INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Fitting the $Ph_t/M_t/s/c$ Time-Dependent Departure Process for Use in Tandem Queueing Networks

Walid W. Nasr, Michael R. Taaffe

To cite this article:
Walid W. Nasr, Michael R. Taaffe (2013) Fitting the $Ph_t/M_t/s/c$ Time-Dependent Departure Process for Use in Tandem Queueing Networks. INFORMS Journal on Computing 25(4):758-773. http://dx.doi.org/10.1287/ijoc.1120.0538

INFORMS Journal on Computing
Vol. 25, No. 4, Fall 2013, pp. 758–773
ISSN 1091-9856 (print) | ISSN 1526-5528 (online)

http://dx.doi.org/10.1287/ijoc.1120.0538

Fitting the $Ph_t/M_t/s/c$ Time-Dependent Departure Process for Use in Tandem Queueing Networks

Walid W. Nasr
Olayan School of Business, American University of Beirut, Riad El-Solh/Beirut 1107 2020, Lebanon, wn12@aub.edu.lb

Michael R. Taaffe
Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, Virginia 24061, taaffe@vt.edu

This paper considers time-dependent $Ph_t/M_t/s/c$ queueing nodes and small tandem networks of such nodes. We examine characteristics of the departure processes from a multiserver queueing node; in particular, we focus on solving for the first two time-dependent moments of the departure-count process. A finite set of partial moment differential equations is developed to numerically solve for the departure-count moments over specified intervals of time $[t_i, t_i + \tau_i)$. We also present a distribution fitting algorithm to match these key characteristics with a $\widetilde{Ph}_t$ process serving as the approximate departure process. A distribution fitting algorithm is presented for time-dependent point processes where a two-level balanced mixture of Erlang distribution is used to serve as the approximating process. We then use the $\widetilde{Ph}_t$ approximating departure process as the approximate composite arrival process to downstream node(s) in a network of tandem queues.

Key words: queues; queueing; tandem queues; algorithms; phase-type distribution; nonstationary processes; queueing networks; time dependent; transient; mixture of Erlangs; count process; moment matching
History: Accepted by Winfried Grassmann, Area Editor for Computational Probability and Analysis; received February 2012; revised July 2012; accepted October 2012. Published online in *Articles in Advance* December 20, 2012.

1. Introduction

We consider tandem queues with time-dependent Markovian (M_t) service distributions, multiple servers, and time-dependent phase-type (Ph_t) arrival distributions. This paper can be divided into two main parts, where the objective of the first part is to present an efficient algorithm to numerically solve for the first two moments of the time-dependent departure-count process over a time interval $[t, t+\tau)$, where $t \geq 0$ and $\tau > 0$. The objective of the second part is to present an algorithm to fit the departure-count moments of the $Ph_t/M_t/s/c$ departure process with an approximate time-dependent phase-type process ($\widetilde{Ph}_t$).

Although the primary motivation of this paper is presenting a fitting algorithm for time-dependent point processes and the analysis of time-dependent queues in tandem, it is also motivated by the larger goal of constructing a queueing network analyzer for time-dependent systems (QNATS). The goal of developing QNATS is to provide accurate and efficient approximations for time-dependent, transient, and stationary systems. A queueing network analyzer (QNA) for stationary systems at steady state is presented in Whitt (1983), where steady-state congestion measures at each node as well as the entire network are approximated. A computationally attractive property of QNA is that each node can be analyzed independently. The computational efficiency of QNA is achieved by decomposing the network and analyzing each node individually. Network decomposition and analysis is achieved in two main steps: (1) At every node in the network, QNA approximates key characteristics of the departure process. This allows each node in the network to be analyzed separately because the arrival process at downstream nodes is the superposition of departure processes from upstream nodes and outside arrivals; (2) Congestion measures at each node are approximated separately as a function of the mean and coefficient of variation of the approximated arrival processes and the user defined service processes. The main steps in the development of approximation methods for time-dependent queueing systems are similar to the main steps in the development of QNA; i.e., decompose the network by approximating the traffic flow among nodes, then analyze each node independently.

Network decomposition significantly reduces the computational effort required. Consider, for example, a tightly coupled two-node tandem network where the arrival process to the first node (upstream node) is a Ph distribution with n_1 phases and the arrival process to the second node (downstream node) is the departure process from the upstream node. The joint number-in-system Markovian representation of such a system has $n_1(c_1+1)(c_2+1)$ states, where c_i is

the capacity of node i for $i = 1, 2$. If the arrival process to the downstream node can be approximated with an approximate $\widetilde{Ph}$ process with n_2 phases, then each node can now be analyzed separately where the number of states in the Markovian representation is $n_1(c_1+1)+n_2(c_2+1)$. If, for example, $n_1 = n_2 = 4$ and $c_1 = c_2 = 100$, then the set of joint number-in-system Kolmogorov forward equations (KFEs) is 40,804 whereas the set of number-in-system KFEs for the decomposed system is 808.

Providing approximations for a queueing node's time-dependent departure process is an integral part of analyzing time-dependent queues in tandem and queueing networks. The composite nodal arrival process can rarely be approximated by a specific distribution or family of distributions and approximating a non-Poisson process with a Poisson process can lead to inaccurate results (Gerhardt and Nelson 2009). In a network with Markovian time-dependent service times, a general $\tilde{G}_t/M_t/s/c$ approximate system represents a node in the network without loss of generality, where $\tilde{G}_t$ is the time-dependent general approximate composite nodal arrival process. Johnson and Taaffe (1988, 1989, 1990) demonstrate algorithms that arbitrarily closely approximate input processes for general models produced by matching moments of the input processes and distributions to moments of approximating phase-type processes and distributions. We conclude that the $G_t/M_t/s/c$ system can be represented and closely approximated by a $Ph_t/M_t/s/c$ system.

1.1. Departure Counts vs. Departure Intervals

Consider the departure process from a single $Ph_t/M_t/s/c$ node. We would like to extract some key characteristics of the departure process from the $Ph_t/M_t/s/c$ node. We then use those characteristics to construct a point process to serve as an accurate approximation for the time-dependent departure process. Without extracting all characteristics of the true departure process (with the exception of a few special cases) any point process constructed from the partial set of characteristics is necessarily an approximation. Consider two approaches to characterizing the true departure process. The first approach is to identify a set of characteristics for the interdeparture time intervals (or perhaps the intervals of time between time t and the time of the kth departure after time t). The second approach is to identify a set of characteristics for the count of the number of departures within a time interval.

An example of the first approach for approximating the departure process might be to obtain the first two (or more) moments of the time-until-the-kth departure starting from time t, and then fitting those moments to a Ph_t process that would then serve as the approximating departure process. If the time-until-the-kth departure distribution were known, then we could compute characteristics of the time between successive departures, such as moments. We do not describe the computational details of the interval-based approach here, but have found that interval-based methods can require significantly more computation for time-dependent processes and have concluded that count-process based methods (described next) are far more efficient.

The second approach considers a time interval $[t, t+\tau)$ for $t \geq 0$ and $\tau > 0$ and obtains the moments of the number of departures over the interval and then constructs an approximating departure process using a Ph distribution. We choose the second approach because it is computationally more friendly and we solve for the first two moments of the number of departures in the interval $[t, t+\tau)$, $t \geq 0$, $\tau > 0$. Notice that matching the second moment until the kth departure or the second moment of the number of departures within a time interval captures autocorrelation between successive interdepartures; see Gusella (1991) and Whitt (1982). Let $D_t(\tau) \equiv$ number of departures in the interval $[t, t+\tau)$, $t \geq 0$, $\tau > 0$. For any time $t \geq 0$ and $\tau > 0$ the first two moments of $D_t(\tau)$ provide partial information about the departure process.

Algorithmic approaches to fit point processes using Markovian distributions have received extensive attention in the literature and a minimal representation of Ph and Markovian arrival processes (MAPs) in Telek and Horvath (2007) allows for more efficient fitting algorithms. The algorithmic approaches to fit stationary systems do not capture time dependency and lead to erroneous results if applied to nonstationary systems. To the best of our knowledge, we could not find efficient algorithmic approaches to fit time-dependent systems with time-dependent Markovian processes.

1.2. Time-Dependent Queueing Systems

Many real-world queueing models exhibit time dependency in their arrival/service processes or in system capacity or number of servers. In a cellular communication network, a geographic cell representing tower coverage can be modeled by a queueing node where the arrival rate to a cell in the network is a result of mobile users crossing from one cell to a neighboring cell (Montenegro and Sengoku 1992). The time dependency in the arrival rate to a node is because of the dependency of vehicle traffic as a function of time of day. Queueing models investigating air traffic can be unrealistic if time dependency is ignored (Koopman 1972). Aircraft arrival rates as well as airport capacity are dependent on the time of day and other factors such as weather (Peterson et al. 1995). Call centers also display time dependency in the rate of incoming calls over the course

of the day (Avramidis et al. 2004). Also, situations exist where stochastic equilibrium cannot be assumed within the entire time period being considered. Many systems that do not exhibit time dependency in the arrival and service processes still require transient analysis (Odoni and Roth 1983).

Time-dependent Markovian queues with exponential interarrival and service distributions have received attention in the literature. Nonstationary Poisson processes provide a realistic fit to many real-world models and ignoring time dependency of the arrival or service process by using an average rate leads to erroneous results (Harrod and Kelton 2006). Margolius (1999) considers the $M_t/M_t/s/c$ system and presents formulas to compute the transient distribution of the number-in-system and discusses extending the analysis to Ph_t arrival or service distributions. The work presented in Margolius (1999) is motivated by a real-world police dispatching problem where the frequency of calls for police assistance is dependent on the time of day. Knessl (2002) considers the single server time-dependent Markovian $M_t/M_t/1$ system where the traffic intensity has a special form. Exact expressions for the time-dependent probability distribution of the number-in-system are derived in Knessl (2002) as an alternative to numerically solving the corresponding differential equations. Bekker and de Bruin (2010) refer to previous work analyzing $M/M/s$ and $M/G/s/s$ systems to model healthcare systems concerned with capacity planning and consider an $M_t/H/s/s$ system where the arrival rate is piecewise constant. Bekker and de Bruin (2010) also present a brief literature review on queueing models with time-dependent Poisson arrivals, and on the impact of time dependency in the arrival process on system performance. A thorough review is presented in Green et al. (2007) on setting staffing requirements for service systems experiencing time-dependent demand where real-world applications include call centers, police stations, bank tellers, and hospital emergency rooms. In Jennings et al. (1996), systems such as call centers, where the number of servers can be made time dependent, are considered to account for the time dependency in the arrival process. Kaczynski et al. (2011) analyze the transient behavior of an $M/M/s$ system with k entities initially present. They derive the probability distribution of the sojourn time of the nth customers as well as present algorithms to calculate the covariance between the sojourn time of customers. We also refer to Kaczynski et al. (2011) for a literature review on cases of queueing systems where transient analysis is investigated. Our model considers the $Ph_t/M_t/s/c$ system as well as the $M_t/M_t/s/c$ special case. We refer to Gerhardt and Nelson (2010) for a very thorough literature investigation on fitting point processes to Ph distributions.

Although Monte Carlo simulation appears to be a straightforward method for estimating time-dependent system performance, constructing and exercising Monte Carlo simulation models can be tedious and computationally prohibitive. To compute system mean performance measures at time t, multiple independent replications of the simulation are required. Further, to estimate system performance at many different values of t, we need to run many independent sets of multiple independent replications to provide independent estimators and their corresponding standard errors. Thus Monte Carlo simulation can be impractical over long time intervals for time-dependent systems. We use Monte Carlo simulation to measure the accuracy of the approximation presented in this paper when solving for estimates on the number-in-system in a two-node tandem network. When testing the quality of our approximations on smaller networks, we solve the joint number-in-system KFEs as an alternative to Monte Carlo simulation. The obvious advantage of this approach is that the numerical integration results are as accurate as we desire (to machine-level accuracy). The drawback can be the computational infeasibility because of the possibly large number of states. The rest of the paper continues as follows. In §2 we present the notation for the Ph_t distribution. In §3 we introduce the Markovian representation of the departure-count process and in §4 we derive the partial moment differential equations (PMDEs) to solve for the first two moments of the departure-count process. In §5 we present an algorithm that fits the departure-count moments to an approximate $\widetilde{Ph}_t$ distribution.

2. The Ph_t Count Process

In this section we present the notation used for the Ph_t distribution (for more details on the Ph_t distribution representation see Rueda and Taaffe 2004 or Nelson and Taaffe 2004).

- Let m_A represent the number of phases in the arrival process.
- For some value of time t let $\boldsymbol{\alpha}(t)$ be a $m_A \times 1$ vector representing the arrival-process initial-state probabilities, where $\alpha_i(t)$ is the conditional probability that an entity will start its arrival process in phase i, given that the arrival process starts at time t.
- Let $\mathbf{a}(t)$ be a matrix of dimension $m_A \times (m_A + 1)$ describing the time-dependent routing probabilities in the Ph arrival process. For example, $a_{ij}(t)$ represents the instantaneous time-dependent probability that at time t an entity that is completing its time in phase i will then proceed to phase j. When $j = m_A + 1$, $a_{ij}(t)$ represents the time-dependent probability that an entity completing its time in phase i will then proceed to the absorbing state; i.e., an entity arrives to the queueing node.

• Let $\boldsymbol{\lambda}(t)$ be a $m_A \times 1$ vector representing the arrival-phase rates (i.e., the reciprocal of the arrival-phase mean resident times). Thus $\lambda_i(t)$ is the departure rate from phase i at time t.

• Let $\{A(t)\colon t \geq 0\}$ be the state of the arrival process at time t, where $A(t) = 1, 2, \ldots, m_A$.

Consider the arrival-count process produced by the time-dependent Ph_t process over time interval $[t, t+\tau)$.

• Let $\{C_t(\tau)\colon t \geq 0\}$ be the number of arrivals in the interval $[t, t+\tau)$, $t \geq 0$, $\tau > 0$. Thus $C_t(\tau)$ is the arrival-count process over the time interval $[t, t+\tau)$.

The Kolmogorov forward and moment differential equations for the joint arrival-and-count process for the Ph_t arrival process over the interval $[t, t+\tau)$ are

$$P'_{c,l} = -\lambda_l P_{c,l} + \sum_{i=1}^{m_A} \lambda_i a_{i,l} P_{c,i} + \sum_{i=1}^{m_A} \lambda_i a_{i,m_A+1} \alpha_l P_{c-1,i},$$
$$\text{where} \quad P_{c,l} = \mathrm{P}\big(C_t(\tau) = c,\ A(t+\tau) = l\big); \quad (1)$$

and

$$\begin{aligned}
&\mathrm{E}'[C_t^n(\tau), A(t+\tau) = l] \\
&\quad = \sum_{c=1}^{\infty} c^n P'_{c,l} = -\lambda_l \mathrm{E}[C_t^n(\tau), A(t+\tau) = l] \\
&\qquad + \sum_{i=1}^{m_A} \lambda_i a_{i,l} \mathrm{E}[C_t^n(\tau), A(t+\tau) = i] \\
&\qquad + \sum_{i=1}^{m_A} \lambda_i a_{i,m_A+1} \alpha_l \left(\sum_{k=1}^{n} \binom{n}{k} \mathrm{E}[C_t^k(\tau), A(t+\tau) = i] \right. \\
&\qquad\qquad \left. + \mathrm{P}(A(t+\tau) = i) \right), \quad (2)
\end{aligned}$$

for $l = 1, \ldots, m_A$ and $n = 1, 2, 3, \ldots$.

For a time interval $[t, t+\tau)$, $t \geq 0$, and $\tau > 0$, the entire probability distribution of the count process can be calculated by numerically solving Equation (1). The nth moment can be calculated numerically using Equation (2). Numerically solving for the nth moment also requires solving for the first $n-1$ moments as well as the probability that the Ph_t arrival process is in phase l at time t, $P(A(t) = l)$, for $l = 1, \ldots, m_A$ and $t \geq 0$. The probability that the Ph_t process is in phase l for $l = 1, \ldots, m_A$ can be calculated by solving

$$\begin{aligned}
\mathrm{P}'(A(t) = l) &= -\lambda_l \mathrm{P}(A(t) = l) + \sum_{i=1}^{m_A} \lambda_i a_{i,l} \mathrm{P}(A(t) = i) \\
&\quad + \sum_{i=1}^{m_A} \lambda_i a_{i,m_A+1} \alpha_l \mathrm{P}(A(t) = i). \quad (3)
\end{aligned}$$

3. The Departure-Count Process: The $Ph_t/M_t/s/c$ Departure Process Kolmogorov Forward Equations

The state of the $Ph_t/M_t/s/c$ system is represented by the number-in-system at time t and the phase of the arrival process at time t. To directly consider the associated departure-count process we augment the usual state space by defining a random variable that counts the number of entities departing the system in the interval $[t, t+\tau)$. Next we define the augmented state space for the $Ph_t/M_t/s/c$ system and list the associated KFEs.

For the $Ph_t/M_t/s/c$ system let the number-in-system be $\{N(t)\colon t \geq 0\}$, where $N(t) = 0, 1, 2, \ldots, c$; and the state of the arrival process be $\{A(t)\colon t \geq 0\}$, where $A(t) = 1, 2, \ldots, m_A$; and let the number of departures in the interval $[t, t+\tau)$ be $\{D_t(\tau)\colon t \geq 0, \tau > 0\}$, where $D_t(\tau) = 0, 1, 2, \ldots$. Define the state probability $P_{i,d,t;l}(\tau) \equiv \mathrm{P}(N(t+\tau) = i, D_t(\tau) = d, A(t+\tau) = l)$. Thus, $P_{i,d,t;l}(\tau)$ is the probability that the number-in-system at time $(t+\tau)$ is i, the arrival process is in phase l at time $(t+\tau)$, and the number of departures in the interval $[t, t+\tau)$ is d. Figure 1 shows the augmented state space of the special case of the $M_t/M_t/s/c$ system. In Figure 1, (i, j) represents a state where the number-in-system at time $(t+\tau)$ is i and the number of departures in the interval $[t, t+\tau)$ is j.

We abbreviate $P_{i,d,t;l}$ with $P_{i,d;l}$, and $P'_{i,d;l}$ is the corresponding derivative with respect to τ. Let μ be the service rate. The number-in-system departure-count process KFEs (NSDC-KFEs) within the interval $[t, t+\tau)$ are given by Equation (4):

$$\begin{aligned}
P'_{0,0;l} &= -\lambda_l P_{0,0;l} + \sum_{j=1}^{m_A} a_{j,l} \lambda_j P_{j,0;0}, \quad \text{for } i = 0; \\
P'_{i,0;l} &= -\lambda_l P_{i,0;l} - \min(i,s)\mu P_{i,0;l} + \sum_{j=1}^{m_A} a_{j,m_A+1} \lambda_j P_{i-1,0;l} \\
&\quad + \sum_{j=1}^{m_A} \alpha_l a_{j,l} \lambda_j P_{i,0;j}, \quad \text{for } 0 < i < c; \\
P'_{0,d;l} &= -\lambda_l P_{0,d;l} + \mu P_{1,d-1;l} \\
&\quad + \sum_{j=1}^{m_A} a_{j,l} \lambda_j P_{0,d;j}, \quad \text{for } d > 0; \\
P'_{i,d;l} &= -\lambda_l P_{i,d;l} - \min(i,s)\mu P_{i,d;l} \\
&\quad + \sum_{j=1}^{m_A} \alpha_l a_{j,m_A+1} \lambda_j P_{i-1,d;j} + \min(i+1,s)\mu P_{i+1,d-1;l} \\
&\quad + \sum_{j=1}^{m_A} a_{j,l} \lambda_j P_{i,d;j}, \quad \text{for } 0 < i < c \text{ and } d > 0; \quad (4) \\
P'_{c,d;l} &= -\lambda_l P_{c,d;l} - s\mu P_{c,d;l} + \sum_{j=1}^{m_A} a_{j,l} \lambda_j P_{c,d;j} \\
&\quad + \sum_{j=1}^{m_A} a_{j,m_A+1} \lambda_j \alpha_l P_{c-1,d;j} \\
&\quad + \sum_{j=1}^{m_A} a_{j,m_A+1} \lambda_j \alpha_l P_{c,d;j}, \quad \text{for } d \geq 0.
\end{aligned}$$

Notice that, theoretically, the number of states is infinite because within a time interval there is no upper limit on the number of arrivals—and therefore

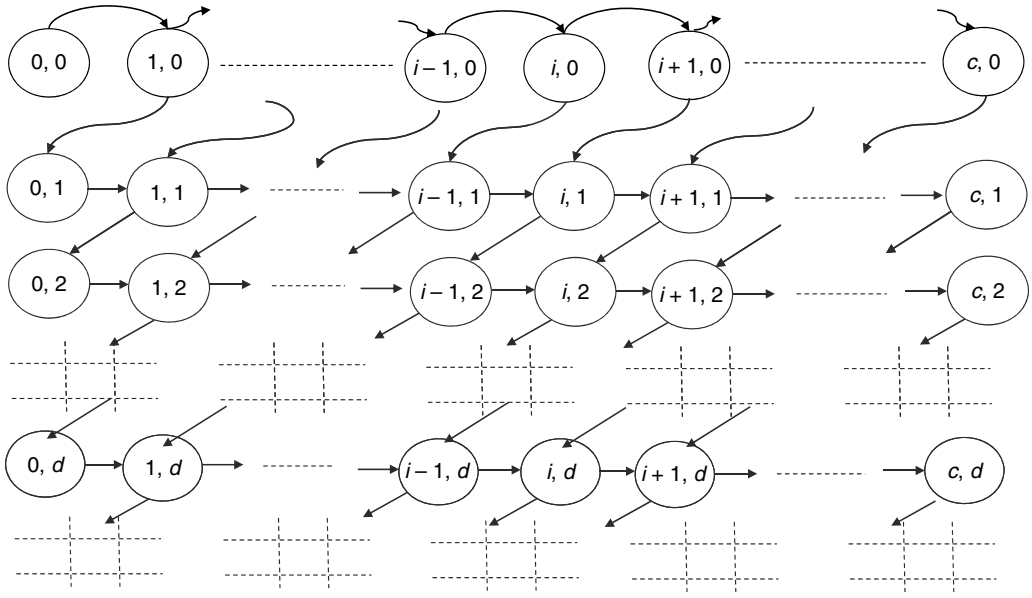

Figure 1 **Number-in-System and Departures States $M_t/M_t/s/c$**

departures; thus d can be infinite and so the number of NSDC-KFEs is infinite. We define a "practical upper limit" on the number of departures, d_m, and should be large enough with respect to the time interval for accurate approximations. Thus, by truncating the state space in this manner the number of NSDC-KFEs is $m_A(c+1)(d_m+1)$.

The choice of d_m could result in less accurate approximations if it is set too small. There exist positive integers d^* and n, where increasing d_m beyond d^* no longer improves the quality of the approximations up to the nth decimal point. So d_m can be the largest positive integer that satisfies $\mathrm{P}(D_t(\tau) \geq d_m) < \epsilon$ for an appropriate choice of ϵ. The NSDC-KFEs are not a computationally practical choice when solving for the first two moments of the departure-count process but are used in the next section to derive a set of finite departure-count PMDEs (DC-PMDEs).

4. The Departure-Count Partial-Moment Differential Equations (DC-PMDEs)

To compute the first two moments of the number of departures over the interval $[t, t+\tau)$, we can solve a set of partial-moment differential equations where the number of equations is not a function of d_m or ϵ. No approximations are required and hence we can obtain numerically exact solutions over the time interval $[t,\ t+\tau)$ for the $Ph_t/M_t/s/c$. In §4.2 we present a simple algorithm to obtain the first two moments of the departure-count process over any time interval.

4.1. The $Ph_t/M_t/s/c$ DC-PMDEs

We next develop a numerically exact computational approach using $2m_A(c+1)$ DC-PMDEs and $m_A(c+1)$ number-in-system KFEs (NS-KFEs). The union of the set of DC-PMDEs with the set of NS-KFEs, (NS-KFE ∪ DC-PMDE), form a closed set of differential equations. This set of (NS-KFE ∪ DC-PMDE) can be numerically integrated (along with an initial condition) to produce numerically exact results. This approach dramatically decreases the number of differential equations, but of course we no longer have the entire time-dependent probability mass function for the state of the system—we just have moments. We obtain the PMDEs by summing the derivatives of the state probabilities. The partial-moment derivative is with respect to τ, where $l = 1, \ldots, m_A$ and $i = 1, \ldots, c$:

$$\mathrm{E}'[D_t(\tau), N(t+\tau) = i, A(t+\tau) = l] \equiv \sum_{d=1}^{\infty} dP'_{i,\,d;\,l}, \quad (5)$$

$$\mathrm{E}'[D_t^2(\tau), N(t+\tau) = i, A(t+\tau) = l] \equiv \sum_{d=1}^{\infty} d^2 P'_{i,\,d;\,l}. \quad (6)$$

The resulting DC-PMDEs together with the NS-KFEs form a finite set of quasi-closed PMDEs, NS-KFE ∪ DC-PMDE, and are the following:

Let $\mathrm{E}[D, i; l] = \mathrm{E}[D_\tau(\tau), N(t+\tau) = i, A(t+\tau) = l]$ and $P_{i;\,l} = \mathrm{P}(N(t+\tau) = i, A(t+\tau) = l)$ for $l = 1, \ldots, m_A$.

The NS-KFEs for $l = 1, \ldots, m_A$,

$$P'_{0;\,l} = -\lambda_l P_{0;\,l} + \sum_{j=1}^{m_A} a_{j,l}\lambda_j P_{0;\,j} + \mu P_{1;\,l},$$

$$P'_{i;\,l} = -\lambda_l P_{i;\,l} - \min(i, s)\mu P_{i;\,l} + \sum_{j=1}^{m_A} a_{j,l}\lambda_j P_{i;\,j} + \min(i+1, s)\mu P_{i+1;\,l}$$

$$+\sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l P_{i-1;\,j} \quad \text{for } 0<i<c, \qquad (7)$$

$$P'_{c;\,l} = -\lambda_l P_{c;\,l} - s\mu P_{c;\,l} + \sum_{j=1}^{m_A} a_{j,\,l}\lambda_j P_{c;\,j} + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l P_{c-1;\,j} + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l P_{c;\,j}.$$

The first moment DC-PMDEs for $l=1,\ldots,m_A$,

$$\begin{aligned}
\mathrm{E}'[D,0;l] &= -\lambda_l \mathrm{E}[D,0;l] + \sum_{j=1}^{m_A} a_{j,\,l}\lambda_j \mathrm{E}[D,0;j] + \mu\mathrm{E}[D,1;l] + \mu P_{1;\,l},\\
\mathrm{E}'[D,i;l] &= -\lambda_l \mathrm{E}[D,i;l] - \min(i,s)\mu\mathrm{E}[D,i;l] + \sum_{j=1}^{m_A} a_{j,\,l}\lambda_j \mathrm{E}[D,i;j] + \min(i+1,s)\mu\mathrm{E}[D,i+1;l]\\
&\quad + \min(i+1,s)\mu P_{i+1;\,l} + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l \mathrm{E}[D,i-1;j] \quad \text{for } 0<i<c, \qquad (8)\\
\mathrm{E}'[D,c;l] &= -\lambda_l \mathrm{E}[D,c;l] - s\mu\mathrm{E}[D,c;l] + \sum_{j=1}^{m_A} a_{j,\,l}\lambda_j \mathrm{E}[D,c;j] + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l \mathrm{E}[D,c-1;j]\\
&\quad + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l \mathrm{E}[D,c;j].
\end{aligned}$$

The second moment DC-PMDEs for $l=1,\ldots,m_A$,

$$\begin{aligned}
\mathrm{E}'[D^2,0;l] &= -\lambda_l \mathrm{E}[D^2,0;l] + \sum_{j=1}^{m_A} a_{j,\,l}\lambda_j \mathrm{E}[D^2,0;j] + \mu\mathrm{E}[D^2,1;l] + 2\mu\mathrm{E}[D,1;l] + \mu P_{1;\,l},\\
\mathrm{E}'[D^2,i;l] &= -\lambda_l \mathrm{E}[D^2,i;l] - \min(i,s)\mu\mathrm{E}[D^2,i;l] + \sum_{j=1}^{m_A} a_{j,\,l}\lambda_j \mathrm{E}[D^2,i;j] + \min(i+1,s)\mu\mathrm{E}[D^2,i+1;l]\\
&\quad + 2\min((i+1),s)\mu\mathrm{E}[D,i+1;l] + \min(i+1,s)\mu P_{i+1;\,l} + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l \mathrm{E}[D^2,i-1;j],\\
&\qquad \text{for } 0<i<c, \qquad (9)\\
\mathrm{E}'[D^2,c;l] &= -\lambda_l \mathrm{E}[D^2,c;l] - s\mu\mathrm{E}[D^2,c;l] + \sum_{j=1}^{m_A} a_{j,\,l}\lambda_j \mathrm{E}[D^2,c;j] + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l \mathrm{E}[D^2,c-1;j]\\
&\quad + \sum_{j=1}^{m_A} a_{j,\,m_A+1}\lambda_j\alpha_l \mathrm{E}[D^2,c;j].
\end{aligned}$$

Notice that the number of NS-KFEs is $m_A(c+1)$, and similarly the number of departure-count first- and second-moment PMDEs is $2m_A(c+1)$. The total number of (NS-KFE ∪ DC-PMDE)'s, Equations (7)–(9), is $3m_A(c+1)$. So the number of (NS-KFE ∪ DC-PMDE)'s to be evaluated is not a function of d_m, the upper bound on the number of departures within an interval. In Table 1 we recap the KFEs and PMDEs presented so far to derive and calculate the departure-count moments.

4.2. The $Ph_t/M_t/s/c$ Numerically Exact Departure-Count Moment Algorithm

In this section we present an algorithm that makes use of the (NS-KFE ∪ DC-PMDE)'s as well as initial conditions and stopping conditions to produce the numerically exact first two moments of the departure counts within a selected time interval. We call this algorithm the numerically exact departure-count moment algorithm (NEDMA). The NEDMA can be called at any time t to compute the departure-count moments for a time interval of width τ; i.e., $\mathrm{E}[D_t^j(\tau)]$, $j=1,2$. We assume here that information about the probabilities of the number-in-system at time t have been computed via the full set of NS-KFEs.

The $Ph_t/M_t/s/c$ NEDMA:

Step 0. Choose a start time, t, and an interval width, τ.

Step 1. Initialize the interval departure-count moments, $\mathrm{E}[D_t^j(0)]=0$, $j=1,2$.

Step 2. Numerically solve Equations (7)–(9) for the time interval $[t, t+\tau)$.

Table 1 Summary Table for Equation Acronyms

Acronym	Expression	Equation reference
NSDC-KFE	Number-in-system departure-count Kolmogorov forward equation	Equation (4)
DC-PMDEs	Departure-count partial moment differential equations	Equations (5), (6) and expanded in Equations (8), (9)
NS-KFE	Number-in-system Kolmogorov forward equation	Equation (7)

Notice that the NEDMA computes the first two moments within an interval, and it can also be used to compute the first two departure-count moments for subintervals of $[t, t+\tau)$. Executing the NEDMA over the interval $[t, t+\tau)$ returns the first two departure-count moments over the interval $[t, t+\tau')$, where $0 < \tau' < \tau$. It also allows the calculation of the first moment departure-count over any time subinterval $[t_1, t_2)$ for any t_1 and t_2, where $[t_1, t_2) \subset [t, t+\tau)$ and $t_1 < t_2$. The first departure-count moment over a subinterval $[t_1, t_2)$ is $\mathrm{E}[D_{t_1}(t_2 - t_1)] = \mathrm{E}[D_t(t_2 - t)] - \mathrm{E}[D_t(t_1 - t)]$.

4.3. The Ph_t Departure-Count-Process Moment Intervals

Define Ph_t as the approximating process for the true departure process, and of course $\widetilde{Ph}_t$ can also serve as an approximating arrival process at the downstream node. We make use of two types of departure-count random variables. The first is a random variable that counts the number of departures over a relatively large interval, and the second is a departure-count random variable over each of a set of subintervals of the larger interval. We use the larger intervals to approximate departure-count second moments and use the smaller intervals to approximate the departure-count first moments. Using larger time intervals for approximating the second moment of the departure count implicitly takes into account autocorrelation in the departure process; see Gusella (1991) and Whitt (1982). We refer to the larger intervals as second moment intervals (SMIs) and the smaller subintervals as first moment intervals (FMIs). For a time interval of length T, the SMIs and FMIs are further described next.

4.3.1. Second-Moment Intervals. Define an interval of size Δ to be a second-moment interval, where $k\Delta = T$ for some positive integer k. These SMIs are intervals during which a $\widetilde{Ph}_t$ of a particular structure will serve as our approximating arrival process. Within an SMI the approximating $\widetilde{Ph}_t$ process can be time dependent, but its structure (number of phases) is not. The second moment of the count process generated by the approximating $\widetilde{Ph}_t$ process over an SMI matches the second moment of the number of departures occurring within an SMI.

4.3.2. First-Moment Intervals. Define an interval of size δ, where $n\delta = \Delta$ for some positive integer n, to be a first-moment interval. FMIs are subintervals of the SMIs. Figure 2 shows a choice of Δ and the corresponding δ's. These FMIs are intervals where the approximating $\widetilde{Ph}_t$ process is stationary (neither the structure nor the parameters change). The fitted $\widetilde{Ph}_t$ process is piecewise constant across the FMIs. The first moment of the count process generated by the approximating $\widetilde{Ph}_t$ process over an FMI matches the first moment of the number of departures occurring within an FMI.

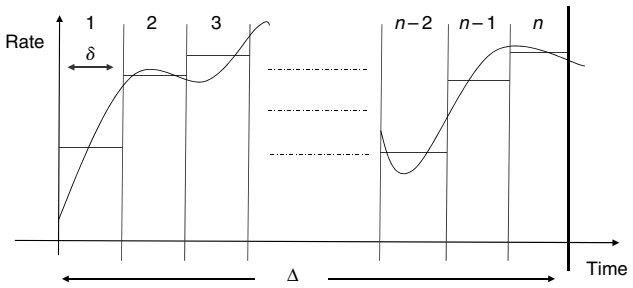

Figure 2 Intervals

4.4. Calculating the Departure-Count Moments Over the Intervals

In this section we make use of the NEDMA as presented in §4.2 to compute the second moments over k consecutive nonoverlapping intervals of length Δ as well as the first moment over $n \times k$ consecutive subintervals of length $\delta = \Delta/n$. Let $S_{(i)}$ be the second moment departure count for the ith SMI for $i = 1, 2, \ldots, k$. Let $F_{(i,j)}$ be the first moment departure-count process over the jth FMI within the ith SMI for $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, k$. We refer to the following algorithm by the moments across intervals algorithm (MIA):

Step 0. Set $i = 1$.

Step 1. Set $t_i = (i-1)\Delta$ and $\tau_i = \Delta$.

Step 2. Execute the NEDMA and set,

- $S_{(i)} = \mathrm{E}[D^2_{(i-1)\Delta}(\Delta)]$
- $F_{(i,j)} = \mathrm{E}[D_{(i-1)\Delta+(j-1)\delta}(\delta)]$ for $j = 1, \ldots, n$.

Step 3. Stop if $i = k$, otherwise set $i = i+1$ and go to Step 1.

This section presented the MIA to numerically solve for the time-dependent departure-count moments for a $Ph_t/M_t/s/c$ queueing node across the FMIs and SMIs. In the next section we assume that the set of first two moments for the departure counts across the set of SMI intervals and FMI subintervals is known; i.e., $S_{(i)}$ and $F_{(i,j)}$ are known for $i = 1, \ldots, k$ and $j = 1, \ldots, n$.

5. Fitting a Ph_t Distribution to the Departure-Count Moments

Given the set of departure-count moments we now describe construction of a point process to approximate the true departure process as a first step in constructing an approximating arrival process for a downstream node in a network. Note that the distribution fitting algorithm presented in this section can be applied to any point process where the count moments can be calculated over specified time intervals.

We change the parameter values across FMIs and SMIs so that the $\widetilde{Ph_t}$ state probabilities are held constant. We accomplish this by specifying the initial condition that all $P(A(t)=i) \Leftarrow \pi_i$, where $P(A(t)=i)$ are the $\widetilde{Ph_t}$ process phase-state probabilities for the ith phase and π_i are the corresponding steady-state probabilities.

Let $D_t(\tau)$ and $C_t(\tau)$ be the number of departures and the count process from the approximating $\widetilde{Ph_t}$ process, respectively, over the time interval $[t, t+\tau)$, $t \geq 0$, $\tau > 0$. We require the following properties when choosing the approximating $\widetilde{Ph_t}$ process over the time interval of length T. Recall that the interval of length T is divided into k subintervals of length Δ, and are in turn divided into n intervals of length δ.

1. The phase state probabilities of the approximating process are constant and equal to their steady-state probabilities throughout the interval of length T.

2. The values of the equilibrium phase state probabilities are not a function of the resulting count process.

3. The first moment of the $\widetilde{Ph_t}$ count process matches the departure-count first moment over the FMIs, $E[C_{(j-1)\delta}(\delta)] = E[D_{(j-1)\delta}(\delta)]$ for $j = 1, \ldots, kn$.

4. The second moment of the $\widetilde{Ph_t}$ count process matches the departure-count second moment over the SMIs, $E[C^2_{(j-1)\Delta}(\Delta)] = E[D^2_{(j-1)\Delta}(\Delta)]$ for $j = 1, \ldots, k$.

Properties 1 and 2 simplify the moment-matching algorithm and eliminate the need to consider conditioning on an arrival at the start of the SMI as one would do in considering the associated Palm process. We use the four properties as well as Equations (2) and (3) to present a distribution fitting algorithm in the next section.

5.1. Fitting Distribution: Balanced Two-Level Mixture of Erlangs

A desired property of the fitted $\widetilde{Ph_t}$ distribution is to have the flexibility to match high and low variability processes over any time interval $[t, t+\tau)$. A departure-count process is defined as having high/low variability over an interval $[t, t+\tau)$ if $E[D_t^2(\tau)](\geq / <) E[D_t(\tau)]^2 + E[D_t(\tau)]$. Let the approximating $\widetilde{Ph_t}$ be a two-level mixture of Erlangs of common order (2-MECO) as presented in Figure 3. The number of phases in each level is m_1 and m_2 for a total of $m_A = m_1 + m_2$ phases.

The shaded box represents the absorbing state. The probabilities of starting in phases 1 and (m_1+1) are α and $(1-\alpha)$, respectively. Let $m_1\lambda_1$ represent the transition rates from phases $1, \ldots, m_1$ and let $m_2\lambda_2$ represent the transition rates from phases $m_1+1, \ldots, m_1+m_2$. Let π_i represent the equilibrium phase-state probabilities for phase i, where $i = 1, \ldots, m_1+m_2$.

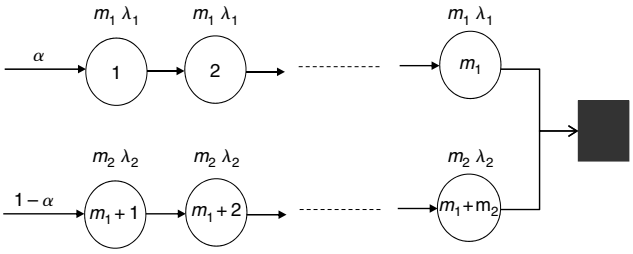

Figure 3 **Balanced 2-MECO—State Numbers and Phase Transition Rates**

We define the 2-MECO as balanced if $\alpha\lambda_1^{-1} = (1-\alpha)\lambda_2^{-1}$, which results in

$$\sum_{i=1}^{m_1} \pi_i = \sum_{i=m_1+1}^{m_1+m_2} \pi_i = \tfrac{1}{2}.$$

Notice that the balanced-means hyperexponential distribution described in Whitt (1982) is a special case of the 2-MECO when $m_1 = m_2 = 1$. The balanced 2-MECO and the balanced hyperexponential special case result in $(1-\alpha)\lambda_1 = \alpha\lambda_2$, or

$$\lambda_1 = \lambda_2\alpha/(1-\alpha). \tag{10}$$

An important characteristic of the balanced 2-MECO is that as long as Equation (10) is satisfied, the phase probabilities are equal to their steady-state values as expressed in Equation (11):

$$\pi_i = \begin{cases} \dfrac{1}{2m_1} & \text{for } i = 1, \ldots, m_1, \\ \dfrac{1}{2m_2} & \text{for } i = m_1+1, \ldots, m_1+m_2. \end{cases} \tag{11}$$

Let $\lambda_2 = \lambda$ and $\lambda_1 = \lambda\alpha/(1-\alpha)$. The balanced 2-MECO system is defined by the number of phases for both Erlang processes (m_1 and m_2) and the two parameters λ and α. In §5.2, we present an algorithm where λ and α are chosen to fit the departure-count first moments over the FMIs and the departure-count second moments of the SMIs for fixed values of m_1 and m_2. In §5.6, we choose the best structure of the 2-MECO distribution by selecting m_1 and m_2.

5.2. The 2-MECO Departure-Count Process

Notice that λ and α are held constant over an interval of length δ. Over an interval of length δ it is possible to represent Equation (2) for $n = 1, 2$ as a closed homogeneous set of differential equations (HSDE). Let $X_i(\tau) = E[C_t(\tau), A(t+\tau) = l]$ for $i = 1, \ldots, m_1 + m_2$ and $Y(\tau) = E[C_t^2(\tau)]$. Using Equation (2), the departure-count first moment is

$$E'[C_t(\tau)] = \sum_{i=1}^{m_1+m_2} X_i(\tau) = \frac{\lambda}{2(1-\alpha)}, \tag{12}$$

$$\Longrightarrow \quad E[C_t(\tau)] = \frac{\lambda\tau}{2(1-\alpha)}. \tag{13}$$

Using Equations (2) and (3), we can express the partial moments, $X_i(\tau)$s, and the second moment, $Y(\tau)$ by

$$X_1'(\tau) = \frac{-\alpha m_1\lambda}{(1-\alpha)}X_1(\tau) + \frac{\alpha^2 m_1\lambda}{(1-\alpha)}X_{m_1}(\tau) + \alpha m_2\lambda X_{m_1+m_2}(\tau) + \frac{\alpha\lambda}{2(1-\alpha)}v(\tau),$$

$$X_i'(\tau) = \frac{\alpha m_1\lambda}{(1-\alpha)}(X_{i-1}(\tau) - X_i(\tau)) \quad \text{for } i = 1,\dots,m_1,$$

$$X_{m_1+1}'(\tau) = -m_2\lambda X_{m_1+1}(\tau) + \alpha m_1\lambda X_{m_1}(\tau) + (1-\alpha)m_2\lambda X_{m_1+m_2}(\tau) + \frac{\lambda}{2}v(\tau), \quad (14)$$

$$X_i'(\tau) = m_2\lambda(X_{i-1}(\tau) - X_i(\tau)) \quad \text{for } i = (m_1+1),\dots,(m_1+m_2),$$

$$Y'(\tau) = \frac{2\alpha^2 m_1\lambda}{(1-\alpha)}X_{m_1}(\tau) + 2m_2\lambda X_{m_1+m_2}(\tau) + \frac{\lambda}{2(1-\alpha)}v(\tau),$$

$$v'(\tau) = 0,$$

where $v(0) = 1$. The partial probability terms of Equation (3) are always constant for the 2-MECO case if the initial values of Equation (3) are the same as their steady-state probabilities (Equation (11)). Equation (3) is not needed to solve the set of partial differential in (14). A HSDE is achieved with an additional dummy variable v where $v'(\tau) = 0$ and $v(0) = 1$. Let $\mathscr{C}$ be an $(m_1+m_2+2)\times(1)$ matrix/vector, where for $i = 1,\dots,m_1+m_2+2$,

$$\mathscr{C}_{(i)} = \begin{cases} X_i(\tau), & \text{for } i = 1,\dots,m_1+m_2; \\ Y(\tau), & \text{for } i = m_1+m_2+1; \\ v(\tau), & m_1+m_2+2. \end{cases}$$

We can represent the HSDE by

$$\mathscr{C}' = A\mathscr{C}, \quad (15)$$

where A is an $(m_1+m_2+2)\times(m_1+m_2+2)$ square matrix. Here we present an algorithm to calculate the second moment departure count over the ith SMI for a fixed α whereas λ is varied within the SMI to match the first moment departure count for every FMI within the SMI. The parameter λ is constant within an FMI but is allowed to change from one FMI to another. So λ is piecewise constant within an SMI. Recall that SMI is divided into n FMIs of length $\delta = \Delta/n$, where $F_{(i,j)} \equiv$ first moment departure count over the jth FMI within the ith SMI. We refer to the following algorithm by the 2-MECO partial moment algorithm (PMA):

Step 0. Set $\mathscr{C}^*$ to be the initial value vector of the partial moments:

- $\mathscr{C}^*_{(i)} = 0$ for $i = 1,\dots,m_1+m_2+1$;
- $\mathscr{C}^*_{(m_1+m_2+2)} = 1$;
- $j = 1$;

(the dimensions of $\mathscr{C}$ are $(m_1+m_2+2)\times 1$ and j represents the jth FMI).

Step 1. Set $\lambda = 2(1-\alpha)F_{(i,j)}/\delta$

(Step 1 follows from Equation (13)).

Step 2. Set A to be a square matrix of zeros with dimension $(m_1+m_2+2)\times(m_1+m_2+2)$.

Step 3. Enter the first moments coefficients in A as presented in Equation (14):

- Coefficients of $X_1(\tau)$: $A_{(1,1)} = -\alpha m_1\lambda/(1-\alpha)$, $A_{(1,m_1)} = \alpha^2 m_1\lambda/(1-\alpha)$, $A_{(1,m_1+m_2)} = \alpha m_2\lambda$ and $A_{(1,m_1+m_2+2)} = \alpha\lambda/(2(1-\alpha))$.
- Coefficients of $X_i(\tau)$ for $i = 1,\dots,m_1$: $A_{(i,i-1)} = \alpha m_1\lambda/(1-\alpha)$ and $A_{(i,i)} = -\alpha m_1\lambda/(1-\alpha)$.
- Coefficients of $X_{m+1}(\tau)$: $A_{(m_1+1,m_1+1)} = -m_2\lambda$, $A_{(m_1+1,m_1)}\alpha m_1\lambda$, $A_{(m_1+1,m_1+m_2)} = (1-\alpha)m_2\lambda$ and $A_{(m_1,m_1+m_2+2)} = \lambda/2$.
- Coefficients of $X_i(\tau)$ for $i = (m_1+1),\dots,(m_1+m_2)$: $A_{(i,i-1)} = m_2\lambda$ and $A_{(i,i)} = -m_2\lambda$.

Step 4. Set the second moments coefficients:

- Coefficients of $Y(\tau)$: $A_{(m_1+m_2+1,m_1+1)} = 2\alpha^2 m_1\lambda/(1-\alpha)$, $A_{(m_1+m_2+1,m_1+m_2)} = 2m_2\lambda$ and $A_{(m_1+m_2+1,m_1+m_2+2)} = \lambda/(2(1-\alpha))$.

Step 5. Solve the HSDE by solving the matrix exponential equation:

- $\mathscr{C} = e^{A\delta}\cdot\mathscr{C}^*$

Step 6. Calculate The first two moments at the end of the current interval:

- First moment departure count $= \sum_{i=1}^{m_1+m_2}\mathscr{C}_{(i)}$.
- Second moment departure count $= \mathscr{C}_{(m_1+m_2+1)}$.

Step 7. Stop if $j > n$, otherwise set $\mathscr{C}^* = \mathscr{C}$, $j = j+1$ and go to Step 1.

Example 5.2. Consider a point process where the time-dependent arrival rate is represented by Figure 4. Consider one SMI of length $\Delta = 4$ and four FMIs within the SMI of length δ. Notice that this example is presented for illustrative purposes and obviously the FMIs need to be much smaller to better capture the time dependency of the rate as presented in Figure 4. In this example we fit a balanced 2-MECO with $m_1 = m_2 = 2$ to the departure-count moments.

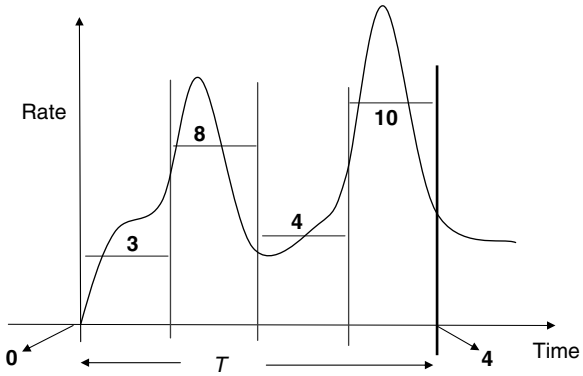

Figure 4 **Time-Dependent Departure Rate—Example 5.2**

Table 2 Moments

t	$X_1(t)$	$X_2(t)$	$X_3(t)$	$X_4(t)$	$\sum_{i=1}^{4} X_i(t)$	$Y(t)$
δ	0.96	0.88	0.72	0.44	3	12.45
2δ	2.96	2.89	2.73	2.42	11	135.21
3δ	3.96	3.89	3.73	3.42	15	244.58
4δ	6.46	6.39	6.23	5.92	25	658.02

Let $\alpha = 0.8$, and the first moment departure counts over the four FMI intervals are $F = [3, 8, 4, 10]$. Executing the PMA, we obtain the partial moments over the intervals shown in Table 2. The second moment of the departure count produced for $\alpha = 0.8$ is 658.02.

Now our objective is to determine α given the set of departure-count first moments over the four FMIs and the departure-count second moment over the SMI. Figure 5 is a plot of α versus the departure-count second moment.

The balanced 2-MECO has the flexibility to match low and high variability. A balanced 2-MECO with $2m$ phases achieves the lowest variability for $\alpha = 0.5$ and $m_1 = m_2 = m$ and is reduced to an Erlang distribution of order m. The lowest attainable departure-count second moment for Example 5.2 is 637.71. If a lower variability distribution is needed, then this can be achieved by increasing m. As illustrated in Figure 5, the plot of the second moment departure count versus α is symmetric around 0.5 and increases as α moves away from 0.5 toward zero or one. If we are looking for an α that corresponds to a departure-count second moment, it is sufficient to perform a simple search on the plot shown in Figure 5 over the interval [0.5, 1). For example, the value of α that satisfies the departure-count first moments $F = [3, 8, 4, 10]$ over the FMIs and a departure-count second moment of 800 for the SMI is 0.957. If S is the target second moment then the following algorithm searches for the α that matches S. We refer to the following algorithm as the fitting algorithm (FA). Let ϵ_1 be the desired accuracy:

Step 0. Set $\alpha = 0.75$ and $jump = 0.25$.

Step 1. Execute PMA and return the second moment departure count $\mathscr{C}_{(m_1+m_2+1)}$.

Step 2. Calculate the resulting error, $e = (\mathscr{C}_{(m_1+m_2+1)} - S)/(\sum_{i=1}^{m_1+m_2} \mathscr{C}_{(i)})$.

Step 3. If $|e| < \epsilon_1$ then return α and stop, otherwise

- if $e > 0$ then $jump = -|jump|/2$;
- if $e < 0$ then $jump = |jump|/2$;
- set $\alpha = \alpha + jump$ and go to Step 1.

Notice that the FA is very efficient because only 10 iterations would result in a fitted α that is within 2^{-10} of the desired α. The FA is executed for every SMI to obtain a fixed α within the SMI, and λ is fixed within every FMI and is calculated in Step 2 of the PMA. In §§5.3 and 5.4 we present results for tightly coupled two-node tandem networks because approximation error is most noticeable for these smaller networks and less noticeable in larger networks. We also consider tandem queues with no outside arrivals to amplify any approximation error resulting from fitting the departure process of the upstream node. The departure process approximation error is not diluted by superposing an external process to the approximate departure process that better demonstrates the quality of our approximations. We also report the execution time of the FA as a measure of the algorithm's efficiency.

5.3. Numerical Example 1: Two-Node Tightly Coupled Queueing Network

Node 1 parameters: $s_1 = 4$, $c_1 = 100$, $\mu_1 = 10(1 - 0.2\sin(0.6t))$.

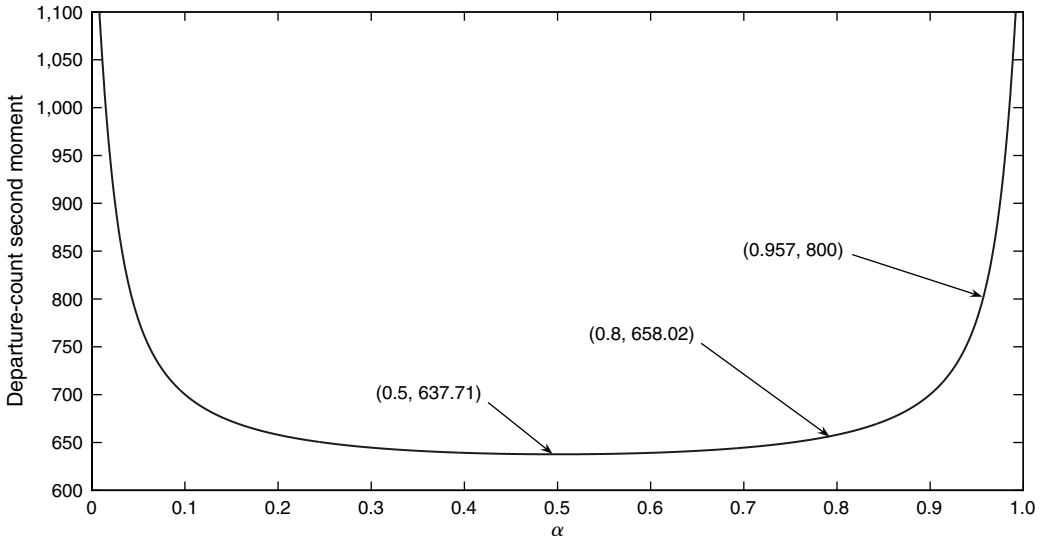

Figure 5 Second Moment Departure-Count vs. α: Example 5.2

Node 2 parameters: $s_2 = 2$, $c_2 = 100$, $\mu_2 = 20(1 - 0.5\sin(0.6t))$.
Outside Ph_t arrival distribution to node 1:

$$a = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{aligned} &\alpha = [0.1, 0.3, 0.6], \\ &\lambda = [10, 25, 50](1 - 0.3\cos(0.6t)). \end{aligned}$$

The system is analyzed over 20 time units that are divided into 10 SMIs of length $\Delta = 2$ time units. Each SMI is divided into 100 FMIs of length $\delta = 0.02$ time units for a total of 10^3 FMIs. The departure-count first two moments over the SMIs and FMIs are calculated using the MIA and are fitted to a piecewise constant balanced 2-MECO distribution with $m_1 = 4$ and $m_2 = 1$. The piecewise constant parameters of the fitted 2-MECO are obtained using the FA. The first and second moments of the number-in-system at node 1 are presented in Figure 6. The approximate first two moments of the number-in-system at node 2 are solved for using the fitted balanced 2-MECO as the input distribution. Solving the joint number-in-system KFEs was not possible because of the large number of states. The two-node system is simulated for 10,000 replications, where for each replication the number-in-system at node 2 for 10 different points in time is reported. The solid plots of Figure 7 show the approximate first two moments of the number-in-system at node 2. Figure 7 also shows the 95% confidence intervals for the first moment at the 10 chosen points in time as calculated by the simulation. The confidence interval bounds for the 10 points in time are denoted by * in the first plot of Figure 7. The second plot of Figure 7 shows the best estimate of the standard deviation obtained using simulation and is also represented by a *.

In Table 3 we take the best estimate for the first moment as obtained from the simulation and compare it with our approximate number-in-system for the 10 selected points in time. The maximum absolute percentage difference as shown in Table 3 is 6.98% at time $t = 8$ and the average of the 10 selected points is 1.56%. Similarly for the number-in-system standard deviation, the maximum absolute percentage difference as shown in Table 4 is 8.35% at $t = 8$ and the average over the 10 selected points is 2.18%.

We also report in Table 5 the number of differential equations solved and the time required to obtain the results. Notice that solving the joint number-in-system would require solving 30,603 KFEs, which is computationally infeasible because of the large computer memory requirement.

5.4. Numerical Example 2: Two-Node Tightly Coupled Queueing Network

Consider the following smaller two-node example system where we compare our results with the joint number-in-system. Node 1 parameters: $s_1 = 2$, $c_1 = 20$, and $\mu_1 = 25$.
Node 2 parameters: $s_2 = 2$, $c_2 = 50$, and $\mu_2 = 18$.

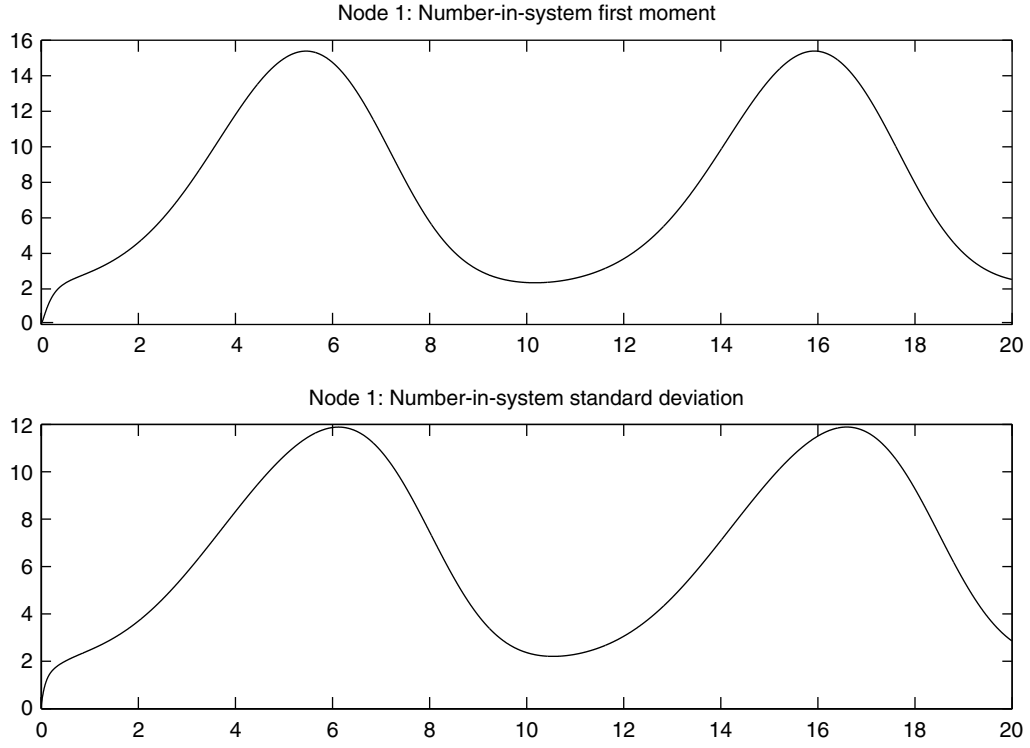

Figure 6 **Node 1: Number-in-System First Moment and Standard Deviation—Example 1**

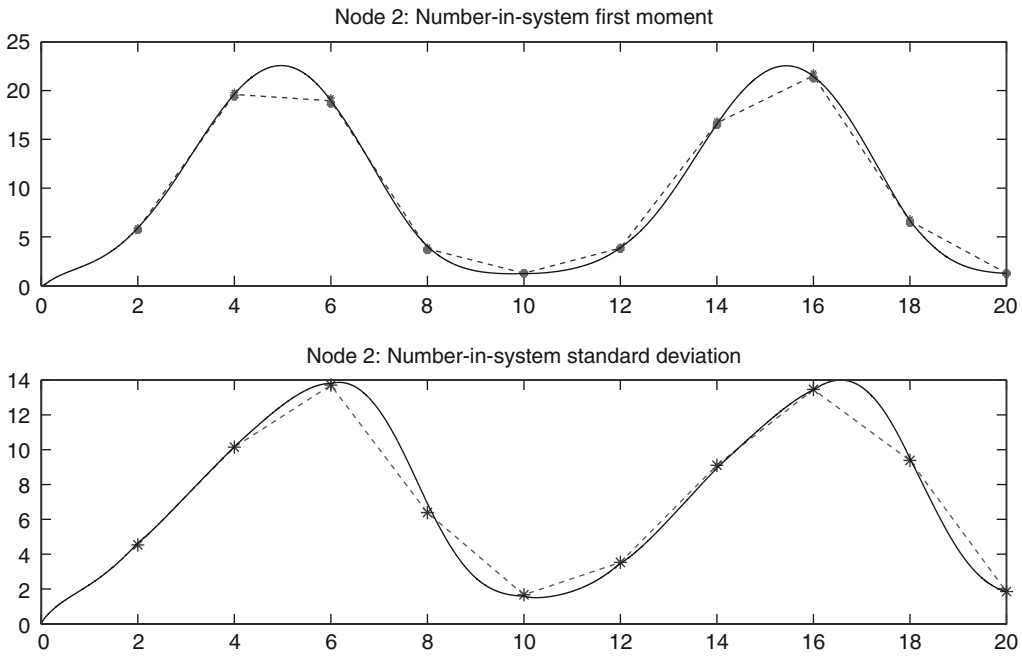

Figure 7 Node 2: Number-in-System First Moment and Standard Deviation—Example 1

Outside Ph_t arrival distribution to node 1:

$$a = \begin{pmatrix} 0 & 0.2 & 0.5 & 0.3 \\ 0.1 & 0 & 0.7 & 0.2 \\ 0.1 & 0.1 & 0 & 0.8 \end{pmatrix}, \quad \begin{array}{l} \alpha = [0.3, 0.4, 0.3], \\ \lambda = [50, 75, 100]. \end{array}$$

The system is analyzed over 40 time units and is divided into four SMIs of length $T = 10$ time units. Each SMI is divided into 10 FMIs of length $\delta = 1$ time units for a total of 40 FMIs. Because of the smaller number of states for the joint number-in-system, we compare the performance of the fitted balanced 2-MECO with solving the joint number-in-system. In Figure 9, the dotted plots represent the approximate first moment and standard deviation of the number-in-system at node 2. The solid plots represent the first moment and standard deviation of the number-in-system at the second node are obtained via solving the joint number-in-system KFEs.

Table 6 shows the number of differential equations solved and the time required to obtain the approximations as well as the joint number-in-system KFEs. The number in phases in the balanced 2-MECO is five where $m_1 = 1$ and $m_2 = 4$.

Table 3 Node 2 First Moment: Approximation and Simulation Results at Specified Points in Time

Time/first moment	2	4	6	8	10	12	14	16	18	20
Simulation	5.81	19.61	18.96	3.79	1.29	3.86	16.68	21.52	6.64	1.28
Approximation	5.93	19.64	18.98	4.05	1.25	3.90	16.55	21.47	6.70	1.27
% difference	−2.01	−0.17	−0.09	−6.98	2.67	−0.98	0.74	0.26	−0.91	0.83

Table 4 Node 2 Standard Deviation: Approximation and Simulation Results at Specified Points in Time

Time/std. dev.	2	4	6	8	10	12	14	16	18	20
Simulation	4.53	10.14	13.71	6.39	1.67	3.53	9.1	13.46	9.39	1.86
Approximation	4.65	10.17	13.81	6.93	1.60	3.47	8.95	13.44	9.52	1.84
% difference	−2.51	−0.25	−0.72	−8.35	4.37	1.65	1.67	0.2	−1.35	0.76

Table 5 Number of PDEs for Different Combinations of Approximation Stages Used—Example 1

Algorithm stages	Number of PDEs	Time to execute (seconds)
MIA	909	4.22
FA time to fit	—	1.06
Node 2 number-in-system KFEs	505	7.44

Table 6 Number of PDEs for Different Combinations of Approximation Stages Used—Example 2

Algorithm stages	Number of PDEs	Time to execute (seconds)
MIA	63	2.34
FA time to fit	—	0.05
Node 2 number-in-system KFEs	255	3.98
Joint number-in-system KFEs	3,213	41.31

Table 7 Node 2 First Moment and Standard Deviation Errors

	Avg. absolute error	Avg. % error	Max. absolute error	Max. % error
First moment	0.15	0.57	0.91	5.46
Std. dev.	0.05	0.32	0.49	2.74

Table 7 shows the average and maximum absolute error over the time interval $[0,40)$. The average percentage error as well as the maximum percentage error are also calculated in Table 7 for the time interval $[2,40)$.

5.5. Comments on the Length of the SMIs and FMIs

The system at node 1 of Example 2 (§5.4) is in steady state during the majority of the time considered as can be seen from the plot of the number-in-system in Figure 8. The fitted balanced 2-MECO has $\alpha = 0.752$ over the first SMI and $\alpha = 0.756$ over the next three SMIs. Fitting the departure process from time 10 to 40 would result in the same balanced 2-MECO, where $\alpha = 0.756$ and $\lambda = 18.64$. For a system in steady state, a recommended condition on the length of the SMI is that the SMI is large enough to assume independence between consecutive SMIs. The lag-1 correlation between the number of departures over the consecutive SMIs $[10, 20)$, $[20, 30)$, and $[30, 40)$ is ≈ 0 and is small enough to assume independence.

Clearly, for a time-dependent departure-count process, smaller SMIs can better capture the time dependency in the second moment departure count. In Example 1 (§5.3), we pair consecutive SMIs and calculate the correlation between departure counts from node 1 over the time intervals, as shown in Table 8.

The average absolute lag-1 correlation over the 10 consecutive intervals is 0.05. The lag-1 correlation between the departure counts over the SMIs is negligible except between intervals 4 and 5 ($[6, 8)$ and $[8, 10)$) as well as intervals 9 and 10 ($[16, 18)$ and $[18, 20)$). A suggested lower limit on the length of the SMI when fitting a time-dependent system is not to allow the average absolute lag-1 correlation to exceed a tolerance level. Here we present the equations for calculating the correlation between consecutive intervals of length Δ. Let $D^{(i)}$ be the number of departures over interval i, then $D^{(i)} = D_{(i-1)\Delta}(\Delta)$. The correlation between the departure-count process of two consecutive intervals $[(i-1)\Delta, i\Delta)$ and $[i\Delta, (i+1)\Delta)$:

$$\mathrm{Corr}(D^{(i)}, D^{(i+1)}) = \frac{\mathrm{E}[D^{(i)}D^{(i+1)}] - \mathrm{E}[D^{(i)}]\mathrm{E}[D^{(i+1)}]}{\sigma_i \sigma_{i+1}}, \quad (16)$$

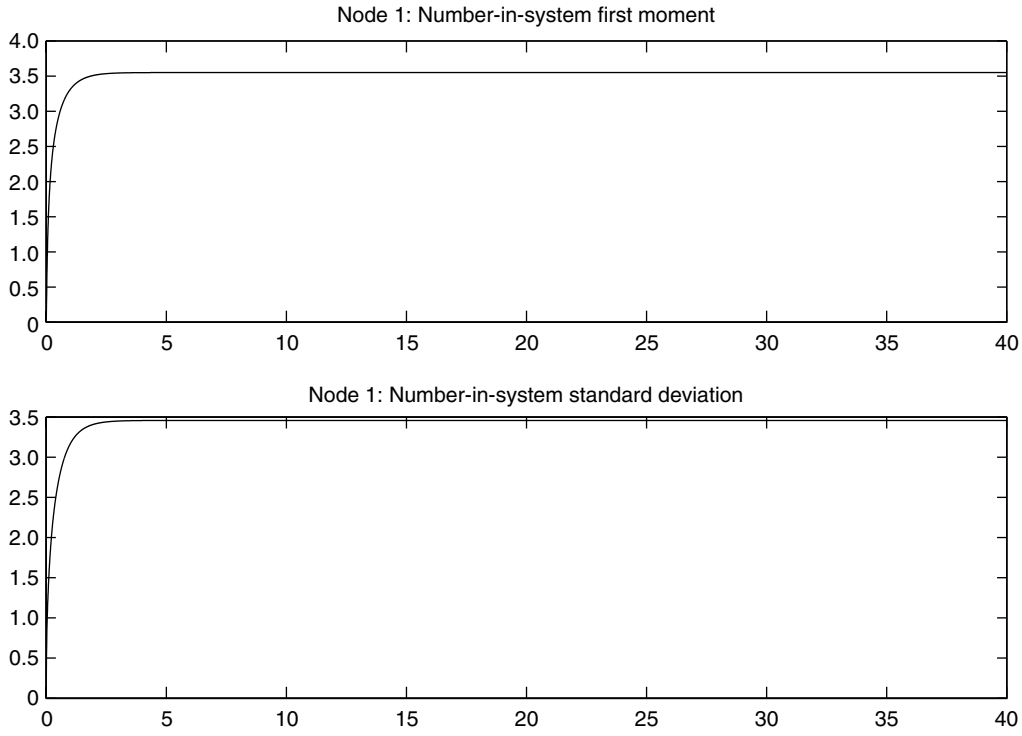

Figure 8 Node 1: Number-in-System First Moment and Standard Deviation—Example 2

Table 8 Correlation Between Departure Counts

Intervals	1, 2	2, 3	3, 4	4, 5	5, 6	6, 7	7, 8	8, 9	9, 10
Correlation	0.03	0.00	0.00	0.16	0.04	0.03	0.01	−0.02	0.15

where σ_i is the standard deviation of $D^{(i)}$ and i is the ith interval. The right-hand side term $\mathrm{E}[D^{(i)}D^{(i+1)}]$ of Equation (16) can be calculated by solving

$$\mathrm{E}[D^{(i)}D^{(i+1)}] = \tfrac{1}{2}\big(\mathrm{E}[(D^{(i)}+D^{(i+1)})^2] - \mathrm{E}[(D^{(i)})^2] - \mathrm{E}[(D^{(i+1)})^2]\big). \quad (17)$$

Notice that solving the NEDMA allows us to calculate the first or second moment departure count over any specified time interval. Equations (16) and (17) can be solved for using the NEDMA.

Choosing a length for an FMI is more straightforward because there is no need to check for correlation. Smaller FMIs better capture the time dependency in the departure rate. The only disadvantage of using small FMIs is that they require more computation when fitting a distribution. In Example 2 of §5.4, larger FMIs can be used without sacrificing accuracy especially over the second, third, and fourth SMIs. In Example 1 of §5.3, smaller FMIs are needed to better capture the change in the departure rate. A suggested measure on whether the length of the FMI is small enough to capture time dependency is

$$\frac{\bar{A}_1}{\bar{A}_2} < \epsilon_2, \quad (18)$$

where $\bar{A}_1 \equiv$ the average departure-count first moment over the FMIs and $\bar{A}_2 \equiv$ is the average of the absolute change in the departure-count first moment between consecutive FMIs. The parameter ϵ_2 represents a tolerance level to approximate the departure rate. A smaller ϵ_2 results in a better fit. Equation (18) is always satisfied for a system in steady state because $\bar{A}_2 = 0$. The ratio of Equation (18) as calculated for Examples 1 and 2 is ≈ 0 for both systems.

5.6. Structure of the Fitted Balanced 2-MECO

We have adjusted the parameters λ and α of the balanced 2-MECO to match key characteristics of the departure process from the $Ph_t/M_t/s/c$ node to obtain an accurate approximation of the time-dependent departure process. As stated previously, with the exception of a few special cases, any point process constructed from the partial set of characteristics is necessarily an approximation. The more characteristics matched, the more accurate the approximate process is.

The structural parameters m_1 and m_2 representing the number of phases for the parallel Erlang distributions can offer more flexibility to the fitting process by matching more characteristics. For example, assume that the fitted distribution is Poisson with rate 1. Over a time interval of length 2 time units, the first and second moment departure counts are two and six, respectively. Any choice of structural parameters, m_1 and m_2, can result in a fit for the first two moments. A balanced 2-MECO with $m_1 = 4$ and $m_2 = 4$, results

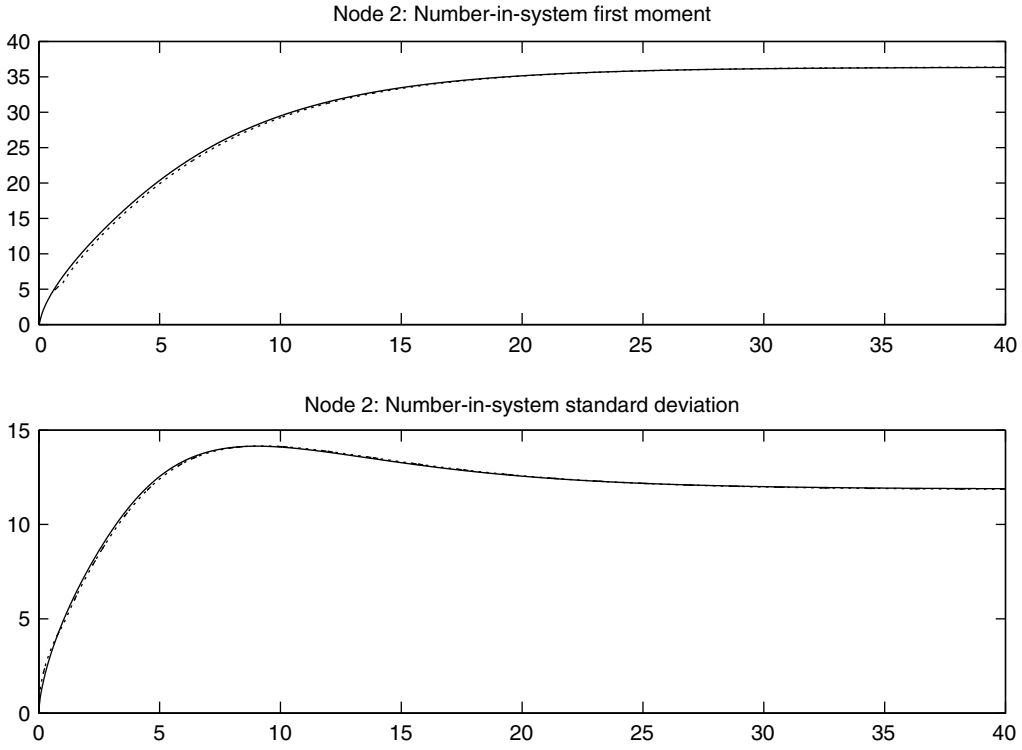

Figure 9 Node 2: Number-in-System First Moment and Standard Deviation—Example 2

in a count process where the first two moments over an interval of length 2 are two and six, respectively, when $\lambda = 0.3368$ and $\alpha = 0.8316$. Obviously the best fit is $m_1 = 1$, $m_2 = 1$, $\lambda = 1$, and $\alpha = 0.5$, which results in a Poisson process. An approach to choose the best combination of m_1 and m_2 is to also fit the first two count moments over an interval of different length. The first two count moments of a Poisson process over an interval of length 1 are one and two, respectively. The best combination that would fit the first two count moments over intervals of length 2 and 1 is when $m_1 = 1$, $m_2 = 1$, $\lambda = 1$, and $\alpha = 0.5$. Although the balanced 2-MECO with $m_1 = m_2 = 4$ matched the first two moments over the interval of length 2, the resulting first second moment over the interval of length 1 is 1.82.

Notice that the FA assumes m_1 and m_2 before calculating the parameters λ and α. An approach for selecting the appropriate structure parameters m_1 and m_2 is to choose two random intervals of different length and find the best fit over the first two moments for different combinations of m_1 and m_2.

Example 5.6. Consider the departure process from a $Ph/M/s/c$ queueing node where the arrival process is Erlang of order 4 and with the following parameters, $s = 1$, $c = 30$.

Outside Ph arrival distribution:

$$a = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{array}{l} \alpha = [1, 0, 0, 0], \\ \lambda = [10, 10, 10, 10]. \end{array}$$

Notice that the traffic intensity is greater than one for $\mu < 2.5$. For the case where $\mu = 1$, the server utilization is ≈ 1 and the probability of an empty queue ≈ 0 (calculated via solving the number-in-system KFEs). In such an overloaded system, the probability of an empty queue is ≈ 0 and the Poisson process provides an accurate approximation for the departure process at steady state. Also for a very large μ, the departure-count process approaches an Erlang order 4. For different values of traffic intensity ρ, the best combination of m_1 and m_2 are shown in Table 9.

The time intervals chosen for Table 9 were of length 4 and 2 at steady state. An approach to choosing m_1 and m_2 after executing the MIA is to randomly select an SMI, $[(j-1)\Delta, j\Delta)$, and the interval $[(j-1)\Delta, j\Delta/2)$, then choose the combination of m_1 and m_2 that provide the best fit over the two intervals.

Table 9 Structure of the Fitted Balanced 2-MECO

ρ	μ	m_1	m_2	λ	α
> 1	1	1	1	1	0.500
0.8	3.125	1	5	1.805	0.639
0.5	5	3	5	1.920	0.616
0.01	250	4	4	2.490	0.502

6. Conclusion

An advantage of considering departure-count processes is that computing the count moments of the superposed independent departure-count processes is straightforward; as would be needed in considering larger networks. For example, let D_1 and D_2 be the departure counts of two independent processes over an interval. Let D be the superposition of the two processes ($D = D_1 + D_2$):

$$\mathrm{E}[D] = \mathrm{E}[D_1] + \mathrm{E}[D_2], \quad \text{and}$$
$$\mathrm{E}[D^2] = \mathrm{E}[D_1^2] + \mathrm{E}[D_2^2] + 2\mathrm{E}[D_1]\mathrm{E}[D_2].$$

Another advantage of considering departure-count processes is that computing moments of the Markov-routed departure-count processes is relatively straightforward as well (as would be needed in larger, more general networks). Consider a process where the arrivals/departures are routed to node i with probability p. Let A be the number of arrivals over a time interval and let A_i be the number of arrivals to nodes i. The first two moments of A_i,

$$\mathrm{E}[A_i] = p\mathrm{E}[A], \quad \text{and}$$
$$\mathrm{E}[A_i^2] = \sum_{n=1}^{\infty} np(np - p + 1)\mathrm{P}(A = n)$$
$$= p^2\mathrm{E}[A^2] - (1-p)p\mathrm{E}[A].$$

The algorithms presented in this paper could be applied on several queues in tandem or on a network. If the network has a tree structure then it is possible to solve each node independently, otherwise each node can be analyzed separately across the SMIs. In this work we lay out algorithmic approaches to calculating key characteristics of the departure-count process and fitting a probability distribution. As stated earlier, this work is part of a larger goal of constructing a QNATS. Future work that builds on what is presented in this paper can take on two directions: (1) Enabling network decomposition to become even more efficient by taking advantage of previously developed closure-type approximation algorithms for the $Ph_t/M_t/s/c$ number-in-system moments (Ong and Taaffe 1987); and develop departure-count moment approximations that are similar in structure, accuracy, and efficiency to number-in-system moment closure approximations. (2) Applying the distribution fitting approach (properties and algorithm) of §5 to distributions with even more parameters/flexibility than a 2-MECO (n-MECO or MAP) to capture more key characteristics.

References

Avramidis AN, Deslauriers A, L'Ecuyer P (2004) Modeling daily arrivals to a telephone call center. *Management Sci.* 50(7): 896–908.

Bekker R, de Bruin AM (2010) Time-dependent analysis for refused admissions in clinical wards. *Ann. Oper. Res.* 178:45–65.

Gerhardt I, Nelson BL (2009) Transforming renewal processes for simulation of nonstationary arrival processes. *INFORMS J. Comput.* 21(4):630–640.

Gerhardt I, Nelson BL (2010) On capturing dependence in point processes: Matching moments and other techniques. Working paper, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL. http://users.iems.northwestern.edu/~nelsonb/Publications/GerhardtNelsonSurvey.pdf.

Green LV, Kolesar PJ, Whitt W (2007) Coping with time-varying demand when setting staffing requirements for a service system. *Production Oper. Management* 16(1):13–39.

Gusella R (1991) Characterizing the variability of arrival processes in indexes of dispersion. *IEEE J. Selected Areas Comm.* 9(2): 203–211.

Harrod S, Kelton WD (2006) Numerical methods for realizing nonstationary Poisson processes with piecewise-constant instantaneous-rate functions. *Simulation* 82(3):147–157.

Jennings OB, Mandelbaum A, Massey WA, Whitt W (1996) Server staffing to meet time-varying demand. *Management Sci.* 42(10): 1383–1394.

Johnson MA, Taaffe MR (1988) The denseness of phase distributions. Working paper, School of Industrial Engineering, Purdue University, West Lafayette, IN.

Johnson MA, Taaffe MR (1989) Matching moments to phase distributions: Mixture of Erlang distributions of common order. *Commun. Statis.-Stochastic Models* 5(4):711–743.

Johnson MA, Taaffe MR (1990) Matching moments to phase distributions: Nonlinear programming approaches. *Stochastic Models* 6(2):259–281.

Kaczynski WH, Leemis LM, Drew JH (2011) Transient queueing analysis. *INFORMS J. Comput.* 24(1):10–28.

Knessl C (2002) An exact solution for an $M(t)/M(t)/1$ queue with time-dependent arrivals and service. *Queueing Systems* 40(3):233–245.

Koopman BO (1972) Air-terminal queues under time-dependent conditions. *Oper. Res.* 20(6):1089–1114.

Margolius BH (1999) A sample path analysis of the $M_t/M_t/c$ queue. *Queueing Systems* 31(59):59–93.

Montenegro G, Sengoku M (1992) Time-dependent analysis of mobile communication traffic in a ring-shaped service area with nonuniform vehicle distribution. *IEEE Trans. Vehicular Tech.* 41(3):243–254.

Nelson BL, Taaffe MR (2004) The $Ph_t/Ph_t/\infty$ queueing system: Part I—The single node. *INFORMS J. Comput.* 16(3):266–274.

Odoni AR, Roth E (1983) An empirical investigation of the transient behavior of stationary queueing systems. *Oper. Res.* 31(3): 432–455.

Ong KL, Taaffe MR (1987) Approximating nonstationary $Ph_t/M_t/s/c$ queueing systems. *Ann. Oper. Res.* 8:103–116.

Peterson MD, Bertsimas DJ, Odoni AR (1995) Models and algorithms for transient queueing congestion at airports. *Management Sci.* 41(8):1279–1295.

Rueda JE, Taaffe MR (2004) The $Ph_t/Ph_t/s/c$ queueing model and approximation. Technical report, The Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA.

Telek M, Horvath G (2007) A minimal representation of Markov arrival processes and a moments matching method. *Performance Eval.* 64(9–12):1153–1168.

Whitt W (1982) Approximating a point process by a renewal process, I: Two basic methods. *Oper. Res.* 30(1):125–147.

Whitt W (1983) The queueing network analyzer. *Bell System Technical J.* 62(9):2779–2815.