On observability and design of observers in timed continuous Petri net systems

Jorge Júlvez, Emilio Jiménez, Laura Recalde, and Manuel Silva

***Abstract*—This paper is devoted both to the study of observability criteria and the design of observers in the continuous Petri net setting. The concept of structural observability, regarding the possibility of estimating the marking of places, i.e., the system state, for any speed of the transitions is introduced and studied for the subclass of Join-Free Petri Nets (JF). For non Join-Free Petri Nets, conditions to compute suitable state estimates are established. The proposed observers are piecewise linear systems that assure the continuity of the estimate even when a switch occurs. The system simulation may allow to estimate even the unobservable space of the net system during a given time period.**

Note to practitioners- Petri nets represent a modeling formalism for discrete dynamical systems that offers a great modeling power. Among others, Petri nets have been successfully applied in the fields of manufacture, communication, logistics and traffic. Continuous Petri nets came up to handle the state explosion problem inherent to highly populated discrete systems. The state of a plant modeled with a continuous Petri net is given by a set of real variables. This way, the initial load of the plant has no effect on the complexity of the analysis techniques to be applied.

Obtaining an accurate knowledge of the state of the plant is a crucial task that can determine the feasibility and reliability of subsequent activities as control. The usual way to catch information about the dynamical system under consideration is through sensors located on the physical plant. Unfortunately, in many real situations some state variables cannot be measured through sensors due to either their inaccessible location, the lack of such sensors or the high cost involved in the installation of the sensors. The good news is that the value of those a priori non measurable variables can be estimated if the plant fulfils some technical conditions. Usually, state estimates can be obtained by building a dynamical system called 'observer' whose output is the estimate for the plant.

The paper can be roughly divided into two parts: In the first part, the conditions required to compute estimates are studied. The second part presents a method to design observers whose state converges to the state of the plant.

***Index Terms*—Continuous Petri nets, Observability, Observers, Piecewise linear systems.**

Jorge Júlvez, Laura Recalde, and Manuel Silva are with the Departamento de Informática e Ingeniería de Sistemas, Universidad de Zaragoza, Maria de Luna 3, E-50015 Zaragoza, Spain, {julvez, lrecalde, silva}@unizar.es.

Emilio Jiménez is with the Departamento de Ingeniería Eléctrica, Universidad de La Rioja, E-26004 Logroño, Spain, emilio.jimenez@die.unirioja.es.

This work has been partially supported by projects CICYT DPI2003-06376, DPI2006-15390, HI2006-0149, and grant B106/2001 from D.G.A.

I. Introduction

Petri nets represent a powerful formalism for the modelling of concurrent systems [1], [2]. In particular, stochastic T-timed Petri nets represent a well-known performance evaluation discrete event model [3], [4]. Under heavy loads, discrete event systems often suffer from the state explosion problem. In the framework of observability in discrete event systems, the number of states of an observer is exponential after the observation of a word of bounded length [5]. One way to tackle this problem is to transform the original discrete model into an easier to analyze model that preserves the properties to be studied.

Fluidification is a relaxation technique in which discrete elements of the system are taken as continuous. The fluidification of a timed Petri net system leads to a deterministic piecewise linear system [6], [7]. At a given instant, the differential equations that rule its evolution uniquely depend on the state of the system (marking). Hence, the switch from one linear differential equation system to another one is triggered by an *internal event*, i.e., by a certain change in the marking of the system.

This work focuses on the study of observability and the design of observers in the framework of continuous Petri net systems [8], [6]. Preliminary results on observability and design of observers for continuous Petri nets can be found in [9], [10].

With respect to the study of observability, our attention is first focused on net systems without synchronizations, named Join Free (JF) systems. For this class of net systems a single linear differential equation system describes the system dynamics, thus classical results on observability of linear systems apply here [11], [12], [13]. For JF systems, an effort has been made to introduce and study the concept of *structural observability*. A system is said to be structurally observable if its marking can be estimated independently of the speeds of the transitions.

Afterwards, general Petri net systems including synchronizations will be considered. Such net systems are a subclass of piecewise linear (PL) systems. In contrast to linear systems, the study of observability in PL is a tough problem. There is no link between the observability of a PL system and the observability of the linear systems composing such PL system [14] (the switching among the different linear systems causes this phenomenon). Fortunately the specific features of continuous Petri net systems allow to obtain interesting observability conditions.

Regarding the design of observers, the starting point is to consider one Luenberger observer for each differential equation system that may rule the evolution of the net system (in a similar way to [15]). Each observer will yield an estimate that will be classified as *suitable* or *non-suitable* with respect to the current system output. In this paper an algorithm is proposed that filters out non-suitable estimates and simulates the net system from a given instant in order to compute an estimate for the marking of the system.

The work is structured as follows. In Section II, continuous Petri nets are introduced. In Section III, the observability problem for continuous Petri nets is stated in a similar way to the observability problem for linear systems. In Section IV, we concentrate on JF systems. For this type of systems, structural conditions of observability are obtained from the output of a fix point algorithm. Section V focuses on the computation of estimates for general Petri net systems. Section VI shows how a set of linear observers can be created for a net system and the different classes of non-suitable observers' estimates that can appear. Finally, Section VII sums up the main presented results.

II. Continuous Petri Net Systems

A. Untimed Continuous Petri Net Systems

A PN system is a pair $\langle \mathcal{N}, \mathbf{m_0} \rangle$, where $\mathcal{N}$ specifies the net structure, $\mathcal{N} = \langle P, T, \mathbf{Pre}, \mathbf{Post} \rangle$ and $\mathbf{m_0}$ is the initial marking. The sets of places and transitions are denoted by P and T, respectively. Matrices $\mathbf{Post}$ and $\mathbf{Pre}$ are the arc weight matrices and $\mathbf{C} = \mathbf{Post} - \mathbf{Pre}$ is the token flow matrix. The set of input (output) places of a given set of transitions V is denoted as ${}^\bullet V$ ($V^\bullet$). Respectively, the set of input (output) transitions of a given set of places W is denoted as ${}^\bullet W$ ($W^\bullet$). A net is Join Free (JF) iff every transition has only one input place (for every $t \in T$, $|{}^\bullet t| = 1$).

In continuous PN systems a transition t is *enabled* at a marking $\mathbf{m}$ iff every input place of t is marked (for every $p \in {}^\bullet t$, $\mathbf{m}[p] > 0$). The *enabling degree* at marking $\mathbf{m}$ of a transition measures the maximum amount in which the transition can be fired in a single occurrence, i.e., $\mathrm{enab}(t, \mathbf{m}) = \min_{p \in {}^\bullet t}\{\mathbf{m}[p]/\mathbf{Pre}[p,t]\}$. The firing of t in an amount α, $0 < \alpha \leq \mathrm{enab}(t, \mathbf{m})$ produces a new marking $\mathbf{m}'$, and it is denoted as $\mathbf{m} \xrightarrow{\alpha t} \mathbf{m}'$, where $\mathbf{m}' = \mathbf{m} + \alpha \cdot \mathbf{C}[P, t]$; hence, the state equation $\mathbf{m} = \mathbf{m_0} + \mathbf{C} \cdot \boldsymbol{\sigma} \geq 0$ summarizes the way the marking evolves, where $\boldsymbol{\sigma}$ is the firing count vector (thus $\boldsymbol{\sigma} \geq 0$).

B. Timed Continuous Petri Net Systems

For the timing interpretation, a first order (or deterministic) approximation of the discrete case [16] will be used, assuming that the delays associated to the firing of the transitions can be approximated by their mean values. Each transition t has associated to it an internal firing speed $\boldsymbol{\lambda}[t] > 0$. The state-transition equation has an explicit dependence on time $\mathbf{m}(\tau) = \mathbf{m_0} + \mathbf{C} \cdot \boldsymbol{\sigma}(\tau)$. Differentiating with respect to time, $\dot{\mathbf{m}}(\tau) = \mathbf{C} \cdot \dot{\boldsymbol{\sigma}}(\tau)$ is obtained. Let us denote $\mathbf{f} = \dot{\boldsymbol{\sigma}}$, since it represents the *flow* through the transitions.

Infinite server semantics will be used in the timed model. Under this semantics the flow of a transition is proportional to its enabling degree. More precisely, the flow of a transition t is computed as the product of $\boldsymbol{\lambda}[t]$ by its enabling degree, i.e., $\mathbf{f}[t] = \boldsymbol{\lambda}[t] \cdot \mathrm{enab}(t, \mathbf{m}) = \boldsymbol{\lambda}[t] \cdot \min_{p \in {}^\bullet t}\{\mathbf{m}[p]/\mathbf{Pre}[p,t]\}$.

In JF systems, transitions have only one input place, and so the computation of the enabling degrees does not require the min operator, i.e., they are linear positive systems. Hence, the flow of the transitions can be expressed as $\mathbf{f} = \boldsymbol{\Psi} \cdot \mathbf{m}$ where $\boldsymbol{\Psi}[t,p] = \boldsymbol{\lambda}[t]/\mathbf{Pre}[p,t]$ if $p = {}^\bullet t$, $\boldsymbol{\Psi}[t,p] = 0$ otherwise [1]. Consequently, the evolution of the marking can be described by an equation in the form $\dot{\mathbf{m}} = \mathbf{C} \cdot \mathbf{f} = \mathbf{A} \cdot \mathbf{m}$, where $\mathbf{A} = \mathbf{C} \cdot \boldsymbol{\Psi}$ is a constant matrix.

For a general PN system, matrix $\mathbf{A}$ is not constant but piecewise-constant. The value of $\mathbf{A}$ at a given instant is determined by the marking $\mathbf{m}$ at that instant. To compute $\mathbf{A}$, it is necessary to know the set of places that is actually enabling the transitions, i.e., the set of places that is giving the minimum in the expression for the enabling degree. Once this set is computed, it is easy to establish a linear relationship between the marking of the places in this set and the flow of the transitions: $\dot{\mathbf{m}} = \mathbf{A}(\mathbf{m}) \cdot \mathbf{m}$, with $\mathbf{A}(\mathbf{m}) = \mathbf{C} \cdot \boldsymbol{\Psi}(\mathbf{m})$ where $\boldsymbol{\Psi}[t,p](\mathbf{m}) = \boldsymbol{\lambda}[t]/\mathbf{Pre}[p,t]$ if $p \in {}^\bullet t$ and $\mathbf{m}[p]/\mathbf{Pre}[p,t] = \min_{q \in {}^\bullet t}\{\mathbf{m}[q]/\mathbf{Pre}[q,t]\}$, $\boldsymbol{\Psi}[t,p](\mathbf{m}) = 0$ otherwise.

The PT-set of a system at marking $\mathbf{m}$ is the set of all the pairs, (p, t), such that the marking of p is restricting the flow of transition t at $\mathbf{m}$.

Definition 1: The PT-set of a given system at marking $\mathbf{m}$ is defined as:

$$\text{PT-set}(\mathbf{m}) = \{(p,t) \mid \mathbf{f}[t] = \boldsymbol{\lambda}[t] \cdot \mathbf{m}[p]/\mathbf{Pre}[p,t]\}$$

Notice that each transition is contained at least in one pair of a PT-set (and will belong to more than one only if several of its input places define the same enabling degree). Thus, any PT-set has at least as many pairs as transitions. Obviously, for JF systems a unique PT-set exists. Otherwise, if the PT-set is known, the system evolves according to $\dot{\mathbf{m}} = \mathbf{A}_1 \cdot \mathbf{m}$ where $\mathbf{A}_1$ depends on $\text{PT-set}(\mathbf{m})$ and the $\boldsymbol{\lambda}$ of the transitions. If at a given instant the PT-set changes, i.e., a transition is restricted by other input place, the system will be ruled by another linear system $\dot{\mathbf{m}} = \mathbf{A}_2 \cdot \mathbf{m}$. That is, every PT-set, k, has associated a square matrix $\mathbf{A}_k$ and a linear system $\Sigma_k : \dot{\mathbf{m}} = \mathbf{A}_k \cdot \mathbf{m}$. The set of PT-sets that will be active during the evolution of the system, i.e., *behavioral PT-sets,* depends on the net structure, on $\boldsymbol{\lambda}$, and on the initial marking. If the initial marking is not known, the net structure defines the set of potential PT-sets, i.e., *structural PT-sets,* that might be active.

In order to illustrate the evolution of a non JF system, let us consider the system in Figure 1 with initial marking $\mathbf{m_0} = (3\ 0\ 0)$ and transition speeds $\boldsymbol{\lambda} = (0.9\ 1\ 1)$. If $\mathbf{m}[p_1] < \mathbf{m}[p_2]$, the flow of transition t_2 will be defined by the marking of p_1 (Σ_1).

Similarly, if $\mathbf{m}[p_1] > \mathbf{m}[p_2]$ the flow of t_2 will be defined by the marking of p_2 (Σ_2).

$$\Sigma_1 : \ \dot{\mathbf{m}} \ = \begin{pmatrix} -1.9 & 0 & 2 \\ -0.1 & 0 & 0 \\ 1.0 & 0 & -1 \end{pmatrix} \cdot \ \mathbf{m}$$

[1] Notice that places and transitions are transposed w.r.t. the incidence matrix.

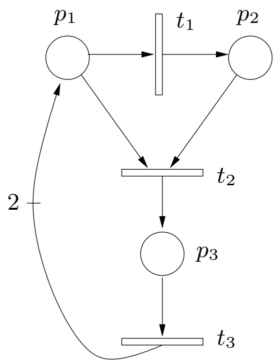

Fig. 1. A non JF net system with two PT-sets.

$$\Sigma_2 : \ \dot{\mathbf{m}} \ = \begin{pmatrix} -0.9 & -1 & 2 \\ 0.9 & -1 & 0 \\ 0.0 & 1 & -1 \end{pmatrix} \cdot \ \mathbf{m}$$

At the time instant in which $\mathbf{m}[p_1] = \mathbf{m}[p_2]$, Σ_1 and Σ_2 behave in the same way and any of them can be taken. Figure 2 shows the evolution of the system along time. At the beginning the system evolves according to Σ_2. Then a switch occurs and the dynamics of the system is described by Σ_1. A second switch turns the system back to Σ_2, the system stabilizes and no more switches take place.

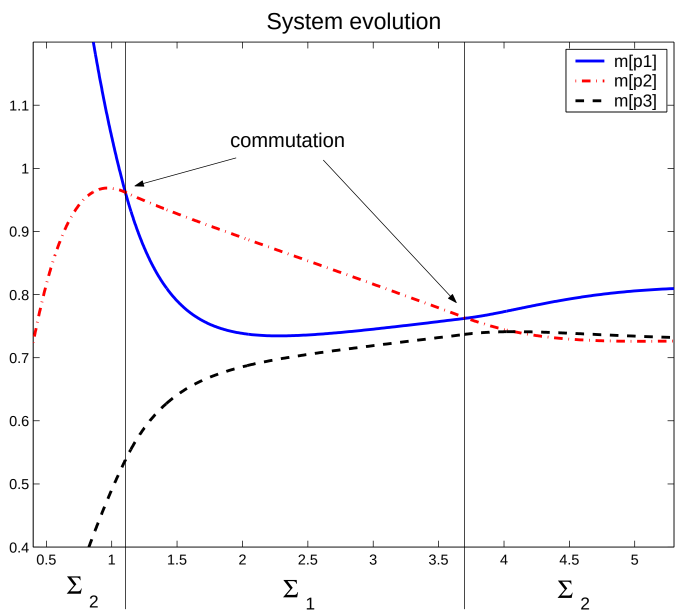

Fig. 2. Marking evolution of the system in Figure 1 with $\mathbf{m_0} = (3\ 0\ 0)$.

III. Observability: Problem Statement

Let us first consider linear time invariant systems, for which observability has been thoroughly studied [13], [17], [11], [12]. An *unforced* linear system (i.e., without inputs) is usually expressed by equations $\dot{\mathbf{x}} = \mathbf{A} \cdot \mathbf{x}, \mathbf{y} = \mathbf{S} \cdot \mathbf{x}$ where $\mathbf{x}$ is the state of the system and $\mathbf{y}$ is the output. Knowing the matrices $\mathbf{A}$ and $\mathbf{S}$ and being able to watch the evolution of $\mathbf{y}$, a linear system is said to be *observable* iff it is possible to compute its initial state, $\mathbf{x}(\tau_0)$. A well-known observability criterion exists that allows to decide whether a continuous (deterministic) time linear system is observable or not: The linear system is observable iff $\vartheta = [\mathbf{S}'\ \mathbf{A}'\mathbf{S}'\ \cdots\ (\mathbf{A}')^{n-1}\mathbf{S}']$ has full rank. Matrix ϑ is called the observability matrix of the system, and the initial state can be deduced from the following equation:

$$\begin{pmatrix} \mathbf{y}(0) \\ \dot{\mathbf{y}}(0) \\ \cdots \\ \mathbf{y}^{(n-1)}(0) \end{pmatrix} = \begin{pmatrix} \mathbf{S} \\ \mathbf{S} \cdot \mathbf{A} \\ \cdots \\ \mathbf{S} \cdot \mathbf{A}^{n-1} \end{pmatrix} \cdot \mathbf{x}_0 = \vartheta \cdot \mathbf{x}_0 \quad (1)$$

A drawback of this theoretical result is that it requires the use of higher order derivatives to compute the initial state. However, different approaches exist in practice to estimate the state of a continuous time linear system, even guaranteeing convergence in finite time [18]. The estimate for the marking of a PN will be denoted as $\tilde{\mathbf{m}}$.

Here it will be assumed that each place is either *measured* or *unmeasured*. It will be said that a place p_i is measured iff there exists a row j in $\mathbf{S}$ such that $\mathbf{S}(j,i) \neq 0$ and $\mathbf{S}(j,k) = 0$ for every $k \neq i$.

Let us define the concept of observability for a continuous Petri net system:

Definition 2: Let $\langle \mathcal{N}, \boldsymbol{\lambda}, \mathbf{m_0} \rangle$ be a continuous PN system and $\mathcal{D}$ the set of measured places.

- A place $p \in P$ is *observable* from $\mathcal{D}$ iff it is possible to compute its initial marking $\mathbf{m_0}[p] = \mathbf{m}(\tau_0)[p]$ by measuring the marking evolution of the places in $\mathcal{D}$.
- $\mathcal{N}$ is *observable* from $\mathcal{D}$ iff every place $p \in P$ is *observable*.

IV. Observability in Join Free Systems

In this section, an observability criterion that is independent of the internal speeds of the transitions, i.e., vector $\boldsymbol{\lambda}$, is obtained.

A. Structural Observability

Definition 3: Let $\mathcal{N}$ be a continuous PN and $\mathcal{D}$ the set of measured places of the system:

- Place p is *structurally observable* from $\mathcal{D}$ iff it is observable from $\mathcal{D}$ for any $\boldsymbol{\lambda} > 0$.
- $\mathcal{N}$ is *structurally observable* from $\mathcal{D}$ iff every place p is *structurally observable*.

In other words, structural observability seeks for observability for any $\boldsymbol{\lambda}$, like structural boundedness seeks for boundedness for any $\mathbf{m_0}$ [2]. Structural observability can be seen as a "robust" observability property because a structurally observable net remains observable for any variation that its vector $\boldsymbol{\lambda}$ may suffer. For instance, let us suppose that the only measured place of the system in Figure 3 is p_3 and that the vector $\boldsymbol{\lambda}$ is known. The variation, i.e., the derivative, of the marking of a place is given by the difference between its input and output flows. For p_3, we have $\dot{\mathbf{m}}[p_3] = \mathbf{f}_2 - \mathbf{f}_3$ where: $\mathbf{f}_2 = \boldsymbol{\lambda}[t_2] \cdot \mathbf{m}[p_2]$ and $\mathbf{f}_3 = \boldsymbol{\lambda}[t_3] \cdot \mathbf{m}[p_3]$, and so $\mathbf{m}[p_2] = (\dot{\mathbf{m}}[p_3] + \boldsymbol{\lambda}[t_3] \cdot \mathbf{m}[p_3])/\boldsymbol{\lambda}[t_2]$. Therefore, from the evolution of $\mathbf{m}[p_3]$, $\mathbf{m}[p_2]$ can be computed. Furthermore, it holds $\dot{\mathbf{m}}[p_2] = \mathbf{f}_1 - \mathbf{f}_2$ and $\mathbf{f}_1 = \boldsymbol{\lambda}[t_2] \cdot \mathbf{m}[p_1]$. Thus, $\mathbf{m}[p_2]$ being computable, $\mathbf{m}[p_1]$ can also be computed. This procedure can be carried out whatever the value of $\boldsymbol{\lambda}$ is, i.e. this net is structurally observable.

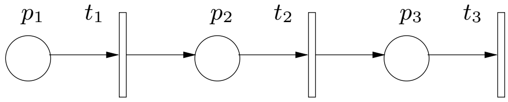

Fig. 3. The system marking can be computed from the observation of p_3.

This result can be generalized as follows:

Theorem 4: Let $\mathcal{N}$ be a continuous PN and $\mathcal{D}$ the set of measured places. If a place p is structurally observable then a forward path from p to $\mathcal{D}$ exists.

Proof: The marking of a place p can only have impact on downstream places and transitions. If there is no forward path from p to $\mathcal{D}$, then the marking of p cannot be observed from the variation of the marking of the places in $\mathcal{D}$. ■

B. Computation Algorithm

A similar approach to the one considered to observe p_1 and p_2 in the system in Figure 3 can be used to observe p_1 and p_2 in the system in Figure 4, where the measured places are p_3, p_4, and p_5. Let us consider a matrix $\mathbf{Post}^u \in \mathbb{R}^{|P|\times|T|}$ containing only the output arc weights of the transitions whose flow is, "in principle", unknown, i.e., the marking of their input places is not known. More formally, in the iterative algorithm that will be proposed, for any $p_i \in t_j{}^\bullet$, $\mathbf{Post}^u[i,j] = 0$ if the marking of the place ${}^\bullet t_j$ is measured or has been computed in previous iterations, and $\mathbf{Post}^u[i,j] = \mathbf{Post}[i,j]$ otherwise.

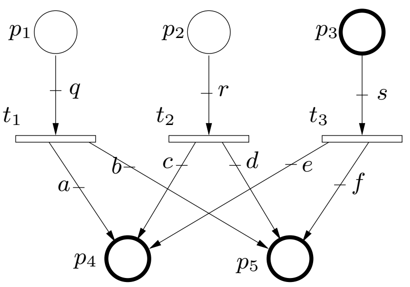

Fig. 4. A JF system whose marking is computable from the evolution of p_3, p_4 and p_5.

The first three rows (that correspond to places p_1, p_2 and p_3) of $\mathbf{Post}^u$ are zeros (because they were zeros in $\mathbf{Post}$) and the forth and fifth rows (that correspond to places p_4 and p_5) are $(a\ c\ 0)$ and $(b\ d\ 0)$, respectively. The marking evolution of places p_4 and p_5 is known (because they are measured) and here it is equal to their input flow. Subtracting the flow coming from p_3, $\mathbf{f}_{i4}^{p_3}$ and $\mathbf{f}_{i5}^{p_3}$, we will obtain the flow coming from the unknown places p_1 and p_2:

$$\begin{pmatrix} \dot{\mathbf{m}}[p_4] - \mathbf{f}_{i4}^{p_3} \\ \dot{\mathbf{m}}[p_5] - \mathbf{f}_{i5}^{p_3} \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\lambda}[t_1] \cdot \dfrac{\mathbf{m}[p_1]}{q} \\ \boldsymbol{\lambda}[t_2] \cdot \dfrac{\mathbf{m}[p_2]}{r} \end{pmatrix} \quad (2)$$

Hence, if the matrix $(a\ c; b\ d)$ has full rank it will be possible to compute the markings of p_1 and p_2 independently of the $\boldsymbol{\lambda}$ of the transitions.

The procedure developed for the above examples can be generalized leading to a fix point algorithm.

Given a set of places $\mathcal{H}$, $\mathbf{Post}^u_{\mathcal{H}}$ denotes a matrix composed by the rows of $\mathbf{Post}^u$ corresponding to the places in $\mathcal{H}$, and whose null columns have been removed.

Algorithm 5

Input ($\mathcal{N}$, $\mathcal{D}$)
Output $\mathcal{Q}$ % places that can be observed $\forall\ \boldsymbol{\lambda} > 0$
  $\mathcal{Q} := \mathcal{D}$
  Compute $\mathbf{Post}^u$
  While $\exists\ \mathcal{H} \subseteq \mathcal{Q}$, such that ${}^\bullet({}^\bullet\mathcal{H}) \not\subseteq \mathcal{Q}$ and
      $\mathbf{Post}^u_{\mathcal{H}}$ has full rank **do**
    $\mathcal{Q} := \mathcal{Q} \cup {}^\bullet({}^\bullet H)$
    Compute $\mathbf{Post}^u$ according to $\mathcal{Q}$
  End_While

Notice that the set $\mathcal{Q}$ increases at each iteration. Since the number of places in the net is finite, so is the final set $\mathcal{Q}$. Hence, the algorithm cannot execute indefinitely.

Proposition 6: Let $\mathcal{N}$ be a JF net, $\mathcal{D}$ the set of measured places and $\mathcal{Q}$ the output of the Algorithm 5 applied on ($\mathcal{N}$, $\mathcal{D}$). Every $p \in \mathcal{Q}$ is structurally observable.

Proof: Since they are measured, all places in $\mathcal{D}$ are trivially structurally observable. At each iteration of algorithm 5 the set $\mathcal{Q}$ increases. The marking of the newly added places to $\mathcal{Q}$ can be computed (Equation 2) from the marking of places that already were in $\mathcal{Q}$. Thus, the marking of every $p \in \mathcal{Q}$ can be computed whatever the value of $\boldsymbol{\lambda}$ is. ■

V. Observability in General Net Systems

A. Structural Observability in non-JF Systems

The PN system in Figure 5 represents a manufacturing system that produces tables (see [19], [16] for details). Transition t_5 has two input places, p_5 and p_6, thus, it is not a JF system. However, if places p_5 and p_6 are measured, i.e., $p_5 \in \mathcal{D}$ and $p_6 \in \mathcal{D}$, Algorithm 5 can be executed in the same way as for JF systems. Let us assume that $\mathcal{D} = \{p_5, p_6\}$. In the first iteration of the algorithm, p_3 is added to the set of observable places. Places p_7 and p_4 are also added to the set of observable places in the second and third iteration, respectively. At that point the algorithm stops. However, if the set of measured places is either $\mathcal{D} = \{p_1, p_5, p_6\}$ or $\mathcal{D} = \{p_2, p_5, p_6\}$ the algorithm will include all places in the set of observable places in four iterations.

B. Computing Estimates

The greatest advantage of the structural observability approach is that it is independent of $\boldsymbol{\lambda}$. The main drawback is that every place in a synchronization should be measured in order to apply the method. If $\boldsymbol{\lambda}$ is known, another way to face the observability problem in general systems consists in computing an estimate for every structural PT-set of the net. The computed estimates can be used to filter out those PT-sets that for sure are not ruling the evolution of the system.

Definition 7:

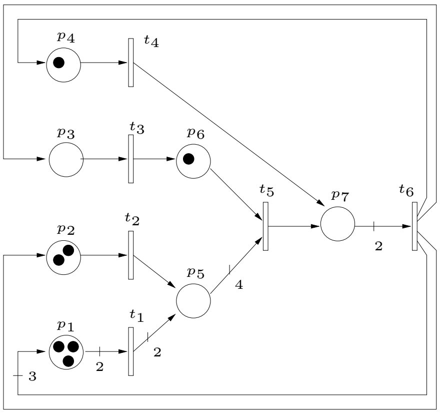

Fig. 5. PN modelling a table factory.

- An estimate $\tilde{\mathbf{m}}$ for a PT-set W is *infeasible* if Equation 1 has no solution when applied for W.
- An estimate $\tilde{\mathbf{m}}$ for a PT-set W is *incoherent* if $W \not\subseteq \text{PT-set}(\tilde{\mathbf{m}})$, i.e, W is not a PT-set of $\tilde{\mathbf{m}}$.

If an estimate is either infeasible or incoherent, then its associated PT-set must be filtered out. Let us consider the subnet composed of p_1, p_2 and t_2 of the system in Figure 1. The net has two structural PT-sets, $W_1 = \{(p_1, t_2)\}$ and $W_2 = \{(p_2, t_2)\}$ (the PT-set $W_3 = \{(p_1, t_2), (p_2, t_2)\}$ is not considered to avoid redundancy, i.e, when $\mathbf{m}(p_1) = \mathbf{m}(p_2)$ all three PT-sets are equivalent). The evolution of the system according to PT-set W_1 is ruled by the matrix $\mathbf{A}_1 = (-1 \;\; 0; -1 \;\; 0)$. The system matrix for PT-set W_2 is $\mathbf{A}_2 = (0 \;\; -1; 0 \;\; -1)$. Considering that the initial marking is $\mathbf{m_0} = (4; 2)$, the initial PT-set for the system is W_2, and so $\mathbf{m}(\tau) = (2 \cdot e^{-\tau} + 2; 2 \cdot e^{-\tau})$.

As external agents of the system we will consider 4 cases depending on: 1) the output of the system, i.e., the measured place, which can be either p_1 or p_2, and 2) the PT-set that is assumed to be ruling the net system, which can be either W_1 or W_2.

In the first two cases $\mathbf{m}[p_1]$ is the output of the system, i.e., $\mathbf{y} = 2 \cdot e^{-\tau} + 2$ (hence, $\mathbf{m}[p_1]$ can be estimated correctly):

Case 1: W_2 is assumed to be the PT-set. For this case $\vartheta = (1 \; 0; 0 \; -1)$ whose rank is 2. The initial marking $\mathbf{m_0} = (4; 2)$ can be recovered.

Case 2: W_1 is assumed to be the PT-set. The observability matrix is $\vartheta = (1 \; 0; -1 \; 0)$. Equation 1 has no solution, thus, an estimate for $\mathbf{m}[p_2]$ cannot be computed.

In this way, by means of an "infeasible" estimate (case 2), it has been detected that the PT-set of the system is W_2.

If p_2 is measured, i.e., $\mathbf{y} = 2 \cdot e^{-\tau}$ (hence, $\mathbf{m}[p_2]$ can be estimated correctly), the two cases are:

Case 3: W_2 is assumed to be the PT-set. $\vartheta = (0 \; 1; 0 \; -1)$, which is not a full rank matrix. Thus, it is only possible to observe the marking of p_2.

Case 4: W_1 is assumed to be the PT-set. The observability matrix is $\vartheta = (0 \; 1; -1 \; 0)$ which has full rank. The solution for Equation 1 is $\tilde{\mathbf{m}}_0 = (2 \; 2)$. Given that $\tilde{\mathbf{m}}_0$ is coherent with W_1, it is a suitable estimate, even if, in this case, p_1 is not estimated correctly.

Finally, let us show when an incoherent estimate may appear for the system in Figure 1. Let us assume that W_1 rules the system evolution. Let us further assume that an estimate $\tilde{\mathbf{m}}_0$ is obtained such that $\tilde{\mathbf{m}}_0[p_1] > \tilde{\mathbf{m}}_0[p_2]$. The PT-set of such estimate is W_2. Hence, $\tilde{\mathbf{m}}_0$ is incoherent with the assumed PT-set W_1, and therefore it cannot be a suitable estimate.

The steps required to obtain suitable estimates are summed up as:

- Compute an estimate per structural PT-set.
- Filter out non-suitable estimates: infeasible or incoherent.
- The remaining estimates represent every potential marking of the net.

VI. Observers and Estimates

The main drawback of the method presented in the previous section is that it is very sensitive to the noise that may appear in the output $\mathbf{y}$. In order to overcome that problem, observers are introduced. For linear systems, Luenberger's observers [13], [11] are widely used. A Luenberger observer for a PN with a single PT-set can be expressed as: $\dot{\tilde{\mathbf{m}}} = \mathbf{A} \cdot \tilde{\mathbf{m}} + \mathbf{K} \cdot (\mathbf{y} - \mathbf{S} \cdot \tilde{\mathbf{m}})$ where $\tilde{\mathbf{m}}$ is the marking estimate, $\mathbf{A}$ and $\mathbf{S}$ (see Section III) are the matrices defining the evolution of the marking of the system and its output in continuous time, $\mathbf{y}$ is the output of the system, and $\mathbf{K}$ is a design matrix of parameters.

The "goodness" of an estimate can be measured by means of a *residual* [15]. Let us use the 1-norm $|| \cdot ||_1$, which is defined as $||\mathbf{x}||_1 = |\mathbf{x}_1| + \ldots + |\mathbf{x}_n|$. The residual at a given instant, $r(\tau)$, is the distance between the output of the system and the output that the observer's estimate, $\tilde{\mathbf{m}}(\tau)$, yields, i.e., $r = ||\mathbf{S} \cdot \tilde{\mathbf{m}}(\tau) - \mathbf{y}(\tau)||_1$.

A. *Filtering estimates*

One (Luenberger) linear observer [11] will be designed per PT-set of the PN. The designed observers will be launched simultaneously. In a similar way to the previous section, observer estimates are expected to verify the following confitions:

- The residual must tend to zero.
- The estimates of the places in a synchronization have to be *coherent* with the PT-set for which they are computed.

Thus, at a given time instant, only coherent estimates are suitable. Moreover, a criterion must be established to decide which coherent estimate is, at a given time instant, the most appropriate. An adequate *heuristics* is to choose the coherent estimate with minimum residual.

B. *Design of a switching observer*

Consider again the continuous PN system in Figure 1. Let its output be the marking of place p_1, i.e., $\mathbf{S} = (1 \; 0 \; 0)$. The net has two PT-sets: let one of the PT-sets be $Z_1 = \{(p_1, t_1), (p_1, t_2), (p_3, t_3)\}$ and the other be $Z_2 = \{(p_1, t_1), (p_2, t_2), (p_3, t_3)\}$. For the PT-set Z_1 the marking of p_2 is unobservable. However, for the PT-set Z_2 the marking of all the places can be estimated. Let $\boldsymbol{\lambda} = (0.9 \; 1 \; 1)$

and $\mathbf{m_0} = (3\ 0\ 0)$. The marking evolution of this system is depicted in Figure 2.

One observer per PT-set will be designed: observer i for PT-set Z_i, $i = 1, 2$. Let the initial state of observer 1 be $\mathbf{e}_{01} = (1\ 2)$ and its eigenvalues be $(-12 + 2 \cdot \sqrt{3} \cdot i,\ -12 - 2 \cdot \sqrt{3} \cdot i)$. Since observer 1 can only estimate p_1 and p_3, the first component of its state vector corresponds to the estimate for $\mathbf{m}[p_1]$, and its second component to the estimate for $\mathbf{m}[p_3]$. For observer 2, let the initial state be $\mathbf{e}_{02} = (1\ 0\ 2)$ and its eigenvalues be $(-15,\ -12 + 2 \cdot \sqrt{3} \cdot i,\ -12 - 2 \cdot \sqrt{3} \cdot i)$. The evolution of the coherent estimate with minimum residual is shown in Figure 6.

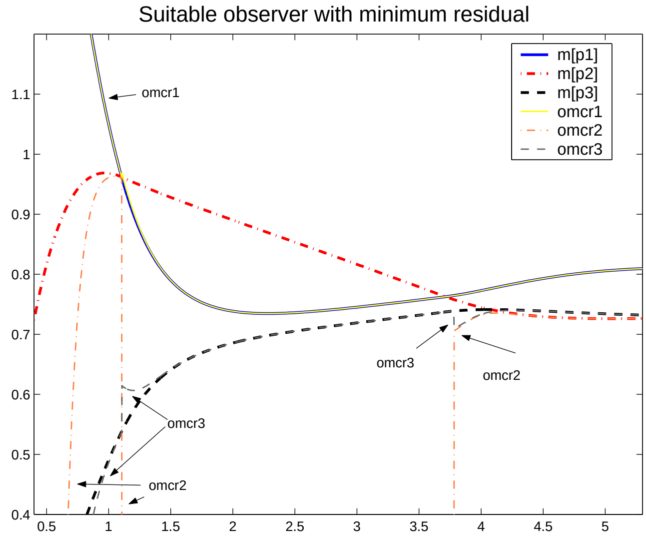

Fig. 6. Minimum residual and coherent observer, $(omcr1,\ omcr2,\ omcr3)$ is the estimate for $(\mathbf{m}[p_1],\ \mathbf{m}[p_2],\ \mathbf{m}[p_3])$.

The resulting estimate can be improved by taking into account some considerations. When the first system switch happens, the estimate becomes discontinuous and, what is more undesirable, the estimate for the marking of p_3 becomes worse. A similar effect happens when the second system switch occurs. Another undesirable phenomenon is that, after the first switch, the estimate of the marking of p_2 just disappears (since $\mathbf{m}[p_2]$ is unobservable in PT-set Z_1).

One way to avoid discontinuities in the resulting estimate, is to use the estimate of the observer that is going to be filtered out to update the estimate of the observer that is not going to be filtered out. This estimate update must be done when a system switch is detected. In order not to lose the estimate of the marking of a place when it was "almost perfectly" estimated (recall the case of p_2 when the first switch happened) a simulation of the system can be launched. The initial marking of this simulation is the estimate of the system just before the observability of the place is lost. Such a simulation can be seen as an estimate for those places that are not observable by the observer being considered. The simulation should only be carried out when an estimate for all the places exists and the residual is not significant. Figure 7 shows the evolution of the estimate obtained by this strategy.

The resulting observer can be seen as a set of switching

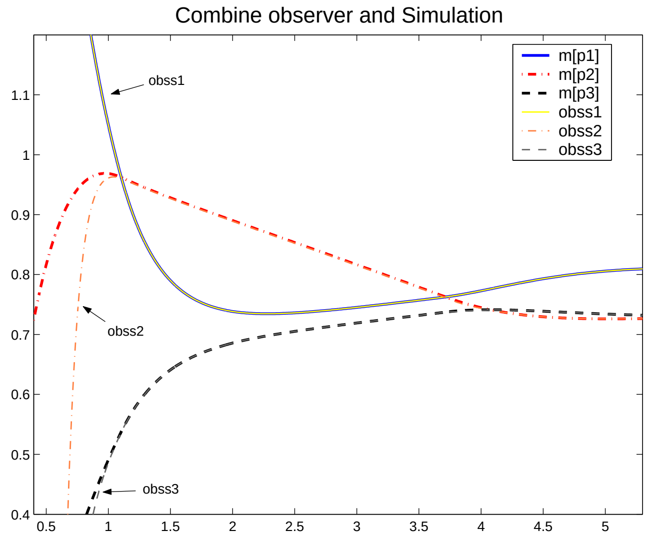

Fig. 7. Resulting observer's estimate that makes use of a simulation, $(obss1,\ obss2,\ obss3)$ is the estimate for $(\mathbf{m}[p_1],\ \mathbf{m}[p_2],\ \mathbf{m}[p_3])$.

linear observers. One of the main advantages is that the residual does not increase sharply when the PT-set of the system changes. Another interesting feature is that the use of a simulation allows to estimate the marking of places that in some PT-sets are in principle not observable: in Figure 7 it can be seen that the marking of p_2 can be estimated, even when it is unobservable due to PT-set Z_1 being active.

VII. Conclusions

The performance model of continuous PNs working under infinite server semantics has been considered. Structural observability has been introduced and studied for continuous Petri net systems without synchronizations (JF systems).

For general (with synchronizations) continuous PN an estimate can be computed for each structural PT-set (linear system). Although this may lead to a large number of estimates, the ones classified as *infeasible* or *incoherent* are non-suitable and can be filtered out.

In order to design an observer for a timed continuous PN, one linear (Luenberger) observer per PT-set has been considered. The estimate yielded by a given observer is not suitable if it produces a *non null residual* or it is *non-coherent*. Based on the idea of choosing the suitable estimate with the smallest residual, a switching observer has been proposed.

References

[1] T. Murata, "Petri nets: Properties, analysis and applications," *Proceedings of the IEEE*, vol. 77, no. 4, pp. 541–580, 1989.

[2] M. Silva, "Introducing Petri nets," in *Practice of Petri Nets in Manufacturing*. Chapman & Hall, 1993, pp. 1–62.

[3] J. Campos and M. Silva, "Structural techniques and performance bounds of stochastic Petri net models," in *Advances in Petri Nets 1992*, ser. Lecture Notes in Computer Science, G. Rozenberg, Ed. Springer, 1992, vol. 609, pp. 352–391.

[4] M. Ajmone Marsan, G. Balbo, G. Conte, S. Donatelli, and G. Franceschinis, *Modelling with Generalized Stochastic Petri Nets*. Wiley, 1995.

[5] C. Özveren and A. Willsky, "Observability of discrete event dynamic systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 7, pp. 797–806, 1990.

[6] M. Silva and L. Recalde, "Petri nets and integrality relaxations: A view of continuous Petri net models," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 32, no. 4, pp. 314–327, 2002.

[7] ——, "On fluidification of Petri net models: from discrete to hybrid and continuous models," *Annual Reviews in Control*, vol. 28, pp. 253–266, 2004.

[8] R. David and H. Alla, *Discrete, Continuous, and Hybrid Petri Nets*. Springer-Verlag, 2004.

[9] J. Júlvez, E. Jiménez, L. Recalde, and M. Silva, "On observability in timed continuous Petri net systems," in *First International Conference on the Quantitative Evaluation of Systems (QEST). IEEE Computer Society.*, Enschede, The Netherlands., 2004.

[10] ——, "Design of observers for timed continuous Petri net systems," in *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, The Hague, The Netherlands, 2004.

[11] K. Ogata, *Modern Control Engineering, 4th ed.* Prentice Hall, 2001.

[12] B. Friedland, *Control system design: An Introduction to State-Space Methods*. McGraw-Hill, 1987.

[13] D. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, December 1971.

[14] A. Bemporad, G. Ferrari-Trecate, and M. Morari, "Observability and controllability of piecewise affine and hybrid systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 10, pp. 1864–1876, 2000.

[15] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, and A. L. Sangiovanni-Vincentelli, "Design of observers for hybrid systems," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, C. J. Tomlin and M. R. Greenstreet, Eds. Berlin Heidelberg New York: Springer-Verlag, 2002, vol. 2289, pp. 76–89.

[16] L. Recalde and M. Silva, "Petri Nets fluidification revisited: Semantics and steady state," *APII-JESA*, vol. 35, no. 4, pp. 435–449, 2001.

[17] T. Kailath, *Linear Systems*. Prentice-Hall, 1980.

[18] R. Engel and G. Kreisselmeier, "A continuous-time observer which converges in finite time," *IEEE Transactions on Automatic Control*, vol. 47, no. 7, pp. 1202–1204, 2002.

[19] E. Teruel, J. M. Colom, and M. Silva, "Choice-free Petri nets: A model for deterministic concurrent systems with bulk services and arrivals," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 27, no. 1, pp. 73–83, 1997.