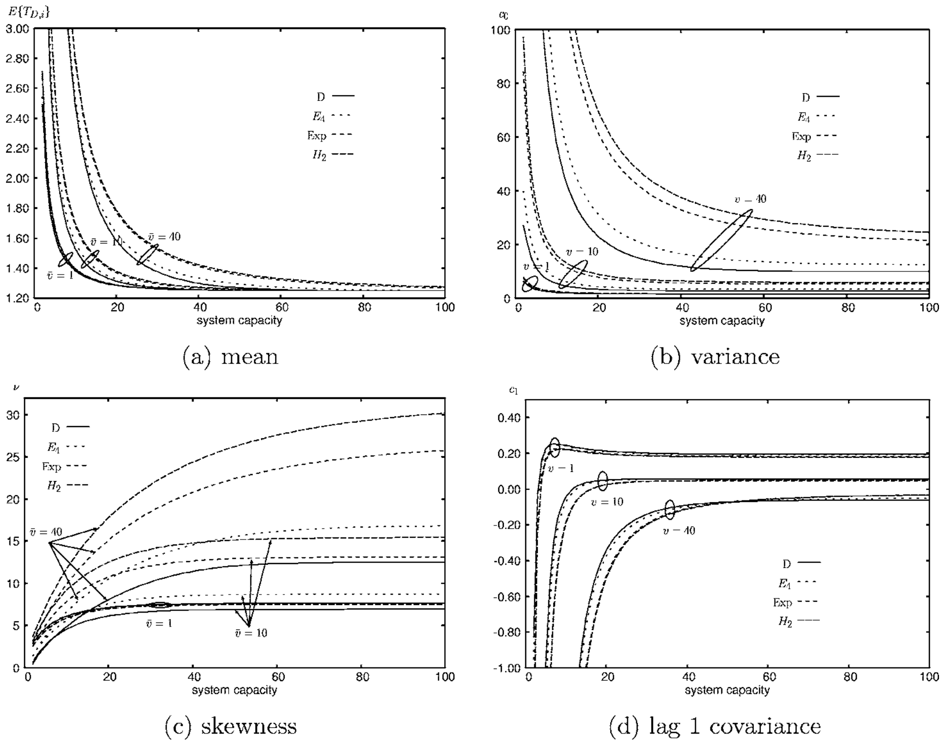



Figure 6. Departure statistics vs. system capacity of a vacation BMAP/D/1/$K$ queue for four different types of vacation distribution, i.e., D, $E_4$, Exp, and $H_2$ with $\bar{h} = 1$, $\bar{v} = 1$, 10 and 40 under $r = 2$, $p = 0.5$, and $\rho = 0.8$: (a) mean $E\{T_{D,i}\}$, (b) variance $c_0$, (c) skewness $\nu$, (d) lag 1 covariance $c_1$.

Influence of system capacity on $c_n$ is further investigated in figures 9(a), (b) for vacationless queues and figures 9(c), (d) for vacation queues under capacity $K = 10$ and 50, respectively. Figure 9(a) ($\rho = 0.2$) and figure 9(b) ($\rho = 0.8$) show that $K = 10$ makes $c_9$ the smallest, while $K = 50$ yields monotonically decreasing trend similar to infinite queues. The above results show that smaller system capacity ($K$) causes $c_n$ to sharply decrease when $1 \leqslant n \leqslant K - 1$, then continues to first go up then slightly drops, and approaches zero for large lags finally.

Unlike $c_n$ of vacationless finite queues, $c_n$ of vacation finite queues behaves differently. Under a small system capacity, say, $K = 10$, larger $\bar{v}$ results in more severe oscillation for small lags. This oscillation gradually diminishes as the lag becomes larger (see figure 9(c) employing a D server and figure 9(d) an $H_2$ server). For larger system capacity, the oscillation vanishes regardless of the value of $\bar{v}$.

## 5. Conclusions

We have analyzed the departure process of both vacationless and vacation BMAP/G/1(/$K$) queues in this paper. Instead of using a recursive method (employed in [23]) to

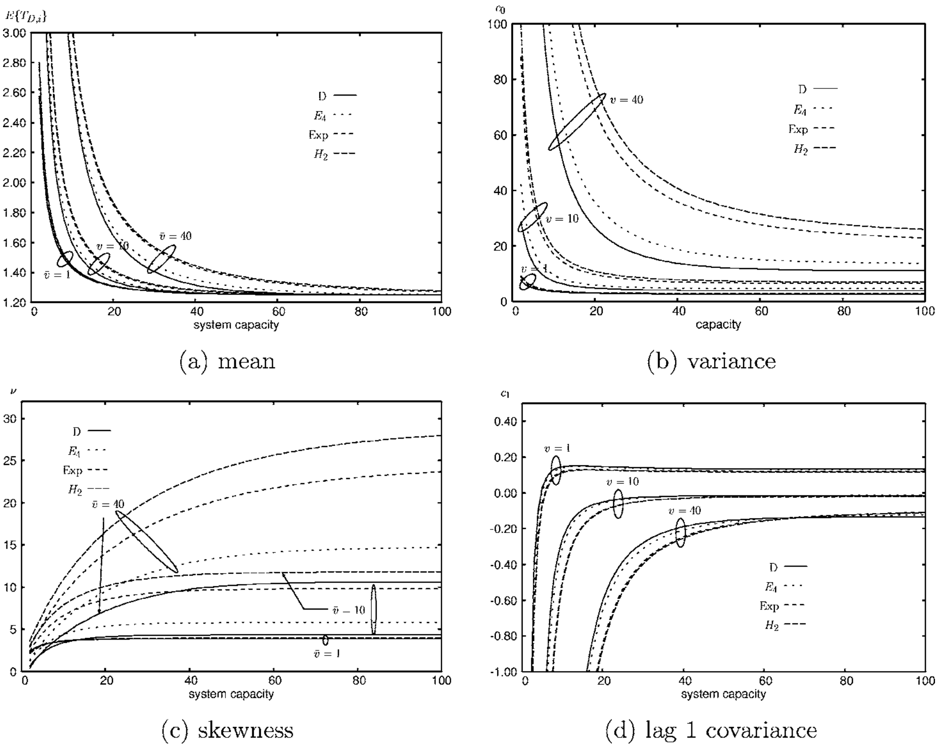



Figure 7. Departure statistics v.s. system capacity of a vacation BMAP/$H_2/1/K$ queue for four different types of vacation distribution, i.e., D, $E_4$, Exp, and $H_2$ with $\bar{h} = 1$, $\bar{v} = 1$, 10 and 40 under $r = 2$, $p = 0.5$, and $\rho = 0.8$: (a) mean $E\{T_{D,i}\}$, (b) variance $c_0$, (c) skewness $\nu$, (d) lag 1 covariance $c_1$.

obtain the lag $n$ ($n \geqslant 1$) covariance, we provided a matrix-form representation to facilitate quick calculation of these statistics. The results obtained in this paper generalize those in [20,22] and have potential applications in contemporary network performance evaluation. Through numerical experiments, several interesting phenomena exhibited by the departure statistics are also investigated.

## Appendix A. Formulae of $A_0$, $A_0^{(1)}$, $A_0^{(2)}$, $V_0$, $V_0^{(1)}$, and $V_0^{(2)}$

From (5) and (6), $A_0$, $A_0^{(1)}$, $A_0^{(2)}$, $V_0$, $V_0^{(1)}$, $V_0^{(2)}$ can be obtained using the following equations:

$$\begin{aligned} A_0 &= \int_0^\infty \mathrm{d}\widetilde{A}_0(x), & V_0 &= \int_0^\infty \mathrm{d}\widetilde{V}_0(x), \\ A_0^{(1)} &= -\int_0^\infty x\,\mathrm{d}\widetilde{A}_0(x), & V_0^{(1)} &= -\int_0^\infty x\,\mathrm{d}\widetilde{V}_0(x), \\ A_0^{(2)} &= \int_0^\infty x^2\,\mathrm{d}\widetilde{A}_0(x), & V_0^{(2)} &= \int_0^\infty x^2\,\mathrm{d}\widetilde{V}_0(x). \end{aligned} \tag{A.1}$$

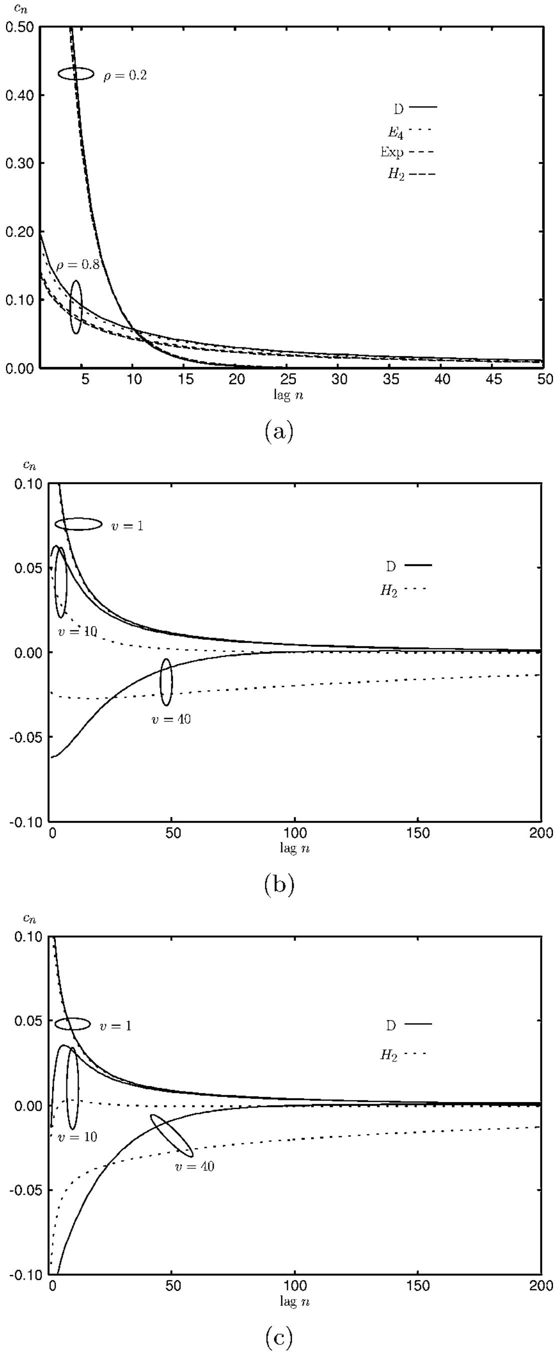



Figure 8. Lag $n$ covariance of infinite queues under $\bar{h} = 1$, $r = 2$, and $p = 0.5$: (a) the vacationless case for four different types of server, (b) the vacation case with D server ($\rho = 0.8$) following D and $H_2$ vacation distributions under $\bar{v} = 1$, 10 and 40, respectively, (c) the vacation case with $H_2$ server ($\rho = 0.8$).

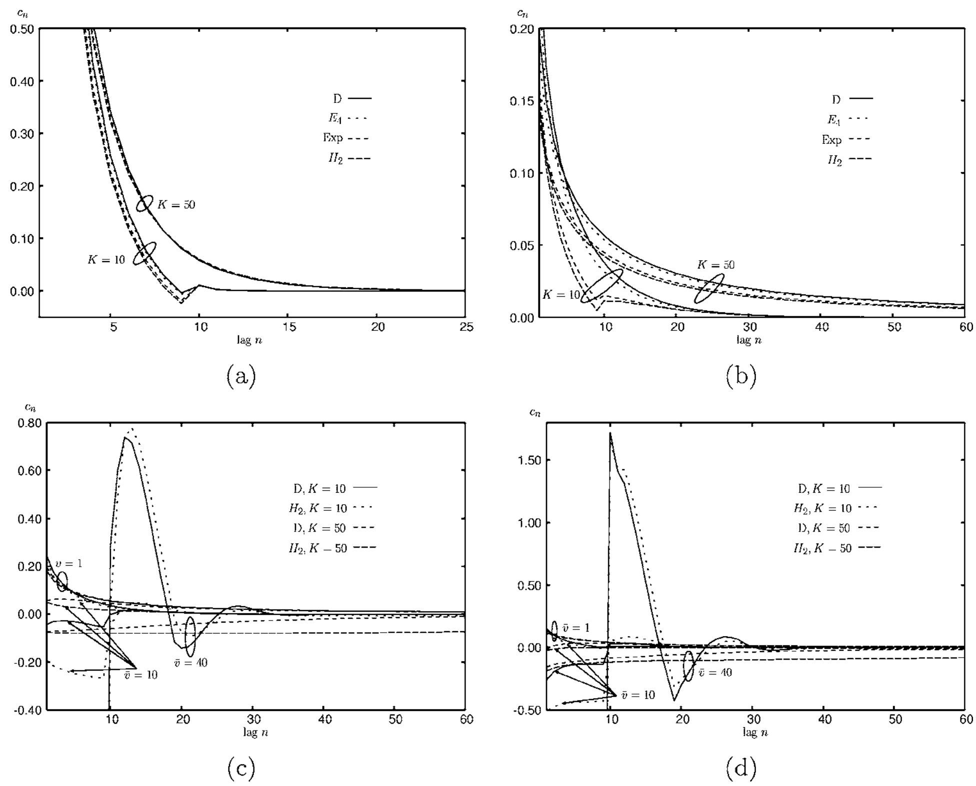


Figure 9. Lag $n$ covariance of finite queues under $\bar{h} = 1$, $r = 2$, and $p = 0.5$: (a) the vacationless case for four different types of server with $\rho = 0.2$, (b) the vacationless case for four different types of server with $\rho = 0.8$, (c) the vacation case with D server ($\rho = 0.8$) following D and $H_2$ vacation distributions under $\bar{v} = 1$, 10 and 40, respectively, (d) the vacation case with $H_2$ server ($\rho = 0.8$).

1. Deterministic service distribution, i.e., $(\mathrm{d}\widetilde{H}(x)/\mathrm{d}x) = \delta(x - h)$:

$$\begin{aligned} \boldsymbol{A}_0 &= \exp[h\boldsymbol{D}_0], & \boldsymbol{V}_0 &= \exp[h\boldsymbol{D}_0], \\ \boldsymbol{A}_0^{(1)} &= -h\boldsymbol{A}_0, & \boldsymbol{V}_0^{(1)} &= -h\boldsymbol{V}_0, \\ \boldsymbol{A}_0^{(2)} &= h^2\boldsymbol{A}_0, & \boldsymbol{V}_0^{(2)} &= h^2\boldsymbol{V}_0. \end{aligned} \tag{A.2}$$

2. $k$-stage Erlangian service distribution, i.e., $\mathrm{d}\widetilde{H}(x)/\mathrm{d}x = \Gamma(x; k, \mu)$ [11,12]:

$$\begin{aligned} \boldsymbol{A}_0 &= \mu^k[\mu\boldsymbol{I} - \boldsymbol{D}_0]^{-k}, \\ \boldsymbol{A}_0^{(1)} &= -k\boldsymbol{A}_0[\mu\boldsymbol{I} - \boldsymbol{D}_0]^{-1}, \\ \boldsymbol{A}_0^{(2)} &= k(k+1)\boldsymbol{A}_0[\mu\boldsymbol{I} - \boldsymbol{D}_0]^{-2}, \\ \boldsymbol{V}_0 &= \mu^k[\mu\boldsymbol{I} - \boldsymbol{D}_0]^{-k}, \\ \boldsymbol{V}_0^{(1)} &= -k\boldsymbol{V}_0[\mu\boldsymbol{I} - \boldsymbol{D}_0]^{-1}, \\ \boldsymbol{V}_0^{(2)} &= k(k+1)\boldsymbol{V}_0[\mu\boldsymbol{I} - \boldsymbol{D}_0]^{-2}. \end{aligned} \tag{A.3}$$

3. Weighted mixture of deterministic and $k$-stage Erlangian distributions: the resultant $\boldsymbol{A}_0$, $\boldsymbol{A}_0^{(1)}$, and $\boldsymbol{A}_0^{(2)}$ ($\boldsymbol{V}_0$, $\boldsymbol{V}_0^{(1)}$, and $\boldsymbol{V}_0^{(2)}$) are the weighted sum of the individual $\boldsymbol{A}_0$, $\boldsymbol{A}_0^{(1)}$, and $\boldsymbol{A}_0^{(2)}$ ($\boldsymbol{V}_0$, $\boldsymbol{V}_0^{(1)}$, and $\boldsymbol{V}_0^{(2)}$).

## Appendix B. Algorithm for calculating $\boldsymbol{A}_n$, $\boldsymbol{A}_n^{(1)}$, $\boldsymbol{A}_n^{(2)}$, $\boldsymbol{V}_n$, $\boldsymbol{V}_n^{(1)}$, and $\boldsymbol{V}_n^{(2)}$, $n \geqslant 0$

Lucantoni [15] has shown that

$$\boldsymbol{P}(n,t) = \sum_{j=0}^{\infty} \mathrm{e}^{-\theta t} \frac{(\theta t)^j}{j!} \boldsymbol{K}_n^{(j)}, \tag{B.1}$$

where $\theta = \max_i\{(-\boldsymbol{D}_0)_{ii}\}$ and $\{\boldsymbol{K}_n^{(j)}\}$ is recursively defined by $\boldsymbol{K}_0^{(0)} = \boldsymbol{I}$, $\boldsymbol{K}_n^{(0)} = \boldsymbol{0}$, $n \geqslant 1$, and

$$\boldsymbol{K}_0^{(j+1)} = \boldsymbol{K}_0^{(j)}\big(\boldsymbol{I} + \theta^{-1}\boldsymbol{D}_0\big), \tag{B.2}$$

$$\boldsymbol{K}_n^{(j+1)} = \theta^{-1} \sum_{i=0}^{n-1} \boldsymbol{K}_i^{(j)} \boldsymbol{D}_{n-i} + \boldsymbol{K}_n^{(j)}\big(\boldsymbol{I} + \theta^{-1}\boldsymbol{D}_0\big) \tag{B.3}$$

in which $\boldsymbol{I}$ and $\boldsymbol{0}$ are $m \times m$ identity and zero matrices. Apply (5), (6), and (B.1) to produce the following expressions for $\boldsymbol{A}_n$, $\boldsymbol{V}_n$, $\boldsymbol{A}_n^{(1)}$, $\boldsymbol{V}_n^{(1)}$, $\boldsymbol{A}_n^{(2)}$, and $\boldsymbol{V}_n^{(2)}$:

$$\begin{aligned}
\boldsymbol{A}_n &= \sum_{j=0}^{\infty} \gamma_{s,j} \boldsymbol{K}_n^{(j)}, & \boldsymbol{V}_n &= \sum_{j=0}^{\infty} \gamma_{v,j} \boldsymbol{K}_n^{(j)}, \\
\boldsymbol{A}_n^{(1)} &= \sum_{j=0}^{\infty} \gamma_{s,j}^{(1)} \boldsymbol{K}_n^{(j)}, & \boldsymbol{V}_n^{(1)} &= \sum_{j=0}^{\infty} \gamma_{v,j}^{(1)} \boldsymbol{K}_n^{(j)}, \\
\boldsymbol{A}_n^{(2)} &= \sum_{j=0}^{\infty} \gamma_{s,j}^{(2)} \boldsymbol{K}_n^{(j)}, & \boldsymbol{V}_n^{(2)} &= \sum_{j=0}^{\infty} \gamma_{v,j}^{(2)} \boldsymbol{K}_n^{(j)},
\end{aligned} \tag{B.4}$$

where

$$\begin{aligned}
\gamma_{s,n} &= \int_0^{\infty} \mathrm{e}^{-\theta x} \frac{(\theta x)^n}{n!}\, \mathrm{d}\widetilde{H}(x), & \gamma_{s,n}^{(1)} &= -\int_0^{\infty} \mathrm{e}^{-\theta x} \frac{\theta^n x^{n+1}}{n!}\, \mathrm{d}\widetilde{H}(x), \\
\gamma_{s,n}^{(2)} &= \int_0^{\infty} \mathrm{e}^{-\theta x} \frac{\theta^n x^{n+2}}{n!}\, \mathrm{d}\widetilde{H}(x), & \gamma_{v,n} &= \int_0^{\infty} \mathrm{e}^{-\theta x} \frac{(\theta x)^n}{n!}\, \mathrm{d}\widetilde{V}(x), \\
\gamma_{v,n}^{(1)} &= -\int_0^{\infty} \mathrm{e}^{-\theta x} \frac{\theta^n x^{n+1}}{n!}\, \mathrm{d}\widetilde{V}(x), & \gamma_{v,n}^{(2)} &= \int_0^{\infty} \mathrm{e}^{-\theta x} \frac{\theta^n x^{n+2}}{n!}\, \mathrm{d}\widetilde{V}(x).
\end{aligned} \tag{B.5}$$

Note that $\sum_{n=0}^{\infty} \gamma_{s,n} = 1$ and $\sum_{n=0}^{\infty} \gamma_{v,n} = 1$. The representation in (B.4) together with the recursion of (B.2) and (B.3) yield a simple and efficient algorithm to compute $\boldsymbol{A}_n$, $\boldsymbol{A}_n^{(1)}$, $\boldsymbol{A}_n^{(2)}$, $\boldsymbol{V}_n$, $\boldsymbol{V}_n^{(1)}$, and $\boldsymbol{V}_n^{(2)}$ for $n \geqslant 0$ as follows:

Step 1: Fix a value for the index $n$, say, $n = 2$.

Step 2: Set $\theta = \max_i\{(-\boldsymbol{D}_0)_{ii}\}$ and $\boldsymbol{K}_0^{(0)} = \boldsymbol{I}$, $\boldsymbol{K}_i^{(0)} = \boldsymbol{0}$ for $1 \leqslant i \leqslant n$.

Step 3: For $l = 0, 1, \ldots, n^*$ ($n^*$ is chosen such that $\sum_{k=0}^{n^*} \gamma_{s,l} > 1 - \varepsilon$ and $\sum_{k=0}^{n^*} \gamma_{v,l} > 1 - \varepsilon$ with $\varepsilon \ll 1$, say $10^{-8}$, specified for truncating the summations in (B.4)), compute $\gamma_{s,l}$, $\gamma_{v,l}$, $\gamma_{s,l}^{(1)}$, $\gamma_{v,l}^{(1)}$, $\gamma_{s,l}^{(2)}$, and $\gamma_{v,l}^{(2)}$ using (B.5).

Step 4: For $j = 0, 1, 2, \ldots, n^*$, compute $\boldsymbol{K}_0^{(j)}, \ldots, \mathrm{K}_n^{(j)}$ using (B.2) and (B.3).

Step 5: Compute $\boldsymbol{A}_n$, $\boldsymbol{A}_n^{(1)}$, $\boldsymbol{A}_n^{(2)}$ and $\boldsymbol{V}_n$, $\boldsymbol{V}_n^{(1)}$, $\boldsymbol{V}_n^{(2)}$ using the following equations:

$$\begin{aligned} \boldsymbol{A}_n &= \sum_{j=0}^{n^*} \gamma_{s,j} \boldsymbol{K}_n^{(j)}, & \boldsymbol{V}_n &= \sum_{j=0}^{n^*} \gamma_{v,j} \boldsymbol{K}_n^{(j)}, \\ \boldsymbol{A}_n^{(1)} &= \sum_{j=0}^{n^*} \gamma_{s,j}^{(1)} \boldsymbol{K}_n^{(j)}, & \boldsymbol{V}_n^{(1)} &= \sum_{j=0}^{n^*} \gamma_{v,j}^{(1)} \boldsymbol{K}_n^{(j)}, \\ \boldsymbol{A}_n^{(2)} &= \sum_{j=0}^{n^*} \gamma_{s,j}^{(2)} \boldsymbol{K}_n^{(j)}, & \boldsymbol{V}_n^{(2)} &= \sum_{j=0}^{n^*} \gamma_{v,j}^{(2)} \boldsymbol{K}_n^{(j)}. \end{aligned} \tag{B.6}$$

## References


[1] ATM Forum, *ATM User-Interface Specification, Version 3.0* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
[2] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1987).
[3] P.J. Burke, The output of a queueing system, Oper. Res. 4 (1956) 699–704.
[4] D.J. Daley, Queueing output processes, Adv. in Appl. Probab. 8 (1976) 395–415.
[5] B.T. Doshi, Queueing systems with vacations – a survey, Queueing Systems 1 (1986) 29–66.
[6] H.W. Ferng and J.F. Chang, The departure process of discrete-time queueing systems with Markovian type inputs, Queueing Systems 36 (2000) 201–220.
[7] W. Fischer and K.S. Meier-Hellstern, The Markov-modulated Poisson process (MMPP) cookbook, Performance Evaluation 18 (1993) 149–171.
[8] W.K. Grassmann, M.I. Taksar and D.P. Heyman, Regenerative analysis and steady state distribution for Markov chains, Oper. Res. 33 (1985) 1107–1116.
[9] P. Harrison and N. Pated, *Performance Modeling of Communication Networks and Computer Architectures* (Addison-Wesley, New York, 1993).
[10] D.P. Heyman and M.J. Sobel, *Handbooks in Operations Research and Management Science, Vol. 2, Stochastic Models* (Elsevier, Amsterdam, 1990), chapter V.
[11] P.G. Hoel, S.C. Port and C.J. Stone, *Introduction to Probability Theory* (Houghton Mifflin, Boston, 1971).
[12] L. Kleinrock, *Queueing Systems, Vol. I: Theory* (Wiley, New York, 1975).
[13] P.J. Kuehn, Remainder on queueing theory for ATM networks, Telecommunication Systems 5 (1996) 1–24.
[14] A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis*, 2nd ed. (McGraw-Hill, New York, 1991).
[15] D.M. Lucantoni, New results on the single server queue with a batch Markovian arrival process, Commun. Statist. Stochastic Models 7(1) (1991) 1–46.
[16] D.M. Lucantoni, K.S. Meier-Hellstern and M.F. Neuts, A single-server queue with server vacations and a class of non-renewal arrival processes, Adv. in Appl. Probab. 22 (1990) 676–705.
[17] D.E. McDysan and D.L. Spohn, *ATM Theory and Application* (McGraw-Hill, New York, 1995).
[18] M.F. Neuts, A versatile Markovian point process, J. Appl. Probab. 16 (1979) 746–779.

[19] M.F. Neuts, *Structured Stochastic Matrices of $M/G/1$ Type and Their Applications* (Marcel Dekker, New York, 1989).

[20] H. Saito, The departure process of an $N$/G/1 queue, Performance Evaluation 11 (1990) 241–251.

[21] L. Takács, *Introduction to the Theory of Queues* (Oxford Univ. Press, Oxford, 1962).

[22] P.C. Yeh, Departure process and queueing analyses for several classes of single server queueing systems, Master thesis, Department of Elect. Engineering, National Taiwan University, Taipei, Taiwan (1998), chapter 3.

[23] P.C. Yeh and J.F. Chang, Characterizing the departure process of a single server queue from the embedded Markov renewal process at departures, Queueing Systems 35 (2000) 381–395.