Observability of continuous Petri nets with infinite server semantics

Cristian Mahulea, Laura Recalde, and Manuel Silva *

October 9, 2009

Abstract

Timed continuous Petri net (contPN) systems with infinite server semantics are non-linear systems, particularly a subclass of piecewise linear (PWL) systems. This paper addresses several problems regarding the state observability of these systems. We assume that the initial marking/state is not known and measuring the marking of some places we want to estimate all the others. First, a study of the different linear systems corresponding to a continuous Petri net system is performed. It is shown that in some cases, some of them are redundant, and so can be disregarded. The notion of *distinguishable modes* is introduced helping to give a necessary and sufficient criterion for the observability in infinitesimal time. Structural observability, i.e., observability for all possible values of firing rates of transitions, is studied and it is proved that in some cases it can be reduced to a linear problem, even if the system is nonlinear. Using results from *linear structured systems*, the concept of *weak structural* or *generic observability* is considered.

To be published as: C. Mahulea, L. Recalde, and M. Silva, "Observability of continuous Petri nets with infinite server semantics", *Nonlinear Analysis: Hybrid Systems*, 2009, doi:10.1016/j.nahs.2009.06.014.
Corresponding author **C. Mahulea**. Tel. +34-976-762517. Fax +34-976-761914.
This work was partially supported by the CICYT - FEDER project DPI2006-15390. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 224498.

*The authors are with are with the Department of Computer Science and System Engineering, University of Zaragoza, Maria de Luna 1, 50018 Zaragoza, Spain {cmahulea, silva@unizar.es}.

1 Introduction

State estimation is a very important part of the control problem when not all the states are directly measurable. In last years, state estimation has been studied by many researchers [1, 4] but many aspects are still open, specially in the case of hybrid systems. In this paper we study *observability in infinitesimal time*, *structural observability* and *generic observability* of timed continuous Petri nets systems with infinite server semantics. Under this semantic, a contPN system with joins corresponds to a set of switching linear systems, where the switches depend on the marking and the rates of the transitions. A brief introduction to timed contPN is given in Section 2 recalling basic concepts and results used along the paper.

While observability is well understood in classical linear system theory [10, 15], it becomes more complex in the case of hybrid systems. In this case, observability has been studied in the literature in the last years considering different classes of systems: piecewise linear, piecewise affine and switched hybrid systems [1, 2, 3, 4, 19]. Being contPN system a PWL system, the results regarding its observability are similar to those obtained in hybrid systems, but they differ for two main reasons: (1) the existence of linear systems that are redundant, i.e., systems that it is not necessary to consider; (2) when the marking is at the border of two regions, more than one linear system can be used indistinctively (thus it is not important which one is taken), which makes harder to distinguish between them.

In Section 3 the concept of *redundant mode* [14] is introduced. An exponential number of linear systems can be embedded in a contPN but not all of them are always fundamental for the evolution. A sufficient and necessary condition for a mode to be redundant is presented. After, in Section 4 the notion of *distinguishable* modes is considered, a concept similar to the one of hybrid systems: distinguishable discrete states. A condition for two modes to be distinguishable is given and, finally, a necessary and sufficient criterium for the observability in infinitesimal time of general contPN systems is proved.

Structural observability is a more general concept of observability, ensuring observability for any (positive) value of firing rates, not just for a particular value. This concept is studied in Section 5. For different subclasses of nets it is shown how can be found the set of places with minimum cardinality that ensures structural observability. This is a different problem from the one in [8] where the idea is to check if a net is structurally observable or not for a given set of measured places. Even if we deal with PWL systems, for some subclases of nets it is proved that the observability can be checked on a reduced join-free net that corresponds to a linear system.

In [12] an interpretation of the loss of observability in the case of join free (JF) contPN systems is given. When the system contains attributions, the observability cannot be checked locally and can be lost for some specific values of the firing rates of the transitions. Moreover, using results on (linear) structured systems [5], the problem of generic observability (here observability for almost all firing rates $\boldsymbol{\lambda}$) is solved in Section 6. In this case, the firing rates of the transitions become parameters and the system is called generically observable if it is observable for almost all values of its firing rates. Hence, the attributions will not create problems for generic observability.

2 Timed Continuous Petri Net Systems

Definition 1 *A contPN system is a pair $\langle \mathcal{N}, \boldsymbol{m_0} \rangle$, where $\mathcal{N} = \langle P, T, \boldsymbol{Pre}, \boldsymbol{Post} \rangle$ is a net structure and $\boldsymbol{m}_0 \in \mathbb{R}_{\geq 0}^{|P|}$ is the initial marking. P is the set of places, T is the set of transitions and $\boldsymbol{Pre}, \boldsymbol{Post} \in \mathbb{R}_{\geq 0}^{|P| \times |T|}$ are the pre and post incidence matrices, respectively.*

Let p_i, $i = 1, \ldots, |P|$ and t_j, $j = 1, \ldots, |T|$ denote the *places* and the *transitions*. For a place $p_i \in P$ and a transition $t_j \in T$, Pre_{ij} and $Post_{ij}$ represent the weight of the arc from p_i to t_j and from t_j to p_i, respectively. Each place $p_i \in P$ has a token load denoted $m_i \in \mathbb{R}_{\geq 0}$. The vector of all token loads is called marking (distributed state) and it is denoted by $\boldsymbol{m} \in \mathbb{R}_{\geq 0}^{|P|}$. The preset and postset of a node $v \in P \cup T$ are denoted by ${}^\bullet v$ and $v^\bullet$, and represents the input and the output nodes of v, respectively.

Example 2 *Let us consider the contPN system in Fig. 1(b). For this net, $P = \{p_1, p_2, p_3\}$, $T = \{t_1, t_2, t_3, t_4\}$,*

$$\boldsymbol{Pre} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{Post} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

For example, $Pre_{13} = 1$ means the existence of an arc from p_1 to t_3 of weight 1, while $Post_{12} = 1$ means the existence of an arc from t_2 to p_1 of weight 1.

A transition t_j is *enabled* at $\boldsymbol{m}$ iff $\forall p_i \in {}^\bullet t_j$, $m_i \geq 0$. Its *enabling degree* is:

$$enab(t_j, \boldsymbol{m}) = \min_{p_i \in {}^\bullet t_j} \left\{ \frac{m_i}{Pre_{ij}} \right\} \qquad (1)$$

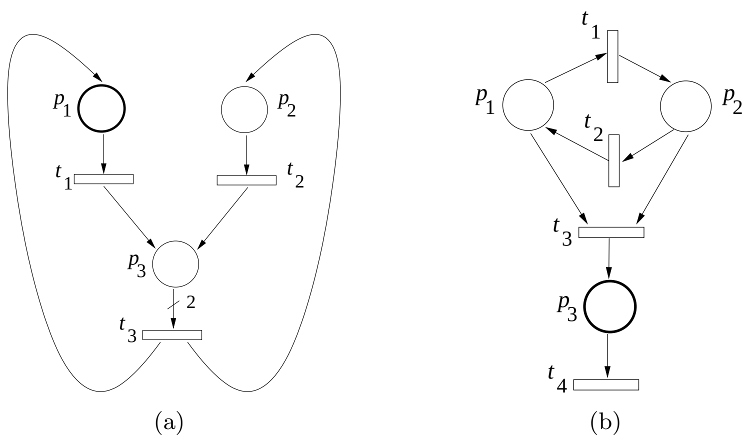

Figure 1: Two timed contPNs

which represent the maximum amount in which t_j can fire at $\boldsymbol{m}$. An enabled transition t_j can fire in any real amount $0 \leq \alpha \leq enab(t_j, \boldsymbol{m})$ leading to $\boldsymbol{m}' = \boldsymbol{m} + \alpha \cdot \boldsymbol{C}_{\cdot j}$ where $\boldsymbol{C} = \boldsymbol{Post} - \boldsymbol{Pre}$ is the *token flow matrix* and $\boldsymbol{C}_{\cdot j}$ denotes its j^{th} column. If $\boldsymbol{m}$ is reachable from $\boldsymbol{m}_0$ through a sequence $\sigma = \alpha_1 t_1 \ldots \alpha_k t_k$, a *state* (or *fundamental*) *equation* can be written: $\boldsymbol{m} = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{\sigma}$, where $\boldsymbol{\sigma} \in \mathbb{R}^{|T|}_{\geq 0}$ is the firing count vector, i.e., σ_j is the cumulative amount of firing of t_j in the sequence σ. The set of all reachable marking from $\boldsymbol{m}_0$ is called *reachability space* and it is denoted by $\mathcal{R}(\mathcal{N}, \boldsymbol{m}_0)$.

When a time interpretation is introduced, the fundamental equation depends on time: $\boldsymbol{m}(\tau) = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{\sigma}(\tau)$. After time differentiation becomes: $\dot{\boldsymbol{m}}(\tau) = \boldsymbol{C} \cdot \dot{\boldsymbol{\sigma}}(\tau)$. The derivative of the firing count vector $\boldsymbol{f}(\tau) = \dot{\boldsymbol{\sigma}}(\tau)$ is the (firing) *flow*. Therefore, the dynamical equation of a contPN results:

$$\dot{\boldsymbol{m}}(\tau) = \boldsymbol{C} \cdot \boldsymbol{f}(\tau) \tag{2}$$

In this paper we consider contPN with *infinite server semantics* that for a broad class of net systems provides a better approximation of the steady state throughput [13] of the discrete net. With this semantics, the flow of a transition t_j is given by:

$$f_j(\tau) = \lambda_j \cdot enab(t_j, \boldsymbol{m}(\tau)) \tag{3}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{|T|}_{>0}$ associate a constant $\lambda_j > 0$ to each transition t_j representing its *firing rate*. Substituting (1) into (3) and the result in (2), it can be easily observed that a timed contPN system with infinite server semantics is a *piecewise linear system* with *polyhedral regions*.

There the dynamical matrix in each $\mathcal{R}_i$ is calculated using a characteristic matrix called $\mathbf{\Pi}^i$ for each mode i (see [11] for more details)

$$\Pi^i_{jh} = \begin{cases} \frac{1}{Pre_{hj}}, \text{ if } \forall \boldsymbol{m} \in \mathcal{R}_i, \frac{m_h}{Pre_{hj}} = enab(t_j, \boldsymbol{m}) \\ 0, \text{ otherwise} \end{cases}$$

Denoting by $\mathbf{\Lambda}$ the matrix having on its diagonal the elements of $\boldsymbol{\lambda}$, the dynamical matrix in $\mathcal{R}_i$ is given by: $\boldsymbol{A}_i = \boldsymbol{C} \cdot \mathbf{\Lambda} \cdot \mathbf{\Pi}^i$. Hence, the dynamics of the markings are given by:

$$\dot{\boldsymbol{m}}(\tau) = \boldsymbol{A}_i \cdot \boldsymbol{m}(\tau),\ \boldsymbol{m} \in \mathcal{R}_i,\ i \in I, \tag{4}$$

where $\boldsymbol{A}_i \in \mathbb{R}^{|P|\times|P|}$, $\mathcal{R}_i$ is a polyhedral set, and I is a set of labels for the *modes* (linear systems) of the piecewise linear system. (See [11] for more details. There the dynamical matrix is calculated using a characteristic matrix called $\mathbf{\Pi}_i$ for each mode i).

Example 3 *Let us consider the contPN system in Fig. 1(b) (see also Ex. 2). The flows of the transitions are given by:* $f_1 = \lambda_1 \cdot m_1$*,* $f_2 = \lambda_2 \cdot m_2$*,*

$$f_3 = \begin{cases} \lambda_3 \cdot m_1 & \textit{if } m_1 \le m_2 \\ \lambda_3 \cdot m_2 & \textit{if } m_2 \le m_1 \end{cases}$$

and $f_4 = \lambda_4 \cdot m_3$*. Let us assume* $\boldsymbol{\lambda} = \mathbf{1}$*. Using* (1)*,* (3) *and* (4)*, the modes of the systems are: (i) the enabling degree of* t_3 *given by* m_1*:* $\mathcal{R}_1 = \{m_1 \le m_2\}$ *with*

$$\boldsymbol{A}_1 = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

and (ii) the enabling degree of t_3 *given by* m_2*:* $\mathcal{R}_2 = \{m_2 \le m_1\}$ *with*

$$\boldsymbol{A}_2 = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

3 Redundant linear systems

The number of regions and modes can be exponential, upper bounded by $\prod_{t_j \in T} |{}^\bullet t_j|$ where $|{}^\bullet t_j|$ is the cardinality of the set ${}^\bullet t_j$. An evident necessary condition for the observability of the PWL systems in infinitesimal time is the observability of all linear systems [4]. So, if the number of linear systems composing the contPN under study is reduced, the complexity analysis of the observability decrease.

In the context of contPN, for a given initial marking $\boldsymbol{m}_0$, some places can be implicit [17] and their marking will never be the unique to give the enabling degree of a transition in (1). In other words, $p_k \in P$ is implicit if for any reachable marking from $\boldsymbol{m}_0$, $\frac{m_i}{Pre_{ij}} \leq \frac{m_k}{Pre_{kj}}$ with $p_i \in {}^\bullet t_j \setminus \{p_k\}$ is satisfied $\forall t_j \in p_k$. This implies that the corresponding regions of the contPN system defined by $\frac{m_i}{Pre_{ij}} \leq \frac{m_k}{Pre_{kj}}$ are either empty or reduced to their borders. On the other hand, for a given initial marking and a given firing rate vector $\boldsymbol{\lambda}$, in [16] it is introduced the notion of *time implicit arc* as those arcs from a place p_i to a transition t_j such that $\frac{m_i}{Pre_{ij}}$ is never the unique quantity that gives the minimum in (1). As in the previous case, the corresponding regions are either empty or reduced to their borders.

Unfortunately, the previous results, based on the net structure, timing and initial marking, cannot be used in the context of state estimation since the initial marking is not known for us and can take any real positive value (our approach will be to identify this marking). In this section we study a stronger concept, only depending on the net structure, valid for all possible initial markings. It may happen that for every initial marking all reachable markings belonging to a region are on the border (an example is illustrated later in Ex. 5). Hence, these markings belong always to other regions, so it is not necessary to consider this mode, obviously neither to check the observability of the corresponding linear system. Therefore, our first step is to structurally characterize these *redundant modes* and to remove them.

Definition 4 *Let $\mathcal{R}_i$, $i \in I$ be a region. If for all $\boldsymbol{m}_0$, $\mathcal{R}_i \subseteq \bigcup_{j \neq i} \mathcal{R}_j$ then $\mathcal{R}_i$ is a* redundant region *and i is a* redundant mode.

Example 5 *Let us consider the subnet in Fig. 2(a). Assume some reachable markings such that the enabling degree of t_1 is given by m_1, i.e., $m_1 \leq m_2$, and the enabling degree of t_2 is given by m_2, i.e., $m_2 \leq m_1$. These markings belong to the following region $\mathcal{R}_1 = \{m_1 \leq m_2, m_2 \leq m_1, \ldots\}$. Taking other markings for which the enabling degree of t_1 and t_2 is given by the same m_1, the corresponding region is $\mathcal{R}_2 = \{m_1 \leq m_2, \ldots\}$. Assume also that the enabling degrees of the other transitions not represented in the figure are given by the marking of the same places, i.e., the other inequalities defining $\mathcal{R}_1$ and $\mathcal{R}_2$ are the same.*

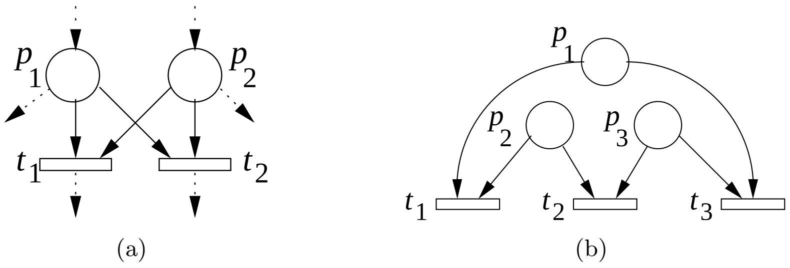

Figure 2: ContPNs with redundant regions.

From the above definition of $\mathcal{R}_1$ and $\mathcal{R}_2$ and from the assumption on the other inequalities, it is obvious that $\mathcal{R}_1 \subseteq \mathcal{R}_2$ for all $\boldsymbol{m}_0 \in \mathbb{R}^{|P|}_{\geq 0}$. In fact, $\mathcal{R}_1$ is a frontier of $\mathcal{R}_2$. Since the linear system in $\mathcal{R}_2$ provides the same time-evolution for the markings $\boldsymbol{m} \in \mathcal{R}_1$, the linear system in $\mathcal{R}_1$ can be ignored in the analysis of the contPN system.

To see if a mode $i \in I$ is non-redundant, check if there exists a marking in the corresponding region $\mathcal{R}_i$ such that the inequalities composing its definition are strictly satisfied. In other words, if for a join t_j with $p_i, p_k \in {}^\bullet t_j$ does not exists $\boldsymbol{m} \in \mathbb{R}^{|P|}_{\geq 0}$ such that $\frac{m_i}{Pre_{ij}} < \frac{m_k}{Pre_{kj}}$ then the linear systems of the regions of the form $\mathcal{R}_f = \left\{ \frac{m_i}{Pre_{ij}} \leq \frac{m_k}{Pre_{kj}}, \dots \right\}$ are redundant.

Proposition 6 *Let $\mathcal{N}$ be a timed contPN system. The region $\mathcal{R}_i$ with the corresponding mode $i \in I$ is redundant iff $\nexists \boldsymbol{m} \in \mathbb{R}^{|P|}_{\geq 0}$ solution of the following system of strictly inequalities of the form $\frac{m_k}{Pre_{kj}} < \frac{m_u}{Pre_{uj}}$, one for each $\frac{m_k}{Pre_{kj}} \leq \frac{m_u}{Pre_{uj}}$ defining $\mathcal{R}_i$.*

Proof: Obviously, if the system has a solution this is an interior point of $\mathcal{R}_i$ and the corresponding linear system cannot be removed.

For the reverse sense, let us assume that the system of strictly inequalities has no solution. This means that for all $\boldsymbol{m}_0 \geq 0$ there exists at least one join transition t_j such that $\frac{m_k}{Pre_{kj}} \geq \frac{m_u}{Pre_{uj}}$. If for all $\boldsymbol{m}$ this inequality is satisfied strictly the region is empty and can be eliminated without problems together with the corresponding linear system. Otherwise, if it is an equality, considering that the flow of t_j is given by m_u (not by m_k) it is clear that the corresponding regions include $\mathcal{R}_i$. Hence, i is a redundant mode. ■

The solution of the system of inequalities in Prop. 6 can be checked solving a linear programming problem (LPP) with a new variable ϵ. For each $\frac{m_k}{Pre_{kj}} \leq \frac{m_u}{Pre_{uj}}$ defining $\mathcal{R}_i$, a constraint of the following form is added: $\frac{m_k}{Pre_{kj}} + \epsilon \leq \frac{m_u}{Pre_{uj}}$. The objective function will be to maximize ϵ. If the resulting LPP is infeasible or has solution $\epsilon = 0$ then $\mathcal{R}_i$ is a redundant region.

It may seem that if a mode is redundant, a set of arcs has to be implicit or timed implicit, since they cannot define the enabling. However, it is not true, since it is not that an arc never defines the enabling, but that a combination of arcs may never define the enabling. For example, in the net in Fig. 2(a), none of the arcs is implicit, although a region is reduced to its borders. In this example, the redundant mode could also have been avoided by fusing transitions t_1 and t_2 into a single one [16]. However, this kind of transformation cannot always be applied, as shown in the following example.

Example 7 *Let us consider the contPN in Fig. 2(b) and let us consider the region* $\mathcal{R}_1 = \{m_2 \leq m_1, m_3 \leq m_2, m_1 \leq m_3\}$ *that it is equivalent to assume that the enabling degree of* t_1 *is given by* m_2*, the one of* t_2 *by* m_3 *and of* t_3 *by* m_1*. Applying Prop. 6 we want to check if* $\mathcal{R}_1$ *is redundant. We have to consider the following system:*

$$\left\{ \begin{array}{ll} m_2 < m_1 & (1) \\ m_3 < m_2 & (2) \\ m_1 < m_3 & (3) \end{array} \right. \tag{5}$$

Combining (5.2) and (5.3) we obtain $m_1 < m_2$ *that is in contradiction with (5.1). Therefore, region* $\mathcal{R}_1$ *and mode* 1 *are redundant.*

4 Observability of unforced timed continuous Petri nets

Let us assume that the marking of some places $P_o \subseteq P$ can be measured, i.e., the token load at every time instant is known, due to some sensors. The problem is to estimate the other marking variables using these measurements. Going back to (4), the system considered here is given by:

$$\left\{ \begin{array}{l} \dot{\boldsymbol{m}}(\tau) = \boldsymbol{A}_i \cdot \boldsymbol{m}(\tau), \boldsymbol{m} \in \mathcal{R}_i, i \in I \\ \boldsymbol{y}(\tau) = \boldsymbol{S} \cdot \boldsymbol{m}(\tau) \end{array} \right. \tag{6}$$

where $\boldsymbol{S}$ is a $|P_o| \times |P|$ matrix, each row of $\boldsymbol{S}$ has all components zero except the one corresponding to the i^{th} measurable place that is 1. Observe that the matrix $\boldsymbol{S}$ is the same for all linear systems since the measured places are characteristic to the contPN system. Here it is considered that all linear systems are deterministic, i.e., noise-free.

Definition 8 *A timed contPN system $\langle \mathcal{N}, \boldsymbol{\lambda} \rangle$ with infinite server semantics is* observable in infinitesimal time *if it is always possible to compute its initial state $\boldsymbol{m}_0$ in any time interval $[0, \epsilon)$, $\forall \epsilon > 0$, by observing only a set of $P_o \subseteq P$ places.*

In the rest of the paper, our attention is focused on observability in infinitesimal time. Thus, if we say that a system is observable we understand that it is observable in infinitesimal time. To study this kind of observability, the following assumptions are considered:

A1. The net structure $\mathcal{N}$ and timing $\boldsymbol{\lambda}$ are known;

A2. The redundant modes are removed.

The observability of a JF contPN systems (a contPN is JF if there is no synchronization, i.e., $\forall t_j \in T, |^{\bullet}t_j| = 1$) has been studied in [8] using the results of classical linear system theory since a JF contPN system is a linear system, i.e., has only one mode. An interesting interpretation at the graph level is given for the state estimation procedure of a contPN system: going backward on path.

Example 9 *Let us consider the contPN in Fig. 1(a) and assume $P_o = \{p_1\}$, i.e., p_1 is measured. So, $m_1(\tau)$ is known at every time instant. Then, the derivative of the marking, i.e., $\dot{m}_1(\tau)$, can be computed, and also the flow of the transition t_1 because $f_1(\tau) = \lambda_1 \cdot m_1(\tau)$ since by assumption (A1.), the vector $\boldsymbol{\lambda}$ is known. Evidently, the flow of t_3 is deduced immediately using that $f_3(\tau) = \dot{m}_1(\tau) + f_1(\tau)$. Thus the marking of p_3 can be computed because, $f_3(\tau) = \lambda_3 \cdot \frac{m_3(\tau)}{2}$. Knowing $m_3(\tau)$ we can derive $\dot{m}_3(\tau)$, but $f_1(\tau)$ and $f_3(\tau)$ are also known, hence $f_2(\tau)$ can be estimated as: $f_2(\tau) = \dot{m}_3(\tau) + 2 \cdot f_3(\tau) - f_1(\tau)$ that permits to estimate $m_2(\tau)$ since $f_2(\tau) = \lambda_2 \cdot m_2(\tau)$.*

Here, we consider the problem of state estimation of the general contPN systems, not only JF. In this case, a very important problem for the observability is the determination of the mode, also called discrete state, i.e, the mode in which the system is. It may happen that the continuous state estimation fits with different discrete states, i.e., observing some places, it may happen that more than one linear system satisfy the observation. If the continuous states are on the border of some regions, it is not important which linear system is assigned, but it may happen that the solution corresponds to interior points of some regions and it is necessary to distinguish between them.

Example 10 *Let us consider the timed contPN in Fig. 1(b) and assume* $\boldsymbol{\lambda} = \mathbf{1}$ *and* $P_o = \{p_3\}$*, i.e.,* $\boldsymbol{S} = [0\ 0\ 1]^T$*. This system has two modes (see Ex. 3) corresponding to the following linear systems:*

$$\Sigma_i = \begin{cases} \dot{\boldsymbol{m}}(\tau) = \boldsymbol{A}_i \cdot \boldsymbol{m}(\tau) \\ \boldsymbol{y}(\tau) = [0\ 0\ 1] \cdot \boldsymbol{m}(\tau) \end{cases} \tag{7}$$

The observability matrices (see the appendix for the classical definition) of these two linear systems are:

$$\vartheta_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ -3 & 1 & 1 \end{bmatrix}; \quad \vartheta_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -3 & 1 \end{bmatrix}$$

Both have full rank, meaning that both linear systems are observable. Let us take $\boldsymbol{m}_1 = [1\ 2\ 0]^T \in \mathcal{R}_1 \setminus \mathcal{R}_2$ *and* $\boldsymbol{m}_2 = [2\ 1\ 0]^T \in \mathcal{R}_2 \setminus \mathcal{R}_1$*. As it is well-known, the corresponding observations are* $\vartheta_i \boldsymbol{m}_i(\tau) = [\boldsymbol{y}(\tau)\ \dot{\boldsymbol{y}}(\tau)\ \ldots]^T$*. Nevertheless, for the selected markings we have that* $\vartheta_1 \cdot \boldsymbol{m}_1 = \vartheta_2 \cdot \boldsymbol{m}_2 = [0\ 1\ -1]^T$*. Therefore, it is impossible to distinguish between* $\boldsymbol{m}_1$ *and* $\boldsymbol{m}_2$*.*

Definition 11 *Two modes* i *and* j *of a contPN system are* distinguishable *if for any* $\boldsymbol{m}_1 \in \mathcal{R}_1 \setminus \mathcal{R}_2$ *and any* $\boldsymbol{m}_2 \in \mathcal{R}_2 \setminus \mathcal{R}_1$ *the observation* $\boldsymbol{y}_1(\tau)$ *for the trajectory through* $\boldsymbol{m}_1$ *and the observation* $\boldsymbol{y}_2(\tau)$ *for the trajectory through* $\boldsymbol{m}_2$ *are* different *on an interval* $[0, \epsilon)$*.*

Remark that we remove the solutions at the border $\mathcal{R}_1 \cap \mathcal{R}_2$ since for those points both linear systems lead to identical behavior, therefore it is not important which one is chosen.

An immediate sufficient condition for being distinguishable is:

Proposition 12 *Let* i $\{i = 1, 2\}$ *be a mode,* ϑ_i *and* $\mathcal{R}_i$ *the corresponding observability matrix and region. If the LPP*

$$\begin{array}{ll} \max & \epsilon \\ s.t. & -\epsilon \cdot \mathbf{1} \leq \boldsymbol{m}_1 - \boldsymbol{m}_2 \leq \epsilon \cdot \mathbf{1} \\ & \vartheta_1 \cdot \boldsymbol{m}_1 - \vartheta_2 \cdot \boldsymbol{m}_2 = 0 \\ & \boldsymbol{m}_1 \in \mathcal{R}_1 \\ & \boldsymbol{m}_2 \in \mathcal{R}_2 \\ & \boldsymbol{m}_1, \boldsymbol{m}_2 \geq \mathbf{0} \end{array} \tag{8}$$

has the solution $\epsilon = 0$*, then the modes* 1 *and* 2 *are distinguishable.*

Proof: First, observe that LPP (8) has always $\epsilon = 0$, i.e., $\boldsymbol{m}_1 = \boldsymbol{m}_2 = \boldsymbol{0}$, as a solution. On the other hand, maximizing ϵ, the infinite norm of two possible markings $\boldsymbol{m}_1 \in \mathcal{R}_1$ and $\boldsymbol{m}_2 \in \mathcal{R}_2$ is obtained. Therefore, if the solution of (8) is $\epsilon = 0$ means that do no exist two markings $(\boldsymbol{m}_1, \boldsymbol{m}_2) \in (\mathcal{R}_1 \setminus \mathcal{R}_2) \times (\mathcal{R}_2 \setminus \mathcal{R}_1)$ for which their outputs, i.e., $\vartheta_1 \cdot \boldsymbol{m}_1$ and $\vartheta_2 \cdot \boldsymbol{m}_2$ are equal. So, given a marking in any of these regions we can determine the mode that governs the evolution of the contPN system. ■

Example 13 *In Ex. 10, for the timed contPN in Fig. 1(b) it is shown that $\vartheta_1 \cdot \boldsymbol{m}_1 = \vartheta_2 \cdot \boldsymbol{m}_2 = [0, 1, -1]^T$. Solving LPP (8), the problem is found to be unbounded, thus according to Prop. 12 we cannot conclude that the modes 1 and 2 are distinguishable. For the interpretation of this result, let us consider the equations that govern the evolution of the system:*

$$f_3 = \lambda_3 \cdot \min\{m_1, m_2\} \tag{9}$$

$$\dot{m}_1 = \lambda_2 \cdot m_2 - \lambda_1 \cdot m_1 - f_3 \tag{10}$$

$$\dot{m}_2 = \lambda_1 \cdot m_1 - \lambda_2 \cdot m_2 - f_3 \tag{11}$$

Summing (10) and (11) and integrating, we obtain

$$(m_1 + m_2)(\tau) = (m_1 + m_2)(0) - 2 \int_0^\tau f_3(\theta) \cdot d\theta \tag{12}$$

Obviously, if p_3 is measured, f_3 can be estimated since $f_3(\tau) = \dot{m}_3(\tau) + \lambda_4 \cdot m_3(\tau)$. Therefore, according to (9), the minimum between m_1 and m_2 is estimated. Moreover, due to (12) their sum is also known. Nevertheless, these two equations are not enough to compute the markings, i.e., we have the values but it is impossible to distinguish which one corresponds to which place.

We use the same contPN system to illustrate that Prop. 12 provides only a sufficient condition. Let us take now $\boldsymbol{\lambda} = [2\ 1\ 1\ 1]^T$. In this case, the dynamical matrices are:

$$\boldsymbol{A}_1 = \begin{bmatrix} -3 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}; \quad \boldsymbol{A}_2 = \begin{bmatrix} -2 & 0 & 0 \\ 2 & -2 & 0 \\ 0 & 1 & -1 \end{bmatrix}.$$

and the observability matrices (assuming also $P_o = \{p_3\}$):

$$\vartheta_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ -4 & 1 & 1 \end{bmatrix}; \quad \vartheta_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & -3 & 1 \end{bmatrix}$$

Let $\boldsymbol{m}_1 = [1\ 5\ 1]^T \in \mathcal{R}_1 \setminus \mathcal{R}_2$ *and* $\boldsymbol{m}_2 = [2\ 1\ 1]^T \in \mathcal{R}_2 \setminus \mathcal{R}_1$. *Making the computations, we have:* $\vartheta_1 \boldsymbol{m}_1 = \vartheta_2 \boldsymbol{m}_2 = [1\ 0\ 2]^T$. *So, we have the same observations for these two markings at a time* τ *but the modes are distinguishable. To see this let us assume the marking at* $\tau + \delta$, *where* δ *is a very small value. Being a small time variation, we can consider that the flow of the transitions are constant during the time interval* $(\tau, \tau + \delta)$ *and we can write:*

$$\boldsymbol{m}_1'(\tau + \delta) = \boldsymbol{m}_1(\tau) + \boldsymbol{A}_1 \boldsymbol{m}_1(\tau)\delta = [1 + 2\delta \ \ 5 - 4\delta \ \ 1]^T$$

and

$$\boldsymbol{m}_2'(\tau + \delta) = \boldsymbol{m}_2(\tau) + \boldsymbol{A}_2 \boldsymbol{m}_2(\tau)\delta = [2 - 4\delta \ \ 1 + 2\delta \ \ 1]^T.$$

The corresponding observations for these markings are: $\vartheta_1 \boldsymbol{m}_1' = [1\ 2\delta\ 2 - 12\delta]^T \neq \vartheta_2 \boldsymbol{m}_2' = [1\ 2\delta\ 2 - 14\delta]^T$. *Since in all linear systems the set of measured places is the same and the firing rates are also the same can be observed immediately that any* $\boldsymbol{m}_1'' \in \mathcal{R}_1$, $\boldsymbol{m}_2'' \in \mathcal{R}_1$ *with* $\vartheta_1 \boldsymbol{m}_1''(\tau) = \vartheta_2 \boldsymbol{m}_2''(\tau)$ *then* $\vartheta_1 \boldsymbol{m}_1''(\tau + \delta) \neq \vartheta_2 \boldsymbol{m}_2''(\tau + \delta)$. *Therefore, according to Def. 11, the modes are distinguishable.*

Remark 14 *Prop. 12 provides only a sufficient condition for two modes to be distinguishable. Considering the particular structure of our systems, i.e., the same places are measured in both systems and the same firing rates are assigned to the transitions, if* (8) *has a solution at time* τ, *the next step is to check the solution of this system after a small variation of time,* $\tau + \delta$. *If there exists no solution, the modes are distinguishable (see previous example).*

Using the notion of distinguishable modes, an immediate criterium for observability in infinitesimal time is:

Theorem 15 *A timed continuous Petri net system* $\langle \mathcal{N}, \boldsymbol{\lambda} \rangle$ *under infinite server semantics is observable in infinitesimal time iff:*

1. *All modes are distinguishable,*
2. *For each region, the associated linear system is observable.*

Proof: Assume that given an observation here are two different markings $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ coherent with it. Since the linear systems are observable, $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ belong to different regions. But the modes are all distinguishable, contradiction.

If the contPN is observable, for any initial marking in any region it must be possible to reconstruct it from observation, hence all the linear systems, i.e., modes, have to be observable. Moreover, the modes have to be distinguishable, since otherwise it would be possible to have two different markings that fit with the observation. ■

5 Structural observability

Observability has been defined for a timed contPN system $\langle \mathcal{N}, \boldsymbol{\lambda} \rangle$, so the firing rates of the transitions are fixed. Since the firing rate vector represents the speed of machines or servers, in many cases, an interesting problem is to study the observability for any value of their rate. Imagine that we want to design an observer and we know that in the future some machines will be replaced but we don't know exactly which one will be bought, hence their speed is not fixed. In this section, we concentrate on the study of the observability of contPNs under infinite server semantics in infinitesimal time for any value of firing rate $\boldsymbol{\lambda}$. Because does not depend on $\boldsymbol{\lambda}$, only on $\mathcal{N}$, we call this problem: *structural observability*. The following assumptions are done:

A1. The net structure $\mathcal{N}$ is known and $\boldsymbol{\lambda}$ is a parameter;

A2. The redundant modes and regions are removed.

Definition 16 *Let $\langle \mathcal{N}, \boldsymbol{m}_0 \rangle$ be a contPN system and $P_o \subseteq P$ the set of measured places.*

- *A place $p_i \in P$ is structurally observable from P_o if for all $\boldsymbol{\lambda} > 0$, $m_i(\tau_0)$ can be computed in $\langle \mathcal{N}, \boldsymbol{\lambda}, \boldsymbol{m}_0 \rangle$ by measuring the marking evolution of the places in P_o.*
- *Let $K(P_o)$ be the set of places structurally observable from P_o. $\mathcal{N}$ is structurally observable from P_o if every place $p_i \in P$ is structurally observable, i.e., $K(P_o) = P$.*

Due to the graphical representation of PNs, the observation procedure has a quite interesting interpretation, going backward on the net (see [8] for more details).

Example 17 *Let us consider the contPN system in Fig. 3(a) and assume that p_2 is measured and $\boldsymbol{\lambda}$ is known. So, the marking in p_2 is known at every time instant as well as the flow of t_4 since $f_4 = \dot{m}_2 + \lambda_2 m_2$ and then the marking of p_4 can be computed because, on the other hand, $f_4 = \lambda_4 \cdot m_4$.*

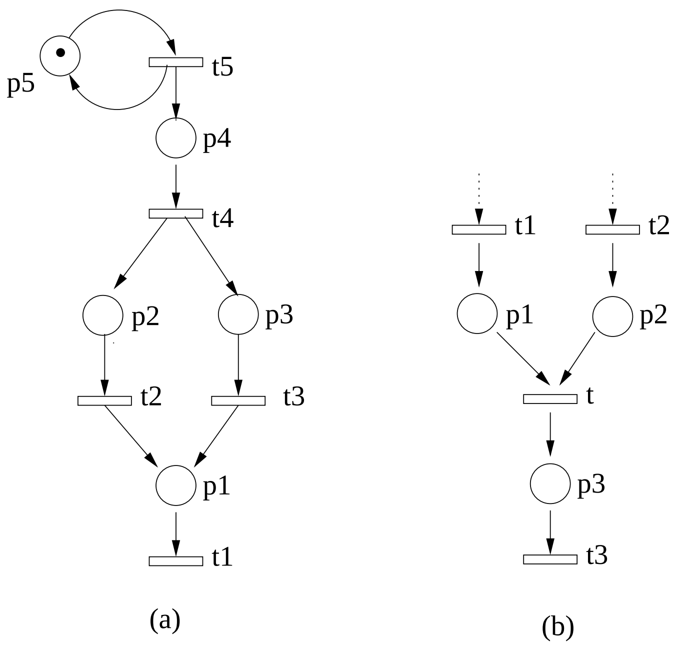

Figure 3: Structural and generic observability: (a) a JF net that if $\lambda_2 = \lambda_3$ then m_2 and m_3 cannot be estimated from the observation of p_1; (b) a PN observable for any initial marking only if all places are measured.

The backward procedure explained in the previous example assumes that all places have only one input transition (attribution free: $p_i \in P$ is an attribution if $|^\bullet p_i| > 1$) and all transitions have only one input place (join free: $t_j \in T$ is a join if $|^\bullet t_j| > 1$).

Definition 18 *Let $\mathcal{N}$ be a contPN.*

- *A place $p_i \in P \setminus P_o$ is* output connected *if there exists a path, denoted $\mathcal{P}_i$, from p_i to a measured place $p_j \in P_o$: $\mathcal{P}_i = \langle p_i, t_i, p_{i+1}, t_{i+1}, \ldots p_{j-1}, t_{j-1}, p_j \rangle$ with $t_i \in p_i^\bullet$, $p_{i+1} \in t_i^\bullet$, ..., $p_j \in t_{j-1}^\bullet$.*
- *$\mathcal{N}$ is output connected if all places are output connected.*

For join attribution free (JA-F) PN, structural observability is solved without difficulty using the basic backward strategy presented above (see [8] for more details).

Lemma 19 *Let $\mathcal{N}$ be a JA-F contPN. A place $p_i \in P$ is structurally observable iff it is output connected.*

Proof: To be observable, obviously p_i should be output connected. On the reverse sense, applying Alg. 5 in [8], going backward from an output to p_i, the rank condition of the algorithm is always satisfied since the matrix required to be full-rank is a 1×1 with a non-null element, so has full rank. According to Prop. 6 in [8], p_i is structurally observable. ∎

Lemma 19 helps us to determine a set of places with minimum cardinality, denoted P_o, that ensures the structural observability of a JA-F contPN. Observe that this is another problem that the one in [8] where the set P_o is given and the problem is to compute the set of structurally observable places. For this, the strongly connected components are used.

Definition 20 *Let $\mathcal{N} = \langle P, T, \boldsymbol{Pre}, \boldsymbol{Post} \rangle$ be a net and $\mathcal{N}' = \langle F, T', \boldsymbol{Pre}', \boldsymbol{Post}' \rangle$ a subnet of $\mathcal{N}$, i.e., $F \subseteq P$, $T' \subseteq T$ and $\boldsymbol{Pre}', \boldsymbol{Post}'$ are the restrictions of $\boldsymbol{Pre}, \boldsymbol{Post}$ to F and T'. $\mathcal{N}'$ is a strongly connected component of $\mathcal{N}$ w.r.t. the places if for all $p_1, p_2 \in F$ there is a path from p_1 to p_2 of the form $\langle p_1, t_1, p_i, t_i, \ldots, t_j, p_j, t_2, p_2 \rangle$ with $t_1 \in p_1^\bullet$, $p_i \in t_1^\bullet$, ..., $p_j \in t_j^\bullet$, $t_2 \in p_j^\bullet$, $p_2 \in t_2^\bullet$.*

Abusing of notation it will be said that a set of places F defines a strongly connected component of $\mathcal{N}$ if $\mathcal{N}'$ is a strongly connected component of $\mathcal{N}$ with $\mathcal{N}'$ the subnet generated by F, i.e., $T' = {}^\bullet F \cup F^\bullet$.

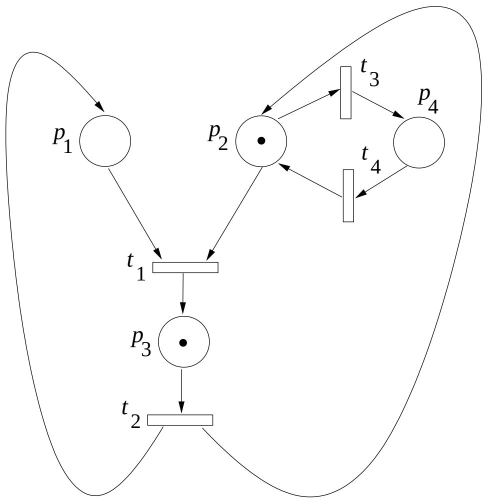

Figure 4: A simple contPN system.

The net in Fig. 4 has only one strongly connected component $F = \{p_1, p_2, p_3, p_4\}$ because a path exists connecting any two places. For example from p_1 to p_4 there is: $\langle p_1, t_1, p_3, t_2, p_2, t_4, p_4 \rangle$. The net in Fig. 3(a) has 5 strongly connected components, each one corresponding to a place, i.e., $F_i = \{p_i\}$, $i = 1 \ldots 5$.

Output connectedness is required for structural observability but also for observability. Obviously, for those places for which there is no path to an output their marking cannot be estimated because they do not affect the observed outputs. Therefore, the *terminal strongly connected components* present a special interest because any place of the net is connected to those components.

Definition 21 *A strongly connected component* $\mathcal{N}' = \langle F, T', \boldsymbol{Pre}', \boldsymbol{Post}' \rangle$ *of a net* $\mathcal{N}$ *is said to be* terminal *if there is no path from a place belonging to* F *to a place not in* F*.*

Strongly-connected components of a PN can be computed immediately, adapting the classical algorithms (for example the one in [6]) to a bipartite graph. The net in Fig. 4 has a unique strongly connected component which is obviously terminal, while the net in Fig. 3.(a) has only one terminal strongly connected component $F_1 = \{p_1\}$.

Proposition 22 *Let $\mathcal{N}$ be a JA-F contPN. $\mathcal{N}$ is structurally observable iff at least one place from each terminal strongly connected component is measured.*

Proof: If $\mathcal{N}$ is structurally observable then every place that is not measured should be output connected to a place that is measured. Therefore at least one place from each terminal strongly connected component should be measured.

On the contrary, if at least one place from each terminal strongly connected component is measured, every place is output connected and the net is structurally observable according to Prop. 19. If $\mathcal{N}$ is structurally observable, it is observable for any particular $\boldsymbol{\lambda}$. ■

Therefore, the minimum number of places to ensure the structural observability of a JA-F contPN is equal to the number of terminal strongly connected components. Let us see what happens when joins appear. According to (3), this introduces nonlinearities into the flow definition due to the minimum functions. This will cause problems in the observation procedure.

Example 23 *Let us consider the contPN subsystem in Fig. 3(b). It is structurally observable iff places $P_o = \{p_1, p_2, p_3\}$ are measured.*

Place p_3 must be measured (it is a terminal strongly-connected component). Using m_3, f_3 is obtained since $f_3 = \lambda_3 \cdot m_3$. Hence, the flow of t is immediately computed as $f_t = \dot{m}_3 + f_3$. On the other hand, this flow is equal to $f_t = \lambda_t \cdot \min\{m_1, m_2\}$. In the last expression, f_t and λ_t are known which implies that the minimum of m_1 and m_2 can be evaluated. If $m_1 \leq m_2$ (place p_2 does not constraint the firing of t), m_1 equals to the minimum and p_2 must be measured. Identically, if at a certain moment $m_2 \leq m_1$, p_1 should be measured. Therefore, if no information regarding how m_1 and m_2 compare is known, then the only solution *for observability is to measure both p_1 and p_2. Moreover, since p_1, p_2 and p_3 are measured, the observability space of this system is the same as that of the system obtained removing the join transition t.*

Proposition 24 *Let $\langle \mathcal{N}, \boldsymbol{\lambda} \rangle$ be a timed AF contPN and assume that for any join t_i there exists no strongly connected component containing all ${}^\bullet t_i$. Let $\mathcal{N}'$ be the net obtained from $\mathcal{N}$ by just removing all join transitions together with its input and output arcs. $\mathcal{N}$ is structurally observable iff $\mathcal{N}'$ is structurally observable.*

Proof: Let us assume, for simplicity, that $\mathcal{N}$ has only one join t_i with $^\bullet t_i = \{p_1, p_2\}$ (the proof can be easily extended). In this case, there exist two modes corresponding to $\mathcal{R}_1 = \left\{ \frac{m_1}{Pre_{1i}} \leq \frac{m_2}{Pre_{2i}} \right\}$ and $\mathcal{R}_2 = \left\{ \frac{m_2}{Pre_{2i}} \leq \frac{m_1}{Pre_{1i}} \right\}$.

"$\Longleftarrow$" $\mathcal{N}'$ structurally observable, hence all places are output connected according to Lemma 19. Obviously, the same paths exist in all modes of $\mathcal{N}$, therefore each one is structurally observable according to the same Lemma 19 being AF the corresponding net of the region. The modes are also distinguishable because $\mathcal{N}'$ is output connected, so p_1 and p_2 are estimated using paths not containing t_i therefore their markings are identified in each mode.

"$\Longrightarrow$" $\mathcal{N}$ is structurally observable then it is observable for any value of $\boldsymbol{\lambda}$. $\langle \mathcal{N}, \boldsymbol{\lambda} \rangle$ is observable iff every linear system is observable and the modes are distinguishable according to Prop. 15.

Let us consider first the linear system of $\mathcal{R}_1$. The enabling degree of t_i is given by the marking of p_1 so the arc (p_2, t_i) is "invisible" in this mode, i.e., the backward procedure explained before cannot use (p_2, t_i). On the other hand, the linear system is observable then p_2 should be output connected according to Lemma 19. Let us denote by $\mathcal{P}_1$ the path from p_2 to one output in this mode. Evidently, as discussed before, the arc (p_2, t_i) does not belong to $\mathcal{P}_1$ but (p_1, t_i) can eventually belong. In this last case, i.e., (p_1, t_i) belongs to $\mathcal{P}_1$, it is obvious that should exists a backward path from p_2 to p_1 not containing t_i.

Analogously, the linear system of $\mathcal{R}_2$ is observable (for this system, the enabling degree of t_i is given by m_2), then a path, denoted $\mathcal{P}_2$, from p_1 to an output should exist but not using the arc (p_1, t_i) that is "invisible" in this mode.

We have three cases:

(i) If $\mathcal{P}_1$ and $\mathcal{P}_2$ do not contain t_i then these paths exist also in $\mathcal{N}'$, so it is observable being all places output connected.

(ii) Assume t_i belongs to $\mathcal{P}_1$ but not to $\mathcal{P}_2$. Let us concentrate first on the linear system of $\mathcal{R}_2$. Since it is observable and (p_1, t_i) and (p_2, t_i) do not belong to $\mathcal{P}_2$, p_1 is output connected without passing through the join transition t_i. On the other hand, the linear system of $\mathcal{R}_1$ is observable, so p_2 is output connected and the corresponding path is $\mathcal{P}_1$. If (p_1, t_1) is not belonging to $\mathcal{P}_1$ then both p_1 and p_2 are output connected to two outputs without t_i and exactly as in (i) $\mathcal{N}'$ is observable. Otherwise, if (p_1, t_1) belongs to $\mathcal{P}_1$ according to the previous remark should exists a backward path from p_2 to p_1 not containing t_i. But p_1 is output connected without passing through t_i so also p_2.

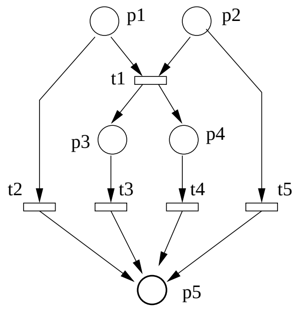

Figure 5: ContPN used in Ex. 25.

(iii) If both $\mathcal{P}_1$ and $\mathcal{P}_2$ contain t_i then there exists a backward path from p_1 to p_2 and also from p_2 to p_1 according to the previous remark. Obviously, p_1 and p_2 belong to a strongly connected component of $\mathcal{N}'$ and the hypothesis is not satisfied. In fact, in this case, the join can "quote" to observe the system (see Ex. 13). ■

The previous theorem does not hold when the net is not AF.

Example 25 *Let us consider the contPN system in Fig. 5 with* $\boldsymbol{\lambda} = [a, 1, 2, 3, 4]^T$*,* $a \in \mathbb{R}_{\geq 0}$ *and* p_5 *measured. This net is not an AF net and has a join in* t_1*. Notice that the linear system obtained removing the join* t_1 *is observable and* p_1 *and* p_2 *do not belong to a strongly connected component. This system has the following two modes:*

$$\Sigma_1 = \begin{cases} \dot{\boldsymbol{m}}(\tau) = \begin{bmatrix} -1-a & 0 & 0 & 0 & 0 \\ -a & -4 & 0 & 0 & 0 \\ a & 0 & -2 & 0 & 0 \\ a & 0 & 0 & -3 & 0 \\ 1 & 2 & 3 & 4 & 0 \end{bmatrix} \cdot \boldsymbol{m}(\tau) \\ \boldsymbol{y}(\tau) = [0,0,0,0,1] \cdot \boldsymbol{m}(\tau) \end{cases}, \quad \Sigma_2 = \begin{cases} \dot{\boldsymbol{m}}(\tau) = \begin{bmatrix} -1 & -a & 0 & 0 & 0 \\ 0 & -4-a & 0 & 0 & 0 \\ 0 & a & -2 & 0 & 0 \\ 0 & a & 0 & -3 & 0 \\ 1 & 2 & 3 & 4 & 0 \end{bmatrix} \cdot \boldsymbol{m}(\tau) \\ \boldsymbol{y}(\tau) = [0,0,0,0,1] \cdot \boldsymbol{m}(\tau) \end{cases}$$

with observability matrices:

$$\vartheta_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 4 & 0 \\ 4 \cdot a - 1 & -8 & -6 & -12 & 0 \\ -(4 \cdot a - 1) \cdot (a+1) - 10 \cdot a & 12 & 36 & 0 & \\ ((4 \cdot a - 1) \cdot (a+1) + 10 \cdot a) \cdot (a+1) + 16 \cdot a & -128 & -24 & -108 & 0 \end{bmatrix}$$

$$\vartheta_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 4 & 0 \\ -1 & 4 \cdot a - 8 & -6 & -12 & 0 \\ 1 & -(4 \cdot a - 8) \cdot (a+4) - 17 \cdot a & 12 & 36 & 0 \\ -1 & ((4 \cdot a - 8) \cdot (a+4) + 17 \cdot a) \cdot (a+4) + 47 \cdot a & -24 & -108 & 0 \end{bmatrix}$$

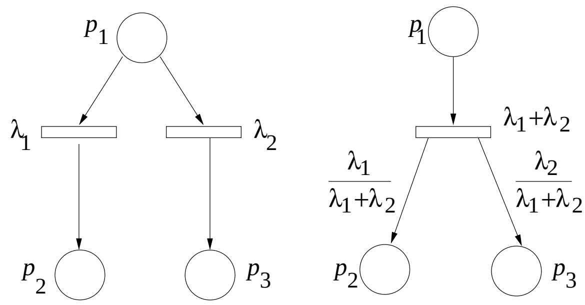

Figure 6: Transforming a CEQ net into a choice free one.

Computing the determinants of the observability matrices, we have: $det(\vartheta_1) = 192 \cdot a^3 - 912 \cdot a^2 + 720 \cdot a + 288$, *which has two positive real roots, and* $\det(\vartheta_2) = -96 \cdot a^3 - 408 \cdot a^2 - 216 \cdot a + 288$, *with one positive real root. Obviously, if* λ_1 *is equal to one of these roots, the contPN system will not be observable since one of the corresponding mode will not be observable.*

Hence, for some particular values of $\boldsymbol{\lambda}$*, the system obtained removing the join is observable but the original system (with join) is not observable.*

Prop. 24 provides the conditions under which the observability of an AF contPN net can be studied using the linear system theory (the joins can be eliminated and a linear system is obtained). In this case, a minimal cardinality set P_o of places that ensures its structural observability can be computed as in JA-F case using Prop. 22, after the joins are eliminated. Notice that due to the elimination of the joins, several unconnected PN can be obtained. In this case, Prop. 22 is applied for each connected component.

Proposition 26 *The structural observability of an AF contPN can be solved in polynomial time at the graph level.*

In the case of continuous equal conflict (CEQ) nets (nets for which if t_1 and t_2 are in conflict, there exists $k > 0$ such that $\boldsymbol{Pre}_{1\cdot} = k \cdot \boldsymbol{Pre}_{2\cdot} \neq 0$, i.e., a generalization of equal conflict relation), observability can be studied also using linear system theory since the joins can be eliminated as in the case of AF nets.

Proposition 27 *Let* $\langle \mathcal{N}, \boldsymbol{\lambda} \rangle$ *be a timed CEQ contPN system and* $\mathcal{N}'$ *obtained from* $\mathcal{N}$ *by just removing all join transitions together with its input and output arcs.* $\mathcal{N}$ *is (structurally) observable iff* $\mathcal{N}'$ *is (structurally) observable.*

Proof: First, notice that the net $\mathcal{N}$ can be transformed into an equivalent choice free fusing the CEQ [9] (see Fig. 6 for an example). Hence, it can be assumed that the net has no choice, i.e., for every join transition t_i and for any $p_j \in {}^\bullet t_i$, t_i is the only output transition of p_j, i.e., $p_j^\bullet = \{t_i\}$.

"$\Longrightarrow$" Let t_i be a join, and $p_1, p_2 \in {}^\bullet t_i$. For those markings in the region defined by $\frac{m_1}{Pre_{1i}} \leq \frac{m_2}{Pre_{2i}}$, the only way to observe p_2 is to measure it, since its only output is t_i and the arc (p_2, t_i) is "invisible" in this mode.

The same can be said for p_1, hence both places have to be measured and removing their output transition cannot affect the (structural) observability of the system.

"$\Longleftarrow$"If $\mathcal{N}'$ is structurally observable it is observable for a given $\boldsymbol{\lambda}$. Since $\mathcal{N}'$ is obtained from $\mathcal{N}$ by removing the joins and $\mathcal{N}$ is CEQ, all input places in the joins of $\mathcal{N}$ have no output transition in $\mathcal{N}'$. But $\mathcal{N}'$ is observable, hence all these places should be measured because cannot be estimated with others measurements. Measuring these places, the corresponding linear systems of $\mathcal{N}$ are distinguishable. Moreover, all linear systems are observable since the observability does not depends on the firing rates of the output transitions of the measured places (Prop. 6 in [8]). According to Th. 15 $\mathcal{N}$ is observable. ■

Unfortunately, the elimination of joins cannot be performed in general, because for nets with attributions the observability should be studied globally, not locally (see Ex. 25). For CEQ nets, that in principle have attributions, all joins can be removed (Prop. 27), but only because their input places must be measured, which is not true in general. In Ex. 25 taking $\boldsymbol{\lambda} = [a, 1, 2, 3, 4]^T$ with a different from the roots of $det(\vartheta_1)$ and $det(\vartheta_2)$ observability is guaranteed. Since t_1, t_2, t_5 are not in CEQ relation, it is not mandatory to measure all their input places. Indeed, this contPN system is observable measuring only p_5.

Summarizing, joins can be removed without affecting the observability for AF nets under some conditions (Prop. 24) and for CEQ nets (Prop. 27). Therefore, *forks* and *choices* do not pose any problem for observability (JA-F case), while *joins* are real "barriers" in the backward procedure. Let us now consider attributions, the only local construction of nets not yet studied. This construction can introduce zeros in the transfer functions, possibly leading to pole-zero cancelations, thus loss of observability.

Example 28 *Let us consider the JF contPN system in Fig. 3(a) (it has an attribution in p_1) and assume p_1 is measured. This system is a continuous linear system. If we consider that the input of the system is the input flow to p_4 and the measured output is m_1, the equivalent linear system* $\dot{\boldsymbol{x}}(\tau) =$

$\boldsymbol{A} \cdot \boldsymbol{x}(\tau), \boldsymbol{y}(\tau) = \boldsymbol{S} \cdot \boldsymbol{x}(\tau)$ *has:*

$$\boldsymbol{A} = \begin{pmatrix} -\lambda_1 & \lambda_2 & \lambda_3 & 0 \\ 0 & -\lambda_2 & 0 & \lambda_4 \\ 0 & 0 & -\lambda_3 & \lambda_4 \\ 0 & 0 & 0 & -\lambda_4 \end{pmatrix}, \quad \boldsymbol{S} = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

The transfer function vector between the input flow in places and the output, using [Equation (19)*] is:*

$$\mathcal{Y}(s) = \frac{1}{(s+\lambda_1)(s+\lambda_2)(s+\lambda_3)(s+\lambda_4)} H^T \tag{13}$$

where:

$$H = \begin{bmatrix} (s+\lambda_2)\cdot(s+\lambda_3)\cdot(s+\lambda_4) \\ \lambda_2\cdot(s+\lambda_3)\cdot(s+\lambda_4) \\ \lambda_3\cdot(s+\lambda_2)\cdot(s+\lambda_4) \\ (\lambda_2\cdot(s+\lambda_3)+\lambda_3\cdot(s+\lambda_2)) \end{bmatrix} \tag{14}$$

In Equations (13) *and* (14)*, if* $\lambda_2 = \lambda_3$ *there is a pole-zero simplification in* all *elements of vector* $\mathcal{Y}(s)$ *leading to the conclusion that the system is not observable [18]. If* $\lambda_2 \neq \lambda_3$*, but* $\lambda_4 = \frac{2\cdot\lambda_2\cdot\lambda_3}{\lambda_2+\lambda_3}$*, there is another simplification and the system is also not observable. Consequently, when an attribution appears, particular values of* $\boldsymbol{\lambda}$ *exist such that the observability is lost. Moreover, it is not a local property, but depends on the whole net structure.*

Usually, if p_j is an attribution place with $t_1, t_2 \in {}^\bullet p_j$ and $\lambda_1 = \lambda_2$, then there exists a pole-zero cancelation and an additional place should be measured. But this is not a general rule as illustrated in the following example.

Example 29 *Let us consider the net in Fig. 7 with* $\boldsymbol{\lambda} = \mathbf{1}$ *and assume that* p_2 *is measured. Then* p_4 *and* p_5 *cannot be estimated directly, but a linear combination of the markings of these places is known (place* p_{45} *in Fig. 7). Going backwards,* p_1 *is estimated and, even although* p_1 *is an attribution, since* p_2 *is measured* p_3 *is also estimated. Using the marking of* p_3*,* p_4 *is estimated and through the linear combination of* p_{45}*,* p_5 *as well. Therefore, the system is observable measuring* p_2*.*

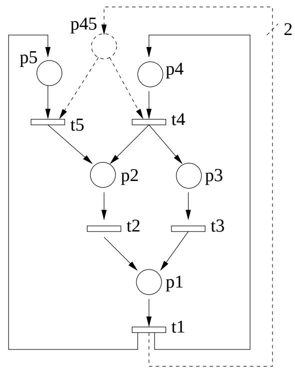

Figure 7: A JF net that is observable measuring the attribution place p_2 even if $\lambda_4 = \lambda_5$.

6 Generic observability

In Ex. 28, the pole-zero cancelation due to the attribution happens for very specific values of $\boldsymbol{\lambda}$. If the firing rates of the transitions are chosen randomly in $\mathbb{R}^+$, the probability to obtain this cancelation is null. Hence, a concept weaker than structural observability can be studied. It is defined following ideas in [5, 7] for linear systems. Hence, we assume:

A1. The net structure $\mathcal{N}$ is known and $\boldsymbol{\lambda}$ is a parameter;

A2. $\mathcal{N}$ is a JF net.

According to the results in the previous section generic observability can be studied also for some AF and CEQ nets. In these cases, as explained before, joins can be removed and the obtained JF net is observable iff the original net is observable. In a JF net, choices are CEQ, thus can be transformed into forks [9], and a JC-F net is obtained. Therefore, we can assume that the nets are JC-F.

Definition 30 *Let $\langle \mathcal{N}, \boldsymbol{\lambda}, \boldsymbol{m_0} \rangle$ be a JF contPN system and P_o the set of measured places. $\mathcal{N}$ is* weakly structural *or* generically observable *from P_o if $\langle \mathcal{N}, \boldsymbol{\lambda}, \boldsymbol{m_0} \rangle$ is observable for all values of $\boldsymbol{\lambda}$ outside a proper algebraic variety of the parameter space.*

Connection between structural and generic observability is obvious. If $\mathcal{N}$ is structurally observable then it is generically observable. The reverse is not true (see Ex. 28).

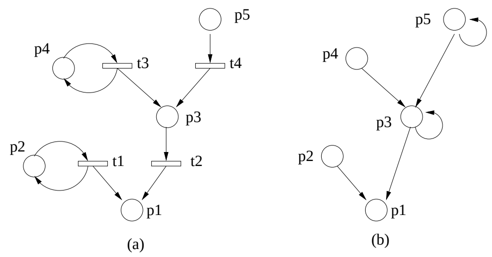

Figure 8: (a) A JF ContPN; (b) Associated graph.

In [5], generic observability is studied for structured linear systems using an *associated graph*; observability is guaranteed when there exists a state-output connection for every state variable (the system is said to be *output connected*) and no *contraction* (defined after) exists.

The associated graph of an unforced linear system (15), $G = (Z, W)$ is defined by a vertex set Z and an edge set W [5]. The vertex set $Z = X \cup Y$ with X the set of state vertices and Y the set of output vertices. Denoting (v, v') for a direct edge from the vertex $v \in Z$ to a vertex $v' \in Z$, the edge set W is described by $W_A \cup W_S$ with $W_A = \{(x_j, x_i)|A[i,j] \neq 0\}$ and $W_S = \{(x_j, y_i)|S[i,j] \neq 0\}$.

The transformation of a JF net into its corresponding *associated directed graph* is straightforward (see Fig. 8). The vertex set Z is given by the set P of places (i.e. $Z = P$). The edge set W is computed as: $W = \{(p_i, p_j)|p_j \in ({p_i}^\bullet)^\bullet \bigwedge p_i \neq p_j\} \cup \{(p_i, p_i)|\exists t \in {p_i}^\bullet, \boldsymbol{Pre}[p_i, t] \neq \boldsymbol{Post}[p_i, t]\}$. The first set adds an edge from a place p_i to all places $({p_i}^\bullet)^\bullet$ since the dynamic matrix has a non null entry and prevents adding an edge in the case of a self-loop. The second subset will add a self-loop in the associated graph for any place with $\boldsymbol{Pre}[p_i, t] \neq \boldsymbol{Post}[p_i, t]$, i.e., the marking of p_i will change firing t, implying that the dynamical matrix has an non zero entry.

Definition 31 *Let $\mathcal{N}$ be a contPN system and $G(\mathcal{N})$ its associated graph with vertex set Z and edge set W. Consider a set S made of k_S state vertices. Denote $E(S)$ the set of vertices w_i for $i = 1, \cdots, l_S$ of Z, such that there exists an edge $(x_j, w_i) \in W$ with $x_j \in S$. S is said to be a* contraction *if $k_S - l_S > 0$.*

Based on the procedure to generate the associated graph (Fig. 8), and using Prop. 1 in [5], the following is true:

Proposition 32 *Let $\mathcal{N}$ be a JF contPN and $G(\mathcal{N})$ its associated graph. $\mathcal{N}$ is generically observable iff:*

1. *$\mathcal{N}$ is output connected*
2. *$G(\mathcal{N})$ contains no contraction.*

Example 33 *Let us consider the contPN in Fig. 8(a) whose associated graph is sketched in Fig. 8(b). Taking $S = \{p_2, p_3, p_4, p_5\}$ ($k_S = 4$), $E(S) = \{p_1, p_3, p_5\}$ ($l_S = 3$). Thus, the net has a contraction ($k_S - l_S = 4 - 3 = 1$), so it is not generically observable. This happens because the flows of the transitions t_1 and t_3 are constant and measuring p_1 it is impossible to distinguish between these two constant incoming flows.* ■

In the case of pure contPN systems, the necessary and sufficient condition of generic observability can be simplified. Since the associated graph of a pure PN has in every node a self-loop (under infinite server semantics, if p_i has at least one output transition t_j the derivative of the marking is: $\dot{m}_i = \cdots - \lambda_j \cdot m_i + \cdots$). Therefore, no contraction can exist and the only remaining condition in Prop. 32 is the output connectedness.

Corollary 34 *Let $\mathcal{N}$ be a pure JF contPN. $\mathcal{N}$ is generically observable iff at least one place from each terminal strongly connected component is measured.*

The previous result can be extended immediately to general contPNs, i.e., it is not true only for JF nets. In Ex. 10 is given a contPN system containing two undistinguishable modes. Then, changing the firing rates of the transitions in Ex. 13, these modes becomes distinguishable. Obviously, two modes are undistinguishable when the path from states (markings) to the outputs are identical in both linear systems. This happens for some particular values of firing rates, e.g., $\lambda_1 = \lambda_2$ in the contPN of Fig. 1(b).

Corollary 35 *Let $\mathcal{N}$ be a pure contPN. $\mathcal{N}$ is generically observable iff at least one place from each terminal strongly connected component is measured.*

Proof: Output connection is a requirement for generic observability of each mode (Corollary 34). On the other hand, the output paths cannot be identically in two modes for all possible values of firing rates. ■

For example, the net in Fig. 1(b). This contPN system is not observable but it is generic observable as explained before.

7 Conclusions and future work

Different aspects of observability have been studied. First, the notion of *redundant modes* is introduced and a necessary and sufficient condition for a mode to be redundant is given. This permits to reduce the number of linear systems (modes) that can govern the evolution of a contPN, number that can be exponential. As already known, observability of a hybrid system requires not only the estimation of the continuous states but also of the discrete ones, i.e., it is important to know in which mode we are. To characterize this, the notion of *distinguishable modes* is introduced and a LPP is given to check if two modes are or are not distinguishable. Then, an observability criteria is given for general contPN systems. *Structural observability*, i.e., observability for all possible values of firing rates of the transitions, has been studied and for some subclasses the procedure to compute the set of places that ensures this kind of observability is given. At the same time, it is proved that the observability for CEQ nets and AF nets under some conditions can be studied on a linear systems by removing joins. Finally, an intermediate concept between observability and structural observability, called *weak structural* or *generic observability*, i.e., observability for almost all possible values of firing rates, has been studied also. It is illustrated how can be computed a set of places of minimum cardinality that ensures this kind of observability. In practice, the markings of some places, or the flow of some transitions, or the token conservation laws can be known. All these information can be used and the observability complexity problem is reduced. For example, by simply introducing the corresponding rows of a basis of P-flows in the observability matrix. Hence the results of this paper are immediately extended to the case of additional hypothesis.

References

[1] M. Babaali and M. Egerstedt. On the observability of piecewise linear systems. In *43rd IEEE Conference on Decision and Control (CDC 2004)*, pages 26–31, Paradise Island, Bahamas, December 2004.

[2] M. Babaali and G. J. Pappas. Observability of switched linear systems in continuous time. In M. Morari and L. Thiele, editors, *Hybrid Systems: Computation and Control, 8th International Workshop, HSCC 2005,*, volume 3414 of *Lecture Notes in Computer Science*, pages 103–117, Zurich, Switzerland, 2005. Springer.

[3] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, and A. L. Sangiovanni-Vincentelli. Observability for hybrid systems. In *Proc. 42nd IEEE Conference on Decision and Control*, Hawaii, USA, December 2003.

[4] P. Collins and J.H. van Schuppen. Observability of piecewise-affine hybrid systems. In R. Alur and G.J. Pappas, editors, *Hybrid Systems: Computation and Control, 7th International Workshop, HSCC 2004,*, volume 2993 of *Lecture Notes in Computer Science*, pages 265–279, Philadelphia, USA, 2004. Springer.

[5] C. Commault, J.M. Dion, and D.H. Trinh. Observability recovering by additional sensor implementation in linear structured systems. In *Proceedings of the* 44^{th} *IEEE Conference on Decission and Control 2005*, pages 7193–7197, Seville, Spain, December 2005.

[6] Thomas H. Cormen, E. Leiserson, Charles, and Ronald L. Rivest. *Introduction to Algorithms.* MIT Press, 1990.

[7] J.M. Dion, C. Commault, and J. van der Woude. Generic properties and control of linear structured systems: a survey. *Automatica*, 39(7):1125–1144, 2003.

[8] J. Júlvez, E. Jiménez, L. Recalde, and M. Silva. On Observability and Design of Observers in Timed Continuous Petri Net Systems. *IEEE Trans. on Autom. Science and Engineering*, 5(3):532–537, July 2008.

[9] J. Júlvez, L. Recalde, and M. Silva. Steady-state performance evaluation of continuous mono-T-semiflow Petri nets. *Automatica*, 41(4):605–616, 2005.

[10] D.G. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16(6):596–602, December 1971.

[11] C. Mahulea, A. Ramírez, L. Recalde, and M. Silva. Steady state control reference and token conservation laws in continuous Petri net systems. *IEEE Trans. on Autom. Science and Engineering*, 5(2):307–320, April 2008.

[12] C. Mahulea, L. Recalde, and M. Silva. Optimal observability for continuous Petri nets. In 16^{th} *IFAC World Congress*, Prague, Czech Republic, July 2005.

[13] C. Mahulea, L. Recalde, and M. Silva. Basic Server Semantics and Performance Monotonicity of Continuous Petri Nets. *Discrete Event Dynamic Systems: Theory and Applications*, 19(2):189–212, June 2009.

[14] C. Mahulea, L. Recalde, and M. Silva. Observability of Timed Continuous Petri Nets: A Class of Hybrid Systems. In 17^{th} *IFAC World Congress*, Seoul, Korea, July 2008.

[15] K. Ogata. *Discrete-Time Control Systems, 2nd. ed.* Prentice Hall, 1995.

[16] L. Recalde, C. Mahulea, and M. Silva. Improving analysis and simulation of continuous Petri nets. In 2^{nd} *IEEE Conference on Automation Science and Engineering*, pages 7–12, Shanghai, China, October 2006.

[17] M. Silva, E. Teruel, and J. M. Colom. Linear algebraic and linear programming techniques for the analysis of net systems. In G. Rozenberg and W. Reisig, editors, *Lectures in Petri Nets. I: Basic Models*, volume 1491 of *Lecture Notes in Computer Science*, pages 309–373. Springer, 1998.

[18] P. K. Sinha. *Multivariable Control: An Introduction.* Marcel Dekker, 1984.

[19] R. Vidal, A. Chiuso, S. Soatto, and S. Sastry. Observability of linear hybrid systems. In O. Maler and A. Pnueli, editors, *Hybrid Systems: Computation and Control: 6th International Workshop, HSCC 2003*, volume 2623 of *Lecture Notes in Computer Science*, pages 526–539, Prague, Czech Republic, 2003. Springer Berlin / Heidelberg.

Appendix: Observability: basic concepts

An *unforced* (i.e., without control inputs) time invariant linear system is expressed by the following equations:

$$\begin{cases} \dot{\boldsymbol{x}}(\tau) = \boldsymbol{A} \cdot \boldsymbol{x}(\tau) \\ \boldsymbol{y}(\tau) = \boldsymbol{S} \cdot \boldsymbol{x}(\tau) \end{cases} \tag{15}$$

where $\boldsymbol{x}(\tau)$ is the state of the system and $\boldsymbol{y}(\tau)$ is the output, i.e., the set of measured variables. Knowing matrices $\boldsymbol{A}$ and $\boldsymbol{S}$, and being able to watch the evolution of $\boldsymbol{y}(\tau)$, a linear system is said to be *observable* if it is always possible to compute its initial state, $\boldsymbol{x}(\tau_0)$ (in fact, since the system is deterministic, knowing the state at the initial time is equivalent to knowing the state at any time).

In Systems Theory a very well-known observability criterion exists which allows to decide whether a continuous time invariant linear system is observable or not. Besides, several approaches exist to compute the initial state of a continuous time linear system that is observable.

Given an unforced linear system (15), the output of the system and the *observability matrix* are:

$$\boldsymbol{y}(\tau) = \boldsymbol{S} \cdot e^{\boldsymbol{A} \cdot \tau} \cdot \boldsymbol{x}(\tau_0) \tag{16}$$

$$\vartheta = \left[\begin{array}{cccc} \boldsymbol{S}^T, & (\boldsymbol{S}\boldsymbol{A})^T, & \cdots & ,(\boldsymbol{S}\boldsymbol{A}^{n-1})^T \end{array} \right]^T \tag{17}$$

Proposition 36 *[10, 15] Eq.* (16) *is solvable* $\forall \boldsymbol{x}(\tau_0), \forall \tau > 0$ *iff the observability matrix* ϑ *has full rank (rank*$(\vartheta) = n$*).*

The initial state can be obtained solving the following system of equations that has a unique solution under the rank condition:

$$\begin{bmatrix} \boldsymbol{y}(0) \\ \frac{d}{dt}\boldsymbol{y}(0) \\ \frac{d^2}{dt^2}\boldsymbol{y}(0) \\ \vdots \\ \frac{d^{n-1}}{dt^{n-1}}\boldsymbol{y}(0) \end{bmatrix} = \vartheta \cdot \boldsymbol{x}(0) \tag{18}$$

An interpretation of complete observability is that there is no simplification in the transfer function between the (actions on) state variables and the output [18]. Considering a single-output system, the transfer functions vector between the state variables and the output is given by:

$$\mathcal{Y}(s) = \boldsymbol{S}(s\boldsymbol{I} - \boldsymbol{A})^{-1} = \frac{1}{\Delta(s)}[q_1(s) \ldots q_n(s)] \tag{19}$$

If $\mathcal{Y}$(s) has a cancelation (all the polynomials $q_i(s)$ and $\Delta(s)$) have a common factor) this canceled mode cannot be observed in the output y.