Characterization and Synthesis of Markovian Workload Models

Giuliano Casale
College of William & Mary
Williamsburg, VA 23187-8795, US
Email: casale@cs.wm.edu

Eddy Z. Zhang
College of William & Mary
Williamsburg, VA 23187-8795, US
Email: eddy@cs.wm.edu

Evgenia Smirni
College of William & Mary
Williamsburg, VA 23187-8795, US
Email: esmirni@cs.wm.edu

***Abstract*—We consider the general problem of workload model generation using Markovian Arrival Processes (MAPs). MAPs are a large class of analytically tractable processes frequently used in communication and computer network modeling. We show that MAP moment and autocorrelation formulas admit a simple scalar form deriving from spectral properties of the MAP defining matrices. This suggests a new approach for studying MAPs, by which we address challenging characterization and fitting problems as well as the open issue of synthesizing processes with prescribed moments and acf for inter-arrival times. A case study illustrates the impact of spectral-based synthesis on sensitivity analysis of network models.**

I. Introduction

Markovian Arrival Processes (MAPs) [12] form a general class of point processes which admits hyper-exponential, Erlang, and Markov Modulated Poisson Processes (MMPPs) as special cases. The most appealing feature of a MAP is the ease of its integration within queueing models, which makes this technology useful for evaluating the performance effects of non-Poisson workloads. Such workloads are prevalent in networking where long-range dependent traffic has been long identified as an important traffic characteristic, but are also recently emerging in systems including disk drives [15] and multi-tiered systems of e-commerce applications [11]. Renewal service processes of high variability have been recently shown inadequate for performance prediction of multi-tiered Internet services if there is dependence in the traffic flows through the various tiers [11]. Models of such systems that are parameterized via MAPs dramatically improve analysis accuracy and applicability to real networks.

Although the MAP technology is rapidly growing and many new results and applications have been recently presented [2], [6], [8], [9], [11], [17], little advances have been obtained in the characterization of the actual capabilities of MAPs. In this work, we provide an analytical characterization of MAPs based on a spectral analysis of the moments and of the autocorrelation function (acf). Our result simplifies the analysis of general MAPs by representing the most important statistics in terms of a few scalar parameters. We illustrate applications of this result to characterization, fitting and synthesis of MAP processes. Sensitivity analysis of network queueing models is shown as an application of spectral-based process synthesis.

A. Markovian Arrival Processes

We point to [12] for background on MAPs, and we limit here to a synthetic overview. A MAP(n) is specified by two $n \times n$ matrices, a stable matrix $\boldsymbol{D_0}$ and a nonnegative matrix $\boldsymbol{D_1}$, that describe transition rates between n states. Each transition in $\boldsymbol{D_1}$ produces a job arrival; $\boldsymbol{D_0}$ describes background transitions not associated with arrivals; $\boldsymbol{Q} = \boldsymbol{D_0} + \boldsymbol{D_1}$ is the infinitesimal generator of the underlying continuous-time Markov chain. We focus on the inter-arrival (or equivalently service) time description of arrival processes [16]. For a MAP(n), inter-arrival time moments and acf are computed using the probability vector $\boldsymbol{\pi}_e$, $\boldsymbol{\pi}_e \boldsymbol{e} = 1$, of the embedded process with irreducible stochastic matrix $\boldsymbol{P} = (-\boldsymbol{D_0})^{-1}\boldsymbol{D_1}$, where $\boldsymbol{e}$ is a column vector of 1's of the appropriate dimension. The MAP inter-arrival times are identically distributed with mean $E[X] = \boldsymbol{\pi}_e(-\boldsymbol{D_0})^{-1}\boldsymbol{e}$, squared coefficient of variation $c_v^2 = 2E[X]^{-2}\boldsymbol{\pi}_e(-\boldsymbol{D_0})^{-2}\boldsymbol{e} - 1$, and k-th moment

$$E[X^k] = k!\boldsymbol{\pi}_e(-\boldsymbol{D_0})^{-k}\boldsymbol{e}, \quad k \geq 0. \qquad (1)$$

The lag-k acf coefficient is computed as

$$\rho_k = \frac{E[X]^{-2}\boldsymbol{\pi}_e(-\boldsymbol{D_0})^{-1}\boldsymbol{P}^k(-\boldsymbol{D_0})^{-1}\boldsymbol{e} - 1}{c_v^2}, \quad k \geq 1. \qquad (2)$$

Throughout the paper we refer to the $(\boldsymbol{D_0}, \boldsymbol{D_1})$ representation as the Markovian representation of a MAP.

B. Paper Organization

The paper is organized as follows. We present in Section II the spectral analysis of MAPs. Characterization and fitting applications are exemplified in Section III on the MAP(2) process. Spectral-based process synthesis is given in Section IV, and applied to a network model in Section V to illustrate the critical impact of non-renewal workloads in models. Finally, Section VI draws conclusions and outlines future work. A preliminary non-copyrighted version of this paper has been recently presented at the MAMA'07 workshop, San Diego.

II. Spectral Representation

We develop a *spectral representation* of MAPs, i.e., a simple scalar representation of (1)-(2) based on spectral properties of $(-\boldsymbol{D_0})^{-1}$ and $\boldsymbol{P}$. The idea is that of representing MAPs moments and acf in terms of a set of few fundamental parameters, rather than by matricial formulas. Applications of this simplified representation are shown in the next sections.

A. Characterization of Moments

We begin by describing the moments (1) in terms of the spectrum of $(-\boldsymbol{D_0})^{-1}$. Recall that the characteristic polynomial $\phi_{\mathbf{A}} \equiv \phi_{\mathbf{A}}(s)$ of a $n \times n$ matrix $\mathbf{A}$ is $\phi_{\mathbf{A}} = s^n + \alpha_1 s^{n-1} + \ldots + \alpha_{n-1}s + \alpha_n$, which is a polynomial in s with roots s_i equal to the eigenvalues of $\mathbf{A}$. We consider the Cayley-Hamilton theorem [7], which states that the powers of $\mathbf{A}$ satisfy $\mathbf{A}^k = -\sum_{j=1}^n \alpha_j \mathbf{A}^{k-j}$, for $k \geq n$, i.e., that matrix powers of $\mathbf{A}$ are linearly dependent according to the coefficients[1] of $\phi_{\mathbf{A}}$. Since MAP moments are computed in (1) from matrix powers of $(-\boldsymbol{D_0})^{-1}$, it is intuitive that they may consequently be linearly dependent as we now prove.

Lemma 1. *In a* MAP(n)*, any* $n+1$ *consecutive moments are linearly dependent, i.e.,*

$$E[X^k] = -\sum_{j=1\ldots n} b_j E[X^{k-j}], \quad E[X^0] = 1, \quad k \geq n, \tag{3}$$

where $b_j = m_j k!/(k-j)!$*, and* m_j *is the coefficient of* s^{n-j} *in* $\phi_{(-\boldsymbol{D_0})^{-1}}$.

Proof: From the Cayley-Hamilton theorem it is $E[X^k] = k!\boldsymbol{\pi}_e(-\boldsymbol{D_0})^{-k}\boldsymbol{e} = -k!\boldsymbol{\pi}_e \sum_{j=1\ldots n} m_j(-\boldsymbol{D_0})^{-k-j})\boldsymbol{e} = -\sum_{j=1\ldots n} \frac{k!\,m_j}{(k-j)!}E[X^{k-j}]$. ■

Observing that the coefficients of a characteristic polynomial $\phi_{\mathbf{A}}$ are functions of the eigenvalues of $\mathbf{A}$, we can derive the relation between eigenvalues of $(-\boldsymbol{D_0})^{-1}$ and moments.

Theorem 1 (Spectral Representation of Moments)**.** *Let* $\theta_t \in \mathbb{C}$*,* $1 \leq t \leq m$*, be the* m *distinct eigenvalues of* $(-\boldsymbol{D_0})^{-1}$*, each with multiplicity* q_t*. Then*

$$E[X^k] = \sum_{t=1\ldots m} k!\,\theta_t^k \sum_{j=1\ldots q_t} M_{t,j} k^{j-1}, \tag{4}$$

$$E[X^0] = \sum_{t=1\ldots m} M_{t,1} = 1, \tag{5}$$

and the constants $M_{t,j}$ *follow imposing* n *arbitrary moments.*

Proof: Equation (3) can be seen as a homogeneous linear recurrence of order n in $E[X^k]/k!$ with constant coefficients m_j. The general solution thus depends on n particular solutions and on the n roots of the associated characteristic equation which are exactly the eigenvalues θ_t of $(-\boldsymbol{D_0})^{-1}$. ■

Observing that the $M_{t,j}$ and θ_t are $2n-2$ parameters, and that the $M_{t,j}$'s are linearly dependent due to the condition $E[X^0] = 1$, we have the following corollary.

Corollary 1 (Independent Moments)**.** *A* MAP(n) *process can fit up to* $2n-1$ *independent moments.*

To appreciate the economicity of the spectral representation (4), consider one of the simplest MAPs, i.e., the MMPP(2) process

$$\boldsymbol{D_0} = \begin{bmatrix} -q_{12}-\mu_1 & q_{12} \\ q_{21} & -q_{21}-\mu_2 \end{bmatrix}, \boldsymbol{D_1} = \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix},$$

[1]The α_j's can be easily computed, e.g., with the MATLAB function `poly`.

TABLE I
FIRST THREE MOMENTS OF A MAP(2)PROCESS

Moment	*Markovian Representation*
$E[X]$	$\dfrac{q_{12}+q_{21}}{\mu_1 q_{21}+q_{12}\mu_2}$
$E[X^2]$	$\dfrac{2(\mu_1 q_{21}+q_{12}\mu_2)(q_{12}+q_{21})(\mu_1 q_{21}+q_{12}\mu_2)^{-1}+2}{(\mu_1\mu_2+\mu_1 q_{21}+q_{12}\mu_2)}$
$E[X^3]$	$\dfrac{6[\mu_1^2 q_{12}+\mu_2^2 q_{21}+(q_{12}+q_{21})(q_{12}\mu_1+q_{21}\mu_2)]}{(\mu_1 q_{21}+q_{12}\mu_2)(\mu_1\mu_2+\mu_1 q_{21}+q_{12}\mu_2)^2}$

and assume that $(-\boldsymbol{D_0})^{-1}$ has distinct eigenvalues. Table I shows the formulas of the first three moments according to the Markovian representation. Using the spectral representation the same moments become

$$E[X] = M_{1,1}\theta_1 + M_{2,1}\theta_1,$$
$$E[X^2] = 2M_{1,1}\theta_1^2 + 2M_{2,1}\theta_2^2,$$
$$E[X^3] = 6M_{1,1}\theta_1^3 + 6M_{2,1}\theta_2^3.$$

The spectral representation is thus able to dramatically simplify formulas with respect to the Markovian representation and clearly reveals the structure of the moments. Furthermore, Corollary 1 shows that only $2n-1$ parameters are needed to impose the maximum number of fittable moments, and stresses the redundancy of the Markovian representation. These simplifications extend also to the acf coefficients.

B. Characterization of Autocorrelation

The spectral characterization can be extended to acf coefficients by considering the spectrum of $\boldsymbol{P}$, which determines the properties of the matrix powers $\boldsymbol{P}^k$ in (2).

Lemma 2. *In a* MAP(n)*, any* $n+1$ *consecutive acf coefficients are linearly dependent, i.e.,*

$$\rho_k = -\sum_{j=1\ldots n} a_j \rho_{k-j}, \quad \rho_0 = \frac{1}{2}\left(1 - \frac{1}{c_v^2}\right), \quad k \geq n, \tag{6}$$

where a_j *is the coefficient of* s^{n-j} *in* $\phi_{\boldsymbol{P}}$ *and* $\sum_{j=1}^n a_j = 0$.

Proof: We wish to prove $\sum_{j=0\ldots n} a_j\rho_{k-j} = 0$, $a_0 = 1$. By definition of ρ_k this is equal to

$$\sum_j a_j(E[X]^{-2}\boldsymbol{\pi}_e(-\boldsymbol{D_0})^{-1}\boldsymbol{P}^{k-j}(-\boldsymbol{D_0})^{-1}\boldsymbol{e} - 1) = 0.$$

Note that the statement is indeed true if we can show that $\sum_{j=0}^n a_j \boldsymbol{P}^{k-j} = 0$ and $\sum_{j=0}^n a_j = 0$. But the first equality is true by the Cayley-Hamilton theorem; the second relation follows by the stochasticity of $\boldsymbol{P}$, as we have that its largest eigenvalue is always $\gamma_1 = 1$ and thus $\phi_{\boldsymbol{P}}(\gamma_1) = 0 \Rightarrow a_1 + a_2 + \ldots + a_n = 0$, which finally proves $\rho_k = -\sum_{j=1\ldots n} a_j\rho_{k-j}$.

The boundary condition $\rho_0 = \frac{1}{2}(1 - 1/c_v^2)$ follows by considering (2) for $k = 0$. We finally have

$$\rho_0 = \frac{2E[X]^{-2}\boldsymbol{\pi}_e(-\boldsymbol{D_0})^{-2}\boldsymbol{e} - 2}{2c_v^2} = \frac{c_v^2 - 1}{2c_v^2} = \frac{1}{2}\left(1 - \frac{1}{c_v^2}\right).$$

■

Using a proof analoguos to that of Theorem 1, it is possible to relate ρ_k with the eigenvalues of $\boldsymbol{P}$ and characterize the maximum number of fittable acf coefficients.

Theorem 2 (Spectral Characterization of Autocorrelation)**.** *Let $\gamma_t \in \mathbb{C}$, $1 \le t \le m$, be an eigenvalue of $\boldsymbol{P}$ with multiplicity r_t. Let also $\gamma_1 = 1$ be the unit eigenvalue of $\boldsymbol{P}$. Then*

$$\rho_k = \sum_{t=2\dots m} \gamma_t^k \sum_{j=1\dots r_t} A_{t,j} k^{j-1}, \tag{7}$$

$$\rho_0 = \sum_{t=2\dots m} A_{t,1}, \quad k \ge 1, \tag{8}$$

where the $A_{t,j}$'s constants can be imposed from $n-2$ independent acf coefficients.

Observing that the distinct $A_{t,j}$ and γ_t in (7) are $2n-2$, and that fixing c_v^2 imposes ρ_0, we have the following corollary.

Corollary 2 (Independent Autocorrelation Coefficients)**.** *A* MAP(n) *process can fit up to $2n-2$ independent acf coefficients ρ_k, $k \ge 1$. With given c_v^2, the maximum number of independent coefficients becomes $2n-3$.*

Similarly to the moments, (7) has a much simpler structure than the corresponding Markovian formula (2).

C. Spectral Representation of MAP*s*

Summarizing, we can describe moments and acf coefficients using the set of parameters $(M_{t,j}, \gamma_t)$ and $(A_{t,j}, \gamma_t)$, respectively. The set $(M_{t,j}, \gamma_t)$ has $2n-1$ degrees of freedom, which once assigned leave $(A_{t,j}, \gamma_t)$ with $2n-3$ degrees of freedom. Therefore, only $4(n-1)$ degrees of freedom have to be assigned in a MAP(n) in order to fix moments and acf. Given that the Markovian representation requires $2n^2-n$ (redundant) parameters, and that a MAP(n) has no more than n^2 degrees of freedom [17], our result unexpectedly indicates that only $4n-4$ degrees of freedom should be spent to impose moments and acf. As we discuss in the next sections, this result can be fruitfully employed in workload characterization, MAP fitting and process design.

We conclude this section with two remarks. First, in the frequent case where $(-\boldsymbol{D_0})^{-1}$ and $\boldsymbol{P}$ have distinct eigenvalues, it can be shown by spectral decomposition [7] that

$$M_{t,1} = \boldsymbol{\pi}_e[(-\boldsymbol{D_0})^{-1}]_t \boldsymbol{e}, \tag{9}$$

$$A_{t,1} = (c_v^2)^{-1} E[X]^{-2} \boldsymbol{\pi}_e (-\boldsymbol{D_0})^{-1} [\boldsymbol{P}]_t (-\boldsymbol{D_0})^{-1} \boldsymbol{e}, \tag{10}$$

where $[\mathbf{A}]_i$ is the i-th spectral projector of matrix A, i.e., the rank-one matrix given by the product of the right and left eigenvectors of $\mathbf{A}$ for the eigenvalue s_i. Due to the direct relation with spectral projectors, we henceforth refer to the $M_{t,j}$ and $A_{t,j}$ constants as moment and acf projectors, respectively. We also remark that the spectral description is also able to represent other statistics of MAPs. For instance, if $(-\boldsymbol{D_0})^{-1}$ has distinct eigenvalues, it follows from (1) and (4) that the cumulative distribution function (cdf) of inter-arrival times [12] is simply

$$F(x) = 1 - \boldsymbol{\pi}_e e^{\boldsymbol{D_0} x} \boldsymbol{e} = 1 - \sum_{t=1}^{n} M_{t,1} e^{-x/\theta_t}. \tag{11}$$

III. MAP(2) Characterization and Fitting

MAP(2)s are used in traffic characterization thanks to their small parametrization space of just six parameters. Several characterization results have been proposed for MAP(2) using diagonalization methods and matrix exponentials [6]. In order to illustrate the simplicity of characterizing MAP(2)s using spectral methods, we immediately derive the structure of MAP(2) moments and acf coefficients. According to (4), the moments of a MAP(2) are

$$E[X^k] = k! M_{1,1} \theta_1^k + k!(1 - M_{1,1}) \theta_2^k, \tag{12}$$

and up to three independent moments may be fitted. Similarly, the acf coefficients are

$$\rho_k = \gamma_2^k A_{2,1} = \gamma_2^k \rho_0 = \frac{\gamma_2^k}{2}\left(1 - \frac{1}{c_v^2}\right), \; k \ge 1, \gamma_2 \in \mathbb{R}, \tag{13}$$

which can fit a single acf coefficient with fixed c_v^2, and admits no more than a few different shapes according to the signs of γ_2 and c_v^2. Note also that ρ_k always converges to zero as $k \to \infty$, unless $\gamma_2 = -1$ which produces obscillations. It can also be shown from criteria in [4] that λ_1 and λ_2 are both reals, and by (12) that $\theta_2 \le E[X] \le \theta_1$ assuming $\theta_2 \le \theta_1$.

Our characterization clearly indicates that the MAP(2) process offers very limited versatility in exploring the impact of non-renewal workloads on systems, since the acf coefficients are always geometrically decaying with rate γ_2, whereas real workloads typically exhibit different decaying rates at low and high lags [11], [13]. Superposition methods [1] can be used to overcome this limitation by creating larger, more flexible, processes but these are limited to the superposition of MMPP(2)s because of the difficulty of assigning the two redundant parameters in the Markovian representation of a MAP(2). Problems of this type are usually tackled with nonlinear optimization methods, which are quite often difficult numerically. Using the spectral characterization, we solved the problem of fitting a MAP(2) in its generality as we show next for a particular case. We point to [18] for a comprehensive discussion.

A. General MAP(2) *fitting*

Given the four independent parameters $\lambda_1, \lambda_2, M_{1,1}, \gamma_2$ which define moments and acf of a MAP(2), we consider the fitting of a MAP(2) with $\boldsymbol{D_0}$ and $\boldsymbol{D_1}$ diagonalized as

$$\boldsymbol{D_0} = X_0 \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} X_0^{-1}, \; X_0 = \begin{bmatrix} m & k \\ 1 & 1 \end{bmatrix},$$

$$\boldsymbol{D_1} = X_1 \begin{bmatrix} \nu_1 & 0 \\ 0 & \nu_2 \end{bmatrix} X_1^{-1}, \; X_1 = \begin{bmatrix} t & y \\ 1 & 1 \end{bmatrix}.$$

where $\lambda_t = -\theta_t^{-1}$, $t = 1, 2$. The Markovian representation can be related to the spectral representation using the relations

$$\lambda_1^{-1} \lambda_2^{-1} \nu_1 \nu_2 = \gamma_2, \quad \nu_2 = \frac{(M_{1,1}(\lambda_1 - \lambda_2) + \lambda_2 + \nu_1)\lambda_1 \lambda_2}{M_{1,1}\nu_1(\lambda_1 - \lambda_2) - \lambda_1(\nu_1 + \lambda_2)},$$

which hold by $\det(\boldsymbol{P}) = \det((-\boldsymbol{D_0})^{-1}) \det(\boldsymbol{D_1})$ and by definition of $M_{1,1}$, respectively.

Given the above definition, the problem of MAP(2) fitting is that of assigning values to the eigenvector entries m, k, t, y such that $(\boldsymbol{D_0}, \boldsymbol{D_1})$ is a valid MAP. Since the infinitesimal generator $\mathbf{Q} = \boldsymbol{D_0} + \boldsymbol{D_1}$ must have rows that sum to zero, we first need to explicitely impose

$$k = \frac{t(y-1)x_2 + y(1-t)y_2}{yx_2 - ty_2 - (x_2 - y_2)},\ m = \frac{t(y-1)x_1 + y(1-t)y_1}{yx_1 - ty_1 - (x_1 - y_1)},$$

where $x_1 = \lambda_2 + \nu_1$, $x_2 = \lambda_1 + \nu_1$, $y_1 = \lambda_2 + \nu_2$, $y_2 = \lambda_1 + \nu_2$. This step reduces the problem of fitting a MAP(2) to assigning y and t so that all rates between states are positive. This is obtained by imposing that the two variables belong to feasible regions bounded by certain hyperbolic and linear constraints [18]. Taking any feasible point (y^*, t^*) inside a region, e.g., its centroid, a MAP(2) can then be defined. E.g., for $x_1 < 0$, $y_1 < 0$, $x_2 < 0$, $y_2 < 0$, $\nu_2 < 0$, a feasible point (y^*, t^*) is $y^* = 0.5[(-\lambda_2)^{-1}\nu_2 + (\nu_1 + \nu_2 + \lambda_1)^{-1}\nu_1]$, $t^* = 0.5[\max(\nu_2 y^* \nu_1{}^{-1}, L_1(y^*)) + \min(\nu_1 y^* \nu_2{}^{-1}, L_2(y^*))]$. where $L_i(y) = [yx_i - (x_i - y_i)]y_i{}^{-1}$, $i = 1, 2$. We point to [18] for a comprehensive discussion of all other cases.

IV. Process Design

Process synthesis is important for design exploration studies, where the analyst wishes to evaluate system response in different scenarios. Non-renewal features can be used to dramatically improve model accuracy, e.g., in the reliability analysis of a system where the acf of service times may describe the temporal dependency of component failures.

Unfortunately, no analytical technique exists for generating higher-order MAPs with prescribed moments and acf of inter-arrival (or equivalently service) times. To address this limitation, we propose a quite general method based on Kronecker products. Since complex acf structure emerge only under non-negligible fluctuations with respect to the mean, we focus on MAPs with $c_v^2 \geq 1$.

A. A Flexible Class of MAP(n)s

We develop a class of MAP(n)s with Markovian representation $(\boldsymbol{D_0}, \boldsymbol{D_1})$ in which $\boldsymbol{D_0} = \mathrm{diag}(-\theta_1^{-1}, -\theta_2^{-1}, \ldots, -\theta_n^{-1})$, $\theta_t > 0$. This property yields two important consequences:

- given an arbitrary stochastic matrix $\boldsymbol{P}$, the process with Markovian representation $(\boldsymbol{D_0}, -\boldsymbol{D_0}\boldsymbol{P})$ is always a valid MAP, since $\boldsymbol{D_1} = \mathrm{diag}(\theta_1^{-1}, \theta_2^{-1}, \ldots, \theta_n^{-1})\boldsymbol{P}$ is always nonnegative. This allows to ignore non-linear feasibility constraints that make MAP fitting a non-trivial optimization problem [8].
- The spectral representation of moments has a particularly simple form, since eigenvalues θ_t are freely assigned with $\boldsymbol{D_0}$; if the θ_t's are chosen distinct, then the projectors $M_{j,t}$ become equal to the elements of the vector $\boldsymbol{\pi}_e$ and the $M_{j,t}$'s vector is stochastic. Hence, $\boldsymbol{\pi}_e$ and $\boldsymbol{D_0}$ uniquely assign moments and cdf. The latter, being $\boldsymbol{D_0}$ diagonal, is in fact $F(x) = 1 - \sum_{t=1}^{n} \pi_e^t e^{-x/\theta_t}$, where π_e^t is the t-th element of $\boldsymbol{\pi}_e$.

Other properties of this class of processes are now discussed.

Fig. 1. Inverse Spectral Characterization of Moments

Step 1. Obtain the n variables m_j's from a system of $n-1$ linear equations (3) for $n \leq k \leq 2n-1$ and the condition $E[X^0] = 1$.
Step 2. Solve $\phi_{(-\boldsymbol{D_0})^{-1}} = s^n + m_1 s^{n-1} + \ldots + m_{n-1}s + m_n$ for the n eigenvalues θ_t.
Step 3. Determine the $M_{t,j}$ constants from the system of linear equations (4) for $k = 0, \ldots, n$, where the θ_t eigenvalues are those obtained in Step 2.

1) Moments: Given a set of $2n-1$ moments, one can easily compute the related $M_{j,t}$ and λ_t values solving an inverse spectral characterization problem as shown in Figure 1. If the set of projectors $M_{j,t}$ is not stochastic or the eigenvalues θ_t are not positive, then the considered set of moments is not exactly fittable by our class. However, an approximation can be obtained, e.g., by the Feldmann-Whitt algorithm [5], which provides an approximate cdf with the same form and parameter ranges of the cdf $F(x)$ of our MAP process.

2) Autocorrelation Coefficients: Moments and acf assignment has been reduced to defining a stochastic matrix $\boldsymbol{P}$ with prescribed spectral properties, e.g., the steady-state probability vector $\boldsymbol{\pi}_e$ which is the left-eigenvector associated to the eigenvalue $\gamma_1 = 1$. The general problem of *exactly* assigning a spectrum to a matrix is known to be hard; however accurate approximations suffice in practice, and we observe that this can be done by exploiting properties of Kronecker products [3].

Recall that given two matrices $\mathbf{A}$ and $\mathbf{B}$, with order p and q, eigenvalues α_i and β_j, and eigenvectors $\mathbf{a}_i$ and $\mathbf{b}_j$, the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is a matrix of order pq with eigenvalues $\alpha_i\beta_j$ and eigenvectors $\mathbf{a}_i \otimes \mathbf{b}_j$. We note that by these definitions, if $\mathbf{A}$ and $\mathbf{B}$ are stochastic, than also $\mathbf{A} \otimes \mathbf{B}$ is stochastic. This suggest the following compositional method for definining $\boldsymbol{P}$. We set $\boldsymbol{P} = \boldsymbol{P}_1 \otimes \boldsymbol{P}_2$, being $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ two small stochastic matrices, so that we can place in $\boldsymbol{P}$ all eigenvalues γ_i^1 of $\boldsymbol{P}_1$ and γ_j^2 of $\boldsymbol{P}_2$. This also inserts a number of spurious eigenvalues $\gamma_i^1\gamma_j^2$, but these vanish quicker than γ_i^1 and γ_j^2 if one of the two eigenvalues is not too big (e.g., $\gamma_j^2 < 0.9$), otherwise $\gamma_i^1\gamma_j^2 \approx \gamma_i^1$ which reinforces the contribute of γ_i^1 (this can also be adjusted by the related $A_{j,t}$ constant). Using Kronecker products, quite complex acf structures can be defined, e.g., by multiple Kronecker products $\boldsymbol{P} = \boldsymbol{P}_1 \otimes \boldsymbol{P}_2 \otimes \cdots \otimes \boldsymbol{P}_k$ or by increasing the order of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$. Note also that the steady-state vector $\boldsymbol{\pi}_e$ is simply the Kronecker product of the steady-state vectors of the defining matrices $\boldsymbol{P}_k$. An example of Kronecker-based process synthesis is given in the next section.

V. Network Model Robustness Analysis

To illustrate our process design methodology, we study a network model response to different temporal dependencies in workloads. We consider a MAP(4) with $\boldsymbol{P} = \boldsymbol{P}_1 \otimes \boldsymbol{P}_2$,

$$\boldsymbol{P}_1 = \begin{bmatrix} a+\alpha & 1-a-\alpha \\ a & 1-a \end{bmatrix}, \boldsymbol{P}_2 = \begin{bmatrix} b+\beta & 1-b-\beta \\ b & 1-b \end{bmatrix},$$

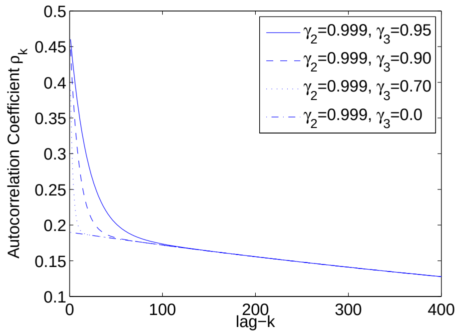

Fig. 2. MAP(4) processes with different acf at low lags.

where $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ have real eigenvalues α and β, respectively; $\boldsymbol{P}$ has thus eigenvalues $\gamma_1 = 1$, $\gamma_2 = \alpha$, $\gamma_3 = \beta$, $\gamma_4 = \alpha\beta$. Without loss of generality, we assume $\gamma_2 \geq \gamma_3 \geq \gamma_4$. We investigate two simple closed queueing networks with two or three queues in series representing different tiers, and we investigate how acf at the first tier impacts on network performance. Tier 1 is modelled as queue with the MAP(4) service process; the remaining tiers are each modeled with a queue having exponential service rate $\mu_2 = \mu_3 = 1$ job/sec.

To have a clear interpretation of results, we fix the moments of the MAP(4) and vary only the acf function. In particular, we leave sufficient degrees of freedom to fit two moments and we keep fixed the higher ones by setting

$$\theta_t = \epsilon_t, \quad t \geq 2,$$

where $0 \leq \epsilon_t \leq \epsilon$, and ϵ is an arbitrarily small constant so that $E[X^k] \to M_{1,1}\theta_1^k$ for $\epsilon \to 0$, which stays constant on all higher moments once assigned $E[X]$ and c_v^2. In the limit $\epsilon \to 0$, also (9)-(10) assume simple expressions that are easy to be inverted analytically for $\alpha, \beta, a, b, \theta_1$, e.g., using symbolic algebra in Maple or Mathematica.We assign these parameters to set $E[X]$ and c_v^2, and to impose the acf asymptote $\rho_k \sim A_{2,1}\gamma_2^k$ as well as γ_3 which is the main determinant of the decay rate at low lags. The family of processes generated in this way is plotted in Figure 2; for all processes $E[X] = 1$, $c_v^2 = 20$ and the acf has oblique asymptote $\rho_k \sim 0.4\rho_0\gamma_2^k$.

Table II shows exact global balance results for network throughput under the different acfs. In all experiments the network population is set to $N = 50$ jobs. The results stress the importance of accounting for non-renewal features, since up to 35% of the throughput can be affected by acf at a single tier. Furthermore, it is quite surprising to observe that low lag deviations from the acf asymptote do not seems responsible for any significant performance degradation in this model, suggesting that medium and high low-lag acf values may have similar performance impact. Observations of this type are impossible using MMPP(2)s or MAP(2)s, and promote our process synthesis methodology to improve understanding of system response to non-renewal workloads.

TABLE II
NETWORK THROUGHPUT [JOB/SEC] UNDER VARYING AUTOCORRELATION AT TIER 1

TIER 1 AUTOCORRELATION	2 TIERS	3 TIERS
$\gamma_2 = 0.000, \gamma_3 = 0.000$	0.762 (renewal)	0.261 (renewal)
$\gamma_2 = 0.700, \gamma_3 = 0.000$	0.701 (-8.0%)	0.254 (-2.7%)
$\gamma_2 = 0.900, \gamma_3 = 0.000$	0.621 (-18.5%)	0.237 (-9.2%)
$\gamma_2 = 0.950, \gamma_3 = 0.000$	0.577 (-24.2%)	0.226 (-13.4%)
$\gamma_2 = 0.999, \gamma_3 = 0.000$	0.507 (-33.4%)	0.211 (-19.2%)
$\gamma_2 = 0.999, \gamma_3 = 0.700$	0.506 (-33.5%)	0.211 (-19.2%)
$\gamma_2 = 0.999, \gamma_3 = 0.900$	0.502 (-34.1%)	0.211 (-19.2%)
$\gamma_2 = 0.999, \gamma_3 = 0.950$	0.496 (-34.9%)	0.209 (-19.9%)

VI. CONCLUSION

We have proposed a spectral characterization of moments and acf that significantly simplifies the analysis of MAP processes. Our method finds natural application in process characterization and synthesis. Ongoing work include the application of spectral-based process synthesis to evaluate network response to correlation in inter-arrival (or service) times, and the impact of the result on capacity planning.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under grant ITR-0428330.

REFERENCES

[1] A.T. Andersen, B.F. Nielsen. A Markovian approach for Modeling Packet traffic with Long-Range Dependence. *IEEE JSAC*, 16(5), 719–732, 1998.
[2] P. Buchholz, A. Panchenko. A Two-Step EM Algorithm for MAP Fitting. in *Proc of ISCIS 2004*, 217–227, 2004, LNCS 3280, Springer.
[3] J.W. Brewer. Kronecker Products and Matrix Calculus in System Theory. IEEE Tran. Circuits and Systems, 25(9), 1978.
[4] M.P. Drazin, E.V. Haynsworth. Criteria for the reality of matrix eigenvalues. *Math. Z.*, (78):449–452, 1962.
[5] A. Feldmann, W. Whitt Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. PEVA, 31(4), 245–279, 1998.
[6] A. Heindl, K. Mitchell, A. van de Liefvoort. *Correlation bounds for second-order MAPs with application to queueing network decomposition*. PEVA, 63(6), 553–577, 2006.
[7] F. E. Hohn. *Elementary Linear Algebra*. Dover, NY, USA, 1973.
[8] G. Horwath, P. Bucholz, M. Telek. A MAP fitting approach with independent approximation of the inter-arrival time distribution and the lag correlation. in *Proc. of QEST 2005*, 124–133, 2005, IEEE Press.
[9] S. H. Kang, Y. H. Kim, D. K. Sung, B. D. Choi. An Application of Markovian Arrival Process (MAP) to Modeling Superposed ATM Cell Streams *IEEE Trans. on Communications*, 50(4), 633–642, 2002.
[10] L.A. Kulkarni, S. Li. Transient behaviour of queueing systems with correlated traffic. *Perform. Eval*, 27-28, 117–145, 1996.
[11] N. Mi, Qi Zhang, A. Riska, E.Smirni, E. Riedel Performance Impacts of Autocorrelation Flows in Systems. To appear in Performance 2007.
[12] M.F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and Their Applications*. Marcel Dekker, New York, 1989.
[13] The Internet traffic archive. http://ita.ee.lbl.gov/index.html.
[14] V. Paxson, S. Floyd. Wide area traffic: the failure of Poisson modelling. *IEEE/ACM Trans. on Networking*, 226–244, 3(3), 1995.
[15] A. Riska, E. Riedel. Disk Drive Level Workload Characterization. in *Proc. of USENIX Annual Technical Conference 2006*, 97–102, 2006.
[16] K. Sriram, W. Whitt Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data. *IEEE JSAC*, vol. SAC-4, No. 6, September 1986, pp. 833–846.
[17] M. Telek, G. Horwath A minimal representation of Markov arrival processes and a moments matching method. To appear in Performance 2007.
[18] E.Z. Zhang, G. Casale, E. Smirni. Analytical Feasible Region for Eigenvalue Based MAP(2) Fitting, Technical Report TR 2007-08, College of William and Mary, 2007.