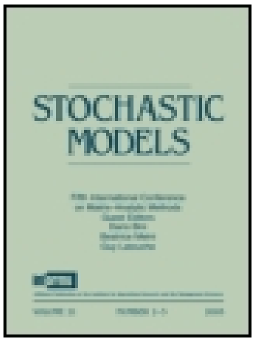

Communications in Statistics. Stochastic Models

ISSN: 0882-0287 (Print) (Online) Journal homepage: http://www.tandfonline.com/loi/lstm19

Markov meco: a simple markovian model for approximating nonrenewal arrival processes

Mary A. Johnson

COMMUN. STATIST.—STOCHASTIC MODELS, 14(1&2), 419–442 (1998)

MARKOV MECO: A SIMPLE MARKOVIAN MODEL FOR APPROXIMATING NONRENEWAL ARRIVAL PROCESSES

Mary A. JOHNSON

Mechanical and Industrial Engineering
University of Illinois at Urbana-Champaign
Urbana, Illinois

ABSTRACT

This paper introduces a special case of the Markovian Arrival Process that can easily be used to approximate both the interarrival-time distribution and the autocorrelation function of an arrival process. This model is labeled a *Markov MECO* (**M**ixture of **E**rlangs of **C**ommon **O**rder). The proposed interarrival-time approximation matches the first three moments of the interarrival time. The Markov MECO autocorrelation function is geometric and determined by a single parameter, once the interarrival-time distribution is fixed. Closed-form expressions are given for matching either the lag-one autocorrelation or the asymptotic index of dispersion for intervals. By applying the resulting arrival-process approximations to a single-server queue with exponential service times, the approximations are empirically evaluated on the basis of the associated error in the steady-state mean number in the system. In these experiments, the original arrival processes are superpositions of independent renewal processes. The numerical results show that the nonrenewal approximations outperform the corresponding renewal approximation that simply ignores autocorrelation, and in some cases, the improvement is an order of magnitude.

1. INTRODUCTION

The behavior of the arrival process of a queueing system has an important effect on the performance of the queueing system. Thus, the modeling of arrival processes is an important component of the problem of developing accurate queueing models. The Poisson assumption so often used for tractability is in many cases simply inadequate for obtaining accurate queueing models of manufacturing systems, computer and communications systems, and the like. For example, the traffic in many communications systems is far more "bursty" than a Poisson process. Also, the traffic in many manufacturing systems is more regular than a Poisson process. The generalization to a renewal arrival process, for which the interarrival times are independent but not necessarily exponential, helps but still does not allow for explicit modeling of dependence among interarrival times. The prevalence of such dependence becomes clear by recalling that the traffic in most queueing networks is nonrenewal. Departure processes are typically nonrenewal (the M/M/s queue in steady state being the obvious exception), and superpositions of independent renewal processes are renewal only if the superposed processes are Poisson. Other evidence of the pervasiveness of nonrenewal arrival processes includes the positive correlation observed in communications-system traffic and in vehicular traffic.

The prevalence of nonrenewal arrival processes in a variety of applications implies that tractable arrival-process models which are flexible both in terms of the interarrival-time distribution and the dependence structure among interarrival times are desirable. In fact, such models are available and widely applied in the queueing literature. Probably the most common example is the Markovian Arrival Process (MAP). The MAP can be interpreted as a generalization of the renewal process with phase-type (PH) interarrival times, and the family of PH probability distributions can be interpreted as a generalization of the exponential distribution. The Markovian assumptions underlying phase-type distributions and MAP's, along with their convenient matrix representations, have led to many tractable queueing models whose solutions are based on matrix-analytic methods. The family of MAP's is a subset of Markov renewal processes, for which some queueing models and associated solutions are also available; see Disney and Kiessler [7]. These arrival-process models are described in greater detail in Section 2 of this paper.

Despite the availability of general, tractable arrival models, our knowledge of what arrival-process descriptors are most important in developing approximations is still limited. In addition to descriptors of the interarrival-time distribution, the most common descriptors of a stationary arrival process are second-order descriptors consisting of or based on the variance of the number of arrivals by time t, $t > 0$, or the variance of the arrival-time of the nth customer, $n = 1, 2, 3, \ldots$. See Cox and Lewis [6] for an overview of such descriptors. Various empirical studies demonstrate the potentially dramatic adverse effect of positive autocorrelation in the arrival process; see, for example, Alfa and Neuts [2], Latouche [17], Livny, Melamed, and Tsiolis [19], and Patuwo, McNickle, and Disney [24]. The concept of burstiness is the basis for alternative descriptors, such as those proposed in Neuts [21], Johnson, Liu, and Narayana [12] and Johnson and Narayana [13].

Another impediment to the full use of general models such as the MAP is the lack of methodology for fitting these models. Some work on this problem has been done, but most of the solutions use only a simple, restrictive arrival-process model. For fitting data, Meier-Hellstern [20] and Rydén [25] developed maximum-likelihood approaches that use a subset of two-state MAP's known as two-state Markov Modulated Poisson processes (MMPP's). Rydén [26] also extended his data-fitting methods to larger MMPP's. Whitt [29] and Albin [1] developed a renewal approximation that adjusts the interarrival-time variance to account for the dependence among the interarrival times of the approximated process. Gusella [9], Heffes [10], and Heffes and Lucantoni [11] used two-state MMPP's to match various properties of the original arrival process. Sitaraman [27] developed an interactive approach that accommodates larger MAP's and a variety of fitting criteria.

Perhaps the approximation approach most closely related to that proposed in this paper is the one proposed in Bitran and Dasu [4]. Bitran and Dasu used a rather complicated numerical procedure to match the first two moments of the interarrival time and a measure of the cumulative effect of the interarrival-time autocorrelation (c_∞^2, as defined in Section 3.1) to a special class of MAP's. The second nonrenewal approximation proposed in this paper is similar but it also matches the third moment of the interarrival time. A difficulty with the method of Bitran and Dasu is that the analysis does not lead to a precise statement of what combinations of descriptor values are feasible for the selected class of MAP's. Also, although the method seems to work very well for approximating superpositions of renewal processes that might be typical of manufacturing systems, it does not work for superpositions of more variable renewal processes. The approximations proposed here are much simpler conceptually and computationally and apply to a wider range of arrival processes.

The primary contribution of this paper is that it extends Johnson and Taaffe's [16] three-moment-matching method to simple arrival-process approximations which crudely approximate the autocorrelation function. These approximations are applied to various superpositions of renewal processes and evaluated on the basis of resulting queueing congestion-measure approximations. The empirical results indicate that the proposed approximations lead to more accurate congestion-measure approximations than a three-moment renewal approximation that ignores autocorrelation. For the new arrival-process approximations, the approximate congestion measure is often within a few percent of the actual measure.

The second contribution of this paper is that its numerical examples give further evidence of the usefulness of approximating the autocorrelation function of the arrival process of a queueing system. The significance of this is that researchers frequently express doubts about the usefulness of the autocorrelation function as an arrival-process descriptor. Conversely, this paper in no way proves that the autocorrelation function is *always* valuable. In fact, statistical analyses of traffic in high-speed communications systems have shown so much variability and dependence that the measure c_∞^2 is inadequate, as it diverges to infinity. See, for example, Beran, Sherman, Taqqu, and Willinger [3], Leland, Taqqu, Willinger, and Wilson [18] and Garrett and Willinger [8].

A third contribution of this paper is that the numerical examples provide insight into the autocorrelation structure of superpositions of independent renewal processes. In particular, for the superpositions considered in this paper, the autocorrelations are moderate. In contrast, examples used to illustrate the substantial effect of autocorrelation on queueing-system performance typically involve more autocorrelation and hence may not be representative of the autocorrelation found in actual systems.

The rest of this paper is organized as follows. Section 2 briefly reviews the fundamentals of PH distributions and the arrival-process models used in this paper. Section 3 introduces the Markov MECO model and shows how it can be used as a basis for two nonrenewal arrival-process approximations. Section 4 presents an empirical investigation into the usefulness of the proposed approximations. Section 5 summarizes the results presented and makes some concluding remarks.

2. DESCRIPTION OF SOME MARKOVIAN MODELS

This section briefly introduces (continuous) PH distributions, (continuous-time) MAP's, and Markov renewal processes.

A PH distribution of dimension m can be interpreted as the absorption time of a continuous-time Markov chain with m transient states and one absorbing state, given a distribution of the initial state. The representation of a PH distribution is $(\alpha, \mathbf{T})$, where row vector α consists of the initial-state probabilities assigned to the m transient states and $m \times m$ matrix $\mathbf{T}$ consists of the transition rates among transient states. The representation of a PH distribution is not unique. A basic reference on PH distributions is Neuts [23].

A MAP of dimension m is represented by $m \times m$ matrices $\mathbf{D}_0$ and $\mathbf{D}_1$, which have the following interpretations. Let $\mathbf{D} = \mathbf{D}_0 + \mathbf{D}_1$. Underlying the MAP is a continuous-time Markov chain with generator $\mathbf{D}$. When the chain is in state i, the probability of an arrival upon a transition to state $j \neq i$ is $(\mathbf{D}_1)_{ij}/(\mathbf{D})_{ij}$, where for any matrix $\mathbf{M}$, $(\mathbf{M})_{ij}$ denotes its (i,j) entry. The probability of no arrival upon a transition to state $j \neq i$ is $(\mathbf{D}_0)_{ij}/(\mathbf{D})_{ij}$ While in state i, arrivals occur according to a Poisson process with rate $(\mathbf{D}_1)_{ii}$. Like PH distributions, MAP's do not have unique representations. Let θ be the stationary probability distribution for $\mathbf{D}$, and throughout this paper let $\mathbf{e}$ represent a column vector of ones with appropriate dimension. Then the interarrival-time of a stationary MAP with representation $(\mathbf{D}_0, \mathbf{D}_1)$ has PH distribution with representation $(\theta_{arr}, \mathbf{D}_0)$, where $\theta_{arr} = (\theta\mathbf{D}_1) / (\theta\mathbf{D}_1\mathbf{e})$.

The superposition of two independent MAP's is also an MAP. Let $\mathbf{I}(m)$ denote the $m \times m$ identity matrix and $\otimes$ the Kronecker product. For independent MAP's with representations $(\mathbf{D}_0(1), \mathbf{D}_1(1))$ and $(\mathbf{D}_0(2), \mathbf{D}_1(2))$ and dimensions m_1 and m_2, respectively, the superposition of these processes has representation $(\mathbf{D}_0, \mathbf{D}_1)$, where

$$\mathbf{D}_i = \mathbf{D}_i(1) \otimes \mathbf{I}(m_2) + \mathbf{I}(m_1) \otimes \mathbf{D}_i(2), \quad (1)$$

Neuts [22]. Observe that this representation of the superposition has dimension

$m_1 m_2$. As additional arrival processes are superposed, the dimension of the representation obtained via (1) quickly becomes computationally prohibitive. In some cases, a more compact representation is available and mitigates the dimensionality problem. In general, the high dimension of the representation of superposed MAP's leads to the need for approximations of the superposition by a more computationally manageable arrival process.

The PH renewal process with representation $(\alpha, \mathbf{T})$ is a renewal process whose interarrival times are PH distributed with representation $(\alpha, \mathbf{T})$. Expressed as a MAP, this process has representation $\mathbf{D}_0 = \mathbf{T}$ and $\mathbf{D}_1 = -\mathbf{T}\mathbf{e}\alpha$.

A Markov renewal arrival process (MRP) has an underlying discrete-time Markov chain. This chain makes a transition upon each arrival. Probability p_{ij} is the probability that the chain will move to state j upon the next arrival, given that the chain is currently in state i. For a transition from i to j, the interarrival time depends on i and j but is otherwise independent of other interarrival times. Let F_{ij} denote the cumulative distribution function (cdf) of the interarrival-time between the nth and (n+1)st arrivals, conditional upon the chain being in state i just after the nth arrival and in state j just after the (n+1)st arrival. A MRP is often represented by its kernel, which is determined by both the transition probabilities of the underlying Markov chain and the conditional interarrival-time cdf's:

$$\mathrm{F}(x) = \begin{bmatrix} p_{11}\,\mathrm{F}_{11}(x) & \cdots & p_{1m}\,\mathrm{F}_{1m}(x) \\ \vdots & \ddots & \vdots \\ p_{m1}\,\mathrm{F}_{m1}(x) & \cdots & p_{mm}\,\mathrm{F}_{mm}(x) \end{bmatrix}.$$

Thus, $p_{ij}\,\mathrm{F}_{ij}(x)$ is the probability that an interarrival-time is no greater than x and that the chain moves to state j upon the next arrival, given that the chain is currently in state i. The MAP with representation $(\mathbf{D}_0, \mathbf{D}_1)$ is a MRP with kernel $\mathbf{H}(x) = [\exp(\mathbf{D}_0 x) - \mathbf{I}]\mathbf{D}_0^{-1}\mathbf{D}_1$.

3. MARKOV MECO: A CONVENIENT ARRIVAL-PROCESS MODEL

This section introduces the *Markov MECO* arrival-process model, determines its autocorrelation structure, shows how to match two descriptors of the autocorrelation function to a Markov MECO that also matches the first three moments of the stationary-interarrival-time distribution, and points out bounds on feasible values of those descriptors. Note that the focus is on describing and approximating a covariance stationary sequence of interarrival times. Thus, if T_1, T_2, T_3,... denotes the sequence of interarrival times of an arrival process, we assume that the mean and variance of T_n are independent of n and that the covariance of T_n and T_{n+j} is independent n. If we interpret time 0 as an arrival epoch and denote the time of the nth arrival after time 0 by S_n, then $S_n = \sum_{j=1}^{n} T_j$. Whether $\{T_n\}$ and $\{S_n\}$ refer to the original arrival process or the approximating Markov MECO is implied by the context.

3.1 Development and Definition of Markov MECO Model

The approximation presented here is an extension of a simple distribution approximation based on PH distributions to an arrival-process approximation based on MAP's. Thus, the distribution approximation is introduced first. For the distribution approximation, we simply match the first three moments of the distribution to a **M**ixture of two **E**rlangs of **C**ommon **O**rder (MECO). See [16] for details. This approximation is attractive for queueing applications, because matching three moments often leads to good approximations of GI/G/1 queues (see Johnson and Taaffe [14, 15] and references therein), and because it has a simple analytical solution. Also, for a given set of first three moments, the minimum feasible order of the mixed Erlangs is easily computed from the closed-form expressions shown below. Let c denote the coefficient of variation (standard deviation divided by the mean) and γ the coefficient of skewness (third central moment divided by the cube of the standard deviation). These are normalized versions of the second and third moments, respectively. For fixed c and γ, the minimum feasible order of the mixed Erlangs is the minimum n that satisfies

$$n > 1/c^2 \tag{2}$$

and

$$n > \frac{-\gamma + 1/c^3 + 1/c + 2c}{\gamma - (c - 1/c)}. \tag{3}$$

Here we assume that the mixed Erlangs are distinct and that each has positive probability, i.e., we do not allow the mixture to be an Erlang. Expressions (2) and (3) are derived from the following bounds on the feasible values of c and γ for a given order n of the mixed Erlangs.

$$c > 1/\sqrt{n} \tag{4}$$

$$\gamma > \frac{1}{1+n}\left(\frac{1}{c^3} + (1-n)\frac{1}{c} + (2+n)c\right) \tag{5}$$

For most common distributions with $c > 1$, mixing Erlangs of order one, i.e., exponentials, is feasible; see Figure 1 in [15]. As c or γ decreases, the minimum feasible order increases.

To extend the three-moment MECO distribution approximation to an arrival-process approximation we proceed as follows.

Step 1: We begin with a three-moment MECO approximation of the interarrival-time of the original process. Let the mixing probabilities be $p_1 > 0$ and $p_2 = 1 - p_1 > 0$ and the mixed Erlang cdf's be F_1 and F_2, respectively. Then the cdf of this approximation is

$$\mathrm{F}(x) = p_1 \mathrm{F}_1(x) + p_2 \mathrm{F}_2(x). \tag{6}$$

Step 2: We approximate the autocorrelation function of the original arrival process by using a MRP with kernel

$$\mathbf{F}(x) = \begin{bmatrix} p_{11} \mathrm{F}_1(x) & p_{12} \mathrm{F}_1(x) \\ p_{21} \mathrm{F}_2(x) & p_{22} \mathrm{F}_2(x) \end{bmatrix}, \tag{7}$$

where the conditional probabilities are chosen such that

$$p_i = p_1 p_{1i} + p_2 p_{2i}, \quad i = 1, 2, \tag{8}$$

so that the distribution of the (marginal) interarrival time of the approximation is still (6).

Beyond satisfying (8), the issue of exactly how to select the conditional probabilities to approximate the autocorrelation function of the original arrival process is addressed at length below. Because the distribution of the time between arrivals in this MRP is a MECO, we label the model a ***Markov MECO.*** Although the author devised and analyzed this model independently, it also appears in [24].

Since F_1 and F_2 are Erlang, the Markov MECO is also a MAP, with the dimension of the MAP being twice the order of the mixed Erlangs. For example, if Erlangs of order two and rate parameters r_1 and r_2 are mixed, the MAP representation of this arrival process is

$$\mathbf{D}_0 = \begin{bmatrix} -r_1 & r_1 & 0 & 0 \\ 0 & -r_1 & 0 & 0 \\ 0 & 0 & -r_2 & r_2 \\ 0 & 0 & 0 & -r_2 \end{bmatrix}, \ \mathbf{D}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ r_1 p_{11} & 0 & r_1 p_{12} & 0 \\ 0 & 0 & 0 & 0 \\ r_2 p_{21} & 0 & r_2 p_{22} & 0 \end{bmatrix}.$$

For any Markov MECO, a renewal-process is obtained by setting $p_{11} = p_{21} = p_1$ and $p_{12} = p_{22} = p_2$. Intuitively, one can see that when $p_{11} > p_1$ and $p_{22} > p_2$, the interarrival times are positively correlated, and when $p_{11} < p_1$ and $p_{22} < p_2$, successive interarrival times are negatively correlated. Other relationships among the probabilities are infeasible. That is, if $p_{11} > p_1$ and $p_{22} < p_2$, or if $p_{11} < p_1$ and $p_{22} > p_2$, then (8) cannot be satisfied.

The primary limitation of the Markov MECO model is that requirement (8) and the identities

$$p_2 = 1 - p_1,\ p_{11} = 1 - p_{12},\ \text{and } p_{22} = 1 - p_{21}, \tag{9}$$

leave just a single degree of freedom in the choice of the conditional probabilities and hence the autocorrelation function. That is, once one conditional probability, say p_{21}, is chosen, the rest follow from (8) and (9). Expressed as a function of p_1 and p_{21}, p_{12} is

$$p_{12} = (1 - p_1)p_{21} / p_1. \tag{10}$$

The remaining probabilities now follow immediately from (9). Also, equation (10) implies an upper bound on feasible values of p_{21}. Since $0 < p_{12} < 1$,

$$p_{21} < \frac{p_1}{1 - p_1}, \tag{11}$$

which is less than one for $p_1 > 0.5$.

At this point we have reduced the problem of selecting the values of the conditional probabilities to the problem of selecting the value of p_{21}. Let $\{\rho_j, j = 1, 2, 3\ldots\}$ denote the autocorrelation function of the original arrival process. How should p_{21} be chosen to approximate $\{\rho_j, j = 1, 2, 3\ldots\}$? One approach to be developed in the following subsections is to match just ρ_1. The second approach to be developed is to match c_∞^2, the limiting value of the index of dispersion for intervals (IDI), defined as

$$c_\infty^2 = \lim_{n\to\infty} \frac{n\,\mathrm{Var}(S_n)}{[\mathrm{E}(S_n)]^2} = \lim_{n\to\infty} \frac{\mathrm{Var}(S_n)}{n[\mathrm{E}(T_1)]^2}. \tag{12}$$

As explained in Sriram and Whitt [28, page 834], this descriptor, along with the mean interarrival time, completely characterizes the arrival process for multi-server, first-come-first-served heavy-traffic queueing systems. Moreover, for given interarrival-time mean and variance, c_∞^2 is a summary descriptor of the autocorrelation function, since

$$\frac{\mathrm{Var}(S_n)}{n} = \mathrm{Var}(T_1)\left[1 + 2\sum_{j=1}^{n-1}(1 - j/n)\rho_j\right]. \tag{13}$$

We caution that for the traffic considered in [3, 8, 18], the autocorrelation function falls off so slowly that c_∞^2 does not converge to a finite value. Thus, the approximations proposed here are not appropriate for such applications.

3.2 Autocorrelation Function of a Markov MECO

In this subsection we completely characterize the behavior of the autocorrelation function of the Markov MECO. This characterization becomes apparent upon pursuing the MRP interpretation of the Markov MECO.

Notice that the embedded Markov chain corresponding to (7) has transition matrix

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}. \tag{14}$$

Let $\mu_k(j)$ be the kth noncentral moment of F_j. Then the kth noncentral moment of the interarrival time of the Markov MECO is

$$\mathrm{E}(T_n^k) = p_1\mu_k(1) + p_2\mu_k(2), \tag{15}$$

and

$$\mathrm{E}(T_n T_{n+j}) = p_1 p_{11}^{(j)} \mu_1^2(1) + p_1 p_{12}^{(j)} \mu_1(1)\mu_1(2) + p_2 p_{21}^{(j)} \mu_1(1)\mu_1(2) + p_2 p_{22}^{(j)} \mu_1^2(2) \tag{16}$$

where $p_{il}^{(j)}$ denotes the j-step transition probability for the embedded Markov chain (14). Since the lag-j autocorrelation of the interarrival times is $\rho_j = \mathrm{Cov}(T_n, T_{n+j}) / \mathrm{Var}(T_n)$, the behavior of the autocorrelation function can be determined by examining the behavior of $\mathrm{Cov}(T_n, T_{n+j})$ as a function of j. To obtain a convenient expression for $\mathrm{Cov}(T_n, T_{n+j})$, it will be helpful to express (15) and (16) in terms of matrices. To this end, we define $\mathbf{p} = (p_1, p_2)$ and $\mathbf{M}_k = \begin{bmatrix} \mu_k(1) & 0 \\ 0 & \mu_k(2) \end{bmatrix}$. Now (15) and (16) can be rewritten as follows.

$$\mathrm{E}(T_n^k) = \mathbf{p}\mathbf{M}_k\mathbf{e} \tag{17}$$

$$\mathrm{E}(T_n T_{n+j}) = \mathbf{p}\mathbf{M}_1\mathbf{P}^j\mathbf{M}_1\mathbf{e} \tag{18}$$

Thus,

$$\begin{aligned} \mathrm{Cov}(T_n T_{n+j}) &= \mathrm{E}(T_n T_{n+j}) - \mathrm{E}(T_n)\mathrm{E}(T_{n+j}) \\ &= \mathbf{p}\mathbf{M}_1\mathbf{P}^j\mathbf{M}_1\mathbf{e} - (\mathbf{p}\mathbf{M}_1\mathbf{e})^2 \\ &= \mathbf{p}\mathbf{M}_1(\mathbf{P}^j - \mathbf{e}\mathbf{p})\mathbf{M}_1\mathbf{e}. \end{aligned} \tag{19}$$

To further characterize $\mathrm{Cov}(T_n, T_{n+j})$, we express $\mathbf{P}$ in terms of its spectral representation. See the appendices of Çinlar [5] for a convenient introduction to spectral representations of Markov Matrices. Matrix $\mathbf{P}$ as defined by (14) has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 1 - (p_{21} / p_1)$, whose absolute value is less than 1. (Since we require p_1 and p_2 to be strictly less than one, the conditional probabilities in $\mathbf{P}$ must be strictly less than one, implying that $\mathbf{P}$ cannot be aperiodic and hence $|\lambda_2| < 1$.) Moreover, the associated spectral representation is $\mathbf{P} = \mathbf{e}\mathbf{p} + \lambda_2\mathbf{A}_2$, where $\mathbf{A}_2 = \left(\frac{-p_{21}}{p_1}\right)\begin{bmatrix} -(1-p_1)p_{21}/p_1 & (1-p_1)p_{21}/p_1 \\ p_{21} & -p_{21} \end{bmatrix}$. Note that

$\mathbf{ep} = \lim_{j\to\infty} \mathbf{P}^j$. Since $\mathbf{P}^j = \mathbf{ep} + \lambda_2^j \mathbf{A}_2$, Theorem 1 below now follows from (19).

Theorem 1. For a Markov MECO,

$$\operatorname{Cov}(T_n, T_{n+j}) = \lambda_2^j (\mathbf{pM}_1\mathbf{A}_2\mathbf{M}_1\mathbf{e}), \tag{20}$$

which can also be expressed as

$$\operatorname{Cov}(T_n, T_{n+j}) = (1 - p_{21}/p_1)^j p_1(1-p_1)(\mu_1(1) - \mu_1(2))^2. \tag{21}$$

From Theorem 1 we can see that $\{\operatorname{Cov}(T_n T_{n+j}), j = 1, 2, 3, \ldots\}$ and hence the autocorrelation function of the Markov MECO are geometric with decay parameter $\lambda_2 = 1 - (p_{21}/p_1)$. When $p_{21} = p_1$, $\lambda_2 = 0$; when $p_{21} < p_1$, $\lambda_2 > 0$; and when $p_{21} > p_1$ $\lambda_2 < 0$. The assertions about the behavior of the autocorrelation function as a function of p_{11} and p_{22}, stated in Section 3.1 on intuitive grounds, are equivalent to these relationships. Theorem 1 also implies that for fixed p_1 and p_{21}, the magnitude of $\operatorname{Cov}(T_n, T_{n+j})$ and ρ_j increases as the difference between the means of the mixed Erlangs increases. Moreover, if the means were allowed to be equal, then $\operatorname{Cov}(T_n, T_{n+j})$ and ρ_j would be zero for all $j \geq 1$.

Theorem 2. For a Markov MECO,

$$c_\infty^2 = \frac{\operatorname{Var}(T_1) + 2\left(\dfrac{p_1 - p_{21}}{p_{21}}\right) p_1(1-p_1)\left(\mu\ (1) - \mu_1(2)\right)^2}{\left[\mathrm{E}(T_1)\right]^2}. \tag{22}$$

Proof: Equations (12) and (13) imply

$$c_\infty^2 = \left[\operatorname{Var}(T_1)\left(1 + 2\lim_{n\to\infty}\sum_{j=1}^{n-1}(1 - j/n)\rho_j\right)\right] \Big/ \left[E(T_1)\right]^2 \tag{23}$$

Since $\rho_j = \operatorname{Cov}(T_n, T_{n+j}) / \operatorname{Var}(T_1)$ and $|\lambda_2| \leq 1$, (20) and (23) imply

$$c_\infty^2 = \left[\operatorname{Var}(T_1) + 2\left(\frac{\lambda_2}{1-\lambda_2}\right)\mathbf{pM}_1\mathbf{A}_2\mathbf{M}_1\mathbf{e}\right] \Big/ \left[\mathrm{E}(T_1)\right]^2.$$

Equation (22) now follows easily.

3.3 Matching ρ_1 and c_∞^2

Now that we are able to obtain ρ_1 and c_∞^2 from the parameters of the Markov MECO, the next task is to express p_{21} in terms of either ρ_1 or c_∞^2 (or their equivalents) and parameters p_1, $\mu_1(1)$, and $\mu_1(2)$ (determined in Step 1 by matching the first three moments of the interarrival time).

Matching ρ_1. Since the interarrival time distribution of the Markov MECO approximation matches the variance of the original interarrival-time distribution, matching ρ_1 is equivalent to matching $\mathrm{Cov}(T_n, T_{n+1})$. Setting j=1 in equation (21) and solving for p_{21} yields

$$p_{21} = p_1 - \frac{\mathrm{Cov}(T_n, T_{n+1})}{(1-p_1)(\mu_1(1)-\mu_1(2))^2}. \tag{24}$$

Matching c_∞^2. To obtain the value of p_{21} that matches c_∞^2, we solve equation (22) for p_{21}. The result is

$$p_{21} = \frac{p_1}{1+\dfrac{[\mathrm{E}(T_1)]^2 c_\infty^2 - \mathrm{Var}(T_1)}{2p_1(1-p_1)(\mu_1(1)-\mu_1(2))^2}}. \tag{25}$$

3.4 Bounds on Feasible Values of ρ_1 and c_∞^2

Recall from Section 3.1 that p_{21} must be in the interval $(0, \min\{p_1/(1-p_1), 1\})$. Suppose p_1, $\mu_1(1)$, and $\mu_1(2)$ are fixed. Then as $p_{21} \to 0$, $\lambda_2 \to 1$ and $c_\infty^2 \to \infty$; as $p_{21} \to \min\{p_1/(1-p_1), 1\}$, ρ_1 and c_∞^2 approach finite lower bounds. In practice, the author restricts p_{21} to the interval $[\varepsilon, \min\{p_1/(1-p_1), 1\} - \varepsilon]$, where $\varepsilon = 0.001$. This restriction imposes finite upper and lower bounds on ρ_1 and c_∞^2. In the experimentation reported in Section 4, these lower bounds frequently prevent exact matching of the target values of ρ_1 and c_∞^2 when $\rho_1 < 0$ (arrival-process Set 1). Fortunately, the numerical results also show that increasing the order of the mixed Erlangs generally mitigates this difficulty. This makes sense, because as the order of the Erlangs increases, the realizations of the two mixed Erlangs become increasingly distinct, allowing high autocorrelation to more easily be achieved.

4. EVALUATION OF MARKOV MECO APPROXIMATIONS IN A QUEUEING SYSTEM

In this section, we empirically evaluate the Markov MECO approximations on the basis of corresponding congestion-measure approximations. That is, for a particular queueing system and a particular congestion measure, we observe how well the congestion measure for the queue with the approximate arrival process approximates that of the queue with the

TABLE 1

Basic Properties of Approximated Arrival Processes

	Stationary interarrival time					dimension	compact
	mean	c^2	γ	ρ_1	c_∞^2	by (1)	dimension
AP1-1	0.750	0.671	1.260	-0.165	0.333	81	15
AP1-2	0.461	0.658	1.409	-0.100	0.472	30	30
AP1-3	0.363	0.619	1.264	-0.129	0.120	24	24
AP1-4	0.341	1.085	1.987	0.013	1.090	162	30
AP1-5	1.333	0.754	1.019	-0.249	0.250	64	20
AP1-6	0.323	0.777	1.476	-0.102	0.443	160	30
AP1-7	0.270	0.711	1.294	-0.112	0.403	160	30
AP1-8	0.319	0.682	1.477	-0.072	0.595	24	24
AP2-1	0.333	1.733	3.213	0.060	2.250	8	4
AP2-2	0.167	1.418	3.025	0.080	2.250	64	7
AP2-3	0.333	4.011	5.962	0.054	6.250	8	4
AP2-4	0.167	2.508	6.116	0.112	6.250	64	7
AP2-5	0.167	1.767	3.959	0.117	4.250	64	16
AP2-6	0.278	1.533	3.045	0.075	2.250	64	16
AP2-7	0.278	2.880	5.671	0.115	6.250	64	16
AP2-8	0.278	1.583	3.202	0.077	2.917	64	16
AP2-9	0.278	2.639	4.970	0.137	5.583	64	16

original arrival process. In all cases, the original arrival process is a MAP that is a superposition of independent renewal processes. The first set of test cases consists of "data sets" 1 through 8 in [4]. For most of the renewal processes superposed in these cases, the coefficient of variation c of the interarrival time is less than one. The second set of arrival processes considered consists of superpositions of renewal processes whose interarrival times are hyperexponential and hence have coefficient of variation greater than one. Both sets are listed in the appendix.

Table 1 shows some basic information for the arrival processes to be approximated. The first set is labeled AP1-1, AP1-2, ..., AP1-8; the second set is labeled AP2-1, AP2-2, ..., AP2-9. The table shows the first three moments of the interarrival-time distribution, the lag-1 autocorrelation, the limiting value of the IDI, the dimension of the MAP representation based on (1) and the dimension obtained by using a more compact representation for superpositions of identical MAP's, as described in [4, pages 910, 911]. We observe that for all arrival processes in Set 1 except AP1-4 (where one of the superposed arrival processes has coefficient of variation greater than 1), $c^2 < 1$, $\rho_1 < 0$, and $c_\infty^2 < c^2$. In contrast, for all arrival processes in Set 2, $c^2 > 1$, $\rho_1 > 0$, and $c_\infty^2 > c^2$. Also, the largest value of $|\rho_1|$ across both arrival-process sets is 0.249.

Tables 2 and 3 present the results of applying Markov-MECO approximations to the arrival process of a MAP/M/1 queue to approximate E(N), the steady-state mean number of customers in the system. The author also observed approximations for the steady-state mean time in the queue, and the results were similar to those shown for E(N). Table 2 shows results for traffic intensity 0.6, and Table 3 show results for traffic intensity 0.9. For each approximated arrival process, Tables 2 and 3 show

1. the value of E(N) corresponding to the original arrival process and
2. for the approximations based on the minimum feasible order of the mixed Erlangs,
    a. the order of the two Erlangs mixed and
    b. the corresponding relative error in E(N) for the three-moment renewal approximation that completely ignores autocorrelation and for the three-moment approximations that match ρ_1 and c_∞^2 as closely as possible.

The minimum order is based on equations (2) and (3), subject to the restriction that the pair (c, γ) does not lie within 0.05 of the bounds of the feasible region defined by (4) and (5). (Numerical difficulties arise close to these bounds.) Relative error is computed as

$$\frac{\text{approximate E}(N) - \text{actual E}(N)}{\text{actual E}(N)}$$

so that the sign of the error indicates whether the approximation is larger or smaller than the actual performance-measure value. An asterisk in the last two columns of Tables 2 and 3 denotes a case for which the target value of ρ_1 or c_∞^2 was not attainable for the specified Erlang order. For some of the arrival processes, Tables 2 and 3 also report the effect of incrementing the order of the mixed Erlangs in the two correlated arrival-process approximations.

TABLE 2

Approximation results, traffic intensity = 0.6

Arrival Process	Order of Erlangs	E(N)	Relative Error in E(N): renewal process	Relative Error in E(N): match ρ_1	Relative Error in E(N): match c_∞^2
AP1-1	3	1.188	0.113	0.037	0.031*
	4			-0.005	-0.020*
AP1-2	2	1.226	0.063	0.048*	0.053*
	3			0.015	-0.022
AP1-3	3	1.200	0.068	0.014	0.005*
	4			0.012	-0.016
AP1-4	4	1.576	-0.015	-0.005	-0.013
AP1-5	3	1.124	0.127	0.087*	0.087*
	4			0.056*	0.056*
	5			0.028	-0.004*
AP1-6	2	1.265	0.095	0.054*	0.054*
	3			0.036	-0.037*
AP1-7	3	1.239	0.089	0.031	-0.006*
	4			0.030	-0.047
AP1-8	2	1.261	0.043	0.022*	0.022*
	3			0.006*	0.004
Average magnitude, Set 1			0.077	0.017	0.020
AP2-1	1	2.085	-0.113	-0.061	-0.024
AP2-2	1	1.955	-0.141	-0.076	-0.043
AP2-3	1	3.743	-0.294	-0.257	-0.140
AP2-4	1	3.133	-0.385	-0.349	-0.285
AP2-5	1	2.669	-0.288	-0.217	-0.142
AP2-6	1	2.002	-0.128	-0.063	-0.019
AP2-7	1	3.373	-0.368	-0.309	-0.213
AP2-8	1	2.077	-0.153	-0.092	0.012
AP2-9	1	3.265	-0.352	-0.262	-0.171
Average magnitude, Set 2			0.247	0.187	0.117

TABLE 3

Approximation results, traffic intensity = 0.9

			Relative Error in E(N)		
Arrival Process	Order of Erlangs	E(N)	renewal process	match ρ_1	match c_∞^2
AP1-1	3	6.259	0.216	0.110	0.101*
	4			0.050	0.030*
AP1-2	2	6.799	0.109	0.091*	-0.096*
	3			0.044	-0.008
AP1-3	3	6.576	0.121	0.046	0.034*
	4			0.041	-0.000
AP1-4	4	9.435	-0.006	-0.007	-0.004
AP1-5	3	5.871	0.221	0.171*	0.171*
	4			0.097*	0.097*
	5			0.081	-0.033*
AP1-6	2	6.758	0.193	0.141*	0.162*
	3			0.113	0.008*
AP1-7	3	6.584	0.183	0.101	0.051*
	4			-0.098	-0.016
AP1-8	2	7.141	0.070	0.044*	0.070*
	3			0.021*	0.018
Average magnitude, Set 1			0.140	0.057	0.015
AP2-1	1	14.171	-0.151	-0.070	-0.006
AP2-2	1	13.928	-0.233	-0.100	-0.015
AP2-3	1	30.523	-0.310	-0.246	-0.025
AP2-4	1	29.352	-0.504	-0.404	-0.098
AP2-5	1	21.565	-0.443	-0.295	-0.047
AP2-6	1	13.993	-0.201	-0.090	-0.003
AP2-7	1	29.702	-0.452	-0.326	-0.044
AP2-8	1	16.099	-0.294	-0.194	0.019
AP2-9	1	27.438	-0.435	-0.260	-0.037
Average magnitude, Set 2			0.336	0.221	0.028

To summarize the information displayed, Tables 2 and 3 also show for each combination of arrival-process set and approximation type, the average magnitude of the relative error in E(N). For the cases in which more than one Erlang order is considered, only the error corresponding to the highest order is included in the average. Tables 2 and 3 lead to the following observations.

1. Except in one case, the renewal approximation is poorer than the approximations that account for autocorrelation. (For the exceptional case, AP1-4 with traffic intensity 0.9, the arrival process has almost no autocorrelation.) Moreover, the improvement attained by accounting for autocorrelation is often substantial, especially for the c_∞^2-matching approximation. For Set 1, when traffic intensity is 0.6, the average relative error, 0.02, shown in Table 2 for the c_∞^2-matching approximation is 74% lower than the average relative error, 0.077, shown for the renewal approximation; when traffic intensity is 0.9, the reduction is 89%. For Set 2, the analogous reductions are 53% and 92%, respectively. The reduction in average relative error for the ρ_1-matching approximations is 78% and 24% for Sets 1 and 2, respectively, when traffic intensity is 0.6, and 59% and 34% for Sets 1 and 2, respectively, when traffic intensity is 0.9.
2. The precise relationship between the ρ_1- and c_∞^2-matching approximations depends on both the arrival-process set and the traffic intensity:
   a. For Set 1, when traffic intensity is 0.6, neither approximation dominates the other. But when traffic intensity is 0.9, the c_∞^2-matching approximation is usually better. The exceptions are all cases in which c_∞^2 cannot be matched exactly.
   b. For Set 2, the c_∞^2-matching approximation is consistently better than the ρ_1-matching approximation. When traffic intensity is 0.9, the difference is often an order of magnitude; when traffic intensity is 0.6, the difference is less dramatic.
3. The relationship between traffic intensity and the magnitude of the relative error in E(N) depends on the approximation type:
   a. With the exception of AP1-4, for the renewal approximation, the magnitude of the relative error in E(N) is larger for traffic intensity 0.9 than for traffic intensity 0.6.
   b. Similarly, with the exception of AP2-3 and AP2-9, for the ρ_1-matching approximation, the magnitude of the relative error in

E(N) is larger for traffic intensity 0.9 than for traffic intensity 0.6.

c. In contrast, with the exception of AP2-8, for the c_∞^2-matching approximation, the magnitude of the relative error in E(N) is larger for traffic intensity 0.6 than for traffic intensity 0.9 for arrival processes in Set 2. The Set 1 cases are mixed.

4. For Set 2, when traffic intensity is 0.6, even the c_∞^2-matching approximations lead to relative errors that are large in magnitude, say in excess of 10%. But when traffic intensity is 0.9, the c_∞^2-matching approximations consistently lead to relative errors under 10%.

5. For Set 1, the renewal and ρ_1-matching approximations usually overestimate E(N). For Set 2, all approximation types underestimate E(N). (The only exception is AP2-8 with the c_∞^2-matching approximation.)

6. For all the cases, the approximation(s) considered have dimension no greater than half the compact dimension of the original arrival process, as listed in the last column of Table 1.

7. For the nonrenewal approximations, incrementing the order of the mixed Erlangs when ρ_1 or c_∞^2 cannot be matched exactly usually reduces the error in E(N). (The exceptions are AP1-3 and AP1-7 with traffic intensity 0.6.)

8. Finally, for Set 1, a comparison with the approximation in Bitran and Dasu [4] is in order. In [4], the average dimension of the approximating arrival-process is 8.125. In this paper, using for each case the approximation of largest dimension shown, the average dimension for the Set-1 approximations is 7.5, very close to that of [4]. However, in [4], the mean relative error in E(N) is only 0.005 when traffic intensity is 0.6 and 0.002 when traffic intensity is 0.9. Thus, for Set 1, the approximations of [4] outperform those presented here. This suggests that when the approximation of Bitran and Dasu is feasible, the extra effort required to obtain it may be worthwhile. Also, since the approximation of Bitran and Dasu matches almost the same set of descriptors as the c_∞^2-matching approximation presented here, a better understanding of why the difference in performance arises would be helpful.

5. SUMMARY AND CONCLUSIONS

This paper provides a simple means of extending the three-moment-matching method in [16] to account for autocorrelation among interarrival times.

Closed-form expressions are given both for matching the ρ_1, lag-one autocorrelation, and for matching c_∞^2, the asymptotic index of dispersion for intervals. The resulting arrival-process model, labeled a Markov MECO, is a MAP. The primary limitation of the Markov MECO approach is that it allows only one degree of freedom for approximating the autocorrelation function. Another consideration is that to match a particular value of ρ_1 or c_∞^2, the orders of the mixed Erlangs may need to be greater than needed to simply match three moments. Otherwise, the ability to approximate the autocorrelation function comes at no increase in the dimension of the MAP used as an arrival-process approximation.

Although the ρ_1-matching and c_∞^2-matching approximations consistently outperform the renewal approximations, the quality of these approximations varies. Usually, the c_∞^2-matching approximation outperforms the ρ_1-matching approximation. However, this dominance does not occur when c_∞^2 is less than c^2, the squared coefficient of variation of the interarrival time, (indicating the presence of negative autocorrelations) and traffic intensity is modest. Among the c_∞^2-matching approximations, the poorest performance is observed for cases in which $c_\infty^2 > c^2$ and traffic intensity is modest. Conversely, matching c_∞^2 is most helpful when $c_\infty^2 > c^2$ and traffic intensity is high. For such cases, the improvement over the renewal approximation is roughly an order of magnetude. Finally, when $c_\infty^2 < c^2$, matching ρ_1 or c_∞^2 exactly may require incrementing the order of the Erlangs mixed; this often improves the approximation of the queueing performance measure.

The improvement in the relative errors shown for the ρ_1-matching and c_∞^2-matching approximations over the renewal approximations adds to existing evidence of the usefulness of the autocorrelation function as a descriptor of dependence among interarrival times. Because the autocorrelation function is often helpful and is so well known, the author believes methods for fitting it should be further pursued, while at the same time a better understanding of its limitations also sought. The author's experimentation suggests that an approximation that allows a more precise fit of the autocorrelation function would be helpful in some cases. For example, for AP1-5, $\rho_1 = -0.249$, $\rho_2 \approx 0.05$, and the autocorrelation for all other lags is approximately zero. This sharp drop in the autocorrelations cannot be closely approximated by the geometric autocorrelation function of a Markov MECO, and this difficulty is reflected in the rather high relative approximation errors for Erlang orders three and four. Fortunately, even for this case, approximating c_∞^2 as closely as possible leads to modest relative errors when the order in incremented to five. The author experimented briefly with an approximation using a three-state MRP and a mixture of three Erlangs. A numerical search was used to obtain conditional probabilities to fit the original autocorrelation function. No positive results were obtained, but the idea was by no means thoroughly investigated. However, any improvement would come at the cost of an approximation that has higher dimension and is more difficult to obtain (probably requiring a numerical search).

Finally, a few comments about the magnitude of autocorrelations in superpositions of independent renewal processes are in order. Studies indicating the need to account for autocorrelation in the arrival process typically make their point by inducing positive autocorrelation into an arrival process and showing the corresponding increase in queueing congestion. See, for example, [17, 19, 24]. This increased congestion becomes dramatic when sufficient autocorrelation is induced. However, the examples presented in this paper all display modest autocorrelation. This suggests that while accounting for autocorrelation is important, the dramatic examples found in the literature may not be found among superpositions of independent renewal processes. On the other hand, the range of examples in this paper is by no mean comprehensive. A far more comprehensive study is needed to obtain a definitive understanding of the autocorrelation functions typically arising from actual applications.

APPENDIX
Arrival Processes

Arrival Processes of Set 1

Recall that all of the arrival processes of Set 1 are superpositions of independent renewal processes. The following notation is used to specify the superposed arrival processes. A renewal process whose interarrival times are Erlang of order k and mean k/λ is denoted by $E(k,\lambda)$. A renewal process whose interarrival times have a hyperexponential distribution that weights the exponential with rate λ_1 by probability p_1 and the exponential with rate λ_2 by probability p_2 is denoted by $Hx(p_1,\lambda_1,p_2,\lambda_2)$. A PH renewal process is denoted by $PH(\alpha,\mathbf{T})$, where α and $\mathbf{T}$ are as defined in Section 2 of this paper.

A superposition of n identical $E(k,\lambda)$ renewal processes is denoted by $nE(k,\lambda)$; likewise for superpositions of renewal processes with hyperexponential and PH interarrival times. A + sign is used to denote the superposition of nonidentical renewal processes.

AP1-1: $4E(3, 1)$

AP1-2: $PH(\alpha_1, \mathbf{T}_1) + PH(\alpha_2, \mathbf{T}_2) + PH(\alpha_3, \mathbf{T}_3)$, where

$\alpha_1 = (1,0,0)$, $\alpha_2 = (1,0)$, $\alpha_3 = (1,0,0,0)$,

$$\mathbf{T}_1 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0.5 \\ 0 & 0 & -1 \end{bmatrix}, \mathbf{T}_2 = \begin{bmatrix} -2 & 2 \\ 0 & -3 \end{bmatrix}, \text{ and } \mathbf{T}_3 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & -5 \end{bmatrix}.$$

AP1-3: $PH(\alpha_1, \mathbf{T}_1) + PH(\alpha_2, \mathbf{T}_2) + PH(\alpha_3, \mathbf{T}_3)$, where the first two processes are as specified for AP1-2, $\alpha_3 = (1,0,0,0)$, and $\mathbf{T}_3 = \begin{bmatrix} -3 & 3 & 0 & 0 \\ 0 & -4 & 4 & 0 \\ 0 & 0 & -5 & 2.5 \\ 0 & 0 & 0 & -6 \end{bmatrix}$.

AP1-4: $4E(3, 1) + Hx(0.5, 1, 0.5, 4)$

AP1-5: $3E(4, 1)$

AP1-6: $5E(2, 1) + E(5, 3)$

AP1-7: $5E(2, 1) + \mathrm{E}(5, 6)$

AP1-8: $PH(\alpha_1, \mathbf{T}_1) + PH(\alpha_2, \mathbf{T}_2) + E(3, 2)$, where $\alpha_1 = (1,0,0,0)$, $\alpha_2 = (1,0)$,

$$\mathbf{T}_1 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & -3 \end{bmatrix}, \text{ and } \mathbf{T}_2 = \begin{bmatrix} -3 & 3 \\ 0 & -4 \end{bmatrix}.$$

Arrival Processes of Set 2

The arrival processes of Set 2 are all superpositions of independent renewal processes whose interarrival-time distributions are two-state hyperexponential distributions. For each hyperexponential distribution, the coefficient of skewness is set to twice the coefficient of variation. (This relationship holds for the family of gamma distributions.) In the detailed specifications below, $H(\mu_1 = a, c = b)$ denotes a renewal process whose interarrival-time distribution is a two-state hyperexponential with mean $\mu_1 = a$, coefficient of variation $c = b$, and coefficient of skewness $\gamma = 2c$; $nH(\mu_1 = a, c = b)$ denotes the superposition of n such processes. Superposition of nonidentical arrival processes is indicated by a + sign.

AP2-1: $3H(\mu_1 = 1, c = 1.5)$
AP2-2: $6H(\mu_1 = 1, c = 1.5)$
AP2-3: $3H(\mu_1 = 1, c = 2.5)$
AP2-4: $6H(\mu_1 = 1, c = 2.5)$
AP2-5: $3H(\mu_1 = 1, c = 1.5) + 3H(\mu_1 = 1, c = 2.5)$
AP2-6: $3H(\mu_1 = 1, c = 1.5) + 3H(\mu_1 = 5, c = 1.5)$
AP2-7: $3H(\mu_1 = 1, c = 2.5) + 3H(\mu_1 = 5, c = 2.5)$
AP2-8: $3H(\mu_1 = 1, c = 1.5) + 3H(\mu_1 = 5, c = 2.5)$
AP2-9: $3H(\mu_1 = 1, c = 2.5) + 3H(\mu_1 = 5, c = 1.5)$

REFERENCES

[1] S.L. Albin, Approximating a Point Process by a Renewal Process, II: Superposition Arrival Processes to Queues, *Operations Research*, 32:5, 1984, 1133-1162.

[2] A.S. Alfa and M.F. Neuts, Modelling Vehicular Traffic Using the Discrete Time Markovian Arrival Process, *Transportation Science*, 29:2, 1995, 109-117.

[3] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, Long-Range Dependence in Variable-Bit-Rate Video Traffic, IEEE Transactions on Communications, 13:2/3/4, 1995, 1566-1579.

[4] G.R. Bitran and S. Dasu, Approximating Nonrenewal Processes by Markov Chains: Use of Super-Erlang (SE) Chains, Operations Research, 41:5, 1993, 903-923.

[5] E. Çinlar, *Introduction to Stochastic Processes*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[6] D.R. Cox and P.A.W. Lewis, *The Statistical Analysis of Series of Events*, Wiley, New York, 1966.

[7] R. L. Disney and P.C. Kiessler, *Traffic Processes in Queueing Networks: A Renewal Theoretic Approach*, Johns Hopkins University Press, Baltimore, Maryland, 1987.

[8] M. W. Garrett and W. Willinger, Analysis, Modeling, and Generation of Self-Similar VBR Video Traffic, Proceedings of SIGCOMM 94, London, England, August 1994, 269-280.

[9] R. Gusella, Characterizing the Variability of Arrival Processes in Indexes

of Dispersion, *IEEE J. on Selected Areas in Communications*, 9:2, 1991, 203-211.

[10] H. Heffes, A Class of Data Traffic Processes Covariance Function Characterization And Related Queueing Results, *Bell System Technical Journal*, 59:6, 1980, 897-929.

[11] H. Heffes and D.M. Lucantoni, A Markov Modulated Characterization of Packetized Voice and Data Traffic And Related Statistical Multiplexer Performance, *IEEE J. on Selected Areas in Communications*, Special Issue on Network Performance Evaluation, 4, 1986, 856-868.

[12] M.A. Johnson, D. Liu, and S. Narayana, Burstiness Descriptors for Markov Renewal Processes and Markovian Arrival Processes, *Stochastic Models,* 1997, 13:3, to appear.

[13] M.A. Johnson and S. Narayana, Descriptors of Arrival-Process Burstiness with Application to the Discrete Markovian Arrival Process, *Queueing Systems*, 9, 1997 1-24.

[14] M.A. Johnson and M.R. Taaffe, A Graphical Investigation of Error Bounds for Moment-Based Queueing Approximations, *Queueing Systems—Theory and Applications*, 6, 1991, 295-312.

[15] M.A. Johnson and M.R. Taaffe, An Investigation of Phase-Distribution Moment-Matching Algorithms for Use in Queueing Models, *Queueing Systems—Theory and Applications*, 6, 1991, 129-147.

[16] M.A. Johnson and M.R. Taaffe, Matching Moments to Phase Distributions: Mixtures of Erlang Distributions of Common Order, *Stochastic Models*, 5:4, 1989, 711-743.

[17] G. Latouche, An Exponential Semi-Markov Process, with Applications to Queueing Theory, *Stochastic Models*, 1:2, 1985, 137-169.

[18] W. E. Leland, M. S. Taqqu, W. Willinger, D. V. Wilson, On the Self-Similar Nature of Ethernet Traffic, Proceedings of SIGCOMM 93, Ithica, New York, September 1993, 183-193.

[19] M. Livny, B. Melamed, and A.K. Tsiolis, The Impact of Autocorrelation on Queueing Systems, *Management Science*, 39:3, 1993, 322-339.

[20] K.S. Meier-Hellstern, A Fitting Algorithm for Markov-Modulated Poisson Processes Having Two Arrival Rates, *European Journal of Operational Research*, 29, 1987, 370-377.

[21] M.F. Neuts, The Burstiness of Point Processes, *Stochastic Models*, 9:3, 1993, 445--466.

[22] M.F. Neuts, Models Based on the Markovian Arrival Process, *IEIC Transactions on Communications*, E75-B:12, 1992, 1255-1265.

[23] M.F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach*, Johns Hopkins University Press, Baltimore, Maryland, 1981.

[24] B.E. Patuwo, D.C. McNickle, and R.L. Disney, The Effect of Correlated Arrivals on Queues, *IIE Transactions*, 25:3, 1993, 105-110.

[25] T. Rydén, Parameter Estimation for Markov Modulated Poisson Processes, *Stochastic Models*, 10:4, 1994, 795-829.

[26] T. Rydén, An EM Algorithm for Multiple State MMPP Parameter Estimation, Technical Report, Department of Mathematical Statistics, University of Lund and Lund Institute of Technology, Sweden, 1993.

[27] H. Sitaraman, Approximation of Some Markov-Modulated Poisson Processes, *ORSA Journal on Computing*, 3:1, 1991, 12-22.

[28] K. Sriram and W. Whitt, Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data, *IEEE Journal on Selected Areas in Communications*, SAC-4:6, 1986, 833-846.

[29] W. Whitt, Approximating a Point Process by a Renewal Process, I: Two Basic Methods, *Operations Research*, 30:1, 1982, 125-147.