UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 09-71
Statistics and Econometrics Series 21
November 2009

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

NON-IDENTIFIABILITY OF THE TWO STATE MARKOVIAN ARRIVAL PROCESS

Pepa Ramírez[1], Rosa E. Lillo[2] and Michael P. Wiper[3]

Abstract

In this paper we consider the problem of identifiability of the two-state Markovian Arrival process (MAP_2). In particular, we show that the MAP_2 is not identifiable and conditions are given under which two different sets of parameters, induce identical stationary laws for the observable process.

Keywords: Batch Markovian Arrival process, Markov Renewal process, Hidden Markov models, Identifiability problems.

[1] Corresponding author: Pepa Ramírez, Département Images et Signaux, UMR CNRS 5216, 961 rue de la Houille Blanche, 38402 Saint Martin d'Hères Cedex, France.
E-mail: Pepa.Ramirezcobo@gipsa-lab.inpg.fr
[2] Rosa E. Lillo, Universidad Carlos III, Departamento de Estadística, e-mail: lillo@est-econ.uc3m.es
[3] Michael P. Wiper, Universidad Carlos III, Departamento de Estadística,e-mail: mwiper@est-econ.uc3m.es

1. Introduction

The *Markovian arrival process* (*MAP*) was defined in [10] and [9] as a generalization of the Poisson arrival process allowing for both dependence between arrivals and non-exponentially distributed interarrival times. The *MAP* is defined by two Markov processes: the first counts the number of arrivals and the second, an underlying Markov process, governs the state changes. At the end of a transition in a *MAP* an arrival may or may not occur and although the transition holding times are exponentially distributed, the interarrival times do not follow an exponential distribution. Special cases of the *MAP* are the phase type renewal processes (which include both Erlang and hyperexponential renewal process) and non-renewal processes such as the Markov modulated Poisson process (*MMPP*). Stationary *MAP*s are dense in the family of all stationary point processes; see [1]. Another important property of *MAP*s is that the superposition of independent *MAP*s is again a *MAP*.

The *MAP* is an challenging process from both a theoretical and applied points of view. From a theoretical perspective, the queueing system where the *MAP* governs the arrival process has been widely studied in the literature, combined with matrix analytic methods (see for example [8]). On the other hand, the *MAP* has been proposed in the literature as a suitable process for modeling teletraffic data, see e.g. [4], [5], [11], [6], and [17]. In this case, the *MAP* is used to fit data where only the interarrival times are observed and neither the underlying Markov chain nor the individual exponential holding times are available and thus, the observed arrival process is a hidden Markov process.

When dealing with inference for hidden Markov processes, it is very common to encounter *identifiability* problems which imply that the likelihood function does not possess a unique maximum. Identifiability conditions for general, hidden Markov processes are studied in [7] and [13]. Identifiability of the *MMPP*, was undertaken in [15]. However, for the *MAP*, to the best of our knowledge, the identifiability problem is still essentially unsolved.

In this paper we address the problem of identifiability of the two-state *MAP*, or MAP_2. We conclude that, on the contrary to the *MMPP*, which is identifiable (see [15]), the MAP_2 is not identifiable.

The paper is organized as follows. In Section 2 we introduce the MAP and its main properties. In Section 3, we study when two MAP_2s have the same interarrival time distributions, a necessary condition for non-identifiability. We call this property *weak equivalence*. In Section 4 we consider the joint distribution of a sequence of interarrival times generated from the MAP_2 and show that there are at least two different parameterizations of the MAP_2 giving rise to the same joint distribution, thus proving the non-identifiability of the MAP_2. Finally in Section 5 we provide conclusions and various possible extensions of this work.

2. The *MAP* and its main properties

Consider an irreducible, continuous, Markov process $J(t)$ with state space $\mathcal{S} = \{1, \ldots, m\}$ and generator matrix D. Let $N(t)$ represent the cumulative number of arrivals in $(0, t]$. A MAP, represented by $\{J(t), N(t)\}$ behaves as follows: the initial state $i_0 \in \mathcal{S}$ is generated according to an initial probability vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ and at the end of an exponentially distributed sojourn time in state i, with mean $1/\lambda_i$, two possible state transitions can occur. Firstly, with probability $0 \leq p_{ij1} \leq 1$ a single arrival occurs and the MAP enters a state $j \in \mathcal{S}$, which may be the same as $(j = i)$ or different from $(j \neq i)$ the previous state. Secondly, with probability $0 \leq p_{ij0} \leq 1$, no arrival occurs and the MAP enters a different state $j \neq i$. Given that from all states a transition must occur to a different state without an arrival or to any state with an arrival, then for $1 \leq i \leq m$,

$$\sum_{j=1, j\neq i}^{m} p_{ij0} + \sum_{j=1}^{m} p_{ij1} = 1.$$

When $m = 2$, we have a two state MAP, denoted by MAP_2. Figure 1 illustrates the different movements that can occur transitions that in this process by means of a transition diagram.

Define the matrices $P_0 = (p_{ij0})_{i,j\in\mathcal{S}}$ and $P_1 = (p_{ij1})_{i,j\in\mathcal{S}}$ where $p_{ii0} = 0$. Then the MAP is defined by the set $\{\boldsymbol{\theta}, \boldsymbol{\lambda}, P_0, P_1\}$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$. Alternatively, the MAP can be characterized by the rate matrices, $D_0 = (d_{ij0})_{i,j\in\mathcal{S}}$ and $D_1 = (d_{ij1})_{i,j\in\mathcal{S}}$, where $d_{ii0} = -\lambda_i$, $d_{ij0} = \lambda_i p_{ij0}$, for $j \neq i$ and $d_{ij1} = \lambda_i p_{ij1}$, for $1 \leq i, j \leq m$. This definition implies that $D \equiv D_0 + D_1$ is the infinitesimal generator of the underlying

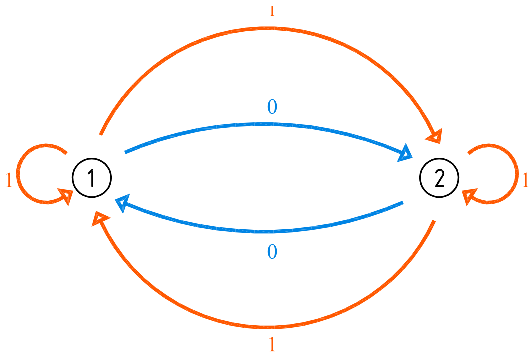

FIGURE 1: Transitions diagram for the MAP_2. 0 and 1 illustrate moves without and with arrivals respectively.

Markov process. Intuitively, D_0 can be thought as governing transitions that do not generate arrivals and D_1 can be thought as governing transitions that do generate arrivals. The stationary probability vector of the Markov process with generator D is $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_m)$, which satisfies

$$\boldsymbol{\pi} D = \mathbf{0}, \quad |\boldsymbol{\pi}| = 1, \tag{2.1}$$

where $|\mathbf{x}|$ denotes the sum of values of vector $\mathbf{x}$. Thus, π_j represents the stationary probability that the process is in state j, for $j = 1, \ldots, m$.

The Markov modulated Poisson process (*MMPP*) can be defined as a simplified *MAP* where the matrix D_1 (and thus, P_1) is diagonal (see [8]). This implies that arrivals can only occur in transitions to the same state. Some important properties of the *MAP* are as follows. Firstly, it is known ([2]) that the *MAP* can be regarded as a Markov renewal process. Let X_n be the state of the *MAP* at the time of the nth arrival, and let T_n be the time between the $(n-1)$st and nth arrival, then $\{X_{n-1}, T_n\}_{n=1}^{\infty}$ is a Markov renewal process with semi-Markovian kernel given by

$$\int_0^t e^{D_0 t} D_1 dt = (I - e^{D_0 t})(-D_0)^{-1} D_1. \tag{2.2}$$

Therefore, $\{X_n\}_{n=1}^{\infty}$ is a Markov chain, where from (2.2) the transition matrix can be computed as

$$P^{\star} = (I - P_0)^{-1} P_1 \tag{2.3}$$

and it can be shown that the stationary distribution, $\boldsymbol{\phi}$, is given by

$$\boldsymbol{\phi} = (\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} D_1. \tag{2.4}$$

See the Appendix A for a proof.

Secondly, let the random variable T_1 denote the time to the first arrival in a *MAP*. Then, from ([2]), the cumulative distribution function (cdf) of T_1 is given by

$$F_{T_1}(t) = \boldsymbol{\theta}(I - e^{D_0 t})(-D_0)^{-1} D_1 \mathbf{e}, \quad \text{for } t \geq 0,$$

where $\mathbf{e}$ is an m column vector of ones. If T represents the stationary interarrival time distribution, it can be found that

$$F_T(t) = \boldsymbol{\phi}(I - e^{D_0 t})(-D_0)^{-1} D_1 \mathbf{e}, \quad \text{for } t \geq 0. \tag{2.5}$$

Finally, the Laplace Stieltjes transform of the interarrival time distribution of a stationary *MAP* is given by

$$f^*_{T;D_0,D_1}(s_1, \ldots, s_n) = \boldsymbol{\phi}(s_1 I - D_0)^{-1} D_1 \ldots (s_n I - D_0)^{-1} D_1 \mathbf{e},$$

or equivalently,

$$f^*_{T;D_0,D_1}(s_1, \ldots, s_n) = \boldsymbol{\phi} \prod_{i=1}^{n} \Delta(s_i) \mathbf{e}, \tag{2.6}$$

where

$$\Delta(s) = (sI - D_0)^{-1} D_1. \tag{2.7}$$

For a more detailed description and further properties of the *MAP* (and *BMAP*) see e.g. [8] or [2].

3. *Weak* equivalence

There have been a number of examples of fitting the *MMPP* to internet data traces. In most applications, the two-state case has been considered (see for example [16], [14] or [3]). The *MMPP*s, despite being simplified *MAP*s (the matrix P_1 is diagonal, and thus they are characterized by two less parameters), are complex processes and usually, two states, at most three, are enough to capture the data behavior.

From now on we consider the two-state *MAP* or MAP_2, characterized by $\mathcal{M} \equiv \{\boldsymbol{\theta}, D_0, D_1\}$ where

$$\boldsymbol{\theta} = (\theta,\ 1-\theta), \quad D_0 = \begin{pmatrix} x & y \\ z & u \end{pmatrix}, \quad D_1 = \begin{pmatrix} w & -x-y-w \\ v & -z-u-v \end{pmatrix}, \tag{3.1}$$

and

$$x = -\lambda_1, \quad y = \lambda_1 p_{120}, \quad w = \lambda_1 p_{111},$$

$$z = \lambda_2 p_{210}, \quad u = -\lambda_2, \quad v = \lambda_2 p_{211}.$$

The stationary probability $\boldsymbol{\phi} = (\phi, 1-\phi)$ (2.4) can be found to be

$$\phi = \frac{wz - vx}{wz - vx - zy - vy + xu + wu}. \tag{3.2}$$

When modeling real data, usually just the times between arrivals are observed, and thus the interest when making inference for the *MAP* is focused on the embedded Markov renewal process $\{X_{n-1}, T_n\}_{n=1}^{\infty}$. As a preliminary step to studying the identifiability of the MAP_2, we study the conditions under which two MAP_2s possess the same marginal interarrival time distributions. For two such MAP_2s, we shall say that they are weakly equivalent.

Definition 3.1. Let $\mathcal{M}$ represent a MAP_2 with parameters $\{\boldsymbol{\theta}, D_0, D_1\}$ as in (3.1). Then we say that another MAP_2, $\widetilde{\mathcal{M}} \equiv \{\widetilde{\boldsymbol{\theta}}, \widetilde{D}_0, \widetilde{D}_1\}$ is weakly equivalent to $\mathcal{M}$ if and only if

$$T_n \stackrel{d}{=} \widetilde{T}_n, \quad \text{for all } n \geq 1, \tag{3.3}$$

where T_n and $\widetilde{T}_n$ represent the times between the $(n-1)$th and nth arrivals under both models.

The term *weak* is employed because equivalence is expressed in a marginal sense. Since the interarrival times in a MAP_2 are not independent, condition (3.3) is necessary but insufficient for non-identifiability. A general condition looking at the joint density of the sequence of interarrival times is analyzed in Section 4.

Given a known MAP_2, $\mathcal{M}$ as in (3.1), define the constant

$$c = z + u - x - y. \tag{3.4}$$

Notice that if $c = 0$, then the rate until an arrival occurs from state 1 coincides with that of state 2. This implies that the observable process (that where arrivals occur) behaves like a Poisson process, with a single arrival rate. Thus, we will assume that $c \neq 0$. In addition, define the matrix $\mathbf{\Phi}$ as that whose rows are composed by the stationary vector $\boldsymbol{\phi}$. Suppose that $P^{\star} = \mathbf{\Phi}$, then it is immediately clear that there are many weakly equivalent MAP_2s for example, any $\widetilde{\mathcal{M}} = \{\widetilde{\theta}, D_0, D_1\}$ is equivalent to $\mathcal{M} = \{\theta, D_0, D_1\}$. We can thus assume too, that $P^{\star} \neq \mathbf{\Phi}$.

Theorem 3.1 gives the conditions for $\widetilde{\mathcal{M}}$ to be weakly equivalent to $\mathcal{M}$.

Theorem 3.1. *Let $\mathcal{M}$ be a MAP_2 as in (3.1) with stationary distribution $\boldsymbol{\phi}$. Given another MAP_2, $\widetilde{\mathcal{M}}$ with stationary distribution $\widetilde{\boldsymbol{\phi}}$ assume that*

A1. $c \neq 0$, $\tilde{c} \neq 0$.

A2. $P^{\star} \neq \mathbf{\Phi}$ *or* $\widetilde{P}^{\star} \neq \widetilde{\mathbf{\Phi}}$.

Then, $\widetilde{\mathcal{M}}$ is weakly equivalent to $\mathcal{M}$ if and only if

B1. $T \stackrel{d}{=} \widetilde{T}$, *and*

B2. $(\boldsymbol{\theta}, \widetilde{\boldsymbol{\theta}}) = (\boldsymbol{\phi}, \widetilde{\boldsymbol{\phi}})$.

The variables T and $\widetilde{T}$ in condition *B1.* represent the interarrival times in the stationary versions of $\mathcal{M}$ and $\widetilde{\mathcal{M}}$. As will be shown in the proof of Theorem 3.1 (see the Appendix B), *B1.* is equivalent to the equality of two rational functions,

$$\frac{\alpha s + \gamma}{s^2 + \beta s + \gamma} = \frac{\tilde{\alpha} s + \tilde{\gamma}}{s^2 + \tilde{\beta} s + \tilde{\gamma}}, \quad \text{for all } s, \tag{3.5}$$

where the coefficients α, β and γ (respectively $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\gamma}$) are given by

$$\begin{aligned} \alpha &= \theta(z + u - x - y) - (z + u), \\ \beta &= -x - u, \\ \gamma &= xu - yz. \end{aligned} \tag{3.6}$$

Theorem 3.1 thus provides a straightforward way to check if two given MAP_2s share the same interarrival distribution.

Example 3.1. As an example, let us consider the MAP_2 defined by

$$D_0 = \begin{pmatrix} -20 & 6 \\ 0.15 & -0.5 \end{pmatrix}, \qquad D_1 = \begin{pmatrix} 12.228 & 1.772 \\ 0.0426 & 0.3074 \end{pmatrix},$$

or alternatively,

$$P_0 = \begin{pmatrix} 0 & 0.3 \\ 0.3 & 0 \end{pmatrix}, \qquad P_1 = \begin{pmatrix} 0.6114 & 0.0886 \\ 0.0852 & 0.6148 \end{pmatrix},$$

and exponential rates $(\lambda_1, \lambda_2) = (20, 0.5)$. Suppose that the initial probability vector is equal to the stationary distribution, $\boldsymbol{\theta} = \boldsymbol{\phi} = (0.496, 0.504)$. The transition probability matrix, $P^\star$ is computed from (2.3):

$$P^\star = \begin{pmatrix} 0.7 & 0.3 \\ 0.2952 & 0.7048 \end{pmatrix} \neq \boldsymbol{\Phi}.$$

Consider another MAP_2 with parameters

$$\widetilde{D}_0 = \begin{pmatrix} -19.7 & 10.835 \\ 0.6146 & -0.8 \end{pmatrix}, \qquad \widetilde{D}_1 = \begin{pmatrix} 7.1452 & 1.7198 \\ 0.1443 & 0.0411 \end{pmatrix},$$

or alternatively,

$$\widetilde{P}_0 = \begin{pmatrix} 0 & 0.5500 \\ 0.7682 & 0 \end{pmatrix}, \qquad \widetilde{P}_1 = \begin{pmatrix} 0.3627 & 0.0873 \\ 0.1804 & 0.0514 \end{pmatrix},$$

and exponential rates $(\widetilde{\lambda}_1, \widetilde{\lambda}_2) = (19.7, 0.8)$. Assume that $\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\phi}} = (0.799, 0.201)$. The transition probability matrix, $\widetilde{P}^\star$ is

$$\widetilde{P}^\star = \begin{pmatrix} 0.8 & 0.2 \\ 0.79 & 0.21 \end{pmatrix} \neq \boldsymbol{\Phi}.$$

It can be seen that $c = 13.65 \neq 0$ and $\tilde{c} = -8.6796 \neq 0$ and (3.5) holds. Therefore, from Theorem 3.1, the processes are weakly equivalent, as shown in Figure 2, which depicts the cdf of the time between two arrivals in the stationary version for both MAP_2s.

More results similar to Theorem 3.1, when the assumptions *A1.* and *A2.* are relaxed (and thus the number of weakly equivalent MAP_2s to a fixed MAP_2, increases), and extensions to the three-state case MAP can be found in [12].

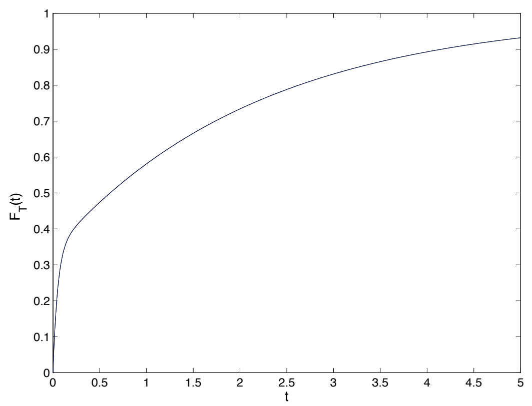

FIGURE 2: CDF of T, time until next arrival in the stationary version, in the Example 1. As $\boldsymbol{\theta} = \boldsymbol{\phi}$, then $T \stackrel{d}{=} T_1$ (similarly, $\widetilde{T} \stackrel{d}{=} \widetilde{T}_1$), and thus $T \stackrel{d}{=} T_1 \stackrel{d}{=} \widetilde{T} \stackrel{d}{=} \widetilde{T}_1$.

4. Non-identifiability of the MAP_2

In this section we now prove that the MAP_2 is a non-identifiable process. Following Theorem 3.1, all MAP_2s will be assumed to be stationary and $\mathcal{M}$ ($\widetilde{\mathcal{M}}$) will denote the set $\{\boldsymbol{\phi}, D_0, D_1\}$ ($\{\tilde{\boldsymbol{\phi}}, \tilde{D}_0, \tilde{D}_1\}$) from now on. Our definition of non-identifiability follows [15],

Definition 4.1. The MAP_2 is a non-identifiable process if for any fixed MAP_2, $\mathcal{M}$, then there exists another MAP_2, $\widetilde{\mathcal{M}}$ such that

$$(T_1, \ldots, T_n) \stackrel{d}{=} (\widetilde{T}_1, \ldots, \widetilde{T}_n), \quad \text{for all } n \geq 1. \tag{4.1}$$

Note that condition (4.1) is equivalent to the equality of the Laplace transforms,

$$f^*_{T;D_0,D_1}(s_1, \ldots, s_n) = f^*_{\widetilde{T};\widetilde{D}_0,\widetilde{D}_1}(s_1, \ldots, s_n), \tag{4.2}$$

for all n, s. We will show that given a MAP_2, $\mathcal{M}$, as in (3.1), then there always exists a differently parameterized $\widetilde{\mathcal{M}}$ such that (4.2) holds for all n, s. Indeed, we will prove that if (4.2) holds for $n = 1, 2$, then it will hold for all n.

Let us first consider the following result which gives the conditions under which equations (4.2) hold for $n = 1, 2$.

Proposition 4.1. *Let $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ be two MAP_2s. Let α, β, γ, δ_1, δ_2 ($\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\gamma}$, $\tilde{\delta}_1$, $\tilde{\delta}_2$)*

be defined as

$$\begin{aligned}
\alpha &= \phi(z+u-x-y)-(z+u),\\
\beta &= -x-u, \qquad\qquad (4.3)\\
\gamma &= xu-yz,\\
\delta_1 &= \phi\left((z+u-x-y)(w-v)+(x+y)(z+u)-(z+u)^2\right)+\\
&\quad (z+u-x-y)v+(z+u)^2,\\
\delta_2 &= \phi(x+y-z-u)(uw-yv-xv+zw)+(x+y-z-u)(xv-zw)-(u+z)\gamma\\
&\quad (x+y-z-u)(xv-zw)-(u+z)\gamma.
\end{aligned}$$

Then, if

$$\tilde{\alpha}=\alpha,\quad \tilde{\beta}=\beta,\quad \tilde{\gamma}=\gamma,\quad \tilde{\delta}_1=\delta_1,\quad \tilde{\delta}_2=\delta_2, \tag{4.4}$$

the equality of Laplace transforms (4.2) holds for all s and for $n=1,2$.

For a proof of Proposition 4.1 see the Appendix C. The following result gives the solutions to equations (4.4).

Proposition 4.2. *Consider a MAP$_2$ as in (3.1). For all $\tilde{u}<0$ and all $\tilde{z}>0$, let $\tilde{x}(\tilde{u},\tilde{z})$, $\tilde{y}(\tilde{u},\tilde{z})$, $\tilde{v}(\tilde{u},\tilde{z})$ and $\tilde{w}(\tilde{u},\tilde{z})$ be defined as*

$$\tilde{x}(\tilde{u},\tilde{z}) = -\tilde{u}+x+u, \tag{4.5}$$

$$\tilde{y}(\tilde{u},\tilde{z}) = -(\tilde{u}^2-\tilde{u}x-\tilde{u}u+xu-zy)/\tilde{z}, \tag{4.6}$$

$$\begin{aligned}
\tilde{v}(\tilde{u},\tilde{z}) = {} & \frac{-\tilde{z}(vx+vy-wz-wu+w\tilde{z}-z\tilde{z}-z\tilde{u})}{(-\tilde{u}u-u\tilde{z}+xu+\tilde{u}^2+2\tilde{u}\tilde{z}+\tilde{z}^2-\tilde{u}x-\tilde{z}x-zy)}+\\
& \frac{-\tilde{z}(-\tilde{u}u+\tilde{z}^2+2\tilde{u}\tilde{z}-u\tilde{z}-v\tilde{z}-\tilde{u}v+w\tilde{u}+\tilde{u}^2)}{(-\tilde{u}u-u\tilde{z}+xu+\tilde{u}^2+2\tilde{u}\tilde{z}+\tilde{z}^2-\tilde{u}x-\tilde{z}x-zy)},
\end{aligned} \tag{4.7}$$

$$\begin{aligned}
\tilde{w}(\tilde{u},\tilde{z}) = {} & \frac{z\tilde{u}x+2\tilde{z}xu-\tilde{z}zy-xu^2+zx\tilde{z}+zu\tilde{z}+\tilde{u}u^2+uzy}{(-\tilde{u}u-u\tilde{z}+xu+\tilde{u}^2+2\tilde{u}\tilde{z}+\tilde{z}^2-\tilde{u}x-\tilde{z}x-zy)}+\\
& \frac{-zxu-2\tilde{z}\tilde{u}x+z^2y+u\tilde{z}v-wzy+zvy-\tilde{z}^2u-3\tilde{u}\tilde{z}u-\tilde{u}\tilde{z}v}{(-\tilde{u}u-u\tilde{z}+xu+\tilde{u}^2+2\tilde{u}\tilde{z}+\tilde{z}^2-\tilde{u}x-\tilde{z}x-zy)}+\\
& \frac{v\tilde{u}x-\tilde{z}vy+2\tilde{u}xu+v\tilde{u}u-w\tilde{u}x-wu\tilde{u}+2\tilde{u}^2\tilde{z}+z\tilde{u}u+wux}{(-\tilde{u}u-u\tilde{z}+xu+\tilde{u}^2+2\tilde{u}\tilde{z}+\tilde{z}^2-\tilde{u}x-\tilde{z}x-zy)}+,\\
& \frac{\tilde{z}^2\tilde{u}-v\tilde{u}^2-\tilde{u}^2x+\tilde{z}wz+\tilde{u}\tilde{z}w-\tilde{u}\tilde{z}z-vxu+w\tilde{u}^2-z\tilde{u}^2+\tilde{u}^3}{(-\tilde{u}u-u\tilde{z}+xu+\tilde{u}^2+2\tilde{u}\tilde{z}+\tilde{z}^2-\tilde{u}x-\tilde{z}x-zy)}+\\
& \frac{-zy\tilde{u}+\tilde{z}u^2-\tilde{z}^2x-2\tilde{u}^2u-\tilde{z}wx}{(-\tilde{u}u-u\tilde{z}+xu+\tilde{u}^2+2\tilde{u}\tilde{z}+\tilde{z}^2-\tilde{u}x-\tilde{z}x-zy)}.
\end{aligned} \tag{4.8}$$

Then, the set $\{\tilde{u}, \tilde{z}, \tilde{x}(\tilde{u}, \tilde{z}), \tilde{y}(\tilde{u}, \tilde{z}), \tilde{v}(\tilde{u}, \tilde{z}), \tilde{w}(\tilde{u}, \tilde{z})\}$ *solves the system of equations given by (4.4).*

The proof of Proposition 4.2 is tedious but straightforward, solving the system of equations (4.4) by conventional methods, and substituting ϕ and $\tilde{\phi}$ from their definitions (3.2). Although Proposition 4.2 gives infinite solutions to the system of equations (4.4), a priori the values of $\tilde{x}$, $\tilde{y}$, $\tilde{v}$ and $\tilde{w}$ may not define a MAP_2. The following theorem shows how to select feasible values of $\tilde{u}$ and $\tilde{z}$ in the sense that $\tilde{x} < 0$, $\tilde{y}, \tilde{v}, \tilde{w} > 0$ and $\tilde{\phi} \in [0, 1]$, that is, provides the way to choose $\tilde{u}$ and $\tilde{z}$ so that they make $\widetilde{\mathcal{M}}$ equivalent to $\mathcal{M}$.

Theorem 4.1. *Define a* MAP_2, $\mathcal{M}$ *as in (3.1), where it is assumed that* $x < u$. *Let* ε *be chosen from*

$$0 < \varepsilon < min \left\{ -x, \frac{u-x}{2}, \frac{z(1-\phi)}{\phi}, \frac{(u-x) + \sqrt{(x-u)^2 + 4zy}}{2} \right\}, \tag{4.9}$$

and define $\tilde{u} \equiv u - \varepsilon$ *and* $\tilde{z} \equiv z + \varepsilon$. *Then there exist an infinite number of* MAP_2*s,* $\widetilde{\mathcal{M}}$, *given by* $\mathcal{F} = \{\tilde{u}, \tilde{z}, \tilde{x}(\tilde{u}, \tilde{z}), \tilde{y}(\tilde{u}, \tilde{z}), \tilde{v}(\tilde{u}, \tilde{z}), \tilde{w}(\tilde{u}, \tilde{z})\}$, *where* $\tilde{x}(\tilde{u}, \tilde{z})$, $\tilde{y}(\tilde{u}, \tilde{z})$, $\tilde{v}(\tilde{u}, \tilde{z})$, *and* $\tilde{w}(\tilde{u}, \tilde{z})$ *are defined by (4.5-4.8), such that (4.4) holds.*

For a proof of Theorem 4.1, see the Appendix D.

As a consequence of Theorem 4.1 the following two corollaries may be derived.

Corollary 4.1. *Consider a* $MAP_2 s$ $\mathcal{M}$ *as in (3.1) and values* $\tilde{u}$, $\tilde{z}$, $\tilde{x}(\tilde{u}, \tilde{z})$, $\tilde{y}(\tilde{u}, \tilde{z})$, $\tilde{v}(\tilde{u}, \tilde{z})$, $\tilde{w}(\tilde{u}, \tilde{z})\}$ *as in Proposition 4.2 characterizing another* MAP_2, $\widetilde{\mathcal{M}}$. *Let* $\Delta(s)$ *be defined as in (2.7), (respectively,* $\widetilde{\Delta}(s)$*). Then,*

$$\boldsymbol{\phi}\Delta(s_1)\mathbf{e} = \tilde{\boldsymbol{\phi}}\widetilde{\Delta}(s_1)\mathbf{e}, \tag{4.10}$$

$$\boldsymbol{\phi}\Delta(s_1)\Delta(s_2)\mathbf{e} = \tilde{\boldsymbol{\phi}}\widetilde{\Delta}(s_1)\widetilde{\Delta}(s_2)\mathbf{e}, \tag{4.11}$$

$$(0,1)\Delta(s_1)\mathbf{e} = (0,1)\widetilde{\Delta}(s_1)\mathbf{e}, \tag{4.12}$$

$$(0,1)\Delta(s_1)\Delta(s_2)\mathbf{e} = (0,1)\widetilde{\Delta}(s_1)\widetilde{\Delta}(s_2)\mathbf{e}, \tag{4.13}$$

for all s_1, s_2.

From (2.6), expressions (4.10)-(4.11) are an alternative way to state that equation (4.2) holds for $n = 1$ and $n = 2$. By substituting the values of $\mathcal{F}$ found in Theorem 4.1 in the expression for $\widetilde{\Delta}(s)$ (2.7), routine, but tedious calculations yield to (4.12)-(4.13).

Corollary 4.2. *Consider a MAP_2s $\mathcal{M}$ as in (3.1) and values $\tilde{u}$, $\tilde{z}$, $\tilde{x}(\tilde{u},\tilde{z})$, $\tilde{y}(\tilde{u},\tilde{z})$, $\tilde{v}(\tilde{u},\tilde{z})$, $\tilde{w}(\tilde{u},\tilde{z})\}$ as in Proposition 4.2 characterizing another MAP_2, $\widetilde{\mathcal{M}}$. Let $\Delta(s)$ be defined as in (2.7), (respectively, $\widetilde{\Delta}(s)$). If $\Delta(s)$ is given by*

$$\Delta(s) = \begin{pmatrix} a(s) & b(s) \\ c(s) & d(s) \end{pmatrix}, \tag{4.14}$$

then the solution to (4.10)-(4.13) is

$$\widetilde{\Delta}(s) = \begin{pmatrix} \widetilde{a}(s) & \widetilde{b}(s) \\ \widetilde{c}(s) & \widetilde{d}(s) \end{pmatrix}, \tag{4.15}$$

where

$$\widetilde{a}(s) = \frac{\phi(a(s)-c(s)) + \widetilde{\phi}c(s)}{\phi}, \tag{4.16}$$

$$\widetilde{b}(s) = \frac{\phi\widetilde{\phi}\ (d(s)+2c(s)-a(s)) + \phi^2\ (a(s)-d(s)+b(s)-c(s)) - \widetilde{\phi}^2 c(s)}{\phi\widetilde{\phi}} \tag{4.17}$$

$$\widetilde{c}(s) = \frac{\widetilde{\phi}c(s)}{\phi} \tag{4.18}$$

$$\widetilde{d}(s) = \frac{\phi\ (c(s)+d(s)) - \widetilde{\phi}c(s)}{\phi}. \tag{4.19}$$

The equations (4.10)-(4.13) form a system with four equations where the unknowns are the elements of (4.15). A trivial verification shows that (4.16)-(4.19) solves the system. The previous Corollary motivates the following definition.

Definition 4.2. Let G and $\widetilde{G}$ be 2×2 matrices where

$$G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \widetilde{G} = \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix}.$$

It will be said that $\widetilde{G}$ is related to G, given the values ϕ and $\widetilde{\phi}$ if and only if

$$\tilde{a} = \frac{\phi(a-c) + \widetilde{\phi}c}{\phi}, \tag{4.20}$$

$$\tilde{b} = \frac{\phi\widetilde{\phi}\ (d+2c-a) + \phi^2\ (a-d+b-c) - \widetilde{\phi}^2 c}{\phi\widetilde{\phi}}, \tag{4.21}$$

$$\tilde{c} = \frac{\widetilde{\phi}c}{\phi}, \tag{4.22}$$

$$\tilde{d} = \frac{\phi\ (c+d) - \widetilde{\phi}c}{\phi}. \tag{4.23}$$

This relation will be noted by $G \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{G}$.

The following result is a direct consequence of the definition (4.20)-(4.23) of $\widetilde{G}$.

Proposition 4.3. *If $G \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{G}$, then*

$$\boldsymbol{\phi} G\mathbf{e} = \widetilde{\boldsymbol{\phi}}\widetilde{G}\mathbf{e},$$

where $\boldsymbol{\phi} = (\phi, 1-\phi)$, and $\widetilde{\boldsymbol{\phi}} = (\widetilde{\phi}, 1-\widetilde{\phi})$.

The proof is straightforward from the definition of $\widetilde{G}$ (4.20-4.23).

The next result, whose proof can be found in the Appendix E, is crucial for proving the non-identifiability of the MAP_2.

Proposition 4.4. *If $G \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{G}$, and $H \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{H}$, then*

$$GH \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{G}\widetilde{H}.$$

Finally, we can prove the general theorem.

Theorem 4.2. *The MAP_2 is not an identifiable process.*

Proof. The proof is based on the fact that given a MAP_2, $\mathcal{M}$ as in (3.1), then any other MAP_2, $\widetilde{\mathcal{M}}$ chosen from the set $\mathcal{F}$ (see Theorem 4.1) satisfies the equality (4.2), for all n.

If $\widetilde{\mathcal{M}}$ is selected from $\mathcal{F}$, then, from Corollary 4.1,

$$\Delta(s_1) \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{\Delta}(s_1) \quad \text{and} \quad \Delta(s_1)\Delta(s_2) \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{\Delta}(s_1)\widetilde{\Delta}(s_2), \quad \text{for all } s_1, s_2.$$

We conclude from Proposition 4.4 that

$$\Delta(s_1)\Delta(s_2)\Delta(s_3) \overset{\phi,\widetilde{\phi}}{\sim} \widetilde{\Delta}(s_1)\widetilde{\Delta}(s_2)\widetilde{\Delta}(s_3), \quad \text{for all } s_1, s_2, s_3,$$

and finally

$$\prod_{i=1}^{n} \Delta(s_i) \overset{\phi,\widetilde{\phi}}{\sim} \prod_{i=1}^{n} \widetilde{\Delta}(s_i),$$

which by Proposition (4.3) is equivalent to (4.2), for all $n \geq 1$.

To illustrate Theorem 4.2, we provide an example which illustrates two MAP_2s that verify (4.2), for all $n \geq 1$.

Example 4.1. Consider the MAP_2 defined by

$$D_0 = \begin{pmatrix} -10 & 2.5 \\ 0.6 & -3 \end{pmatrix}, \qquad D_1 = \begin{pmatrix} 4 & 3.5 \\ 1.35 & 1.05 \end{pmatrix},$$

or alternatively,

$$P_0 = \begin{pmatrix} 0 & 0.25 \\ 0.2 & 0 \end{pmatrix}, \qquad P_1 = \begin{pmatrix} 0.4 & 0.35 \\ 0.45 & 0.35 \end{pmatrix},$$

with exponential rates $(\lambda_1, \lambda_2) = (10, 3)$, and stationary probability $\boldsymbol{\phi} = (0.5474, 0.4526)$.

It can be seen that ε, defined in (4.9), has to be chosen from

$$0 < \varepsilon < min\{10, 3.5, 0.4961, 7.2081\}.$$

Let $\varepsilon = 0.3$. Then, according to the equations given in Theorem 4.1,

$$\widetilde{D}_0 = \begin{pmatrix} -9.7 & 3.9 \\ 0.9 & -3.3 \end{pmatrix}, \qquad \widetilde{D}_1 = \begin{pmatrix} 4.675 & 1.1250 \\ 2.025 & 0.375 \end{pmatrix},$$

or

$$\widetilde{P}_0 = \begin{pmatrix} 0 & 0.4020 \\ 0.2727 & 0 \end{pmatrix}, \qquad \widetilde{P}_1 = \begin{pmatrix} 0.4820 & 0.1160 \\ 0.6137 & 0.1136 \end{pmatrix},$$

with exponential rates $(\widetilde{\lambda}_1, \widetilde{\lambda}_2) = (9.7, 3.3)$ and stationary distribution $\widetilde{\boldsymbol{\phi}} = (0.8217, 0.1783)$. Because of the proof of Theorem (4.2), (4.2) holds for all s, and n, that is both MAP_2s will possess the same joint interarrival time distribution.

Three remarks need to be made here. First, let us point out that as has been described in the previous proof, given a fixed MAP_2, $\mathcal{M}$, then any other MAP_2, $\widetilde{\mathcal{M}}$ chosen from the set $\mathcal{F}$, will verify (4.2) for all s and n, and thus both MAP_2s will have the same joint interarrival time distribution.

Our second remark is connected with the *MMPP*. Its definition implies that $w \equiv -x - y$, and $v \equiv 0$. As Rydén shows in [15], the *MMPP* is an identifiable process up to permutations of the states, or equivalently, for the two-states case, if the set $\{x, y, z, u, w = -x - y, v = 0\}$ defines a $MMPP_2$, then the only $MMPP_2$ with the same likelihood will be given by

$$\{\tilde{x} = u, \tilde{y} = z, \tilde{z} = y, \tilde{u} = x, \tilde{w} = -z - u, \tilde{v} = 0\}. \tag{4.24}$$

It can be verified that when $\tilde{v} = v = 0$, $\tilde{w} = -\tilde{x} - \tilde{y}$, and $w = -x - y$ then, Proposition 4.2 provides just two solutions, the original $MMPP_2$ and its permuted version, given by (4.24). Thus, our results are equivalent to Rydén's.

Finally, from the non-identifiability of the MAP_2 and the identifiability of the $MMPP_2$, one could wonder if given a MAP_2 there exist an equivalent $MMPP_2$. This has been tested from the equations given by Proposition 4.2, and the answer is that this is not true, in general. The following example illustrates this fact.

Example 4.2. Let us consider the same MAP_2 defined in Example 2:

$$D_0 = \begin{pmatrix} -10 & 2.5 \\ 0.6 & -3 \end{pmatrix}, \qquad D_1 = \begin{pmatrix} 4 & 3.5 \\ 1.35 & 1.05 \end{pmatrix}.$$

Then, according to the solutions given in Proposition 4.2 (having previously fixed $w = -x - y$ and $v = 0$), the only equivalent $MMPP_2$ is given by

$$\widetilde{D}_0 = \begin{pmatrix} -7.4231 & 2.2712 \\ 5.6788 & -5.5769 \end{pmatrix}, \qquad \widetilde{D}_1 = \begin{pmatrix} 5.1519 & 0 \\ 0 & -0.1019 \end{pmatrix},$$

which, since $d_{221} < 0$, does not define a real $MMPP_2$. Thus, in this case, there does not exist a $MMPP_2$ equivalent to the given MAP_2.

However, there do exist MAP_2s which can be reduced to $MMPP_2$s. For example,

$$D_0 = \begin{pmatrix} -20 & 8 \\ 3.5 & -5 \end{pmatrix}, \qquad D_1 = \begin{pmatrix} 11.5 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

is equivalent to the $MMPP_2$ defined by

$$\widetilde{D}_0 = \begin{pmatrix} -19.4211 & 7.8973 \\ 4.6027 & -5.5789 \end{pmatrix}, \qquad \widetilde{D}_1 = \begin{pmatrix} 11.5238 & 0 \\ 0 & 0.9762 \end{pmatrix}.$$

5. Conclusions and Extensions

In spite of the good properties that the *MAP*s present, which make them very suitable processes for modeling non-exponential events, there exist few works dealing with identifiability of the *MAP* which is of crucial importance when inference for the process is considered.

This present work is novel in two aspects: firstly, we have proven that the MAP_2 is not an identifiable process, providing a procedure to build an equivalent MAP_2 to a fixed one. We have shown that if two MAP_2s have equal LST for one and two data, then (and despite the great complexity involved in the equations), their LST will be equal for any set of points $(s_1, \ldots, s_n)$. Secondly, unlike other purely theoretical works, the results presented in this paper are illustrated by numerical examples. Many calculations have been carried out using Matlab and all codes utilized in the examples are available from the authors on request.

A number of extensions are possible. Firstly, we could consider the $BMAP_2$ where different numbers of batch arrivals are possible. Furthermore, we could extend this analysis to MAPs or $BMAP$s with more than two states. Finally, it is of practical interest to consider what happens when there is missing data, i.e. when a full sequence of interarrival times is not considered. Work on these problems is underway.

Appendix A. Proof of expression (2.4)

As $\boldsymbol{\phi}$ is the unique solution to $\boldsymbol{\phi} P^\star = \boldsymbol{\phi}$, we need to show that

$$(\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} D_1 P^\star = (\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} D_1.$$

Define $\Lambda = diag\{\lambda_1, \ldots, \lambda_m\}$, then $D_1 = \Lambda P_1$ and $D_0 = \Lambda(P_0 - I)$. From (2.3), $P^\star = (-\Lambda^{-1} D_0)^{-1} \Lambda^{-1} D_1$, and thus

$$\begin{aligned}
(\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} D_1 P^\star &= (\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} D_1 (-\Lambda^{-1} D_0)^{-1} \Lambda^{-1} D_1 \\
&= -(\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} D_1 D_0^{-1} D_1 \\
&= -(\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} (D - D_0) D_0^{-1} D_1 \\
&= (\boldsymbol{\pi} D_1 \mathbf{e})^{-1} (\boldsymbol{\pi} D - \boldsymbol{\pi} D_0) D_0^{-1} D_1 \\
&= (\boldsymbol{\pi} D_1 \mathbf{e})^{-1} \boldsymbol{\pi} D_1.
\end{aligned}$$

Appendix B. Proof of Theorem 3.1

Firstly, we provide some lemmas that will be necessary for the proof.

Lemma B.1. *Let T_n and $\widetilde{T}_n$ denote the times between the $(n-1)$th and nth arrival* *in two MAP_2s, $\mathcal{M}$ and $\widetilde{\mathcal{M}}$. Then,*

$$T_n \stackrel{d}{=} \widetilde{T}_n, \quad \textit{for all } n \geq 1,$$

if and only if

$$\boldsymbol{\theta}(P^\star)^{(n-1)}(sI - D_0)^{-1}D_1\mathbf{e} = \widetilde{\boldsymbol{\theta}}(\widetilde{P}^\star)^{(n-1)}(sI - \widetilde{D}_0)^{-1}\widetilde{D}_1\mathbf{e}, \tag{B.1}$$

for all $n \geq 1$, and for all s.

Proof. The variables T_n and $\widetilde{T}_n$ are equally distributed if and only if their Laplace transforms are the same. These are given by (B.1), where $\boldsymbol{\theta}(P^\star)^{(n-1)}$, and $\widetilde{\boldsymbol{\theta}}(\widetilde{P}^\star)^{(n-1)}$ represent the "initial" probabilities after $n-1$ arrivals.

A similar result to Lemma B.1, that provides a different characterization of Condition B1. in Theorem 3.1 is shown next.

Lemma B.2. *Let T and $\widetilde{T}$ denote the interarrival times of two stationary MAP_2s $\mathcal{M}$ and $\widetilde{\mathcal{M}}$, with stationary probabilities $\boldsymbol{\phi}$, and $\widetilde{\boldsymbol{\phi}}$. Then,*

$$T \stackrel{d}{=} \widetilde{T},$$

if and only if

$$\boldsymbol{\phi}(sI - D_0)^{-1}D_1\mathbf{e} = \widetilde{\boldsymbol{\phi}}(sI - \widetilde{D}_0)^{-1}\widetilde{D}_1\mathbf{e}. \tag{B.2}$$

Proof. We proceed as in the proof of Lemma B.1, but taking into account that

$$\lim_{n\to\infty}(P^\star)^n = \boldsymbol{\Phi}, \quad \lim_{n\to\infty}(\widetilde{P}^\star)^n = \widetilde{\boldsymbol{\Phi}}$$

or equivalently, $\lim_{n\to\infty}\boldsymbol{\theta}(P^\star)^n = \boldsymbol{\phi}$ and $\lim_{n\to\infty}\widetilde{\boldsymbol{\theta}}(\widetilde{P}^\star)^n = \widetilde{\boldsymbol{\phi}}$.

Lemma B.3. *Let $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ denote two MAP_2s, and let $\boldsymbol{\rho} = (\rho, 1-\rho)$, $\widetilde{\boldsymbol{\rho}} = (\tilde{\rho}, 1-\tilde{\rho})$ be any probability vectors. If*

$$\boldsymbol{\rho}(sI - D_0)^{-1}D_1\mathbf{e} = \widetilde{\boldsymbol{\rho}}(sI - \widetilde{D}_0)^{-1}\widetilde{D}_1\mathbf{e},$$

then

$$c\rho + \tilde{c}\tilde{\rho} + d = 0, \tag{B.3}$$

where c and $\tilde{c}$ are as defined in (3.4) and $d = \tilde{z} + \tilde{u} - z - u$.

Proof. Substituting D_0, D_1, $\widetilde{D}_0$ and $\widetilde{D}_1$ by their values (3.1), it can be seen that

$$\boldsymbol{\rho}(sI - D_0)^{-1}D_1\mathbf{e} = \frac{\alpha s + \gamma}{s^2 + \beta s + \gamma}, \tag{B.4}$$

where the coefficients α, β and γ are expressed in terms of (3.1) as:

$$\begin{aligned} \alpha &= \rho(z + u - x - y) - (z + u), \\ \beta &= -x - u, \\ \gamma &= xu - yz. \end{aligned} \tag{B.5}$$

Similarly,

$$\widetilde{\boldsymbol{\rho}}(sI - D_0)^{-1}D_1\mathbf{e} = \frac{\tilde{\alpha} s + \tilde{\gamma}}{s^2 + \tilde{\beta} s + \tilde{\gamma}},$$

where

$$\begin{aligned} \tilde{\alpha} &= \tilde{\rho}(\tilde{z} + \tilde{u} - \tilde{x} - \tilde{y}) - (\tilde{z} + \tilde{u}), \\ \tilde{\beta} &= -\tilde{x} - \tilde{u}, \\ \tilde{\gamma} &= \tilde{x}\tilde{u} - \tilde{y}\tilde{z}. \end{aligned} \tag{B.6}$$

Next, if

$$\frac{\alpha s + \gamma}{s^2 + \beta s + \gamma} = \frac{\tilde{\alpha} s + \tilde{\gamma}}{s^2 + \tilde{\beta} s + \tilde{\gamma}}, \tag{B.7}$$

then it can be easily seen that

$$\alpha = \tilde{\alpha},$$

which is equivalent, given the definitions of $\alpha, \tilde{\alpha}$, to

$$c\rho + \tilde{c}\tilde{\rho} + d = 0$$

with c and $\tilde{c}$ as in (3.4) and $d = \tilde{z} + \tilde{u} - z - u$.

Lemma B.4. *Let $P^\star$ be the transition probability matrix in a MAP_2 with vector of stationary probabilities $\boldsymbol{\phi}$. If all the rows of $P^\star$ are equal, then $P^\star = \Phi$, the matrix with rows are equal to $\boldsymbol{\phi}$.*

Proof. The proof is straightforward once the equation $\boldsymbol{\phi}P^\star = \boldsymbol{\phi}$ is solved, where it is assumed that all the rows of $P^\star$ are equal.

Finally, if $P^\star$ is transition matrix with vector of stationary probabilities $\boldsymbol{\phi}$, where

$$P^\star = \begin{pmatrix} p^\star_{11} & p^\star_{12} \\ p^\star_{21} & p^\star_{22} \end{pmatrix},$$

and $\boldsymbol{\phi} = (\phi, 1-\phi)$, then it is straightforward to check that

$$\phi = \frac{p^\star_{21}}{1 - p^\star_{11} + p^\star_{21}}. \tag{B.8}$$

Proof of the Theorem 3.1:

1. *B1.* and *B2.* $\rightarrow$ Weak equivalence.

   Let us first assume that both *B1.* and *B2.* hold. We want to prove equivalence given by (B.1), for all s and $n \geq 1$. As $(\boldsymbol{\theta}, \widetilde{\boldsymbol{\theta}}) = (\boldsymbol{\phi}, \widetilde{\boldsymbol{\phi}})$ and since *B2.* is equivalent to (B.2), then all equivalence conditions (B.1) hold because $\boldsymbol{\theta} P^\star = \boldsymbol{\phi} P^\star = \boldsymbol{\phi}$ and $\widetilde{\boldsymbol{\theta}} \widetilde{P}^\star = \widetilde{\boldsymbol{\phi}} \widetilde{P}^\star = \widetilde{\boldsymbol{\phi}}$.

2. Weak equivalence $\rightarrow$ *B1.* and *B2.*

   If two given MAP_2s are weakly equivalent, (B.1) holds, for all s and $n \geq 1$. If $n \rightarrow \infty$, then from Lemma B.2, (B.2) holds, and thus *B1.* holds too. Let us deduce *B2.* from weak equivalence; since (B.2) holds, then the pair $(\boldsymbol{\phi}, \widetilde{\boldsymbol{\phi}})$ verifies equation (B.3) (where $\rho = \phi$ and $\tilde{\rho} = \tilde{\phi}$),

   $$c\phi + \tilde{c}\tilde{\phi} + d = 0.$$

   Because of weak equivalence, (B.1) holds for $n = 1$, and thus the pair $(\boldsymbol{\theta}, \widetilde{\boldsymbol{\theta}})$ also satisfies (B.3),

   $$c\theta + \tilde{c}\tilde{\theta} + d = 0.$$

   Both equations imply,

   $$c\phi + \tilde{c}\tilde{\phi} = c\theta + \tilde{c}\tilde{\theta},$$

   or equivalently, using (B.8),

   $$c \, \frac{p^\star_{21}}{1 - p^\star_{11} + p^\star_{21}} + \tilde{c} \, \frac{\tilde{p}^\star_{21}}{1 - \tilde{p}^\star_{11} + \tilde{p}^\star_{21}} = c\theta + \tilde{c}\tilde{\theta}. \tag{B.9}$$

Again because of weak equivalence, and taking $n = 2$ in condition (B.1), then

$$c\,\frac{p^\star_{21}}{1 - p^\star_{11} + p^\star_{21}} + \tilde{c}\,\frac{\tilde{p}^\star_{21}}{1 - \tilde{p}^\star_{11} + \tilde{p}^\star_{21}} = c\theta^{(1)} + \tilde{c}\tilde{\theta}^{(1)},$$

where

$$\boldsymbol{\theta}^{(1)} = \boldsymbol{\theta} P^\star = \left(\theta^{(1)}, 1 - \theta^{(1)}\right), \quad \widetilde{\boldsymbol{\theta}}^{(1)} = \widetilde{\boldsymbol{\theta}}\widetilde{P}^\star = \left(\tilde{\theta}^{(1)}, 1 - \tilde{\theta}^{(1)}\right).$$

It can be checked that

$$\theta^{(1)} = \theta(p^\star_{11} - p^\star_{21}) + p^\star_{21}, \qquad \tilde{\theta}^{(1)} = \tilde{\theta}(\tilde{p}^\star_{11} - \tilde{p}^\star_{21}) + \tilde{p}^\star_{21},$$

and thus we need to solve for $(\theta, \tilde{\theta})$ in the following system of linear equations:

$$\begin{aligned} c\theta + \tilde{c}\tilde{\theta} &= c\,\frac{p^\star_{21}}{1 - p^\star_{11} + p^\star_{21}} + \tilde{c}\,\frac{\tilde{p}^\star_{21}}{1 - \tilde{p}^\star_{11} + \tilde{p}^\star_{21}} \\ c(p^\star_{11} - p^\star_{21})\theta + \tilde{c}(\tilde{p}^\star_{11} - \tilde{p}^\star_{21})\tilde{\theta} &= c\left(\frac{p^\star_{21}}{1 - p^\star_{11} + p^\star_{21}} - p^\star_{21}\right) + \tilde{c}\left(\frac{\tilde{p}^\star_{21}}{1 - \tilde{p}^\star_{11} + \tilde{p}^\star_{21}} - \tilde{p}^\star_{21}\right), \end{aligned} \tag{B.10}$$

whose coefficient matrix is

$$C = \begin{pmatrix} c & \tilde{c} \\ c(p^\star_{11} - p^\star_{21}) & \tilde{c}(\tilde{p}^\star_{11} - \tilde{p}^\star_{21}) \end{pmatrix}.$$

It can be easily seen that $\boldsymbol{\theta} = \boldsymbol{\phi}$, and $\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\phi}}$ solves the system. We need to determine the uniqueness of this solution. This comes from *A1.*, *A2.*, and Lemma B.4; since $P^\star \neq \boldsymbol{\Phi}$ or $\widetilde{P}^\star \neq \widetilde{\boldsymbol{\Phi}}$, then by Lemma B.4 either the rows of $P^\star$ or that of $\widetilde{P}^\star$ are not equal. This implies that $p^\star_{11} - p^\star_{21} \neq 0$ or $\tilde{p}^\star_{11} - \tilde{p}^\star_{21} \neq 0$. In addition, since $c, \tilde{c} \neq 0$ the rank of C is 2, and the solution is unique: $\boldsymbol{\theta} = \boldsymbol{\phi}$, and $\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\phi}}$.

Appendix C. Proof of Proposition 4.1

Let us first consider the case where $n = 1$. It is known from the proof of Lemma B.2 that the equality of Laplace transforms $f^*_{T;D_0,D_1}(s) = f^*_{\widetilde{T};\widetilde{D}_0,\widetilde{D}_1}(s)$ in the stationary version is equivalent to (B.7), where $\boldsymbol{\rho} = \boldsymbol{\phi}$, and $\widetilde{\boldsymbol{\rho}} = \widetilde{\boldsymbol{\phi}}$. If $\tilde{\alpha} = \alpha$, $\tilde{\beta} = \beta$, and $\tilde{\gamma} = \gamma$, or equivalently

$$\tilde{\phi}(\tilde{z}+\tilde{u}-\tilde{x}-\tilde{y})-(\tilde{z}+\tilde{u}) = \alpha, \quad (C.1)$$

$$-\tilde{x}-\tilde{u} = \beta, \quad (C.2)$$

$$\tilde{x}\tilde{u}-\tilde{y}\tilde{z} = \gamma, \quad (C.3)$$

then (B.7) holds, and thus $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ are *weakly* equivalent.

In the two data case, it can be shown that

$$f^*_{T;D_0,D_1}(s_1,s_2)=\frac{\delta_1 s_1 s_2+\delta_2 s_2+\alpha\gamma s_1+\gamma^2}{s_1^2 s_2^2+\gamma s_1^2+\gamma s_2^2+\beta s_1^2 s_2+\beta s_1 s_2^2+\beta^2 s_1 s_2+\beta\gamma s_2+\gamma^2},$$

where

$$\begin{aligned}
\delta_1 &= \phi\left((z+u-x-y)(w-v)+(x+y)(z+u)-(z+u)^2\right)+\\
&\quad (z+u-x-y)v+(z+u)^2,\\
\delta_2 &= \phi(x+y-z-u)(uw-yv-xv+zw)+(x+y-z-u)(xv-zw)-(u+z)\gamma\\
&\quad (x+y-z-u)(xv-zw)-(u+z)\gamma.
\end{aligned}$$

If the equations (C.1)-(C.3) are satisfied, and in addition

$$\tilde{\phi}\left((\tilde{z}+\tilde{u}-\tilde{x}-\tilde{y})(\tilde{w}-\tilde{v})+(\tilde{x}+\tilde{y})(\tilde{z}+\tilde{u})-(\tilde{z}+\tilde{u})^2\right)+ +(\tilde{z}+\tilde{u}-\tilde{x}-\tilde{y})\tilde{v}+(\tilde{z}+\tilde{u})^2 = \delta_1, \quad (C.4)$$

$$\tilde{\phi}(\tilde{x}+\tilde{y}-\tilde{z}-\tilde{u})(\tilde{u}\tilde{w}-\tilde{y}\tilde{v}-\tilde{x}\tilde{v}+\tilde{z}\tilde{w})+ +(\tilde{x}+\tilde{y}-\tilde{z}-\tilde{u})(\tilde{x}\tilde{v}-\tilde{z}\tilde{w})-(\tilde{u}+\tilde{z})\tilde{\gamma} = \delta_2, \quad (C.5)$$

then (4.2) holds for $n=1$ and $n=2$.

Appendix D. Proof of Theorem 4.1

The set $\mathcal{F}$ is contained in the set of solutions given by Proposition 4.2. To prove that the set $\mathcal{F}$ provides feasible solutions (real MAP_2s) let us first assume that $x<u$. Notice that this just orders the states; given a MAP_2 defined by $\{x,y,z,u,w,v\}$ then if $x>u$, then the same MAP_2 can be parameterized by changing state 1 by state 2 as $\{x',y',z',u',w',v'\}=\{u,z,y,x,-z-u-v,-x-y-w\}$. Let ε be defined as in (4.9):

$$0<\varepsilon<min\left\{-x,\frac{u-x}{2},\frac{z(1-\phi)}{\phi},\frac{(u-x)+\sqrt{(x-u)^2+4zy}}{2}\right\}.$$

It is easily checked that $\min\{-x, \frac{u-x}{2}, \frac{z(1-\phi)}{\phi}, \frac{(u-x)+\sqrt{(x-u)^2+4zy}}{2}\} > 0$. Because of that,

$$\tilde{u} = u - \varepsilon < 0, \quad \tilde{z} = z + \varepsilon > 0.$$

Moreover, since $\varepsilon < \frac{u-x}{2}$, this assures that $\tilde{x} < \tilde{u}$, and thus the parameterization of $\widetilde{\mathcal{M}}$ is different from that of $\mathcal{M}$ with permuted states. Next,

$$\frac{(u-x) - \sqrt{(x-u)^2 + 4zy}}{2} < 0 < \varepsilon < \frac{(u-x) + \sqrt{(x-u)^2 + 4zy}}{2},$$

implies

$$\tilde{y}(\tilde{u}, \tilde{z}) \equiv \frac{-(\varepsilon^2 + (x-u)\varepsilon - zy)}{z + \varepsilon} > 0.$$

In addition,

$$\tilde{w}(\tilde{u}, \tilde{z}) \equiv \frac{wz + v\varepsilon}{z} > 0, \quad \tilde{v} \equiv \frac{v(z+\varepsilon)}{z} > 0,$$

and finally, since $\varepsilon < z(1-\phi)/\phi$,

$$\tilde{\phi} \equiv \frac{(z+\varepsilon)\phi}{z} \in [0, 1].$$

Appendix E. Proof of Proposition 4.4

Let us assume that

$$G = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad H = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{and} \quad GH = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Now it is straightforward to verify that if $\widetilde{G}$ and $\widetilde{H}$ are defined by equations (4.20)-(4.23) (with respect to the elements of G and H, respectively), then

$$\widetilde{G}\widetilde{H} = \begin{pmatrix} \frac{\phi(A-C)+\widetilde{\phi}C}{\phi} & \frac{\phi\widetilde{\phi}\ (D+2C-A)+\phi^2\ (A-D+B-C)-\widetilde{\phi}^2C}{\phi\widetilde{\phi}} \\ \frac{\widetilde{\phi}C}{\phi} & \frac{\phi\ (C+D)-\widetilde{\phi}C}{\phi} \end{pmatrix}.$$

Acknowledgements

This work has been partially supported by the Spanish research projects 2007/04438/001, 2008/00052/001 and 2009/00147/001.

References

[1] Asmussen, S. and Koole, G. (1993). Marked point processes as limits of Markovian arrival streams. *Journal of Applied Probability* **30,** 365–372.

[2] Chakravarthy, S. (2001). The batch markovian arrival process: A review and future work. In *Advances in Probability and Stochastic Processes*.

[3] Fearnhead, P. and Sherlock, C. (2006). An exact Gibbs sampler for the Markov-modulated poisson process. *Journal of the Royal Statististical Society: Series B* **65,** 767–784.

[4] Heffes, H. (1980). A class of data traffic processes-covariance function characterization and related queueing results. *Bell Systems Technical Journal* **59,** 897–929.

[5] Heffes, H. and Lucantoni, D. (1986). A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE J.Select. Areas Communications* **4,** 856–868.

[6] Heyman, D. and Lucantoni, D. (2003). Modeling multiple IP traffic streams with rate limits. *IEEE/ACM Transactions on Networking* **11,** 948–958.

[7] Leroux, B. (1992). Maximum-likelihood estimation for hidden Markov models. *Stochastic Processes and Their Applications* **40,** 127–143.

[8] Lucantoni, D. (1993). The $BMAP/G/1$ queue: A tutorial. In *Models and Techniques for Performance Evaluation of Computer and Communication Systems*. ed. L. Donatiello and R. Nelson. Springer, New York pp. 330–358.

[9] Lucantoni, D., Meier-Hellstern, K. and Neuts, M. (1990). A single-server queue with server vacations and a class of nonrenewal arrival processes. *Advances in Applied Probability* **22,** 676–705.

[10] Neuts, M. F. (1979). A versatile markovian point process. *Journal of Applied Probability* **16,** 764–779.

[11] Nielsen, B. (2000). Modeling long-range dependent and heavy-tailed phenomena by matrix analytic methods. In *Advances in Algorithmic methods for Stochastic Models*. pp. 265–278.

[12] Ramirez, P., Lillo, R. and Wiper (2008). On identifiability of *MAP* processes. *Technical report.* Statistics and Econometrics Working Papers ws084613, Universidad Carlos III de Madrid.

[13] Rydén, T. (1994). Consistency and asymptotically normal parameter estimates for hidden Markov models. *Annals of Statistics* **22,** 1884–1895.

[14] Rydén, T. (1996). An em algorithm for estimation in markov-modulated poisson processes. *Computational Statistics and Data Analysis* **21,** 431–447.

[15] Rydén, T. (1996). On identifiability and order of continous-time aggregated Markov chains, Markov-modulated Poisson processes, and phase-type distributions. *Journal of Applied Probability* **33,** 640–653.

[16] Scott, S. (1999). Bayesian analysis of the two state markov modulated poisson process. *Journal of Computational and Graphical Statistics* **8,** 662–670.

[17] Scott, S. and Smyth, P. (2003). The Markov Modulated Poisson Process and Markov Poisson Cascade with applications to web traffic modeling. *Bayesian Statistics* **7,** 1–10.