# Two Size Measures for Timed Languages

Eugene Asarin, Aldric Degorre

# Two Size Measures for Timed Languages

Eugene Asarin
LIAFA – Université Paris Diderot and CNRS, France, eugene.asarin@liafa.jussieu.fr

Aldric Degorre
Université Libre de Bruxelles, Belgium, aldric.degorre@ulb.ac.be

October 20, 2010

**Abstract**

Quantitative properties of timed regular languages, such as information content (growth rate, entropy) are explored. The approach suggested by the same authors is extended to languages of timed automata with punctual (equalities) and non-punctual (non-equalities) transition guards. Two size measures for such languages are identified: mean dimension and volumetric entropy. The former is the linear growth rate of the dimension of the language; it is characterized as the spectral radius of a max-plus matrix associated to the automaton. The latter is the exponential growth rate of the volume of the language; it is characterized as the logarithm of the spectral radius of a matrix integral operator on some Banach space associated to the automaton. Relation of the two size measures to classical information-theoretic concepts is explored.

## 1 Introduction

In a previous work [4, 3], we have formulated the problem of measuring the size (or information content) of a timed regular language. There, we have associated with a language $L$ the volume $\mathcal{V}(L_n)$ of all its words of size $n$. This volume grows (or vanishes) exponentially as $n \to \infty$, and its rate (i.e. $\lim_n \log \mathcal{V}(L_n)/n$) is referred to as entropy $\mathcal{H}(L)$ of the language. In [4, 3], we characterize this entropy as spectral radius of an integral operator and give some methods to approximately compute it.

The volume-based definition of entropy has, however, some weaknesses when the automaton contains "punctual" transitions guarded by clock constraints of the form $x = c$. Indeed, as soon as a run of the automaton includes such a punctual transition, the volume corresponding to this run becomes 0. Hence, the information content of such a run is disregarded. For example, in the automaton on Fig. 1A, intuitively there are more runs on $ba^*$ than on $a^*$, however, according to our previous definitions, the volume of all the runs starting by $b$ is 0 (because they should cross a punctual edge), and they are disregarded.

Worse, if all the runs of some automaton include punctual transitions, the entropy becomes $\log 0 = -\infty$ and does not adequately represent the information content.

In this paper, we address the problem of adequately measuring the language size/ information content of a timed language accepted by a timed automaton with punctual and non-punctual transitions. Our solution is freely inspired by some ideas of symbolic dynamics [10], and especially by Gromov's mean dimension, see [11].

The first difficulty is conceptual: the language (up to $n$ events) corresponding to a timed automaton from a geometrical standpoint is a set of polyhedra in $\mathbb{R}^n$. Without punctual transitions, all these polyhedra are full-dimensional, and their cumulated volume is a good aggregate measure of the language. Whenever we allow punctual transitions, the dimension of polyhedra varies from 0 to $n$, and it becomes more difficult to find their total size. For example, in the automaton of Fig. 1B, the geometric set contains some (exponentially many) $n$-dimensional rectangles of volume

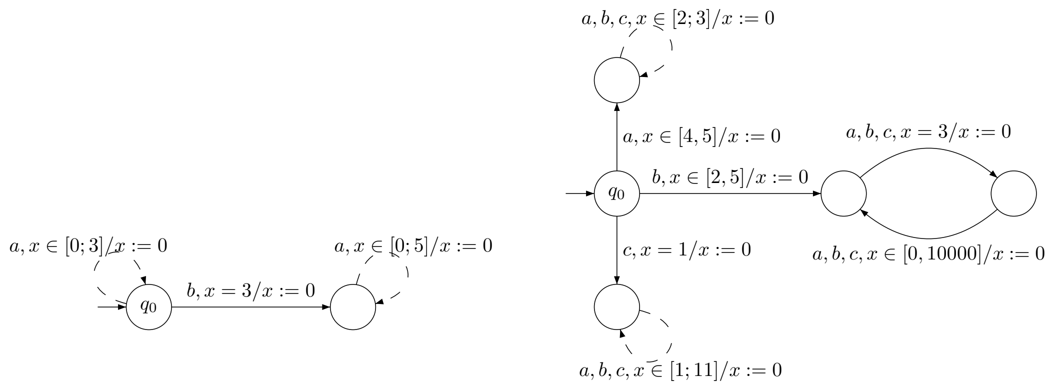


Figure 1: A. Is it reasonable not to go right? B. Three paths. Who will win?

1 corresponding to the words $a\Sigma^*$, some $n-1$-dimensional rectangles of volume $10^{n-1}$, and some $(n-1)/2$-dimensional rectangles of volume $3 \cdot 100^{n-1}$. A priori it is not clear how to sum up all these volumes. To address this difficulty, we measure the size of a multidimensional polyhedral set $S$ using a variant of $\varepsilon$-entropy from [8], which corresponds to the amount of information (in bits) needed to specify any point of $S$ with precision $\varepsilon$.

Applying this approach to a regular timed language $L$, we show that a typical timed word of $L_n$, whenever time is measured with precision $\varepsilon$, contains $n(\alpha \log(1/\varepsilon) + \mathcal{H})$ bits of information. Thus, the size (growth rate, information production rate) of $L$ is characterized by two numbers $(\alpha, \mathcal{H})$ referred to as *mean dimension and v-entropy*[1]. Roughly, $L_n$ resembles to an $\alpha n$-dimensional subset of $\Sigma^n \times \mathbb{R}^n$ of a volume $2^{n\mathcal{H}}$.

The main result of this paper is a characterization of $\alpha$ and $\mathcal{H}$ of the language $L$ accepted by a timed automaton $\mathcal{A}$ which proceeds as follows. After pre-processing the automaton (splitting its states in regions and removing unreachable states), we obtain a timed automaton $\mathcal{A}'$ and associate with it a max-plus matrix $\Phi$ (a kind of adjacency matrix of $\mathcal{A}'$). The mean dimension $\alpha(L)$ is the max-plus spectral radius of $\Phi$. The eigenspace corresponding to this spectral radius leads to identification of several *critical sub-automata* $\mathcal{A}_c$, where all the paths have the same mean dimension $\alpha$. For each such subautomaton $\mathcal{A}_c$, we build an integral operator $\Psi_c$ acting on a space of functions on the state space of $\mathcal{A}_c$. The $v$-entropy is the logarithm of the largest (among all the critical components) spectral radius of $\Psi_c$.

The paper is structured as follows. In Section 2, we introduce and illustrate the notion of $\varepsilon$-entropy and define mean dimension and $v$-entropy of a timed language. In Section 3, we make some assumptions on timed automata and describe how to preprocess them, and characterize volumes and dimensions of polyhedra in a timed regular language. In Section 4, we obtain the characterization of mean dimension as spectral radius of a natural max-plus matrix $\Phi$ (Theorem 9). We also describe the construction of the "critical sub-automata" $\mathcal{A}_c$. In Section 5, we associate to each $\mathcal{A}_c$ a Banach space and a positive linear operator $\Psi_c$ on this space. We characterize $v$-entropy in terms of its spectral radius in Theorem 14. We also discuss how this spectral radius can be computed in practice. In Section 6, we give an information-theoretic interpretation of $\alpha$ and $\mathcal{H}$ in terms of $\varepsilon$-entropy, which can be considered as correctness result for our algorithms. We conclude in Section 7, where we discuss related work and perspectives.

[1] "v" for volumetric.

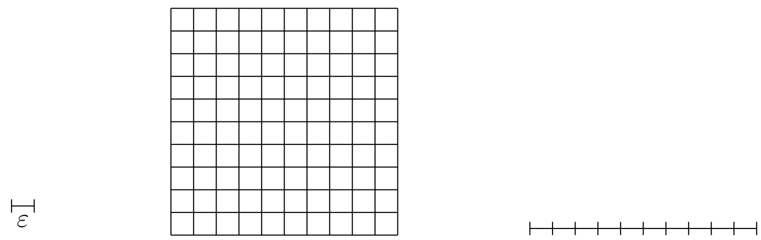

Figure 2: Adding meters to square meters: two polyhedra and their minimal $\varepsilon$-partitions.

# 2 Timed languages and their size measures

## 2.1 Some geometric terminology

A *convex polyhedron* $P \subset \mathbb{R}^d$ is a bounded finite intersection of half-spaces. If it is a subset of some $k$-dimensional affine subspace, but not of any $k-1$-dimensional one, we say that $\dim P = k$. A *polyhedral set* $P$ is a finite union of convex polyhedra. It can be decomposed into polyhedral components $P^m$ of various dimensions $m$ from 0 to $\dim P$. Such a decomposition could be non-unique, but we will always use a greedy algorithm: find maximal polyhedral subset of maximal dimension - this is the first component. Remove it from $P$, and repeat the procedure.

The notions above easily extend to subsets of $S \times \mathbb{R}^d$, with $S$ a finite set. Given such a subset $P$, we denote its component corresponding to an $s \in S$ by $P_s$, that is $P_s = \{x | (s, x) \in P\} \subset \mathbb{R}^d$. We call a subset $P$ polyhedral, if every component $P_s$ is a polyhedral set. The dimension of a polyhedral set $P$ is the maximum of dimensions of its components. $P_s$ can be further decomposed into subcomponents of different dimensions $P_s^m$, with $m \leq d$.

In this paper, we will use the well-known $\infty$-metric on $\mathbb{R}^d$ defined as follows:

$$d(\mathbf{x}, \mathbf{x}') = \max_i |x_i - x'_i|,$$

balls in this metric are cubes. It can be naturally extended to $S \times \mathbb{R}^d$:

$$d((s, \mathbf{x}), (s', \mathbf{x}')) = \begin{cases} d(\mathbf{x}, \mathbf{x}'), & \text{if } s = s'; \\ \infty, & \text{otherwise.} \end{cases}$$

## 2.2 Size of multidimensional sets

The key to measuring such multidimensional sets is provided by Kolmogorov and Tikhomirov's theory of $\varepsilon$-capacity and $\varepsilon$-entropy [8]. We will use a "diametric" variant of the notion of $\varepsilon$-entropy, following [12]. Given a compact metric space $X$, and a $\varepsilon > 0$, we define the $\varepsilon$-entropy of $X$ as logarithm of the minimum cardinality of a partition of $X$ into (Borel) subsets of a diameter $\leq \varepsilon$. The $\varepsilon$-entropy can be seen as the amount of information (in bits) that is necessary to represent an arbitrary point in $X$ with precision $\varepsilon$.

In particular, for an $m$-dimensional polyhedron $P$ of a volume $V$, the minimum cardinality of an $\varepsilon$-partition is close to $V/\varepsilon^m$, thus its logarithm is

$$h_\varepsilon(P) \approx \log V - m \log \varepsilon.$$

For a disjoint union of finitely many polyhedra $P_i$ of dimension $m_i$ and volume $V_i$, its $\varepsilon$-partition has a size close to $\sum_i V_i/\varepsilon^{m_i}$, (for small $\varepsilon$) and its logarithm can be considered again as information content. Fig. 2 illustrates this simple fact.

This justifies the following:

**Definition 1.** Formal $\varepsilon$-entropy of a finite disjoint family of polyhedra $P_i$ of dimension $m_i$ and volume $V_i$ is defined as:

$$\mathfrak{h}_\varepsilon(P) = \log \sum_i V_i/\varepsilon^{m_i}.$$

## 2.3 Timed languages and their polyhedra

Given a finite alphabet $\Sigma$, a *timed word* is a sequence $t_1a_1t_2a_2\dots t_na_n$ with *events* $a_i \in \Sigma$ and *delays* $t_i \in [0;\infty)$. A *timed language* is just a set of timed words. We will use a natural geometrical interpretation of timed words and languages. Thus a timed word $w = t_1a_1t_2a_2\dots t_na_n$ can be seen as a couple of a discrete word $\eta(w) = a_1a_2\dots a_n$ (called *untiming* of $w$) and a point $\theta(w) = (t_1, t_2, \dots, t_n) \in \mathbb{R}^n$ called its *timing*, or equivalently as a point in $\Sigma^n \times \mathbb{R}^n$. Similarly, we associate with a timed language $L$ and a natural $n$, a subset $L_n \subset \Sigma^n \times \mathbb{R}^n$.

For $L$ a timed regular language, all the geometrical sets described above are polyhedral.

In this paper, we will explore dimensionality and volume characteristics of $L$, such as $\dim(L_n)$, the dimension of the set of $n$-event timed words in $L$, and $\mathcal{V}(L_n^m)$, the volume of the $m$-dimensional component of this set (clearly, it can be non zero only for $m \le \dim(L_n)$).

## 2.4 Main definitions

The precise aim of this article is to explore the asymptotic behavior of

- $\dim(L_n)$ as $n \to \infty$ (we will show that it is linear);
- $\mathcal{V}(L_n^m)$ as $n \to \infty$ and $m \approx \dim(L_n)$ (we will show that it is exponential).

We will characterize this behavior by two rates (which are just real numbers):

**Definition 2.** Two size measures of a timed regular language are defined as follows:

- *Mean dimension* $\alpha(L) = \lim_{n\to\infty} \dim(L_n)/n$;
- *v-entropy* $\mathcal{H}(L) = \lim_{n\to\infty} \log \max\{\mathcal{V}(L_n^m) \mid \alpha n - d < m < \alpha n + d\}/n$ (with a constant $d$ specified below).

The definition of $\alpha(L)$ is very natural, it says that $L_n$ has dimension $\approx \alpha n$. The definition of $\mathcal{H}(L)$ saying that the volume of $L_n^m$ is approximately $2^{n\mathcal{H}}$ for dimension $m$ close to its maximal possible value, also seems plausible, the only possible doubt is related to the choice of $m$. We will justify these definitions by Theorem 18, which relates $\alpha$ and $\mathcal{H}$ to the formal $\varepsilon$-entropy $\mathfrak{h}_\varepsilon(L_n)$.

## 2.5 Example

To illustrate the notions above, consider again the timed language of the automaton of the Fig. 1B. Here we have the choice to explore three different areas of the automaton, depending on the choice of the first symbol between $a, b$ and $c$, yielding the following language: $L_B = [4;5]a([2;3]\Sigma)^* + [2;5]b(3\Sigma[0;10000]\Sigma)^* + 1c([1;11]\Sigma)^*$. The first branch produces $3^{n-1}$ copies of the $n$-dimensional rectangle $[4;5] \times [2;3]^{n-1}$, the second the same number of rectangles $[2;5] \times (\{3\} \times [0;10000])^{(n-1)/2}$ of dimension $(n+1)/2$ (we suppose $n$ odd). The third branch has as many $n-1$-dimensional rectangles $\{1\} \times [1;11]^{n-1}$. Clearly $\dim(L_n) = n$, thanks to the first branch. Hence, the mean dimension $\alpha(L) = 1$. Consider now the volumes:

$$\mathcal{V}(L_n^n) = 3^{n-1}; \quad \mathcal{V}(L_n^{(n+1)/2}) = 300^{n-1} \cdot 3; \quad \mathcal{V}(L_n^{n-1}) = 30^{n-1}.$$

According to Def. 2, $v$-entropy takes into account only the first and the third volumes (with $m \approx \alpha n = n$, and not the second one). The third volume wins the race, hence the $v$-entropy $\mathcal{H}(L) = \log 30 \approx 4.9$.

In order to understand why the huge second volume has been disqualified, let us consider the formal $\varepsilon$-entropy of $L_n$. The sum under logarithm would be:

$$\sum_i V_i/\varepsilon^{m_i} = 3^{n-1}\varepsilon^{-n} + 300^{n-1} \cdot 3\varepsilon^{(1-n)/2} + 30^{n-1}\varepsilon^{1-n}$$
$$= \varepsilon^{-n}\left(3^{n-1} + (300\varepsilon^{1/2})^{n-1} \cdot 3\varepsilon + 30^{n-1}\varepsilon\right).$$

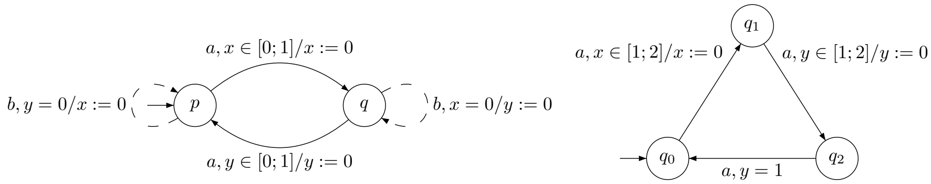


Figure 3: Two interesting automata

It shows that for $\varepsilon < \frac{1}{100}$, the third term, coming from the sublanguage $1c([1;11]\Sigma)^*$ is the preponderant one, despite the huge interval in $[2;5]b(3\Sigma[0;10000]\Sigma)^*$, which appears only every 2 events. Only terms with dimension $m$ close to the maximum $\alpha n$ can contribute.

In the example considered, we were able to compute mean dimension, $v$-entropy and formal $\varepsilon$-entropy directly from definitions. To convince the reader that it is not always possible, and advanced methods could be useful, we propose to consider the two automata on Fig. 3. We believe that their size parameters (at least $v$-entropy) cannot be obtained using elementary methods. We will use the automaton on the right of the figure as a running example.

# 3 Good timed automata and their pre-processing

We will be able to compute $\alpha$ and $\mathcal{H}$ for a subclass of timed automata, and, before proceeding we have to put the automaton in some normal form. The subclass and the normal form are very close to those from [3], the main difference is that punctual transitions are allowed.

We consider the following variant of Alur and Dill's timed automata (see [1] for original definition). A *timed automaton* (TA) is a tuple $\mathcal{A} = (Q, \Sigma, C, \Delta, q_0)$. Its elements are respectively the set of locations, the alphabet, the set of clocks, the transition relation, and the initial location (we do not need to specify accepting states since all the states are accepting, neither do we need any invariants). A generic state of $\mathcal{A}$ is a pair $(q, \mathbf{x})$ of a control location and a vector of clock values. A generic element of $\Delta$ is written as $\delta = (q, a, \mathfrak{g}, \mathfrak{r}, q')$ meaning a transition from $q$ to $q'$ with label $a$, guard $\mathfrak{g}$ and reset $\mathfrak{r}$. We spare the reader the definitions of a run of $\mathcal{A}$ and of acceptance, but we are obliged to fix some notations. Given an automaton $\mathcal{A}$, we write $L(\mathcal{A})$ or just $L$ for its accepted language, $L_p(\mathbf{x})$ for the language accepted by the runs starting at $(p, \mathbf{x})$; if we want to also specify the last state, we write $L_{pq}(\mathbf{x})$; finally, for the language accepted along some fixed path $\pi = \delta_1 \dots \delta_n$ in $\mathcal{A}$, we write $L_\pi(\mathbf{x})$. This notation will be freely combined with two indices for dimension, thus $L_{pn}^m(\mathbf{x})$ is the $m$-dimensional component of the set of traces of $n$-event runs starting at $p$ with clock values $\mathbf{x}$.

We say that a deterministic timed automaton with all accepting states and all guards bounded by a constant $M$ is *good* if the following Assumption holds:

A1. There exists a $D \in \mathbb{N}$ such that on every run segment of $D$ transitions, every clock is reset at least once.

We say that a good TA $\mathcal{A} = (Q, \Sigma, C, \delta, q_0)$ is in a *region-split form* if the following properties hold:

B1. Each location and each transition of $\mathcal{A}$ is visited by some run starting at $(q_0, 0)$.

B2. For every location $q \in Q$, a unique clock region $\mathbf{r}_q$ (called its *entry region*) exists, such that the set of clock values with which $q$ is entered is exactly $\mathbf{r}_q$. For the initial location $q_0$, its entry region is the singleton $\{0\}$.

B3. The guard $\mathfrak{g}$ of every non-punctual transition $\delta = (q, a, \mathfrak{g}, \mathfrak{r}, q') \in \Delta$ is just one clock region.

B4. The guard $\mathfrak{g}$ of every punctual transition $\delta = (q, a, \mathfrak{g}, \mathfrak{r}, q') \in \Delta$ has a form $x_i = c$.

| | Punctual $\delta$ | Non-punctual $\delta$ |
|---|---|---|
| $\delta$ | $(q, a, y = c, \mathfrak{r}, q')$ | $(q, a, \mathfrak{g}, \mathfrak{r}, q')$ |
| $P_{\delta\pi}(\mathbf{x})$ | $\{c - y\} \times P_\pi(\mathfrak{r}(\mathbf{x} + c - y))$ | $\bigcup_{\tau : \mathbf{x}+\tau \in \mathfrak{g}} \{\tau\} \times P_\pi(\mathfrak{r}(\mathbf{x} + \tau))$ |
| $\dim P_{\delta\pi}(\mathbf{x}) = \dim P_\pi(\mathbf{x}) + \phi_\delta$ | | |
| $\phi_\delta$ | 0 | 1 |
| $\mathcal{V}(P_{\delta\pi}(\mathbf{x})) = (\psi_\delta \mathcal{V} P_\pi)(\mathbf{x})$ | | |
| $(\psi_\delta v)(\mathbf{x})$ | $v(\mathfrak{r}(\mathbf{x} + c - y))$ | $\int_{\mathbf{x}+\tau \in \mathfrak{g}} v(\mathfrak{r}(\mathbf{x} + \tau))\, d\tau$ |

Table 1: Recurrence equations for polyhedra, dimensions, and volumes. $\psi_\delta$ is an operator transforming functions to functions.

Similarly to [3], it is easy to prove the following:

**Proposition 3.** *Given a good TA $\mathcal{A}$, a region-split TA $\mathcal{A}'$ accepting the same language can be constructed.*

### 3.1 Recurrent formulas

Given a region-split automaton $\mathcal{A}$, and a path $\pi = \delta_1 \dots \delta_n$ in this automaton, the timed language $L_\pi(\mathbf{x})$ corresponds to one convex polyhedron in $\Sigma^n \times \mathbb{R}^n$, we will denote its projection on $\mathbb{R}^n$ by $P_\pi(\mathbf{x})$. We will compute this polyhedron, its dimension and volume by recurrence on $\pi$, starting with an empty path and adding transitions at its beginning one by one.

For the base case (empty path $\epsilon$), we put

$$P_\epsilon(\mathbf{x}) = \{\mathbf{x}\}, \quad \dim P_\epsilon(\mathbf{x}) = 0, \quad \mathcal{V}(P_\epsilon(\mathbf{x})) = 1.$$

Suppose that we know $P = P_\pi(\mathbf{x})$. The recurrent formulas for $P_{\delta\pi}(\mathbf{x})$, its dimension and volume look differently for punctual and non-punctual $\delta$. We summarize them in Table 1 and deduce the following result:

**Proposition 4.** *The dimension and the volume of $P_\pi(\mathbf{x})$ for $\pi = \delta_1 \dots \delta_n$ are as follows:*

$$\dim P_\pi(\mathbf{x}) = \sum_{i=1}^{n} \phi_{\delta_i}; \tag{1}$$

$$\mathcal{V} P_\pi(\mathbf{x}) = (\psi_{\delta_1} \dots \psi_{\delta_n} 1)(\mathbf{x}). \tag{2}$$

Knowing the volume and the dimension for every path, we can in principle compute[2] them (for given $n$ and $m$) for the whole language and its sublanguages, in particular

$$\dim(L_n) = \max \left\{ \sum_{i=1}^{n} \phi_{\delta_i} | \delta_1 \dots \delta_n \text{ a path from } (q_0, 0) \right\}; \tag{3}$$

$$\mathcal{V}(L_n^m) = \sum \left\{ \psi_{\delta_1} \dots \psi_{\delta_n}(1) | \delta_1 \dots \delta_n \text{ a path from } (q_0, 0) \text{ s.t. } \sum_{i=1}^{n} \phi_{\delta_i} = m \right\}. \tag{4}$$

In two subsequent sections we will determine the asymptotical behavior of these quantities.

# 4 Max-plus and mean dimension

In this section, we show that $\dim L_n$ is approximately linear wrt $n$ and compute the rate of this dependence. We also clean up the automaton removing the paths that do not give the maximal dimension. The techniques used come from max-plus algebra (see [5]).

[2] Iterated integrals in the chain of $\psi_\delta$ only lead to polynomials and can be easily computed symbolically.

$$A = \begin{pmatrix} -\infty & -\infty & \mathbf{1} & 1 & -\infty & -\infty \\ \mathbf{2} & \mathbf{1} & -\infty & -\infty & -\infty & 1 \\ -\infty & \mathbf{0} & -\infty & -\infty & -\infty & -\infty \\ -\infty & -\infty & -\infty & -\infty & -\infty & -\infty \\ -\infty & 0 & -\infty & 1 & \mathbf{1} & -\infty \\ -\infty & -\infty & -\infty & -\infty & 0 & -\infty \end{pmatrix}$$

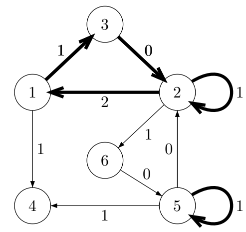


Figure 4: A matrix and a graph; critical edges represented in thick lines, $\alpha = 1$.

## 4.1 Recalling max-plus

Consider the set $\mathbb{R}_{\max} = \mathbb{R} \cup \{-\infty\}$ endowed with two operations: max ("addition" denoted $\oplus$) and $+$ ("multiplication" denoted $\otimes$). Operations $\oplus$ and $\otimes$ are extended in a natural way to vectors and matrices; $A^n$ denotes the $n$-th max-plus power of a matrix $A$.

Similarly to usual linear algebra, for a matrix $A$ we say that $\lambda \in \mathbb{R}_{\max}$ and $\mathbf{x} \in \mathbb{R}^n_{\max}$ are respectively an *eigenvalue* and an *eigenvector* of $A$ if $A \otimes \mathbf{x} = \lambda \otimes \mathbf{x}$. The highest eigenvalue called *spectral radius* of $A$ admits an interpretation in terms of paths in weighted graphs. An $n \times n$ matrix A corresponds to a weighted graph $G$ with vertices $1, \ldots, n$. Whenever $A_{ij} > -\infty$, there is an edge $(i, j)$ in the graph, its weight is $A_{ij}$. For a path $\pi$ in $G$, its weight $w(\pi)$ is the sum of weights of edges, and its mean weight is just $w(\pi)/|\pi|$. It is easy to see that $A^n_{ij}$ is the maximal weight of a path of $n$ edges from $i$ to $j$.

As for the spectral radius $\alpha$, it can be characterized as the maximum of mean weights for all circuits in the graph $G$. All the circuits having the same mean weight $\alpha$ are called critical. The critical subgraph $G_{\text{crit}} \subset G$ contains all the vertices and the edges of critical circuits (see an example on Fig. 4). We will need the following well-known results on weights of paths and powers of max-plus matrices:

**Proposition 5** (see [5])**.** *Let $A$ be a max-plus matrix, $\alpha(A)$ its spectral radius, $G_{crit}$ its critical subgraph. Then,*

- *$\alpha$ and $G_{crit}$ can be computed in $O(n^3)$ using Karp's algorithm;*
- *$G_{crit}$ is a union of several disjoint strongly connected graphs (critical components);*
- *for $i$ and $j$ in the same critical component, all the $n$-paths in $G_{crit}$ from $i$ to $j$ have the same weight equal to the maximal weight in the original graph: $= A^n_{ij}$, and close to $\alpha n$: more precisely for some constant $d$ we have $|A^n_{ij} - \alpha n| < d$;*
- *for arbitrary $i$ and $j$ only the upper bound holds: $A^n_{ij} < \alpha n + d$.*

We see that $\alpha$ is the optimal asymptotic mean weight, and that it is attained by any path in (an SCC of) the critical graph. On the other hand, any path visiting often enough non-critical edges has a lesser weight:

**Proposition 6.** *For any path $\pi$ in a matrix of $\{-\infty, 0, 1\}^{n \times n}$ with $m$ non-critical and $k$ critical edges, the following upper bound holds:*

$$w(\pi) < d + \alpha k + \beta m,$$

*with some constant $\beta < \alpha$ which depends only on the matrix $A$.*

### 4.2 Matrix $\Phi$

The theory described above applies almost directly to dimension of timed polyhedra. Indeed, let us define a max-plus matrix $\Phi$ such that $\Phi_{pq} = \max\{\phi_\delta \mid \delta \text{ from } p \text{ to } q\}$. In other words,

$$\Phi_{pq} = \begin{cases} -\infty & \text{if there is no transition from } p \text{ to } q; \\ 0 & \text{if there is a transition and all transitions are punctual;} \\ 1 & \text{if there is a non-punctual transition.} \end{cases}$$

Then, the following is almost immediate from (3).

**Proposition 7.** *$\Phi^n_{pq}$ is the dimension of $L_{pqn}(\mathbf{x})$ (for any clock valuation $\mathbf{x}$).*

Thus, we can apply the general theory, compute the spectral radius $\alpha(\Phi)$, the critical subgraph and its SCC decomposition. This way, for any SCC $c$ we obtain a *critical subautomaton* $\mathcal{A}_c$ whose control states are the ones corresponding to vertices of $c$ and whose transitions are those of $\mathcal{A}$, going along critical edges, and having maximal dimension for each such edge[3]. Applying Prop. 5 to $\mathcal{A}_c$ we obtain:

**Corollary 8.** *In each critical subautomaton $\mathcal{A}_c$, for any $n$-path $\pi$ from $p$ to $q$ , the dimension of $P_\pi(\mathbf{x})$ does not depend on the choice of the path. It is close to $\alpha n$: more precisely for some constant $d$ we have $|\dim P_\pi(\mathbf{x}) - \alpha n| < d$.*

And we deduce the main result of this section:

**Theorem 9.** *For any good region-split automaton $\mathcal{A}$, let $\alpha > -\infty$ be the spectral radius of its matrix $\Phi$. Then, for some constant $d$ and all $n$, the language $L$ of $\mathcal{A}$ satisfies:*

$$|\dim L_n - \alpha n| < d.$$

*Thus $\alpha$ is the mean dimension of $L$. (In the degenerate case when $\alpha = -\infty$, the automaton is acyclic and $L_n$ is empty for $n$ large enough.)*

**Example 10.** We consider the automaton on the right of Fig. 3 and put it in region-split form. We call $p_0$, $p_1$ and $p_2$ the region-split states corresponding to $q_0$, $q_1$ and $q_2$ with respective entry regions $[y = 1 < x < 2]$, $[0 = x < 1 < y]$ and $[y = 1 < x < 2]$. The only critical cycle is then $p_0 \xrightarrow{1<x<2,\ x:=0} p_1 \xrightarrow{1<y<2,\ y:=0} p_2 \xrightarrow{y=1} p_0$. As it yields two non-punctual and one punctual transition, its mean dimension, the max-plus spectral radius of $\Phi$, is $\frac{2}{3}$. Therefore, in this example, $\alpha = \frac{2}{3}$.

## 5 Functional analysis and $v$-entropy

Exploration of the asymptotic behavior of the volume goes along the same lines as for dimension, but instead of a max-plus matrix, a matrix of integral operators is iterated. As a result, while dimension's asymptotics is linear, the volume evolves exponentially.

### 5.1 Recalling functional analysis

In order to characterize and compute the $v$-entropy $\mathcal{H}$, we will use the approach introduced in [3], based on functional analysis (see e.g. [13]) and in particular the theory[4] of positive linear operators on Banach spaces (see [9]). We will need the following result:

**Theorem 11** (see [9])**.** *Given a positive linear bounded compact operator $\Theta$ with a spectral radius $\rho > 0$, defined on a Banach space ordered by a generating cone [5], the following holds:*

[3] We do not define the initial state for $\mathcal{A}_c$.

[4] A generalisation of the classical Perron-Frobenius theory to infinite dimension.

[5] The space $\mathcal{F}$ of continuous function considered below straightforwardly satisfies these properties.

1. *$\rho$ is an eigenvalue;*

2. *there exists a non-negative eigenvector $f \geq 0$ corresponding to this eigenvalue;*

3. *Gelfand's formula holds: $\lim_n ||\Theta^n||^{1/n} = \rho$.*

### 5.2 Banach space and operator $\Psi$

Similarly to [3], and to the previous section, we will represent equation (4) as iteration of some positive operator, and apply Theorem 11. However, we must take into account the two parameters $n$ and $m$ of the volume $\mathcal{V}(L_n^m)$. Our solution is as follows: we will consider the critical subgraph introduced above, and thus concentrate on the polyhedra of maximal dimension $m \approx \alpha n$ (which corresponds to the definition of $v$-entropy).

Given a good timed automaton in the region-split form, we first restrict to one critical subautomaton $\mathcal{A}_c$ (we denote its state space by $Q_c$). We define the Banach space $\mathcal{F}_c$ as the set of continuous bounded functions on the set $\{(q, \mathbf{x}) | q \in Q_c, \mathbf{x} \in \mathbf{r}_q\}$. The norm is defined by $||f|| = \sup_{q,\mathbf{x}} |f(q, \mathbf{x})|$. An element $f \in \mathcal{F}$ can be seen as a vector of functions $f_q$ with simplicial domains $\mathbf{r}_q$.

For any $p, q \in Q_c$, we define an operator $\psi_{pq}$ which maps functions over $\mathbf{r}_q$ to functions over $\mathbf{r}_p$ as the sum of all the operators $\psi_\delta$ (defined in Table 1) for the transitions going from $p$ to $q$. Next we define an operator $\Psi_c$ on $\mathcal{F}$ with the matrix $(\psi)_{pq}$:

$$(\Psi_c f)(p, \mathbf{x}) = \sum_q (\psi_{pq} f_q)(\mathbf{x}).$$

The operator $\Psi_c$ provides a simple form to formulas (4) restricted to critical paths and maximal dimension. Let $L^c$ (with subscripts and arguments interpreted in the standard way) denote the language of the critical subautomaton $\mathcal{A}_c$.

**Proposition 12.** *For any $p, q \in Q_c$:*

$$\mathcal{V}(L_{pqn}^c(\mathbf{x})) = (\Psi_c^n 1_q)_p(x),$$

*where the function $1_q$ equals one on the state $q$ (more precisely, on $\{q\} \times \mathbf{r}_q$), and zero elsewhere.*

### 5.3 Characterization of $v$-entropy

In Prop. 12, we have characterized the volume of $L_{pqn}^c(\mathbf{x})$ in terms of $n$th iteration of the matrix integral operator $\Psi_c$. On the other hand, this operator almost satisfies the hypotheses of Theorem 11:

**Proposition 13.** *$\Psi_c$ is a bounded linear positive operator. $\Psi_c^{D+1}$ is compact with $D$ as in Assumption A1.*

This allows to prove, using Theorem 11:

**Theorem 14.** *Let $\mathcal{A}$ be a good region-split TA, $\mathcal{A}_c$ its critical subautomaton, $\Psi_c$ the operator described above for $\mathcal{A}_c$, and $\rho_c$ the spectral radius of this operator. Then the following holds:*

- *for any $\sigma > 0$, and $n$ big enough, for any state $(q, \mathbf{x})$ of $\mathcal{A}_c$, the upper bound:*

$$\mathcal{V}(L_{qn}^c(\mathbf{x})) < (\rho_c + \sigma)^n;$$

- *and for some state $q$ of $\mathcal{A}_c$, some open set $O$ inside $\mathbf{r}_q$ and some $\gamma > 0$, for all $\mathbf{x} \in O$, for all $n$, the lower bound:*

$$\mathcal{V}(L_{qn}^c(\mathbf{x})) > \gamma \rho_c^n.$$

The theorem says that in a critical $\mathcal{A}_c$, the volume grows roughly as $\rho_c^n$, or, in exponential form, as $2^{n \log \rho_c}$. We can deduce now the required characterization of the $v$-entropy of the whole language of $\mathcal{A}$.

**Theorem 15.** *Let $\mathcal{A}$ be a good region-split TA. Then its $v$-entropy $\mathcal{H}$ is the maximum of $\log \rho_c$ over all its critical subautomata $\mathcal{A}_c$.*

**Example 16.** On the only critical component of the automaton on the right of Fig. 3 (containing only the three transitions of the critical cycle we mentioned earlier in Ex. 10), the operators $\psi_{pq}$ (defining $\Psi_c$) have the following expressions:

$$(\psi_{p_0p_1} f)(x,y) = \int_{\tau=0}^{2-x} f(0, y+\tau) d\tau;$$
$$(\psi_{p_1p_2} f)(x,y) = \int_{\tau=0}^{2-y} f(x+\tau, 0) d\tau;$$
$$(\psi_{p_2p_0} f)(x,y) = f(x+1, 1).$$

A close examination shows that the integral system $\Psi_c f = \lambda f$ can be rewritten using only one real variable, then differentiated twice and finally solved symbolically as a linear ordinary differential equation. Doing so, we find that the $\lambda$ having the highest absolute value such that the system still has non-trivial solutions is $\left(\frac{2}{\pi}\right)^{2/3}$. Therefore $\mathcal{H} = \frac{2}{3} \log \frac{2}{\pi}$.

### 5.4 Algorithmic aspects

Practical computation of the spectral radius of an operator $\Psi$ represented by a matrix of integral operators is a nontrivial task. However, the two methods proposed in [3] can be applied almost without change. We refer the reader to [3, 2], and only sketch the two methods:

- The first one applies to the subclass of "1 1/2 clocks" automata such that all the regions $\mathbf{r}_q$ are of dimension 0 or 1 (it means that all the clocks but one should be reset when taking a transition). For such an automaton it is possible to transform the integral eigenvalue equation $\Psi v = \lambda v$ to a system of linear ordinary differential equations (indeed, unknown functions $v_q$ are functions of scalar arguments), solve it symbolically and thus obtain a closed-form equation on the largest eigenvalue $\rho$.

- The second (numerical) method uses iterations of operator $\Psi$. It is based on the following fact on positive operators from [9]:

  **Proposition 17.** *Let $v_n = \Psi^n 1$, and $\alpha = \min_{q,x} \frac{v_{n+1}(q,x)}{v_n(q,x)}$; $\beta = \max_{q,x} \frac{v_{n+1}(q,x)}{v_n(q,x)}$. Then $\alpha \leq \rho(\Psi) \leq \beta$.*

  The bounds $\alpha$ and $\beta$ can be obtained by a straightforward symbolic computation.

## 6 Size versus information

**Theorem 18.** *Let $L$ be a timed language of a good automaton, $\alpha$ and $\mathcal{H}$ its mean dimension and $v$-entropy. If $\alpha > 0$ and $\mathcal{H} > -\infty$, then the formal $\varepsilon$-entropy of $L_n$ satisfies the inequality, for all $\eta > 0$, for $\varepsilon$ small enough and $n$ large enough:*

$$n(-\alpha \log \varepsilon + \mathcal{H} - \eta) \leq \mathfrak{h}_\varepsilon(L_n) \leq n(-\alpha \log \varepsilon + \mathcal{H} + \eta).$$

The second term of this inequality is logarithm of $\sum_m \mathcal{V}(L_n^m)\varepsilon^{-m}$ where the sum is computed over all the dimensions $m$. The first and third terms are close to the logarithm of the similar sum computed only for terms with $m \approx \alpha n$. The proof of the upper bound in Theorem 18 is based on technical estimates showing that contribution of terms with $m < \alpha n$ in the sum can be reasonably upper bounded.

# 7 Conclusions

This paper reports progress achieved recently in the information-theoretical studies of timed regular languages, especially with punctual transitions. Two components of information content have been identified and characterized: one represents the dimensionality, another the volume. Both are characterized by spectral radii of linear operators, matrices of both operators reproduce the structure of the automaton, but they are still very different in nature (one is max-plus and finite-dimensional, other is a classical integral operator). Our understanding of the role of punctual transitions in timed languages has substantially improved. As ongoing and future work we are interested in relating these results to other information measures, such as Kolmogorov complexity, and topological entropy, in the spirit of symbolic dynamics. Our work went in parallel and in interaction with the MSc thesis [6] supervised by one of us and exploring in depth symbolic dynamics of timed automata. These two research lines would eventually merge. On the other hand, we believe that improved measures of information content for a larger class of timed languages introduced here are more suitable for eventual implementation and applications.

## Acknowledgments

We are thankful to Dominique Perrin for encouraging discussions, to Mike Boyle for attiring our attention to mean dimension, and to Nicolas Basset for sharing his inspiring ideas on timed symbolic dynamics.

## References

[1] R. Alur and D. L. Dill. A theory of timed automata. *Theoretical Computer Science*, 126:183–235, 1994.

[2] E. Asarin and A. Degorre. Volume and entropy of regular timed languages. Preprint, 2009. http://hal.archives-ouvertes.fr/hal-00369812/.

[3] E. Asarin and A. Degorre. Volume and entropy of regular timed languages: Analytic approach. In *FORMATS'09*, LNCS 5813, pages 13–27, 2009.

[4] E. Asarin and A. Degorre. Volume and entropy of regular timed languages: Discretization approach. In *Concur'09*, LNCS 5710, pages 69–83, 2009.

[5] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat. *Synchronization and Linearity.* 2001. http://www-rocq.inria.fr/metalau/cohen/documents/BCOQ-book.pdf.

[6] N. Basset. Dynamique symbolique et langages temporisés. Master's thesis, Master Parisien de la Recherche Informatique, 2010.

[7] G. A. Edgar, editor. *Classics on Fractals.* Addison-Wesley, 1993.

[8] A. Kolmogorov and V. Tikhomirov. $\varepsilon$-entropy and $\varepsilon$-capacity of sets in function spaces. *Uspekhi Mat. Nauk*, 14(2):3–86, 1959. Russian, partial English translation in [7].

[9] M. Krasnosel'skij, E. Lifshits, and A. Sobolev. *Positive Linear Systems: the Method of Positive Operators.* Heldermann Verlag, Berlin, 1989.

[10] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding.* Cambridge University Press, 1995.

[11] E. Lindenstrauss and B. Weiss. Mean topological dimension. *Israel J. of Math.*, 115:1–24, 2000.

[12] E. C. Posner and E. R. Rodemich. Epsilon entropy and data compression. *Ann. Math. Statist*, 42(6):2079–2125, 1971.

[13] B. Rynne and M. Youngson. *Linear Functional Analysis*. Springer, 2008.

# A Appendix: some proofs

## A.1 Non-critical paths and dimension defect

*Proposition* (6). For any path $\pi$ in a matrix of $\{-\infty, 0, 1\}^{n \times n}$ with $m$ non-critical and $k$ critical edges, the following upper bound holds:

$$w(\pi) < d + \alpha k + \beta m,$$

with some constant $\beta < \alpha$ which depends only on the matrix $A$.

*Proof.* Let $\beta$ be the maximum mean weight of a simple non-critical circuit in $A$, by definition $\beta < \alpha$. If Let $d$ be the maximum weight of an acyclic path in $A$. Consider a path $\pi$ having $m$ non-critical edges and $k$ critical edges.

Let us assume that $\pi$ is not acyclic. Then it has at least one simple cycle $c_1$ as a factor: $\exists p, s$ s.t. $\pi = pc_1 s$. We define $\pi_1 \triangleq ps$. Obviously $w(\pi) = w(\pi_1) + w(c_1)$. If $\pi_1$ is not acyclic, we can repeat the operation. That way we obtain an acyclic path $\pi_l$ and a sequence of simple cycles $c_1, \ldots, c_l$ such that $w(\pi) = w(\pi_l) + \sum_{i=1}^{l} w(c_i)$. Thus, in the case when $\beta > -\infty$, we have the following chain of inequalities (where by $|\bullet|_c$ and $|\bullet|_{nc}$, we denote the amount of resp. critical and non-critical edges in a path):

$$\begin{aligned}
w(\pi) = w(\pi_l) + \sum_{c_i \text{critical}} w(c_i) + \sum_{c_i \text{non-critical}} w(c_i) \\
\leq d + \sum_{c_i \text{critical}} \alpha |c_i| + \sum_{c_i \text{non-critical}} \beta |c_i| \\
\leq d + \sum_{c_i \text{critical}} \alpha |c_i|_c + \sum_{c_i \text{non-critical}} \beta (|c_i|_c + |c_i|_{nc}) \\
\leq d + \sum_{c_i \text{critical}} \alpha |c_i|_c + \sum_{c_i \text{non-critical}} (\alpha |c_i|_c + \beta |c_i|_{nc}) \\
\leq d + \alpha k + \sum_{c_i \text{non-critical}} \beta |c_i|_{nc} \leq d + \alpha k + \beta m.
\end{aligned}$$

In the degenerate case when $\beta = -\infty$, there is no noncritical circuits and

$$w(\pi) = w(\pi_l) + \sum_{c_i \text{critical}} w(c_i) = d + \sum_{c_i \text{critical}} \alpha |c_i| \leq d + \alpha k + 0m. \quad \square$$

## A.2 Operator $\Psi$ and $v$-entropy

### A.2.1 Basic properties

*Proposition* (13). $\Psi_c$ is a bounded linear positive operator. $\Psi_c^{D+1}$ is compact with $D$ as in Assumption A1.

Proof idea: almost the same as in [3] with slight modifications to take punctual guards into account.

*Theorem* (14). Let $\mathcal{A}$ be a good region-split TA, $\mathcal{A}_c$ its critical subautomaton, $\Psi_c$ the operator described above for $\mathcal{A}_c$, and $\rho_c$ the spectral radius of this operator. Then the following holds:

- for any $\sigma > 0$, and $n$ big enough, for any state $(q, \mathbf{x})$ of $\mathcal{A}_c$, the upper bound:

$$\mathcal{V}(L_{qn}^c(\mathbf{x})) < (\rho_c + \sigma)^n;$$

- and for some state $q$ of $\mathcal{A}_c$, some open set $O$ inside $\mathbf{r}_q$ and some $\gamma > 0$, for all $\mathbf{x} \in O$, for all $n$, the lower bound:

$$\mathcal{V}(L_{qn}^c(\mathbf{x})) > \gamma \rho_c^n.$$

*Proof.* For the first bullet: first, we write the following upper bound $\mathcal{V}(L^c_{qn}(\mathbf{x})) = (\Psi_c^n 1)(q, \mathbf{x}) \leq ||\Psi_c^n||$. Then we use Gelfand's formula (see Theorem 11) for operator $\Psi_c$: $\lim_{n\to\infty} ||\Psi_c^n||^{\frac{1}{n}} = \rho_c$. Put together, it implies $\mathcal{V}(L^c_{qn}(\mathbf{x})) = o((\rho_c + \sigma)^n)$ for $n$ large enough.

For the second bullet: we use the existence of the eigenvector $v^*$ (which follows from Prop. 13 and Thm. 11) and the fact it is smaller than $\frac{1}{\lambda}1$ for some $\lambda > 0$. The following holds: $\mathcal{V}(L^c_{qn}(\mathbf{x})) = \Psi_c^n 1_q(\mathbf{x}) > \lambda \Psi_c^n v_q^*(\mathbf{x}) = \lambda \rho_c^n v_q^*(\mathbf{x})$. Since $v^* \neq 0$, it is positive for some $(q, \mathbf{x})$, and since it is a continuous function, it is also above some constant $\gamma > 0$ on an open set $O$ containing $\mathbf{x}$. Thus, for all $\mathbf{x} \in O$, $\mathcal{V}(L^c_{qn}(\mathbf{x})) > \lambda\gamma\rho_c^n$. □

**Lemma 19.** *Let $f : \mathbf{r}_{q'} \to \mathbb{R}$, non-negative, and above some constant $c' > 0$ on an open set $O' \subseteq \mathbf{r}_{q'}$, and $\delta$ be a transition from $q$ to $q'$, then $\psi_\delta f$ is non-negative, and is above some constant $\lambda > 0$ on some open set $O \subseteq \mathbf{r}_q$.*

*Proof.* If $\delta$ is punctual (test: $y = c$): $\psi_\delta f(\mathbf{x}) = f(\mathfrak{r}(\mathbf{x} + c - y))$. But the set of values of $\mathbf{x}$ such that $\mathfrak{r}(\mathbf{x} + c - y) \in O'$ is an open set too, and the value of $\psi_\delta f(\mathbf{x})$ for such $\mathbf{x}$'s is above $\lambda$.

If $\delta$ is not punctual (guard: region $\mathfrak{g}$): $\psi_\delta f(\mathbf{x}) = \int_{\mathbf{x}+\tau\in\mathfrak{g}} f(\mathfrak{r}(\mathbf{x}+\tau))d\tau$. Here, by usual properties of regions, for some $\mathbf{x}_0 \in \mathbf{r}_q, \tau_0 \geq 0$, we have that $\mathfrak{r}(\mathbf{x}+\tau) \in O'$ and thus $f(\mathfrak{r}(\mathbf{x}+\tau)$. An integral of a non-negative continuous function which takes some positive value is positive, thus $\psi_\delta f(\mathbf{x}_0) > 0$. By continuity of $\psi_\delta f$, there is an open $\Omega \ni \mathbf{x}_0$, and a constant $\lambda > 0$ such that $\forall \mathbf{x} \in \Omega\ \psi_\delta f(\mathbf{x}) > \lambda$. □

### A.2.2 Path and language decomposition

In the proofs of theorems 15 and 18 we will use the following method:

- Let $\rho = \max_c \rho_c$ the maximal spectral radius of all the $\Psi_c$. We fix some $\sigma$ and find $n_0$ such that for any critical path of length $n \geq n_0$ (in any SCC $c$) it holds that

  $$||\Psi_c||^n \leq \rho + \sigma^n. \tag{5}$$

  This is possible due to Gelfand formula.

- We split every path in the automaton into green, and yellow parts as follows:

  **Definition 20.** In a path, say a transition is *green* if it belongs a contiguous block of at least $n_0$ critical transitions. We say it is *yellow* otherwise.

  **Proposition 21.** *The obtained coloring has the following properties:*

  – *for the path (of length $n > n_0$) with $m$ non-critical transitions, there are at most $2n_0m$ yellow ones;*

  – *the green part consists of several contiguous fragments of lengths $\geq n_0$, each such fragments has only critical transitions, moreover, it stays in some critical subautomaton $\mathcal{A}_c$.*

  Indeed, the yellow part belongs to $n_0$-neighborhood of non-critical transitions.

- We split the set of all the paths (of some length $n$) in the automaton, according to:

  – which positions are yellow, which are green,

  – and which transitions are taken at each yellow position in the path.

  We formalize it as follows:

  **Definition 22.** A *path pattern* is a mapping $\Gamma : \{1 \ldots n\} \to \Delta \cup \{\perp\}$. Its length is $n$, its weight (number of yellow transitions) is $\#\{i | \Gamma(i) \neq \perp\}$. The pattern is $n_0$-regular if all $\perp$ values are grouped in contiguous blocks of size $\geq n_0$.

  **Definition 23.** A path $\pi$ of length $n$ *conforms* to a path pattern $\Gamma$ of size $n$, if

– for all $i$ with $\Gamma(i) \neq \perp$, the transition is as specified by the pattern: $\pi[i] = \Gamma(i)$;
– for all $i$ with $\Gamma(i) = \perp$, the transition $\pi[i]$ is critical.

We denote the set of $\pi$ conforming to $\Gamma$ by $P_\Gamma$, and the sum of the volumes of such paths by $V_\Gamma$.

- The number of path patterns can be upper bounded by simple combinatorics:

  **Proposition 24.** *The number of path patterns of size $n$ and weight $k$ does not exceed $C_n^k|\Delta|^k$.*

- The volume corresponding to a regular path pattern can be upper bounded using (5), but this requires some work.

  **Proposition 25.** *If $\Gamma$ is an $n_0$-regular path pattern of length $n$ and weight $k$, and starting from the initial state $q_0$, then $V_\Gamma < (\rho+\sigma)^{n-k}$.*

  *Proof.* The volume of one path conforming to $\Gamma$ is $(\psi_{\delta_1}\psi_{\delta_2}\dots\psi_{\delta_n}1)(q_0, 0)$, so taking all such $\pi \in P_\Gamma$, the volume of the union is

  $$V_\Gamma = ([\psi_{\delta_1 1}]\dots[\psi\delta_{1i_1}][\Psi_{c_1}^{k_1}][\psi_{\delta_{21}}]\dots[\psi_{\delta_{2i_2}}][\Psi_{c_2}^{k_2}]\dots 1)(q_0, 0),$$

  where the $\delta_{ij}$ are the non-critical transitions, and the $c_i$ the critical SCC that are visited (during $k_i \geq n_0$ transitions) by this set of paths, and for an operator $\psi$ applying to a subset of the state space, $[\psi]$ denotes the same operator lifted to the whole state space of $\mathcal{A}$. For every $\delta$, $||\psi_\delta|| \leq 1$, so

  $$V_\Gamma \leq \prod_i ||\Psi_{c_i}^{k_i}|| \leq \prod_i (\rho_{c_i}+\sigma)^{k_i} \leq (\rho+\sigma)^{n-k}. \quad \square$$

### A.2.3 Characterization of $\mathcal{H}$

*Theorem* (15). Let $\mathcal{A}$ be a good region-split TA. Then its $v$-entropy $\mathcal{H}$ is the maximum of $\log\rho_c$ over all its critical subautomata $\mathcal{A}_c$.

*Proof.* As before let $\rho = \max_c \rho_c$, let $\sigma > 0$ be an arbitrary small number and $n_0$ as in Prop. 25.

**Proof of $\mathcal{H} \leq \log\rho$.** By definition 2,

$$\mathcal{H} = \lim_{n\to\infty} \log\max\{\mathcal{V}(L_n^m) \mid \alpha n - d < m < \alpha n + d\}/n.$$

Let us fix $m \in [\alpha n - d; \alpha n + d]$. All the accepting paths for words in $L_n^m$, according to Prop. 6, have at most $d_1 = 2d/(\alpha-\beta)$ non-critical transitions, and their coloring, by Prop. 21 involves no more than $d_2 = 2n_0 d_1$ yellow transitions. Thus, such a path conforms to some $n_0$-regular path pattern $\Gamma$ of size $n$ and weight not exceeding $d_2$. Then we have $\mathcal{V}(L_n^m) = \sum_{\text{relevant }\Gamma} V_\Gamma$. By Prop. 24 there is at most $C_n^k|\Delta|^k \leq (n|\Delta|)^k$ potential $\Gamma$s for each $k \leq d_2$; and in total

$$\#\{\text{relevant }\Gamma\} = O(n^{d_2+1}).$$

For each such $\Gamma$ of a weight $k \leq d_2$ we apply Prop. 25 to estimate the volume of its paths:

$$V_\Gamma < (\rho+\sigma)^{n-k} < (\rho+\sigma)^n.$$

Multiplying two last estimates we obtain:

$$\mathcal{V}(L_n^m) = \sum_{\text{relevant}\Gamma} V_\Gamma < (\rho+\sigma)^n O(n^{d_2+1}) < (\rho+2\sigma)^n$$

for $n$ large enough.

Therefore (since it holds for any $\sigma > 0$) we conclude $\mathcal{H} \leq \log(\rho)$.

**Proof of $\mathcal{H} \geq \log \rho$.** We apply the second bullet of Thm. 14 to an SCC realizing the maximal $\rho_c = \rho$. There exists $q$ in that component and an open $O \subseteq \mathbf{r}_q$ such that for all $\mathbf{x} \in O$, $\mathcal{V}(L^c_{qn}(\mathbf{x})) > \gamma \rho_c^n$. But this state is reachable from $(q_0, 0)$, using some simple path $\pi = \delta_1 \dots \delta_k$. Consider a sublanguage $L'$ of traces of all the runs beginning by $\pi$ and staying in SCC $c$ afterwards. The dimension along such paths satisfies $m > \alpha n - d$. Obviously, for one of such $m$, $\mathcal{V}(L^m_{q_0 n}) \geq \mathcal{V}(L'_n)^m \geq \psi_{\delta_1} \dots \psi_{\delta_k} \mathcal{V}(L_{q(n-k)}) \geq \psi_{\delta_1} \dots \psi_{\delta_k} \gamma \rho^{n-k} 1_O$

By applying $k$ times Lemma 19, we get that there is some open on which $\psi_{\delta_1} \dots \psi_{\delta_k} 1_O$ is above some constant $c > 0$ (but as this function is defined on a singleton, that means it is above $c$ on $(q_0, 0)$). As $\psi_{\delta_1} \dots \psi_{\delta_k}$ is a linear operator, it also implies that $\mathcal{V}(L^m_{q_0 n}) \geq c\gamma\rho^{n-k}$. If we take the limit of the logarithm, this yields the inequality we wanted to prove: $\mathcal{H} \geq \log \rho$. □

## A.3 Information identity

In this section we will prove Thm 18 using the same method, based on path coloring and path patterns, as the proof of Thm 15 above. However, here we have to consider paths of all dimensions, non necessarily $m \approx \alpha n$ which makes computations more difficult.

Combining propositions 6 and 21 we relate the number of yellow transition in a path with its dimension defect:

**Corollary 26.** *There exists $K$ such that for any long enough path,*

$$\# \text{ yellows} \leq \frac{d + \alpha n - m}{K};$$
$$\# \text{ greens} \geq n - \frac{d + \alpha n - m}{K}.$$

We will also rephrase Prop. 25 as follows:

**Corollary 27.** *For a path pattern $\Gamma$ of length $n > n_0$ and weight (number of yellows) $k$, the volume $V_\Gamma$ of conforming paths can be bounded is smaller than $2^{(n-k)(\mathcal{H}+\eta)}$.*

*Proof.* We just have to chose $\sigma$ in Prop. 25 small enough (and thus $n_0$ large enough) so that $\log(\rho + \sigma) < \mathcal{H} + \eta = \log \rho + \eta$. Then we get the desired upper bound. □

We will use the following simple fact:

**Proposition 28.** *The function $x \mapsto \left(\frac{\lambda}{x}\right)^x$ is increasing on $[0, \frac{\lambda}{e}]$ (with convention $\left(\frac{1}{0}\right)^0 = 1$), has a maximal value $e^{\frac{\lambda}{e}}$, realized for $x = \frac{\lambda}{e}$, and is decreasing on $[\frac{\lambda}{e}, \infty)$.*

*Proof.* The logarithm of that function is $x \mapsto x \log \lambda - x \log x$. Its derivative is $x \mapsto \log \lambda - 1 - \log x$, positive for $x < \frac{\lambda}{e}$, zero realized for $x = \frac{\lambda}{e}$, and negative for $x > \frac{\lambda}{e}$.

This proves that the maximum of $x \mapsto \left(\frac{1}{x}\right)^x$ is realized for $x = \frac{\lambda}{e}$. And by replacing $x$, we obtain the maximum value: $e^{\frac{\lambda}{e}}$. □

Now we can prove the main theorem on information in timed languages:

*Theorem* (18). Let $L$ be a timed language of a good automaton, $\alpha$ and $\mathcal{H}$ its mean dimension and $v$-entropy. If $\alpha > 0$ and $\mathcal{H} > -\infty$, then the formal $\varepsilon$-entropy of $L_n$ satisfies the inequality, for all $\eta > 0$, for $\varepsilon$ small enough and $n$ large enough:

$$n(-\alpha \log \varepsilon + \mathcal{H} - \eta) \leq \mathfrak{h}_\varepsilon(L_n) \leq n(-\alpha \log \varepsilon + \mathcal{H} + \eta).$$

*Proof.* **The lower bound** $\mathfrak{h}_\varepsilon(L_n) \geq n(-\alpha \log \varepsilon + \mathcal{H} + r)$ is easy. Indeed:

$$\begin{aligned}
\mathfrak{h}_\varepsilon(L_n) &= \log \sum_{m=0}^{\alpha n+d} \mathcal{V}(L_n^m)\varepsilon^{-m} \\
&\geq \log \max\{\mathcal{V}(L_n^m)\varepsilon^{-m} \mid 0 \leq m \leq \alpha n + d\} \\
&\geq \log \max\{\mathcal{V}(L_n^m)\varepsilon^{-m} \mid \alpha n - d \leq m \leq \alpha n + d\} \\
&\geq \log \varepsilon^{-(\alpha n - d)} \max\{\mathcal{V}(L_n^m) \mid \alpha n - d \leq m \leq \alpha n + d\} \\
&\geq n\left(-\alpha \log \varepsilon + \frac{1}{n} \log \max\{\mathcal{V}(L_n^m) \mid \alpha n - d \leq m \leq \alpha n + d\} + \frac{d}{n} \log \varepsilon\right).
\end{aligned}$$

By definition of $v$-entropy, $\lim_{n\to\infty} \frac{1}{n} \log \max\{\mathcal{V}(L_n^m) \mid \alpha n - d \leq m \leq \alpha n + d\} = \mathcal{H}(L)$, thus for $n$ large enough:

$$\frac{1}{n} \log \max\{\mathcal{V}(L_n^m) \mid \alpha n > \mathcal{H}(L) - \eta/2.$$

Similarly, for any $\varepsilon$, for $n$ large enough $\frac{d}{n} \log \varepsilon > -\eta/2$.

Therefore $\mathfrak{h}_\varepsilon(L_n) \geq n(-\alpha \log \varepsilon + \mathcal{H}(L) - \eta)$.

**The upper bound** for $\mathfrak{h}_\varepsilon(L_n)$ is more involved. First we write down the formula for the entropy (from definition 1), and split the sum into two parts:

$$2^{\mathfrak{h}_\varepsilon(L_n)} = \sum_{m=0}^{\alpha n} \mathcal{V}(L_n^m)\varepsilon^{-m} = \sum_{m=0}^{\alpha n - \gamma n} \mathcal{V}(L_n^m)\varepsilon^{-m} + \sum_{m=\alpha n - \gamma n}^{\alpha n} \mathcal{V}(L_n^m)\varepsilon^{-m} = S_1 + S_2,$$

where $\gamma$ $(0 \leq \gamma \leq \alpha)$, which will be chosen later, separates the sum between the low dimension paths (in $S_1$) and those whose dimension is close to that of the critical paths (in $S_2$).

First we look at $S_1$, that is the lower dimension paths.

Because guards in the region-split automaton have length at most 1, then $|\Sigma|^n$ is always an upper bound to $\mathcal{V}(L_n^m)$. Thus, for $m < (\alpha - \gamma)n$, we have that $\mathcal{V}(L_n^m)\varepsilon^{-m} \leq |\Sigma|^n \varepsilon^{-(\alpha-\gamma)n}$, which implies:

$$S1 \leq (\alpha - \gamma)n|\Sigma|^n \varepsilon^{-(\alpha-\gamma)n}.$$

Dividing by the main term $2^{n(\mathcal{H}+\eta)}\varepsilon^{-\alpha n}$, we obtain the bound

$$\frac{S_1}{2^{n(\mathcal{H}+\eta)}\varepsilon^{-\alpha n}} \leq (\alpha - \gamma)n \left(\frac{\varepsilon^\gamma |\Sigma|}{2^{\mathcal{H}+\eta}}\right)^n.$$

This term is $O(1)$ for $\varepsilon$ small enough.

Now, we look at $S_2$, the volume of the paths of higher dimension.

For the higher dimension paths:

$$
\begin{aligned}
S_2 &= \sum_{m=\alpha n-\gamma n}^{\alpha n} \sum_{y \le \frac{d+\alpha n-m}{K}} (\text{vol. paths with } y \text{ yellows}) \varepsilon^{-m} \\
&= \sum_{m} \sum_{y \le \frac{d+\alpha n-m}{K}} (\#\text{ patterns } \Gamma \text{ of weight } y) V_\Gamma \varepsilon^{-m} \\
&\le \sum_{m} \sum_{y \le \frac{d+\alpha n-m}{K}} C_n^y |\Delta|^y 2^{(n-y)(\mathcal{H}+\eta)} \varepsilon^{-m} && \text{(Prop. 24, Cor. 27)} \\
&\le 2^{n(\mathcal{H}+\eta)} \varepsilon^{-\alpha n} \sum_m \varepsilon^{\alpha n-m} \sum_{y \le \frac{d+\alpha n-m}{K}} C_n^y \left(\frac{|\Delta|}{2^{\mathcal{H}+\eta}}\right)^y && \left(\text{using } C_n^k \le \left(\frac{ne}{k}\right)^k\right) \\
&\le 2^{n(\mathcal{H}+\eta)} \varepsilon^{-\alpha n} \sum_m \varepsilon^{\alpha n-m} \sum_{y \le \frac{d+\alpha n-m}{K}} \left(\frac{\frac{ne}{y}|\Delta|}{2^{\mathcal{H}+\eta}}\right)^y \\
&\le 2^{n(\mathcal{H}+\eta)} \varepsilon^{-\alpha n} \sum_m \varepsilon^{\alpha n-m} \sum_{y \le \frac{d+\alpha n-m}{K}} \left(\frac{An}{y}\right)^y && \text{(A: constant).}
\end{aligned}
$$

To upper bound the inner sum we replace it by (the number of terms)$\times$(the last term). This is legitimate since the (by virtue of Prop. 28) the summand is increasing if $y \le \frac{An}{e}$, which is true if we choose $\gamma \le \frac{KA}{e}$.

$$
\begin{aligned}
\frac{S_2}{2^{n(\mathcal{H}+\eta)}\varepsilon^{-\alpha n}} &\le \sum_{m \ge \alpha n-\gamma n} \varepsilon^{\alpha n-m} \sum_{y \le \frac{d+\alpha n-m}{K}} \left(\frac{An}{y}\right)^y \\
&\le \sum_{m \ge \alpha n-\gamma n} \varepsilon^{-d} \varepsilon^{d+\alpha n-m} \left(\frac{d+\alpha n-m}{K}\right) \left(\frac{KAn}{d+\alpha n-m}\right)^{\frac{\alpha n-m}{K}} \\
&\le \varepsilon^{-d} \sum_{s=0}^{\gamma n} \frac{s}{K} \left(\frac{KAn\varepsilon^K}{s}\right)^{\frac{s}{K}} \\
&\le \varepsilon^{-d} \sum_{s=0}^{\gamma n} \frac{s}{K} e^{\frac{KAn\varepsilon^K}{eK}} \quad \text{(bound from Prop. 28)} \\
&\le \varepsilon^{-d} \frac{(\gamma n)^2}{K} e^{\frac{An\varepsilon^K}{e}}.
\end{aligned}
$$

$\varepsilon$ can be chosen small enough so that $e^{\frac{An\varepsilon^K}{e}}$ grows more slowly than $2^{\eta n}$.

To sum it up, for every $\eta > 0$, for $\varepsilon > 0$ small enough, and for $n$ large enough (depending on $\varepsilon$),

$$2^{\mathfrak{h}_\varepsilon(L_n)} = S_1 + S_2 \le \left(O(1) + O(\varepsilon^{-d} 2^{\eta n})\right) 2^{n(\mathcal{H}+\eta)} \varepsilon^{-\alpha n} \le 2^{n(\mathcal{H}+3\eta)} \varepsilon^{-\alpha n},$$

and by taking logarithm we obtain the desired bound. $\square$