University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Mehra, S;Taylor, PG

Title:
Open networks of infinite server queues with non-homogeneous multivariate batch Poisson arrivals

Date:
2023-12-01

Citation:
Mehra, S. & Taylor, P. G. (2023). Open networks of infinite server queues with non-homogeneous multivariate batch Poisson arrivals. Queueing Systems, 105 (3-4), pp.171-187. https://doi.org/10.1007/s11134-023-09891-x.

Persistent Link:
https://hdl.handle.net/11343/344392

License:
cc-by

Queueing Systems (2023) 105:171–187
https://doi.org/10.1007/s11134-023-09891-x

Open networks of infinite server queues with non-homogeneous multivariate batch Poisson arrivals

Somya Mehra[1] · **Peter G. Taylor**[1]

Received: 28 March 2023 / Revised: 20 August 2023 / Accepted: 1 September 2023 /
Published online: 7 October 2023
© The Author(s) 2023

Abstract
In this paper, we consider the occupancy distribution for an open network of infinite server queues with multivariate batch arrivals following a non-homogeneous Poisson process, and general service time distributions. We derive a probability generating function for the transient occupancy distribution of the network and prove that it is necessary and sufficient for ergodicity that the expected occupancy time for each batch be finite. Further, we recover recurrence relations for the transient probability mass function formulated in terms of a distribution obtained by compounding the batch size with a multinomial distribution.

Keywords Infinite server queue · Open network · Multivariate batch arrivals

1 Introduction

Infinite server queues with Poisson batch arrivals—denoted $M_t^X/G/\infty$ queues in Kendall's notation—have elicited significant research attention, from the canonical work of [2, 20, 22] in the 1960 s, to studies by Liu and Templeton [3, 13] and the recent treatise of [6]. The independence between customers within the queue—termed the 'non-interacting property' by [13]—allows for an analytic characterisation of the numbers of busy servers and departures, typically in the form of a probability generating function (PGF) [2, 3, 6].

A natural generalisation is an open network of infinite server queues, with multivariate batch arrivals following a non-homogeneous Poisson process and general service time distributions. With customers arriving one at a time, networks of $M_t/G/\infty$ queues have been studied by refs [7, 9, 14], amongst others. In contrast, networks with multivariate batch arrivals have received comparatively little attention. Models have been

✉ Somya Mehra
somya.mehra@unimelb.edu.au

[1] School of Mathematics and Statistics, The University of Melbourne, Parkville, Australia

proposed for specific applications such as private-line telecommunication services [15] and malarial relapses [16], while stochastic fluid networks have been shown to arise as scaling limits of appropriate batch-arrival infinite server queueing networks [10]. To the best of our knowledge, results for more general networks of $\cdot/G/\infty$ queues with multivariate batch arrivals are yet to appear in the literature.

Here, we analyse the occupancy distribution in a network of infinite server queues, with multivariate batches arriving according to a non-homogeneous Poisson process. Our model assumptions are set out in Sect. 2. We derive the PGF for the transient queue length distribution in Theorem 3.1, extending the construction we adopted in [16]; similar arguments have been presented previously by [2, 3, 7] and others.

While infinite server queues with single arrivals are necessarily stable when the expected service time is finite, the queue length Markov chain can be transient or null-recurrent when batch arrivals are permitted. Stability conditions for the $\mathrm{M}^{\mathrm{X}}/\mathrm{G}/\infty$ queue, accommodating heterogenous customers within each batch, have been characterised by [5]. When service times are exponential, [30] have shown that it is necessary and sufficient for an infinite server queue with a batch Markovian arrival process to be stable that the batch size distribution has a bounded logarithmic moment. For networks with multivariate batch arrivals, we extend these results in Sect. 4 to show that a necessary and sufficient condition for ergodicity is that the expected occupancy time for each batch be finite (Theorem 4.1). When the customer occupancy time distributions have exponential tails, this is equivalent to the logarithmic moment condition of [5, 30] (Corollaries 4.1.2 and 4.1.3).

We recover recurrence relations for the probability mass function (PMF) of the time-dependent queue length distribution, formulated with respect to the PMF of the batch size compounded with a multinomial distribution in Theorem 5.1. The utility of these formulae is constrained by the underlying compound distribution. In the context of the $\mathrm{M}^{\mathrm{X}}/\mathrm{G}/\infty$ queue, [27–29] found that for distributions within 'Sundt and Jewell's family'—encompassing the binomial, Poisson, logarithmic, geometric and negative binomial distributions, and zero-inflated analogues thereof [26]—the recurrence relations may be computationally tractable. Here, we find that univariate batch sizes within 'Sundt and Jewell's family' [26] may likewise yield tractable formulae when the initial queue allocation for each customer is independent and identically distributed (i.i.d.) (Sect. 5.1.1).

2 Model structure

Consider an open network of J infinite server queues, indexed by $j = 1, \ldots, J$, such that

- Arrivals occur at points $T_1, T_2, \ldots$ of a non-homogeneous Poisson process with rate $\lambda(t)$.
- At time point T_i, for $j = 1, \ldots, J$, Σ_{ij} customers arrive at queue j. The vectors $\Sigma_i = (\Sigma_{i1}, \ldots, \Sigma_{iJ})$ are independent vector-valued random variables with non-negative integer components selected from a multivariate distribution with probability generating function

$$G_{\mathbf{S}}(z_1, \ldots, z_J) := \mathbb{E}\left[\prod_{j=1}^{J} z_j^{S_j}\right] \tag{1}$$

defined on a subset $\mathcal{D} \subseteq \mathcal{C}^J$ that contains the J-dimensional unit polydisc with $|z_j| \leq 1$. Note that the distribution of Σ_i does not depend on i. We use the generic random variable $\mathbf{S}$ to denote a random variable with this distribution.

- The ℓth customer that enters queue j at time point T_i is marked with a stochastic process $X_{i,\ell}^j(t)$ where $X_{i,\ell}^j(t) = k$ if the customer will be in queue k at time $T_i + t$. For each i and $\ell = 1, \ldots, \Sigma_{ij}$, the $X_{i,\ell}^j(t)$ are selected independently with a law such that $P(X_{i,\ell}^j(t) = k) = q_k^j(t)$.

Under this model, the routes of customers through the network (that is, the sequence of visited nodes and associated service times) are mutually independent and identically distributed, conditional on the node of entry. This 'non-interacting property' [13] enables explicit analysis. Here, we focus on the number of customers in the network at time t, denoted by the random vector $\mathbf{N}(t) = (N_1(t), \ldots, N_J(t))$.

Remark 2.1 We make no assumptions on the probability $q_k^j(t)$, barring the constraint that it is non-negative and $\sum_k q_k^j(t) \leq 1$ is a monotonically decreasing function of t. A common example may be where each customer follows a continuous time Markov chain with transition rate matrix $\mathbf{Q}$ through a set of compartments, in which case $q_k^j(t)$ is the $(j,k)^{\text{th}}$ entry of the matrix exponential $e^{\mathbf{Q}t}$ (see, for example, the model of hypnozoite dynamics in [16]). However, the quantity $q_k^j(t)$ could be more complicated, with a customer staying in each node for arbitrary service times that may be correlated; we require the independence of customer paths, but not along customer paths. In generality, a straightforward expression for $q_k^j(t)$ may not be readily available.

For notational convenience, we introduce the random vector $\mathbf{C}_i(t) = (C_{i,1}(t), \ldots, C_{i,J}(t))$, with

$$C_{i,k}(t) = \sum_{j=1}^{J} \sum_{\ell=1}^{\Sigma_{ij}} \mathbb{1}\{X_{i,\ell}^j(t) = k\} \tag{2}$$

giving the number of customers from the ith batch that are in queue k at time $T_i + t$. It follows from our assumptions above that the distribution of $\mathbf{C}_i(t)$ is independent of i and we denote by $\mathbf{C}(t)$ a generic random variable with this distribution.

3 The transient queue length distribution

We first characterise the time-dependent PGF for the state of the network, generalising the construction of [16] (which was used to analyse an open network of infinite server queues tailored to within-host superinfection and hypnozoite dynamics for *Plasmodium vivax* malaria).

Theorem 3.1 *(Multivariate PGF) For* $z = (z_1, \ldots, z_J) \in \mathcal{D}$

$$\phi_t(\mathbf{z}) := \mathbb{E}\left[\prod_{k=1}^{J} z_j^{N_j(t)}\right]$$
$$= \exp\left\{-\int_0^t \lambda(\tau)\left[1 - G_{\mathbf{S}}\left(1 + \sum_{k=1}^{J}(z_k - 1)q_k^1(t-\tau), \ldots, 1 + \sum_{k=1}^{J}(z_k - 1)q_k^J(t-\tau)\right)\right]\mathrm{d}\tau\right\}. \quad (3)$$

Proof We use similar reasoning to [16]. In short, we condition on the multivariate batch size, and the sequence of arrival times. For related systems, similar proofs have been previously presented by others, including [2, 3, 7].

We begin by deriving the PGF for the generic random vector $\mathbf{C}(t)$ described in Equation (2). Since $X_{i,\ell}^j(t)$ are independent random variables taking values in $\{1, \ldots, J\}$, the number of customers originating in queue j that are in queue k at time $T_i + t$ has a multinomial distribution with parameters $q_k^j(t)$. Therefore, conditional on a multivariate batch of size $\Sigma_{i1} = n_1, \ldots, \Sigma_{iJ} = n_J$,

$$\mathbb{E}\left[\prod_{k=1}^{J} z_k^{C_k(t)} \,\middle|\, \Sigma_{i1} = n_1, \ldots, \Sigma_{iJ} = n_J\right]$$
$$= \prod_{j=1}^{J}\left(1 + \sum_{k=1}^{J}(z_k - 1)q_k^j(t)\right)^{n_k}.$$

By the law of total expectation, it follows that

$$\mathbb{E}\left[\prod_{k=1}^{J} z_k^{C_k(t)}\right] = \sum_{n_1=0}^{\infty}\cdots\sum_{n_J=0}^{\infty} P(\Sigma_{i1} = n_1, \ldots, \Sigma_{iJ} = n_J)$$
$$\times \mathbb{E}\left[\prod_{k=1}^{J} z_k^{C_k(t)} \,\middle|\, \Sigma_{i1} = n_1, \ldots, \Sigma_{iJ} = n_J\right]$$
$$= G_{\mathbf{S}}\left(1 + \sum_{k=1}^{J}(z_k - 1)q_k^1(t), \ldots, 1 + \sum_{k=1}^{J}(z_k - 1)q_k^J(t)\right). \quad (4)$$

Equation (4) is the PGF for the occupancy of the network time t after the arrival of a single batch.

Under the assumption that arrivals follow a non-homogeneous Poisson process, the number of arrivals $M(t)$ in the interval $[0, t)$

$$M(t) \sim \text{Poisson}(\Lambda(t))$$

where

$$\Lambda(t) = \int_0^t \lambda(\tau)\mathrm{d}\tau.$$

Following [7], conditional on the event $\{M(t) = m\}$, the distribution of $T_1, \ldots, T_m$ is the same as the distribution of the order statistics of m i.i.d. variables with density

$$f(\tau) = \frac{\lambda(\tau)}{\Lambda(t)} \mathbb{1}_{\{\tau \in [0,t)\}}.$$

Therefore, given $M(t) = m$ arrivals in the interval $[0, t)$, we can apply the law of total probability over the arrival times $T_1, \ldots, T_m$ to yield the conditional PGF

$$\mathbb{E}\left[\prod_{k=1}^{J} z_k^{N_k(t)} \middle| M(t) = m\right] = \left(\int_0^t \frac{\lambda(\tau)}{\Lambda(t)} \cdot \mathbb{E}\left[\prod_{k=1}^{J} z_k^{C_k(t-\tau)}\right] d\tau\right)^m.$$

Using the law of total expectation, we thus deduce

$$\begin{aligned}\mathbb{E}\left[\prod_{j=1}^{J} z_j^{N_j(t)}\right] &= \sum_{m=0}^{\infty} \mathbb{E}\left[\prod_{j=1}^{J} z_j^{N_j(t)} \middle| M(t) = m\right] \cdot P(M(t) = m) \\ &= \exp\left\{-\int_0^t \lambda(\tau)\left(1 - \mathbb{E}\left[\prod_{j=1}^{J} z_j^{C_j(t-\tau)}\right]\right) \mathrm{d}\tau\right\}. \end{aligned} \tag{5}$$

Substituting Eq. (4) into Eq. (5) yields expression (3). □

4 Necessary and sufficient conditions for ergodicity

Here, we characterise the stability, that is, ergodicity, of the network, assuming a homogeneous arrival rate $\lambda(t) = \lambda$. For M/G/∞ queues, a necessary and sufficient condition for ergodicity is that the occupation time distribution has finite expectation. The analogous constraint in this setting is that at least one customer from each batch will be present somewhere in the network for a time W that has finite expectation.

A customer who arrives to queue j at time T has left the network before time $T + t$ with probability

$$1 - Q_j(t) = 1 - \sum_{k=1}^{J} q_k^j(t).$$

The expected time that at least one customer from a batch is present in the network can then be written

$$\begin{aligned}\mathbb{E}[W] &= \sum_{n_1=0}^{\infty} \cdots \sum_{n_J=0}^{\infty} P(\mathbf{S} = (n_1, \ldots, n_J)) \cdot \int_0^{\infty} \left(1 - \prod_{j=1}^{J} \left[1 - Q_j(t)\right]^{n_j}\right) d\tau \\ &= \int_0^{\infty} \left[1 - G_{\mathbf{S}}\Big(1 - Q_1(\tau), \ldots, 1 - Q_J(\tau)\Big)\right] d\tau. \end{aligned} \tag{6}$$

where we have used Tonelli's theorem to swap the order of summation and integration; and noted that the resultant integrand can be written as a sum of two absolutely convergent series. This expression is central to the ergodicity of the network.

Theorem 4.1 *The network is ergodic if and only if* $\mathbb{E}[W] < \infty$

Proof Let $\Omega \in \mathcal{C}^J$ denote the open J-dimensional polydisc with $|z_j| < 1$ for each $j \in \{1, \ldots, J\}$. By the dominated convergence theorem, for all $\mathbf{z} \in \Omega$,

$$\lim_{t\to\infty} \mathbb{E}\left[\prod_{k=1}^{J} z_k^{N_j(t)}\right] = \mathbb{E}\left[\prod_{k=1}^{J} z_k^{\lim_{t\to\infty} N_j(t)}\right]$$

since $P(\mathbf{N}(t) = \mathbf{n}) \leq 1$ for all $\mathbf{n} \in \mathbb{Z}^J$, $t \geq 0$.

Observe that the transient PGF $\phi_t(\mathbf{z})$ for $\mathbf{N}(t)$ defines a family of functions $\{\phi_t\}_{t\geq 0}$ that is both analytic and bounded on Ω, with $|\phi_t(\mathbf{z})| \leq 1$ for all $t \geq 0$. By Montel's theorem, there exists a subsequence ϕ_{t_i} which converges uniformly on all compact subsets $K \subset \Omega$ (Theorem 1.7.3 of [11]). It follows from a theorem of Weierstrass (Theorem 1.7.1 of [11]) that $\lim_{t\to\infty} \phi_t$ is analytic on Ω, with the uniform convergence of partial derivatives

$$\lim_{t_i\to\infty} \left\{ \frac{\partial^{v_1+\cdots+v_n}}{\prod_{k=1}^n \partial z_k^{v_k}} \phi_{t_i}(\mathbf{z}) \right\} = \frac{\partial^{v_1+\cdots+v_n}}{\prod_{k=1}^n \partial z_k^{v_k}} \left\{ \lim_{t\to\infty} \phi_t(\mathbf{z}) \right\}$$

on all compact subsets $K \subset \Omega$. In particular, for all $\mathbf{v} \in \mathbb{Z}^J$, we note that

$$\lim_{t_i\to\infty} P(\mathbf{N}(t_i) = \mathbf{v}) = \frac{\partial^{v_1+\cdots+v_n}}{\prod_{k=1}^n \partial z_k^{v_k}} \left\{ \lim_{t\to\infty} \phi_t(\mathbf{z}) \right\}\Big|_{\mathbf{z}=\mathbf{0}}.$$

It is therefore sufficient to consider the infinite time limit of the multivariate PGF (3)

$$\lim_{t\to\infty} \phi_t(\mathbf{z}) = \exp\left\{ -\lambda \int_0^\infty \left[1 - G_{\mathbf{S}}\Big(1 + \sum_{k=1}^J (z_k - 1) q_k^1(\tau), \ldots, 1 + \sum_{k=1}^J (z_k - 1) q_k^J(\tau)\Big)\right] \mathrm{d}\tau \right\}.$$

As in [5], it is necessary and sufficient for the limiting function to be a PGF that the multivariate limit

$$\lim_{z_k \uparrow 1,\, k\in\{1,\ldots,J\}} \left\{ \lim_{t\to\infty} \phi_t(\mathbf{z}) \right\} = 1,$$

where $(z_1, \ldots, z_J) \to (1, \ldots, 1)$ along any path through Ω. Since the generating function is analytic inside Ω, this is equivalent to the condition

$$\lim_{z\uparrow 1} \int_0^\infty \Big[1 - G_{\mathbf{S}}\Big(1 + (z-1)Q_1(\tau), \ldots, 1 + (z-1)Q_J(\tau)\Big)\Big] \mathrm{d}\tau \Big\} = 0. \tag{7}$$

For notational convenience, we denote the integrand

$$H(z, \tau) := 1 - G_{\mathbf{S}}\big(1 + (z-1)Q_1(\tau), \ldots, 1 + (z-1)Q_J(\tau)\big),$$

and observe from Eq. (6) that

$$\mathbb{E}[W] = \int_0^\infty H(0, \tau)\mathrm{d}\tau.$$

Since G_S is a multivariate PGF, for fixed τ, $H(z, \tau)$ is a decreasing function of z in the domain $[0, 1]$. In particular, for each $\tau \geq 0$ we have pointwise convergence $H(z, \tau) \to 0$ as $z \to 1$, with $|H(z, \tau)| \leq H(0, \tau)$. Suppose $\mathbb{E}[W] < \infty$. Then by the dominated convergence theorem,

$$\lim_{z\uparrow 1} \int_0^\infty H(z, \tau)\mathrm{d}\tau = \int_0^\infty \lim_{z\uparrow 1} H(z, \tau)\mathrm{d}\tau = 0.$$

We thus see that $\mathbb{E}[W] < \infty$ is a sufficient condition for Eq. (7) to hold.

Now, suppose that the network is ergodic. Then the limiting probability of the queue being empty can be written

$$\lim_{t\to\infty} P(\mathbf{N}(t) = \mathbf{0}) = e^{-\lambda\,\mathbb{E}[W]} > 0,$$

which necessarily implies $\mathbb{E}[W] < \infty$. □

In direct analogy to Lemma 1 of [5], under the constraint that the time spent in the network by each customer has finite expectation—which is necessary for batch occupancy time to have finite expectation, that is, $\mathbb{E}[W] < \infty$ —we recover a sufficient condition for ergodicity.

Corollary 4.1.1 *Suppose the network occupation time distributions $Q_j(t)$ have finite expectation. Then the network is ergodic if $\mathbb{E}[S_1 + \cdots + S_J] < \infty$.*

Proof Similarly to [5], we observe that

$$1 - \prod_{j=1}^{J} [1 - Q_j(\tau)]^{n_j} \leq \sum_{j=1}^{J} n_j Q_j(\tau).$$

Therefore, from Eq. (6), it follows that

$$\begin{aligned}\mathbb{E}[W] &\leq \sum_{n_1=0}^{\infty} \cdots \sum_{n_J=0}^{\infty} P(\mathbf{S} = (n_1, \ldots, n_J)) \int_0^\infty \sum_{j=1}^{J} n_j Q_j(\tau) d\tau \\ &= \sum_{j=1}^{J} \mathbb{E}[S_j] \cdot \int_0^\infty Q_j(\tau) d\tau.\end{aligned}$$

A finite expected network occupation time for each customer $\int_0^\infty Q_j(\tau)d\tau < \infty$ and a finite mean batch size $\mathbb{E}[S_j] < \infty$ for each $j \in \{1, \ldots, J\}$ is thus a sufficient condition for $\mathbb{E}[W] < \infty$ and, by Theorem 4.1, the ergodicity of the network. □

For $M^X/M/\infty$ queues, [5, 30] have shown that a necessary and sufficient condition for stability is the batch size S has a finite logarithmic moment, that is,

$$\mathbb{E}[\log(S+1)] < \infty.$$

Given an exponential service time of mean duration μ per customer, the expected occupancy time for a batch can written

$$\mathbb{E}[W] = \sum_{j=1}^{\infty} P(S=j) \sum_{k=1}^{j} \frac{1}{\mu k}.$$

Since

$$\sum_{j=1}^{n} \frac{1}{j} - \log(n+1) \to \gamma_e \text{ as } n \to \infty \tag{8}$$

where γ_e denotes the Euler–Mascheroni constant [25], in the case of the $M^X/M/\infty$ queue

$$\mathbb{E}[\log(S+1)] < \infty \iff \mathbb{E}[W] < \infty.$$

We can generalise this observation to recover a sufficient condition for ergodicity when the network occupancy time distributions $Q_j(t)$ have exponentially-bounded tails, that is, there exists $\delta > 0, t_0 > 0$ such that

$$Q_j(\tau) \le e^{-\delta\tau} \text{ for all } \tau \ge t_0 \text{ and } j \in \{1, \ldots, J\}.$$

Corollary 4.1.2 *Suppose the occupancy time distributions $Q_j(t)$ have exponentially bounded tails. Then the network is ergodic if $\mathbb{E}[\log(S_1 + \cdots + S_J + 1)] < \infty$.*

Proof By assumption, there exist $\delta > 0, t_0 > 0$ such that

$$\begin{aligned} 0 &< G_{\mathbf{S}}\big(1 - e^{-\delta\tau}, \ldots, 1 - e^{-\delta\tau}\big) \\ &\le G_{\mathbf{S}}\big(1 - Q_1(\tau), \ldots, 1 - Q_J(\tau)\big) \le 1 \text{ for all } \tau \ge t_0 \end{aligned}$$

since $G_{\mathbf{S}}$ is a generating function. Equation (6) then yields the bound

$$\mathbb{E}[W] \le t_0 + \int_{t_0}^{\infty} \Big[1 - G_{\mathbf{S}}\big(1 - e^{-\delta\tau}, \ldots, 1 - e^{-\delta\tau}\big)\Big] d\tau.$$

Setting $S = S_1 + \cdots + S_J$, we observe that

$$G_{\mathbf{S}}\big(1 - e^{-\delta\tau}, \ldots, 1 - e^{-\delta\tau}\big) = \sum_{n=0}^{\infty} P(S=n)(1 - e^{-\delta\tau})^n.$$

Therefore, a *sufficient* condition for $\mathbb{E}[W] < \infty$—and, by Theorem 4.1, the ergodicity of the network—is that

$$Y := \int_0^\infty \sum_{n=0}^{\infty} P(S=n)\big[1-(1-e^{-\delta\tau})^n\big]\mathrm{d}\tau < \infty.$$

Using the binomial expansion and identity 0.155.4 of [8], we compute.

$$Y = \sum_{n=1}^{\infty} P(S=n) \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k+1}}{\delta k} = \sum_{n=1}^{\infty} P(S=n) \sum_{k=1}^{n} \frac{1}{\delta k},$$

so from Eq. (8) [25]

$$Y < \infty \iff \mathbb{E}[\log(S+1)] < \infty.$$

Therefore, $\mathbb{E}[\log(S+1)] < \infty$ is a sufficient condition for the ergodicity of the network. □

For Markovian queueing networks, where the time spent by a customer is each node is exponentially-distributed, we can strengthen Corollary 4.1.2 to recover a necessary and sufficient condition for ergodicity.

Corollary 4.1.3 *For each $j \in \{1,\dots,J\}$, suppose there exists $\delta_j > 0$ such that $Q_j(t) = \Theta(e^{-\delta_j \tau})$ in the limit $\tau \to \infty$. Then the network is ergodic if and only if $\mathbb{E}[\log(S_1 + \cdots + S_J + 1)] < \infty$.*

Proof Sufficiency follows directly from Corollary 4.1.2. By assumption, there exists $\eta > \delta_j$ and $t_0 > 0$ such that

$$Q_j(\tau) \geq e^{-\eta\tau} \text{ for all } \tau \geq t_0 \text{ and } j \in \{1,\dots,J\}.$$

Set $S = S_1 + \cdots + S_J$. Adopting much the same reasoning as the proof of Corollary 4.1.2, we obtain a lower bound for the expected batch occupancy time

$$\mathbb{E}[W] + t_0 \geq \int_0^\infty \Big[1 - G_{\mathbf{S}}\big(1-e^{-\eta\tau},\dots,1-e^{-\eta\tau}\big)\Big]d\tau = \sum_{n=1}^{\infty} P(S=n) \sum_{k=1}^{n} \frac{1}{\eta k}.$$

From Eq.(8) [25], we observe that

$$\begin{aligned} \mathbb{E}[\log(S+1)] = \infty &\implies \sum_{n=1}^{\infty} P(S=n) \sum_{k=1}^{n} \frac{1}{\eta k} = \infty \\ &\implies \mathbb{E}[W] = \infty, \end{aligned}$$

so by Theorem 4.1, $\mathbb{E}[\log(S+1)] < \infty$ is a necessary condition for ergodicity. □

In a similar vein, we can recover a sufficient condition for ergodicity in the case of fat-tailed customer occupancy time distributions.

Corollary 4.1.4 *Suppose there exists $\alpha > 1$ and $t_0 > 0$ such that $Q_j(t) \leq t^{-\alpha}$ for all $j \in \{1, \ldots, J\}$ and $t \geq t_0$. Then the network is ergodic if $\mathbb{E}[(S_1 + \cdots + S_J)^{1/\alpha}] < \infty$.*

Proof Using similar reasoning to Corollary 4.1.2, we obtain the lower bound

$$\mathbb{E}[W] \leq t_0 + \int_{t_0}^{\infty} \sum_{n=0}^{\infty} P(S = n)\left[1 - (1 - \tau^{-\alpha})^n\right] d\tau.$$

Performing a change of variables $u = 1 - \tau^{-\alpha}$, we find that the condition

$$Z := \sum_{n=0}^{\infty} P(S = n) \int_0^1 \left[\sum_{i=0}^{n-1} (1-u)^{-\frac{1}{\alpha}} u^i \right] \mathrm{d}u < \infty \tag{9}$$

is sufficient to ensure $\mathbb{E}[W] < \infty$, where the interchanging of the order of summation and integration is justified by Tonelli's theorem. Using the integral representation of the beta function and identity 8.384.1 of [8], we compute

$$\begin{aligned} \int_0^1 \left[\sum_{i=0}^{n-1} (1-u)^{-1/\alpha} u^i \right] \mathrm{d}u &= \sum_{k=0}^{n-1} B\left(1 - \frac{1}{\alpha}, k+1\right) \\ &= \Gamma\left(1 - \frac{1}{\alpha}\right) \sum_{k=0}^{n-1} \frac{\Gamma(k+1)}{\Gamma\left(k + 2 - \frac{1}{\alpha}\right)}. \end{aligned} \tag{10}$$

Substituting Eq. (10) in (9) and interchanging the order of summation (which is justified by Tonelli's theorem) yields

$$Z = \Gamma\left(1 - \frac{1}{\alpha}\right) \sum_{k=0}^{\infty} \frac{\Gamma(k+1)}{\Gamma\left(k + 2 - \frac{1}{\alpha}\right)} P(S \geq k+1).$$

By identity 6.1.46 of [1]

$$\frac{\Gamma(k+1)}{\Gamma\left(k + 2 - \frac{1}{\alpha}\right)} \sim k^{\frac{1}{\alpha} - 1} \text{ in the limit } k \to \infty.$$

Therefore,

$$Z < \infty \iff \sum_{k=1}^{\infty} k^{\frac{1}{\alpha}-1} P(S \geq k+1) = \sum_{n=2}^{\infty} P(S = n) \sum_{k=1}^{n-1} k^{\frac{1}{\alpha}-1} < \infty$$

where we have interchanged the order of summation. Noting that

$$\sum_{k=1}^{n} k^{\frac{1}{\alpha}-1} \sim \frac{1}{\alpha} n^{\frac{1}{\alpha}} \text{ in the limit } n \to \infty$$

(identity 0.121 of [8]), we further deduce that

$$\mathbb{E}\left[S^{\frac{1}{\alpha}}\right] < \infty \iff \sum_{n=1}^{\infty} P(S=n) \sum_{k=1}^{n} k^{\frac{1}{\alpha}-1} < \infty.$$

Therefore, $\mathbb{E}\left[S^{\frac{1}{\alpha}}\right] < \infty \implies \mathbb{E}[W] < \infty$ and the claim follows from Theorem 4.1. □

5 Recurrence relations for the PMF of the transient occupancy distribution

For a single batch arrival ($M^X/G/\infty$) queue with a homogeneous arrival process, [27–29] recognised the PGF governing the number of customers at a fixed time t as that of a compound Poisson distribution. They observed that application of the Panjer–Adleson recursion scheme [19, 24] yields recurrence relations for the PMF of the time-dependent queue length distribution, formulated with respect to the PMF of a distribution obtained by compounding the batch size with a binomial distribution. Using the transient PGF given by Eq. (3), we can likewise recover a recurrence relation for the PMF of the time-dependent occupancy distribution, formulated in terms of the PMF of $\mathbf{C}(t)$ (Eq. (2)), using an analogous argument to [19].

Theorem 5.1 *(A recurrence relation for the multivariate PMF.)*
For $v \in \{2, \ldots, J\}$ such that $n_v \geq 1$,

$$\begin{aligned} P\big[\mathbf{N}(t) = (n_1, \ldots, n_v, 0, \ldots 0)\big] \\ = \sum_{i_1=0}^{n_1} \cdots \sum_{i_{v-1}=0}^{n_{v-1}} \sum_{i_v=1}^{n_v} \left[\frac{i_v}{n_v} \cdot P\big[\mathbf{N}(t) = (n_1 - i_1, \ldots, n_v - i_v, 0, \ldots, 0)\big] \right. \\ \left. \cdot \int_0^t \lambda(\tau) \cdot P\big[\mathbf{C}(t-\tau) = (i_1, \ldots, i_v, 0, \ldots, 0)\big] \mathrm{d}\tau \right]. \end{aligned} \tag{11}$$

Proof Following the proof of Theorem 3.1, given that there are $M(t) = m$ arrivals in the interval $[0, t)$, the PMF for the number of customers in each queue $1, \ldots, J$ at time t attributable to each arrival event is i.i.d. with PMF

$$f(i_1, \ldots, i_J) = \int_0^t \frac{\lambda(\tau)}{\Lambda(t)} P\big[\mathbf{C}(t-\tau) = (i_1, \ldots, i_J)\big] \mathrm{d}\tau.$$

Denote by $f^{(m)}$ the m-fold convolution of f. Then by the law of total probability, the PMF of $\mathbf{N}(t)$ can be written in the form

$$P\big[\mathbf{N}(t) = (n_1, \ldots, n_J)\big] = \sum_{m=0}^{\infty} \frac{e^{-\Lambda(t)}\Lambda(t)^m}{m!} f^{(m)}(n_1, \ldots, n_J). \tag{12}$$

Now, by symmetry,

$$\sum_{i_1=0}^{n_1} \cdots \sum_{i_J=0}^{n_J} \frac{i_v}{n_v} f(i_1, \ldots, i_J) \cdot f^{(m-1)}(n_1 - i_1, \ldots, n_J - i_J) = \frac{f^{(m)}(n_1, \ldots, n_J)}{m}; \tag{13}$$

that is, given that the sum of m i.i.d. random vectors, each with PMF f, is $(n_1, \ldots, n_J)$, the conditional mean of the vth element is n_v/m (Relation II of [19]).

Using Eqs. (12) and (13), for $(n_1, \ldots, n_J) \neq \mathbf{0}$, we obtain the expression

$$\begin{aligned} &P\big[\mathbf{N}(t) = (n_1, \ldots, n_J)\big] \\ &\quad = \sum_{i_1=0}^{n_1} \cdots \sum_{i_J=0}^{n_J} \frac{i_v}{n_v} f(i_1, \ldots, i_j) \sum_{m=1}^{\infty} \frac{e^{-\Lambda(t)}\Lambda(t)^m}{(m-1)!} f^{(m-1)}(n_1 - i_1, \ldots, n_J - i_J) \\ &\quad = \sum_{i_1=0}^{n_1} \cdots \sum_{i_J=0}^{n_J} \frac{i_v}{n_v} \Lambda(t) f(i_1, \ldots, i_j) P\big[\mathbf{N}(t) = (n_1 - i_1, \ldots, n_J - i_J)\big] \end{aligned} \tag{14}$$

where interchanging the order of summation is justified since the series is absolutely convergent. Sequential application of Eq. (14) along the lower triangular integer lattice yields precisely the recurrence relation (11). The expression (11) can also be obtained by computing partial derivatives of the multivariate PGF (3), using the formulae provided in [18]. □

Through an application of Faa di Bruno's formula, we noted in [16] that the transient distribution for the $\mathrm{M}_t^X/\mathrm{G}/\infty$ queue can also be formulated in terms of Bell polynomials [4], which yield an analogous recurrence structure. For queueing networks with multivariate batch arrivals, alternative representations of the transient PMF in terms of multivariate Bell polynomials [21] can also be recovered.

Likewise, we can derive recurrence relations for the joint factorial moments of $\mathbf{N}(t)$, formulated with respect to those of $\mathbf{C}(t)$.

5.1 Some special cases

The recurrence relations derived in Theorem 3.1 are formulated in terms of the PMF for the distribution of $\mathbf{C}(t)$. The tractability of these formulae are therefore constrained by the multivariate PMF for $\mathbf{C}(t)$. Here, we provide explicit formulae for several special cases.

5.1.1 Univariate batch sizes

For $M^X/G/\infty$ queues, [27, 28] observed that batch sizes within 'Sundt and Jewell's family of discrete distributions' [23, 26] may yield tractable formulae. This class of distributions has a probability mass function that is characterised by a recurrence relation of the form

$$P(S = n) = P(S = n-1) \cdot \left(a + \frac{b}{n}\right)$$

for fixed a, b and $n \geq 2$. It was shown in Theorem 1 of [26] that this class includes Poisson, negative binomial and logarithmic distributions, and zero-inflated analogues thereof (that is, mixtures of the relevant distribution and a point mass at zero).

In the context of queueing networks, we restrict our attention to multivariate batch size distributions $\mathbf{S}$ that can be expressed in the form

$$\mathbb{E}\left[\prod_{j=1}^{J} z_j^{S_j}\right] = G_S\left(\sum_{j=1}^{J} p_j z_j\right)$$

where S is a univariate random variable within Sundt and Jewell's family [23, 26] and $p_j \geq 0$ with $\sum_j p_j = 1$. This corresponds to a univariate batch, with each incoming customer independently assigned an entry point $j \in \{1, \ldots, J\}$. For notational convenience, we set

$$q_k(t) = \sum_{j=1}^{J} p_j \cdot q_k^j(t)$$

to be the probability that a customer is situated in queue k time t after their arrival into the network. In the corollaries below, we state explicit formulae for binomial, Poisson, negative binomial and logarithmic batch sizes.

Corollary 5.1.1 *Suppose $S \sim Binomial(N, \alpha)$. Then for $v \in \{2, \ldots, J\}$ and $n_v \geq 1$*

$$\begin{aligned} &P\big[\mathbf{N}(t) = (n_1, \ldots, n_v, 0, \ldots 0)\big] \\ &\quad = \sum_{i_1=0}^{n_1} \cdots \sum_{i_{v-1}=0}^{n_{v-1}} \sum_{i_v=1}^{n_v} \left[\frac{i_v}{n_v} \cdot P\big[\mathbf{N}(t) = (n_1 - i_1, \ldots, n_v - i_v, 0, \ldots 0)\big] \cdot \mathbb{1}_{\sum_{j=1}^{v} i_j \leq N}\right. \\ &\quad \left. \cdot \int_0^t \lambda(\tau) \frac{N!}{\left(N - \sum_{k=1}^{v} i_k\right)!} \Big(1 - \alpha \sum_{k=1}^{v} q_k(t-\tau)\Big)^{N - \sum_{k=1}^{v} i_k} \prod_{k=1}^{v} \frac{(\alpha q_k(t-\tau))^{i_k}}{i_k!} d\tau\right]. \end{aligned}$$

Corollary 5.1.2 *Suppose $S \sim Poisson(\mu)$. Then for $v \in \{2, \ldots, J\}$ and $n_v \geq 1$*

$$\begin{aligned} &P\big[\mathbf{N}(t) = (n_1, \ldots, n_v, 0, \ldots 0)\big] \\ &\quad = \sum_{i_1=0}^{n_1} \cdots \sum_{i_{v-1}=0}^{n_{v-1}} \sum_{i_v=1}^{n_v} \left[\frac{i_v}{n_v} \cdot P\big[\mathbf{N}(t) = (n_1 - i_1, \ldots, n_v - i_v, 0, \ldots, 0)\big]\right. \end{aligned}$$

$$\cdot \int_0^t \lambda(\tau)\left[\prod_{k=1}^{v}\frac{(q_k(t-\tau)\mu)^{i_k}}{i_k!}\right]e^{-\mu}\mathrm{d}\tau\Bigg].$$

Corollary 5.1.3 *Suppose S follows a negative binomial distribution with PGF*

$$G(z) := \mathbb{E}\left[z^S\right] = \frac{1}{\left[1+\nu(1-z)\right]^r}.$$

Then for $v \in \{2,\ldots,J\}$ and $n_v \geq 1$

$$\begin{aligned}
&P\left[\mathbf{N}(t) = (n_1,\ldots,n_v,0,\ldots 0)\right] \\
&\quad = \sum_{i_1=0}^{n_1}\cdots\sum_{i_{v-1}=0}^{n_{v-1}}\sum_{i_v=1}^{n_v}\left[\frac{i_v}{n_v}\cdot P\left[\mathbf{N}(t) = (n_1-i_1,\ldots,n_v-i_v,0,\ldots 0)\right]\right. \\
&\quad\quad \left.\cdot\int_0^t \lambda(\tau)\frac{\Gamma\left(r+\sum_{k=1}^{v} i_k\right)}{\Gamma(r)\prod_{k=1}^{v} i_k!}\frac{\nu^{\sum_{k=1}^{v} i_k}}{\left[1+\nu\sum_{k=1}^{J} q_k(t-\tau)\right]^{r+\sum_{k=1}^{v} i_k}}\prod_{k=1}^{v} q_k(t-\tau)^{i_k}\mathrm{d}\tau\right].
\end{aligned}$$

Corollary 5.1.4 *Suppose $S \sim Log(\rho)$, with PMF*

$$P(S=n) = -\frac{1}{\log(1-\rho)}\frac{\rho^k}{k},\ k \in \mathbb{Z}^+.$$

Then for $v \in \{2,\ldots,J\}$ and $n_v \geq 1$

$$\begin{aligned}
&P\left[\mathbf{N}(t) = (n_1,\ldots,n_v,0,\ldots 0)\right] \\
&\quad = \sum_{i_1=0}^{n_1}\cdots\sum_{i_{v-1}=0}^{n_{v-1}}\sum_{i_v=1}^{n_v}\left[\frac{i_v}{n_v}\cdot P\left[\mathbf{N}(t) = (n_1-i_1,\ldots,n_v-i_v,0,\ldots 0)\right]\right. \\
&\quad\quad \left.\cdot\int_0^t -\lambda(\tau)\frac{\rho^{\sum_{k=1}^{v} i_k}}{\log(1-\rho)}\left(\prod_{k=1}^{v}\frac{q_k(t-\tau)^{i_k}}{i_k!}\right)\frac{\left(-1+\sum_{v=1}^{v} i_k\right)!}{\left[1-\rho\left(1-\sum_{k=1}^{J} q_k(t-\tau)\right)\right]^{\sum_{k=1}^{v} i_k}}d\tau\right].
\end{aligned}$$

5.1.2 Constant batch sizes

Suppose $\mathbf{S} = \mathbf{s}$ is constant. Then $\mathbf{C}(t-\tau)$ is the sum of independent, but non-identical categorical random variables, that is, $\mathbf{C}(t-\tau)$ follows a Poisson multinomial distribution [12]. Using the exact DFT-CT method proposed by [12], we can recover the PMF for $\mathbf{C}(t-\tau)$, which can then be plugged into the recurrence relation (11) to compute the transient occupancy distribution.

5.1.3 A model of superinfection and hypnozoite dynamics for vivax malaria

In [16], we constructed an open network of infinite server queues with exponential service times to model within-host superinfection and hypnozoite dynamics for *Plasmodium vivax* malaria. A schematic of the queueing network, with nodes labelled $\{1,\ldots,k, NL, A, D, C, P, PC\}$, is replicated from Figure 3 of [16] below (Fig. 1).

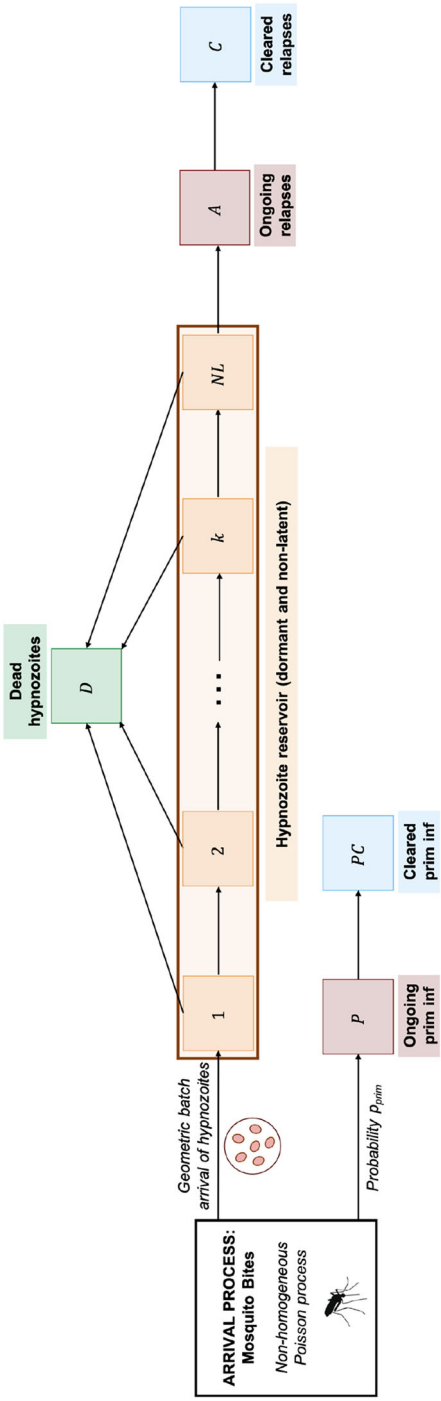


Fig. 1 Model schematic replicated from Figure 3 of [16]

The batch takes the form of a single arrival into queue P with probability p_{prim} (representing a primary blood-stage infection) and a geometrically-distributed arrival into queue 1 (representing the establishment of dormant hypnozoites). Each hypnozoite then undergoes a compartmental process, accounting for death (queue D), or progression through successive latency compartments (queues $2, \ldots, k$) and subsequent activation giving rise to a blood-stage relapse (queue A). We further account for the clearance of each primary infection (queue P to PC) and relapse (queue A to C). Analysis of the queueing network allows for the characterisation of quantities of epidemiological interest, with applications to population-level transmission modelling [17].

6 Conclusion

We have extended results for single infinite-server queues with batch arrivals and open networks of infinite-server queues with single arrivals to a general network of infinite server queues with multivariate batch arrivals arriving according to a non-homogeneous Poisson process. Theorem 3.1 gives an expression for the PGF of the transient queue length distribution, Theorem 4.1 a necessary and sufficient condition for ergodicity and Theorem 5.1 a recurrence relation for the multivariate probability mass function.

Acknowledgements The authors would like to acknowledge funding from the Australian Research Council through Laureate Fellowship FL130100039.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Abramowitz, M., Stegun, I.A., Romer, R.H.: Handbook of mathematical functions with formulas, graphs, and mathematical tables. Am. J. Phys. **56**, 958 (1988)
2. Brown, M., Ross, S.M.: Some results for infinite server Poisson queues. J. Appl. Probab. **63**, 604–611 (1969)
3. Chatterjee, U., Mukherjee, S.: On the non-homogeneous service system MX/G/∞. Euro. J. Op. Res. **38**(2), 202–207 (1989)
4. Comtet, L.: Advanced combinatorics: the art of finite and infinite expansions. Springer, Berlin (2012)
5. Cong, T.D.: On the $M^X/G/\infty$ queue with heterogeneous customers in a batch. J. Appl. Probab. **31**, 280–286 (1994)
6. Daw, A., Pender, J.: On the distributions of infinite server queues with batch arrivals. Queueing Syst. **91**(3), 367–401 (2019)

7. Harrison, J.M., Lemoine, A.J.: A note on networks of infinite-server queues. J. Appl. Probab. **18**, 561–567 (1981)
8. Jeffrey, A., Zwillinger, D.: Table of integrals, series, and products. Elsevier, Amsterdam (2007)
9. Keilson, J., Servi, L.D.: Networks of non-homogeneous $M/G/\infty$ systems. Tech. rep Massachusetts Institute of Technology, Cambridge (1990)
10. Kella, O., Whitt, W.: Linear stochastic fluid networks. J. Appl. Probab. **36**(1), 244–260 (1999)
11. Korevaar, J., Wiegerinck, J.: Several complex variables. Korteweg-de Vries Institute for Mathematics, Amsterdam (2017)
12. Lin, Z., Wang, Y., Hong , Y.: The poisson multinomial distribution and its applications in voting theory, ecological inference, and machine learning. arXiv preprint arXiv:2201.04237 (2022)
13. Liu, L., Templeton, J.: The $GR_n^X/Gn/\infty$ system: system size. Queueing Syst. **8**(1), 323–356 (1991)
14. Massey, W.A., Whitt, W.: Networks of infinite-server queues with nonstationary Poisson input. Queueing Syst. **13**(1), 183–250 (1993)
15. McCalla, C., Whitt, W.: A time-dependent queueing-network model to describe the lifecycle dynamics of private-line telecommunication services. Telecommun. Syst. **19**(1), 9–38 (2002)
16. Mehra, S., Stadler, E., Khoury, D., McCaw, J.M., Flegg, J.A.: Hypnozoite dynamics for Plasmodium vivax malaria: the epidemiological effects of radical cure. J. Theor. Biol. **537**, 111014 (2022)
17. Mehra, S., Taylor, P.G., McCaw, J.M., Flegg, J. A.: A hybrid transmission model for Plasmodium vivax accounting for superinfection, immunity and the hypnozoite reservoir. arXiv preprint arXiv:2208.10403 (2022)
18. Miatto, F.M.: Recursive multivariate derivatives of $e^{f(X1,...,Xn)}$ of arbitrary order. arXiv preprint arXiv:1911.11722 [cs, math] (2019)
19. Panjer, H.H.: Recursive evaluation of a family of compound distributions. ASTIN Bull. J. IAA **12**(1), 22–26 (1981)
20. Reynolds, J.F.: Some results for the bulk-arrival infinite-server Poisson queue. Op. Res. **16**(1), 186–189 (1968)
21. Schumann, A.: Multivariate bell polynomials and derivatives of composed functions. arXiv preprint arXiv:1903.03899 (2019)
22. Shanbhag, D.: On infinite server queues with batch arrivals. J. Appl. Probab. **3**(1), 274–279 (1966)
23. Sundt, B., Jewell, W.S.: Further results on recursive evaluation of compound distributions. ASTIN Bull. J. IAA **12**(1), 27–39 (1981)
24. Tijms, H.C.: A first course in stochastic models. Wiley, London (2003)
25. Whittaker, E.T., Watson, G.N.: A course of modern analysis. Courier Dover Publications, Mineola (2020)
26. Willmot, G.: Sundt and Jewell's family of discrete distributions. ASTIN Bull. J. IAA **18**(1), 17–29 (1988)
27. Willmot, G.E., Drekic, S.: On the transient analysis of the $M^X/M/\infty$ queue. Op. Res. Lett. **28**(3), 137–142 (2001)
28. Willmot, G.E., Drekic, S.: Transient analysis of some infinite server queues. Recent Adv. Stat. Methods (2002). https://doi.org/10.1142/9781860949531_0029
29. Willmot, G.E., Drekic, S.: Time-dependent analysis of some infinite server queues with bulk Poisson arrivals. INFOR Inform. Syst. Op. Res. **47**(4), 297–303 (2009)
30. Yajima, M., Phung-Duc, T., Masuyama, H.: The stability condition of BMAP/M/∞ queues. In: Proceedings of the 11th international conference on queueing theory and network applications, pp. 1–6 (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.