

Asset Analytics

Performance and Safety Management

Series Editors: Ajit Kumar Verma · P. K. Kapur · Uday Kumar

Kusum Deep

Madhu Jain

Said Salhi *Editors*

Performance Prediction and Analytics of Fuzzy, Reliability and Queuing Models

Theory and Applications



Springer

Asset Analytics

Performance and Safety Management

Series editors

Ajit Kumar Verma, Western Norway University of Applied Sciences, Hagesund, Rogaland Fylke, Norway

P. K. Kapur, Center for Interdisciplinary Research, Amity University, Noida, India

Uday Kumar, Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden

The main aim of this book series is to provide a floor for researchers, industries, asset managers, government policy makers and infrastructure operators to cooperate and collaborate among themselves to improve the performance and safety of the assets with maximum return on assets and improved utilization for the benefit of society and the environment.

Assets can be defined as any resource that will create value to the business. Assets include physical (railway, road, buildings, industrial etc.), human, and intangible assets (software, data etc.). The scope of the book series will be but not limited to:

- Optimization, modelling and analysis of assets
- Application of RAMS to the system of systems
- Interdisciplinary and multidisciplinary research to deal with sustainability issues
- Application of advanced analytics for improvement of systems
- Application of computational intelligence, IT and software systems for decisions
- Interdisciplinary approach to performance management
- Integrated approach to system efficiency and effectiveness
- Life cycle management of the assets
- Integrated risk, hazard, vulnerability analysis and assurance management
- Adaptability of the systems to the usage and environment
- Integration of data-information-knowledge for decision support
- Production rate enhancement with best practices
- Optimization of renewable and non-renewable energy resources

More information about this series at <http://www.springer.com/series/15776>

Kusum Deep · Madhu Jain · Said Salhi
Editors

Performance Prediction and Analytics of Fuzzy, Reliability and Queuing Models

Theory and Applications

Editors

Kusum Deep
Department of Mathematics
Indian Institute of Technology Roorkee
Roorkee, Uttarakhand, India

Madhu Jain
Department of Mathematics
Indian Institute of Technology Roorkee
Roorkee, Uttarakhand, India

Said Salhi
Kent Business School, Centre for Logistics
and Heuristic Optimization (CLHO)
University of Kent
Canterbury, Kent, UK

ISSN 2522-5162

Asset Analytics

ISBN 978-981-13-0856-7

<https://doi.org/10.1007/978-981-13-0857-4>

ISSN 2522-5170 (electronic)

ISBN 978-981-13-0857-4 (eBook)

Library of Congress Control Number: 2018950199

© Springer Nature Singapore Pte Ltd. 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

Busy Period Analysis of $GI/G/c$ and $MAP/G/c$ Queues	1
Srinivas R. Chakravarthy	
Solving LP Models for Multi-objective Matrix Games with I-Fuzzy Goals	33
Sandeep Kumar	
Fuzzy Integrated Super-Efficiency Slack Based Measure Model	43
Alka Arya and Shiv Prasad Yadav	
Prioritizing Factors Affecting the Adoption of Mobile Financial Services in Emerging Markets—A Fuzzy AHP Approach	55
Kriti Priya Gupta and Rishi Manrai	
Design of Reliability Single Sampling Plan by Attributes Based on Exponentiated Exponential Distribution	83
A. Loganathan and M. Gunasekaran	
Availability Prediction of Repairable Fault-Tolerant System with Imperfect Coverage, Reboot, and Common Cause Failure	93
Madhu Jain and Pankaj Kumar	
Software Reliability Growth Model in Distributed Environment Subject to Debugging Time Lag	105
Ritu Gupta, Madhu Jain and Anuradha Jain	
Imperfect Software Reliability Growth Model Using Delay in Fault Correction	119
Bhoopendra Pachauri, Ajay Kumar and Sachin Raja	
F-Policy for M/M/1/K Retrial Queueing Model with State-Dependent Rates	127
Madhu Jain and Sudeep Singh Sanga	

Time-Shared Queue with Nopassing Restriction for the Loss-Delay Customers and Additional Server	139
Madhu Jain, Shalini Shukla and Rakesh Kumar Meena	
The Effect of Vacation Interruptions Policy on the Queueing System with Cost Optimization	153
Anupama, Anjana Solanki and Chandan Kumar	
Balking Strategies for a Working Vacation Priority Queueing System with Two Classes of Customers	165
Anamika Jain and Madhu Jain	
$M^X/G/1$ Queue with Optional Service and Server Breakdowns	177
Charan Jeet Singh and Sandeep Kaur	
Performance Analysis of Series Queue with Customer's Blocking	191
Sreekanth Kolledath and Kamlesh Kumar	
Markovian Multi-server Queue with Reneging and Provision of Additional Removable Servers	203
Madhu Jain, Shivani Kumari, Rashika Qureshi and Roly Shankaran	
Analysis of Queues with Impatient Clients: An Application to Online Shopping	219
Yogesh Shukla, Nasir Khan and Sonia Shivhare	
Designing Bulk Arrival Queue Model to an Interdependent Communication System with Fuzzy Parameters	225
Reeta Bhardwaj, T. P. Singh and Vijay Kumar	
Transient Analysis of Markov Feedback Queue with Working Vacation and Discouragement	235
Madhu Jain, Shobha Rani and Mayank Singh	
Transient and Steady-State Behavior of a Two-Heterogeneous Servers' Queuing System with Balking and Retention of Reneging Customers	251
Rakesh Kumar, Sapana Sharma and Gulab Singh Bura	
Mehar Methods to Solve Intuitionistic Fuzzy Linear Programming Problems with Trapezoidal Intuitionistic Fuzzy Numbers	265
Sukhpreet Kaur Sidhu and Amit Kumar	

About the Editors

Dr. Kusum Deep is a professor in the Department of Mathematics, Indian Institute of Technology Roorkee. Her research interests include numerical optimization, nature-inspired optimization, computational intelligence, genetic algorithms, parallel genetic algorithms, and parallel particle swarm optimization.

Dr. Madhu Jain is an associate professor in the Department of Mathematics, Indian Institute of Technology Roorkee. Her research interests include computer communications networks, performance prediction of wireless systems, mathematical modeling, and biomathematics.

Dr. Said Salhi is Director of the Centre for Logistics and Heuristic Optimization (CLHO) at Kent Business School, University of Kent, UK. Prior to his appointment at Kent in 2005, he served at the University of Birmingham's School of Mathematics for 15 years, where in the latter years he acted as Head of the Management Mathematics Group. He obtained his B.Sc. in Mathematics from the University of Algiers and his M.Sc. and Ph.D. in Operational Research at Southampton (Institute of Mathematics) and Lancaster (School of Management), respectively. He has edited six special journal issues and chaired the European Working Group in Location Analysis in 1996 and recently the International Symposium on Combinatorial Optimisation (CO2016) in Kent from September 1 to 3, 2016. He has published over 100 papers in academic journals.

Busy Period Analysis of $GI/G/c$ and $MAP/G/c$ Queues



Srinivas R. Chakravarthy

Abstract The busy period analysis of queueing systems, in general, is very involved and complicated. Even for the simplest queueing model, namely $M/M/1$, the probability density function of the busy period is obtained in terms of modified Bessel function. A number of approaches using complex analysis, combinatorics, lattice path, and matrix-analytic methods have been applied to study some selected queueing models. While the steady-state analysis involving queue length and waiting times of queueing models, in general, has been receiving considerable and significant attention in the literature from both analytical and algorithmic points of view, the same cannot be said (relatively speaking) about busy period analysis. This is inherent in the nature of the busy period more than by choice. In this paper, after establishing the complexity involved in the study of the busy period, we record some interesting observations on the busy period under a wide variety of scenarios through simulation approach. The main purpose is to help researchers to look for novel theoretical and/or numerical approach to solving functional equations which naturally arise in the study of busy periods and use the simulated results here as one of the ways to confirm/validate their results.

Keywords Queueing · Busy period · Matrix-analytic method · Algorithmic probability · Simulation

1 Introduction and Notation

In this paper, we define the busy period (BP) to be the duration of the time interval that begins with an arrival of a customer to an empty system and ends with the system becoming empty again at the departure of a customer. This will be the case even for a multi-server queueing system. In the literature (see, e.g., [1, 2]), several authors recourse to full and partial busy periods when dealing with multiple-server system.

S. R. Chakravarthy (✉)

Departments of Industrial and Manufacturing Engineering & Mathematics,
Kettering University, Flint, MI 48504, USA
e-mail: schakrav@kettering.edu

© Springer Nature Singapore Pte Ltd. 2019

K. Deep et al. (eds.), *Performance Prediction and Analytics of Fuzzy, Reliability and Queueing Models*, Asset Analytics, https://doi.org/10.1007/978-981-13-0857-4_1

Our definition here for multiple-server system is referred to as partial busy period. A full busy period is the one that starts with all servers becoming busy until at least one server becomes free. Note that in a single-server queueing system, the partial and full busy periods are the same.

The busy period analysis in queueing systems, in general, is very involved and complicated (see, e.g., [3–6]). Even for the simplest queueing model, namely $M/M/1$, the probability density function of the busy period is obtained in terms of modified Bessel function. A number of approaches using complex analysis, combinatorics, lattice path, and matrix-analytic methods have been applied to study some selected queueing models. While the steady-state analysis involving queue length and waiting times of queueing models, in general, has been receiving considerable and significant attention in the literature from both analytical and algorithmic points of view, the same cannot be said (relatively speaking) about busy period analysis. This is inherent in the nature of the busy period more than by choice. In fact, the busy period analysis got a new focus since the introduction of matrix-analytic methods by Neuts [7, 8] in the context of $M/G/1$ and $GI/M/1$ paradigms. In this paper, after establishing the complexity involved in the study of the busy period, we record some interesting observations on the busy period under a wide variety of scenarios through simulation approach.

The purpose of this paper is twofold. First one is to show the complexity involved in the study of the busy period. Secondly, we want to record some interesting observations on the busy period of queueing systems in general context through simulation approach. This will help researchers to look for novel theoretical and/or numerical approach to solving functional equations which naturally arise in the study of busy periods.

In the following, we will denote by $f(\cdot)$ and $F(\cdot)$, respectively, the probability density and probability distribution function of the inter-arrival times. Similarly, we define by $h(\cdot)$ and $H(\cdot)$ to be, respectively, the probability density and probability distribution function of the service times. We will also assume that means of $F(\cdot)$ and $H(\cdot)$ exist and are given by

$$\frac{1}{\lambda} = \int_0^\infty [1 - F(t)]dt \quad \text{and} \quad \frac{1}{\mu} = \int_0^\infty [1 - H(t)]dt, \quad (1)$$

so that λ denotes the rate of arrivals to the system and μ gives the rate of services.

Let Y denote the busy period of the queueing system under study, and let $\Phi(\cdot)$ and $\phi(\cdot)$ denote, respectively, the probability distribution and the density function of Y . We will denote by N_Y the number of customers served during the busy period, Y .

The Laplace–Stieltjes transforms (LST) of $F(\cdot)$, $H(\cdot)$, and $\Phi(\cdot)$ are defined as

$$\begin{aligned} f^*(s) &= \int_0^\infty e^{-st} dF(t), \\ h^*(s) &= \int_0^\infty e^{-st} dH(t), \end{aligned} \quad (2)$$

and

$$\phi^*(s) = \int_0^\infty e^{-st} d\Phi(t), \quad \text{Re}(s) \geq 0.$$

The probability generating function, $N(z) = E[z^{N_Y}]$, is

$$N(z) = \sum_{n=1}^{\infty} z^n P(N_Y = n), \quad |z| < 1. \quad (3)$$

The rest of the paper is organized as follows. In Sect. 2, we present known key results for the busy period for the classical $M/G/1$ - and $M/G/1$ -type queues. The corresponding known results for the classical $GI/M/1$ - and $GI/M/1$ -type queues are presented in Sect. 3. In Sect. 4, we look at $GI/G/1$ queues, and in Sect. 5, we look at multi-server queueing systems. A brief summary of some known papers dealing with algorithmic analysis of busy periods is presented in Sect. 6. Validation of our simulated results against some queueing models for which numerical results are reported is done in Sect. 7. Some interesting and illustrative simulated examples are discussed in Sect. 8. Finally, in Sect. 9, some concluding remarks are presented.

2 $M/G/1$ -Type Queue

In this section, we will look at the classical (scalar) $M/G/1$ - and the $M/G/1$ -type queues.

2.1 Classical (Scalar) $M/G/1$ Queue

Here we look at the case of Poisson arrivals, and the service times following a general probability distribution function. That is,

$$F(t) = 1 - e^{-\lambda t}, \quad t \geq 0, \quad \text{and} \quad f^*(s) = \frac{\lambda}{\lambda + s}, \quad \text{Re}(s) \geq 0. \quad (4)$$

In this case, Takacs [9] introduced a novel idea in obtaining the *LST* of the busy period (as it does not depend on the type of service discipline unlike the waiting time distribution) and showed (see also [6]) that $\phi^*(s)$ and $N(z)$ satisfy the following equations.

$$\phi^*(s) = h^*[s + \lambda - \lambda \phi^*(s)], \quad \text{Re}(s) \geq 0, \quad \text{and} \quad N(z) = h^*[\lambda(1 - N(z))], \quad |z| < 1. \quad (5)$$

Moments of busy period: Suppose that $\mu_Y^{(k)}$ denotes the k th moment of the busy period and σ_h^2 denotes the variance of the service time. Let $\rho = \frac{\lambda}{\mu}$. The following can be easily verified.

$$\begin{aligned}\mu_Y^{(1)} &= -\left.\frac{d\phi^*(s)}{ds}\right]_{s=0} = \frac{1}{\mu(1-\rho)}, \\ \mu_Y^{(2)} &= -\left.\frac{d^2\phi^*(s)}{ds^2}\right]_{s=0} = \frac{\sigma_h^2 + \frac{1}{\mu^2}}{(1-\rho)^3}, \quad \sigma_Y^2 = \frac{\sigma_h^2 + \rho\frac{1}{\mu^2}}{(1-\rho)^3}.\end{aligned}\quad (6)$$

The mean and the variance of the number of customers served during a busy period can be verified to be

$$\mu(N_Y) = \frac{1}{1-\rho} \quad \text{and} \quad \sigma_{N_Y}^2 = \frac{\rho(1-\rho) + \lambda^2(\sigma_h^2 + \frac{1}{\mu^2})}{(1-\rho)^3}. \quad (7)$$

2.2 M/M/1-Queue

Here we look at the case of Poisson arrivals, and the service times are exponentially distributed. That is,

$$\begin{aligned}F(t) &= 1 - e^{-\lambda t}, \quad t \geq 0, \quad H(t) = 1 - e^{-\mu t}, \quad t \geq 0, \\ f^*(s) &= \frac{\lambda}{\lambda + s}, \quad h^*(s) = \frac{\mu}{\mu + s}, \quad \operatorname{Re}(s) \geq 0.\end{aligned}\quad (8)$$

It is easy to verify that

- (i) The *LST* of the busy period is given by [9]

$$\begin{aligned}\phi^*(s) &= \frac{\mu}{\mu + s + \lambda(1 - \phi^*(s))} \rightarrow \phi^*(s) \\ &= \frac{1}{2\lambda} \left[(\lambda + \mu + s) - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu} \right].\end{aligned}\quad (9)$$

(ii) $N(z) = z \frac{\mu}{\mu + \lambda(1 - N(z))}.$

- (iii) The density of the number served can be obtained explicitly as (see, e.g., [6, 9])

$$P(N_Y = n) = \frac{1}{n} \binom{2n-2}{n-1} \rho^{n-1} (1+\rho)^{1-2n}, \quad n \geq 1. \quad (10)$$

As part of transient analysis of this queueing system, Leguesdron et al. [10] obtained explicit expressions for the probability distribution function of Y and N_Y . Defining

$$\theta = \lambda + \mu, \quad a = \frac{\mu}{\theta}, \quad b = 1 - a, \quad (11)$$

the probability distribution function of Y is obtained using Bessel function as

$$P(Y \leq t) = \sum_{n=1}^{\infty} e^{-\theta t} \frac{(\theta t)^k}{k!} \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{2k}{k} \frac{a^{k+1} b^k}{k+1}, \quad t \geq 0, \quad (12)$$

and

$$P(N_Y = n) = \binom{2n-2}{n-1} \frac{a^n b^{n-1}}{n}, \quad n \geq 1. \quad (13)$$

Note that (10) and (13) are identical, as it should be. Also, we refer to [3, 11–14] for different approaches to getting the *LST* of the busy period.

2.3 $M/G/1$ -Type Queues

Understanding the important role of the busy period in the classical (scalar) $M/G/1$ queue, Neuts generalized it to the $M/G/1$ -type queues using matrix formalism. We will briefly summarize the key results pertinent to our discussion here and refer the reader to [7, 8, 15–17] for full details. Consider a Markov renewal process (*MRP*) with transition probability matrix given by

$$Q(x) = \begin{bmatrix} B_0(x) & B_1(x) & B_2(x) & B_3(x) & \cdots \\ A_0(x) & A_1(x) & A_2(x) & A_3(x) & \cdots \\ & A_0(x) & A_1(x) & A_2(x) & \cdots \\ & & A_0(x) & A_1(x) & \cdots \\ & & & A_0(x) & \cdots \\ & & & & \ddots \end{bmatrix}, \quad (14)$$

where the entries are block matrices and govern transitions within each level. While the matrices $A_k(x)$, $k \geq 0$, $x \geq 0$, representing possibly defective probability distributions on $[0, \infty)$ are square, the others are rectangular, and in some applications, they may also be square. The matrix, $A(x) = \sum_{i=0}^{\infty} A_i(x)$, is a stochastic semi-Markov matrix, and $A(\infty)$ is a stochastic matrix. Further, $Q(\infty)$ is a stochastic matrix. Note that the levels may represent the number of customers in the system and the auxiliary variable within the level may represent the phase of the service.

Due to simple boundary conditions (see [7, 8] for more complex boundary cases), the fundamental period (to be defined below) will be the busy period as defined in

this paper. We will assume that the *MRP* is irreducible, which is the case in most applications.

Let $T(i + r, j; i, k)$ denote the first passage time from state $(i + r, j)$ to the state (i, k) , for $i, r \geq 1, 1 \leq j, k \leq m$. That is, $T(i + r, j; i, k)$ is the duration that the semi-Markov process $Q(\cdot)$ starting in state $(i + r, j)$ visits level i for the first time by entering the state (i, k) . Let $V(i + r, j; i, k)$ denote the number of transitions involved in the *MRP* during the first passage time $T(i + r, j; i, k)$.

The joint probability function of $T(i + r, j; i, k)$ and $V(i + r, j; i, k)$ plays a key role in $M/G/1$ -type queues. Suppose that we define the matrix $G^{(r)}(n, x)$ such that its (j, k) th entry gives the joint probability as follows (note that these matrices do not depend on i due to the structure of $Q(x)$):

$$g_{j,k}^{(r)}(n, x) = P\{T(i + r, j; i, k) \leq x, V(i + r, j; i, k) = n\}, \quad r \geq 0, 1 \leq j, k \leq m, n \geq 0, \quad (15)$$

where we take

$$g_{j,k}^{(0)}(n, x) = \begin{cases} 1, & j = k, n = 0, x = 0, \\ 0, & \text{elsewhere.} \end{cases} \quad (16)$$

It can be verified that $G^{(r)}(n, x) = G^r(n, x)$. Also note that $G(n, x) = G^{(1)}(n, x)$, $n \geq 1, x \geq 0$, is such that $g_{j,k}(n, x)$ is the conditional probability that the first passage time from $(i + 1, j)$ to (i, k) , for $i \geq 1, 1 \leq j, k \leq m$, occurs in exactly n transitions and no later than x .

Denoting $A_r^*(s)$ to be the *LST* of $A_r(\cdot)$, the joint *LST*, $G(z, s)$, defined as

$$G(z, s) = \sum_{n=1}^{\infty} z^n \int_0^{\infty} e^{-sx} dG(n, x), \quad |z| \leq 1, \operatorname{Re}(s) \geq 0, \quad (17)$$

satisfies [8]

$$G(z, s) = z \sum_{r=0}^{\infty} A_r^*(s) G^r(z, s). \quad (18)$$

The substochastic matrix, $G(x) = \sum_{n=0}^{\infty} G(n, x)$, gives the matrix distribution for the fundamental period. That is, $g_{j,k}(x)$ is the conditional probability that the first passage time from $(i + 1, j)$ to (i, k) , for $1 \leq j, k \leq m$, occurs no later than x .

The sequence, $\{\hat{G}(n) = G(n, \infty)\}$, $n \geq 1$, of matrices gives the matrix-mass functions for the number of transitions during a fundamental period. That is, $\hat{g}_{j,k}(n)$, $1 \leq j, k \leq m, n \geq 1$, is the conditional probability that exactly n transitions occur during the first passage time from $(i + 1, j)$ to (i, k) .

The matrix $G = G(\infty) = \sum_{n=1}^{\infty} \hat{G}(n)$ is such that its (j, k) th entry gives the conditional probability that the *MRP* eventually visits the state (i, k) for the first time starting in state $(i + 1, j)$. This matrix plays an important role in $M/G/1$ -type queues like R in $GI/M/1$ -type queues (see [7, 8, 15]).

3 $GI/M/1$ -Type Queues

In this section, we will look at the classical (scalar) $GI/M/1$ - and the $GI/M/1$ -type queues.

3.1 Classical (Scalar) $GI/M/1$ Queue

Here we look at the general independent and identically distributed inter-arrival times and exponential services. That is,

$$H(t) = 1 - e^{-\mu t}, \quad t \geq 0, \quad h^*(s) = \frac{\mu}{\mu + s}, \quad \text{Re}(s) \geq 0. \quad (19)$$

Let $\{t_k : k \geq 0\}$ with $t_0 = 0$ denote the time points at which k th arrival occurs. Thus, for $k \geq 0$, $(t_{k+1} - t_k)$, denotes the duration of the time between k th and $(k + 1)$ st arrivals. Note that the nonnegative random variable $\tau_{k+1} = t_{k+1} - t_k$ has distribution function $F(\cdot)$ that is independent of k . Let $N(t)$ denote the number of customers in the system at time t , and let $N_k = N(t_k - 0)$, $k \geq 0$, denote the number of customers in the system just before the k th arrival.

The process $\{(N_n, \tau_n) : n \geq 0\}$ is a Markov renewal process (MRP) with TPM given by

$$Q(x) = \begin{pmatrix} b_0(x) & a_0(x) & & \cdots \\ b_1(x) & a_1(x) & a_0(x) & \cdots \\ b_2(x) & a_2(x) & a_1(x) & a_0(x) & \cdots \\ b_3(x) & a_3(x) & a_2(x) & a_1(x) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (20)$$

where $a_k(x) = \int_0^x e^{-\mu t} \frac{(\mu t)^k}{k!} dF(t)$, $k \geq 0$, gives the probability that k departures (or service completions) occur during an arrival that occurs at or before time x , and $b_k(x) = 1 - \sum_{i=0}^k a_i(x)$, $k \geq 0$, $t \geq 0$.

Thus, we need to look at the imbedded Markov chain to study $GI/M/1$ queue.

For use in sequel, we define the taboo probability, ${}_i P_{i,i+1}^{(k)}(t)$, as the conditional probability that the MRP with TPM given in (20) starting in state i at time 0 makes k transitions by avoiding state i and visits state $i + 1$ at the k th transition which occurs no later than time t . Note that due to the structure of the TPM , this conditional probability does not depend on i .

Define $r(t)$, $t \geq 0$, as

$$r(t) = \sum_{k=1}^{\infty} {}_i P_{i,i+1}^{(k)}(t). \quad (21)$$

Note that $r(t)$ gives the expected number of visits to state $i + 1$ during $\min\{t, \tau\}$ units of time before first return to state i given that the *MRP* started in state i . Here τ is the epoch at which the first return to state i occurs starting from state i .

Then it is easy to see (see, e.g., [9, 18]) that $\phi^*(s)$ satisfies

$$\phi^*(s) = \frac{\mu(1 - r^*(s))}{s + \mu(1 - r^*(s))}, \quad \operatorname{Re}(s) \geq 0, \quad (22)$$

where $r^*(s)$ is obtained as the solution to

$$r^*(s) = \sum_{n=0}^{\infty} (r^*(s))^n \int_0^{\infty} e^{-(s+\mu)t} \frac{(\mu t)^n}{n!} dF(t) = f^*[s + \mu(1 - r^*(s))], \quad \operatorname{Re}(s) \geq 0. \quad (23)$$

3.2 *M/M/1-Queue*

We will now revisit *M/M/1* queue via *GI/M/1* approach. Here we take $F(t) = 1 - e^{-\lambda t}$, $t \geq 0$. Then, we have

- (i) $f^*(s) = \frac{\lambda}{s + \lambda}$
- (ii) It is easy to verify that

$$\begin{aligned} r^*(s) &= \frac{\lambda}{\lambda + s + \mu(1 - r^*(s))} \rightarrow r^*(s) \\ &= \frac{1}{2\mu} \left[(\lambda + \mu + s) - \sqrt{[(\lambda + \mu + s)^2 - 4\lambda\mu]} \right], \end{aligned} \quad (24)$$

from which it can be seen that

$$\begin{aligned} \phi^*(s) &= \frac{\mu}{\lambda} r^*(s) \rightarrow \phi^*(s) \\ &= \frac{1}{2\lambda} \left[(\lambda + \mu + s) - \sqrt{[(\lambda + \mu + s)^2 - 4\lambda\mu]} \right], \end{aligned} \quad (25)$$

which (as it should be) is same as the one we got earlier [see (9)].

3.3 *GI/M/1-Type Queues*

Here we will look at the *GI/M/1*-type queues that were introduced and studied extensively by Neuts [7]. Consider a Markov renewal process (*MRP*) with transition probability matrix given by

$$Q(x) = \begin{bmatrix} B_0(x) & A_0(x) & & \cdots \\ B_1(x) & A_1(x) & A_0(x) & \cdots \\ B_2(x) & A_2(x) & A_1(x) & A_0(x) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (26)$$

where the matrices $A_n(x)$ and $B_n(x)$, for $n \geq 0$ and for $x \geq 0$, are square matrices of order m representing possibly defective probability distributions on $[0, \infty)$. Further, $Q(\infty)$ is a stochastic matrix.

The *MRP* of the type given in (26) occurs naturally in many stochastic models including the well-known ones such as $GI/PH/1$ and $SM/M/1$. Also, the one given in (26) has a simple boundary condition and more complex boundary conditions also occur naturally, and we refer the reader to [7] for more details.

Before we proceed further, we will summarize some key concepts needed for discussion in the sequel.

Recall that if P is the *TPM* of a *DTMC*, then the (i, j) th entry of the matrix $\sum_{n=0}^{\infty} P^n$ gives the expected number of visits to state j starting in state i .

Suppose that H is a subset of the state space, Δ , of a *DTMC*. The taboo probability denoted by ${}_H p_{i,j}^{(n)}$ is the probability that starting in state i , the *DTMC* visits state j at time n without visiting any of the states in H . The taboo probabilities play an important in stochastic modeling, and Neuts used this concept extensively in the development of matrix-analytic methods.

Referring to the *TPM* given in (26), we define the level (or set) of states as follows.

$$\mathbf{i} = \{(i, j) : 1 \leq j \leq m\}. \quad (27)$$

Suppose we define the matrix $R^{(k)}(t) = \{r_{j,l}^{(k)}(t)\}$, $1 \leq j, l \leq m, k \geq 0$, such that (by convention we take $R^{(0)}(t) = I$, an identity matrix of dimension m)

$$r_{j,l}^{(k)}(t) = \sum_{n=0}^{\infty} {}_i p_{(i,j),(i+k,l)(t)}^{(n)}. \quad (28)$$

Note that ${}_i p_{(i,j),(i+k,l)}^{(n)}(t)$ gives the conditional probability that the *MRP*, $Q(\cdot)$, starting at time 0 in state (i, j) makes n transitions during $(0, t]$ by avoiding the level $\mathbf{i}$ and that the n th transition results in *MRP* being in state $(i+k, l)$. Observe from the structure of $Q(\cdot)$, the taboo probabilities do not depend on i but rather depend only on the submatrix obtained from Q by deleting all rows and columns with indices (l', j') , $l' \leq i, 1 \leq j' \leq m$. Also, note that these submatrices are identical for all $i \geq 0$.

Also note that $r_{j,l}^{(k)}(t)$ gives the expected number of visits during the time interval $(0, \min(t, \tau)]$ (here τ is the epoch of the first return to level $\mathbf{i}$) to the state $(i+k, l)$ starting in state (i, j) before the first return to level $\mathbf{i}$.

We will denote $R^{(1)}(t)$ by $R(t)$. Let $R^*(s)$ denote the *LST* of $R(t)$. Suppose that $A_n^*(s)$ denote the *LST* of $A_n(x)$. Then, it is known (see [18]) that $R(s)$ is the minimal nonnegative solution (in the class of all such matrices whose spectral radius is at most one) to the matrix nonlinear equation:

$$R^*(s) = \sum_{n=0}^{\infty} (R^*(s))^n A_n^*(s), \quad Re(s) \geq 0. \quad (29)$$

This matrix, $R^*(0) = R(\infty)$, is the well-known rate matrix (see [7]).

3.4 *GI/PH/1 Queue*

As seen in earlier sections, the busy period analysis tends to be very complicated. The first paper, to our knowledge, that dealt with non-exponential services is that of Conolly [19] who studied *GI/E_k/1* queue. The methodology (based on difference equations) used in [19] is that of the one applied by the same author [20] in the context of *GI/M/1* queue.

After more than three decades since the study of *GI/E_k/1* queue, Ramaswami [18] used probabilistic arguments (developed and popularized by Neuts through his matrix-analytic methods) to develop transform-free methods for analyzing busy period for a large class of queues that possess matrix-geometric steady-state probability vector. We will briefly outline the key steps here and refer the reader to Ramaswami [18] for full details.

Let the service times be of phase type (*PH*) with representation (β, S) of order n (see [7]). Recall that the service rate is given by $\mu = [\beta(-S)^{-1}\mathbf{e}]^{-1}$. In the sequel, we will assume that the queue is stable implying that $\lambda < \mu$. The probability density function, the distribution function, and the *LST* of service times are given by

$$h(t) = \beta e^{St} \mathbf{S}^0, \quad H(t) = 1 - \beta e^{St} \mathbf{e}, \quad t \geq 0, \quad h^*(s) = \beta(sI - S)^{-1} \mathbf{S}^0, \quad Re(s) \geq 0. \quad (30)$$

Suppose that $\{N(t)\}$ denotes the counting process associated with a *PH*-renewal process. That is, $\{N(t)\}$ denotes the number of renewals if the times between renewals follow a *PH*-distribution with representation (β, S) . Define the matrix $P(n, t)$, whose (i, j) th element is given by $p_{ij}(n, t) = P[N(t) = n, J(t) = j | N(0) = 0, J(0) = i]$. That matrix $P(n, t)$ satisfies (see, e.g., [7])

$$P'(0, t) = P(0, t)S, \quad P'(n, t) = P(n, t)S + P(n-1, t)\mathbf{S}^0\beta, \quad n \geq 1, \quad t \geq 0, \quad (31)$$

with $P(n, 0) = \delta_{n0}I$, where δ_{n0} is the Kronecker delta.

The matrices, $A_k(t)$ and $B_k(t)$, in the $GI/PH/1$ case are given by

$$A_k(t) = \int_0^t P(k, u) dF(u), \quad k \geq 0, \quad B_k(t) = \sum_{r=k+1}^{\infty} \int_0^t P(k, u) \mathbf{e} \boldsymbol{\beta} dF(u), \quad k \geq 0. \quad (32)$$

Suppose that the (i, j) th element of the matrix $C_k(t)$ of dimension m denotes the conditional probability that the busy period starts at time 0 with an arrival which sees k customers already present in the system and the service phase at time 0 is i , the busy period ends by time t with exactly $k + 1$ services, and that the next busy period starts in phase j . It can be verified (see, e.g., [18]) that

$$C_k(t) = \int_0^t P(k, u) \mathbf{e} \boldsymbol{\beta} [1 - F(u)] du, \quad k \geq 0. \quad (33)$$

Let $C_n^*(s)$ denote the *LST* of $C_k(t)$. One can get $C_n^*(s)$ in terms of $A_k^*(s)$ as follows (see, e.g., [18])

$$\begin{aligned} C_0^*(s) &= [I - A_0^*(s)](sI - S)^{-1} \mathbf{S}^0 \boldsymbol{\beta}, \\ C_k^*(s) &= [C_{k-1}^*(s) - A_k^*(s)](sI - S)^{-1} \mathbf{S}^0 \boldsymbol{\beta}, \quad k \geq 1. \end{aligned} \quad (34)$$

Let $g_i(t)$, $1 \leq i \leq m$, denote the distribution function of the busy period that starts with an arrival of a customer whose service phase will be in state i . Defining $\mathbf{g}(t)$ to be the vector of dimension m whose i th component is given by $g_i(t)$ and $\mathbf{g}^*(s)$ to be the *LST* of $\mathbf{g}(t)$, we register the explicit expression for the (vector) *LST* of the busy period distribution as obtained by Ramaswami [18] as follows.

$$\mathbf{g}^*(s) = \sum_{n=0}^{\infty} [R^*(s)]^n C_n^*(s) \mathbf{e} = [I - R^*(s)][I - h^*(s)R^*(s)]^{-1} (sI - S)^{-1} \mathbf{S}^0, \quad \text{Re}(s) \geq 0. \quad (35)$$

We will look at the simplifications of Ramaswami's expression given in (35) for two special queueing models.

3.4.1 $GI/M/1$ Queue

In this case, the expression given in (35) reduces to (note that $\mathbf{g}^*(s)$ and $R^*(s)$ are scalars)

$$g^*(s) = \frac{\mu(1 - R^*(s))}{s + \mu(1 - R^*(s))}, \quad \text{Re}(s) \geq 0, \quad (36)$$

and $R^*(s)$ is obtained as the solution to

$$R^*(s) = \sum_{n=0}^{\infty} (R^*(s))^n \int_0^{\infty} e^{-(s+\mu)t} \frac{(\mu t)^n}{n!} dF(t) = f^*[s + \mu(1 - R^*(s))], \quad \text{Re}(s) \geq 0, \quad (37)$$

which agrees with (23).

3.4.2 $M/M/1$ Queue

In this case, verify that

- (i) $f^*(s) = \frac{\lambda}{s+\lambda}$.
- (ii) $R^*(s)$ is obtained as the solution to

$$R^*(s) = \frac{\lambda}{\lambda + s + \mu(1 - R^*(s))} \rightarrow R^*(s) = \frac{1}{2\mu} \left[(\lambda + \mu + s) - \sqrt{[(\lambda + \mu + s)^2 - 4\lambda\mu]} \right], \quad (38)$$

from which it can be seen that

$$g^*(s) = \frac{\mu}{\lambda} R^*(s) \rightarrow g^*(s) = \frac{1}{2\lambda} \left[(\lambda + \mu + s) - \sqrt{[(\lambda + \mu + s)^2 - 4\lambda\mu]} \right], \quad (39)$$

which agrees with (9).

4 $GI/G/1$ -Queue

The earliest known study on the busy period of $GI/G/1$ queue is by Finch [21] in which expressions for the LST of the busy period and the density of the number served during a busy period are obtained using combinatorial approach. The expressions are not only complicated but also numerically unstable due to alternating signs appearing in those expressions. For the sake of completeness, we will reproduce the results (just to show the complexity involved in the busy period analysis), and for details, we refer the reader to [21].

Suppose that $\xi_n(t)$, $n \geq 1$, $t \geq 0$ denotes the joint probability of Y and N_Y such that

$$\xi_n(t) = P(Y \leq t, N_Y = n), \quad n \geq 1, \quad t \geq 0, \quad (40)$$

and $\xi_n^*(s)$ denote the LST of $\xi_n(t)$. Denoting by $F^{(n)}(\cdot)$ to be the n -fold convolution of $F(\cdot)$ with itself, and $H^{(n)}(\cdot)$ to be that of $H(\cdot)$, and

$$a_n^*(s) = \int_0^{\infty} e^{-st} [1 - F^{(n)}(t)] dH^{(n)}(t), \quad n \geq 1, \quad \text{Re}(s) \geq 0, \quad (41)$$

Finch showed that the LST of the joint probability function, $\xi_n(t)$, to be

$$\xi_n^*(s) = \sum \frac{(-1)^{k_1+k_2+\dots+k_n+1}}{k_1!k_2!\dots k_n!} (a_1^*(s))^{k_1} \left(\frac{a_2^*(s)}{2}\right)^{k_2} \dots \left(\frac{a_n^*(s)}{n}\right)^{k_n}, \quad (42)$$

where $n \geq 1$, $Re(s) \geq 0$ and the summation is over all nonnegative integers k_r such that $\sum_{r=1}^n rk_r = n$.

From (42), it can be verified that, for $n \geq 1$, $Re(s) \geq 0$,

$$P(N_Y = n) = \xi_n^*(0) = \sum \frac{(-1)^{k_1+k_2+\dots+k_n+1}}{k_1!k_2!\dots k_n!} (a_1)^{k_1} \left(\frac{a_2}{2}\right)^{k_2} \dots \left(\frac{a_n}{n}\right)^{k_n}, \quad (43)$$

where $a_n = a_n^*(0) = \int_0^\infty [1 - F^{(n)}(t)] dH^{(n)}(t)$, $n \geq 1$.

Bertsimas et al. [22] analyzed the busy period by formulating it as Hilbert factorization problem using two-dimensional Lindley process. Subsequently, Bertsimas and Nakazato [23] applied the method of stages by assuming the arrivals and services are of mixed generalized Erlang distributions and obtained the LST for the busy period.

It should be noted that Pakes [24], using duality results for $GI/G/1$ queue, derived expressions for the (a) probability of number served during a busy period of a $GI/G/1$ queue and (b) LST of $GI/M/1$ queue in which the first customer starting the busy period has a different service-time distribution compared to the other customers in that busy period.

In Baltrunas et al. [25, 26], the authors study the tail behavior of the busy period of a stable $GI/G/1$ queue with subexponential services.

5 Multi-server Queues

So far, we looked at single-server queueing systems under various scenarios. In this section, we will briefly summarize a few models dealing with multi-server case for which results are reported.

- Chae and Lim [27] derive the joint transform of the length of a busy period, the number of customers served during the busy period, and the remaining inter-arrival times at the instant the busy period ends for $GI/M/c$ queue with n -policy. By taking $n = 1$, their model reduces to the classical $GI/M/c$ queue. Some numerical results are reported.
- Natvig [28] derives the first- and second-order moments of the (partial) busy period as well as the distribution of the number of customers served by looking at a general birth-and-death queueing model with multiple servers.

- Omahen and Marathe [29] apply the technique of decomposition of busy periods to $M/M/c$ queueing system and derive recursive formulas for computing the LST of the (partial) busy period as well as the first two moments of the busy period.
- Artalejo and Lopez-Herrero [1] present an algorithmic analysis of the busy period in the context of $M/M/c$ queueing model. They obtain the LST as the solution of a finite system of linear equations. Further, they provide recurrent relations for computing the moments of the distribution of the length of the busy period as well as the number of customers served during a busy period. Some numerical examples are presented.
- Ghahramani and Wolff [30] provide a probabilistic proof for conditions that will guarantee (full) busy periods to have finite moments.

6 Algorithmic Analysis of Busy Periods

With a genuine concern for algorithmic feasibility of solutions of stochastic models to be useful in practice, Neuts [7] developed phase-type distribution, (batch) Markovian arrival processes, matrix-geometric methods, and later on matrix-analytic methods in stochastic modeling. In queueing theory, for what Neuts will be remembered (among many things) most in years to come, is the introduction and the development of the matrix-analytic methods (*MAMs*) for the solution of a wide variety of practical problems.

In the invited article published by European Journal of Operational Research [31], Neuts passionately says, “...The history of the matrix methods (so called for brevity) is short, but worth telling... I tackled a number of models involving embedded Markov renewal processes, evidently with some measure of success, since the papers were published in noted journals and some academic recognition came my way. It privately bothered me that, as the papers grew longer and the analysis more complex, the explicit or qualitative results in them became fewer and fewer.” He continues further, “...In the history of mathematics, a similarity of formalism has always indicated similarity of structure and an ultimate level of understanding is that of unifying structure.”

Stochastic modeling occurs naturally in many walks of life. The mathematical tools needed to solve a specific problem vary depending on the application of the stochastic model. Telecommunications area first adopted *MAMs* soon after their introduction. Its well-known self-similarity property is seen in ethernet traffic, WWW traffic, signaling traffic, multi-media traffic, and other high-speed network traffic, and hence, Poisson/exponential distributions are not well suited for modeling such traffic. The benefit of using Markov-modulated Poisson process (which is a special class of Markovian arrival processes (MAP)) has been well documented.

Since the introduction of *MAMs* by Neuts, the users of this methodology have grown from within queueing community to other areas due to solutions that are transparent, implementable, and probabilistic in nature. The practitioners benefiting from Neuts’ contributions come from various fields such as health care, computer

and communications engineering, production and manufacturing, industrial engineering, electrical engineering, actuarial science, transportation, and wireless sensor networks. Neuts' work has inspired many researchers from all over the world as can be seen from constant publications of research papers dealing with *MAMs* in many diverse fields of applications.

While the steady-state analysis of many queueing models under a variety of scenarios can be performed both analytically and algorithmically using *MAMs*, only very few papers exist that address the busy period analysis in the same spirit. Again, this is inherent in the complexity of the busy period analysis as opposed to the choice of the methodology. As seen in previous sections, the busy period analysis is very complex. Even for the simplest queue in continuous time, namely $M/M/1$, the expression for the probability distribution function of the busy period is quite complex. Due to this complexity, many authors have resorted to deriving expressions, some of which are not algorithmically suited under a wide variety of parameters of the queueing model under study. Since most expressions are given in terms of *LST*, Abate and Whitt [32, 33] pioneered the numerical methods for computing *LST* expressions in the context of queueing models.

Some notable early papers, in addition to the ones mentioned earlier, dealing with computational aspects of the busy period include the following.

- Conolly [20] applied difference equations technique to derive probability distributions associated with the busy period for $GI/M/1$ queue and performed a few numerical comparisons in the case of $M/M/1$ and $D/M/1$ queues.
- Abate et.al. [34] show that the probability density of the busy period can be numerically inverted without the need to use iterative procedure for solving Kendall's functional equation. They apply their technique to $M/G/1$ queue with gamma service-time distribution.
- Garikiparthi et al. [35] derive the joint *LST* for the busy period and the number of customers served during a busy period for a finite *QBD*-process and propose algorithms for computing the moments of the busy period and the number served. Also, illustrative examples are presented.
- Artalejo and Gomez-Corral [36] using catastrophe method derived the *LST* of the busy period for an $M/G/1$ retrial queue with finite orbit size and discuss illustrative examples with the help of numerical inversion of transforms.
- Using lattice path approach, the authors in [37–40] study the busy period. They also discuss illustrative numerical examples for special cases under a wide range of values for the parameters of the model.
- In the context of $M/E_k/1$ queue, Baek et al. [41] give a closed form expression for the queue length within a busy period and discuss some illustrative numerical examples. It should be pointed out that the authors do not give expressions for the density of the busy period but rather the queue length during that busy period.
- Novak et al. [42, 43] provide analysis of the distribution of the number of arrivals in a subinterval of a busy period of an $M/D/1$ queue as well as for $M/G/1$ queue using Takacs' ballot theorem and its generalization. The authors compare simulated results to the actual density function.

- Assuming the inter-arrival and service times of a customer are correlated, Langaris [44] gives a closed form expression for the joint transform of the busy period and the number served during that period and discuss a few numerical results.

7 Validation of the Simulated Model

In the literature, to the best of our knowledge, there are very few exact and approximate results for the tail probabilities as well as for complementary distribution function for the busy period are available and that too for limited queueing models. We will use these results to validate/compare our simulated results.

We used ARENA [45] to get our simulated results for the busy period for various queueing systems. We simulated the model using 5 replications and for 1,000,000 units (which in our case is minutes) for each replicate.

7.1 Abate and Whitt for $M/G/1$ Queueing Model

In this section, we will compare the results given in Abate and Whitt [33] wherein the authors provide exact (through transform inversion) tail probabilities for various time points. These time points depend on the type of queueing model as well as the value of ρ . They fix the mean service time to be 1 in all cases and take the arrival rate so as to obtain a given value for ρ . Two queueing models, $M/E_4/1$ and $M/\Gamma(2, 0.5)/1$, are considered in [33] under three scenarios by varying ρ . We denote by $\Gamma(\alpha, \beta)$, a gamma distribution with shape parameter given by α and the scale parameter is β . Recall that the density function, $f(t)$, of $\Gamma(\alpha, \beta)$ is given by

$$f(t) = \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{-t/\beta}, \quad \alpha > 0, \beta > 0, t \geq 0,$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$.

In Table 1, we display the (absolute) error percentages of $P(Y > t)$ by comparing the results given in [33] with our simulated ones. The absolute error percentage, here and elsewhere, is defined as

$$100 \left| \frac{\text{reported} - \text{simulated}}{\text{reported}} \right| \%.$$

By looking at the entries in Table 1, we notice that, in general, our simulated results agree very well with the ones reported in [33]. The ones that have somewhat large error percentages (e.g., the largest one is 7.28%) we noticed that the underlying probabilities are very small and generally the values differ in the fifth or sixth decimal

Table 1 Absolute error percentages for $M/G/1$ queue

$\rho = 0.50$	t	0.5	1	2	3	5	9	12	15	20	32
	$M/E_4/1$	0.51%	0.15%	0.1%	0%	0.15%	0.21%	1.19%	1.27%	0.55%	5.14%
	t	0.1	1	2	5	8	15	20	30	40	60
	$M/\Gamma(2, 0.5)/1$	6.46%	0.6%	1.77%	3.3%	3.28%	2.29%	1.42%	1.77%	3.65%	13.4%
$\rho = 0.75$	t	0.5	1	2	6	10	15	30	40	80	120
	$M/E_4/1$	0.36%	0.09%	0.03%	0.06%	0.22%	0.41%	1.22%	0.25%	1.12%	6.75%
	t	0.1	1	5	8	15	30	60	80	120	250
	$M/\Gamma(2, 0.5)/1$	6.39%	1.76%	1.58%	1.98%	1.57%	1.35%	0.55%	1.07%	1.54%	6.39%
$\rho = 0.90$	t	0.5	1	5	15	30	60	120	200	400	600
	$M/E_4/1$	0.36%	0.04%	0.04%	0.32%	0.05%	0.19%	1.52%	0.93%	2.72%	3.49%
	t	0.1	1	5	10	20	50	100	250	500	1000
	$M/\Gamma(2, 0.5)/1$	6.31%	1.58%	2.19%	2.23%	2.24%	1.55%	1.68%	2.13%	6.41%	7.28%

places. For example, corresponding to the scenario, $\rho = 0.9$, $M/\Gamma(2, 0.5)/1$, the simulated value for $P(Y > 32)$ is 0.00043577, whereas the reported value (see [33]) is 0.000470. So, the actual (absolute) difference is 0.00003423, which is a small number, but the percentage-wise it results in 7.28%. This is the case for other percentages too.

7.2 Adan and Resing for $M/M/1$ Queueing Model

In their book on queueing theory, Adan and Resing [46] report results on selected tail probabilities for $M/M/1$ queue by considering three values for the traffic intensities, $\rho = 0.8, 0.9, 0.95$, and fixing the mean service time to be 1. The tail probabilities are reported using two decimal places, and hence, we did the same so as to compare the results properly. In Table 2, we display the error percentages of our simulated results compared to the ones in [46]. Obviously, we notice that the results agree very much in all cases. In fact, only one error percentage is different from zero.

Table 2 Absolute error percentages for $M/M/1$

t	1	2	4	8	16	40	80
$\rho = 0.80$	0%	0%	0%	0%	0%	0%	0%
$\rho = 0.90$	0%	0%	0%	0%	0%	0%	0%
$\rho = 0.95$	0%	2.7%	0%	0%	0%	0%	0%

Table 3 Absolute error percentages for $M/M/5$

ρ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
μ_Y	0.02%	0.13%	0.10%	0.11%	0.14%	0.17%	0.02%	0.07%	0.31%
$\frac{\sigma_Y}{\mu_Y}$	0.17%	0.16%	0.06%	0.02%	0.07%	0.01%	0.23%	0.23%	0.08%
$E(N_Y)$	0.10%	0.08%	0.09%	0.11%	0.16%	0.16%	0.04%	0.07%	0.32%
$\frac{\sigma_{N_Y}}{E(N_Y)}$	0.21%	0.11%	0.05%	0.08%	0.05%	0.04%	0.27%	0.25%	0.06%

7.3 Artalejo and Lopez-Herrero for $M/M/5$ Queueing Model

Here we look at a multi-server system. Artalejo and Lopez-Herrero [1] consider $M/M/5$ queueing system by fixing the mean service time to be 1 and vary ρ from 0.1 through 0.9. Unlike the previous sections, the authors here report the various moments of the busy period and the number of customers served during a busy period. In Table 3, we display the (absolute) error percentages for the (a) mean busy period, μ_Y ; (b) coefficient of variation of the busy period, $\frac{\sigma_Y}{\mu_Y}$; (c) mean number of customers served during a busy period, $E(N_Y)$; and (d) coefficient of variation of the number served during a busy period, $\frac{\sigma_{N_Y}}{E(N_Y)}$.

It is clear from the entries in the above table that simulated results agree very well with the ones reported in [1].

7.4 Blanc for $M/G/1$ and $GI/M/1$ Queueing Models

In [47], the author replaces the original contour integral involving implicitly known functions with alternative contour integrals with known transforms and numerically inverts the transforms. The author provides numerical examples for $GI/M/1$ queueing model by considering Erlang and gamma arrivals and $M/G/1$ queueing model by considering Erlang and gamma distributions for services. For all scenarios, the arrival rate is fixed at 0.8 and the service rate to be 1, so that the traffic intensity is set at $\rho = 0.8$. In Table 4, we display the (absolute) error percentages of our simulated results with the ones provided in [47].

A quick look at the entries in Table 4 indicates that generally our simulated results seem to be closer to the ones reported in [47] in all scenarios except for a few scenarios involving $M/\Gamma(8, 0.125)/1$ and $\Gamma(10, 0.125)/M/1$ queues. For example, for the $\Gamma(8, 0.125)$ services, we notice the error percentages to be very large for a few tail probabilities. We increased the simulation run for each replicate from 1,000,000 units to 10,000,000 units to see whether a longer run is warranted and thus possibly explain the higher values for the error percentages. That increase in the simulation run pretty much yielded the same results. In other comparisons (see Tables 1 and 5) involving service distributions having gamma distribution, we never noticed that

Table 4 Absolute error percentages for $M/G/1$ and $GI/M/1$ queues

	t	0.1	0.5	1	2	5	10	25	50	100
$M/G/1$	$M/E_8/1$	0.00%	0.06%	0.16%	0.11%	0.05%	0.00%	0.26%	0.88%	5.26%
	$M/E_2/1$	0.79%	0.39%	0.17%	0.13%	0.10%	0.44%	0.72%	1.36%	0.00%
	$M/M/1$	0.00%	0.02%	0.02%	0.00%	0.11%	0.46%	0.23%	0.00%	0.00%
	$M/\Gamma(2, 0.5)/1$	6.38%	3.82%	1.61%	0.07%	1.61%	2.03%	3.37%	3.90%	2.74%
	$M/\Gamma(8, 0.125)/1$	43.95%	27.62%	18.51%	9.84%	1.57%	2.63%	6.76%	8.02%	8.33%
$G/M/1$	$E_8/M/1$	0.02%	0.21%	0.10%	0.04%	0.40%	0.76%	1.36%	0.00%	0.00%
	$E_2/M/1$	0.02%	0.14%	0.04%	0.00%	0.00%	0.12%	0.95%	0.88%	0.00%
	$M/M/1$	0.00%	0.02%	0.02%	0.00%	0.11%	0.46%	0.23%	0.00%	0.00%
	$\Gamma(2.5, 0.5)/M/1$	0.48%	1.95%	2.38%	2.28%	1.20%	0.68%	0.15%	0.00%	2.78%
	$\Gamma(10, 0.125)/M/1$	2.18%	6.56%	9.08%	11.22%	11.64%	10.66%	8.62%	7.70%	8.11%

Table 5 Absolute error percentages for $\Gamma(5, 0.4)/M/c$

c	μ_Y	σ_Y	$E(N_Y)$	σ_{N_Y}
1	0.06%	0.21%	0.00%	0.34%
2	0.08%	0.17%	0.04%	0.03%
3	0.11%	0.09%	0.00%	0.00%

high error percentages making us believe that the results in [47] may need to be checked for possible numerical inversion problems.

7.5 Chae and Lim for $GI/M/c$ Queueing Model

Looking at $GI/M/c$ queue, Chae and Lim [27] present numerical results by considering inter-arrivals to follow a gamma distribution with shape parameter to be 5 and scale parameter to be 0.4 so that the mean time between arrivals is 2 and the mean service time to be 1. They vary c from 1 to 3 and present the mean and standard deviation of the (a) busy period and (b) number served during a busy period. In Table 5, we display the (absolute) error percentages of our simulated results with the ones reported in [27].

Once again, the entries in Table 5 reveal that the simulated results agree very well with the reported ones.

8 Illustrative Examples Based on Simulation

In this section, we will discuss a few illustrative and interesting examples obtained through simulation. We look at several classes of queueing systems of the type $GI/G/c$ and $MAP/G/c$. In all our simulation examples, we used 5 replicates and each replicate of length 1,000,000 units (in our case, minutes). Unless otherwise specified, we will fix the service mean to be 1. That is, we fix $\mu = 1$, and vary λ so that a given traffic intensity, $\rho = \frac{\lambda}{c\mu}$, is achieved.

8.1 $GI/G/c$

Here we look at four scenarios for arrival processes, Erlang, Poisson, hyperexponential, and two-parameter Weibull. For ease of reference, we list the type of arrival processes (TAP) and the type of service times (TS) with appropriate labels.

TAP 1: Erlang (ERA): Here we consider an Erlang distribution of order 2 with rate 2λ .

TAP 2: Exponential (EXA): This corresponds to the classical Poisson process with rate λ .

TAP 3: Hyperexponential (HEA): We look at a mixture of two exponentials with rates 1.9λ and 0.19λ , respectively, with probabilities 0.9 and 0.1.

TAP 4: Weibull (WEA): We consider a two-parameter Weibull whose *CDF* is given by

$$F_{WB}(x) = \begin{cases} 1 - e^{-(0.5\lambda x)^{0.5}}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

TS 1: Erlang (ERS): This is Erlang of order 2 with rate 2 in each stage.

TS 2: Exponential (EXS): This is an exponential distribution with mean 1.

TS 3: Hyperexponential (HES): Here we look at mixture of two exponentials with rates 1.9 and 0.19, respectively, with mixing probabilities 0.9 and 0.1.

TS 4: Shifted exponential (SXP): The shifted exponential with a shift of magnitude 0.2 one with *CDF* given by

$$F_{SE}(x) = \begin{cases} 1 - e^{-1.25(x-0.2)}, & x \geq 0.2, \\ 0, & x < 0.2. \end{cases}$$

TS 5: Weibull (WES): We consider a two-parameter Weibull whose *CDF* is given by

$$F_{WB}(x) = \begin{cases} 1 - e^{-(2x)^{0.5}}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

First note that the coefficient of variation of the four arrival processes labeled *ERA*, *EXA*, *HEA* and *WBA* are, respectively, 0.7071, 1, 2.2447, and 2.2361. Similarly, the coefficient of variation of the 5 service times labeled *ERS*, *EXS*, *HES*, *SXP*, and *WES* are, respectively, 0.7071, 1, 2.2447, 0.8, and 2.2361.

Recall from queueing literature (see, e.g., [3, 5, 6]) that the mean waiting time in the system is known to increase as the variability in the arrival (or services) increases (assuming that all other parameters are fixed). So, we decided to explore whether such a behavior is seen for the mean busy period.

In Table 6, we display the mean busy period under various scenarios. A quick look at the table reveals the following interesting observations.

- While for $c = 1$ and $c = 2$, we notice that the mean busy period appears to increase with increasing variability in the arrival processes, for $c = 5$, we notice a different trend. For example, by looking at the mean busy period when there are 5 servers in the system, a larger variability in the services such as *HES* appears to yield a smaller value as compared to *ERS* which has a smaller variability. This appears to be the case for $\rho = 0.80$ and $\rho = 0.95$.
- We notice the mean busy period to be very large when $c = 5$. This is not surprising since we use partial busy period in that the busy period starts when an arriving