Profinite Techniques for Probabilistic Automata and the Markov Monoid Algorithm*

Nathanaël Fijalkow

University of Oxford, United Kingdom

Abstract. We consider the value 1 problem for probabilistic automata over finite words: it asks whether a given probabilistic automaton accepts words with probability arbitrarily close to 1. This problem is known to be undecidable. However, different algorithms have been proposed to partially solve it; it has been recently shown that the Markov Monoid algorithm, based on algebra, is the most correct algorithm so far. The first contribution of this paper is to give a characterisation of the Markov Monoid algorithm.
The second contribution is to develop a profinite theory for probabilistic automata, called the prostochastic theory. This new framework gives a topological account of the value 1 problem, which in this context is cast as an emptiness problem. The above characterisation is reformulated using the prostochastic theory, allowing us to give a simple and modular proof.

Keywords: Probabilistic Automata; Profinite Theory; Topology

1 Introduction

Rabin [11] introduced the notion of probabilistic automata, which are finite automata with randomised transitions. This powerful model has been widely studied since then and has applications, for instance in image processing, computational biology and speech processing. This paper follows a long line of work that studies the algorithmic properties of probabilistic automata. We consider the value 1 problem: it asks, given a probabilistic automaton, whether there exist words accepted with probability arbitrarily close to 1.

This problem has been shown undecidable [7]. Different approaches led to construct subclasses of probabilistic automata for which the value 1 problem is decidable; the first class was $\sharp$-acylic automata [7], then concurrently simple automata [2] and leaktight automata [5]. It has been shown in [4] that the so-called Markov Monoid algorithm introduced in [5] is the most correct algorithm of the three algorithms. Indeed, both $\sharp$-acylic and simple automata are strictly subsumed by leaktight automata, for which the Markov Monoid algorithm correctly solves the value 1 problem.

* A preliminary version appeared in the proceedings of STACS'2016 [3]. This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1.

Yet we were missing a good understanding of the computations realised by the Markov Monoid algorithm. The aim of this paper is to provide such an insight by giving a characterisation of this algebraic algorithm. We show the existence of *convergence speeds* phenomena, which can be polynomial or exponential. Our main technical contribution is to prove that the Markov Monoid algorithm captures exactly *polynomial behaviours*.

Proving this characterisation amounts to giving precise bounds on convergences of non-homogeneous Markov chains. Our second contribution is to define a new framework allowing us to rephrase this characterisation and to give a modular proof for it, using techniques from topology and linear algebra. We develop a profinite approach for probabilistic automata, called prostochastic theory. This is inspired by the profinite approach for (classical) automata [1,10,6], and for automata with counters [12].

Section 3 is devoted to defining the Markov Monoid algorithm and stating the characterisation: it answers "YES" if, and only if, the probabilistic automaton accepts some polynomial sequence.

In Section 4, we introduce a new framework, the prostochastic theory, which is used to restate and prove the characterisation. We first construct a space called the free prostochastic monoid, whose elements are called prostochastic words. We define the acceptance of a prostochastic word by a probabilistic automaton, and show that the value 1 problem can be reformulated as the emptiness problem for probabilistic automata over prostochastic words. We then explain how to construct non-trivial prostochastic words, by defining a limit operator ω, leading to the definition of polynomial prostochastic words. The above characterisation above reads in the realm of prostochastic theory as follows: the Markov Monoid algorithm answers "YES" if, and only if, the probabilistic automaton accepts some polynomial prostochastic word.

Section 5 concludes by showing how this characterisation, combined with an improved undecidability result, supports the claim that the Markov Monoid algorithm is *in some sense* optimal.

Acknowledgments

This paper and its author owe a lot to Szymon Toruńczyk's PhD thesis and its author, to Sam van Gool for his expertise on Profinite Theory, to Mikołaj Bojańczyk for his insightful remarks and to Jean-Éric Pin for his numerous questions and comments. The opportunity to present partial results on this topic in

several scientific meetings has been a fruitful experience, and I thank everyone that took part in it. Last but not least, the reviewers greatly participated in improving the quality of this paper.

2 Probabilistic Automata and the Value 1 Problem

Let Q be a finite set of states.

A (probability) distribution over Q is a function $\delta : Q \to [0,1]$ such that $\sum_{q \in Q} \delta(q) = 1$. We denote $\mathcal{D}(Q)$ the set of distributions over Q, which we often consider as vectors indexed by Q.

For $E \subseteq \mathbb{R}$, we denote $\mathcal{M}_{Q \times Q}(E)$ the set of (square) matrices indexed by Q over E. We denote I the identity matrix. A matrix $M \in \mathcal{M}_{Q \times Q}(\mathbb{R})$ is stochastic if each row is a distribution over Q; the subset consisting of stochastic matrices is denoted $\mathcal{S}_{Q \times Q}(E)$. The space $\mathcal{S}_{Q \times Q}(\mathbb{R})$ is equipped with the norm $|| \cdot ||$ defined by

$$||M|| = \max_{s \in Q} \sum_{t \in Q} |M(s,t)|.$$

This induces the standard Euclidean topology on $\mathcal{S}_{Q \times Q}(\mathbb{R})$. The following classical properties will be useful:

Fact 1 *(Topology of the stochastic matrices)*

- *For all matrix $M \in \mathcal{S}_{Q \times Q}(\mathbb{R})$, we have $||M|| = 1$,*
- *For all matrices $M, M' \in \mathcal{M}_{Q \times Q}(\mathbb{R})$, we have $||M \cdot M'|| \leq ||M|| \cdot ||M'||$,*
- *The monoid $\mathcal{S}_{Q \times Q}(\mathbb{R})$ is (Hausdorff) compact.*

Definition 1. *(Probabilistic automaton) A probabilistic automaton $\mathcal{A}$ is given by a finite set of states Q, a transition function $\phi : A \to \mathcal{S}_{Q \times Q}(\mathbb{R})$, an initial state $q_0 \in Q$ and a set of final states $F \subseteq Q$.*

Observe that it generalises the definition for classical deterministic automata, in which transitions functions are $\phi : A \to \mathcal{S}_{Q \times Q}(\{0,1\})$. We allow here the transition functions of probabilistic automata to have arbitrary real values; when considering computational properties, we assume that they are rational numbers.

A transition function $\phi : A \to \mathcal{S}_{Q \times Q}(\mathbb{R})$ naturally induces a morphism $\phi : A^* \to \mathcal{S}_{Q \times Q}(\mathbb{R})$[1].

We denote $P_{\mathcal{A}}(s \xrightarrow{w} t)$ the probability to go from state s to state t reading w on the automaton $\mathcal{A}$, *i.e.* $\phi(w)(s,t)$. We extend the notation: for a subset T of the set of states, $P_{\mathcal{A}}(s \xrightarrow{w} T)$ is defined by $\phi(w)(s,T) = \sum_{t \in T} \phi(w)(s,t)$.

[1] Note that we use "morphism" for "monoid homomorphism" throughout the paper.

The *acceptance probability* of a word $w \in A^*$ by $\mathcal{A}$ is $P_{\mathcal{A}}(q_0 \xrightarrow{w} F)$, denoted $P_{\mathcal{A}}(w)$. In words, the above is the probability that a run starting from the initial state q_0 ends in a final state (*i.e.* a state in F). The *value* of a probabilistic automaton $\mathcal{A}$ is $\text{val}(\mathcal{A}) = \sup\{P_{\mathcal{A}}(w) \mid w \in A^*\}$, the supremum over all words of the acceptance probability.

Definition 2. *(Value* 1 *Problem) The value* 1 *problem is the following decision problem: given a probabilistic automaton* $\mathcal{A}$ *as input, determine whether* $val(\mathcal{A}) = 1$*, i.e. whether there exist words whose acceptance probability is arbitrarily close to* 1*.*

Equivalently, the value 1 problem asks for the existence of a sequence of words $(u_n)_{n\in\mathbb{N}}$ such that $\lim_n P_{\mathcal{A}}(u_n) = 1$.

3 Characterisation of the Markov Monoid Algorithm

The Markov Monoid algorithm was introduced in [5], we give here a different yet equivalent presentation. Consider a probabilistic automaton $\mathcal{A}$, the Markov Monoid algorithm consists in computing, by a saturation process, the Markov Monoid of $\mathcal{A}$.

It is a monoid of Boolean matrices: all numerical values are projected to Boolean values. So instead of considering $M \in \mathcal{S}_{Q\times Q}(\mathbb{R})$, we are interested in $\pi(M) \in \mathcal{M}_{Q\times Q}(\{0,1\})$, the Boolean matrix such that $\pi(M)(s,t) = 1$ if $M(s,t) > 0$, and $\pi(M)(s,t) = 0$ otherwise. Hence to define the Markov Monoid, one can consider the underlying non-deterministic automaton $\pi(\mathcal{A})$ instead of the probabilistic automaton $\mathcal{A}$. Formally, $\pi(\mathcal{A})$ is defined as $\mathcal{A}$, except that its transitions are given by $\pi(\phi(a))$ for the letter $a \in A$.

The Markov Monoid of $\pi(\mathcal{A})$ contains the transition monoid of $\pi(\mathcal{A})$, which is the monoid of Boolean matrix generated by $\{\pi(\phi(a)) \mid a \in A\}$. Informally speaking, the transition monoid accounts for the Boolean action of every finite word. Formally, for a word $w \in A^*$, the element $\langle w\rangle$ of the transition monoid of $\pi(\mathcal{A})$ satisfies the following: $\langle w\rangle(s,t) = 1$ if, and only if, there exists a run from s to t reading w on $\pi(\mathcal{A})$.

The Markov Monoid extends the transition monoid by introducing a new operator, the stabilisation. On the intuitive level first: let $M \in \mathcal{S}_{Q\times Q}(\mathbb{R})$, it can be interpreted as a Markov chain; its Boolean projection $\pi(M)$ represents the structural properties of this Markov chain. The stabilisation $\pi(M)^\sharp$ accounts for $\lim_n M^n$, *i.e.* the behaviour of the Markov chain M in the limit.

To give the formal definition of the stabilisation operator, we need a few more notations. As a convention, M denotes a matrix in $\mathcal{S}_{Q\times Q}(\mathbb{R})$, and m a Boolean matrix. Note that when considering stochastic matrices we compute in

the real semiring, and when considering Boolean matrices, we compute products in the Boolean semiring, leading to two distinct notions of idempotent matrices.

The following definitions mimick the notions of recurrent and transient states from Markov chain theory.

Definition 3. *(Idempotent Boolean matrix, recurrent and transient state) Let m be a Boolean matrix. It is idempotent if $m \cdot m = m$.*

Assume m is idempotent. We say that:

- *the state $s \in Q$ is m-recurrent if for all $t \in Q$, if $m(s,t) = 1$, then $m(t,s) = 1$, and it is m-transient if it is not m-recurrent,*
- *the m-recurrent states $s, t \in Q$ belong to the same recurrence class if $m(s,t) = 1$.*

Definition 4. *(Stabilisation) Let m be a Boolean idempotent matrix.*

The stabilisation of m is denoted $m^\sharp$ and defined by:

$$m^\sharp(s,t) = \begin{cases} 1 & \text{if } m(s,t) = 1 \text{ and } t \text{ is } m\text{-recurrent,} \\ 0 & \text{otherwise.} \end{cases}$$

The definition of the stabilisation matches the intuition that in the Markov chain $\lim_n M^n$, the probability to be in non-recurrent states converges to 0.

Definition 5. *(Markov Monoid) The Markov Monoid of an automaton $\mathcal{A}$ is the smallest set of Boolean matrices containing $\{\pi(\phi(a)) \mid a \in A\}$ closed under product and stabilisation of idempotents.*

On an intuitive level, a Boolean matrix in the Markov Monoid reflects the asymptotic behaviour of a sequence of finite words.

The Markov Monoid algorithm computes the Markov Monoid, and looks for *value* 1 *witnesses*:

Definition 6. *(Value 1 witness) Let $\mathcal{A}$ be a probabilistic automaton.*

A Boolean matrix m is a value 1 witness if: for all states $t \in Q$, if $m(q_0, t) = 1$, then $t \in F$.

The Markov Monoid algorithm answers "YES" if there exists a value 1 witness in the Markov Monoid, and "NO" otherwise.

Our main technical result is the following theorem, which is a characterisation of the Markov Monoid algorithm. It relies on the notion of *polynomial sequences of words*.

We define two operations for sequences of words, mimicking the operations of the Markov Monoid.

ALGORITHM 1: The Markov Monoid algorithm.

Data: A probabilistic automaton.
$\mathcal{M} \leftarrow \{\pi(\phi(a)) \mid a \in A\} \cup \{I\}$.
repeat
  if *there is* $m, m' \in \mathcal{M}$ *such that* $m \cdot m' \notin \mathcal{M}$ **then**
    add $m \cdot m'$ to $\mathcal{M}$
  end
  if *there is* $m \in \mathcal{M}$ *such that* m *is idempotent and* $m^\sharp \notin \mathcal{M}$ **then**
    add $m^\sharp$ to $\mathcal{M}$
  end
until *there is nothing to add*;
if *there is a value* 1 *witness in* $\mathcal{M}$ **then**
  return YES;
else
  return NO;
end

- the first is concatenation: given $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$, the concatenation is the sequence $(u_n \cdot v_n)_{n\in\mathbb{N}}$,
- the second is iteration: given $(u_n)_{n\in\mathbb{N}}$, its iteration is the sequence $(u_n^n)_{n\in\mathbb{N}}$; the n^{th} word is repeated n times.

Definition 7. *(Polynomial sequence) The class of polynomial sequences is the smallest class of sequences containing the constant sequences* $(a)_{n\in\mathbb{N}}$ *for each letter* $a \in A$ *and* $(\varepsilon)_{n\in\mathbb{N}}$*, closed under concatenation and iteration.*

A typical example of a polynomial sequence is $((a^n b)^n)_{n\in\mathbb{N}}$, and a typical example of a sequence which is not polynomial is $\left((a^n b)^{2^n}\right)_{n\in\mathbb{N}}$.

We proceed to our main result:

Theorem 1. *(Characterisation of the Markov Monoid algorithm) The Markov Monoid algorithm answers "YES" on input* $\mathcal{A}$ *if, and only if, there exists a polynomial sequence* $(u_n)_{n\in\mathbb{N}}$ *such that* $\lim_n P_{\mathcal{A}}(u_n) = 1$.

This result could be proved directly, without appealing to the prostochastic theory developed in the next section. The proof relies on technically intricate calculations over non-homogeneous Markov chains; the prostochastic theory allows to simplify its presentation, making it more modular. We will give the proof of Theorem 1 in Subsection 4.5, after restating it using the prostochastic theory.

A second advantage of using the prostochastic theory is to give a more natural and robust definition of polynomial sequences, which in the prostochastic theory correspond to polynomial prostochastic words.

A direct corollary of Theorem 1 is the absence of false negatives:

Corollary 1. *(No false negatives for the Markov Monoid algorithm)*
If the Markov Monoid algorithm answers "YES" on input $\mathcal{A}$, then $\mathcal{A}$ has value 1.

4 The Prostochastic Theory

In this section, we introduce the prostochastic theory, which draws from profinite theory to give a topological account of probabilistic automata. We construct the free prostochastic monoid in Subsection 4.1.

The aim of this theory is to give a topological account of the value 1 problem; we show in Subsection 4.2 that the value 1 problem can be reformulated as an emptiness problem for prostochastic words.

In Subsection 4.3 we define the notion of polynomial prostochastic words.

The Subsection 4.4 is devoted to a technical proof, about the powers of stochastic matrices.

The characterisation given in Section 3 is stated and proved in this new framework in Subsection 4.5.

4.1 The Free Prostochastic Monoid

The purpose of the prostochastic theory is to construct a (Hausdorff) compact[2] monoid $\mathcal{P}A^*$ together with an injective morphism $\iota : A^* \to \mathcal{P}A^*$, called the free prostochastic monoid, satisfying the following universal property:

"Every morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ extends uniquely to a continuous morphism $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$."

Here, by "$\widehat{\phi}$ extends ϕ" we mean $\phi = \widehat{\phi} \circ \iota$.

We give two statements about $\mathcal{P}A^*$, the first will be weaker but enough for our purposes in this paper, and the second more precise, and justifying the name "free prostochastic monoid". The reason for giving two statements is that the first avoids a number of technical points that will not play any further role, so the reader interested in the applications to the Markov Monoid algorithm may skip this second statement.

Theorem 2. *(Existence of the free prostochastic monoid – weaker statement) For every finite alphabet A, there exists a compact monoid $\mathcal{P}A^*$ and an injective morphism $\iota : A^* \to \mathcal{P}A^*$ such that every morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ extends uniquely to a continuous morphism $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$.*

[2] Following the French tradition, here by compact we mean Hausdorff compact: distinct points have disjoint neighbourhoods.

We construct $\mathcal{P}A^*$ and ι. Consider $X = \prod_{\phi:A^*\to\mathcal{S}_{Q\times Q}(\mathbb{R})} \mathcal{S}_{Q\times Q}(\mathbb{R})$, the product of several copies of $\mathcal{S}_{Q\times Q}(\mathbb{R})$, one for each morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$. An element m of X is denoted $(m(\phi))_{\phi:A^*\to\mathcal{S}_{Q\times Q}(\mathbb{R})}$: it is given by an element $m(\phi)$ of $\mathcal{S}_{Q\times Q}(\mathbb{R})$ for each morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$. Thanks to Tychonoff's theorem, the monoid X equipped with the product topology is compact.

Consider the map $\iota : A \to X$ defined by $\iota(a) = (\phi(a))_{\phi:A\to\mathcal{P}}$, it induces an injective morphism $\iota : A^* \to X$. To simplify notation, we sometimes assume that $A \subseteq X$ and denote a for $\iota(a)$.

Denote $\mathcal{P}A^* = \overline{A^*}$, the closure of A^* in X. Note that it is a compact monoid: the compactness follows from the fact that it is closed in X. By definition, an element $\overline{u}$ of $\mathcal{P}A^*$, called a *prostochastic word*, is obtained as the limit in $\mathcal{P}A^*$ of a sequence $\mathbf{u}$ of finite words. In this case we write $\lim \mathbf{u} = \overline{u}$ and say that $\mathbf{u}$ induces $\overline{u}$.

Note that by definition of the product topology on X, a sequence of finite words $\mathbf{u}$ converges in X if, and only if, for all morphisms $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, the sequence of stochastic matrices $\phi(\mathbf{u})$ converges.

We say that two converging sequences of finite words $\mathbf{u}$ and $\mathbf{v}$ are equivalent if they induce the same prostochastic word, *i.e.* if $\lim \mathbf{u} = \lim \mathbf{v}$. Equivalently, two converging sequences of finite words $\mathbf{u}$ and $\mathbf{v}$ are equivalent if, and only if, for all morphisms $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, we have $\lim \phi(\mathbf{u}) = \lim \phi(\mathbf{v})$.

Proof. We prove that $\mathcal{P}A^*$ satisfies the universal property. Consider a morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, and define $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ by $\widehat{\phi}(\overline{u}) = \lim \phi(\mathbf{u})$, where $\mathbf{u}$ is *some* sequence of finite words inducing $\overline{u}$. This is well defined and extends ϕ. Indeed, consider two equivalent sequences of finite words $\mathbf{u}$ and $\mathbf{v}$ inducing $\overline{u}$. By definition, for all $\psi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, we have $\lim \psi(\mathbf{u}) = \lim \psi(\mathbf{v})$, so in particular for ϕ this implies $\lim \phi(\mathbf{u}) = \lim \phi(\mathbf{v})$, and $\widehat{\phi}$ is well defined. Both continuity and uniqueness are clear.

We prove that $\widehat{\phi}$ is a morphism. Consider

$$D = \{(\overline{u}, \overline{v}) \in \mathcal{P}A^* \times \mathcal{P}A^* \mid \widehat{\phi}(\overline{u} \cdot \overline{v}) = \widehat{\phi}(\overline{u}) \cdot \widehat{\phi}(\overline{v})\}.$$

To prove that $\widehat{\phi}$ is a morphism, we prove that $D = \mathcal{P}A^* \times \mathcal{P}A^*$. First of all, $A^* \times A^* \subseteq D$. Since $A^* \times A^*$ is dense in $\mathcal{P}A^* \times \mathcal{P}A^*$, it suffices to show that D is closed. This follows from the continuity of both product functions in $\mathcal{P}A^*$ and in $\mathcal{S}_{Q\times Q}(\mathbb{R})$ as well as of $\widehat{\phi}$. ∎

We give a second, stronger statement about $\mathcal{P}A^*$, which in particular justifies the name "free prostochastic monoid".

From now on, by "monoid" we mean "compact topological monoids". The term topological means that the product function is continuous:

$$\begin{aligned} \mathcal{P} \times \mathcal{P} &\to \mathcal{P} \\ (s,t) &\mapsto s \cdot t \end{aligned}$$

A monoid is profinite if any two elements can be distinguished by a continuous morphism into a finite monoid, *i.e.* by a finite automaton. (Formally speaking, this is the definition of residually finite monoids, which coincide with profinite monoids for compact monoids, see [1].) To define prostochastic monoids, we use a stronger distinguishing feature, namely probabilistic automata.

Definition 8. *(Prostochastic monoid) A monoid $\mathcal{P}$ is prostochastic if for all elements $s \neq t$ in $\mathcal{P}$, there exists a continuous morphism $\psi : \mathcal{P} \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ such that $\psi(s) \neq \psi(t)$.*

There are many more prostochastic monoids than profinite monoids. Indeed, $\mathcal{S}_{Q\times Q}(\mathbb{R})$ is prostochastic, but not profinite in general.

The following theorem extends Theorem 2. The statement is the same as in the profinite theory, replacing "profinite monoid" by "prostochastic monoid".

Theorem 3. *(Existence of the free prostochastic monoid – stronger statement) For every finite alphabet A,*

1. *There exists a prostochastic monoid $\mathcal{P}A^*$ and an injective morphism $\iota : A^* \to \mathcal{P}A^*$ such that every morphism $\phi : A^* \to \mathcal{P}$, where $\mathcal{P}$ is a prostochastic monoid, extends uniquely to a continuous morphism $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{P}$.*
2. *All prostochastic monoids satisfying this universal property are homeomorphic.*

The unique prostochastic monoid satisfying the universal property stated in item 1. is called the free prostochastic monoid, and denoted $\mathcal{P}A^*$.

Proof. We prove that $\mathcal{P}A^*$ satisfies the stronger universal property, along the same lines as for the weaker one. Consider a morphism $\phi : A^* \to \mathcal{P}$, and define $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{P}$ by $\widehat{\phi}(\overline{u}) = \lim \phi(\mathbf{u})$, where $\mathbf{u}$ is *some* sequence of finite words inducing $\overline{u}$.

To see that this is well defined, we use the fact that $\mathcal{P}$ is prostochastic. Consider two equivalent sequences of finite words $\mathbf{u}$ and $\mathbf{v}$ inducing $\overline{u}$. Consider a continuous morphism $\psi : \mathcal{P} \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, the composition $\psi \circ \phi$ is a continuous morphism from A^* to $\mathcal{S}_{Q\times Q}(\mathbb{R})$, so since $\mathbf{u}$ and $\mathbf{v}$ are equivalent it follows that

$\lim(\psi\circ\phi)(\mathbf{u}) = \lim(\psi\circ\phi)(\mathbf{v})$, *i.e.* $\lim\psi(\phi(\mathbf{u})) = \lim\psi(\phi(\mathbf{v}))$. Since ψ is continuous, this implies $\psi(\lim\phi(\mathbf{u})) = \psi(\lim\phi(\mathbf{v}))$. We proved that for all continuous morphisms $\psi : \mathcal{P} \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, we have $\psi(\lim\phi(\mathbf{u})) = \psi(\lim\phi(\mathbf{v}))$; since $\mathcal{P}$ is prostochastic, it follows that $\lim\phi(\mathbf{u}) = \lim\phi(\mathbf{v})$, and $\widehat{\phi}$ is well defined.

Clearly $\widehat{\phi}$ extends ϕ. Both continuity and uniqueness are clear. We prove that $\widehat{\phi}$ is a morphism. Consider

$$D = \{(\overline{u}, \overline{v}) \in \mathcal{P}A^* \times \mathcal{P}A^* \mid \widehat{\phi}(\overline{u}\cdot\overline{v}) = \widehat{\phi}(\overline{u})\cdot\widehat{\phi}(\overline{v})\}.$$

To prove that $\widehat{\phi}$ is a morphism, we prove that $D = \mathcal{P}A^* \times \mathcal{P}A^*$. First of all, $A^* \times A^* \subseteq D$. Since $A^* \times A^*$ is dense in $\mathcal{P}A^* \times \mathcal{P}A^*$, it suffices to show that D is closed. This follows from the continuity of both product functions in $\mathcal{P}A^*$ and in $\mathcal{P}$ as well as of $\widehat{\phi}$.

We prove that $\mathcal{P}A^*$ is prostochastic. Let $\overline{u} \neq \overline{v}$ in $\mathcal{P}A^*$. Consider two sequences of finite words $\mathbf{u}$ and $\mathbf{v}$ inducing respectively $\overline{u}$ and $\overline{v}$, there exists a morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ such that $\lim\phi(\mathbf{u}) \neq \lim\phi(\mathbf{v})$. Thanks to the universal property proved in the first point, this induces a continuous morphism $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ such that $\widehat{\phi}(\mathbf{u}) \neq \widehat{\phi}(\mathbf{v})$, finishing the proof that $\mathcal{P}A^*$ is prostochastic.

We now prove that there is a unique prostochastic monoid satisfying the universal property, up to homeomorphism. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two prostochastic monoids satisfying the universal property, together with two injective morphisms $\iota_1 : A^* \to \mathcal{P}_1$ and $\iota_2 : A^* \to \mathcal{P}_2$. Thanks to the universal property, ι_1 and ι_2 are extended to continuous morphisms $\widehat{\iota_1} : \mathcal{P}_2 \to \mathcal{P}_1$ and $\widehat{\iota_2} : \mathcal{P}_1 \to \mathcal{P}_2$, and $\widehat{\iota_1}\circ\iota_2 = \iota_1$ and $\widehat{\iota_2}\circ\iota_1 = \iota_2$. This implies that $\widehat{\iota_1}\circ\widehat{\iota_2}\circ\iota_1 = \iota_1$; thanks to the universal property again, there exists a unique continuous morphism θ such that $\theta\circ\iota_1 = \iota_1$, and since both $\widehat{\iota_1}\circ\widehat{\iota_2}$ and the identity morphism on $\mathcal{P}_1$ satisfy this equality, it follows that they are equal. Similarly, $\widehat{\iota_2}\circ\widehat{\iota_1}$ is equal to the identity morphism on $\mathcal{P}_2$. It follows that $\widehat{\iota_1}$ and $\widehat{\iota_2}$ are mutually inverse homeomorphisms between $\mathcal{P}_1$ and $\mathcal{P}_2$. ∎

Remark 1. We remark that the free prostochastic monoid $\mathcal{P}A^*$ contains the free profinite monoid $\widehat{A^*}$. To see this, we start by recalling some properties of $\widehat{A^*}$, which is the set of *converging* sequences up to *equivalence*, where:

- a sequence of finite words $\mathbf{u}$ is converging if, and only if, for every deterministic automaton $\mathcal{A}$, the sequence is either ultimately accepted by $\mathcal{A}$ or ultimately rejected by $\mathcal{A}$, *i.e.* there exists $N \in \mathbb{N}$ such that either for all $n \geq N$, the word u_n is accepted by $\mathcal{A}$, or for all $n \geq N$, the word u_n is rejected by $\mathcal{A}$,

– two sequences of finite words $\mathbf{u}$ and $\mathbf{v}$ are equivalent if for every deterministic automaton $\mathcal{A}$, either both sequences are ultimately accepted by $\mathcal{A}$, or both sequences are ultimately rejected by $\mathcal{A}$.

Clearly:

– if a sequence of finite words is converging with respect to $\mathcal{P}A^*$, then it is converging with respect to $\widehat{A^*}$, as deterministic automata form a subclass of probabilistic automata,
– if two sequences of finite words are equivalent with respect to $\mathcal{P}A^*$, then they are equivalent with respect to $\widehat{A^*}$.

Each profinite word induces at least one prostochastic word: by compactness of $\mathcal{P}A^*$, each sequence of finite words $\mathbf{u}$ contains a converging subsequence with respect to $\mathcal{P}A^*$. This defines an injection from $\widehat{A^*}$ into $\mathcal{P}A^*$. In particular, this implies that $\mathcal{P}A^*$ is uncountable. Since it is defined as the topological of a countable set, it has the cardinality of the continuum.

4.2 Reformulation of the Value 1 Problem

The aim of this subsection is to show that the value 1 problem, which talks about sequences of finite words, can be reformualted as an emptiness problem over prostochastic words.

Definition 9. *(Prostochastic language of a probabilistic automaton) Let $\mathcal{A}$ be a probabilistic automaton, ϕ is the transition function of $\mathcal{A}$. The prostochastic language of $\mathcal{A}$ is:*

$$L(\mathcal{A}) = \{\overline{u} \mid \widehat{\phi}(\overline{u})(q_0, F) = 1\}.$$

We say that $\mathcal{A}$ accepts a prostochastic word $\overline{u}$ if $\overline{u} \in L(\mathcal{A})$.

Theorem 4. *(Reformulation of the value 1 problem) Let $\mathcal{A}$ be a probabilistic automaton. The following are equivalent:*

– *val*$(\mathcal{A}) = 1$,
– $L(\mathcal{A})$ *is non-empty.*

Proof. Assume val$(\mathcal{A}) = 1$, then there exists a sequence of words $\mathbf{u}$ such that $\lim P_{\mathcal{A}}(\mathbf{u}) = 1$. We see $\mathbf{u}$ as a sequence of prostochastic words. By compactness of $\mathcal{P}A^*$ it contains a converging subsequence. The prostochastic word induced by this subsequence belongs to $L(\mathcal{A})$.

Conversely, let $\overline{u}$ in $L(\mathcal{A})$, *i.e.* such that $\widehat{\phi}(\overline{u})(q_0, F) = 1$. Consider a sequence of finite words $\mathbf{u}$ inducing $\overline{u}$. By definition, we have $\lim \phi(\mathbf{u})(q_0, F) = 1$, *i.e.* $\lim P_{\mathcal{A}}(\mathbf{u}) = 1$, implying that val$(\mathcal{A}) = 1$. ∎

4.3 The Limit Operator, Fast and Polynomial Prostochastic Words

We show in this subsection how to construct non-trivial prostochastic words, and in particular the polynomial prostochastic words. To this end, we need to better understand *convergence speeds phenomena*: different limit behaviours can occur, depending on how fast the underlying Markov chains converge.

We define a limit operator ω. Consider the function $f : \mathbb{N} \to \mathbb{N}$ defined by $f(n) = k!$, where k is maximal such that $k! \leq n$. The function f grows linearly; the choice of n is arbitrary, one could replace n by any polynomial, or even by any subexponential function, see Remark 2.

The operator ω takes as input a sequence of finite words, and outputs a sequence of finite words. Formally, let $\mathbf{u}$ be a sequence of finite words, define:

$$\mathbf{u}^\omega = (u_n^{f(n)})_{n\in\mathbb{N}}.$$

It is not true in general that if $\mathbf{u}$ converges, then $\mathbf{u}^\omega$ converges. We will show that a sufficient condition is that $\mathbf{u}$ is fast.

We say that a sequence $(M_n)_{n\in\mathbb{N}}$ converges exponentially fast to M if there exists a constant $C > 1$ such that for all n large enough, $||M_n - M|| \leq C^{-n}$.

Definition 10. *(Fast sequence) A sequence of finite words $\mathbf{u}$ is fast if it converges (we denote $\overline{u}$ the prostochastic word it induces), and for every morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, the sequence $(\phi(u_n))_{n\in\mathbb{N}}$ converges exponentially fast.*

A prostochastic word is *fast* if it is induced by *some* fast sequence. We denote $\mathcal{P}A^*_f$ the set of fast prostochastic words. Note that a priori, not all prostochastic words are induced by some fast sequence.

We first prove that $\mathcal{P}A^*_f$ is a submonoid of $\mathcal{P}A^*$.

Lemma 1. *(The concatenation of two fast sequences is fast) Let $\mathbf{u}, \mathbf{v}$ be two fast sequences.*

The sequence $\mathbf{u} \cdot \mathbf{v} = (u_n \cdot v_n)_{n\in\mathbb{N}}$ is fast.

Proof. Consider a morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ and $n \in \mathbb{N}$.

$$\begin{aligned}
&||\phi(u_n \cdot v_n) - \widehat{\phi}(\overline{u} \cdot \overline{v})|| \\
&= ||\phi(u_n) \cdot \phi(v_n) - \widehat{\phi}(\overline{u}) \cdot \widehat{\phi}(\overline{v})|| \\
&= ||\phi(u_n) \cdot (\phi(v_n) - \widehat{\phi}(\overline{v})) - (\widehat{\phi}(\overline{u}) - \phi(u_n)) \cdot \widehat{\phi}(\overline{v})|| \\
&\leq ||\phi(u_n)|| \cdot ||\phi(v_n) - \widehat{\phi}(\overline{v})|| + ||\widehat{\phi}(\overline{u}) - \phi(u_n)|| \cdot ||\widehat{\phi}(\overline{v})|| \\
&= ||\phi(v_n) - \widehat{\phi}(\overline{v})|| + ||\widehat{\phi}(\overline{u}) - \phi(u_n)||.
\end{aligned}$$

Since $\mathbf{u}$ and $\mathbf{v}$ are fast, the previous inequality implies that $\mathbf{u} \cdot \mathbf{v}$ is fast. ∎

Let $\overline{u}$ and $\overline{v}$ be two fast prostochastic words, thanks to Lemma 1, the prostochastic word $\overline{u} \cdot \overline{v}$ is fast.

The remainder of this subsection is devoted to proving that ω is an operator $\mathcal{P}A_f^* \to \mathcal{P}A_f^*$. This is the key technical point of our characterisation. Indeed, we will define polynomial prostochastic words using concatenation and the operator ω, mimicking the definition of polynomial sequences of finite words. The fact that ω preserves the fast property of prostochastic words allows to obtain a perfect correspondence between polynomial sequences of finite words and polynomial prostochastic words.

The main technical tool is the following theorem, stating the exponentially fast convergence of the powers of a stochastic matrix.

Theorem 5. *(Powers of a stochastic matrix) Let $M \in \mathcal{S}_{Q\times Q}(\mathbb{R})$. Denote $P = M^{|Q|!}$. Then the sequence $(P^n)_{n\in\mathbb{N}}$ converges exponentially fast to a matrix M^ω, satisfying:*

$$\pi(M^\omega)(s,t) = \begin{cases} 1 & \text{if } \pi(P)(s,t) = 1 \text{ and } t \text{ is } \pi(P)\text{-recurrent,} \\ 0 & \text{otherwise.} \end{cases}$$

The proof of Theorem 5 is given in Subsection 4.4.

The following lemma shows that the ω operator is well defined for fast sequences. The second item shows that ω commutes with morphisms.

Lemma 2. *(Limit operator for fast sequences) Let $\mathbf{u}, \mathbf{v}$ be two equivalent fast sequences, inducing the fast prostochastic word $\overline{u}$. Then the sequences $\mathbf{u}^\omega$ and $\mathbf{v}^\omega$ are fast and equivalent, inducing the fast prostochastic word denoted $\overline{u}^\omega$.*

Furthermore, for every morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, we have $\widehat{\phi}(\overline{u}^\omega) = \widehat{\phi}(\overline{u})^\omega$.

Proof. Let $\phi : A \to \mathcal{S}_{Q\times Q}(\mathbb{R})$.

The sequence $(\widehat{\phi}(\overline{u})^{f(n)})_{n\in\mathbb{N}}$ is a subsequence of $(\widehat{\phi}(\overline{u})^{|Q|!\cdot n})_{n\in\mathbb{N}}$, so Theorem 5 implies that it converges exponentially fast to $\widehat{\phi}(\overline{u})^\omega$. It follows that there exists a constant $C_1 > 1$ such that for all n large enough, we have $||\widehat{\phi}(\overline{u})^{f(n)} - \widehat{\phi}(\overline{u})^\omega|| \leq C_1^{-f(n)}$.

We proceed in two steps, using the following inequality, which holds for every n:

$$||\phi(u_n^{f(n)}) - \widehat{\phi}(\overline{u})^\omega|| \leq ||\phi(u_n)^{f(n)} - \widehat{\phi}(\overline{u})^{f(n)}|| + ||\widehat{\phi}(\overline{u})^{f(n)} - \widehat{\phi}(\overline{u})^\omega||.$$

For the left summand, we rely on the following equality, where x and y may not commute:

$$x^N - y^N = \sum_{k=0}^{N-1} x^{N-k-1} \cdot (x-y) \cdot y^k.$$

Leting $N = f(n)$, this gives:

$$\begin{aligned}
&||\phi(u_n)^N - \widehat{\phi}(\overline{u})^N|| \\
&= || \sum_{k=0}^{N-1} \phi(u_n)^{N-k-1} \cdot (\phi(u_n) - \widehat{\phi}(\overline{u})) \cdot \widehat{\phi}(\overline{u})^k|| \\
&\leq \sum_{k=0}^{N-1} ||\phi(u_n)^{N-k-1}|| \cdot ||\phi(u_n) - \widehat{\phi}(\overline{u})|| \cdot ||\widehat{\phi}(\overline{u})^k|| \\
&\leq \sum_{k=0}^{N-1} \underbrace{||\phi(u_n)||^{N-k-1}}_{=1} \cdot ||\phi(u_n) - \widehat{\phi}(\overline{u})|| \cdot \underbrace{||\widehat{\phi}(\overline{u})||^k}_{=1} \\
&= N \cdot ||\phi(u_n) - \widehat{\phi}(\overline{u})||.
\end{aligned}$$

Since $\mathbf{u}$ is fast, there exists a constant $C_2 > 1$ such that $||\phi(u_n) - \widehat{\phi}(\overline{u})|| \leq C_2^{-n}$. Altogether, we have

$$||\phi(u_n^{f(n)}) - \widehat{\phi}(\overline{u})^\omega|| \leq f(n) \cdot C_2^{-n} + C_1^{-f(n)}.$$

To conclude, observe that for all n large enough, we have $\frac{n}{\log(n)} \leq f(n) \leq n$. It follows that the sequence $\mathbf{u}^\omega$ is fast, and that $\phi(\mathbf{u}^\omega)$ converges to $\widehat{\phi}(\overline{u})^\omega$.

Furthermore, since $\mathbf{u}$ and $\mathbf{v}$ are equivalent, we have $\lim \phi(\mathbf{u}) = \lim \phi(\mathbf{v})$, *i.e.* $\widehat{\phi}(\overline{u}) = \widehat{\phi}(\overline{v})$, so $\widehat{\phi}(\overline{u})^\omega = \widehat{\phi}(\overline{v})^\omega$, *i.e.* $\lim \phi(\mathbf{u}^\omega) = \lim \phi(\mathbf{v}^\omega)$, This implies that $\mathbf{u}^\omega$ and $\mathbf{v}^\omega$ are equivalent. ∎

Let $\overline{u}$ be a fast prostochastic word, we define the prostochastic word $\overline{u}^\omega$ as induced by $\mathbf{u}^\omega$, for some sequence $\mathbf{u}$ inducing $\overline{u}$. Thanks to Lemma 2, the prostochastic word $\overline{u}^\omega$ is well defined, and fast.

We can now define polynomial prostochastic words.

First, ω-expressions are described by the following grammar:

$$E \quad \longrightarrow \quad a \quad | \quad E \cdot E \quad | \quad E^\omega.$$

We define an interpretation $\overline{\cdot}$ of ω-expressions into fast prostochastic words:

- $\overline{a}$ is prostochastic word induced by the constant sequence of the one letter word a,
- $\overline{E_1 \cdot E_2} = \overline{E_1} \cdot \overline{E_2}$,
- $\overline{E^\omega} = \overline{E}^\omega$.

The following definition of polynomial prostochastic words is in one-to-one correspondence with the definition of polynomial sequences of finite words.

Definition 11. *(Polynomial prostochastic word) The set of polynomial prostochastic words is* $\{\overline{E} \mid E \text{ is an } \omega\text{-expression}\}$.

Remark 2. Why the term polynomial?

Consider an ω-expression E, say $(a^\omega b)^\omega$, and the prostochastic word $\overline{(a^\omega b)^\omega}$, which is induced by the sequence of finite words $((a^{f(n)}b)^{f(n)})_{n\in\mathbb{N}}$. The function f grows linearly, so this sequence represents a polynomial behaviour. Furthermore, the proofs above yield the following robustness property: all converging sequences of finite words $((a^{g(n)}b)^{h(n)})_{n\in\mathbb{N}}$, where $g, h : \mathbb{N} \to \mathbb{N}$ are subexponential functions, are equivalent, so they induce the same polynomial prostochastic word $\overline{(a^\omega b)^\omega}$. We say that a function $g : \mathbb{N} \to \mathbb{N}$ is subexponential if for all constants $C > 1$ we have $\lim_n g(n) \cdot C^{-n} = 0$; all polynomial functions are subexponential.

This justifies the terminology; we say that the polynomial prostochastic words represent all polynomial behaviours.

4.4 Powers of a Stochastic Matrix

In this subsection, we prove Theorem 5.

$$\left(\begin{array}{c|ccc} T & \multicolumn{3}{c}{R} \\ \hline 0 & J_1 & 0 & \\ & 0 & J_2 & J \\ & & & \ddots \\ & & & & J_r \end{array}\right) \left(\begin{array}{c|ccc} T^n & \multicolumn{3}{c}{\sum_{k=0}^{n-1} T^k \cdot R \cdot J^{n-1-k}} \\ \hline 0 & J_1^n & 0 & \\ & 0 & J_2^n & J^n \\ & & & \ddots \\ & & & & J_r^n \end{array}\right)$$

Fig. 1. Decomposition of P (on the left) and of P^n (on the right).

Let $M \in \mathcal{S}_{Q\times Q}(\mathbb{R})$, consider $P = M^{|Q|!}$. It is easy to see that $\pi(P)$ is idempotent. We decompose P as illustrated in Figure 1, by indexing states in the following way:

- first, $\pi(P)$-transient states,
- then, $\pi(P)$-recurrent states, grouped by recurrence class.

In this decomposition, we have the following properties:

- for all m-transient states $s \in Q$, we have $\sum_{t\ m\text{-transient}} T(s,t) < 1$, so $||T|| < 1$,
- the matrices J_i are irreducible: for all states $s, t \in Q$ corresponding to the same J_i, we have $J_i(s,t) > 0$.

The power P^n of P is represented in Figure 1.

This decomposition allows to treat separately the three blocks:

1. the block T^n: thanks to the observation above $||T|| < 1$, which combined with $||T^n|| \leq ||T||^n$ implies that $(T^n)_{n\in\mathbb{N}}$ converges to 0 exponentially fast,
2. the block $\sum_{k=0}^{n-1} T^k \cdot R \cdot J^{n-1-k}$,
3. the block J^n: it is handled by Lemma 3.

We first focus on item 3., and show that the sequence $(J^n)_{n\in\mathbb{N}}$ converges exponentially fast. Each block J_i is handled separately by the following lemma.

Lemma 3. *(Powers of an irreducible stochastic matrix) Let $J \in \mathcal{S}_{Q\times Q}(\mathbb{R})$ be irreducible: for all states $s, t \in Q$, we have $J(s,t) > 0$. Then the sequence $(J^n)_{n\in\mathbb{N}}$ converges exponentially fast to a matrix J^∞.*

Furthermore, J^∞ is irreducible.

This lemma is a classical result from Markov chain theory, sometimes called the Convergence Theorem; see for instance [8].

We now consider item 2., and show that the sequence $(\sum_{k=0}^{n-1} T^k \cdot R \cdot J^{n-1-k})_{n\in\mathbb{N}}$ converges exponentially fast. Observe that since $||T|| < 1$, the matrix $I - T$ is invertible, where I is the identity matrix with the same dimension as T. Denote $N = (I-T)^{-1}$, it is equal to $\sum_{k\geq 0} T^k$. Denote $J^\infty = \lim_n J^n$, which exists thanks to Lemma 3.

We have:

$$
\begin{aligned}
&||\sum_{k=0}^{n-1} T^k \cdot R \cdot J^{n-1-k} - N \cdot R \cdot J^\infty|| \\
&= ||\sum_{k=0}^{n-1} \left[T^k \cdot R \cdot \left(J^{n-1-k} - J^\infty\right) + T^k \cdot R \cdot J^\infty\right] - N \cdot R \cdot J^\infty|| \\
&= ||\sum_{k=0}^{n-1} T^k \cdot R \cdot \left(J^{n-1-k} - J^\infty\right) + \left(\sum_{k=0}^{n-1} T^k - N\right) \cdot R \cdot J^\infty|| \\
&\leq ||\sum_{k=0}^{n-1} T^k \cdot R \cdot \left(J^{n-1-k} - J^\infty\right)|| + ||\left(\sum_{k=0}^{n-1} T^k - N\right) \cdot R \cdot J^\infty||.
\end{aligned}
$$

We first consider the right summand:

$$
\begin{aligned}
&|| \left(\sum_{k=0}^{n-1} T^k - N \right) \cdot R \cdot J^\infty || \\
&= || \left(\sum_{k \geq n} T^k \right) \cdot R \cdot J^\infty || \\
&\leq || \sum_{k \geq n} T^k || \cdot \underbrace{||R||}_{\leq 1} \cdot \underbrace{||J^\infty||}_{=1} \\
&= ||T^n \cdot N|| \\
&\leq ||N|| \cdot ||T||^n.
\end{aligned}
$$

The first equality follows from the fact that $\sum_{k=0}^{n-1} T^k - N = \sum_{k \geq n} T^k$. Thus, this term converges exponentially fast to 0.

We next consider the left summand. Thanks to Lemma 3, there exists a constant $C > 1$ such that for all $p \in \mathbb{N}$, we have $||J^p - J^\infty|| \leq C^{-p}$.

$$
\begin{aligned}
&|| \sum_{k=0}^{n-1} T^k \cdot R \cdot \left(J^{n-1-k} - J^\infty \right) || \\
&\leq \sum_{k=0}^{n-1} ||T||^k \cdot \underbrace{||R||}_{\leq 1} \cdot ||J^{n-1-k} - J^\infty|| \\
&\leq \sum_{k=0}^{n-1} ||T||^k \cdot ||J^{n-1-k} - J^\infty|| \\
&= \sum_{k=0}^{\lfloor n/2 \rfloor} \underbrace{||T||^k}_{\leq 1} \cdot ||J^{n-1-k} - J^\infty|| + \sum_{k=\lfloor n/2 \rfloor + 1}^{n-1} ||T||^k \cdot \underbrace{||J^{n-1-k} - J^\infty||}_{\leq 2} \\
&\leq \frac{C^{-(\lfloor n/2 \rfloor + 1)} - C^{-n}}{1 - C} + 2 \cdot \frac{||T||^{\lfloor n/2 \rfloor + 1} - ||T||^n}{1 - ||T||} \\
&\leq 2 \cdot \left(\frac{C^{-(\lfloor n/2 \rfloor + 1)}}{1 - C} + \frac{||T||^{\lfloor n/2 \rfloor + 1}}{1 - ||T||} \right).
\end{aligned}
$$

Thus, this term converges exponentially fast to 0.

We proved that $(P^n)_{n \in \mathbb{N}}$ converges exponentially fast to a matrix M^ω. We conclude the proof of Theorem 5 by observing that:

$$
\pi(M^\omega)(s,t) = \begin{cases} 1 & \text{if } \pi(P)(s,t) = 1 \text{ and } t \text{ is } \pi(P)\text{-recurrent,} \\ 0 & \text{otherwise.} \end{cases}
$$

Assume first that $\pi(M^\omega)(s,t) = 1$, *i.e.* $M^\omega(s,t) > 0$. This already implies that t is $\pi(P)$-recurrent, looking at the decomposition of P^n. Since $M^\omega = \lim_n P^n$, it follows that for n large enough, we have $P^n(s,t) > 0$. The matrix $\pi(P)$ is idempotent, so we have for all $n \in \mathbb{N}$ the equality $\pi(P^n) = \pi(P)$, implying that $P(s,t) > 0$, *i.e.* $\pi(P)(s,t) = 1$.

Conversely, assume that $\pi(P)(s,t) = 1$ and t is $\pi(P)$-recurrent. Observe that for all $n \in \mathbb{N}$ we have $P^{n+1}(s,t) \geq P(s,t) \cdot P^n(t,t)$. For n tending to infinity, this implies $M^\omega(s,t) \geq P(s,t) \cdot M^\omega(t,t)$. Note that $P(s,t) > 0$, and $M^\omega(t,t) > 0$ since t is $\pi(P)$-recurrent and thanks to Lemma 3. It follows that $M^\omega(s,t) > 0$, *i.e.* $\pi(M^\omega)(s,t) = 1$.

4.5 Reformulating the Characterisation

For proof purposes, we give an equivalent presentation of the Markov Monoid through ω-expressions. Given a probabilistic automaton $\mathcal{A}$, we define an interpretation $\langle\cdot\rangle$ of ω-expressions into Boolean matrices:

- $\langle a \rangle$ is $\pi(\phi(a))$,
- $\langle E_1 \cdot E_2 \rangle$ is $\langle E_1 \rangle \cdot \langle E_2 \rangle$,
- $\langle E^\omega \rangle$ is $\langle E \rangle^\sharp$, only defined if $\langle E \rangle$ is idempotent.

Then the Markov Monoid of $\mathcal{A}$ is $\{\langle E \rangle \mid E \text{ an } \omega\text{-expression}\}$.

The following theorem is a reformulation of Theorem 1, using the prostochastic theory. It clearly implies Theorem 1: indeed, a polynomial prostochastic word induces a polynomial sequence, and vice-versa.

Theorem 6. *(Characterisation of the Markov Monoid algorithm) The Markov Monoid algorithm answers "YES" on input $\mathcal{A}$ if, and only if, there exists a polynomial prostochastic word accepted by $\mathcal{A}$.*

The proof relies on the notion of reification, used in the following proposition, from which follows Theorem 6.

Definition 12. *(Reification) Let $\mathcal{A}$ be a probabilistic automaton.*

A sequence $(u_n)_{n\in\mathbb{N}}$ of words reifies a Boolean matrix m if for all states $s,t \in Q$, the sequence $\left(P_\mathcal{A}(s \xrightarrow{u_n} t)\right)_{n\in\mathbb{N}}$ converges and:

$$m(s,t) = 1 \iff \lim_n P_\mathcal{A}(s \xrightarrow{u_n} t) > 0.$$

Proposition 1. *(Characterisation of the Markov Monoid algorithm) For every ω-expression E, for every $\phi : A \to \mathcal{S}_{Q\times Q}(\mathbb{R})$, we have*

$$\pi(\widehat{\phi}(\overline{E})) = \langle E \rangle.$$

Consequently, for every probabilistic automaton $\mathcal{A}$:

– *any sequence inducing the polynomial prostochastic word $\overline{E}$ reifies $\langle E \rangle$,*
– *the element $\langle E \rangle$ of the Markov Monoid is a value 1 witness if, and only if, the polynomial prostochastic word $\overline{E}$ is accepted by $\mathcal{A}$.*

Proof. We prove the first part of Proposition 1 by induction on the ω-expression E, which essentially amounts to gather the results from Section 4.

The base case of $a \in A$ is clear.

The product case: let $E = E_1 \cdot E_2$, and $\phi : A \to \mathcal{S}_{Q \times Q}(\mathbb{R})$.

By definition $\overline{E} = \overline{E_1} \cdot \overline{E_2}$, so $\widehat{\phi}(\overline{E}) = \widehat{\phi}(\overline{E_1}) \cdot \widehat{\phi}(\overline{E_2})$ because $\widehat{\phi}$ is a morphism, and $\pi(\widehat{\phi}(\overline{E})) = \pi(\widehat{\phi}(\overline{E_1})) \cdot \pi(\widehat{\phi}(\overline{E_2}))$. Also by definition, we have $\langle E \rangle = \langle E_1 \rangle \cdot \langle E_2 \rangle$, so $\pi(\widehat{\phi}(\overline{E_1 \cdot E_2})) = \langle E_1 \cdot E_2 \rangle$.

The iteration case: let $E = F^\omega$, and $\phi : A \to \mathcal{S}_{Q \times Q}(\mathbb{R})$.

By definition, $\overline{E} = \overline{F}^\omega$, so $\widehat{\phi}(\overline{E}) = \widehat{\phi}(\overline{F}^\omega)$, which is equal to $\widehat{\phi}(\overline{F})^\omega$ thanks to Lemma 2. Now, $\pi(\widehat{\phi}(\overline{F})^\omega) = \pi(\widehat{\phi}(\overline{F}))^\sharp$ thanks to Theorem 5. By induction hypothesis, $\pi(\widehat{\phi}(\overline{F})) = \langle F \rangle$, so $\pi(\widehat{\phi}(\overline{F}^\omega)) = \langle F^\omega \rangle$.

We prove the second part. Consider a sequence $\mathbf{u}$ inducing the polynomial prostochastic word $\overline{E}$. Thanks to the first item, $\pi(\widehat{\phi}(\overline{E})) = \langle E \rangle$, implying that $\pi(\lim \phi(\mathbf{u})) = \langle E \rangle$, which means that $\mathbf{u}$ reifies $\overline{E}$.

We prove the third part.

Assume that $\langle E \rangle$ is a value 1 witness, *i.e.* for all states $t \in Q$, if $\langle E \rangle(q_0, t) = 1$, then $t \in F$. So for $t \notin F$, we have $\lim \phi(\mathbf{u})(q_0, t) = 0$. Since we have $\lim \phi(\mathbf{u})(q_0, Q) = 1$, it follows that $\lim \phi(\mathbf{u})(q_0, F) = 1$, so $\widehat{\phi}(\overline{E})(q_0, F) = 1$, *i.e.* the polynomial prostochastic word $\overline{E}$ is accepted by $\mathcal{A}$.

Conversely, assume that the polynomial prostochastic word $\overline{E}$ is accepted by $\mathcal{A}$. Since it is induced by $\mathbf{u}$, it follows that $\lim \phi(\mathbf{u})(q_0, F) = 1$. Consider a state $t \in Q$ such that $\langle E \rangle(q_0, t) = 1$. It follows that $\pi(\lim \phi(\mathbf{u}))(q_0, t) = 1$, so $\lim \phi(\mathbf{u})(q_0, t) > 0$. Since $\lim \phi(\mathbf{u})(q_0, F) = 1$, this implies that $t \in F$, hence $\langle E \rangle$ is a value 1 witness. ∎

5 Towards an Optimality Argument

In this section, we build on the characterisation obtained above to argue that the Markov Monoid algorithm is in some sense optimal.

5.1 Undecidability of the Two-Tier Value 1 Problem

Theorem 7. *(Undecidability of the two-tier value 1 problem) The following problem is undecidable: given a probabilistic automaton $\mathcal{A}$, determine whether there exist two finite words u, v such that $\lim_n P_{\mathcal{A}}((u \cdot v^n)^{2^n}) = 1$.*

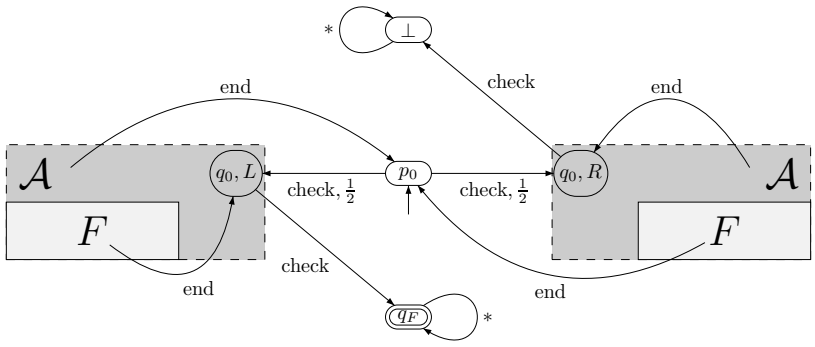

Fig. 2. Reduction.

The two-tier value 1 problem seems easier than the value 1 problem as it restricts the set of sequences of finite words to very simple sequences. We call such sequences two-tier, because they exhibit two different behaviours: the word v is repeated a linear number of times, namely n, while the word $u{\cdot}v^n$ is repeated an exponential number of times, namely 2^n.

The proof is obtained using the same reduction as for the undecidability of the value 1 problem, from [7], with a refined analysis.

Proof. We construct a reduction from the emptiness problem for probabilistic automata to the two-tier value 1 problem. For technical reasons, we will assume that the probabilistic automata have transition probabilities 0, $\frac{1}{2}$, or 1.

Let $\mathcal{A}$ be a probabilistic automaton. We construct a probabilistic automaton $\mathcal{B}$ such that the following holds:

$$\begin{array}{c}\text{there exists a finite word } w \text{ such that } P_{\mathcal{A}}(w) > \frac{1}{2} \text{ if, and only if,}\\ \text{there exist two finite words } u, v \text{ such that } \lim_n P_{\mathcal{B}}((u \cdot v^n)^{2^n}) = 1.\end{array}$$

The emptiness problem for probabilistic automata has been shown undecidable in [9]. We refer to [7] for a simple proof of this result.

Without loss of generality we assume that the initial state q_0 of $\mathcal{A}$ has no ingoing transitions.

The alphabet of $\mathcal{B}$ is $B = A \uplus \{\text{check}, \text{end}\}$, its set of states is $Q_{\mathcal{B}} = Q \times \{L, R\} \uplus \{p_0, \bot, q_F\}$, its transition function is ϕ', the only initial state is p_0 and

the only final state is q_F. We describe ϕ' as a function $\phi' : Q_{\mathcal{B}} \times B \to \mathcal{D}(Q_{\mathcal{B}})$:

$$\begin{cases} \phi'(p_0, \text{check}) & = \frac{1}{2} \cdot (q_0, L) + \frac{1}{2} \cdot (q_0, R) \\ \phi'((q, d), a) & = (\phi(q, a), d) \text{ for } a \in A \text{ and } d \in \{L, R\} \\ \phi'((q_0, L), \text{check}) & = q_F \\ \phi'((q, L), \text{end}) & = q_0 \text{ if } q \in F \\ \phi'((q, L), \text{end}) & = p_0 \text{ if } q \notin F \\ \phi'((q_0, R), \text{check}) & = \bot \\ \phi'((q, R), \text{end}) & = p_0 \text{ if } q \in F \\ \phi'((q, R), \text{end}) & = q_0 \text{ if } q \notin F \\ \phi'(q_F, *) & = q_F \end{cases}$$

where as a convention, if a transition is not defined, it leads to $\bot$.

Assume that there exists a finite word w such that $P_{\mathcal{A}}(w) > \frac{1}{2}$, then we claim that $\lim_n P_{\mathcal{B}}((\text{check} \cdot (w \cdot \text{end})^n)^{2^n}) = 1$. Denote $x = P_{\mathcal{A}}(w)$.

We have

$$P_{\mathcal{A}}(p_0 \xrightarrow{\text{check}\cdot(w\cdot\text{end})^n} (q_0, L)) = \frac{1}{2} \cdot x^n,$$

and

$$P_{\mathcal{A}}(p_0 \xrightarrow{\text{check}\cdot(w\cdot\text{end})^n} (q_0, R)) = \frac{1}{2} \cdot (1 - x)^n.$$

We fix an integer N and analyse the action of reading $(\text{check} \cdot (w \cdot \text{end})^n)^N$: there are N "rounds", each of them corresponding to reading $\text{check} \cdot (w \cdot \text{end})^n$ from p_0. In a round, there are three outcomes: winning (that is, remaining in (q_0, L)) with probability $p_n = \frac{1}{2} \cdot x^n$, losing (that is, remaining in (q_0, R)) with probability $q_n = \frac{1}{2} \cdot (1 - x)^n$, or going to the next round (that is, reaching p_0) with probability $1 - (p_n + q_n)$. If a round is won or lost, then the next check leads to an accepting or rejecting sink; otherwise it goes on to the next round, for N rounds. Hence:

$$\begin{aligned} & P_{\mathcal{A}}((\text{check} \cdot (w \cdot \text{end})^n)^N) \\ & = \sum_{i=1}^{N-1} (1 - (p_n + q_n))^{i-1} \cdot p_n \\ & = p_n \cdot \frac{1 - (1 - (p_n + q_n))^{N-1}}{1 - (1 - (p_n + q_n))} \\ & = \frac{1}{1 + \frac{q_n}{p_n}} \cdot \left(1 - (1 - (p_n + q_n))^{N-1}\right) \end{aligned}$$

First, $\frac{q_n}{p_n} = (\frac{1-x}{x})^n$ converges to 0 as n goes to infinity since $x > \frac{1}{2}$.

Denote $N = f(n)$ and consider the sequence $((1 - (p_n + q_n))^{N-1})_{n\in\mathbb{N}}$; if $(f(n) \cdot x^n)_{n\in\mathbb{N}}$ converges to ∞, then the above sequence converges to 0. Since $x > \frac{1}{2}$, this holds for $f(n) = 2^n$. It follows that the acceptance probability converges to 1. Consequently:

$$\lim_n P_{\mathcal{A}}((\text{check} \cdot (w \cdot \text{end})^n)^{2^n}) = 1.$$

Conversely, assume that for all finite words w, we have $P_{\mathcal{A}}(w) \leq \frac{1}{2}$. We claim that every finite word in B^* is accepted by $\mathcal{B}$ with probability at most $\frac{1}{2}$. First of all, using simple observations we restrict ourselves to words of the form

$$w = \text{check} \cdot w_1 \cdot \text{end} \cdot w_2 \cdot \text{end} \cdots w_n \cdot \text{end} \cdot w',$$

with $w_i \in A^*$ and $w' \in B^*$. Since $P_{\mathcal{A}}(w_i) \leq \frac{1}{2}$ for every i, it follows that in $\mathcal{B}$, after reading the last letter end in w before w', the probability to be in (q_0, L) is smaller or equal than the probability to be in (q_0, R). This implies the claim. It follows that the value of $\mathcal{B}$ is not 1, and in particular for two finite words u, v, we have $\lim_n P_{\mathcal{B}}((u \cdot v^n)^{2^n}) < 1$. ∎

Note that the proof actually shows that one can replace in the statement of Theorem 7 the value 2^n by $f(n)$ for any function f such that $f(n) \geq 2^n$.

5.2 Combining the Two Results

It was shown in [4] that the Markov Monoid algorithm subsumes all previous known algorithms to solve the value 1 problem. Indeed, it was proved that it is correct for the subclass of leaktight automata, and that the class of leaktight automata strictly contains all subclasses for which the value 1 problem has been shown to be decidable.

At this point, the Markov Monoid algorithm is the best algorithm *so far*. But can we go further? If we cannot, then what is an optimality argument? It consists in constructing a maximal subclass of probabilistic automata for which the problem is decidable. We can reverse the point of view, and equivalently construct an optimal algorithm, *i.e.* an algorithm that correctly solves a subset of the instances, such that no algorithm correctly solves a superset of these instances. However, it is clear that no such strong statement holds, as one can always from any algorithm obtain a better algorithm by precomputing finitely many instances.

Since there is no strong optimality argument, we can only give a subjective argument. We argue that the combination of our characterisation from Section 3

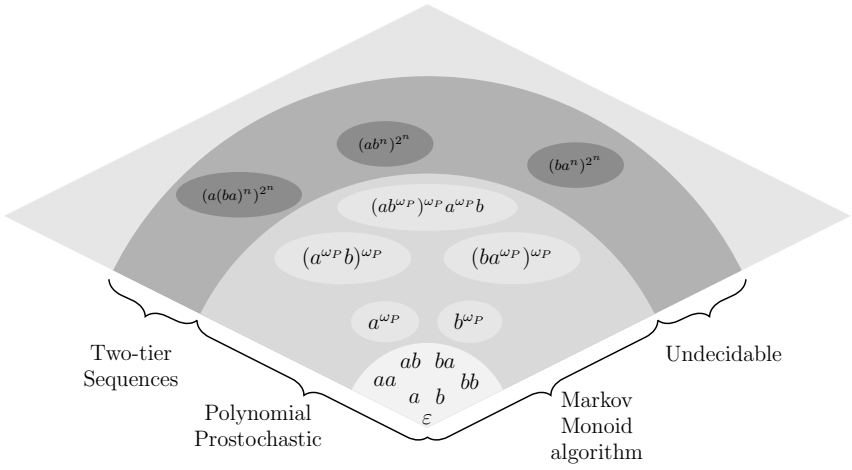

Fig. 3. Optimality of the Markov Monoid algorithm.

and of the undecidability of the two-tier value 1 problem from Subsection 5.1 supports the claim that the Markov Monoid algorithm is *in some sense* optimal:

- The characterisation says that the Markov Monoid algorithm captures exactly all polynomial behaviours.
- The undecidability result says that the undecidability of the value 1 problem arises when polynomial and exponential behaviours are combined.

So, the Markov Monoid algorithm is optimal in the sense that it captures a *large* set of behaviours, namely polynomial behaviours, and that no algorithm can capture both polynomial and exponential behaviours.

References

1. Jorge Almeida. Profinite semigroups and applications. *Structural Theory of Automata, Semigroups, and Universal Algebra*, 207:1–45, 2005.
2. Krishnendu Chatterjee and Mathieu Tracol. Decidable problems for probabilistic automata on infinite words. In *LICS*, pages 185–194, 2012.
3. Nathanaël Fijalkow. Characterisation of an algebraic algorithm for probabilistic automata. In *STACS*, pages 34:1–34:13, 2016.
4. Nathanaël Fijalkow, Hugo Gimbert, Edon Kelmendi, and Youssouf Oualhadj. Deciding the value 1 problem for probabilistic leaktight automata. *Logical Methods in Computer Science*, 11(1), 2015.
5. Nathanaël Fijalkow, Hugo Gimbert, and Youssouf Oualhadj. Deciding the value 1 problem for probabilistic leaktight automata. In *LICS*, pages 295–304, 2012.
6. Mai Gehrke, Serge Grigorieff, and Jean-Éric Pin. A topological approach to recognition. In *ICALP (2)*, pages 151–162, 2010.

7. Hugo Gimbert and Youssouf Oualhadj. Probabilistic automata on finite words: Decidable and undecidable problems. In *ICALP (2)*, pages 527–538, 2010.
8. David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.
9. Azaria Paz. *Introduction to Probabilistic Automata*. Academic Press, 1971.
10. Jean-Éric Pin. Profinite methods in automata theory. In *STACS*, pages 31–50, 2009.
11. Michael O. Rabin. Probabilistic automata. *Information and Control*, 6(3):230–245, 1963.
12. Szymon Toruńczyk. *Languages of Profinite Words and the Limitedness Problem*. PhD thesis, University of Warsaw, 2011.