Fitting Second-Order Acyclic Marked Markovian Arrival Processes

Andrea Sansottera*, Giuliano Casale†‡, Paolo Cremonesi*
*Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Via Ponzio 34/5, 20133, Milan
†Imperial College London, Department of Computing, 180 Queen's Gate, SW7 2AZ, London
‡Basque Center for Applied Mathematics (BCAM), Al. de Mazarredo 14, 48009, Bilbao, Basque Country, Spain
sansottera@elet.polimi.it, g.casale@imperial.ac.uk, paolo.cremonesi@polimi.it

***Abstract*—Markovian Arrival Processes (MAPs) are a tractable class of point-processes useful to model correlated time series, such as those commonly found in network traces and system logs used in performance analysis and reliability evaluation. Marked MAPs (MMAPs) generalize MAPs by further allowing the modeling of multi-class traces, possibly with cross-correlation between multi-class arrivals.**

In this paper, we present analytical formulas to fit second-order acyclic MMAPs with an arbitrary number of classes. We initially define closed-form formulas to fit second-order MMAPs with two classes, where the underlying MAP is in canonical form. Our approach leverages forward and backward moments, which have recently been defined, but never exploited jointly for fitting. Then, we show how to sequentially apply these formulas to fit an arbitrary number of classes. Representative examples and trace-driven simulation using storage traces show the effectiveness of our approach for fitting empirical datasets.

***Keywords*-Multi-class workload, point process, dependence**

I. Introduction

Marked Markovian Arrival Processes (MMAPs) have recently gained attention as a promising tool for performance and reliability analysis, useful in particular to model multi-class arrivals to queueing systems [1], [2]. A MMAP generalizes a Markovian Arrival Process (MAP) by featuring the ability to mark arrivals with a class label [2], [3]. This gives the ability to describe temporal dependent multi-class processes, e.g., cross-correlated non-Poisson flows. Compared to other point processes, MMAPs enable describing the statistical correlations between classes using a representation that is composable with continuous-time Markov chains. However, fitting methods for MMAPs are in the early stages and more work is needed to understand and fully exploit the expressive power of marked processes.

In this work, we develop novel analytical formulas for fitting MMAPs, restricting our attention to the important case of models with two states and an arbitrary number of classes. Only few methods are known to fit the parameters of MMAPs [1], [4]. To estimate the parameters of Batch Markovian Arrival Processes, defined similarly to MMAPs, an expectation maximization approach has been presented in [8], [9]. Expectation maximization, however, is computationally intensive and unsuited to analyze traces with a large number of samples.

Compared to these works, our fitting approach leverages closed-form analytical formulas that are specific to second-order models. Models with few states are the only ones analytically tractable, even when just studying the properties of the underlying PH (phase-type distribution) or MAP [5]–[7]. However, despite their apparent simplicity, second-order MMAPs are non-trivial to fit due to the non-linear equations involved, the multiple options for choosing what to fit, and the many degrees of freedom they offer. Moreover, our work may serve as a building block to approximate more complex traces, through the generalization of approaches such as superposition. Our main technical results to tackle these problems may be summarized as follows:

- We develop a fitting methodology that first fits the underlying MAP, and then matches first-order multi-class moments of the types proposed in [1], [4], possibly together with probabilities characterizing the marking process. A useful result is the observation that resorting to an underlying MAP in canonical form enables analytical tractability for the second-order MMAP, impacting in a limited manner on the available degrees of freedom.
- When fitting MMAP parameters to trace data, the resulting process is not necessarily a Markov process unless some constraints are satisfied [4]. Stemming from this, the complex relationship between multi-class moments and feasibility regions is illustrated. For approximate moment matching, we develop a characterization of the feasible region for the two-state MMAP which allows us to project a general moment set into one that can be fitted exactly by a second-order MMAP.
- We show how to sequentially apply our methodology to fit second-order MMAPs with an arbitrary number of classes. Here, classes are fitted one after the other, exposing the detail for each class one-at-a-time. This method is found to yield accurate results also if the dataset is not drawn from a second-order MMAP.

This study is also motivated by the fact that MMAPs are needed to model real systems where cross-correlation between workloads arises in a natural manner, such as storage systems. For instance, upon issuing read and write requests to a remote storage drive, the operating system can reschedule and merge requests to optimize storage performance. This leads to situations where sequences of successive reads are interleaved by sequences of successive writes, in such a way that the assumption of independence between the two flows is significantly violated. In these cases, we show that MMAPs appear promising for describing correlations between non-Poisson flows. Furthermore, thanks to their Markovian representation, the resulting queueing systems can be efficiently evaluated by matrix-analytic models without resorting to simulation [10]. MMAPs may also be useful in simulation-based dependability evaluation. For example, one could use MMAPs to model the inter-arrival times between events of different types for the system under study. Also, Markov-modulated processes may be helpful to generate multi-class traces that are stochastically similar to an existing measured trace [11].

The remainder of this paper is organized as follows. In Section II, we give basic notation and definitions. Properties of multi-class moments are examined in Section III. Section IV introduces fitting methods for second-order MMAPs with two classes. The sequential fitting approach for arbitrary number of classes is introduced in Section V. Results are given in Section VI for both random instances and real-world traces. Conclusions are drawn in Section VII.

II. Background and Notation

We consider an interval-stationary sequence of arrivals, each marked with a class label. Let $X_i \in \mathbb{R}^+$ be the inter-arrival time between arrivals $i-1$ and i, for $i \geq 1$. Moreover, let $\mathcal{C} = \{1, \dots, m\}$ be the set of class labels and denote by $C_i \in \mathcal{C}$ the class of the i-th arrival. Our goal is to fit the marked arrival sequence (X_i, C_i). Below, we provide required definitions needed for this task.

A. Multi-Class Moments

Differently from single-class processes, in marked point-processes several statistical moments of inter-arrival times may be defined. We relate our approach with existing work [1], [4] by considering similar moments. We define

$$M_j = E[X_i^j] \tag{1}$$

as the *ordinary* j-th moment of the inter-arrival times between arrivals of *any* class. To capture the behavior of marked processes, we can also define other types of moments conditioning on the C_i random variable. We refer to these moments as *multi-class moments*. The moments of the inter-arrival time *preceding* an arrival of class c are defined as

$$B_{j,c} = E[X_i^j | C_i = c] \tag{2}$$

and referred to in the rest of this paper as *backward moments*. Let $p_c = \Pr[C_i = c]$ be the fraction of class-c arrivals, then the scaled moments $B_{j,c}p_c$ correspond to moments used by Horvath *et al.* in [1]. Furthermore, we consider the moments of the inter-arrival time *following* an arrival of class c, which are defined by Buchholz *et al.* in [4] as

$$F_{j,c} = E[X_i^j | C_{i-1} = c] \tag{3}$$

and referred to as *forward moments*.

B. Marked Markovian Arrival Process (MMAP)

In this paper, we use MMAPs as a fitting tool for marked arrival sequences (X_i, C_i), $i \geq 1$. A MMAP of order n with m classes, called an MMAP$[m]$ of order n, may be seen as a hidden-Markov process capable of producing samples $(\tilde{X}_i, \tilde{C}_i)$, $i \geq 1$, where $\tilde{X}_i$ is an inter-arrival time drawn from a MAP with n transient states and $\tilde{C}_i \in \{1, \dots, m\}$ is a class label. The MMAP moment matching problem is to find a MMAP such that the moments of $(\tilde{X}_i, \tilde{C}_i)$ match, either exactly or approximately, a set of empirical moments of (X_i, C_i) computed on a sufficient large, but finite, number of samples. It is shown in [1], [4] that the degrees of freedom of a MMAP$[m]$ of order n are at most mn^2. For example, in a second-order MMAP[2], the available degrees of freedom for fitting are at most eight.

A MMAP$[m]$ of order n is specified by a set of $n \times n$ matrices $(D_0, D_{1,1}, D_{1,2}, \dots, D_{1,m})$, called a *representation*, where D_0 is the sub-generator matrix of the underlying PH of order n, and $D_{1,c} = (-D_0)P_c$ where P_c is a sub-stochastic matrix with elements (i, j) representing the probability that, knowing that an arrival started in phase i, the arrival is of class c and is followed by an arrival (of any class) that starts in phase j. The interval-stationary version of the MMAP is obtained by initializing the process in a phase chosen at random according to the equilibrium probability vector α of the discrete-time Markov chain $P = (-D_0)^{-1}D_1$, where $D_1 = \sum_c D_{1,c}$. Such vector is the left eigenvector of P associated to its unitary eigenvalue, i.e., $\alpha P = \alpha,\ \alpha\mathbb{1} = 1$, where $\mathbb{1}$ is a column vector of ones.

Note that (D_0, D_1) is the order-n MAP (i.e., MAP(n)), obtained by considering only the inter-arrival time sequence $\widetilde{X}_i$ defined by the MMAP$[m]$ of order n and ignoring class markings. We refer to this process as the MAP underlying the MMAP. Similarly, the underlying distribution of inter-arrival times is a PH with representation (α, D_0), where $\alpha = \alpha(-D_0)^{-1}D_1$. We refer to this distribution as the PH underlying the MMAP.

Let $p_c = \alpha P_c \mathbb{1}$ be the probability of a class c arrival, where $P_c = (-D_0)^{-1}D_{1,c}$. The multi-class moments defined in the previous subsection are readily computed as:

$$F_{j,c} = j! p_c^{-1} \alpha P_c (-D_0)^{-j} \mathbb{1} \tag{4}$$

$$B_{j,c} = j! p_c^{-1} \alpha (-D_0)^{-j} P_c \mathbb{1} \tag{5}$$

where p_c^{-1} is used to condition on the class c arrival.

III. Moment Dependencies

Forward and backward moments are in general *dependent* statistical descriptors. This raises concerns on the actual degrees of freedom that are available when attempting to fit them simultaneously. In this section, we try to answer this question by illustrating some relationships between multi-class moments. To illustrate the dependencies between backward and forward moments, we introduce the moments of the inter-arrival time between successive arrivals of classes c and k

$$M_{j,c,k} = E[X_i^j | C_{i-1} = c, C_i = k] \tag{6}$$

which we refer to as *cross-moments* of classes c and k. Such moments may be computed as

$$M_{j,c,k} = j! p_{c,k}^{-1} \alpha P_c (-D_0)^{-j} P_k \mathbb{1} \tag{7}$$

where the class transition probability is

$$p_{c,k} = \Pr[C_{i-1} = c, C_i = k] = \alpha P_c P_k \mathbb{1}. \tag{8}$$

We first observe that ordinary moments may be readily written as weighted sums of multi-class moments, i.e.,

$$M_j = \sum_{c=1}^{m} p_c B_{j,c} = \sum_{c=1}^{m} p_c F_{j,c} = \sum_{c=1}^{m} \sum_{k=1}^{m} p_{c,k} M_{j,c,k} \tag{9}$$

This follows immediately by the definitions. Similarly, we can observe that

$$B_{j,c} = p_c^{-1} \sum_k p_{k,c} M_{j,k,c} \tag{10}$$

$$F_{j,c} = p_c^{-1} \sum_k p_{c,k} M_{j,c,k} \tag{11}$$

which again follows by definition and it is subject to the following constraints:

$$\sum_{c=1}^{m} p_c = 1, \qquad \sum_{k=1}^{m} p_{c,k} = p_c, \qquad \sum_{c=1}^{m} p_{c,k} = p_k \tag{12}$$

Equations (10)-(11) define a system of linear equations that characterizes the dependencies between multi-class moments. For example, for a MMAP[2] (10)-(11) become

$$B_{j,1} = p_1^{-1} \left(p_{1,1} M_{j,1,1} + p_{2,1} M_{j,2,1} \right) \tag{13}$$

$$B_{j,2} = p_2^{-1} \left(p_{1,2} M_{j,1,2} + p_{2,2} M_{j,2,2} \right) \tag{14}$$

$$F_{j,1} = p_1^{-1} \left(p_{1,1} M_{j,1,1} + p_{1,2} M_{j,1,2} \right) \tag{15}$$

$$F_{j,2} = p_2^{-1} \left(p_{2,1} M_{j,2,1} + p_{2,2} M_{j,2,2} \right) \tag{16}$$

Combining with (12), analyzing the rank of the coefficient matrix for $j = 1$ and assuming the variables $M_{j,c,k}$ as unknowns, we find that the system of linear equations has rank 2, provided that probabilities and forward and backward moments are known. Adding the constraint (9) raises the rank to 3, where the missing information depends on the products $P_c(-D_0)^{-j} P_k$ in (7) which are not expressible by $p_{c,k}$ or by the forward and backward moments.

The above analysis is useful for a justification of the fitting strategy pursued in this paper. One problem that commonly arises in analytically inverting the formulas for the MMAP parameters is that the non-linearities make it difficult to find explicit fitting formulas. The strategy we propose to tackle this problem is to focus on fitting the underlying MAP first, and then first-order moments $B_{1,c}$ or $F_{1,c}$, possibly together with some probabilities p_c or $p_{c,k}$. With this approach, expressions are considerably simplified and amenable to analytical treatment.

However, the fact that (13) is under-determined for $j = 1$ suggests that also a cross-moment may be needed to uniquely determine a general second-order MMAP[2]. Still, we find that assuming that the underlying MAP is in canonical form we can slightly reduce the available degrees of freedom, making it possible the unique determination of the second-order MMAP[2] up to seven degrees of freedom. This is illustrated in the fitting procedures developed next.

IV. Acyclic Second-Order MMAP[2] Fitting

From now on, we develop fitting algorithms focusing on the case of second-order MMAP[2] models where we assume an acyclic structure for the underlying MAP. We refer to these models as second-order AMMAP[2] models. The generalization to the second-order AMMAP[m] model is provided in Section V. For ease of notation and unless otherwise specified, from now on we omit the number of states in MMAPs and assume a model with two states.

A. Acyclic Marked PH Renewal Process (AMPH[2])

We being with illustrating our fitting methodology for the special case of a phase-type renewal process with marked arrivals. This is a marked process where the inter-arrival time sequence X_k is independent and identically distributed according to a PH distribution. Since we are focusing on two-state models, we here consider the underlying PH to be in acyclic canonical form, called APH(2), since it is well-known that second-order APH and PH have the same expressive power [6]. The resulting marked PH renewal process is here called a AMPH process. While choosing an underlying PH that is acyclic is helpful to remove redundant parameters, we show that care should be taken in choosing the right APH, since the decision may affect the feasibility region of the AMPH. Recall that an APH(2) is a canonical form for any second-order PH, and it is defined as follows:

$$D_0 = \begin{bmatrix} -\frac{1}{h_1} & \frac{r_1}{h_1} \\ 0 & -\frac{1}{h_2} \end{bmatrix}, \qquad D_1 = \begin{bmatrix} \frac{(1-r_1)}{h_1} & 0 \\ \frac{1}{h_2} & 0 \end{bmatrix}$$

where h_1, r_1, and h_2 are parameters. To define a AMPH[2], we use coefficients $0 \le q_1, q_2 \le 1$, such that

$$D_{1,1} = \begin{bmatrix} \frac{q_1(1-r_1)}{h_1} & 0 \\ \frac{q_2}{h_2} & 0 \end{bmatrix}, \qquad D_{1,2} = \begin{bmatrix} \frac{(1-q_1)(1-r_1)}{h_1} & 0 \\ \frac{1-q_2}{h_2} & 0 \end{bmatrix}$$

which ensure that $D_{1,1}$ and $D_{1,2}$ have non-negative elements and $D_{1,1} + D_{1,2} = D_1$, where D_1 is assigned given that the underlying APH(2) is assumed fitted.

Notice that an AMPH[2] renewal process leads us to leverage only five degrees of freedom out of a theoretical maximum of eight [4], however this gap is mainly due to the renewal assumption which we relax in the next section. To fit the above AMPH[2] we proceed as follows. First, we assume that the (unmarked) inter-arrival times X_k are fitted with an APH(2) distribution, such that the parameters h_1, r_1, and h_2 are assigned. This can be done by fitting three ordinary moments (1) of the inter-arrival time distribution, using for example the formulas proposed in [7]. Then, we seek for values of q_1 and q_2 to fit the AMPH[2] by matching multi-class moments or class probabilities. By expanding (4), (5), and (8) we find, after simple algebraic manipulations, that

$$\begin{aligned} p_1 &= q_1(1-r_1) + r_1 q_2 \\ p_{1,2} &= p_1 p_2 \\ B_{1,1} &= h_1 + h_2 r_1 q_2 / p_1 \\ F_{1,1} &= F_{1,2} = M_1 = h_1 + h_2 r_1 \\ F_{2,1} &= F_{2,2} = M_2 = 2h_1^2 + 2r_1 h_1 h_2 + 2r_1 h_2^2 \end{aligned}$$

The above formulas lead to the non-trivial observation that, if the underlying PH is acyclic, only the backward moments can be used to fit the AMPH[2]. In fact, the forward moments degenerate into the ordinary moments M_j. Also, if any probability term is known, all the other probabilities are readily obtained using (12) and $p_{c,k} = p_c p_k$. Fitting p_1 and $B_{1,1}$ we then obtain the closed-form fitting expressions:

$$q_1 = p_1 \frac{h_1 - B_{1,1} + h_2}{h_2(1-r_1)}, \quad q_2 = p_1 \frac{B_{1,1} - h_1}{h_2 r_1} \tag{17}$$

which can be used as moment matching formulas to fit a AMPH[2] from $B_{1,1}$ and p_1.

Feasibility Region. The fitting formulas (17) are valid subject to

$$p_1 B_{1,1} \leq \min((h_1 + h_2)p_1, h_1 p_1 + h_2 r_1) \tag{18}$$

$$p_1 B_{1,1} \geq \max((h_1 + h_2)p_1 - h_2(1-r_1), h_1 p_1) \tag{19}$$

derived from the constraints $0 \leq q_1, q_2 \leq 1$.

In general, multiple APH(2) exist that can fit the distribution of the renewal process, and since each of the APH(2) has different values of the parameters h_1, h_2 and r_1, we can investigate the impact of the choice of the underlying APH(2) on the feasible region of the AMPH[2] parameters. This is illustrated in the example in Figure 1. There, we consider the APH(2) with $h_1 = \frac{2}{5}$, $h_2 = \frac{4}{5}$, $r_1 = \frac{3}{4}$, which has ordinary moments $M_1 = 1$, $M_2 = \frac{44}{25}$, $M_3 = \frac{552}{125}$. The same moments are also fitted by the APH(2) with $h_1 = \frac{4}{5}$, $h_2 = \frac{2}{5}$, $r_1 = \frac{1}{2}$. Figure 1 shows the feasible regions of $p_1 B_{1,1}$, which is the un-normalized backward moment in [1], and p_1 for the AMPHs that use these two APH(2) processes. The feasible region of the first AMPH[2] (dashed) contains the feasible region of the second AMPH[2] (dotted). Hence, the first APH(2) leads to a AMPH[2] which has greater flexibility for fitting purposes. Since the number of APH(2) that can fit a given moment set is at most two, the choice of the best one amounts to evaluating both cases to check if a fitting AMPH[2] can be found in any of the two.

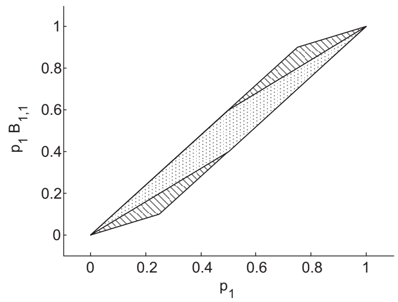

Figure 1. Feasible regions for the multi-class moments for two AMPH(2) using different underlying APH(2) representations fitting the same three moments.

Approximate fitting. When an exact fit in the feasible region cannot be found, priority should be given to fitting the p_1 class probability first. Indeed, in queueing analysis, fitting p_1 and $p_2 = 1 - p_1$ is crucial to match the utilizations of each class since $\lambda_1 = p_1/M_1$ and $\lambda_2 = p_2/M_1$ are the arrival rates of the two classes, assuming that the MMAP is used to describe a job arrival process. Thus, when no exact fit can be found, we suggest to adjust the value of $B_{1,1}$ so that a valid AMPH[2] matching the arrival rates can be found. The feasible $\widetilde{B}_{1,1}$ closest to $B_{1,1}$ is obtained as follows:

- if $p_1 B_{1,1}$ is above the upper bound (18),

$$\widetilde{B}_{1,1} = p_1^{-1} \min((h_1 + h_2)p_1, h_1 p_1 + h_2 r_1) \tag{20}$$

- if $p_1 B_{1,1}$ is below the lower bound (19),

$$\widetilde{B}_{1,1} = p_1^{-1} \max((h_1 + h_2)p_1 - h_2(1-r_1), h_1 p_1) \tag{21}$$

When two underlying APH(2) are feasible, we evaluate the best feasible backward moment $\widetilde{B}_{1,1}$ for both and choose the APH(2) representation which leads to the smallest possible error $|\widetilde{B}_{1,1} - B_{1,1}|$. This provides improved accuracy since one of the two representations has a larger feasible region for the backward moment, as shown in Figure 1.

Degenerate cases. We also remark that, for some values of the parameters of the underlying PH, a degree of freedom might be lost. For example, for $r_1 = 1$ one degree of freedom is lost and the process becomes:

$$D_0 = \begin{bmatrix} -\frac{1}{h_1} & \frac{1}{h_1} \\ 0 & -\frac{1}{h_2} \end{bmatrix} \tag{22}$$

$$D_{1,1} = \begin{bmatrix} 0 & 0 \\ \frac{q_2}{h_2} & 0 \end{bmatrix}, D_{1,2} = \begin{bmatrix} 0 & 0 \\ \frac{1-q_2}{h_2} & 0 \end{bmatrix} \tag{23}$$

In this case, the only parameter is q_2 and imposes that the arrival rates of the classes are properly fitted. The ordinary moments, the forward moments and the backward moments coincide. For instance, the first-order moments are

$$F_{1,1} = F_{1,2} = B_{1,1} = B_{1,2} = M_1 = h_1 + h_2. \quad (24)$$

B. AMMAP[2] with positive autocorrelation decay

In this and in the next section, we relax the restriction of the underlying MAP(2) being a PH renewal process. Recall from [12] that any MAP(2) can be represented with one of two different acyclic forms, called AMAP(2)s, where the specific acyclic form chosen depends on whether the autocorrelation decay rate is positive or negative. These two cases are considered in this section and in Section IV-C, respectively, in which we discuss the parameter fitting for the resulting acyclic process, called a AMMAP[2].

The AMMAP[2] process with positive autocorrelation decay in the underlying MAP has representation:

$$D_0 = \begin{bmatrix} -\frac{1}{h_1} & \frac{r_1}{h_1} \\ 0 & -\frac{1}{h_2} \end{bmatrix} \quad (25)$$

$$D_{1,1} = \begin{bmatrix} \frac{q_{1,1}(1-r_1)}{h_1} & 0 \\ \frac{q_{2,1}(1-r_2)}{h_2} & \frac{q_{2,2}r_2}{h_2} \end{bmatrix} \quad (26)$$

$$D_{1,2} = \begin{bmatrix} \frac{(1-q_{1,1})(1-r_1)}{h_1} & 0 \\ \frac{(1-q_{2,1})(1-r_2)}{h_2} & \frac{(1-q_{2,2})r_2}{h_2} \end{bmatrix} \quad (27)$$

which offers seven degrees of freedom out of a theoretical maximum of eight. Thus, we can match up to four parameters in the underlying MAP, the theoretical maximum, and then spend three degrees of freedom to match multi-class moments and class probabilities. The multi-class moments and class probabilities of the AMMAP[2] are as follows

$$\begin{aligned} p_1 =& \frac{r_1(q_{1,1} - q_{2,2} - r_1 q_{1,1} + r_1 q_{2,1})}{(r_1-1)(p_2(r_1-1)+1)} - \\ & \frac{q_{1,1} - r_1 q_{1,1} + r_1 q_{2,1} - r_1 q_{2,2}}{r_1 - 1} \\ p_{1,1} =& p_1(1-r_2)(q_{1,1} - r_1 q_{1,1} - r_2 q_{1,1} + r_1 q_{2,1} + \\ & r_1 r_2^2 q_{2,1} - r_1 r_2^2 q_{2,2} + r_1 r_2 q_{1,1} - 2 r_1 r_2 q_{2,1}) \\ B_{1,1} =& h_1 + \\ & \frac{r_1(h_2 - h_1 r_2)(q_{2,1} - r_2 q_{2,1} + r_2 q_{2,2})}{r_1 q_{2,1} - q_{1,1}(r_1 + r_2 - r_1 r_2 - 1) - r_1 r_2 q_{2,1} + r_1 r_2 q_{2,2}} \\ F_{1,1} =& h_1 + h_2 p_1 - \\ & \frac{r_1 r_2 q_{2,2}(h_1 - h_2 + h_2 r_1)}{r_1 q_{2,1} - q_{1,1}(r_1 + r_2 - r_1 r_2 - 1) - r_1 r_2 q_{2,1} + r_1 r_2 q_{2,2}} \end{aligned}$$

Since the degrees of freedom for multi-class moment matching are three, only a subset of p_1, $p_{1,1}$, $B_{1,1}$, and $F_{1,1}$ may be matched exactly. Assuming again that the probabilities p_1 and $p_2 = 1 - p_1$ are required to impose the mean arrival rate of each class, we may consider three cases:

1) fitting p_1, $B_{1,1}$ and $F_{1,1}$, described in Section IV-B1;
2) fitting p_1, $F_{1,1}$ and $p_{1,1}$, described in Section IV-B2;
3) fitting p_1, $B_{1,1}$ and $p_{1,1}$, described in Section IV-B3.

For each case, we proceed similar to the AMPH[2] model in which we start from the expressions of multi-class moments and class probabilities to obtain explicit expressions for the $q_{1,1}$, $q_{1,2}$, $q_{2,2}$ parameters. In general, the resulting expressions are non-linear but elementary enough to be solvable symbolically. Finally in Section IV-B4, we discuss degenerate cases in which special care should be taken when fitting the second-order AMMAP[2], due to zeros occurring in D_0 or D_1 of the underlying AMAP(2).

1) Fitting p_1, $B_{1,1}$ and $F_{1,1}$: Solving the expressions of p_1, $B_{1,1}$ and $F_{1,1}$ for the parameters $q_{1,1}$, $q_{1,2}$, $q_{2,2}$ yields

$$\begin{aligned} q_{1,1} =& -\frac{p_1(r_1 r_2 - r_2 + 1)(B_{1,1} - h_1 - h_2 + h_1 r_2)}{(h_2 - h_1 r_2)(r_1 - 1)(r_2 - 1)} \\ q_{2,1} =& \frac{p_1(r_1 r_2 - r_2 + 1)}{(r_1 - 1)(r_2 - 1)} - \frac{p_1(B_{1,1} - h_1)(r_1 r_2 - r_2 + 1)}{r_1(h_2 - h_1 r_2)(r_2 - 1)} \\ & - \frac{p_1(r_1 r_2 - r_2 + 1)(h_1 - F_{1,1} + B_{1,1} r_1)}{r_1(h_1 + h_2(r_1 - 1))(r_1 - 1)(r_2 - 1)} \\ q_{2,2} =& \frac{p_1(r_1 r_2 - r_2 + 1)(h_1 - F_{1,1} + h_2 r_1)}{r_1 r_2(h_1 - h_2 + h_2 r_1)} \end{aligned}$$

Feasible region. The feasible region of the forward and backward moments, for different values of p_1, is shown in Figure 2. The figure shows the feasible region of the four possible AMAP(2)s that can be used to fit the same set of moments and autocorrelation decay rate. Such AMAP(2)s differ for the ordering of the diagonal elements in D_0 and for the chosen root of the quadratic equation arising in the fitting procedure [7]. We find that the feasible region strongly depends on the specific AMAP(2) chosen, even though a common region exists which is insensitive to this choice. The special case $B_{1,1} = F_{1,1} = M_1$ is always feasible and corresponds to the choice of the parameters $q_{1,1} = q_{2,1} = q_{2,2} = p_1$. While for the AMPH[2] processes one of the feasible regions is a superset of the other, as shown in Figure 1, for the AMMAP[2] the situation is more complex, since the best choice depends on the actual values of the sample backward and forward moments. Furthermore, we observe that as the fraction p_1 of class-1 arrival increases, the class-1 moments become increasingly difficult to match, as we can see from the fact that the feasible region progressively collapses. This is due to the fact that ordinary moments are convex combinations of forward and backward moments, as shown in (9), thus as the value of p_1 grows large $B_{1,1}$ and $M_{1,1}$ become the dominant terms in the linear combination and both tend to M_1. This suggests the fact that, for queueing analysis purposes, it appears more flexible to fit the statistical properties of a class of jobs that arrive rarely in the arrival stream rather than the properties of a class that arrives frequently. This holds under the assumption that the total arrival rate $\lambda = 1/M_1$ is fixed.

Approximate fitting. When the first-order forward and backward are outside the feasible region, we fit the feasible point $(\widetilde{B}_{1,1}, \widetilde{F}_{1,1})$ closest to $(B_{1,1}, F_{1,1})$ in Euclidean norm. The explicit expressions for the AMMAP[2] parameters are linear with respect to the un-normalized forward and backward moments $p_1 F_{1,1}$ and $p_1 B_{1,1}$. Hence, the feasibility for the first-order forward and backward moments is determined by six linear inequalities, derived by imposing the conditions $0 \leq q_{i,j} \leq 1$ for the above expressions. To find the feasible point $(\widetilde{B}_{1,1}, \widetilde{F}_{1,1})$ closest in Euclidean norm to $(B_{1,1}, F_{1,1})$, we enumerate all the vertices and edges of the polytope and find the closest point on the boundary.

2) Fitting forward moments and class probabilities: While, for the second-order AMMAP[2], the probabilities $p_{c,k}$ are uniquely determined once we fix p_c, $F_{1,c}$ and $B_{1,c}$, for approximate fitting purposes, depending on the underlying model, it might be valuable to fit p_c, $F_{1,c}$ and $p_{c,k}$. We have obtained explicit analytical expression to derive the parameters from this set of characteristics. First, for convenience of notation, we define

$$\begin{aligned} w_1 &= F_{1,1} - h_1 - F_{1,1}r_2 + h_1 r_2 - h_2 r_1 + F_{1,1}r_1 r_2 \\ w_2 &= p_1(r_1 - 1)w_1 \\ w_3 &= r_1(r_2 - 1)(h_1 - h_2 + h_2 r_1)w_1 \end{aligned}$$

The parameters of the second-order AMMAP[2] are then

$$\begin{aligned} q_{1,1} =& w_2^{-1}(p_1^2 h_1 - {p_1}^2 F_{1,1} - p_{1,1}h_1 r_1 + p_{1,1}h_2 r_1) + \\ & w_2^{-1}(p_1^2 F_{1,1} r_1 + p_1^2 F_{1,1} r_2 - p_1^2 h_1 r_2 - p_{1,1} h_2 r_1^2) + \\ & w_2^{-1}(-2p_1^2 F_{1,1} r_1 r_2 + p_1^2 h_1 r_1 r_2 + p_1^2 F_{1,1} r_1^2 r_2) \end{aligned} \tag{28}$$

$$q_{2,1} = -(y_1 + y_2 + y_3 + y_4 p_1^2 \textstyle\sum_{i=1}^{11} x_i)(p_1 w_3)^{-1} \tag{29}$$

$$q_{2,2} = p_1(r_1 r_2 - r_2 + 1)(h_1 - F_{1,1} + h_2 r_1)u_1 \tag{30}$$

where the y_i, x_1, and u_1 are given in the appendix.

Feasible region and approximate fitting. Constraints are obtained by imposing $0 \leq q_{i,j} \leq 1$ to the explicit parameter expression (28), (29), (30), which are nonlinear with respect to $F_{1,1}$ and $p_{1,1}$. Recall that p_1 is always feasible, hence no adjustment is necessary and the arrival of the two classes can be always matched exactly. When the provided forward moments and class transition probabilities are unfeasible, we fit the feasible point $(\widetilde{F}_{1,1}, \widetilde{p}_{1,1})$ closest in Euclidean norm to $(F_{1,1}, p_{1,1})$. This point can be obtained by solving a nonlinearly constrained least squares problem in which we minimize the norm of $\Delta = [\Delta_{F_{1,1}} \Delta_{p_{1,1}}]$, with

$$\Delta_{F_{1,1}} = p_1(\widetilde{F}_{1,1} - F_{1,1}), \quad \Delta_{p_{1,1}} = (\widetilde{p}_{1,1} - p_{1,1})$$

3) Fitting backward moments and class probabilities: For convenience of notation we define

$$\begin{aligned} w_4 &= B_{1,1} - h_1 - B_{1,1}r_2 + h_1 r_2 - h_2 r_1 + B_{1,1} r_1 r_2 \\ w_5 &= r_1(h_2 - h_1 r_2)(r_2 - 1)w_4 \end{aligned}$$

The parameters are then computed as follows

$$q_{1,1} = -\frac{p_1(r_1 r_2 - r_2 + 1)(B_{1,1} - h_1 - h_2 + h_1 r_2)}{(h_2 - h_1 r_2)(r_1 - 1)(r_2 - 1)} \tag{31}$$

$$\begin{aligned} q_{2,2} =& \frac{p_{1,1}h_2 - p_1^2 h_2 - p_1^2 B_{1,1} r_2^2 - p_{1,1} h_1 r_2}{p_1 r_2 w_5} \\ &+ \frac{-p_{1,1} h_2 r_2 + p_1^2 B_{1,1} r_2 + p_1^2 h_2 r_2 + p_{1,1} h_1 r_2^2}{p_1 r_2 w_5} \\ &+ \frac{-p_1^2 h_2 r_1 r_2 + p_1^2 B_{1,1} r_1 r_2^2}{p_1 r_2 w_5} \end{aligned} \tag{32}$$

$$q_{2,1} = \frac{y_1' + y_2' + p_1^2 \sum_{i=1}^{9} x_i'}{p_1 w_5} \tag{33}$$

where x_i' and y_i' are defined in the appendix.

Feasible region and approximate fitting. Similarly to the previous case, the feasible region is bounded by nonlinear constraints, obtained by imposing $0 \leq q_{i,j} \leq 1$ in (31), (32), (33). We fit the feasible point closest in Euclidean norm to $(B_{1,1}, p_{1,1})$ by solving a nonlinearly constrained least squares problem.

4) Degenerate cases: We remark that the following degenerate cases have to be considered upon fitting a second-order AMMAP[2] with positive autocorrelation decay.

- If $r_1 = 0$ or $r_2 = 1$, the AMAP(2) is exponential, because the stationary probability of for state 1 or 2, respectively, is equal to 1. In this degenerate case, only a marked Poisson process can be fitted with $D_0 = [-1/h]$, $D_{1,1} = [q/h]$, $D_{1,2} = [(1-q)/h]$. The probability of an arrival of the first class is $p_1 = q$.
- If $h_2 - h_1 r_2 = 0$ the AMAP(2) is also exponential. This case is interesting: even if the matrix $D_1 = D_{1,1} + D_{1,2}$ has three non-zero elements, the AMMAP[2] cannot be fitted, since the denominator of the expressions of $q_{1,1}$, $q_{1,2}$, and $q_{2,2}$ becomes null. Hence, like in the previous case, we can only fit a marked Poisson process.
- If $r_1 = 1$ and $r_2 \neq 0$ the underlying AMAP(2) the AMMAP[2] loses one degree of freedom, because the coefficient of $q_{1,1}$ in $D_{1,2}$ is null. In this case, expressions simplify to

$$p_1 = q_{2,1} - r_2 q_{2,1} + r_2 q_{2,2} \tag{34}$$

$$F_{1,1} = h_2 - \frac{h_1 q_{2,1}(r_2 - 1)}{r_2 q_{2,2} - q_{2,1}(r_2 - 1)} \tag{35}$$

$$B_{1,1} = h_1 + h_2 - h_1 r_2. \tag{36}$$

Inverting this formulas for fitting p_1 and $F_{1,1}$ we obtain:

$$q_{2,1} = \frac{p_1 h_2 - p_1 F_{1,1}}{h_1(r_2 - 1)} \tag{37}$$

$$q_{2,2} = \frac{p_1 h_1 + p_1 h_2 - p_1 F_{1,1}}{h_1 r_2} \tag{38}$$

- If $r_2 = 0$, the underlying MAP process degenerates into an APH(2) renewal process. Therefore, the AMPH fitting approach defined in Section IV-A applies.

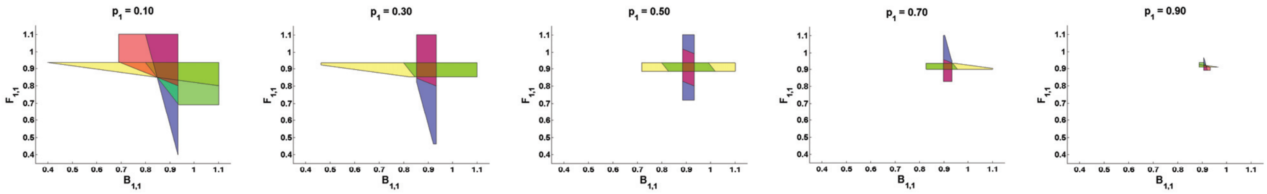

Figure 2. Feasible regions for the multi-class moments of AMMAP[2] processes based on four AMAP(2) with the same first three moments and autocorrelation decay. Each figure refers to a different value of the probability of arrivals of class one. The regions shaded in different colors represent the feasible values of the conditional first-order moments.

C. AMMAP[2] with negative autocorrelation decay

The AMMAP[2] process with negative autocorrelation decay is:

$$D_0 = \begin{bmatrix} -\frac{1}{h_1} & \frac{r_1}{h_1} \\ 0 & -\frac{1}{h_2} \end{bmatrix} \tag{39}$$

$$D_{1,1} = \begin{bmatrix} 0 & \frac{q_{1,1}(1-r_1)}{h_1} \\ \frac{q_{2,1}(1-r_2)}{h_2} & \frac{q_{2,2}r_2}{h_2} \end{bmatrix} \tag{40}$$

$$D_{1,2} = \begin{bmatrix} 0 & \frac{(1-q_{1,1})(1-r_1)}{h_1} \\ \frac{(1-q_{2,1})(1-r_2)}{h_2} & \frac{(1-q_{2,2})r_2}{h_2} \end{bmatrix}. \tag{41}$$

Multi-class moments and class probabilities of the AMMAP[2] are as follows:

$$\begin{aligned} p_1 =& q_{1,1} - \frac{q_{1,1} - q_{2,1} + r_2 q_{2,1} - r_2 q_{2,2}}{r_1(r_2-1) - r_2 + 2} \\ p_{1,1} =& 2q_{1,1}q_{2,1} + r_2 q_{2,1}^2 - q_{2,1}^2 - \\ & \frac{(2q_{2,1} - r_2 q_{2,1} + r_2 q_{2,2})(q_{1,1} - q_{2,1} + r_2 q_{2,1} - r_2 q_{2,2})}{r_1(r_2-1) - r_2 + 2} - \\ & r_2 q_{1,1} q_{2,1} + r_2 q_{1,1} q_{2,2} - r_2 q_{2,1} q_{2,2} \\ B_{1,1} =& h_1 - \frac{(r_{2,1} - r_2 q_{2,1} + r_2 q_{2,2})(h_1 - h_2 - h_1 r_1 + h_1 r_1 r_2)}{q_{2,1} - r_2 q_{2,1} + r_2 q_{2,2} - q_{1,1}(r_1 + r_2 - r_1 r_2 - 1)} \\ F_{1,1} =& h_2 - \frac{q_{2,1}(r_2 - 1)(h_1 - h_2 + h_2 r_1)}{q_{2,1} - r_2 q_{2,1} + r_2 q_{2,2} - q_{1,1}(r_1 + r_2 - r_1 r_2 - 1)} \end{aligned}$$

Analogously to the results for the first canonical form discussed in Section IV-B, we provide closed analytical formulas for fitting:

- p_1, $B_{1,1}$ and $F_{1,1}$ in Section IV-C1;
- p_1, $F_{1,1}$ and $p_{1,1}$ in Section IV-C2;
- p_1, $B_{1,1}$ and $p_{1,1}$ in Section IV-C3.

For each of this case we provide the closed explicit formulas to be used for fitting. In Section IV-C4 we discuss degenerate cases that require special attention due to the loss of one or two degrees of freedom because of zeros appearing in the D_0 and D_1 of the underlying AMAP(2).

1) Fitting p_1, $B_{1,1}$ and $F_{1,1}$: Solving the expressions of p_1, $B_{1,1}$ and $F_{1,1}$ for the parameters $q_{1,1}$, $q_{1,2}$, $q_{2,2}$ yields

$$\begin{aligned} q_{1,1} =& -\frac{p_1(B_{1,1} - h_2 - h_1 r_1 + h_1 r_1 r_2)(r_1 + r_2 - r_1 r_2 - 2)}{(r_1 - 1)(r_2 - 1)(h_1 - h_2 - h_1 r_1 + h_1 r_1 r_2)} \\ q_{2,1} =& \frac{p_1(F_{1,1} - h_2)(r_1 + r_2 - r_1 r_2 - 2)}{(r_2 - 1)(h_1 - h_2 + h_2 r_1)} \\ q_{2,2} =& \frac{p_1(B_{1,1} - h_1)(r_1 + r_2 - r_1 r_2 - 2)}{r_2(h_1 - h_2 - h_1 r_1 + h_1 r_1 r_2)} - \frac{p_1(r_1 + r_2 - r_1 r_2 - 2)}{r_2(r_1 - 1)} \\ & + \frac{p_1(h_1 - F_{1,1} + F_{1,1} r_1)(r_1 + r_2 - r_1 r_2 - 2)}{r_2(h_1 + h_2(r_1 - 1))(r_1 - 1)} \end{aligned}$$

Feasible region. When the autocorrelation decay is negative, there are at most two feasible fits for the underlying AMAP(2). Similarly to the case with positive autocorrelation decay, the choice of the underlying APH(2) impacts the feasible region of the first forward and backward moments, as shown in Figure 3.

Approximate fitting. Similarly to Section IV-B, when the values of $F_{1,1}$ or $p_{1,1}$ are infeasible, we find the closest feasible values. When multiple underlying AMAP(2) are feasible we select the one which leads to the smallest error. However, for the AMAP(2) with negative correlation decay only two forms, instead of four, exist.

2) Fitting forward moments and class probabilities: For convenience of notation we define

$$\begin{aligned} w_6 &= h_1 + h_2 - h_1 r_2 + F_{1,1}(r_1 + r_2 - r_1 r_2 - 2) \\ w_7 &= r_2(h_1 - h_2 + h_2 r_1) w_6 \end{aligned}$$

The explicit formulas for the parameters of the second-order AMMAP[2] as a function of p_1, $F_{1,1}$ and $p_{1,1}$ are:

$$\begin{aligned} q_{1,1} &= (p_1(r_1 - 1)w_6)^{-1} \textstyle\sum_{i=1}^{3} x_i'' \\ q_{2,1} &= \frac{p_1(F_{1,1} - h_2)(r_1 + r_2 - r_1 r_2 - 2)}{(r_2 - 1)(h_1 - h_2 + h_2 r_1)} \\ q_{2,2} &= w_7^{-1} \left[p_1 \left(\textstyle\sum_{i=1}^{6} y_i'' - h_1 \sum_{i=7}^{9} y_i'' \right) + p_1^{-1} \sum_{i=1}^{4} u_i'' \right] \end{aligned}$$

where the y_i'' and u_i'' definitions are given in the appendix.

Feasible region and approximate fitting. Similarly to the case with positive autocorrelation decay discussed in Section IV-B4, the feasible region is defined by nonlinear inequalities. When $(F_{1,1}, p_{1,1})$ is unfeasible, we find the closest point in the feasible region, solving a nonlinearly constrained least squares problem.

3) Fitting backward moments and class probabilities: For the sake of compactness we define

$$\begin{aligned} w_8 &= h_1 - h_2 - h_1 r_1 + h_1 r_1 r_2 \\ w_9 &= h_1 + h_2 - h_1 r_2 + B_{1,1}(r_1 + r_2 - r_1 r_2 - 2) \\ w_{10} &= B_{1,1} - h_2 - h_1 r_1 + h_1 r_1 r_2 \\ w_{11} &= p_1 r_2 w_8 w_9 \end{aligned}$$

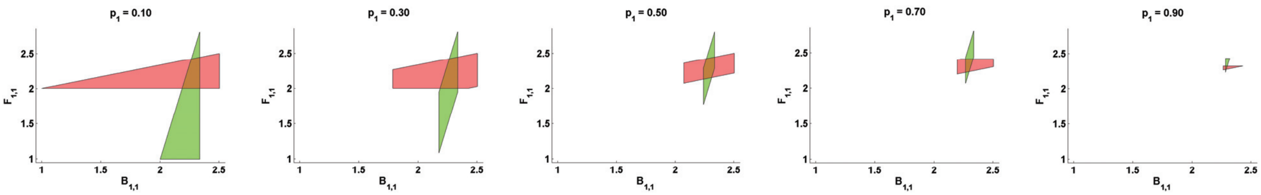

Figure 3. Feasible regions for the multi-class moments of AMMAP[2] processes based on four AMAP(2) with the same first three moments and autocorrelation decay. Each figure refers to a different value of the probability of arrivals of class one. The regions shaded in different colors represent the feasible values of the conditional first-order moments.

The explicit expression for the parameters of the AMMAP[2] are obtained as a function of p_1, $B_{1,1}$ and $p_{1,1}$ as follows:

$$
\begin{aligned}
q_{1,1} &= -\frac{p_1 w_{10}(r_1 + r_2 - r_1 r_2 - 2)}{(r_1-1)(r_2-1)w_8} \\
q_{2,1} &= -p_1^{-1} w_9^{-1} \left(x_1''' + x_2''' + x_3''' + p_{1,1} h_1 r_1 r_2\right) \\
q_{2,2} &= w_{11}^{-1} \left[\sum_{i=1}^{6} y_i''' + p_1 \left(\sum_{i=7}^{15} y_i''' + h_2 \sum_{i=16}^{19} y_1'''\right)\right]
\end{aligned}
$$

where x_i''' and y_i''' are defined in the final appendix.

Feasible region and approximate fitting. When $(B_{1,1}, p_{1,1})$ is unfeasible, analogously to the previous case, we find the closest point in the feasible region, solving a nonlinearly constrained least squares problem.

4) Degenerate cases: The following degenerate cases arise when fitting a second-order AMMAP[2] with negative autocorrelation decay.

- If $r_2 = 1$, we fit a marked Poisson process.
- If $r_2 = 0$ and $r_1 = 1$ two degrees of freedom are lost and the process becomes a AMPH[2] with a single degree of freedom. This is the degenerate case for the AMPH[2] discussed in Section IV-A.
- If $r_2 = 0$ and $r_1 \neq 1$ one degree of freedom is lost and the process becomes a degenerate case of the AMMAP[2]. We define the following auxiliary variables, for the sake of compactness,

$$
\begin{aligned}
z_1 &= h_2 r_1 r_2 - 1 \\
z_2 &= h_1 - h_2 + h_2 r_1 - h_1 h_2 r_2 \\
z_3 &= r_1 - h_2 r_2 + h_2 r_1 r_2 - 2 \\
z_4 &= h_2 r_2 + h_2^2 r_2^2 - h_2^2 r_1 r_2^2 - h_2 r_1 r_2 + 1
\end{aligned}
$$

  and obtain

$$
\begin{aligned}
q_{1,1} &= -\frac{p_1 z_1 (h_1 - F_{1,1} + h_2 r_1 + F_{1,1} h_2 r_1 r_2) z_3}{(r_1 - 1) z_2 z_4} \\
q_{2,1} &= -\frac{p_1 z_1 (h_2 - F_{1,1} + h_1 h_2 r_2 + F_{1,1} h_2 r_1 r_2) z_3}{(h_2 r_2 - h_2 r_1 r_2 + 1) z_2}
\end{aligned}
$$

- If $r_1 = 1$, the AMMAP[2] loses one degree of freedom. This case in common with the degenerate case of the first canonical form described in Section IV-B4 because the D_1 matrix has the same structure.

V. MMAP[m] SEQUENTIAL FITTING

Given a marked arrival sequence (X_i, C_i), $i \geq 1$, we define the aggregated sequence for class c as a two-class marked arrival sequence (X_i, C_i'), where $C_i' = 1$ if $C_i = c$, and $C_i' = 2$ otherwise. In a MMAP[m] of order n, the aggregated process for class c is defined as the MMAP[2] of order n with representation

$$(D_0, D_{1,1}^c, D_{1,2}^c) = (D_0, D_{1,c}, \textstyle\sum_{k \neq c} D_{1,k})$$

We now investigate methods to exploit the aggregation mechanism for fitting second-order AMMAP[m] with an arbitrary number of classes. Recall that the fitting process we have defined for a second-order AMMAP[2] starts with a single-class canonical AMAP(2) and successively obtains a second-order AMMAP[2] that fits two classes. The method we propose for the second-order AMMAP[m] is similar: we start with a single class AMAP(2) and perform m fitting steps, one for each class. At step i, we fit the aggregated sequence for class i using any of the fitting algorithms presented in the previous section and obtain the representation $(D_0, D_{1,1}^i, D_{1,2}^i)$ subject to $D_{1,1}^i + D_{1,2}^i = D_1$. We then store $D_{1,1}^i$ and increase i. Finally, we return the AMMAP $(D_0, D_{1,1}^1, D_{1,1}^2, \ldots, D_{1,1}^m)$ which is obtained by assembling all the matrices $D_{1,c}$ obtained in the sequential fitting.

Notice that in this approach we exploit the important property of the forward and backward moments to be insensitive to aggregation. This follows immediately by (4)-(5), since all the terms appearing in the formulas are not affected by aggregation. Thus, for a n-order MMAP[m], $F_{j,c}$ and $B_{j,c}$ are preserved in both the aggregated marked sequence and in the aggregated process for class c. Thus, this approach is particularly appealing for fitting AMPH[m] and AMMAP[m] where we have shown that first-order forward and backward moments uniquely determine the process.

Approximate fitting. When some of the first and backward moments are infeasible, independent adjustments for each class might lead to an invalid MMAP[m], i.e. $\sum_{k=1}^{m} D_{1,k} \neq D_1$. To avoid, this problem, after fitting $D_{1,c}$ (i.e., $D_{1,1}^c$), we check if $\sum_{k=1}^{c} D_{1,k}$ is greater than D_1. If this condition happens the $D_{1,j}$ for $j \geq c$ are set as follows

$$D_{1,j} = \frac{p_j}{\sum_{j=c}^{m} p_j} \left(D_1 - \sum_{k=1}^{c-1} D_{1,k} \right)$$

Table I
FITTING RESULTS FOR EXAMPLE 1

	p_1	$p_{1,1}$	$p_{2,2}$	$p_{1,2}$
Model	5.00e-01	0.00e+00	0.00e+00	5.00e-01
Trace	5.00e-01	0.00e+00	0.00e+00	5.00e-01
Fitted	5.00e-01	1.15e-02	1.15e-02	4.89e-01
	$F_{1,1}$	$F_{1,2}$	$F_{2,1}$	$F_{2,2}$
Model	1.00e+00	5.00e-01	2.00e+00	5.00e-01
Trace	1.00e+00	4.99e-01	2.02e+00	4.97e-01
Fitted	1.00e+00	4.99e-01	2.01e+00	5.04e-01
	$B_{1,1}$	$B_{1,2}$	$B_{2,1}$	$B_{2,2}$
Model	5.00e-01	1.00e+00	5.00e-01	2.00e+00
Trace	4.99e-01	1.00e+00	4.97e-01	2.02e+00
Fitted	4.99e-01	1.00e+00	5.04e-01	2.01e+00

Table II
FITTING RESULTS FOR EXAMPLE 2

	p_1	$p_{1,1}$	$p_{2,2}$	$p_{1,2}$
Model	4.35e-01	1.84e-01	3.14e-01	2.51e-01
Trace	4.36e-01	1.85e-01	3.12e-01	2.51e-01
Fitted	4.36e-01	1.87e-01	3.14e-01	2.50e-01
	$F_{1,1}$	$F_{1,2}$	$F_{2,1}$	$F_{2,2}$
Model	1.10e+00	9.21e-01	2.41e+00	1.84e+00
Trace	1.11e+00	9.18e-01	2.44e+00	1.84e+00
Fitted	1.11e+00	9.18e-01	2.43e+00	1.84e+00
	$B_{1,1}$	$B_{1,2}$	$B_{2,1}$	$B_{2,2}$
Model	9.88e-01	1.01e+00	2.05e+00	2.12e+00
Trace	9.91e-01	1.01e+00	2.08e+00	2.12e+00
Fitted	9.91e-01	1.01e+00	2.07e+00	2.12e+00

This procedure assures that the class probabilities p_j are matched exactly for all the classes and the forward and backward moments $F_{1,j}$ and $B_{1,j}$ are optimally approximated for $j = 1, \ldots, c-1$. The adjustment for the remaining classes is not optimal, but it turns out to be good in practice if we take special care in ordering the classes, i.e., we fit first the classes whose forward and backward moments are closer to M_1 so that the class is more likely to be fitted exactly.

Cross-moments. The reasoning behind the sequential fitting can be easily extended to cross-moments when applied to the n-order MMAP$[m]$ process, where aggregation may be defined to leave a subset of classes explicit. For such classes, their cross moments would be unaffected, thus reducing the complexity fitting problem provided that a fitting scheme for cross-moments is defined. We leave this interesting research direction for future work.

VI. VALIDATION

In this section, we show the accuracy of the proposed fitting method on traces generated by MMAP$[m]$ of different orders. Our validation considers both exact and approximate fitting. We also illustrate queueing performance results for data obtained from the SNIA repository for the Microsoft Live Maps Back End (*http://iotta.snia.org*).

If not otherwise stated, the number of samples used to fit a trace is 10^5. Also, unless specified, the fitting formulas are always the one for second-order AMMAP$[2]$ for fitting p_1, $B_{1,1}$ and $F_{1,1}$.

Table III
FITTING RESULTS FOR EXAMPLE 3

	p_1	$p_{1,1}$	$p_{2,2}$	$p_{1,2}$
Model	5.32e-01	2.81e-01	2.16e-01	2.52e-01
Trace	5.32e-01	2.80e-01	2.17e-01	2.51e-01
Fitted	5.32e-01	2.81e-01	2.18e-01	2.51e-01
	$F_{1,1}$	$F_{1,2}$	$F_{2,1}$	$F_{2,2}$
Model	8.57e-01	8.00e-01	1.65e+00	1.43e+00
Trace	8.57e-01	7.97e-01	1.66e+00	1.42e+00
Fitted	8.57e-01	7.97e-01	1.63e+00	1.45e+00
	$B_{1,1}$	$B_{1,2}$	$B_{2,1}$	$B_{2,2}$
Model	8.10e-01	8.54e-01	1.49e+00	1.62e+00
Trace	8.09e-01	8.51e-01	1.48e+00	1.62e+00
Fitted	8.09e-01	8.51e-01	1.49e+00	1.62e+00

Table IV
FITTING RESULTS FOR EXAMPLE 4

	p_1	p_2	p_3	p_4
Model	1.12e-01	3.26e-01	4.62e-01	9.97e-02
Trace	1.12e-01	3.27e-01	4.62e-01	9.94e-02
Fitted	1.12e-01	3.27e-01	4.62e-01	9.94e-02
	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$
Model	1.39e-02	3.09e-02	5.47e-02	1.28e-02
Trace	1.33e-02	3.05e-02	5.52e-02	1.27e-02
Fitted	1.38e-02	3.11e-02	5.43e-02	1.26e-02
	$F_{1,1}$	$F_{1,2}$	$F_{1,3}$	$F_{1,4}$
Model	2.80e-02	2.24e-02	1.69e-02	2.47e-02
Trace	2.77e-02	2.21e-02	1.67e-02	2.51e-02
Fitted	2.77e-02	2.21e-02	1.67e-02	2.51e-02
	$B_{1,1}$	$B_{1,2}$	$B_{1,3}$	$B_{1,4}$
Model	2.35e-02	1.64e-02	2.22e-02	2.45e-02
Trace	2.34e-02	1.65e-02	2.20e-02	2.43e-02
Fitted	2.34e-02	1.65e-02	2.20e-02	2.43e-02

A. Representative Examples

Example 1: Degenerate MMAP$[2]$ of order 2. Consider the following MMAP, discussed in [4], which generates a marked arrival sequence where the class markings alternate deterministically:

$$D_0 = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}, D_{1,1} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, D_{1,2} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad (42)$$

This model is an interesting test for our fitting algorithms since it is a degenerate case: due to the alternating nature of the arrivals of the two classes, the backward moment of the first class coincides with the forward moment of the second class. Fitting results are shown in Table I. Differently from the original model, in the fitted process the probability $p_{c,k}$ is not null, but quite small. The error is mainly due to the difficulty of estimating the autocorrelation decay rate for the underlying MAP, which has the degenerate value -1 and thus it is difficult to estimate from the empirical trace.

Example 2: Non-Acyclic MMAP$[2]$ of order 2. In this example we show that a second-order AMMAP$[2]$ can successfully capture the behavior of a second-order *non-acyclic* MMAP$[2]$. The MMAP$[2]$ is as follows:

$$D_0 = \begin{bmatrix} -1 & 1/3 \\ 1/4 & -2 \end{bmatrix}, D_{1,1} = \begin{bmatrix} 1/6 & 1/12 \\ 3/4 & 1/10 \end{bmatrix}, D_{1,2} = \begin{bmatrix} 1/6 & 1/4 \\ 1/2 & 2/5 \end{bmatrix}$$

Fitting results are shown in Table II. Even though we are only fitting first-order forward and backward moments, this example shows that the second-order multi-class moments are fitted with a relative error smaller than 1%.

Example 3: Non-Acyclic MMAP[2] *of order* 4. In this example we show that even higher order processes can be well approximated by a second-order AMMAP[2]. We consider the following non-acyclic MMAP[2] of order 4:

$$D_0 = \begin{bmatrix} -1 & 1/8 & 1/8 & 1/8 \\ 0 & -2 & 2/5 & 2/5 \\ 1/5 & 1/5 & -3 & 1 \\ 1/2 & 1/4 & 3/4 & -4 \end{bmatrix},$$

$$D_{1,1} = \begin{bmatrix} 1/8 & 1/40 & 1/56 & 1/10 \\ 3/10 & 2/25 & 1/10 & 1/5 \\ 1/25 & 4/25 & 9/25 & 8/25 \\ 2/3 & 1/4 & 9/20 & 1/14 \end{bmatrix},$$

$$D_{1,2} = \begin{bmatrix} 1/8 & 1/10 & 6/56 & 1/40 \\ 1/10 & 3/25 & 1/10 & 1/5 \\ 4/25 & 6/25 & 6/25 & 2/25 \\ 1/12 & 1/2 & 6/20 & 5/28 \end{bmatrix}$$

This is a case where the autocorrelation decay in the underlying canonical AMAP(2) is positive. Fitting results are shown in Table III. Both the first-order forward and backward moments are matched exactly. Also second-order multi-class moments are fitted accurately with a 2% error.

Example 4: Acyclic MMAP[4] *of order* 2. We now apply the sequential fitting approach discussed in Section V. We consider a trace generated by a second-order AMMAP[4] with negative autocorrelation decay in the underlying canonical AMAP(2). This AMMAP[4] is defined as follows:

$$D_0 = \begin{bmatrix} -2 & 1/2 \\ 0 & -5 \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 0 & 3/2 \\ 1 & 4 \end{bmatrix} \tag{43}$$

Let $d_{i,j}$ be the element in column j and row i of matrix D_1. The matrices $D_{1,c}$ are obtained from matrices $D^*_{1,c} = [d^*_{i,j,c}]$ whose elements are defined as follows

$$d^*_{i,j,c} = d_{i,j} \left\{ \left[(n \cdot i + m \cdot j)^c \mod 2c+1 \right] + 1 \right\} / G \tag{44}$$

where n is the number of states, m is the number of classes, and G is a constant ensuring that $\sum_c d^*_{i,j,c} = d_{i,j}$. Fitting results in Table IV shows that the forward and backward moments of the first order are matched for all classes. The probabilities $p_{c,k}$ are well approximated with a relative error of just 4%.

Example 5: Non-Acyclic MMAP[4] *of order* 4. We now apply sequential fitting to a trace generated by a fourth-order MMAP[4], with D_0 defined as in Example 3 and

$$D_1 = \begin{bmatrix} 2/8 & 1/8 & 1/8 & 1/8 \\ 2/5 & 1/5 & 1/5 & 2/5 \\ 1/5 & 2/5 & 3/5 & 2/5 \\ 3/4 & 3/4 & 3/4 & 1/14 \end{bmatrix} \tag{45}$$

The matrices $D_{1,c}$ are obtained as in the previous example. Results in Table V show that the first-order forward and backward moments are fitted exactly for all classes. The probabilities $p_{c,k}$ are matched with a relative error of less than 3% .

Table V
FITTING RESULTS FOR EXAMPLE 5

	p_1	p_2	p_3	p_4
Model	1.45e-01	2.18e-01	3.57e-01	2.79e-01
Trace	1.47e-01	2.19e-01	3.56e-01	2.78e-01
Fitted	1.47e-01	2.19e-01	3.56e-01	2.78e-01
	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$
Model	2.15e-02	3.20e-02	5.09e-02	4.10e-02
Trace	2.20e-02	3.33e-02	5.13e-02	4.04e-02
Fitted	2.17e-02	3.25e-02	5.19e-02	4.09e-02
	$F_{1,1}$	$F_{1,2}$	$F_{1,3}$	$F_{1,4}$
Model	6.80e-02	6.93e-02	6.43e-02	6.51e-02
Trace	6.90e-02	6.87e-02	6.43e-02	6.63e-02
Fitted	6.90e-02	6.87e-02	6.43e-02	6.63e-02
	$B_{1,1}$	$B_{1,2}$	$B_{1,3}$	$B_{1,4}$
Model	6.82e-02	6.90e-02	6.31e-02	6.68e-02
Trace	6.82e-02	6.92e-02	6.35e-02	6.73e-02
Fitted	6.82e-02	6.92e-02	6.35e-02	6.73e-02

Table VI
RANDOM AMMAP[2] MODELS RESULTS.

Samples	States	Forward Error	Backward Error
10^3	2	9.262e-02	7.831e-02
10^4	2	1.071e-02	3.117e-02
10^5	2	2.618e-02	4.137e-02
10^3	4	6.422e-02	7.209e-02
10^4	4	2.278e-02	2.813e-02
10^5	4	4.923e-02	2.895e-02
10^3	8	1.101e-01	6.157e-02
10^4	8	1.833e-02	1.459e-02
10^5	8	1.126e-02	8.592e-03

B. Random Campaign

We now evaluate the fitting approach on traces from randomly generated MMAPs. The number of samples in each trace varies between 10^3 and 10^5, while the number of states n of the generating MMAP varies between 2 and 8; the number of classes is kept fixed at $m = 4$. Considering one of the randomly generated MMAPs, from which 10^3 arrivals were sampled, the 95% confidence intervals of the first-order forward and backward moments of the first class are $[5.7866, 7.8527]$ and $[4.4802, 5.8292]$. For another randomly generated MMAP, for which 10^5 arrivals were samples, the 95% confidence intervals are significantly tighter: $[16.6778, 17.1908]$ and $[12.3978, 12.8830]$ for the first-order forward and backward moments of class one, respectively. For each sample size and value of n we perform 10 trials. Results are shown in Table VI. For each test case we report the median, over 10 trials, of the maximum relative error, over all the workload classes, of the first forward and backward moments. The results indicate that the approach is very accurate.

C. Real Traces and Queueing Results

As a final case study, we consider a real-world trace gathered from a storage system and obtained from the SNIA repository (*http://iotta.snia.org*). The trace represents an hour of the original Microsoft Live Maps Back End, relatively to a subset collected between 10:30AM and 11:29AM on February 21, 2008. A description of this trace can be found in [13]. The trace contains the inter-issue times of read and write requests to a storage drive.

We model read and writes as separate classes and study the survival function of the queue length for a queueing system with exponentially distributed service times, first-come first-served scheduling policy, infinite buffer capacity and a single server. We have simulated this queue by replaying the trace to generate the arrivals and collecting a total of 10^8 samples. The comparison is performed by evaluating MAP(2), second-order AMPH$[2]$ and AMMAP$[2]$ processes fitted to the trace. For the AMMAP$[2]$, we have considered the fitting methods discussed in Section IV, respectively denoted as $F+B$ (Section IV-B1), $F+p$ (Section IV-B2), and $B+p$ (Section IV-B3). Simulation results are then compared with the results obtained using the MMAP/M/1 queue evaluated by the QMAM toolbox [10].

In Figure VI-C, we consider the arrival process to a disk and show the survival functions for different choices for the service rates of the two classes. The service rates are chosen such that the utilization of the system is 70%. We have considered two different cases (from left to right). In the first figure the utilization is equally split between read and write requests. In the second figure, 7% of the utilization is due to read requests and 63% is due to write requests. The figure indicates that MMAPs provide a better fit than MAPs and renewal MPHs for this trace. This is because this trace is strongly temporal dependent with lag-1 autocorrelation $\rho_1 = 0.1674$ and estimated geometric autocorrelation decay rate of 0.9956. We find that for this specific example $F+p$ and $B+p$ provide the best results, whereas $F+B$ improves over the MAP, but less than the other two fitting methods. We have further investigated into this and found that in this case the MMAP$[2]$ is able to fit well the backward moment, but the feasibility region prevents accurate fitting of the forward moment. This illustrates the additional value of leveraging three different fitting schemes, which can mutually compensate for such problems. These results justify the choice of MMAPs as a model for the arrival processes in real systems. With a modest increase in the number of parameters, MMAPs provide more fitting flexibility that the simpler models such as MAPs or MPHs, striking a good balance between model complexity and accuracy.

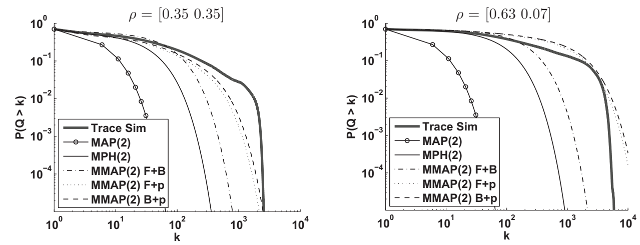

Figure 4. $Trace/M/1$ queueing results for a SNIA trace: Microsoft Live Maps Back End Disk Trace. Read and write requests are modeled as different classes with different service times. Utilization is 70%.

VII. Conclusion

In this paper, we have presented a methodology for analytical fitting of second-order acyclic marked Markovian arrival processes (AMMAPs). We have developed exact fitting formulas that are capable of fitting any second-order AMMAPs with two classes that have an underlying MAP in canonical form. Our approach focuses in particular on fitting backward and forward moments proposed in the recent literature [1], [4] as well as probabilities that describe the class marking process. Based on these results, we have introduced a sequential fitting algorithm that generalizes the approach to second-order AMMAPs with an arbitrary number of classes. Validation results argue for the effectiveness of the approach even using traces extracted from MMAPs of large order and from real-world storage systems.

Our analysis suggests that future work should investigate the application of sequential fitting to MMAPs of larger order, possibly in combination with fitting of cross moments.

Acknowledgement

The research of leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement no. 318484.

References

[1] A. Horváth, G. Horváth, and M. Telek, "A traffic based decomposition of two-class queueing networks with priority service," *Comp. Netw.*, vol. 53, no. 8, 1235 – 1248, 2009.

[2] Q.-M. He, "The versatility of MMAP[K] and the MMAP[K]/G[K]/1 queue," *QUESTA*, vol. 38, no. 4, 397–418, 2001.

[3] J. Artalejo, A. Gomez-Corral, and Q. He, "Markovian arrivals in stochastic modelling: a survey and some new results," *Stat. and Oper. Res. Trans.*, 34(2), 101–144, 2010.

[4] P. Buchholz, P. Kemper, and J. Kriege, "Multi-class Markovian arrival processes and their parameter fitting," *Perform. Eval.*, 67(11), 1092 – 1106, 2010.

[5] G. Horváth, M. Telek, "On the canonical representation of phase type distributions," *Perf. Eval.*, 66(8), 396–409, 2009.

[6] L. Bodrog, A. Heindl, G. Horváth, and M. Telek, "A markovian canonical form of second-order matrix-exponential processes," *EJOR*, 190, no. 2, 459–477, 2008.

[7] A. Heindl, G. Horváth, and K. Gross, "Explicit inverse characterizations of acyclic MAPs of second order," in *EPEW*, ser. Lecture Notes in Computer Science, A. Horváth and M. Telek, Eds., 4054. Springer, 2006, 108–122.

[8] L. Breuer, "An EM algorithm for batch markovian arrival processes and its comparison to a simpler estimation procedure," *Annals of Oper. Res.*, 112, 123–138, 2002.

[9] A. Klemm, C. Lindemann, and M. Lohmann, "Modeling IP traffic using the batch markovian arrival process," *Perform. Eval.*, 54(2), 149 – 173, 2003, Modelling Techniques and Tools for Computer Perf. Evaluation.

[10] D. Bini, B. Meini, S. Steffé, J. F. Pérez, and B. V. Houdt, "SMCSolver and Q-MAM: tools for matrix-analytic methods," *ACM Perf. Eval. Rev.*, 39(4), p. 46, 2012.

[11] J. Kriege and P. Buchholz, "Correlated phase-type distributed random numbers as input models for simulations," *Perform. Eval.*, 68(11), 1247–1260, 2011.

[12] L. Bodrog, A. Heindl, G. Horvath, and M. Telek, "A Markovian canonical form of second-order matrix-exponential processes," *EJOR*, 190(2), 459 – 477, 2008.

[13] S. Kavalanekar, B. L. Worthington, Q. Zhang, and V. Sharda, "Characterization of storage workload traces from production Windows servers," in *IISWC*, 2008, 119–128.

APPENDIX

The following symbols appear in Section IV.

A. AMMAP[2] with positive autocorrelation decay

- Fitting forward moments and class probabilities

$$\begin{aligned}
x_1 =& F_{1,1}^2 r_1 r_2 \left(-r_1 r_2 + 2r_2 + 2\right) \\
x_2 =& F_{1,1} \left(F_{1,1} r_2^2 - 2F_{1,1} r_2 + F_{1,1} - h_1 r_1^2 r_2^2\right) \\
x_3 =& F_{1,1} h_1 r_1 r_2 \left(r_1 + 3r_2 - 4\right) \\
x_4 =& F_{1,1} h_1 \left(r_1 - 2r_2^2 + 4r_2\right) \\
x_5 =& F_{1,1} \left(-2h_1 - h_2 r_1^3 r_2^2 + h_2 r_1^3 r_2\right) \\
x_6 =& F_{1,1} h_2 r_1^2 \left(2r_2^2 - 5r_2 + 1\right) \\
x_7 =& F_{1,1} h_2 r_1 \left(-r_2^2 + 4r_2 - 3\right) \\
x_8 =& h_1^2 \left(-r_1 r_2^2 + r_1 r_2 + r_2^2 - 2r_2\right) \\
x_9 =& h_1 \left(h_1 - h_2 r_1^2 r_2^2 + h_2 r_1^2 r_2 + h_2 r_1 r_2^2\right) \\
x_{10} =& h_2 r_1 \left(-2h_1 r_2 + h_1 + h_2 r_1 r_2\right) \\
x_{11} =& h_2^2 r_1 \left(1 - r_2\right) \\
y_1 =& p_{1,1} r_1 \left(2h_2^2 r_1 - h_2^2 r_1^2 - h_1^2\right) \\
y_2 =& p_{1,1} r_1 \left(-h_2^2 - 2h_1 h_2 r_1 + h_1^2 r_2\right) \\
y_3 =& p_{1,1} h_2^2 r_1 r_2 \left(1 - 2r_1 + r_1^2\right) \\
y_4 =& 2p_{1,1} h_1 h_2 r_1 \left(1 - r_2 + r_1 r_2\right) \\
u_1 =& \left[r_1 r_2 (h_1 - h_2 + h_2 r_1)\right]^{-1}
\end{aligned}$$

- Fitting backward moments and class probabilities

$$\begin{aligned}
y_1' =& p_{1,1} r_1 h_1 r_2^2 \left(h_1 r_2 - h_1 - 2\right) \\
y_2' =& p_{1,1} r_1 h_2 \left(-2h_1 r_2 - h_2 r_2 + h_2\right) \\
x_1' =& -B_{1,1}^2 r_1 r_2 \left(r_1 r_2 - 2r_2 + 2\right) \\
x_2' =& B_{1,1} \left(B_{1,1} r_2^2 - 2B_{1,1} r_2 + B_{1,1} + h_1 r_1^2 r_2^3\right) \\
x_3' =& B_{1,1} h_1 r_1 r_2^2 \left(4 - r_1 - r_2\right) \\
x_4' =& B_{1,1} h_1 r_2 \left(4 - 3r_1 - 2r_2\right) \\
x_5' =& -B_{1,1} \left(2h_1 + h_2 r_1^2 r_2^2 + h_2 r_1^2 r_2\right) \\
x_6' =& B_{1,1} h_2 r_1 \left(r_2 - 1\right) \left(r_2 + 1 - h_1^2 r_1 r_2\right) \\
x_7' =& h_1 \left(r_2 - 1\right) \left[h_1 \left(r_2 - 1\right) - h_2 r_1^2 r_2\right] \\
x_8' =& h_2 r_1 \left(h_1 r_2^2 - 2h_1 r_2 + h_1 + h_2 r_1 r_2\right) \\
x_9' =& h_2^2 r_1 \left(1 - r_2\right)
\end{aligned}$$

B. AMMAP[2] with negative autocorrelation decay

- Fitting forward moments and class probabilities

$$\begin{aligned}
x_1'' &= p_{1,1} h_1 - p_{1,1} h_2 + 2p_1^2 F_{1,1} - 2p_1^2 h_1 + p_1^2 F_{1,1} r_1^2 \\
x_2'' &= p_{1,1} h_2 r_1 - 3p_1^2 F_{1,1} r_1 - p_1^2 F_{1,1} r_2 + p_1^2 h_1 r_1 \\
x_3'' &= p_1^2 r_2 \left(h_1 + 2F_{1,1} r_1 - h_1 r_1 - F_{1,1} r_1^2\right) \\
y_1'' &= F_{1,1}^2 r_1 \left(r_1 r_2^2 - 2r_1 r_2 + r_1 - 2r_2^2\right) \\
y_2'' &= F_{1,1}^2 \left(6r_1 r_2 - 4r_1 + r_2^2 - 4r_2\right) \\
y_3'' &= F_{1,1} \left(4F_{1,1} - h_2 r_1^2 r_2^2 + h_2 r_1^2 r_2\right) \\
y_4'' &= F_{1,1} h_2 r_1 \left(2r_2^2 - 5r_2 + 2\right) \\
y_5'' &= h_2 \left(-F_{1,1} r_2^2 + 4F_{1,1} r_2 - 4F_{1,1} + h_2 r_1^2 r_2\right) \\
y_6'' &= h_2^2 r_1 \left(2 - r_1 - r_2\right) \\
y_7'' &= 2\left(2F_{1,1} - 2h_2 - F_{1,1} r_1 - 2F_{1,1} r_2 + h_2 r_1\right) \\
y_8'' &= r_2 \left(4h_2 + F_{1,1} r_2 - h_2 r_2 + h_2 r_1 r_2\right) \\
y_9'' &= r_1 r_2 \left(3F_{1,1} - 3h_2 - F_{1,1} r_2\right) \\
u_1'' &= p_{1,1} \left(h_1^2 + h_2^2 + h_2^2 r_1^2 - 2h_1 h_2\right) \\
u_2'' &= -p_{1,1} \left(h_1^2 r_2 + 2h_2^2 r_1 + h_2^2 r_2\right) \\
u_3'' &= p_{1,1} h_2 r_1 \left(2h_2 r_2 - h_2 r_1 r_2 + 2h_1\right) \\
u_4'' &= 2p_{1,1} h_1 h_2 r_2 \left(1 - r_1\right)
\end{aligned}$$

- Fitting backward moments and class probabilities

$$\begin{aligned}
x_1''' &= p_{1,1} h_1 - p_{1,1} h_2 + 2p_1^2 B_{1,1} - 2p_1^2 h_1 \\
x_2''' &= -p_{1,1} h_1 r_1 - p_1^2 B_{1,1} r_1 - p_1^2 B_{1,1} r_2 + p_1^2 h_1 r_1 \\
x_3''' &= p_1^2 h_1 r_2 + p_1^2 B_{1,1} r_1 r_2 - p_1^2 h_1 r_1 r_2 \\
y_1''' &= -p_{1,1} h_1^2 r_1^2 r_2^3 + 3p_{1,1} h_1^2 r_1^2 r_2^2 - 3p_{1,1} h_1^2 r_1^2 r_2 \\
y_2''' &= p_{1,1} h_1^2 r_1^2 - 2p_{1,1} h_1^2 r_1 r_2^2 + 4p_{1,1} h_1^2 r_1 r_2 \\
y_3''' &= -2p_{1,1} h_1^2 r_1 - p_{1,1} h_1^2 r_2 + p_{1,1} h_1^2 \\
y_4''' &= 2p_{1,1} h_1 h_2 r_1 r_2^2 - 4p_{1,1} h_1 h_2 r_1 r_2 \\
y_5''' &= 2p_{1,1} h_1 h_2 r_1 + 2p_{1,1} h_1 h_2 r_2 - 2p_{1,1} h_1 h_2 \\
y_6''' &= -p_{1,1} h_2^2 r_2 + p_{1,1} h_2^2 \\
y_7''' &= -B_{1,1}^2 r_1^2 r_2^2 + 2B_{1,1}^2 r_1^2 r_2 - B_{1,1}^2 r_1^2 \\
y_8''' &= 2B_{1,1}^2 r_1 r_2^2 - 6B_{1,1}^2 r_1 r_2 + 4B_{1,1}^2 r_1 \\
y_9''' &= -B_{1,1}^2 r_2^2 + 4B_{1,1}^2 r_2 - 4B_{1,1}^2 \\
y_{10}''' &= B_{1,1} h_1 r_1^2 r_2^3 - 2B_{1,1} h_1 r_1^2 r_2^2 + B_{1,1} h_1 r_1^2 r_2 \\
y_{11}''' &= -B_{1,1} h_1 r_1 r_2^3 + 2B_{1,1} h_1 r_1 r_2^2 \\
y_{12}''' &= B_{1,1} h_1 r_1 r_2 - 2B_{1,1} h_1 r_1 + B_{1,1} h_1 r_2^2 \\
y_{13}''' &= -4B_{1,1} h_1 r_2 + 4B_{1,1} h_1 - h_1^2 r_1^2 r_2^3 \\
y_{14}''' &= 3h_1^2 r_1^2 r_2^2 - 3h_1^2 r_1^2 r_2 + h_1^2 r_1^2 + h_1^2 r_1 r_2^3 \\
y_{15}''' &= -4h_1^2 r_1 r_2^2 + 5h_1^2 r_1 r_2 - 2h_1^2 r_1 \\
y_{16}''' &= 4B_{1,1} - 4h_1 - 2B_{1,1} r_1 - 4B_{1,1} r_2 \\
y_{17}''' &= 2h_1 r_1 + 4h_1 r_2 + B_{1,1} r_2^2 - h_1 r_2^2 \\
y_{18}''' &= h_1 r_1 r_2^2 + 3B_{1,1} r_1 r_2 - 3h_1 r_1 r_2 \\
y_{19}''' &= -B_{1,1} r_1 r_2^2
\end{aligned}$$