

# Approximate Throughput Computation of Stochastic Weighted T-systems <sup>1</sup>

Carlos J. Pérez-Jiménez, Javier Campos and Manuel Silva

Dpto. Informática e Ingeniería de Sistemas, Universidad de Zaragoza

María de Luna 1, E-50018 Zaragoza, Spain

e-mail: cjperez, jcampos, silva @unizar.es.

**Abstract:** A general iterative technique for approximate throughput computation of stochastic live and bounded weighted T-systems is presented. It generalizes a previous technique on stochastic marked graphs. The approach has two basic foundations. First, a deep understanding of the qualitative behaviour of weighted T-systems leads to a general decomposition technique. Second, after the decomposition phase, an iterative response time approximation method is applied for the throughput computation. Existence of convergence points for the iterative approximation method can be proved. Experimental results generally have an error of less than 5 %. The state space is usually reduced by more than one order of magnitude; therefore, the analysis of otherwise intractable systems is possible.

**Keywords:** Petri nets, Stochastic systems, Approximation methods, Structural decomposition, Performance evaluation

## 1 Introduction

*Weighted T-systems* (WTS) [25] are the weighted generalization of marked graphs (MG's). *Stochastic Weighted T-systems* (SWTS's) are a subclass of *stochastic Petri net* (SPN) models [3]. SWTS's allow the modelling of concurrency, synchronization and bulk movements of jobs but not decisions. SWTS's can be used for the modelling of assembly–disassembly

---

<sup>1</sup>This work has been developed within the projects TIC2003-05226 and DPI2003-06376 of the Spanish science and technology ministry.

systems. These systems can assemble several pieces to produce more complex ones or disassemble pieces in several simpler ones.

In this paper, we consider live and bounded SWTS's with time and marking independent *exponentially distributed* service times associated with transitions. Liveness is a necessary condition for non null throughput of transitions (so the throughput approximation makes sense), and boundedness is necessary (and sufficient for live SWTS) to obtain finite continuous time Markov chains (CTMC's). A live and bounded SWTS is strongly connected, consistent and reversible [24] so its isomorphic CTMC is ergodic [3] and hence has a unique *steady-state probability distribution*. For this class of models, several computation techniques have been presented in the literature. Exact performance results can be obtained from the numerical solution of the isomorphic CTMC [3], but the *state explosion problem* makes intractable the evaluation of large systems. The efficient computation of exact performance indices of SWTS's cannot be done analytically because *local balance property* does not hold in general [15]. The alternative approach of *bounds computation* has been studied by several authors using different techniques (see, e. g., [6, 8]).

Concerning approximation techniques, several proposals have been done. In [4], a method is proposed for nets that admit a *time scale decomposition* based on *near-complete decomposability* of Markov chains. Near decomposability properties are also used in [11] for an iterative approximate solution of weakly connected nets. In [7], some particular queueing networks with subnetworks having *population constraints* are analyzed using *flow equivalent aggregation* (i.e., a non-iterative technique) and Marie's method [19]. In [9], given an arbitrary cut (subset of places producing a net partition), a *structural decomposition* technique is developed that allows to split a strongly connected MG in two subnets. With the subnets

three *aggregated systems* are computed; two *low level systems* and a *basic skeleton* or *high level system*. Each low level system is composed by one subnet and an aggregation of the other one. The basic skeleton is composed by the aggregation of the two subnets. Each aggregated subnet summarizes the behaviour of one subnet. With the aggregated systems, by means of an iterative response time approximation algorithm, the throughput of the transitions of the MG is approximated. The structural decomposition and iterative algorithm presented in [9] is included in TimeNet tool [27]. In [26], an iterative approximation algorithm is developed for some subclass of stochastic process algebras (SPA's) and the TIPP tool [16] includes an algorithm for response time approximation for another subclass of SPA's.

In this paper we *extend* the technique in [9] in several ways: (1) The approximation technique do apply to the class of SWTS's; (2) it will be possible to decompose the original net in an arbitrary finite number of subnets (not only two); (3) the structure of the aggregated systems is reduced (in terms of number of nodes) directly by the decomposition technique; (4) and a formal proof for the existence of convergence points for the iterative algorithm is included (proof also valid for the subclass considered in [9]). A first incomplete approach to the technique was presented in [21]. This paper improves it by the reduction of the aggregated systems structure and the proof of existence of convergence points for the numerical approximation algorithm.

In the case of MG's the aggregated systems were exact projections (in terms of reachable markings and firing sequences projected on the preserved nodes) of the original net. This is not the case with WTS's. With the intend of reducing the structure of the aggregated systems, they preserve a weaker set of structural properties (liveness and boundedness included) but it is sufficient to obtain good throughput approximations. Experimental results

on several examples show fast convergence (3 to 5 iterations) with a final error of less than 5%. The state space is usually reduced by more than one order of magnitude.

The paper is organized as follows. In section 2, basic notations and fundamental properties on WTS's and implicit places are presented. In section 3, the technique presented in [9] is exposed by means of a simple example. Also the problems to extend the technique to the class of SWTS's will be arisen. Section 4 will be devoted to the study of structural properties of WTS's (gain, weighted marking and resistance). Section 5 includes the structural decomposition of WTS's used in the rest of the paper. The iterative technique to compute the throughput approximation is described in section 6. Section 7 includes some examples to illustrate the introduced technique. Finally, concluding remarks are presented in section 8.

## 2 Basics notations

We assume the reader is familiar with concepts of P/T nets. Here we only present notations used in later sections. For further extensions the reader is referred to [20, 14].

A P/T net is a 4-tuple (graph oriented definition)  $\mathcal{N} = \langle P, T, F, W \rangle$ , where  $P$  and  $T$  are disjoint non empty sets of *places* and *transitions* ( $|P| = n$ ,  $|T| = m$ ),  $F \subseteq (P \times T) \cup (T \times P)$  is the set of directed arcs and  $W : F \rightarrow \mathbb{N}$  assigns a natural weight to each arc.  $P \cup T$  is the set of *nodes*. An alternative matrix oriented definition of P/T nets is a 4-tuple  $\mathcal{N} = \langle P, T, \mathbf{Pre}, \mathbf{Post} \rangle$  with  $P, T$  the same as in the previous definition and  $\mathbf{Pre}$  ( $\mathbf{Post}$ ) the pre- (post-) incidence non negative integer matrix of size  $|P| \times |T|$ . *Ordinary* nets are P/T nets whose arcs have weight 1. The *pre-* and *post-set* of a node  $v \in P \cup T$  are defined respectively as  $\bullet v = \{u \in P \cup T | (u, v) \in F\}$  and  $v \bullet = \{u \in P \cup T | (v, u) \in F\}$ . The *incidence matrix* of the P/T net is defined as  $\mathbf{C} = \mathbf{Post} - \mathbf{Pre}$ .

A vector  $\mathbf{m} \in \{\mathbb{Z}^+\}^{|P|}$  which assigns a non negative integer to each place is called *marking*. A *P/T system*, or *marked Petri net*  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$ , is a P/T net  $\mathcal{N}$  with an *initial marking*  $\mathbf{m}_0$ . A transition  $t \in T$  is *enabled* at marking  $\mathbf{m}$  if  $\mathbf{m} \geq \mathbf{Pre}[P, t]$ . A transition  $t$  enabled at  $\mathbf{m}$  can *fire* yielding a new marking  $\mathbf{m}' = \mathbf{m} + \mathbf{C}[P, t]$  (*reached marking*). It is denoted by  $\mathbf{m} \xrightarrow{t} \mathbf{m}'$ . A sequence of transitions  $\tau = \{t_i\}_{i=1}^n$  is a *firing sequence* in  $\mathcal{S}$  if there exists a sequence of markings such that  $\mathbf{m}_0 \xrightarrow{t_1} \mathbf{m}_1 \xrightarrow{t_2} \mathbf{m}_2 \dots \xrightarrow{t_n} \mathbf{m}_n$ . In this case, marking  $\mathbf{m}_n$  is said to be *reachable* from  $\mathbf{m}_0$  by firing  $\tau$ , and this is denoted by  $\mathbf{m}_0 \xrightarrow{\tau} \mathbf{m}_n$ . The *firing count vector*  $\sigma$  of the firable sequence  $\tau$  is a vector such that  $\sigma[t]$  represents the number of occurrences of  $t \in T$  in  $\tau$ . If  $\mathbf{m}_0 \xrightarrow{\tau} \mathbf{m}$ , then we can write in vector form  $\mathbf{m} = \mathbf{m}_0 + \mathbf{C} \cdot \sigma$ , which is referred to as the *linear state equation* of the net. The set  $L(\mathcal{S}) = \{\tau \mid \tau \text{ firable from } \mathbf{m}_0\}$  is the *language of firing sequences* of  $\mathcal{S}$ .  $\text{RS}(\mathcal{S})$  is the set of all reachable markings from  $\mathbf{m}_0$ .  $\text{RG}(\mathcal{S})$  is the reachability graph of  $\mathcal{S}$  (a graph with  $\text{RS}(\mathcal{S})$  as set of vertices and whose set of edges are the firing sequences of length 1 between vertices).

A P/T system is *live* when every transition can ultimately occur from every reachable marking. A place  $p \in P$  is said to be *k-bounded* if  $\mathbf{m}[p] \leq k$  for all  $\mathbf{m} \in \text{RS}(\mathcal{S})$ . A P/T system is said to be (marking) *k-bounded* if every place is *k-bounded*, and *bounded* if there exists some  $k$  for which it is *k-bounded*. A marking  $\mathbf{m}$  is a *home state* in  $\mathcal{S}$  if it is reachable from every reachable marking and  $\mathcal{S}$  is *reversible* if  $\mathbf{m}_0$  is a home state.  $\mathcal{S}$  is *deadlock-free* when at least one transition is enabled at every reachable marking.

A *path* in  $\mathcal{N}$  is a sequence  $\{x_1, \dots, x_n\}$  of nodes such that  $(x_i, x_{i+1}) \in F$  for  $1 \leq i < n$ . A *circuit* is a path such that  $(x_n, x_1) \in F$ . A path (circuit) is called *simple* if all its nodes are different. We denote by  $\mathcal{P}(x, y)$ , the set of simple paths from  $x$  to  $y$ .  $\mathcal{N}$  is *strongly connected* if for every two nodes  $x, y$  of  $\mathcal{N}$  there exist a path from  $x$  to  $y$ .

A *p-semiflow* is a vector  $\mathbf{y} \geq \mathbf{0}$ ,  $\mathbf{y} \neq \mathbf{0}$  such that  $\mathbf{y} \cdot \mathbf{C} = \mathbf{0}$ . A *t-semiflow* is a vector  $\mathbf{x} \geq \mathbf{0}$ ,  $\mathbf{x} \neq \mathbf{0}$  such that  $\mathbf{C} \cdot \mathbf{x} = \mathbf{0}$ . A net is *consistent* if it has a t-semiflow  $\mathbf{x} \geq \mathbf{1}$ . A net is *conservative* if it has a p-semiflow  $\mathbf{y} \geq \mathbf{1}$ .

A *weighted T-system* (WTS) or weighted marked graph is a P/T net such that each place has exactly one input transition and exactly one output transition. From a queueing network perspective, WTS's are a mild generalization of *Fork-Join Queueing Networks with Blocking* where bulk movements of jobs are allowed. Even if some results for WTS's are essentially parallel to those for the ordinary (non-weighted) case [25], there are interesting differences that play an important role in the decomposition of WTS models.

### 3 Overview of the approximation technique

In this section we are going to expose the technique in [9] for throughput approximation in MG's by means of a simple example. Then we will see the problems to extend the technique to the class of WTS's.

In Fig. 1.a an MG is depicted. The approximation technique in [9] is based in a structural decomposition of the original net. Given an arbitrary *cut* (subset of places producing a net partition), the original net is split in two subnets. In Fig. 1 the cut is composed by places  $B_1$  and  $B_2$ . With the subnets three aggregated systems are computed; two low level systems ( $\mathcal{LS}_1$  and  $\mathcal{LS}_2$  in Fig. 1.b and Fig. 1.c respectively) and a basic skeleton ( $\mathcal{BS}$ ) or high level system (Fig. 1.d). Each low level system is composed by one subnet and an aggregation of the other one. The  $\mathcal{BS}$  is composed by the aggregation of the two subnets. The aggregation of subnets is done by means of *implicit places*. Each aggregated subnet summarizes the behaviour of one subnet. After that, with the aggregated systems, by means of an iterative

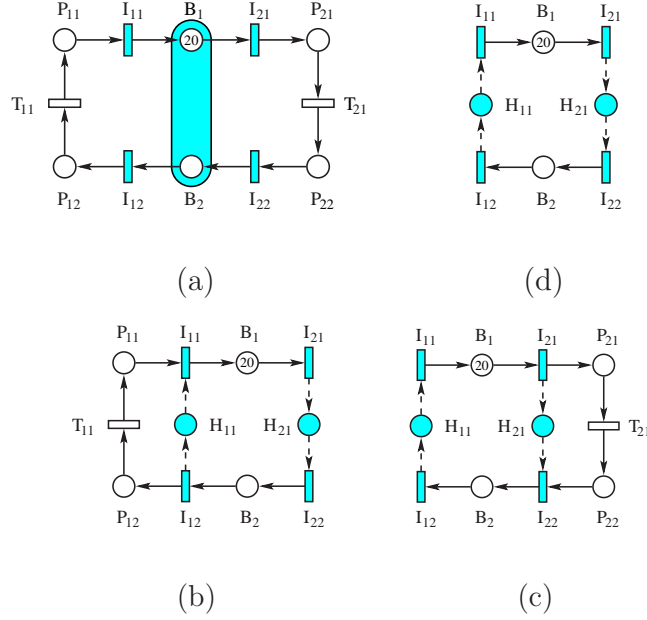


Figure 1: (a) An MG and its (b)  $\mathcal{LS}_1$ , (c)  $\mathcal{LS}_2$  and (d)  $\mathcal{BS}$ .

response time approximation algorithm, the throughput of the transitions of the MG is approximated. For example, in Fig. 1, assuming all transition rates equal to 1.0 with single server semantics, the exact steady-state throughput of all transitions on the original net is 0.8. The original MG has **53130** reachable states,  $\mathcal{LS}_1$  and  $\mathcal{LS}_2$  have **10626** and  $\mathcal{BS}$  **1771**. Applying the iterative algorithm of [9] to the aggregated systems, we obtain an approximation for the throughput of transitions equal to 0.804093 (+0.51% of relative error).

Now, we are going to include weights in the example. In Fig. 2.a a WTS is depicted. Given the same cut as in the previous example, the original net is split in two subnets. If we try to apply the technique presented in [9] to this WTS, we can use the same implicit places to “aggregate” the subnets, producing the low level systems  $\mathcal{LS}_1$  (Fig. 2.b) and  $\mathcal{LS}_2$  (Fig. 2.c) and the basic skeleton  $\mathcal{BS}$  (Fig. 2.d). The original WTS has **462** reachable states, but  $\mathcal{LS}_1$  and  $\mathcal{LS}_2$  have **10626**, and  $\mathcal{BS}$  has **1771**. So, we have not reduced at all the state space in the “aggregated” systems. Instead of that, we have increased their sizes. Obviously,

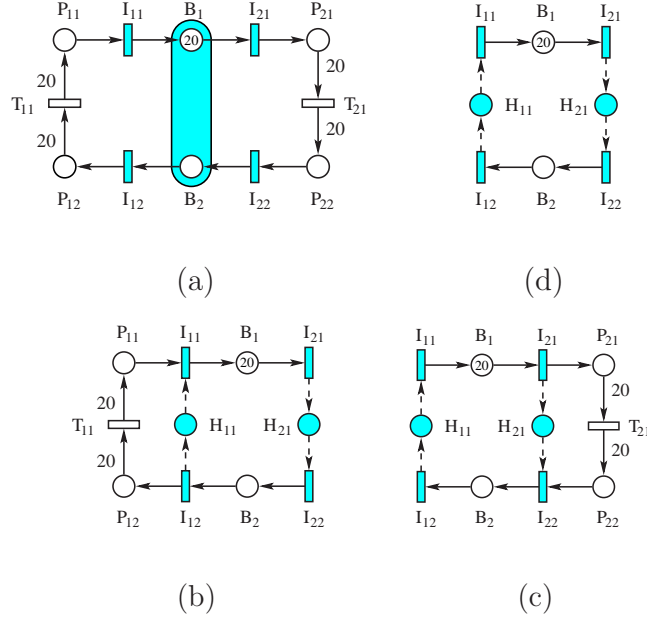


Figure 2: (a) A WTS and its (b)  $\mathcal{LS}_1$ , (c)  $\mathcal{LS}_2$  and (d)  $\mathcal{BS}$ .

with “aggregated” systems bigger than the original one, no good approximations can be expected. Assuming all transition rates equal to 1.0 with single server semantics, the exact steady-state throughput of all shadow transitions on the original net is 0.384396. Applying the iterative algorithm of [9] to the “aggregated” systems, we obtain an approximation for the throughput of transitions equal to 0.205574, in other words a  $-46.52\%$  of relative error (unacceptable for an approximation technique). This example shows that it is not trivial to directly apply the technique presented in [9] for MG’s to WTS’s.

Which is the problem? In our opinion, the problem deals with the deficient treatment of the arc weights. The weights of the arcs induce synchronizations on transitions. In Fig. 2.a the arcs with weight 20 induce a synchronization on the internal transitions (the white ones). When a subnet is reduced by means of an implicit place these internal synchronizations disappear in the aggregated systems leading to aggregated systems with more reachable states than the original net and a poor approximation. To improve the approximations a



deeper understanding on the behaviour of WTS's is needed. The synchronizations induced by the weights of the arcs need to be taken into account.

## 4 Structural properties

The subsection 4.1 is devoted to the study of the structural properties of WTS's needed for the rest of the paper. In subsection 4.2 implicit places in WTS are presented.

### 4.1 Gain, weighted marking and resistance

To decompose a WTS some basic properties on the behaviour of WTS's must be known. These properties are defined on paths of the WTS's connecting transitions; they are the *gain*, *weighted marking* and *resistance*. Gain and weighted marking must be preserved in order to obtain live and bounded aggregated systems. Resistance is proposed to improve the numerical results of the approximate analysis technique overviewed in section 3.

The gain was introduced in [25] for weighted circuits. In a live and bounded WTS any weighted circuit has gain equal to 1 (i.e. it is *neutral*). In this paper, the concept of gain is extended to weighted paths connecting transitions. The gain of a weighted path represents the mean firing ratio between the last transition and the first one.

**Definition 1 Gain of a path.** Let  $\langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a WTS,  $\pi = (t_0 p_1 t_1 p_2 \dots p_n t_n)$  a path in  $\mathcal{N}$  from transition  $t_0$  to  $t_n$ . Let  $r_i = W(t_{i-1}, p_i)$ ,  $s_i = W(p_i, t_i)$  with  $1 \leq i \leq n$ . The gain of  $\pi$  is  $G(\pi) = \prod_{i=1}^n \frac{r_i}{s_i} \in \mathbb{Q}^+ \setminus \{0\}$ .

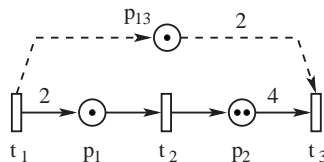


Figure 3: Path with  $G = 1/2$ ,  $M = 3/2$  and its reduction preserving  $G$  and  $M$ .

The gain of a path (see Fig. 3) depends only on the weights of its arcs. In an MG's any path has gain equal to 1. From Def. 1 immediately follows that if a path  $\pi_3 = \pi_1 \circ \pi_2$  is the concatenation of two given paths, then  $G(\pi_3) = G(\pi_1) \cdot G(\pi_2)$ .

To reduce subnets in MG's the sum of tokens in a path played an important role (see [9]). A weighted marking is the natural extension to WTS's of the sum of tokens in MG's. The weighted marking  $M(\pi, \mathbf{m})$  of a path  $\pi$  at marking  $\mathbf{m}$  can be interpreted as the number of firings of the first transition needed to reach  $\mathbf{m}$  in  $\pi$  being initially empty (if rational tokens and rational number of firings were allowed).

**Definition 2 Weighted marking of a path.** *Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a WTS,  $\pi = (t_0 p_1 t_1 \dots p_n t_n)$  a path in  $\mathcal{N}$  from  $t_0$  to  $t_n$ . Let  $r_i = W(t_{i-1}, p_i)$ ,  $s_i = W(p_i, t_i)$  with  $1 \leq i \leq n$ . For each  $\mathbf{m} \in \text{RS}(\mathcal{S})$ , the weighted marking of  $\pi$  in  $\mathbf{m}$  is  $M(\pi, \mathbf{m}) = \sum_{i=1}^n \left( \frac{\mathbf{m}[p_i]}{r_i} \prod_{j=1}^{i-1} \frac{s_j}{r_j} \right) \in \mathbb{Q}^+$ .*

Weighted marking (see Fig. 3) depends not only on the structure (weights of the arcs) but also on the marking of places. In the case of MG's, the weighted marking of any path is the sum of tokens in its places. From Def. 1 and 2 immediately follows that if a path  $\pi_3 = \pi_1 \circ \pi_2$  is the concatenation of two given paths then  $M(\pi_3, \mathbf{m}) = M(\pi_1, \mathbf{m}) + \frac{M(\pi_2, \mathbf{m})}{G(\pi_1)}$ . From Def. 2 immediately follows that the weighted marking of a path  $\pi$  only changes if the initial or the final transition in the path fire ( $M(\pi, \mathbf{m})$  is increased by 1 if the initial transition is fired and is decreased by  $1/G(\pi)$  if the final transition is fired). For live and bounded WTS's we have:

**Theorem 3** *(Proof in App. 1) Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS and  $\pi$  a circuit in  $\mathcal{N}$ . Then  $G(\pi) = 1$  (see [25]) and  $M(\pi, \mathbf{m}) = M(\pi, \mathbf{m}_0) \geq 1$  for all  $\mathbf{m} \in \text{RS}(\mathcal{S})$ .*

This theorem establishes that the weighted marking in any circuit is constant (greater than or equal to 1) for any reachable marking.

**Theorem 4** (Proof in App. 1) Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS,  $t_i, t_o$  two transitions in  $\mathcal{N}$  and  $\pi_1, \pi_2$  two paths in  $\mathcal{N}$  from  $t_i$  to  $t_o$ . Then  $G(\pi_1) = G(\pi_2)$  and there exists  $q \in \mathbb{Q}$  such that  $\forall \mathbf{m} \in \text{RS}(\mathcal{S}), M(\pi_1, \mathbf{m}) - M(\pi_2, \mathbf{m}) = q$ .

This theorem shows that in a live and bounded WTS, all paths connecting the same ordered pair of transitions must have the same gain and the difference of their weighted markings is the same in all reachable markings. Moreover, the path with greater weighted marking has always “trapped tokens” due to the synchronizations inside the path. As any live and bounded system is strongly connected, the concept of gain between two transitions is well defined by the previous theorem.

**Definition 5 Gain between transitions.** Let  $\mathcal{S}$  be a live and bounded WTS and  $t_i, t_o$  two transitions. The gain from  $t_i$  to  $t_o$ , denoted by  $G(t_i, t_o)$ , is the gain of any path from  $t_i$  to  $t_o$ .

Applying the relation between gain of concatenated paths, immediately follows that in any live and bounded WTS  $G(t_a, t_c) = G(t_a, t_b) \cdot G(t_b, t_c)$  for any  $t_a, t_b, t_c$  transitions in the WTS. Let  $\mathbf{g}_k = [G(t_k, t_1), \dots, G(t_k, t_n)]$ , be the gain vector from  $t_k$ .

**Theorem 6** (Proof in App. 1) Let  $\langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS,  $\mathbf{C}$  its incidence matrix,  $t_k \in T$  and  $\mathbf{g}_k$  the gain vector from  $t_k$ . Then  $\mathbf{C} \cdot \mathbf{g}_k = 0$  (i. e.  $\mathbf{g}_k$  is a T-semiflow).

This theorem establishes that the gain vector is a T-semiflow. With the gain and the weighted marking it is possible to reduce subnets in a WTS preserving liveness and boundedness, so we could now extend the technique in [9] to live and bounded WTS’s. But simple examples (see example of Fig. 2 in section 3) show that the throughput approximations may be very poor (−46.52% error in that example). To solve this problem we propose to take into

account the synchronizations of tokens due to the weights of the arcs. These synchronizations increase the response time of the subnets in WTS's. For example, in Fig. 2.a, we have paths between shadow transitions in the same subnet with gain equal to 1, but the weights induce a synchronization in the middle transition. It is needed to fire twenty times the first transition in order to fire twenty times the final transition. Then, liveness and boundedness preservation is not enough to obtain good approximations.

We introduce another concept, the *resistance* of a path, to take into account its internal synchronizations. The resistance of a path is independent of the gain and more complex to introduce. First we will study the case of an isolated path (concept of *resistance*) and later the case of several paths between transitions (concept of *behavioural resistance*). The resistance of a path is a structural concept representing the mean number of firings of the first transition needed to fire the last transition assuming the path initially empty.

**Definition 7 Resistance of a path.** Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a WTS,  $\pi = (t_0 p_1 t_1 p_2 \dots p_n t_n)$  a path in  $\mathcal{N}$  from  $t_0$  to  $t_n$ . Let  $r_i = W(t_{i-1}, p_i)$ ,  $s_i = W(p_i, t_i)$  with  $1 \leq i \leq n$ . The resistance of  $\pi$  is:  $RE(\pi) = \max_{1 \leq j \leq n} \{ \prod_{i=1}^j \frac{s_i}{r_i} \} \in \mathbb{Q}^+ \setminus \{0\}$ .

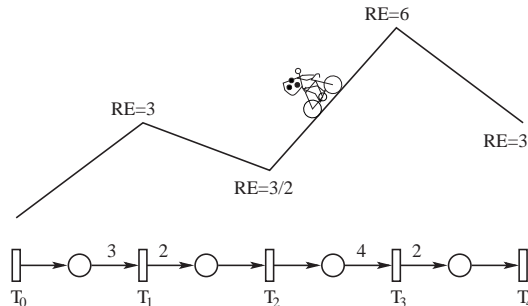


Figure 4: Path with resistance 6.

The resistance of a path only depends on the weights of the arcs and it is a positive

rational number. The resistance can be seen as the effort of a cyclist in a mountain stage in the “French Tour”. A place with output weight bigger than the input one supposes an extra resistance to the path because more firings of the initial transitions are needed to fire the final one. For instance, the resistance of the path in Fig. 4 is 6 (the maximal difficulty in the path is reached in  $T_3$ , so on this case  $RE(\pi) = \max_{1 \leq j \leq 4} \{\prod_{i=1}^j \frac{s_i}{r_i}\} = \prod_{i=1}^3 \frac{s_i}{r_i}$ ). Given a path  $\pi$  from  $t_0$  to  $t_n$ , there exists a transition  $t_i$  such that  $RE(\pi) = \prod_{i=1}^j s_i/r_i$  ( $T_3$  in Fig. 4), so the subpath  $\pi_i$  from  $t_0$  to  $t_i$  has  $RE(\pi_i) = RE(\pi) = 1/G(\pi_i)$ . From Def. 7 and Def. 1 immediately follows that if a path  $\pi_3 = \pi_1 \circ \pi_2$  then  $RE(\pi_3) = \max \left\{ RE(\pi_1), \frac{RE(\pi_2)}{G(\pi_1)} \right\}$ .

**Proposition 8** (*Proof in App. 1*) *Let  $\langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS. Let  $\pi$  be a circuit in  $\mathcal{N}$ . Then  $M(\pi, \mathbf{m}_0) \geq RE(\pi)$ .*

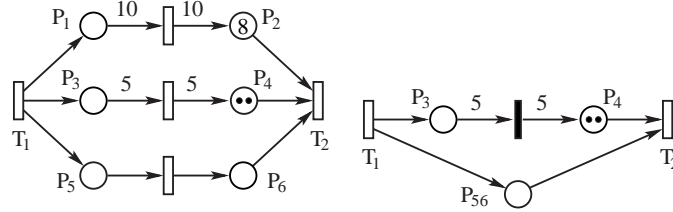


Figure 5: (a) Behavioural resistance between transitions and (b) its reduction.

In a WTS several paths joining two transitions may appear. For example, in Fig. 5.a, there are 3 paths from  $T_1$  to  $T_2$ . If the WTS is live and bounded, all these paths must have the same gain (1 in this case) but they can have different resistance (10; 5 and 1 in this case). The weighted marking of the paths are 8; 2 and 0. By Th. 4, the path with minimal weighted marking (the third one) traps 8 tokens on the first path and 2 tokens on the second one. The trapped tokens in a path decrease its “behavioural resistance” (with respect to its resistance) to fire the last transition. The trapped tokens can be seen as a permanent provisioning in race mitigating the effort of the cyclist (see Fig. 4). So, the trapped tokens

must be taken into account in order to define the behavioural resistance. The behavioural resistance of a path is defined to be the resistance of the path (i. e. the resistance of the isolated path) decreased by the trapped weighted marking of the path (quantity depending on the rest of the WTS). The formal definition of the behavioural resistance will be done by means of an algorithm in the section 5. In the case of Fig. 5.a, the first path has behavioural resistance  $10 - 8 = 2$ , the second one  $5 - 2 = 3$  and the third one  $1 - 0 = 1$ . So the path with maximum behavioural resistance is the second one. In Fig. 5.b the proposed reduction of Fig. 5.a is depicted (this reduction will be exposed in section 5). The first path is one with maximum behavioural resistance (it is not unique in general) and the second one has the minimal weighted marking (it is not unique in general), to preserve the possible trapped tokens in the first one (in this way the behavioural resistance of the first path is maintained). If we changed the initial marking of Fig. 5.a, deleting the tokens in  $P_2$  and  $P_4$ , then the behavioural resistance of the paths would be  $10 - 0 = 10$ ;  $5 - 0 = 5$  and  $1 - 0 = 1$  so in this case the first path would be the one with maximum behavioural resistance.

## 4.2 Implicit places and WTS's

An *implicit place* never is the unique restricting the firing of its output transitions. Let  $\mathcal{N}$  be any net and  $\mathcal{N}^p$  the net resulting from adding a place  $p$  to  $\mathcal{N}$ . If  $\mathbf{m}_0$  is an initial marking of  $\mathcal{N}$ ,  $\mathbf{m}_0^p$  denotes the initial marking of  $\mathcal{N}^p$  and  $\mathbf{m}_0[p] = \mathbf{m}_0^p[p]$ . The incidence vector of place  $p$  is denoted as  $\mathbf{C}_p$ .

**Definition 9 [22] Implicit place (IP).** *Let  $\mathcal{S}$  a P/T system and  $p$  a place to be added (provided with  $\mathbf{m}_0[p]$  tokens). Then  $p$  is an implicit place (IP) with respect to  $\mathcal{S}$  (or equivalently, it is an implicit place in  $\mathcal{S}^p$ ) if  $L(\mathcal{S}^p) = L(\mathcal{S})$ .*

A place is an IP depending on the initial marking  $\mathbf{m}_0$ . Places which can be implicit for any  $\mathbf{m}_0$  are said to be *structurally implicit* (SIP). Inside the class of SIP's we are interested in the so called *marking structurally implicit places* (MSIP) whose structural characterization is given in the following result.

**Theorem 10** [22, 12] **Marking structurally implicit place (MSIP).** *Let  $\mathcal{N}$  a P/T net and  $p \notin P$  a place with incidence vector  $\mathbf{C}_p$ . The place  $p$  is an MSIP in  $\mathcal{N}^p$  if and only if there exists  $\mathbf{y} \geq \mathbf{0}$  such that  $\mathbf{C}_p = \mathbf{y} \cdot \mathbf{C}$ .*

From this characterization of an MSIP  $p$ , a method to compute an initial marking of  $p$  making it implicit with respect to  $\mathcal{S}$  is presented in [12]. In the following, we characterize a special class of MSIP's within the class of live and bounded WTS's called *TT-MSIP's* (transition to transition MSIP). These places have only one input arc and one output arc, so  $\mathcal{S}^p$  (the P/T system resulting from the addition of a MSIP  $p$  to  $\mathcal{S}$ ) is also a WTS. The incidence vector of a TT-MSIP is a linear combination of the incidence vectors of the places in any path from the input transition to the output transition of the place.

**Theorem 11** (*Proof in App. 1*) *Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS and  $p \notin P$  a place to be added with  $\bullet p = t$ ,  $p \bullet = t'$  and  $G(t, t') = \frac{g_1}{g_2}$ . If  $W(t, p) = g_1$  and  $W(p, t') = g_2$  then  $p$  is a TT-MSIP with respect to  $\mathcal{S}$ .*

The following result gives an initial marking of a TT-MSIP in order to make implicit the place. This marking is computed from the weighted marking of the existing paths from the input transition of  $p$  to its output transition.

**Theorem 12** (*Proof in App. 1*) *Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS and  $t, t'$*

two transitions such that  $G(t, t') = \frac{g_1}{g_2}$ . Let  $p \notin P$  be a TT-MSIP with  $W(t, p) = g_1$  and  $W(p, t') = g_2$ . If  $\mathbf{m}_0[p] \geq \mathbf{m}_0^{\min}(p) = \lfloor g_1 \cdot \min\{M(\pi, \mathbf{m}_0) \mid \pi \in \mathcal{P}(t, t')\} \rfloor$  then  $p$  is implicit.

The computation of the initial marking of a TT-MSIP needed to make the place implicit can be done with a slight modification of the *all shortest path Floyd algorithm* [2] that will be explained in the next section. Next theorem gives a sufficient condition to reduce the number of TT-MSIP's needed to resume subnets in a WTS.

**Theorem 13** (*Proof in App. 1*) Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS;  $t_i, t_j, t_k$  three transitions in  $\mathcal{N}$ , and  $p_{ik}, p_{kj}, p_{ij}$  three TT-MSIP's from  $t_i$  to  $t_k$ ,  $t_k$  to  $t_j$  and  $t_i$  to  $t_j$  respectively, with the initial marking function  $\mathbf{m}_0^{\min}$  for  $p_{ik}, p_{kj}, p_{ij}$  given in theorem 12. If there exists a minimal weighted marking path in  $\mathcal{N}$  from  $t_i$  to  $t_j$  through  $t_k$ , then  $p_{ij}$  is an implicit place with respect to the system generated by places  $p_{ik}, p_{kj}$ .

## 5 Structural decomposition of WTS's

The basic idea is the following: a live and bounded WTS is split into  $K$  subnets by a *cut*  $B$  defined through some places (see in Fig. 6.a a cut in  $K = 2$  subnets). From the cut and the subnets we define  $K + 1$  aggregated systems:  $K$  low level systems ( $\mathcal{LS}_1, \dots, \mathcal{LS}_K$ ; see Figs. 6.b and 6.c) and a high level system or *basic skeleton* ( $\mathcal{BS}$ ; see Fig. 6.d). These systems will be obtained by substitution (what we call “aggregation”) of the so called subnets, defined from the cut  $B$ , by a set of nodes. Each  $\mathcal{LS}_i$  has one subnet obtained from the cut and the rest subnets are aggregated. In the  $\mathcal{BS}$  all subnets are aggregated. The aggregation of subnets for WTS supposes an extension of that in [9] for MG in the sense that this aggregation applied to an MG produces exactly the same aggregated systems of [9].



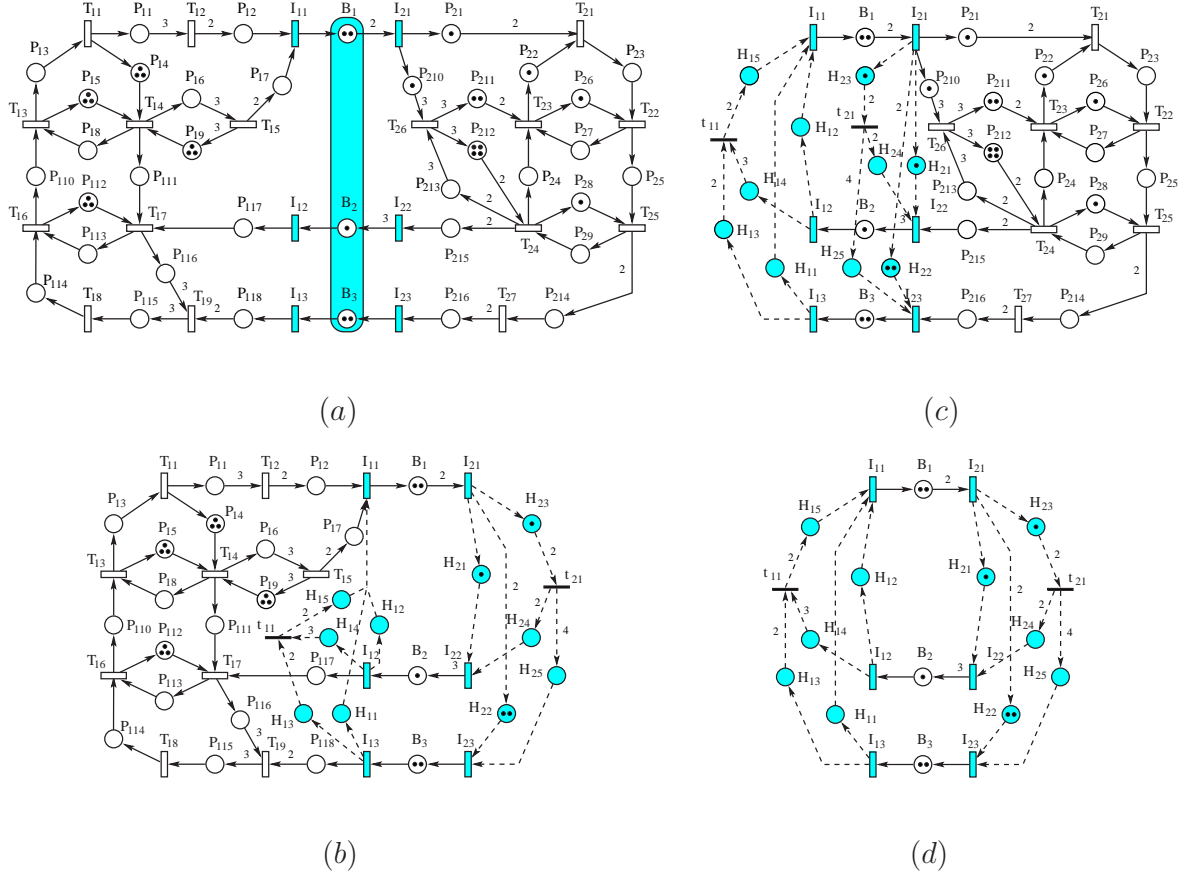


Figure 6: (a) A system and its (b)  $\mathcal{LS}_1$ , (c)  $\mathcal{LS}_2$  and (d)  $\mathcal{BS}$ .

**Definition 14** Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a WTS with  $\mathcal{N} = \langle P, T, F, W \rangle$ . A subset of places  $B \subseteq P$  is said to be a  $K$ -cut ( $K \geq 2$ ) of  $\mathcal{N}$  if there exists  $K$  subnets  $\mathcal{N}_i = \langle P_i, T_i, F_i, W_i \rangle$  with  $1 \leq i \leq K$  of  $\mathcal{N}$  such that:

- i)  $\bigcup_{i=1}^K T_i = T$  and  $T_i \cap T_j = \emptyset$  for  $1 \leq i, j \leq K, i \neq j$ .
- ii)  $P_i = \bullet T_i \cup T_i \bullet$  for  $1 \leq i \leq K$ .
- iii)  $\bigcup_{i=1}^K P_i = P$  and  $\bigcup_{i \neq j} (P_i \cap P_j) = B$  for  $1 \leq i, j \leq K$ .
- iv)  $F_i = F \cap ((P_i \times T_i) \cup (T_i \times P_i))$  and  $W_i = W|_{F_i}$  for  $1 \leq i \leq K$ .

The transitions in  $\text{TI} = \bullet B \cup B \bullet$  are the interface transitions. The other ones are the internal transitions.

Depending on the selection of the cut, different levels of aggregation may be obtained. So, the efficiency of the entire technique is directly affected by the cut selection. It is beyond the scope of this paper to develop a detailed analysis on the cut selection procedure. But from the intensive experiments already achieved, the intuitive “cut the model into equal size pieces” idea (in terms of reachable markings) seems to work very well in practice.

Once a cut is defined on a WTS, the aggregated systems must be computed. To do that a set of nodes will be computed to reduce each subnet of the cut. For each subnet the interface transitions (shadow transitions in Fig.6) will be preserved (these transitions will resume the response time of the subnet). To reduce a subnet in a WTS it is necessary to preserve the gain between interface transitions (to obtain live and bounded aggregated systems) and the maximum behavioural resistance between interface transitions (to get “good” throughput approximations). To compute these gains and behavioural resistances we will translate the nets language to the graph language. The net structure of a WTS can be represented as a directed graph whose vertices are the transitions and the edges are the places joining them.

The first computation to be done is the gain vector  $\mathbf{g}_1$  (Th. 6) such that  $\mathbf{g}_1(t_i) = G(t_1, t_i)$ . To simplify notation in this section, we will denote  $\mathbf{g}_1$  as  $\mathbf{g}$ . To compute  $\mathbf{g}$  a classical *breadth first* search of the WTS (graph form) can be done [13]. We omit the complete exposition of the algorithm for conciseness. Any live and bounded WTS is strongly connected, so the search visits all vertices. Any new vertex is connected to a visited vertex by an edge whose gain can be computed from the arc weights. The algorithm begins the search on  $t_1$  so  $\mathbf{g}(t_i) = G(t_1, t_i)$  and  $G(t_i, t_j) = \mathbf{g}(t_j)/\mathbf{g}(t_i)$  for all  $t_i, t_j \in T$ .

The next computation to be done is the minimal weighted marking and the maximum behavioural resistance between transitions. In order to include information about the gain,

weighted marking and resistance in the graph, labels to the edges will be associated. Each label has three components. The first one represents the resistance (see Def. 7) of the edge, the second one represents the weighted marking of the edge and the third one the gain of the edge. From the labels of the edges, by applying simple operations, it is possible to compute the minimal weighted marking and the maximum behavioural resistance between any pair of transitions in the WTS. The operations are the following.

**Definition 15** *Let be  $S = (\mathbb{Q}^+ \times \mathbb{Q}^+ \times \mathbb{Q}^+ \setminus \{0\}) \cup \bar{0}$  ( $\bar{0}$  is a special element), and  $\bar{1} = (0, 0, 1)$ .*

*The following operators are defined on  $S$ :*

*i) Extension operator  $\odot$ :  $(r_1, w_1, g_1) \odot (r_2, w_2, g_2) = (\max\{r_1, r_2/g_1\}, w_1 + w_2/g_1, g_1g_2)$ ;*

$$(r, w, g) \odot \bar{0} = \bar{0} \odot (r, w, g) = \bar{0} \odot \bar{0} = \bar{0}.$$

*ii) Summary operator  $\oplus$ :  $(r_1, w_1, g) \oplus (r_2, w_2, g) = (\max_{i=1,2}\{r_i - w_i\} + \min_{i=1,2}\{w_i\}, \min_{i=1,2}\{w_i\}, g)$ ;*

$$(r, w, g) \oplus \bar{0} = \bar{0} \oplus (r, w, g) = (r, w, g) \text{ and } \bar{0} \oplus \bar{0} = \bar{0}.$$

The set  $S$  is the domain of possible labels, the extension operator  $\odot$  computes the label of a path from the labels of subpaths and the summary operator  $\oplus$  selects between two paths joining the same two transitions the path with maximum behavioural resistance (and the path with minimal weighted marking in the case of equal behavioural resistance). In this way we have a recursive definition of the gain, behavioural resistance and weighted marking of a path.  $\bar{1}$  is the label of the empty path and  $\bar{0}$  is the label for the non-existent path.

To compute the paths with maximum behavioural resistance and the paths with minimal weighted marking we will follow the general framework for solving path problems in directed graphs (see [13], pp. [570–575] for the details). This framework is based on the algebraic structure of semirings. A *closed semiring* is a 5-tuple  $(S, \oplus, \odot, \bar{0}, \bar{1})$ , where  $S$  is a set of

elements,  $\oplus$  (summary operator) and  $\odot$  (extension operator) binary operations in  $S$  and  $\bar{0}$ ,  $\bar{1}$  elements of  $S$  satisfying:  $(S, \oplus, \bar{0})$  and  $(S, \odot, \bar{1})$  are monoids (i.e. they satisfy associative property and they have identity element);  $\bar{0}$  is an annihilator for  $\odot$ ;  $\oplus$  is commutative and idempotent;  $\odot$  distributes over  $\oplus$ ;  $\oplus$  is well defined for countable sequences of elements of  $S$ ; associativity, commutativity and idempotence apply to infinite summaries;  $\odot$  distributes over infinite summaries of elements of  $S$ . In the special case of WTS's the summary operator always applies to labels of paths with the same origin and destination, so with the same gain (th. 4). Thus, some conditions of closed semiring can be relaxed. A *relaxed closed semiring* is a closed semiring in which the summary operator  $\oplus$  only applies to labels of paths with the same origin and destination and the empty or non-existent path.

The next theorem gives us the proof of the existence of a dynamic programming algorithm to compute in polynomial time the path with minimal weighted marking and the path with maximum behavioural resistance between any pair of transitions in a WTS.

**Theorem 16** (*Proof in App. 2*) *Let  $S = (\mathbb{Q}^+ \times \mathbb{Q}^+ \times \mathbb{Q}^+ \setminus \{0\}) \cup \bar{0}$  be the domain of possible labels and  $\odot, \oplus$  the operators of Def. 15. Then  $\langle S, \oplus, \odot, \bar{0}, \bar{1} \rangle$  is a relaxed closed semiring.*

With this algebraic environment, following the general framework for solving path problems in directed graphs presented in [13], we can assure that there exists a dynamic programming algorithm to compute the path with maximum resistance joining any pair of vertices in a graph. This algorithm is nothing more than a generalization of the well-known Floyd's *all pairs shortest path* algorithm [2].

**Algorithm 17 Max. behavioural resistances and min. WM computation.**

**input:**  $G := (V, A)$ ,  $\mathbf{m}_0$  WTS (graph form) and initial marking; Gain vector  $\mathbf{g}$

**R** matrix of max. behav. resist. Init:  $\mathbf{R}(i, j) := W(p_{ij}, t_j)/W(t_i, p_{ij}) \Leftrightarrow \exists p_{ij} \in P, i \neq j$ ;  
 $\mathbf{R}(i, j) = 0$  otherwise.  
**L** matrix of min. weighted markings. Init:  $\mathbf{L}(i, j) := \mathbf{m}_0[p_{ij}]/W(t_i, p_{ij}) \Leftrightarrow \exists p_{ij} \in P, i \neq j$ ;  
 $\mathbf{L}(i, i) := 0; \mathbf{L}(i, j) = \infty \Leftrightarrow \nexists p_{ij} \in P$   
**K** = **0** matrix of maximum behavioural resistance paths vertices.  
**for**  $k := 1$  **to**  $n$  **do**  
  **for**  $i := 1$  **to**  $n, i \neq k$  **do**  
    **for**  $j := 1$  **to**  $n, j \neq k$  **do**  
       $l := \mathbf{L}(i, k) + \frac{\mathbf{L}(k, j)\mathbf{g}(i)}{\mathbf{g}(k)}$ ;  $m := \min\{\mathbf{L}(i, j), l\}$ ;  $r := \max\left\{\mathbf{R}(i, k), \frac{\mathbf{R}(k, j)\mathbf{g}(i)}{\mathbf{g}(k)}\right\} - l + m$   
       $\mathbf{R}(i, j) := \mathbf{R}(i, j) - \mathbf{L}(i, j) + m$   
      **if**  $m < \mathbf{L}(i, j)$  **then**  $\mathbf{L}(i, j) := m$   
      **if**  $r > \mathbf{R}(i, j)$  **then**  $\mathbf{R}(i, j) := r; \mathbf{K}(i, j) := k$   
    **end for**  
  **end for**  
**end for**  
**output:** Matrices **R**, **L**, **K**

This algorithm assumes that the  $n$  transitions are numbered beginning with the  $n'$  internal transitions (see Def. 14). Initially, the elements  $\mathbf{R}(i, i) = \mathbf{L}(i, i) = 0$  are the result of  $\bar{1} \oplus \lambda(i, i) = (0, 0, 1)$  (the empty path has label  $\bar{1} = (0, 0, 1)$  and  $\lambda(i, i) = (r, w, 1)$ ,  $w \geq r$  by prop. 8). The elements  $\mathbf{R}(i, j), \mathbf{L}(i, j)$  with  $i \neq j$  are initialized with the resistance and weighted marking of the edges of the graph. If  $(t_i, t_j) \notin A$  then we put  $\mathbf{R}(i, j) = 0$ ,  $\mathbf{L}(i, j) = \infty$  representing the minimal possible resistance and maximum possible weighted marking. The three nested loops compute, for any pair of transitions  $(t_i, t_j)$ , one path of maximum behavioural resistance from  $t_i$  to  $t_j$  (in **R** and **K**) and the minimal weighted marking of any path from  $t_i$  to  $t_j$  (in **L**). The external loop indicates the vertex  $k$  to explore. For the values  $i, j, k$  of the vertices, the algorithm compares the optimal paths already computed from  $i$  to  $j$  with a new path formed by the optimum path from  $i$  to  $k$  concatenated with the optimum path from  $k$  to  $j$ . In the case of the behavioural resistance,

the algorithm selects, according to the operator  $\oplus$ , the one with more behavioural resistance. In the case of the weighted marking, the algorithm selects the one with less weighted marking. This weighted marking is also the not trapped weighted marking of the paths from  $t_i$  to  $t_j$  (see Theorem 4). Then, after the iteration  $k$  of the external loop,  $\mathbf{R}(i, j)$  is the maximum behavioural resistance and  $\mathbf{L}(i, j)$  is the minimal weighted marking of the paths from  $t_i$  to  $t_j$  passing through vertices of index less than or equal to  $k$ . Thus, the output of the algorithm gives the maximum behavioural resistances ( $\mathbf{R}$ ) and minimal weighted markings ( $\mathbf{L}$ ) between any pair of transitions of the WTS. The matrix  $\mathbf{K}$  (see Floyd algorithm [2]) keeps track of the vertices of the optimal paths in terms of behavioural resistance. The complexity of this algorithm is  $O(n^3)$  in time and  $O(n^2)$  in memory ( $n = |T|$ ).

The reduction of the minimal weighted marking path from  $t_i$  to  $t_j$  ( $G(t_i, t_j) = p/q$ ) can be done (Th. 11 and 12) by means of a TT-MSIP  $p_{ij}$  with  $W(t_i, p_{ij}) = p$ ,  $W(p_{ij}, t_j) = q$  and  $\mathbf{m}_0[p_{ij}] = \lfloor p \cdot \mathbf{L}(i, j) \rfloor$  (see place  $p_{13}$  in Fig. 3 reducing the path from  $t_1$  to  $t_3$ ). Taking into account theorem 13, the number of TT-MSIP's needed to reduce the subnets can be reduced if a minimal weighted marking path between any pair of interface transitions passes through another interface transition. So, with an additional  $n \times n$  matrix  $\mathbf{B}$  in the last algorithm we can keep track if the optimum path between a given pair of transitions contains another interface transition. Moreover, the algorithm can be slightly modified to select among all the optimal paths between a given pair of transitions one of them passing through another interface transition in order to reduce the number of TT-MSIP's.

With respect to the behavioural resistance, given any pair  $(t_0, t_n)$  of transitions with several paths between them, only one with minimal weighted marking and only one with maximal behavioural resistance will be kept (see in Fig. 5 an example of reduction of sev-

eral paths). The path with maximum behavioural resistance can be reduced by means of a immediate transition (replica of the transition  $t$  where the path reaches its maximum resistance) and two TT-MSIP's (one from  $t_0$  to  $t$  and another one from  $t$  to  $t_n$ ). The path with minimal weighted marking is kept only to maintain the trapped tokens in the path with maximum behavioural resistance and it can be reduced to a TT-MSIP from  $t_0$  to  $t_n$ . Now, the aggregated systems can be computed.

**Definition 18** *Let  $\langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS,  $B \subseteq P$  a  $K$ -cut of  $\mathcal{N}$ . The extended system  $\mathcal{ES} = \langle \mathcal{EN}, \mathbf{m}_0^{\mathcal{EN}} \rangle$  is the net system obtained from  $\langle \mathcal{N}, \mathbf{m}_0 \rangle$  by adding the places and transitions needed to reduce the minimal weighted marking paths and the maximum behavioural resistance paths between interface transitions. The low level system  $\mathcal{LS}_i = \langle \mathcal{LN}_i, \mathbf{m}_0^i \rangle$ ,  $1 \leq i \leq K$  is the net system obtained from  $\mathcal{ES}$  by deleting the places of  $\bigcup_{j \neq i} P_j \setminus B$  and the transitions of  $\bigcup_{j \neq i} T_j \setminus \text{TI}$ . The high level system (or basic skeleton)  $\mathcal{BS} = \langle \mathcal{BN}, \mathbf{m}_0^{\mathcal{BN}} \rangle$  is the net system obtained from  $\mathcal{ES}$  by deleting the places of  $\bigcup_{i=1}^K P_i \setminus B$  and the transitions of  $\bigcup_{i=1}^K T_i \setminus \text{TI}$ .*

Each  $\mathcal{LS}_i$  maintains the subnet  $\mathcal{N}_i$  and reduces the other ones (see  $\mathcal{LS}_1$  and  $\mathcal{LS}_2$  in Fig. 6.b and 6.c). In the  $\mathcal{BS}$  all subnets are reduced (see Fig. 6.d). The important point is that proceeding in this way, projections of reachable markings and firing sequences of the original system are preserved in the aggregated systems, what seems good untimed properties to keep in order to obtain good approximations.

**Theorem 19** *(Proof in App. 1) Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS,  $B \subseteq P$  a  $K$ -cut of  $\mathcal{N}$ . The  $\mathcal{ES}$ ,  $\mathcal{LS}_i$ ,  $1 \leq i \leq K$  and  $\mathcal{BS}$  are live and bounded WTS;  $L(\mathcal{ES})|_T = L(\mathcal{S})$  and  $RS(\mathcal{ES})|_P = RS(\mathcal{S})$ .*

## 6 Iterative technique

### 6.1 The pelota algorithm

The technique for an approximate computation of the throughput that we present now is, basically, a *response time approximation* method [1, 9]. The interface transitions of  $\mathcal{N}_j$  in  $\mathcal{LS}_i$  ( $j \neq i$ ) approximate the response time of all the subsystem  $\mathcal{N}_j$ .

#### Algorithm 20 Pelota algorithm

**input:**  $\langle \mathcal{S}, w \rangle$  live and bounded stochastic WTS  
 select a  $K$ -cut  $B \subseteq P$  of  $\mathcal{N}$  and derive  $\mathcal{LS}_i, \mathcal{BS}$  for  $1 \leq i \leq K$   
 select initial service rates  $\mu_i^{(0)}$  for  $t \in T_j \cap \text{TI}$ ,  $2 \leq j \leq K$  (their rates in  $\langle \mathcal{S}, w \rangle$ )  
 $n := 0$  (counter for iteration steps)  
**repeat**  
    $n := n + 1$   
   **for**  $i := 1$  **to**  $K$  **do**  
     **solve**  $\mathcal{LS}_i$  with:  
       In: rates  $\mu_l^{(n)}$  of transitions  $T_l \cap \text{TI}$ ,  $1 \leq l < i$   
           rates  $\mu_l^{(n-1)}$  of transitions  $T_l \cap \text{TI}$ ,  $i < l \leq K$   
       Out: relative rates  $\mu_i$  and throughput  $\chi_i^{(n)}$  of transitions  $T_i \cap \text{TI}$   
     **solve**  $\mathcal{BS}$  (find scale factor  $\lambda$ ) with:  
       In: rates  $\mu_l^{(n)}$  of transitions  $T_l \cap \text{TI}$ ,  $1 \leq l < i$   
           rates  $\mu_l^{(n-1)}$  of transitions  $T_l \cap \text{TI}$ ,  $i < l \leq K$   
           relative rates  $\mu_i$  and throughput  $\chi_i^{(n)}$  of transitions  $T_i \cap \text{TI}$   
       Out: rates  $\mu_i^{(n)} = \lambda \cdot \mu_i$  of transitions  $T_i \cap \text{TI}$  such that  $\chi_{\mathcal{BS}}^{(n)} = \chi_i^{(n)}$   
     **end for**  
**until** convergence of  $\{\chi_i^{(n)}\}_{i=1}^K$   
**output:** throughput  $\{\chi_i^{(n)}\}_{i=1}^K$  of transitions  $\langle \mathcal{S}, w \rangle$

In live and bounded WTS's the relative throughput of transitions are determined by the structure of the net. Thus, with an approximation of the throughput of any transition we can obtain approximations for the other ones with the same percentage error. The proposed algorithm is basically the same that this in [9].



In the above procedure, once a  $K$ -cut has been selected and given some initial values  $\boldsymbol{\mu}_t^{(0)}$  for service rates of all interface transitions except those in  $\mathcal{N}_1$ , the isomorphic CTMC of aggregated subsystem  $\mathcal{LS}_1$  is solved. The selection of the initial values of interface transitions rates does not affect (under our experience) to the accuracy of the method. A simple option is putting the initial rate of the transitions in the model. From the solution of that CTMC, the first estimation  $\boldsymbol{\chi}_1^{(1)}$  of the throughput of transitions  $T_1 \cap \text{TI}$  can be computed. Then, the initial estimated values of service rates of interface transitions  $T_1 \cap \text{TI}$  must be derived. To do that, we take the initial values  $\boldsymbol{\mu}_t^{(0)}$  for service rates of transitions in  $T_1 \cap \text{TI}$  and we search in the  $\mathcal{BS}$  a *scale factor* for all these rates such that the throughput of  $\mathcal{BS}$  and the throughput of  $\mathcal{LS}_1$ , computed before, are equal. There is no problem with the different throughput of interface transitions in  $\mathcal{LS}_i$  and  $\mathcal{BS}$  because by construction the relative throughput of transitions in the aggregated systems are the same (theorem 6). The same procedure is executed for each  $\mathcal{LS}_i$  in a cyclic way. Each time we solve the  $\mathcal{LS}_i$  we obtain in the  $\mathcal{BS}$  a new estimation of the rates of  $T_i \cap \text{TI}$ . The previous steps are repeated until convergence is achieved (for instance, when the difference of the estimated throughput of transitions in two iterations is below 0.1%).

With regard to the computation of the ratios among interface transition rates in  $\mathcal{LS}_i$ , the mean enabling degree  $e_t$  of transition  $t \in T_i \cap \text{TI}$  can be computed (infinite server of transitions assumed). Then we put  $\boldsymbol{\mu}_i^{(n)}[t] = \boldsymbol{\chi}_i^{(n)}[t]/e_t$ . To compute the mean enable degree only the nodes to resume the subnet  $\mathcal{N}_i$  are taken into account. In this way, we compute a performance index related with the same interface transition in the  $\mathcal{BS}$ . For example, in Fig. 6.c, to compute the relative rate of interface transition  $I_{22}$  we compute the mean enabling degree of transition  $I_{22}$  taking into account only the implicit places  $H_{21}, H_{24}$ . So,

the relative rate of transition  $I_{22}$  is equal to (infinite server semantics)  $\frac{\mathcal{X}_{\mathcal{L}S_2}[I_{22}]}{E[\max\{M[H_{21}], M[H_{24}]\}]}$ .

For single server semantics, the mean enabling degree of a given transition is substituted by the probability to be enabled. For example, the relative rate of transition  $I_{22}$  is equal to (single server semantics)  $\frac{\mathcal{X}_{\mathcal{L}S_2}[I_{22}]}{P\{M[H_{21}]>0, M[H_{24}]>0\}}$ .

The computation of the scale factor in the  $\mathcal{BS}$  can be implemented with a linear search. Now the  $\mathcal{BS}$  has considerably fewer states than the original one. In each iteration of this linear search, the isomorphic CTMC of the  $\mathcal{BS}$  is solved. Note that only in the first iteration the CTMC is completely derived. For later iterations only some values must be changed.

## 6.2 About convergence of the method

The next step will be to prove the existence of convergence points for the numerical algorithm. To do that, we need some results given by other authors. In [18] continuity of the steady-state probability vector  $\boldsymbol{\pi}$  on the entries of the infinitesimal generator  $\mathbf{Q}$  of the CTMC of any SPN is proved. As the elements of  $\mathbf{Q}$  are continuous on the transition rates and many steady-state performance indices (such that throughput or mean marking of places) can be expressed as continuous functions on  $\boldsymbol{\pi}$ , continuity of throughput of any transition on the entries of  $\mathbf{Q}$  and on the transitions rates can be assured for any SPN. In [5] monotonicity of throughput transitions on transition rates and on the topology of the WTS is proved. Monotonicity on transition rates means that if we increase (decrease) the rate of any transition, the throughput of any other transition will not be decreased (increased). Monotonicity on the topology means that if we add new places and/or transitions to a WTS we do not increase the throughput of any transition. Finally, in [10] by means of a linear programming problem, an upper bound for the mean firing time of a transition in any live and bounded SPN is obtained. Taking

the inverse of the upper bound for the mean firing time of a transition we can obtain a lower bound for the throughput of the transition.

**Theorem 21** *(Proof in App. 1) The linear search on the  $\mathcal{BS}$  in Alg. 20 always converges.*

To prove the existence of convergence points for the algorithm 20, a functional form of the algorithm will be used. Let  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_K)$  be the vector of rates for interface transitions. In each iteration  $n$  of the algorithm, a new vector  $\boldsymbol{\lambda}^{(n)}$  is computed, so the algorithm can be expressed in the following form:

**Algorithm 22 Functional form for pelota algorithm.**

**input:**  $\boldsymbol{\lambda}^{(0)}$  initial rates for interface transitions  
 $n := 0$  (counter for iteration steps)  
**repeat**  
     $n := n + 1$   
     $\boldsymbol{\lambda}^{(n)} := A(\boldsymbol{\lambda}^{(n-1)})$   
**until** convergence of  $\boldsymbol{\lambda}^{(n)}$   
**output:** throughput  $\{\chi_i(\boldsymbol{\lambda}^{(n)})\}$  of transitions  $\langle \mathcal{S}, w \rangle$

**Theorem 23** *(Proof in App. 1) The function  $A : (\mathbb{R}^+)^M \longrightarrow (\mathbb{R}^+)^M$  ( $M = |IT|$ ) of algorithm 22 is continuous and there exist a non empty, convex and compact set  $S \subset (\mathbb{R}^+)^M$  such that  $A(S) \subseteq S$ . Then, there exists  $x \in S$  such that  $A(x) = x$ .*

This theorem establishes that the iterative algorithm has convergence points. To assure the convergence of the algorithm or the uniqueness of the solution, stronger properties of function  $A$  must be proved, contraction for example, but there is not success for the moment. With regard to the accuracy of the results, no formal proof gives positive answer to the question, but an extensive battery of numerical experiments has shown us that the algorithm converges in 3 to 5 iteration steps and the error is less than 5% in the worst cases.

## 7 Examples

In this section, we present several numerical results of the application of the iterative technique previously introduced. Among all the tested cases, we have selected the following two WTS's. The first one (already introduced in Fig. 6.a) is structurally asymmetric. Assuming all transition rates equal to 1.0 and single server semantics, the exact throughput of the reference transition  $I_{11}$  is 0.119353. The net has **319680** reachable states. The aggregated systems  $\mathcal{LS}_1$ ,  $\mathcal{LS}_2$  and  $\mathcal{BS}$  are depicted in Fig.6.(b), (c) and (d) respectively and they have **27362**, **10640** and **783** reachable states respectively.

In Table 1 we present the iteration steps of the method for this case. Columns  $\chi(I_{11})$  are the estimated values for throughput of transition  $I_{11}$  at each iteration step. Columns  $\mu_{ij}$  are the estimated relative rates of interface transition  $I_{ij}$  computed in  $\mathcal{LS}_i$  and columns  $\lambda$  are the scale factors modifying the previous relative rates, computed in the  $\mathcal{BS}$ . Convergence of the method is obtained in this case in 3 iteration steps. The error for this example was +1.32%.

As a second example, let us consider the WTS depicted in Fig. 7.a. Any splitting of the net will generate strongly coupled aggregated systems. This fact makes the system difficult to study and puts our method through a rigorous test.

We select the following cut:  $B = \{B_i\}_{i=1}^8$ . The corresponding  $\mathcal{ES}$  is depicted in Fig. 7.b. The low level systems are depicted in Figs. 8.a and 8.b. The  $\mathcal{BS}$  is that of Fig. 8.c. The

Service rates 1.0									
$\mathcal{LS}_1$					$\mathcal{LS}_2$				
$\mathcal{X}(I_{11})$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\lambda$	$\mathcal{X}(I_{21})$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\lambda$
0.135420	0.864370	0.203130	0.135420	1.007942	0.059609	0.059609	0.178827	0.245863	0.765510
0.120868	0.773185	0.181302	0.120868	1.027901	0.060469	0.060469	0.181408	0.249409	0.767521
0.120930	0.773566	0.181394	0.120930	1.027788	0.060464	0.060464	0.181395	0.249391	0.767521
0.120929	0.773565	0.181394	0.120929	1.027788	0.060464	0.060464	0.181394	0.249390	0.767521
$\mathcal{X}(I_{11})$ exact: 0.119353									Error: +1.32%

Table 1: Iteration results for the WTS in Fig. 6.

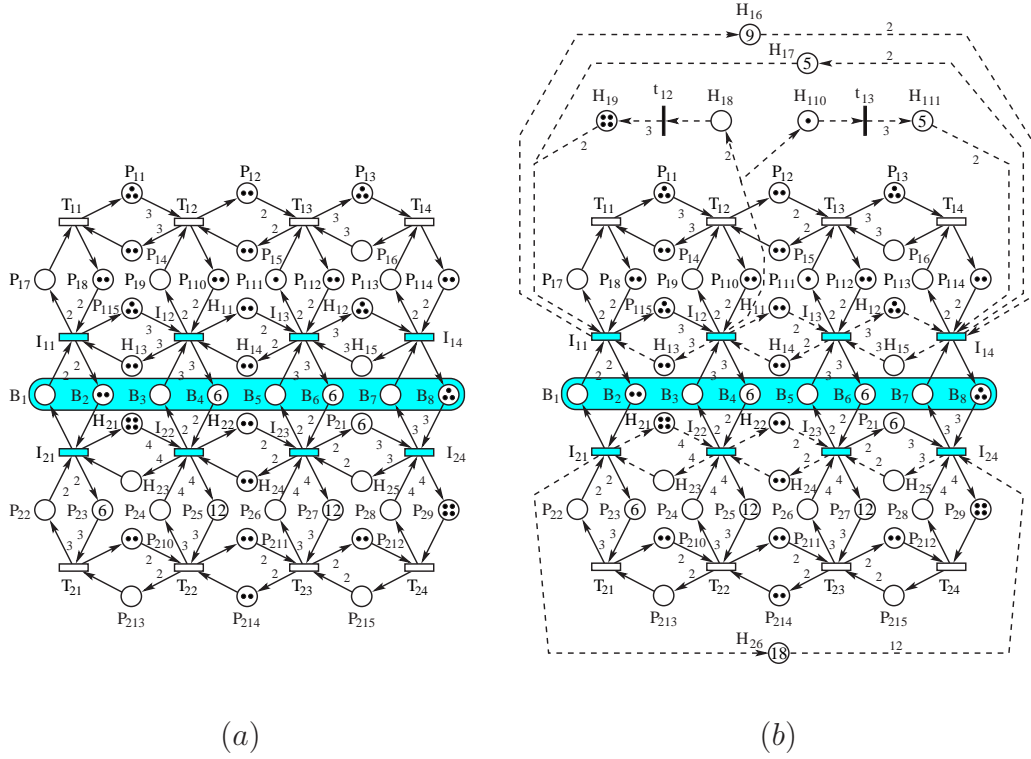


Figure 7: (a) A WTS's and (b) its  $\mathcal{ES}$ .

isomorphic CTMC of the original WTS has **90171** states (the same as the  $\mathcal{ES}$ ), while the low level systems and the  $\mathcal{BS}$  have **10751**, **9256** and **989** states, respectively.

We consider three different situations from different transition service rates (we assume single server semantics in all cases). In the first case, arbitrarily chosen service rates (from left to right and from top to bottom) 2,3,2,3,6,3,2,4,10,8,2,4,4,6,8,5. The exact throughput of transition  $I_{11}$  is 0.778787 in this case. In Table 2 we present the iteration steps of the method for this case. The legend is the same as in the table 1. Convergence of the method is usually obtained in 3 to 5 iteration steps. The error for this example was  $-0.22\%$ .

In the second case, we suppose all transitions rates equal to one in the original system (representing a symmetric timing case, in the sense that both sides of the system have response time of the same order of magnitude). The exact throughput of  $I_{11}$  is 0.217662.

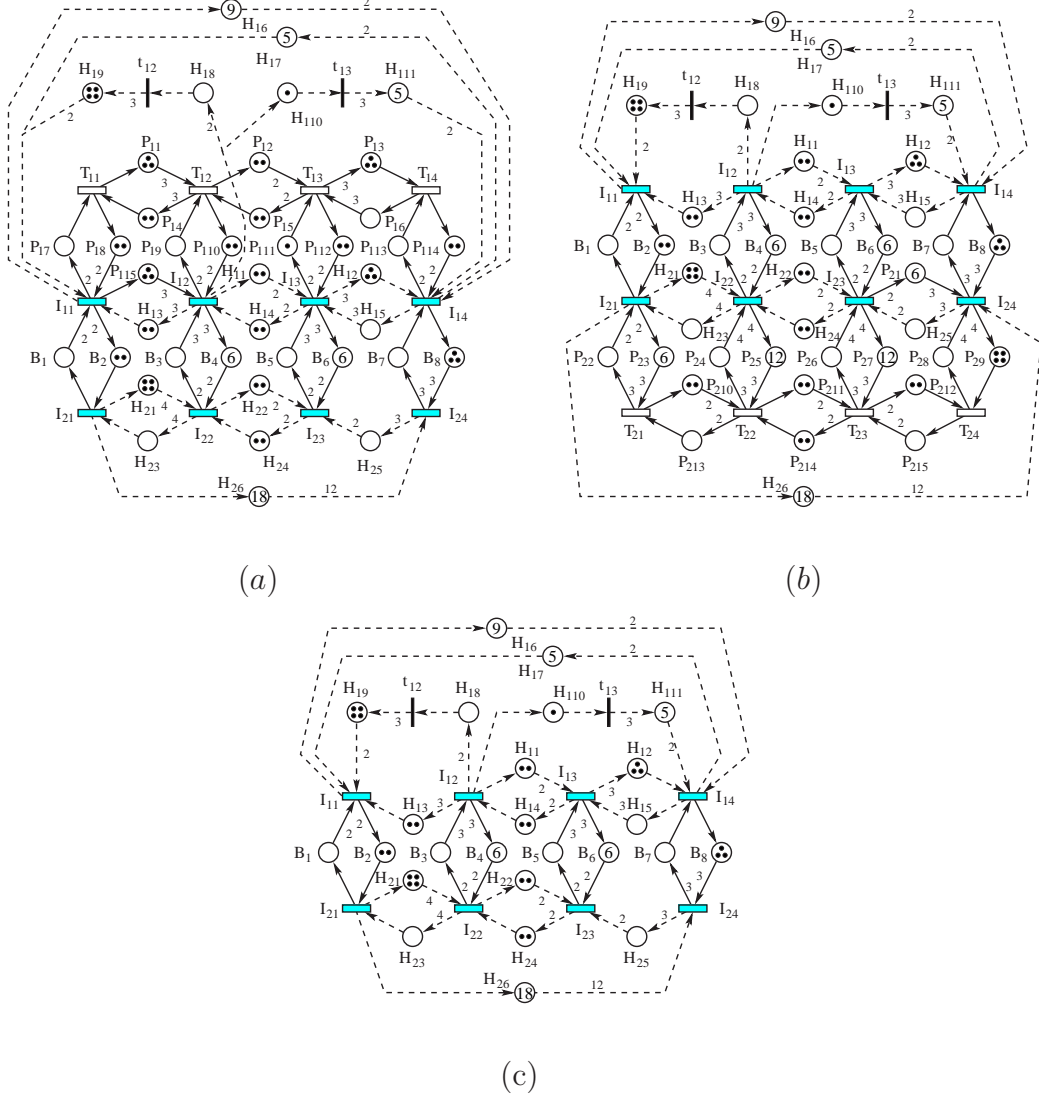


Figure 8: (a)  $\mathcal{LS}_1$ , (b)  $\mathcal{LS}_2$  and (c)  $\mathcal{BS}$  corresponding to the WTS in Fig 7.

The error of the approximated value is  $-0.45\%$ . On the other hand, if a very asymmetric timing is considered (service rates differ in three orders of magnitude for both subsystems, as given in the third case of Table 2), the exact throughput of  $I_{11}$  is 0.027686 leading up to an error of  $+0.02\%$ .

With regard to the accuracy of the method in other cases, no formal proof gives positive answers, but extensive testing allow us to assure that the error of the method is usually below 5%.

Variable service rates											
$\mathcal{LS}_1$						$\mathcal{LS}_2$					
$\mathcal{X}(I_{11})$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{14}$	$\lambda$	$\mathcal{X}(I_{21})$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{24}$	$\lambda$
0.8091	0.8276	0.0949	0.0674	0.6525	1.0388	1.5125	0.1741	0.1690	0.2379	0.1205	1.1279
0.7803	0.7939	0.0897	0.0650	0.5583	1.0523	1.5461	0.1785	0.1735	0.2516	0.1355	1.1067
0.7775	0.7881	0.0893	0.0648	0.5459	1.0541	1.5531	0.1795	0.1744	0.2534	0.1377	1.1042
0.7771	0.7873	0.0893	0.0648	0.5445	1.0544	1.5540	0.1796	0.1745	0.2536	0.1379	1.1040
0.7771	0.7872	0.0893	0.0648	0.5443	1.0544	1.5541	0.1796	0.1745	0.2536	0.1379	1.1039
$\mathcal{X}(I_{11})$ exact: 0.778787										Error: -0.22%	
Service rates equal to 1											
$\mathcal{LS}_1$						$\mathcal{LS}_2$					
$\mathcal{X}(I_{11})$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{14}$	$\lambda$	$\mathcal{X}(I_{21})$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{24}$	$\lambda$
0.2448	0.4630	0.0826	0.0408	0.5676	1.0145	0.4274	0.4575	0.2525	0.2140	0.1958	1.0523
0.2187	0.4078	0.0735	0.0365	0.4187	1.0323	0.4324	0.4636	0.2559	0.2216	0.2150	1.0423
0.2168	0.4028	0.0728	0.0361	0.4086	1.0354	0.4333	0.4647	0.2565	0.2220	0.2165	1.0418
0.2167	0.4025	0.0728	0.0361	0.4081	1.0355	0.4334	0.4647	0.2565	0.2220	0.2166	1.0417
0.2167	0.4024	0.0728	0.0361	0.4081	1.0355	0.4334	0.4647	0.2565	0.2220	0.2166	1.0417
$\mathcal{X}(I_{11})$ exact: 0.217662										Error: -0.45%	
Service rates 10, 1											
$\mathcal{LS}_1$						$\mathcal{LS}_2$					
$\mathcal{X}(I_{11})$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{14}$	$\lambda$	$\mathcal{X}(I_{21})$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{24}$	$\lambda$
0.0399	0.0040	0.0013	0.0007	0.2429	1.0118	0.0554	0.6568	0.3415	0.2232	0.2990	1.0409
0.0277	0.0028	0.0009	0.0005	0.1095	1.5438	0.0554	0.6571	0.3416	0.2233	0.3002	1.0396
0.0277	0.0028	0.0009	0.0005	0.1093	1.5438	0.0554	0.6571	0.3416	0.2233	0.3002	1.0396
0.0277	0.0028	0.0009	0.0005	0.1093	1.5746	0.0554	0.6572	0.3416	0.2233	0.3002	1.0396
$\mathcal{X}(I_{11})$ exact: 0.027686										Error: +0.02%	

Table 2: Iteration results for the WTS in Fig. 7.

## 8 Conclusions

The approximation technique presented here is mainly based on decomposition and aggregation of the obtained submodels. In order to aggregate subnets in weighted T-systems it is necessary to preserve some complex structural characteristics in order to achieve good approximation results. Our election was to preserve the gain, the minimal weighted marking and the maximum behavioural resistance between each pair of interface transitions. The preservation of the gain and minimal weighted marking assures liveness and boundedness of the obtained submodels. The preservation of the maximum behavioural resistance reflects part of the response time of the aggregated subnets that is due to the synchronization imposed by the weights of the arcs. Its preservation improves the approximation results. A *dynamic programming* algorithm was presented to compute in polynomial time on the net size the three selected characteristics (gain, weighted marking and behavioural resistance).

This decomposition algorithm supposes a generalization with respect to the previous one for marked graphs [9] in the following terms: First, a more general class of nets (with weights) is treated and the decomposition technique in this paper applied to a marked graph produces the same aggregated systems that the previous technique for marked graphs; second the original model can be decomposed in a finite number of subnets, not only two; finally, the number of nodes needed to reduce the subnets is directly reduced by the reduction algorithm. The efficiency of the proposed technique is directly influenced by the selected cut. Further research should be done in order to optimize this efficiency. However, from extensive experiments already achieved, the intuitive “cut the model into equal size pieces” approach (in terms of reachable markings) seems to give the best results.

With respect to the throughput approximation algorithm, we used basically the same *response time approximation* method used for marked graphs in [9], called *pelota* algorithm, that iteratively solves the isomorphic CTMC of each aggregated system. Existence of convergence points for the pelota algorithm have been proved (proof also valid for the algorithm for marked graphs in [9]). Extensive numerical experiments using this method showed fast convergence (3 to 5 iteration steps), small error (less than 5%) and the approximate computation of throughput can be achieved with a considerable saving of time and memory (more than one order of magnitude in many cases).

## References

- [1] S. C. Agrawal, J. P. Buzen, and A. W. Shum. Response time preservation: A general technique for developing approximate algorithms for queueing networks. In *Proceedings of the 1984 ACM Sigmetrics Conference on Measurement and Modeling of Computer Systems*, pages 63–77, Cambridge, MA, August 1984.



- [2] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *Data Structures and Algorithms*. Addison-Wesley, 1983.
- [3] M. Ajmone Marsan, G. Balbo, G. Conte, S. Donatelli, and G. Franceschinis. *Modelling with Generalized Stochastic Petri Nets*. John Wiley & Sons, 1995.
- [4] H. H. Ammar and S. M. R. Islam. Time scale decomposition of a class of generalized stochastic Petri net models. *IEEE Transactions on Software Engineering*, 15(6):809–820, June 1989.
- [5] F. Baccelli, G. Cohen, and B. Gaujal. Recursive equations and basic properties of timed Petri nets. *Discrete Event Dynamic Systems: Theory and Applications*, 1(4):415–439, 1992.
- [6] F. Baccelli and Z. Liu. Comparison properties of stochastic decision free Petri nets. *IEEE Transactions on Automatic Control*, 37(12):1905–1920, December 1992.
- [7] B. Baynat and Y. Dallery. Approximate techniques for general closed queueing networks with subnetworks having population constraints. Rapport MASI 90.49, Laboratoire MASI, Université Paris VI, Paris, France, October 1990.
- [8] J. Campos, G. Chiola, and M. Silva. Ergodicity and throughput bounds of Petri nets with unique consistent firing count vector. *IEEE Transactions on Software Engineering*, 17(2):117–125, February 1991.
- [9] J. Campos, J. M. Colom, H. Jungnitz, and M. Silva. Approximate throughput computation of stochastic marked graphs. *IEEE Transactions on Software Engineering*, 20(7):526–535, July 1994.
- [10] J. Campos and M. Silva. Embedded product-form queueing networks and the improvement of performance bounds for Petri net systems. *Performance Evaluation*, 18(1):3–19, July 1993.
- [11] G. Ciardo and K. S. Trivedi. A decomposition approach for stochastic Petri nets models. In *Proceedings of the 4th International Workshop on Petri Nets and Performance Models*, pages 74–83, Melbourne, Australia, December 1991. IEEE Computer Society Press.
- [12] J. M. Colom and M. Silva. Improving the linearly based characterization of P/T nets. In G. Rozenberg, editor, *Advances in Petri Nets 1990*, volume 483 of *Lecture Notes in Computer Science*, pages 113–145. Springer-Verlag, Berlin, 1991.
- [13] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press and McGraw-Hill, 1990.
- [14] F. DiCesare, G. Harhalakis, J. M. Proth, M. Silva, and F.B. Vernadat. *Practice of Petri Nets in Manufacturing*. Chapman & Hall, London, 1993.

- [15] S. Haddad, P. Moreaux, M. Sereno, and M. Silva. Structural characterization and qualitative properties of product form stochastic petri nets. In *22th International Conference on Application and Theory of Petri Nets*, Newcastle Upon Tyne, U. K., June 2001.
- [16] H. Hermanns, U. Herzog, U. Klehmet, V. Mertsiotakis, and M. Siegle. Compositional performance modelling with the TIPPTool. *Performance Evaluation*, 39(1-4):5–35, February 2000.
- [17] G. Klambauer. *Aspects of Calculus*. Springer-Verlag, Berlin, 1986.
- [18] V. Mainkar and K. Trivedi. Fixed point iteration using stochastic reward nets. In *Proceedings of the 6<sup>th</sup> International Workshop on Petri Nets and Performance Models*, pages 21–30, Durham, North Carolina, USA, October 1995. IEEE-Computer Society Press.
- [19] R. A. Marie. An approximate analytical method for general queueing networks. *IEEE Transactions on Software Engineering*, 5(5):530–538, September 1979.
- [20] T. Murata. Petri nets: Properties, analysis, and applications. *Proceedings of the IEEE*, 77(4):541–580, April 1989.
- [21] C. J. Pérez-Jiménez, J. Campos, and M. Silva. On approximate performance evaluation of manufacturing systems modelled with weighted T-systems. In *Proceedings of the IMACS/IEEE-SMC Multiconference on Computational Engineering in Systems Applications (CESA '96)*, pages 201–207, Lille, France, July 1996.
- [22] M. Silva. *Las Redes de Petri en la Automática y la Informática*. Editorial AC, Madrid, 1985. In Spanish.
- [23] D. R. Smart. *Fixed Point Theorems*. Cambridge University Press, 1974.
- [24] E. Teruel. *Structure Theory of Weighted Place/Transition Net Systems: The Equal Conflict Hiatus*. PhD thesis, Departamento de Ingeniería Eléctrica e Informática, Universidad de Zaragoza, Spain, June 1994.
- [25] E. Teruel, P. Chrzastowski-Wachtel, J. M. Colom, and M. Silva. On weighted T-systems. In K. Jensen, editor, *Application and Theory of Petri Nets 1992*, volume 616 of *Lecture Notes in Computer Science*, pages 348–367. Springer-Verlag, Berlin, 1992.
- [26] N. Thomas, J. T. Bradley, and D. J. Thornley. An approximate solution of PEPA models using component substitution. *IEE Proceedings - Computers and Digital Techniques*, 150(2):67–74, March 2003.
- [27] A. Zimmermann, J. Freiheit, and A. Huck. A Petri net based design engine for manufacturing systems. *International Journal of Production Research*, 39(2):225–253, January 2001.

# Appendix 1

**Theorem 3** Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS and  $\pi$  a circuit in  $\mathcal{N}$ . Then  $G(\pi) = 1$  (see [25]) and  $M(\pi, \mathbf{m}) = M(\pi, \mathbf{m}_0) \geq 1$  for all  $\mathbf{m} \in \text{RS}(\mathcal{S})$ .

**Proof:** In [25] there is a proof for  $G(\pi) = 1$ . Then, the firing of  $t_n$  modifies the weighted marking by  $1 - 1/G(\pi) = 0$ . So the weighted marking is the same in any reachable marking. Now, let  $\mathbf{m} \in \text{RS}(\mathcal{S})$  such that  $t_n$  is enabled in  $\mathbf{m}$  (so  $\mathbf{m}[p_n] \geq s_n$ ). By Def. 2,  $M(\pi, \mathbf{m}) \geq \frac{\mathbf{m}[p_n]}{r_n} \prod_{i=1}^{n-1} \frac{s_i}{r_i} \geq \prod_{i=1}^n \frac{s_i}{r_i} = 1/G(\pi) = 1 \forall \mathbf{m} \in \text{RS}(\mathcal{S})$ .  $\diamond$

**Theorem 4** Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS,  $t_i, t_o$  two transitions in  $\mathcal{N}$  and  $\pi_1, \pi_2$  two paths in  $\mathcal{N}$  from  $t_i$  to  $t_o$ . Then  $G(\pi_1) = G(\pi_2)$  and there exists  $q \in \mathbb{Q}$  such that  $\forall \mathbf{m} \in \text{RS}(\mathcal{S}), M(\pi_1, \mathbf{m}) - M(\pi_2, \mathbf{m}) = q$ .

**Proof:** If  $G(\pi_1) \neq G(\pi_2)$  we can suppose without loss of generality that  $G(\pi_1) < G(\pi_2)$ . Firing any other transition than  $t_o$  does not modify  $M(\pi_2) - M(\pi_1)$  and each firing of  $t_o$  increases  $M(\pi_2) - M(\pi_1)$  by  $1/G(\pi_1) - 1/G(\pi_2) > 0$  which contradicts boundedness (by liveness). So  $G(\pi_1) = G(\pi_2)$  and  $M(\pi_2) - M(\pi_1)$  is constant for any reachable marking. Let be  $q = M(\pi_1, \mathbf{m}_0) - M(\pi_2, \mathbf{m}_0) \in \mathbb{Q}$ . Then  $M(\pi_1, \mathbf{m}) - M(\pi_2, \mathbf{m}) = q$  for all  $\mathbf{m} \in \text{RS}(\mathcal{S})$ .  $\diamond$

**Theorem 6** Let  $\langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS,  $\mathbf{C}$  its incidence matrix,  $t_k \in T$  and  $\mathbf{g}_k \in (\mathbb{Q}^+)^{|T|}$  the gain vector from  $t_k$ . Then  $\mathbf{C} \cdot \mathbf{g}_k = 0$  (i. e.  $\mathbf{g}_k$  is a  $T$ -semiflow).

**Proof:** We will prove that  $\mathbf{C}[p] \cdot \mathbf{g}_k = 0$  for all  $p \in P$ . Let  $t_i = \bullet p, t_o = p \bullet, r = W(t_i, p)$  and  $s = W(p, t_o)$ . Then  $G(t_i, t_o) = \frac{r}{s}$  and  $\mathbf{C}[p] \cdot \mathbf{g}_k = r\mathbf{g}_k(t_i) - s\mathbf{g}_k(t_o) = rG(t_k, t_i) - sG(t_k, t_o) = s(G(t_k, t_i) \frac{r}{s} - G(t_k, t_o)) = s(G(t_k, t_i) \cdot G(t_i, t_o) - G(t_k, t_o)) = s(G(t_k, t_o) - G(t_k, t_o)) = 0$ .  $\diamond$

**Proposition 8** Let  $\langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS. Let  $\pi$  be a circuit in  $\mathcal{N}$ . Then  $M(\pi, \mathbf{m}_0) \geq RE(\pi)$ .

**Proof:** Let  $\{p_i\}_{i=1}^n, \{t_i\}_{i=1}^n$  be the places and transitions of  $\pi$  and  $r_i, s_i$  with  $1 \leq i \leq n$  the weights of the arcs of  $\pi$  ( $r_i$  for transition to place arcs and  $s_i$  for place to transition arcs). Let  $i$  be the

integer such that  $RE(\pi) = \prod_{j=1}^i \frac{s_j}{r_j}$ . We will observe the firing of transition  $t_i$ . Let  $\mathbf{m} \in \text{RS}(\mathcal{S})$  s. t.  $t_i$  is enabled in  $\mathbf{m}$ . Then  $\mathbf{m}[p_i] \geq s_i$  so  $M(\pi, \mathbf{m}) \geq \frac{\mathbf{m}[p_i]}{r_i} \prod_{j=1}^{i-1} \frac{s_j}{r_j} \geq \prod_{j=1}^i \frac{s_j}{r_j} = RE(\pi)$ . By theorem 3,  $M(\pi, \mathbf{m}_0) = M(\pi, \mathbf{m})$  so  $M(\pi, \mathbf{m}_0) \geq RE(\pi)$ .  $\diamond$

**Theorem 11** *Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS and  $p \notin P$  a place to be added with  $\bullet p = t$ ,  $p \bullet = t'$  and  $G(t, t') = \frac{g_1}{g_2}$ . If  $W(t, p) = g_1$  and  $W(p, t') = g_2$  then  $p$  is a TT-MSIP with respect to  $\mathcal{S}$ .*

**Proof:**  $\mathcal{S}$  is live and bounded, so strongly connected. Let  $\pi$  be any path from  $t$  to  $t'$ . By Def. 5  $G(\pi) = G(t, t')$ . Let  $\bar{p}$  the complementary place of  $p$ . Then  $\pi' = \pi \circ \bar{p}$  is a circuit of gain 1, so it is a conservative component (see [25]). Then, there exists  $\mathbf{y} \geq \mathbf{0}$  with  $\mathbf{y}|_{\pi'} \geq \mathbf{1}$  such that  $\mathbf{y} \cdot \mathbf{C}^{\bar{p}} = 0$  ( $\mathbf{C}^{\bar{p}}$  is the incidence matrix of  $\mathcal{N}^{\bar{p}}$ ). Let be  $\mathbf{y}' = \mathbf{y}|_{\mathcal{N}}$  and  $y = \mathbf{y}[\bar{p}] \geq 1$ . Then  $\mathbf{0} = \mathbf{y} \cdot \mathbf{C}^{\bar{p}} = \mathbf{y}' \cdot \mathbf{C} + y \cdot \mathbf{C}_{\bar{p}} = \mathbf{y}' \cdot \mathbf{C} - y \cdot \mathbf{C}_p \Rightarrow \mathbf{C}_p = (\mathbf{y}'/y) \cdot \mathbf{C}$ . Now,  $\mathbf{y}' \geq \mathbf{0}$  and  $y \geq 1$ , so  $p$  is a TT-MSIP (by theorem 10).  $\diamond$

**Theorem 12** *Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS and  $t, t'$  two transitions such that  $G(t, t') = \frac{g_1}{g_2}$ . Let  $p \notin P$  be a TT-MSIP with  $W(t, p) = g_1$  and  $W(p, t') = g_2$ . If  $\mathbf{m}_0[p] \geq \mathbf{m}_0^{\min}[p] = \lfloor g_1 \cdot \min\{M(\pi, \mathbf{m}_0) \mid \pi \in \mathcal{P}(t, t')\} \rfloor$  then  $p$  is implicit.*

**Proof:** Let  $\pi \in \mathcal{P}(t_i, t_o)$ . Then  $G(\pi) = G(t_i, t_o)$ . Let  $\{p_i\}_{i=1}^n, \{t_i\}_{i=0}^n$  be the nodes of  $\pi$  ( $t_0 = t_i, t_n = t_o$ ) and  $r_i = W(t_{i-1}, p_i), s_i = W(p_i, t_i), 1 \leq i \leq n$ .  $G(\pi) = G(t_i, t_o) = \frac{g_1}{g_2} = \prod_{i=1}^n \frac{r_i}{s_i}$ . We can suppose that  $g_1 = \prod_{i=1}^n r_i$  and  $g_2 = \prod_{i=1}^n s_i$ , so  $\mathbf{m}_0^{\min}[p] = g_1 \cdot \min\{M(\pi, \mathbf{m}_0) \mid \pi \in \mathcal{P}(t_i, t_o)\} \in \mathbb{N}$ . We will prove that  $L(\mathcal{S}) = L(\mathcal{S}^p)$ .

$L(\mathcal{S}^p) \subseteq L(\mathcal{S})$ .  $\mathcal{S}$  is constructed from  $\mathcal{S}^p$  by removing constraints for firing transitions (place  $p$ ).  $L(\mathcal{S}) \subseteq L(\mathcal{S}^p)$ . We will prove that  $\forall \mathbf{m}^p \in \text{RS}(\mathcal{S}^p)$  s. t.  $\mathbf{m}^p[q] \geq W(q, t_o) \forall q \in \bullet t_o \setminus \{p\}$  then  $\mathbf{m}^p[p] \geq g_2$ . Let  $\pi \in \mathcal{P}(t_i, t_o)$  in  $\mathcal{N}$  with minimal weighted marking ( $\pi$  always exists because a non simple path has greater weighted marking than the contained simple one (th. 3) and  $\mathcal{P}(t_i, t_o)$  is finite). Let  $\pi' = t_i p t_o$  be a path in  $\mathcal{N}^p$ .  $\pi, \pi' \in \mathcal{P}(t_i, t_o)$  in  $\mathcal{N}^p$ . By construction,  $G(\pi) = G(\pi')$  and  $M(\pi, \mathbf{m}_0) = M(\pi, \mathbf{m}_0^p) = M(\pi', \mathbf{m}_0^p)$ , so  $M(\pi, \mathbf{m}^p) = M(\pi', \mathbf{m}^p) \forall \mathbf{m}^p \in \text{RS}(\mathcal{S}^p)$ . Let be  $\mathbf{m}^p \in \text{RS}(\mathcal{S}^p)$  s. t.  $\mathbf{m}^p[q] \geq W(q, t_o) \forall q \in \bullet t_o \setminus \{p\}$ . Then  $\mathbf{m}^p[p_n] \geq s_n$ , so  $M(\pi, \mathbf{m}^p) \geq 1/G(\pi) = \frac{g_2}{g_1}$ ,

in other words,  $M(\pi', \mathbf{m}^p) = M(\pi, \mathbf{m}^p) \geq \frac{g_2}{g_1}$ . As  $M(\pi', \mathbf{m}^p) = \frac{\mathbf{m}^p[p]}{g_1}$ , then  $\mathbf{m}^p[p] \geq g_2$ . Thus, place  $p$  is implicit for  $\mathbf{m}_0[p] = \mathbf{m}_0^{\min}[p] = \lfloor g_1 \cdot \min\{M(\pi, \mathbf{m}_0) \mid \pi \in \mathcal{P}(t_i, t_o)\} \rfloor$ . Increasing  $\mathbf{m}_0[p]$  only traps tokens in  $p$  (by th. 4, because the minimal weighted marking from  $t_i$  to  $t_o$  does not change), so  $p$  is implicit for  $\mathbf{m}_0[p] \geq \mathbf{m}_0^{\min}[p]$ .  $\diamond$

**Theorem 13** *Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS;  $t_i, t_j, t_k$  three transitions in  $\mathcal{N}$ , and  $p_{ik}, p_{kj}, p_{ij}$  three TT-MSIP's from  $t_i$  to  $t_k$ ,  $t_k$  to  $t_j$  and  $t_i$  to  $t_j$  respectively, with the initial marking  $\mathbf{m}_0^{\min}$  given in theorem 12. If there exists a minimal weighted marking path in  $\mathcal{N}$  from  $t_i$  to  $t_j$  through  $t_k$ , then  $p_{ij}$  is an implicit place with respect to the system generated by places  $p_{ik}, p_{kj}$ .*

**Proof:** Let  $\pi_3$  be a minimal weighted marking path from  $t_i$  to  $t_j$  through  $t_k$  and  $\pi_1, \pi_2$  the sub-paths of  $\pi_3$  from  $t_i$  to  $t_k$  and  $t_k$  to  $t_j$  respectively.  $\pi_1$  and  $\pi_2$  are minimal weighted marking paths from  $t_i$  to  $t_k$  and  $t_k$  to  $t_j$  respectively. Let  $r_i, s_i$  with  $1 \leq i \leq n$  be the weights of the arcs of  $\pi_1$  and  $r'_i, s'_i$  with  $1 \leq i \leq m$  the weights of the arcs of  $\pi_2$ . Let  $r = \prod_{i=1}^n r_i$ ,  $s = \prod_{i=1}^m s_i$ ,  $r' = \prod_{i=1}^m r'_i$  and  $s' = \prod_{i=1}^n s'_i$ . We can suppose, without loss of generality that  $W(t_i, p_{ik}) = r$ ,  $W(p_{ik}, t_k) = s$  and  $\mathbf{m}_0[p_{ik}] = \lfloor rM(\pi_1, \mathbf{m}_0) \rfloor = rM(\pi_1, \mathbf{m}_0)$  (because  $rM(\pi_1, \mathbf{m}_0) \in \mathbb{Z}^+$ ). Similarly,  $W(t_k, p_{kj}) = r'$ ,  $W(p_{kj}, t_j) = s'$ ,  $\mathbf{m}_0[p_{kj}] = r'M(\pi_2, \mathbf{m}_0)$ ,  $W(t_i, p_{ij}) = rr'$ ,  $W(p_{ij}, t_j) = ss'$  and  $\mathbf{m}_0[p_{ij}] = rr'M(\pi_3, \mathbf{m}_0)$ . Let  $\pi$  be the path  $(t_i p_{ik} t_k p_{kj} t_j)$ . By theorem 4,  $G(\pi) = G(\pi_3) = \frac{rr'}{ss'}$  and  $M(\pi, \mathbf{m}_0) = \frac{\mathbf{m}_0[p_{ik}]}{r} + \frac{\mathbf{m}_0[p_{kj}]}{r'} = M(\pi_1, \mathbf{m}_0) + \frac{M(\pi_2, \mathbf{m}_0)}{G(\pi_1)} = M(\pi_3, \mathbf{m}_0)$ . So  $\pi$  is another minimal weighted marking path from  $t_i$  to  $t_j$ . By theorem 12,  $p_{ij}$  is also implicit with respect to the path  $\pi$ .  $\diamond$

**Theorem 19** *Let  $\mathcal{S} = \langle \mathcal{N}, \mathbf{m}_0 \rangle$  be a live and bounded WTS,  $B \subseteq P$  a  $K$ -cut of  $\mathcal{N}$ . The  $\mathcal{ES}$ ,  $\mathcal{LS}_i$ ,  $1 \leq i \leq K$  and  $\mathcal{BS}$  are live and bounded WTS;  $L(\mathcal{ES})|_T = L(\mathcal{S})$  and  $RS(\mathcal{ES})|_P = RS(\mathcal{S})$ .*

**Proof:**  $\mathcal{ES}$  is constructed from  $\mathcal{S}$  by adding some places (TT-MSIP's) and immediate transitions. So,  $\mathcal{ES}$  is a WTS. All these nodes are the reduction (preserving the gain and weighted marking) of existing paths in  $\mathcal{S}$ . So  $L(\mathcal{ES})|_T = L(\mathcal{S})$  and  $RS(\mathcal{ES})|_P = RS(\mathcal{S})$ . Thus,  $\mathcal{ES}$  is live and bounded.

Now, the  $\mathcal{LS}_i$ ,  $1 \leq i \leq K$  and  $\mathcal{BS}$  are constructed from  $\mathcal{ES}$  by reducing some nodes, thus removing constraints to the firing of transitions, so the aggregated systems are also live WTS. To prove boundedness, in [25] it is proved that a WTS is live and bounded if and only if all cycles are live and they have gain equal to 1. As  $\mathcal{ES}$  is live and bounded, any cycle is live and has gain equal to 1. By construction, any cycle in the aggregated systems is also a cycle in  $\mathcal{ES}$ , so all cycles in the aggregated systems are live and have gain equal to 1, so the aggregated systems are live and bounded WTS.  $\diamond$

**Theorem 21** *The linear search on the  $\mathcal{BS}$  in algorithm 20 always converges.*

**Proof:** Given the aggregated system  $\mathcal{LS}_i$  and arbitrarily chosen vectors of rates  $\lambda_j$  with  $1 \leq j \leq K$ ,  $j \neq i$  for transitions in  $T_j \cap \text{TI}$ , the algorithm 20 computes for  $\mathcal{LS}_i$  the vectors  $\chi_i$  and  $\mu_i$  of throughput and relative rates of transitions in  $T_i \cap \text{TI}$ . We select a reference transition  $t \in T_i \cap \text{TI}$ . Let  $\chi_{\mathcal{BS}}(\lambda) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  be the function that computes the throughput of  $t$  in  $\mathcal{BS}$  with  $(\lambda_1, \dots, \lambda\mu_i, \dots, \lambda_K)$  as vector of transition rates. We will prove the existence of  $\lambda \in \mathbb{R}^+$  such that  $\chi_{\mathcal{BS}}(\lambda) = \chi_i[t]$ . Obviously  $\chi_{\mathcal{BS}}(\lambda)$  is a continuous non decreasing function in  $\mathbb{R}^+$ .  $\chi_{\mathcal{BS}}(0) = 0$  because if a transition in a WTS has rate 0, then it never fires.  $\chi_{\mathcal{BS}}(\lambda)$  is a bounded function because if  $\lambda$  is arbitrarily increased the transitions in  $\text{TI} \setminus T_i$  maintain the throughput bounded. Taking the initial rates  $w_i$  of transitions in  $T_i \cap \text{TI}$  in  $\mathcal{LS}_i$ , there exists  $\lambda' \in \mathbb{R}^+$  such that  $w_i \leq \lambda'\mu_i$ , so by monotonicity on the topology and on the transition rates  $\chi_{\mathcal{BS}}(\lambda') \geq \chi_i[t]$ . Thus,  $\chi_{\mathcal{BS}}(\lambda)$  is continuous on  $[0, \lambda']$  and  $\chi_{\mathcal{BS}}(0) < \chi_i[t] \leq \chi_{\mathcal{BS}}(\lambda')$ . By Darboux property [17] there exists  $\lambda \in (0, \lambda']$  such that  $\chi_{\mathcal{BS}}(\lambda) = \chi_i[t]$ .  $\diamond$

**Theorem 23** *The function  $A : (\mathbb{R}^+)^M \rightarrow (\mathbb{R}^+)^M$  ( $M = |\text{IT}|$ ) of algorithm 22 is continuous and there exist a non empty, convex and compact set  $S \subset (\mathbb{R}^+)^M$  such that  $A(S) \subseteq S$ . Then, there exists  $x \in S$  such that  $A(x) = x$ .*

**Proof:** Given the rates  $\lambda$  of interface transitions, the algorithm computes (by means of  $\mathcal{LS}_1$ )  $\chi_1$  and  $\mu_1$ , the vector of throughput and relative rates of transitions in  $T_1 \cap \text{TI}$ .  $\chi_1$  is continuous on transitions rates, so it is continuous on  $\lambda$ . Given  $t \in T_1 \cap \text{TI}$ ,  $\mu_1[t]$  is the quotient between

the throughput  $\chi_1[t]$  (continuous on  $\lambda$ ) and the enabling degree  $e_t$  of  $t$  in  $\mathcal{LS}_1$  by means of the nodes of reduction. This enabling degree can be expressed as a continuous function on the steady-state probability vector of  $\mathcal{LS}_1$ , so it is continuous on  $\lambda$ . Also, by liveness of  $\mathcal{LS}_1$   $e_t > 0$ , so  $\chi_1[t]$  is continuous on  $\lambda$  for all  $t \in T_1 \cap \text{TI}$ . The proof of theorem 21 shows that the function  $\chi_{\mathcal{BS}}(\lambda) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  such that  $\chi_{\mathcal{BS}}(\lambda) = \chi_1[t]$  is continuous, bounded and monotonic for any  $t \in T_1 \cap \text{TI}$ , so there exists inverse function and it is continuous, thus the scale factor  $\lambda$  computed in the  $\mathcal{BS}$  is continuous on  $\lambda$ , so the new rates  $\lambda\chi_1$  of transitions in  $T_1 \cap \text{TI}$  are continuous on  $\lambda$ . By induction on the low level systems,  $\lambda^{(1)} = A(\lambda)$  is continuous.

Now, Taking  $\lambda^{(0)} = (\infty, \dots, \infty)$ , the throughput of transitions in  $\mathcal{LS}_i$  are finite and strictly positive because the internal transitions in  $\mathcal{N}_i$  have finite rates. Let  $\chi$  the maximum throughput obtained for a transition in a  $\mathcal{LS}_i$  in this manner. The maximum throughput obtained for any transition in the iterative algorithm is  $\chi$ . In  $\mathcal{BS}$ , by putting  $\lambda = 1/\chi \in (0, \infty)$  as the rate for a transition, and making immediate the other ones, we obtain  $\chi$  as the throughput in  $\mathcal{BS}$  for this transition. So, the maximum rate obtained for any transition in the iterative algorithm is  $\lambda$ . Let  $S = [0, \lambda]^M \subseteq (\mathbb{R}^+)^M$ . Clearly,  $S$  is non empty ( $\lambda > 0$ ), compact convex set and  $A(S) \subseteq S$ . Now, by applying the Brouwer fixed point theorem [23], there exists  $x \in S$  such that  $A(x) = x$ .  $\diamond$

## Appendix 2

**Theorem 16** Let  $S = (\mathbb{Q}^+ \times \mathbb{Q}^+ \times \mathbb{Q}^+ \setminus \{0\}) \cup \bar{0}$  be the domain of possible labels and  $\odot, \oplus$  the operators of definition 15. Then  $\langle S, \oplus, \odot, \bar{0}, \bar{1} \rangle$  is a relaxed closed semiring.

**Proof:** *i.1)*  $\langle S, \oplus, \bar{0} \rangle$  is a monoid.

a) Closure:  $s \oplus \bar{0}, \bar{0} \oplus s \in S$  for all  $s \in S$ . Now,  $(r_1, w_1, g) \oplus (r_2, w_2, g)$  has  $g \in \mathbb{Q}^+ \setminus \{0\}$  as third component and  $\min\{w_1, w_2\} \in \mathbb{Q}^+$  as second component. The first component is clearly a rational number. Now, let  $w_j = \min\{w_1, w_2\}$  (with  $j = 1$  or  $2$ ). Then  $\max_{i=1,2}\{r_i - w_i\} + \min_{i=1,2}\{w_i\} \geq r_j - w_j + w_j = r_j \geq 0$ , so the first component of  $(r_1, w_1, g) \oplus (r_2, w_2, g)$  is a non negative rational.

b) Identity: By definition.

c) Associativity: Let  $s_1, s_2, s_3 \in S$ . Two cases:

c.1) Some  $s_i$  is equal  $\bar{0}$ . As  $\bar{0}$  is identity element for  $\oplus$ , we have:  $\bar{0} \oplus (s_2 \oplus s_3) = s_2 \oplus s_3 = (\bar{0} \oplus s_2) \oplus s_3$ ;  $s_1 \oplus (\bar{0} \oplus s_3) = s_1 \oplus s_3 = (s_1 \oplus \bar{0}) \oplus s_3$ ;  $s_1 \oplus (s_2 \oplus \bar{0}) = s_1 \oplus s_2 = (s_1 \oplus s_2) \oplus \bar{0}$ .

c.2)  $s_i \neq \bar{0}$ ,  $i = 1, 2, 3$ . This case can be proved by associativity of  $\min$  and  $\max$  in  $\mathbb{Q}^+$ .

*i.2)*  $\langle S, \odot, \bar{1} \rangle$  is a monoid.

a) Closure:  $g \in \mathbb{Q}^+ \setminus \{0\}$  so  $1/g \in \mathbb{Q}^+ \setminus \{0\}$ . Now,  $\max, +$  are internal in  $\mathbb{Q}^+$  and  $\cdot$  is internal in  $\mathbb{Q}^+ \setminus \{0\}$ .

b) Identity:  $\bar{1} = (0, 0, 1)$ , so  $\bar{1} \odot \bar{0} = \bar{0} \odot \bar{1} = \bar{0}$ ;  $(r, w, g) \odot \bar{1} = (r, w, g) \odot (0, 0, 1) = (\max\{r, 0/g\}, w + 0/g, g \cdot 1) = (r, w, g)$ ;  $\bar{1} \odot (r, w, g) = (0, 0, 1) \odot (r, w, g) = (\max\{0, r/1\}, 0 + w/1, 1 \cdot g) = (r, w, g)$ .

c) Associativity: Let  $s_1, s_2, s_3 \in S$ . Two cases:

c.1) Some  $s_i = \bar{0}$ . As  $\bar{0}$  is annuller for  $\odot$ , we have:  $\bar{0} \odot (s_2 \odot s_3) = \bar{0}$  and  $(\bar{0} \odot s_2) \odot s_3 = \bar{0} \odot s_3 = \bar{0}$ ;  $s_1 \odot (\bar{0} \odot s_3) = s_1 \odot \bar{0} = \bar{0}$  and  $(s_1 \odot \bar{0}) \odot s_3 = \bar{0} \odot s_3 = \bar{0}$ ;  $s_1 \odot (s_2 \odot \bar{0}) = s_1 \odot \bar{0} = \bar{0}$  and  $(s_1 \odot s_2) \odot \bar{0} = \bar{0}$ .

c.2)  $s_i = (r_i, w_i, g_i) \neq \bar{0}$ ,  $i = 1, 2, 3$ .  $s_1 \odot (s_2 \odot s_3) = s_1 \odot \left( \max \left\{ r_2, \frac{r_3}{g_2} \right\}, w_2 + \frac{w_3}{g_2}, g_2 g_3 \right) = \left( \max \left\{ r_1, \frac{r_2}{g_1}, \frac{r_3}{g_1 g_2} \right\}, w_1 + \frac{w_2}{g_1} + \frac{w_3}{g_1 g_2}, g_1 (g_2 g_3) \right) = \left( \max \left\{ r_1, \frac{r_2}{g_1} \right\}, w_1 + \frac{w_2}{g_1}, g_1 g_2 \right) \odot (r_3, w_3, g_3) = (s_1 \odot s_2) \odot s_3$ .

*ii), iii), iv)*  $\bar{0}$  is annuller for  $\odot$ ,  $\oplus$  is commutative and idempotent by definition.

*v)*  $\odot$  distributes over  $\oplus$ . Let  $s_1, s_2, s_3 \in S$ . Two cases:



v.1) Some  $s_i = \bar{0}$ :  $\bar{0} \odot (s_2 \oplus s_3) = \bar{0}$  and  $(\bar{0} \odot s_2) \oplus (\bar{0} \odot s_3) = \bar{0} \oplus \bar{0} = \bar{0}$ ;  $(s_2 \oplus s_3) \odot \bar{0} = \bar{0}$  and  $(s_2 \odot \bar{0}) \oplus (s_3 \odot \bar{0}) = \bar{0} \oplus \bar{0} = \bar{0}$ ;  $s_1 \odot (\bar{0} \oplus s_3) = s_1 \odot s_3$  and  $(s_1 \odot \bar{0}) \oplus (s_1 \odot s_3) = \bar{0} \oplus (s_1 \odot s_3) = s_1 \odot s_3$ ;  $(\bar{0} \oplus s_3) \odot s_1 = s_3 \odot s_1$  and  $(\bar{0} \odot s_1) \oplus (s_3 \odot s_1) = \bar{0} \oplus (s_3 \odot s_1) = s_3 \odot s_1$ ;  $s_1 \odot (s_2 \oplus \bar{0}) = s_1 \odot s_2$  and  $(s_1 \odot s_2) \oplus (s_1 \odot \bar{0}) = (s_1 \odot s_2) \oplus \bar{0} = s_1 \odot s_2$ ;  $(s_2 \oplus \bar{0}) \odot s_1 = s_2 \odot s_1$  and  $(s_2 \odot s_1) \oplus (\bar{0} \odot s_1) = (s_2 \odot s_1) \oplus \bar{0} = s_2 \odot s_1$ .

v.2)  $s_i = (r_i, w_i, g_i) \neq \bar{0}$ ,  $i = 1, 2, 3$ . In order for the summaries to make sense, there must be  $g_2 = g_3 = g$ .

$$s_1 \odot (s_2 \oplus s_3) = (r_1, w_1, g_1) \odot \left( \max_{i=2,3} \{r_i - w_i\} + \min_{i=2,3} \{w_i\}, \min_{i=2,3} \{w_i\}, g \right) = \left( \max \left\{ r_1, \frac{1}{g_1} \left( \max_{i=2,3} \{r_i - w_i\} + \min_{i=2,3} \{w_i\} \right) \right\}, w_1 + \frac{1}{g_1} \min_{i=2,3} \{w_i\}, g_1 g \right).$$

Moreover

$$(s_1 \odot s_2) \oplus (s_1 \odot s_3) = \left( \max \left\{ r_1, \frac{r_2}{g_1} \right\}, w_1 + \frac{w_2}{g_1}, g_1 g \right) \oplus \left( \max \left\{ r_1, \frac{r_3}{g_1} \right\}, w_1 + \frac{w_3}{g_1}, g_1 g \right) = \left( \max \left\{ \max \left\{ r_1, \frac{r_2}{g_1} \right\} - w_1 - \frac{w_2}{g_1}, \max \left\{ r_1, \frac{r_3}{g_1} \right\} - w_1 - \frac{w_3}{g_1} \right\} + \min \left\{ w_1 + \frac{w_2}{g_1}, w_1 + \frac{w_3}{g_1} \right\}, \min \left\{ w_1 + \frac{w_2}{g_1}, w_1 + \frac{w_3}{g_1} \right\}, g_1 g \right).$$

We will prove that the two expressions are equal. The third components are  $g_1 g$ ;  $w_1 + \frac{1}{g_1} \min_{i=2,3} \{w_i\} = w_1 + \min \left\{ \frac{w_2}{g_1}, \frac{w_3}{g_1} \right\} = \min \left\{ w_1 + \frac{w_2}{g_1}, w_1 + \frac{w_3}{g_1} \right\}$ , so the second components are also equal. Only the equality of the first components must be proven:

$$\begin{aligned} & \max \left\{ \max \left\{ r_1, \frac{r_2}{g_1} \right\} - w_1 - \frac{w_2}{g_1}, \max \left\{ r_1, \frac{r_3}{g_1} \right\} - w_1 - \frac{w_3}{g_1} \right\} + \min \left\{ w_1 + \frac{w_2}{g_1}, w_1 + \frac{w_3}{g_1} \right\} = \\ & \max \left\{ \max \left\{ r_1, \frac{r_2}{g_1} \right\} - \frac{w_2}{g_1}, \max \left\{ r_1, \frac{r_3}{g_1} \right\} - \frac{w_3}{g_1} \right\} - w_1 + w_1 + \frac{1}{g_1} \min \{w_2, w_3\} = \\ & \max \left\{ r_1 - \frac{w_2}{g_1}, \frac{r_2 - w_2}{g_1}, r_1 - \frac{w_3}{g_1}, \frac{r_3 - w_3}{g_1} \right\} + \frac{1}{g_1} \min_{i=2,3} \{w_i\} = \max \left\{ r_1 + \frac{1}{g_1} \max_{i=2,3} \{-w_i\}, \frac{1}{g_1} \max_{i=2,3} \{r_i - w_i\} \right\} + \\ & \frac{1}{g_1} \min_{i=2,3} \{w_i\} = \max \left\{ r_1 - \frac{1}{g_1} \min_{i=2,3} \{w_i\}, \frac{1}{g_1} \max_{i=2,3} \{r_i - w_i\} \right\} + \frac{1}{g_1} \min_{i=2,3} \{w_i\} = \\ & \max \left\{ r_1, \frac{1}{g_1} \left( \max_{i=2,3} \{r_i - w_i\} + \min_{i=2,3} \{w_i\} \right) \right\}. \end{aligned}$$

So  $s_1 \odot (s_2 \oplus s_3) = (s_1 \odot s_2) \oplus (s_1 \odot s_3)$ .

Now:  $(s_2 \oplus s_3) \odot s_1 = \left( \max_{i=2,3} \{r_i - w_i\} + \min_{i=2,3} \{w_i\}, \min_{i=2,3} \{w_i\}, g \right) \odot (r_1, w_1, g_1) = \left( \max \left\{ \max_{i=2,3} \{r_i - w_i\} + \min_{i=2,3} \{w_i\}, \frac{r_1}{g} \right\}, \min_{i=2,3} \{w_i\} + \frac{w_1}{g}, g g_1 \right)$ . Moreover

$$(s_2 \odot s_1) \oplus (s_3 \odot s_1) = \left( \max \left\{ r_2, \frac{r_1}{g} \right\}, w_2 + \frac{w_1}{g}, g g_1 \right) \oplus \left( \max \left\{ r_3, \frac{r_1}{g} \right\}, w_3 + \frac{w_1}{g}, g g_1 \right) = \left( \max \left\{ \max \left\{ r_2, \frac{r_1}{g} \right\} - w_2 - \frac{w_1}{g}, \max \left\{ r_3, \frac{r_1}{g} \right\} - w_3 - \frac{w_1}{g} \right\} + \min \left\{ w_2 + \frac{w_1}{g}, w_3 + \frac{w_1}{g} \right\}, \min \left\{ w_2 + \frac{w_1}{g}, w_3 + \frac{w_1}{g} \right\}, g g_1 \right).$$

We will prove that the two expressions are equal. The third components are  $g_1 g$ ; the second ones are also equal:  $\min_{i=2,3} \{w_i\} + \frac{w_1}{g} = \min \left\{ w_2 + \frac{w_1}{g}, w_3 + \frac{w_1}{g} \right\}$ . Only the equivalence of the first components

must be proven:

$$\begin{aligned}
& \max \left\{ \max \left\{ r_2, \frac{r_1}{g} \right\} - w_2 - \frac{w_1}{g}, \max \left\{ r_3, \frac{r_1}{g} \right\} - w_3 - \frac{w_1}{g} \right\} + \min \left\{ w_2 + \frac{w_1}{g}, w_3 + \frac{w_1}{g} \right\} = \\
& \max \left\{ \max \left\{ r_2, \frac{r_1}{g} \right\} - w_2, \max \left\{ r_3, \frac{r_1}{g} \right\} - w_3 \right\} - \frac{w_1}{g} + \frac{w_1}{g} + \min_{i=2,3} \{w_i\} = \\
& \max \left\{ r_2 - w_2, r_3 - w_3, \frac{r_1}{g} - w_2, \frac{r_1}{g} - w_3 \right\} + \min_{i=2,3} \{w_i\} = \max \left\{ \max_{i=2,3} \{r_i - w_i\}, \frac{r_1}{g} + \max_{i=2,3} \{-w_i\} \right\} + \\
& \min_{i=2,3} \{w_i\} = \max \left\{ \max_{i=2,3} \{r_i - w_i\}, \frac{r_1}{g} - \min_{i=2,3} \{w_i\} \right\} + \min_{i=2,3} \{w_i\} = \max \left\{ \max_{i=2,3} \{r_i - w_i\} + \min_{i=2,3} \{w_i\}, \frac{r_1}{g} \right\}.
\end{aligned}$$

$$\text{So } (s_2 \oplus s_3) \odot s_1 = (s_2 \odot s_1) \oplus (s_3 \odot s_1).$$

*vi)* Summary of sequences is well defined. Let  $\{s_i\}_{i=1}^{\infty} \subseteq S$  be a sequence of labels. If  $\exists s_i = \bar{0}$  it can be removed from the summary. So, if  $\{s_i \mid s_i \neq \bar{0}\}_{i=1}^{\infty}$  is finite then  $\bigoplus_{i=1}^{\infty} s_i \in S$  is well defined (by closure of  $\oplus$ ). If  $\{s_i \mid s_i \neq \bar{0}\}_{i=1}^{\infty}$  is infinite, in order for the summary to make sense, the labels  $\{s_i\}_{i=1}^{\infty}$  must be associated to paths  $\pi_i$  with the same origin ( $t$ ) and destination ( $t'$ ). Then  $s_i = (r_i, w_i, g)$  for all  $i \in \mathbb{N}$  and the summary has the same gain  $g$ . The second component of the summary is  $\min_{i=1}^{\infty} \{w_i\}$ . We will prove that this minimum exists and it is a non negative rational number. All  $w_i \geq 0$  so  $\min_{i=1}^{\infty} \{w_i\} \geq 0$ . The numbers  $w_i$  are weighted markings of paths in the WTS from  $t$  to  $t'$ . Given a transition  $t$ , any edge  $p_j$  of the WTS (graph form) adds a fixed quantity  $q_j \in \mathbb{Q}^+$  to the weighted marking of any path with the edge. Given any rational number  $w$ , for any edge  $p_j$  in the graph there exist a finite natural number  $n_j$  such that  $n_j q_j \leq w$ . Then, given  $w_1$  as rational number, there exist only a finite number of paths from  $t$  to  $t'$  with weighted marking less or equal than  $w_1$ . So,  $\min_{i=1}^{\infty} \{w_i\}$  exists and it is a non negative rational number.

The first component of the summary is  $R = \max_{i=1}^{\infty} \{r_i - w_i\} + \min_{i=1}^{\infty} \{w_i\}$ . Let  $\pi_j$  be a path of the summary with minimal weighted marking  $w_j$ . Then  $R = \max_{i=1}^{\infty} \{r_i - w_i\} + w_j \geq r_j - w_j + w_j = r_j \geq 0$ . So, the first component is non negative. It must be proven that there exist  $\max_{i=1}^{\infty} \{r_i - w_i\} \in \mathbb{Q}$ . By definition of  $\oplus$ , the behavioural resistance of a path is the resistance decreased by the trapped weighted marking due to synchronizations with other paths. Let  $RE_i = RE(\pi_i)$  (resistance) and  $W_i = RE_i - r_i$  the trapped weighted marking of the path  $\pi_i$  for  $i \in \mathbb{N}$ . We will prove that  $\{RE_i\}_{i=1}^{\infty}$ ,  $\{W_i \mid W_i \leq RE_i\}_{i=1}^{\infty}$  and  $\{w_i \mid w_i \leq r_i\}_{i=1}^{\infty}$  are finite sets.

A similar reasoning shows that  $\{RE_i\}_{i=1}^{\infty}$  is a finite set. Given an initial transition, any edge  $p_j$  of the graph gives a fixed level of resistance  $u_j \in \mathbb{Q}^+ \setminus \{0\}$  to any path passing through  $p_j$  and

$RE_i = \max_{p_j \in \pi_i} \{u_j\}$ . Now,  $\{p_j\}$  is finite, so  $\{u_j\}$  is finite, thus  $\{RE_i\}_{i=1}^\infty \subseteq \mathbb{Q}^+ \setminus \{0\}$  is also finite. Let  $RE = \max_{i=1}^\infty \{RE_i\} \in \mathbb{Q}^+ \setminus \{0\}$ . Now  $r_i = RE_i - W_i \leq RE_i \leq RE \forall i \in \mathbb{N}$  so  $\{w_i \mid w_i \leq r_i\}_{i=1}^\infty \subseteq \{w_i \mid w_i \leq RE\}_{i=1}^\infty$ . As  $\{w_i \mid w_i \leq RE\}_{i=1}^\infty$  is finite then  $\{w_i \mid w_i \leq r_i\}_{i=1}^\infty$  is also finite. Finally,  $W_i$  is the trapped weighted marking in  $\pi_i$ , so the difference of weighted markings of paths in the summary, thus  $\{W_i \mid W_i \leq RE\}_{i=1}^\infty$  is finite. As  $\{W_i \mid W_i \leq RE_i\}_{i=1}^\infty \subseteq \{W_i \mid W_i \leq RE\}_{i=1}^\infty$  then  $\{W_i \mid W_i \leq RE_i\}_{i=1}^\infty$ , is also finite. Summarizing  $\{RE_i\}_{i=1}^\infty$  and  $\{W_i \mid W_i \leq RE_i\}_{i=1}^\infty$  are finite sets, so  $\{RE_i - W_i \mid RE_i - W_i \geq 0\}_{i=1}^\infty = \{r_i \mid r_i \geq 0\}_{i=1}^\infty$  is finite;  $\{w_i \mid w_i \leq r_i\}_{i=1}^\infty$  is finite, so  $\{r_i - w_i \mid r_i - w_i \geq 0\}_{i=1}^\infty$  is also a finite non empty set (a path with minimal weighted marking is in this set). Then  $\max_{i=1}^\infty \{r_i - w_i\} = \max\{r_i - w_i \mid r_i - w_i \geq 0\}$ . Moreover,  $r_i - w_i \in \mathbb{Q}$  for all  $i \in \mathbb{N}$ , so  $\max_{i=1}^\infty \{r_i - w_i\} \in \mathbb{Q}^+$ .

*vii)* Commutativity, associativity and idempotence applies to the summary of sequences: Labels  $\bar{0}$  can be removed (this is equivalent to a summary with a unique label  $\bar{0}$ ). In other case, applying commutativity e idempotence of  $\oplus$ , the labels  $(r, w, g)$  can be commuted and the repeated instances can be removed.

*viii)*  $\odot$  distributes over summary of sequences. Let  $a, b_i \in S$  with  $i \in \mathbb{N}$ . Two cases:

a)  $b_i = \bar{0}$  for all  $i \in \mathbb{N}$ . Then  $a \odot (\bigoplus_{i=1}^\infty b_i) = a \odot (\bigoplus_{i=1}^\infty \bar{0}) = a \odot \bar{0} = \bar{0}$  and  $\bigoplus_{i=1}^\infty (a \odot b_i) = \bigoplus_{i=1}^\infty (a \odot \bar{0}) = \bigoplus_{i=1}^\infty \bar{0} = \bar{0}$ ;  $(\bigoplus_{i=1}^\infty b_i) \odot a = (\bigoplus_{i=1}^\infty \bar{0}) \odot a = \bar{0} \odot a = \bar{0}$  and  $\bigoplus_{i=1}^\infty (b_i \odot a) = \bigoplus_{i=1}^\infty (\bar{0} \odot a) = \bigoplus_{i=1}^\infty \bar{0} = \bar{0}$ .

b)  $\exists b_i \neq \bar{0}$ . By *vii*, we can suppose  $b_i \neq \bar{0} \forall i \in \mathbb{N}$ . By induction on part *v*,  $\odot$  distributes over finite summaries. For  $\bigoplus_{i=1}^\infty b_i$  (see part *vii*)  $\exists n, m \in \mathbb{N}$  s. t.  $r_n - w_n = \max_{i=1}^\infty \{r_i - w_i\}$  and  $w_m = \min_{i=1}^\infty \{w_i\}$ . Then  $\exists k_1 \in \mathbb{N}$  ( $k_1 = \max\{n, m\}$ ) s. t.  $\bigoplus_{i=1}^\infty b_i = \bigoplus_{i=1}^{k_1} b_i \forall k \geq k_1$ . Similarly,  $\exists k_2 \in \mathbb{N}$  s. t.  $\bigoplus_{i=1}^\infty (a \odot b_i) = \bigoplus_{i=1}^{k_2} (a \odot b_i) \forall k \geq k_2$ . Let  $k = \max\{k_1, k_2\}$ . Then  $a \odot (\bigoplus_{i=1}^\infty b_i) = a \odot (\bigoplus_{i=1}^k b_i) = \bigoplus_{i=1}^k (a \odot b_i) = \bigoplus_{i=1}^\infty (a \odot b_i)$ ;  $(\bigoplus_{i=1}^\infty b_i) \odot a = (\bigoplus_{i=1}^k b_i) \odot a = \bigoplus_{i=1}^k (b_i \odot a) = \bigoplus_{i=1}^\infty (b_i \odot a)$ .  $\diamond$