Profinite Techniques for Probabilistic Automata and the Optimality of the Markov Monoid Algorithm

Nathanaël Fijalkow

Profinite Techniques for Probabilistic Automata
and the Optimality of the Markov Monoid Algorithm

Nathanaël Fijalkow

LIAFA, Paris 7, France
University of Warsaw, Poland

Abstract. We consider the value 1 problem for probabilistic automata over finite words. This problem is known to be undecidable. However, different algorithms have been proposed to partially solve it. The aim of this paper is to prove that one such algorithm, called the Markov Monoid algorithm, is optimal.
To this end, we develop a profinite theory for probabilistic automata. This new framework gives a topological account by constructing the free prostochastic monoid. We use it in two ways. First, to characterize the computations realized by the Markov Monoid algorithm, and second to prove its optimality.

1 Introduction

In 1963 Rabin [Rab63] introduced the notion of probabilistic automata, which are finite automata with randomized transitions. This powerful model has been widely studied ever since and has applications, for instance in image processing [CK97], computational biology [DEKM99] and speech processing [Moh97]. This paper follows a long line of work that studies the algorithmic properties of probabilistic automata. For instance, Schützenberger [Sch61] proved in 1961 that *language equivalence* is decidable in polynomial time, and even faster with randomized algorithms, which led to applications in software verification [KMO$^+$11].

However, many natural decision problems are *undecidable*; for example the *emptiness*, the *isolation* and the *value* 1 problems are undecidable, as shown in [Paz71,BMT77,GO10]. To overcome such untractability results, a lot of effort went into finding subclasses of probabilistic automata for which natural decision problems become decidable. For instance, Chadha et al. and Korthikanti et al. look at restrictions implying a decidable model-checking problem against ω-regular specifications [KVAK10,CKV$^+$11], and investigates whether assuming isolated cut-points leads to decidability for the emptiness problem [CSV13].

So far, no optimality argument has ever been given for the proposed subclasses. In other words, for the natural decision problems we still lack a good understanding of the decidability barrier.

The aim of this paper is to draw such a decidability barrier for 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 [GO10], but attracted a lot of attention recently (see, for instance, [BBG12,CT12,FGO12,FGHO14,FGKO14]).

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. Hence our optimality argument has to be weaker.

We show that the Markov Monoid algorithm is *in some sense* optimal, by showing that no algorithm can correctly solve *substantially* more instances than the Markov Monoid algorithm. To this end, we first characterize the computations of the Markov Monoid algorithm: roughly speaking, it captures exactly all *polynomial behaviours*. We then show that no algorithm can capture both polynomial and *exponential behaviours*, supporting the claim that the Markov Monoid algorithm is optimal.

To make sense of the notion of *convergence speeds*, we rely on topological techniques. We develop a profinite theory for probabilistic automata, called prostochastic theory. This is inspired by the profinite approach for (classical) automata [Pin09,GGP10], and for distance automata as developed in Szymon Toruńczyk's PhD thesis [Tor11].

The Section 3 is devoted to constructing the free prostochastic monoid and showing some of its properties. In particular, we define the acceptance of a prostochastic word by a probabilistic automaton, and show that the value 1 problem reformulates as the emptiness problem for probabilistic automata over prostochastic words. The free prostochastic monoid is represented in Figure 1, as a diamond. It contains the set of finite words, represented on the bottom of the picture. It is contained in the set of polynomial prostochastic words, itself contained in the set of exponential prostochastic words.

Our main result is the following (the missing definitions are given in Section 2 and 4):

Theorem 1. *[Optimality of the Markov Monoid algorithm]*

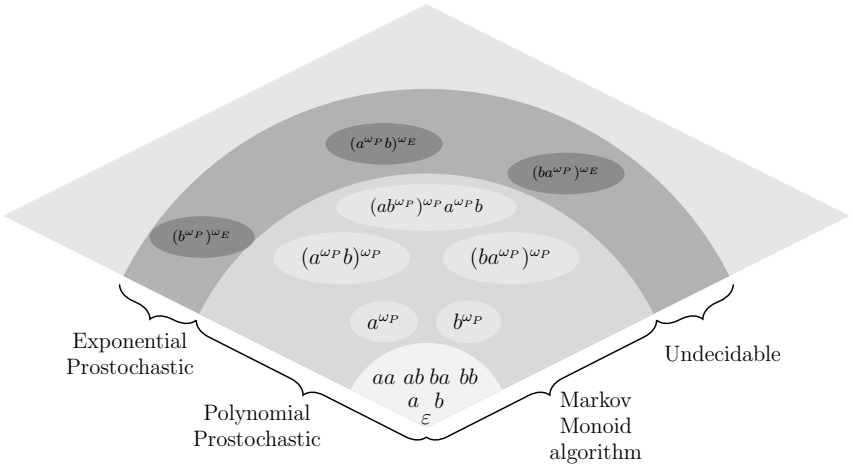

Fig. 1. The Free Prostochastic Monoid.

1. *(Characterization) The Markov Monoid algorithm answers "YES" on input $\mathcal{A}$ if, and only if, there exists a polynomial prostochastic word accepted by $\mathcal{A}$,*
2. *(Undecidability) The following problem is undecidable: given a probabilistic automaton $\mathcal{A}$ as input, determine whether there exists an exponential prostochastic word accepted by $\mathcal{A}$.*

To construct prostochastic words, we define two limit operators: an operator ω_P, where P stands for "polynomial", and an operator ω_E, where E stands for "exponential".

The polynomial prostochastic words are built using concatenation and the operator ω_P. On an intuitive level, this does not allow for different convergence speeds to compete. Indeed, part of the proof consists in showing that the polynomial prostochastic words are fast, a notion made precise in Section 4.3. On the other hand, the exponential prostochastic words are built using both operators ω_P and ω_E, which allows for two convergence speeds to interfere, leading to undecidability.

The Section 4 is devoted to proving the optimality of the Markov Monoid algorithm. Specifically, we prove the first half of the theorem above in Section 4.4, and the second half in Section 4.5.

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.

2 Probabilistic Automata and the Value 1 Problem

We work with finite words over a fixed finite alphabet A. We denote by $\mathbb{D}$ the set of dyadic rationals, *i.e.* numbers of the form $\frac{a}{2^b}$, for $a \in \mathbb{Z}$ and $b \in \mathbb{N}$. The set of real numbers is denoted $\mathbb{R}$.

A matrix is stochastic if every entry is non-negative and each line sums up to 1. For a finite set Q (thought of as a set of states) and $E \subseteq \mathbb{R}$, we denote $\mathcal{M}_{Q\times Q}(()E)$ the set of matrices over E. The restriction to stochastic matrices is denoted $\mathcal{S}_{Q\times Q}(()E)$. We consider the ℓ_1-norm $||\cdot||$ defined by $||M|| = \max_j \sum_i M(i,j)$. The following classical properties will be useful:

Fact 1

- *For every $M \in \mathcal{S}_{Q\times Q}(\mathbb{R})$, we have $||M|| = 1$,*
- *For $M, M' \in \mathcal{M}_{Q\times Q}(\mathbb{R})$, we have $||M \cdot M'|| \leq ||M|| \cdot ||M'||$,*
- *The space $\mathcal{S}_{Q\times Q}(\mathbb{R})$ is compact (so also complete).*

Definition 1 (Probabilistic automaton). *A probabilistic automaton is given by a finite set of states Q, a transition function $\phi : A \to \mathcal{M}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$, a stochastic vector of initial states I and a boolean vector of final states F.*

A transition function $\phi : A \to \mathcal{M}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$ naturally induces a morphism $\phi : A^* \to \mathcal{M}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$. We denote by $P_{\mathcal{A}}(s \xrightarrow{w} t)$ the probability to go from state s to state t reading w on the automaton $\mathcal{A}$, *i.e.* $\phi(w)(s,t)$.

The *acceptance probability* of a word $w \in A^*$ by $\mathcal{A}$ is $I \cdot \phi(w) \cdot F$, which we denote by $P_{\mathcal{A}}(w)$. In words, it is the probability that a run ends in a final state from F, starting from the initial distribution given by I.

Definition 2 (Value). *The* value *of a probabilistic automaton $\mathcal{A}$, denoted by* $\mathrm{val}(\mathcal{A})$*, is the supremum acceptance probability over all input words:*

$$\mathrm{val}(\mathcal{A}) = \sup_{w\in A^*} P_{\mathcal{A}}(w) \ .$$

We are interested in the following decision problem:

Problem 1 (Value 1 Problem). Given a probabilistic automaton $\mathcal{A}$, determine whether $\mathrm{val}(\mathcal{A}) = 1$.

An equivalent formulation of the value 1 problem is as follows: given a probabilistic automaton $\mathcal{A}$, is it true that for all $\varepsilon > 0$, there exists a word w such that $P_{\mathcal{A}}(w) \geq 1 - \varepsilon$?

The value 1 problem can also be reformulated using the notion of *isolated cut-point* introduced by Rabin in his seminal paper [Rab63]: an automaton has value 1 if and only if the cut-point 1 is *not* isolated.

Unfortunately:

Theorem 2 ([GO10]). *The value* 1 *problem is undecidable.*

A series of papers [GO10,CT12,FGO12,FGHO14,FGKO14] tackled the following question, with different approaches and techniques:

"To what extent is the value 1 problem undecidable?"

One line of work was to construct algorithms to solve the problem on some subclass of probabilistic automata [GO10,CT12,FGO12,FGKO14]. As proved in [FGKO14], the Markov Monoid algorithm is the most correct algorithm of all the algorithms proposed in these papers: all subclasses considered are included in the subclass of leaktight automata, for which the Markov Monoid algorithm correctly solves the value 1 problem.

Another route was to consider variants of the problem, by abstracting away the numerical values [FGHO14], but this does not lead to decidability.

In this paper, our aim is different. The objective is to draw a decidability barrier for the value 1 problem, through a precise understanding of both the Markov Monoid algorithm and the undecidability proof.

3 The Prostochastic Theory

In this section, we develop a profinite theory for probabilistic automata. The main point here is to construct the free prostochastic monoid, which allows to reformulate the value 1 problem as an emptiness problem over prostochastic words. The prostochastic theory is then used as a formalism to prove the optimality of the Markov Monoid algorithm, in the next section.

3.1 The Free Prostochastic Monoid

A profinite monoid is a monoid for which two elements can be distinguished by a morphism into a finite monoid, *i.e.* essentially by a finite automaton. To

define prostochastic monoids, we use a stronger distinguishing feature, namely probabilistic automata, which mathematically correspond to stochastic matrices over the reals.

Definition 3 (Prostochastic Monoid). *We say that a monoid $\mathcal{P}$ is prostochastic if for every $s \neq t \in \mathcal{P}$, there exists a morphism $\psi : \mathcal{P} \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ such that $\psi(s) \neq \psi(t)$.*

A prostochastic monoid $\mathcal{P}$ is naturally equipped with the prostochastic topology, which is the smallest that makes continuous every morphism $\psi : \mathcal{P} \to \mathcal{S}_{Q\times Q}(\mathbb{R})$.

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

Lemma 1 (Prostochastic Monoids are Compact and Topological).

Every prostochastic monoid $\mathcal{P}$ is compact and topological, i.e. the product function

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

is continuous.

Proof. By definition, a prostochastic monoid $\mathcal{P}$ can be seen as a closed subset of $\prod_{\psi:\mathcal{P}\to\mathcal{S}_{Q\times Q}(\mathbb{R})} \mathcal{S}_{Q\times Q}(\mathbb{R})$. The latter is compact as a product of compact spaces, thanks to Tychonoff's theorem. It follows that $\mathcal{P}$ is compact.

Thanks to this observation, the continuity of the product function follows from the continuity of the product function in $\mathcal{S}_{Q\times Q}(\mathbb{R})$. ∎

The main theorem of the prostochastic theory is the existence and unicity of a space, called the free prostochastic monoid, that satisfies a Universal Property.

Theorem 3 (Existence of the Free Prostochastic Monoid).

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

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

The unicity argument (item 2.) is a consequence of the Universal Property (item 1.), following standard arguments. The remainder of this subsection focuses on the existence part of this theorem (item 1.): we first construct a set $\mathcal{P}A^*$, and then show that it has the desired properties.

Our first move is to equip A^* with an appropriate distance function.

Definition 4 (Stochastic Distance). *We say that two words u and v are (N, η)-separated if there exists $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$ such that*

$$|Q| \leq N \quad \text{and} \quad ||\phi(u) - \phi(v)|| \geq \eta \, .$$

Define $d(u, v)$ as 2^{-N}, where N is minimal such that u and v are $(N, 2^{-N})$-separated.

In the profinite theory, two words are *close* if there exists a morphism into a *small* monoid that distinguishes them.

In the prostochastic theory, two words are *close* if there exists a morphism into a *small* stochastic matrix monoid over the reals, that separates them by a *large* value. Hence the distance involves a threshold between the size of the monoid, which should be small, and the separation between the values, which should be large.

Lemma 2. *The function d is a metric.*

Proof. Observe that any two distinct words can be separated, so $d(u, v) = 0$ if, and only if, $u = v$.

We now focus on the triangle inequality. The key (simple) observation is that if u and v are (N, η)-separated, then for every w, there exists $\eta_1 + \eta_2 \geq \eta$ such that:

- u and v are (N, η_1)-separated,
- v and w are (N, η_2)-separated.

Indeed, assume that there exists $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$ such that $|Q| \leq N$ and $||\phi(u) - \phi(v)|| \geq \eta$. It follows that $\eta \leq ||\phi(u) - \phi(w)|| + ||\phi(w) - \phi(v)||$, implying the result for $\eta_1 = ||\phi(u) - \phi(w)||$ and $\eta_2 = ||\phi(w) - \phi(v)||$.

Now, let u, v, w such that $d(u, v) = 2^{-N}$. The words u and v are $(N, 2^{-N})$-separated, so there exists $\eta_1+\eta_2 \geq 2^{-N}$ such that u and v are (N, η_1)-separated, and v and w are (N, η_2)-separated. Observe that $\phi(u) - \phi(w)$ and $\phi(w) - \phi(v)$ are dyadic numbers, so η_1, η_2 can be chosen dyadic, and we fall in either of the following three cases: *(i)* $\eta_1 \geq 2^{-N}$, or *(ii)* $\eta_2 \geq 2^{-N}$, or *(iii)* both $\eta_1 \geq 2^{-(N+1)}$ and $\eta_2 \geq 2^{-(N+1)}$ hold. In the three cases, we have $d(u, v) \leq d(u, w) + d(w, v)$. ∎

From now on, we see A^* as a metric space equipped with the distance d.

Fact 2 *The product function $\cdot$ defined by*

$$\begin{cases} A^* \times A^* \to A^* \\ (u, v) \mapsto u \cdot v \end{cases}$$

is uniformly continuous.

Proof. Let u, v, u', v' in A^*. We have:

$$\begin{aligned}
&||\phi(u \cdot v) - \phi(u' \cdot v')|| = ||\phi(u) \cdot \phi(v) - \phi(u') \cdot \phi(v')|| \\
&\leq ||\phi(u) - \phi(u')|| \cdot ||\phi(v)|| + ||\phi(v) - \phi(v')|| \cdot ||\phi(u')|| \\
&= ||\phi(u) - \phi(u')|| + ||\phi(v) - \phi(v')|| \,.
\end{aligned}$$

Assume $d(u \cdot v, u' \cdot v') = 2^{-N}$. The words $u \cdot v$ and $u' \cdot v'$ are $(N, 2^{-N})$-separated, so thanks to the above inequality, $2^{-N} \leq ||\phi(u) - \phi(u')|| + ||\phi(v) - \phi(v')||$. Observe that $||\phi(u) - \phi(u')||$ and $||\phi(v) - \phi(v')||$ are dyadic numbers, so we fall in either of the following three cases: *(i)* $||\phi(u) - \phi(u')|| \geq 2^{-N}$, or *(ii)* $||\phi(v) - \phi(v')|| \geq 2^{-N}$, or *(iii)* both $||\phi(u) - \phi(u')|| \geq 2^{-(N+1)}$ and $||\phi(v) - \phi(v')|| \geq 2^{-(N+1)}$ hold. In the three cases, we have $d(u \cdot v, u' \cdot v') \leq d(u, u') + d(v, v')$.

It follows that the product function is uniformly continuous. ∎

We can now construct $\mathcal{P}A^*$, as the topological completion of the metric space A^*. Formally, $\mathcal{P}A^*$ is the set of Cauchy sequences of A^*, up to the following equivalence relation. We denote sequences of finite words by $\mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots$, implicitely assuming that $\mathbf{u} = (u_n)_{n \in \mathbb{N}}$.

Definition 5 (Equivalence of Cauchy Sequences). *We say that two Cauchy sequences of words* $\mathbf{u}$ *and* $\mathbf{v}$ *are equivalent if the interleave sequence defined as* $(u_0, v_0, u_1, v_1, \ldots)$ *is Cauchy.*

Lemma 3 (Characterization of the Cauchy sequences). *Let* $\mathbf{u}$ *be a sequence of words. The following are equivalent:*

1. *the sequence* u *is Cauchy:*

$$\forall \varepsilon > 0,\ \exists N,\ \forall p, q \geq N,\ d(u_p, u_q) \leq \varepsilon \,. \tag{1}$$

2. *for every* $\phi : A \to \mathcal{S}_{Q \times Q}(\{0, \frac{1}{2}, 1\})$*, the sequence of stochastic matrices* $\phi(\mathbf{u})$ *converges (in* $\mathcal{S}_{Q \times Q}(\mathbb{R})$*).*

Note that we restricted ϕ to take values in $\mathcal{S}_{Q \times Q}(\{0, \frac{1}{2}, 1\})$. This is crucial, as it implies that they are finitely such functions if $|Q|$ is bounded.

Proof. Assume that u is Cauchy, and let $\phi : A \to \mathcal{S}_{Q \times Q}(\{0, \frac{1}{2}, 1\})$, we show that $\phi(\mathbf{u})$ converges in $\mathcal{S}_{Q \times Q}(\mathbb{R})$. Since this space is complete, it is enough to show that it is Cauchy:

$$\forall \varepsilon > 0,\ \exists M,\ \forall p, q \geq M,\ ||\phi(u_p) - \phi(u_q)|| \leq \varepsilon \,. \tag{2}$$

Let $\varepsilon = 2^{-P}$ for some P, and N given by Equation 1. (Without loss of generality, $|Q| \leq P$.) Denote by M the maximum of $|Q|$, N and P. For $p, q \geq M$, we have $d(u_p, u_q) \leq \varepsilon$ thanks to Equation 1 and $M \geq N$. By definition of the distance, this implies that u_p and u_q are not (P, ε)-separated. In particuar ϕ does not separate them; since $|Q| \leq P$, it follows that $||\phi(u_p) - \phi(u_q)|| < \varepsilon$. So Equation 2 holds.

Conversely, assume that for every $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$, the sequence $\phi(\mathbf{u})$ converges. We show that Equation 1 holds. Let $\varepsilon = 2^{-P}$ for some P. We consider every $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$ for $|Q| \leq P$; since $\phi(\mathbf{u})$ converges, it is Cauchy: there exists N_ϕ such that for $p, q \geq N_\phi$, we have $||\phi(u_p) - \phi(u_q)|| \leq \varepsilon$. The key observation is that there are finitely many such ϕ. Let N be the maximum of all the N_ϕ. For $p, q \geq N$, none of the above mentioned ϕ can (N, ε)-separate u_p and u_q, so $d(u_p, u_q) \leq \varepsilon$. This proves that Equation 1 holds, and concludes the proof. ∎

Lemma 4 (Equivalent Cauchy sequences). *Two Cauchy sequences of words* $\mathbf{u}$ *and* $\mathbf{v}$ *are equivalent if and only if for every* $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$*, we have* $\lim_n \phi(u_n) = \lim_n \phi(v_n)$.

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ two Cauchy sequences, denote by $\mathbf{w}$ the interleave sequence.

First assume that $\mathbf{u}$ and $\mathbf{v}$ are equivalent. Then $\mathbf{w}$ is Cauchy, so thanks to Lemma 3, for every $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$, the sequence $\phi(\mathbf{w})$ converges. Since $\phi(\mathbf{u})$ and $\phi(\mathbf{v})$ are subsequences of $\phi(\mathbf{w})$, it follows that they converge as well, and $\lim_n \phi(u_n) = \lim_n \phi(v_n)$.

Conversely, assume that for every $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$, we have $\lim_n \phi(u_n) = \lim_n \phi(v_n)$. Then for such ϕ, the sequence $\phi(\mathbf{w})$ converges, so it is Cauchy, and thanks to Lemma 3 this implies that $\mathbf{u}$ and $\mathbf{v}$ are equivalent. ∎

The elements in $\mathcal{P}A^*$ are called prostochastic words. Formally, such a word is an equivalence class of Cauchy sequences of finite words. In practice, we will often abuse the notations and identify Cauchy sequences and prostochastic words.

Now the objective is to prove that $\mathcal{P}A^*$ satisfies the Universal Property stated in Theorem 3. To this end, we rely on the following lemma from basic topology.

Lemma 5 (Uniformly Continuous Extensions in Complete Spaces). *Let* E *and* F *be metric spaces, such that* F *is complete. Denote by* $\widehat{E}$ *the topological completion of* E.

Every uniformly continuous function $\phi : E \to F$ *extends uniquely to a uniformly continuous function* $\widehat{\phi} : \widehat{E} \to F$.

Fact 3 *Let $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$. It uniquely extends to a uniformly continuous morphism $\phi : A^* \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$.*

Proof. The fact that it uniquely extends to a morphism is clear. We prove that the extension is uniformly continuous:

$$\forall \varepsilon > 0,\ \exists \eta,\ \forall u, v \in A^*,\ d(u, v) \leq \eta \Longrightarrow ||\phi(u) - \phi(v)|| \leq \varepsilon\ .$$

Let $\varepsilon > 0$. Consider $\eta = 2^{-P}$ such that $\eta \leq \varepsilon$ and $|Q| \leq P$. Let u, v such that $d(u, v) \leq \eta$. Since $|Q| \leq P$, it follows that $||\phi(u) - \phi(v)|| \leq \eta \leq \varepsilon$. ∎

We are now read to prove the Universal Property.

Lemma 6 (Universal Property of $\mathcal{P}A^*$). *Every $\phi : A \to M$, where M is a prostochastic monoid, extends uniquely to a uniformly continuous morphism $\widehat{\phi} : \mathcal{P}A^* \to M$.*

Proof. By construction, $\mathcal{P}A^*$ is the topological completion of A^*, and following the Fact 3, ϕ is uniformly continuous, hence the existence of a unique uniformly continuous extension $\widehat{\phi} : \mathcal{P}A^* \to M$ thanks to Lemma 1 and 5. Consider

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

To prove that $\widehat{\phi}$ is a morphism, it is enough to show that $D = \mathcal{P}A^* \times \mathcal{P}A^*$. Since $\widehat{\phi}$ extends the morphism ϕ on A^*, we already have that $A^* \times A^* \subseteq D$. Furthermore, A^* is dense in its topological completion $\mathcal{P}A^*$, so it suffices to show that D is closed. Indeed, $D = \bigcup_{m\in M} f^-(\{m\}) \cap g^-(\{m\})$, where:

$$f : \begin{cases} \mathcal{P}A^* \times \mathcal{P}A^* \to & \mathcal{P} \\ (\mathbf{u}, \mathbf{v}) & \mapsto \widehat{\phi}(\mathbf{u} \cdot \mathbf{v}) \end{cases} \qquad g : \begin{cases} \mathcal{P}A^* \times \mathcal{P}A^* \to & \mathcal{P} \\ (\mathbf{u}, \mathbf{u}) & \mapsto \widehat{\phi}(\mathbf{u}) \cdot \widehat{\phi}(\mathbf{v}) \end{cases}$$

Both functions are continuous as compositions of $\widehat{\phi}$, the product $\cdot$ in $\mathcal{P}A^*$ and the product $\cdot$ in $\mathcal{P}$. It follows that D is closed, so $D = \mathcal{P}A^* \times \mathcal{P}A^*$, implying that $\widehat{\phi}$ a morphism. ∎

The Universal Property will be instrumental in the following developments, for instance, it is used in Lemma 7 and in Theorem 4.

Lemma 7 ($\mathcal{P}A^*$ is Prostochastic). *$\mathcal{P}A^*$ is prostochastic.*

Proof. Let $\mathbf{u}$ and $\mathbf{u}$ be two Cauchy sequences that are not equivalent. We prove that there exists $\psi : \mathcal{P}A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$ such that $\psi(\mathbf{u}) \neq \psi(\mathbf{v})$.

Thanks to Lemma 3, since $\mathbf{u}$ and $\mathbf{v}$ are not equivalent, there exists $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$ such that $\lim_n \phi(u_n) \neq \lim_n \phi(v_n)$. Thanks to the Universal

Property stated in Lemma 6, ϕ extends to a uniformly continuous morphism $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$. We claim that $\widehat{\phi}(\mathbf{u}) \neq \widehat{\phi}(\mathbf{v})$. Indeed, by definition of $\widehat{\phi}$, we have that the sequence $\phi(\mathbf{u})$ converges to $\widehat{\phi}(\mathbf{u})$, and similarly the sequence $\phi(\mathbf{v})$ converges to $\widehat{\phi}(\mathbf{v})$. This concludes the proof. ∎

The proof of Theorem 3 is complete: we constructed $\mathcal{P}A^*$, and showed that it is a prostochastic monoid satisfying the Universal Property.

3.2 Reformulation of the Value 1 Problem

The aim of this subsection is to reformulate the value 1 problem, which talks about sequences of finite words, into an emptiness problem over prostochastic words.

Definition 6 (Prostochastic Language of a Probabilistic Automaton). *Let $\mathcal{A}$ be a probabilistic automaton and $\mathbf{u}$ a prostochastic word. We say that $\mathbf{u}$ is accepted by $\mathcal{A}$ if $\mathbf{u}$ is represented by some Cauchy sequence $\mathbf{u}$ such that $\lim_n P_{\mathcal{A}}(u_n) = 1$.*

We denote by $L(\mathcal{A})$ the set of prostochastic words accepted by $\mathcal{A}$.

Note that thanks to Lemma 3, $\mathbf{u}$ is accepted by $\mathcal{A}$ if and only if all Cauchy sequences $\mathbf{u}$ representing $\mathbf{u}$ satisfy $\lim_n P_{\mathcal{A}}(u_n) = 1$: it does not depend on the chosen representative.

Theorem 4 (The Value 1 Problem and the Emptiness Problem over Prostochastic Words). *Let $\mathcal{A}$ be a probabilistic automaton. The following are equivalent:*

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

Proof. Assume $\mathrm{val}(\mathcal{A}) = 1$, then there exists a sequence of words $\mathbf{u}$ such that $\lim_n P_{\mathcal{A}}(u_n) = 1$, *i.e.* $\lim_n I \cdot \phi(u_n) \cdot F = 1$. Thanks to the Universal Property stated in Lemma 6, ϕ extends to a uniformly continuous morphism $\widehat{\phi} : \mathcal{P}A^* \to \mathcal{S}_{Q\times Q}(\mathbb{R})$. We see the sequence $\mathbf{u}$ as a sequence of prostochastic words. By compactness of $\mathcal{P}A^*$ it contains a converging subsequence, which without loss of generality we assume is $\mathbf{u}$ itself. By continuity of $\widehat{\phi}$ and of the matrix product, we have $\lim_n I \cdot \widehat{\phi}(u_n) \cdot F = 1$, so $\lim_n P_{\mathcal{A}}(u_n) = 1$, *i.e.* the prostochastic word $\mathbf{u}$ belongs to $L(\mathcal{A})$.

Conversely, let $\mathbf{u}$ be a prostochastic word accepted by $\mathcal{A}$. Consider a sequence $\mathbf{u}$ representing $\mathbf{u}$. By definition, we have $\lim_n P_{\mathcal{A}}(u_n) = 1$, implying that $\mathrm{val}(\mathcal{A}) = 1$. ∎

4 Optimality of the Markov Monoid Algorithm

In this section, we use the prostochastic theory developed in the previous section to prove the optimality of the Markov Monoid algorithm. We first present the algorithm in Subsection 4.1, introduced in [FGO12]. We introduce two limit operators in Subsection 4.2, and our main technical tool, the fast Cauchy sequences, in Subsection 4.3. We give in Subsection 4.4 a characterization of the Markov Monoid algorithm using polynomial prostochastic words, and the Subsection 4.5 shows an undecidability result for exponential prostochastic words.

4.1 The Algorithm

The Markov Monoid algorithm was introduced in [FGO12]. The presentation that we give here is different yet equivalent. Consider $\mathcal{A}$ a probabilistic automaton, the Markov Monoid algorithm consists in computing, through a saturation process, the Markov Monoid of $\mathcal{A}$.

It is a monoid of boolean matrices: all numerical values are projected away to boolean values. Formally, for $M \in \mathcal{S}_{Q\times Q}(\mathbb{R})$, define its boolean projection $[M]$, as the boolean matrix such that $[M](s,t) = 1$ if $M(s,t) > 0$, and $[M](s,t) = 0$ otherwise. Hence to define the Markov Monoid, one can consider the underlying non-deterministic automaton $[\mathcal{A}]$ instead of the probabilistic automaton $\mathcal{A}$.

The Markov Monoid of $[\mathcal{A}]$ contains the transition monoid of $[\mathcal{A}]$, which is the monoid generated by $\{[\phi(a)] \mid a \in A\}$ and closed under (boolean matrix) products. 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 $[\mathcal{A}]$ satisfies the following: $\langle w\rangle(s,t) = 1$ if, and only if there exists a run from s to t reading w on $[\mathcal{A}]$.

The Markov Monoid generalizes the transition monoid by introducing a new operator, the stabilization. On the intuitive level first: let $M \in \mathcal{S}_{Q\times Q}(R)$, it can be interpreted as a Markov chain; its boolean projection $[M]$ give the structural properties of this Markov chain. The stabilization $[M]^\sharp$ accounts for $\lim_n M^n$, *i.e.* the behaviour of the Markov chain M in the limit. The formal definition of the stabilization operator relies on basic concepts from Markov chain theory.

Definition 7 (Stabilization). *Let M be a boolean matrix. It is said idempotent if $M \cdot M = M$.*

Assume M is idempotent, then we say that $t \in Q$ is M-recurrent if for all $s \in Q$, if $M(s,t) = 1$, then $M(t,s) = 1$.

The stabilization operator is defined only on idempotent elements:

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

The definition of the stabilization matches the intuition that in the Markov chain $\lim_n M^n$, the probability to be in non-recurrent states converges to 0. This will be made precise in Subsection 4.4.

Definition 8 (Markov Monoid). *The Markov Monoid of $\mathcal{A}$ is the smallest set of boolean matrices containing $\{[\phi(a)] \mid a \in A\}$ and closed under product and stabilization of idempotents.*

We give an equivalent presentation through ω-expressions, described by the following grammar:

$$u \quad \longrightarrow \quad a \quad | \quad u \cdot u \quad | \quad u^\omega \,.$$

We define an interpretation $\langle\cdot\rangle$ of ω-expressions into boolean matrices:

- $\langle a \rangle$ is $[\phi(a)]$,
- $\langle u_1 \cdot u_2 \rangle$ is $\langle u_1 \rangle \cdot \langle u_2 \rangle$,
- $\langle u^\omega \rangle$ is $\langle u \rangle^\sharp$, only defined if $\langle u \rangle$ is idempotent.

Then the Markov Monoid is $\{\langle u \rangle \mid u \text{ an } \omega\text{-expression}\}$.

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

Definition 9 (Value 1 Witnesses). *A boolean matrix M is a value 1 witness if: for all $s \in I$, $t \in Q$, if $M(s,t) = 1$, then $t \in F$.*

The Markov Monoid algorithm answers "YES" if there exists a value 1 witness in the Markov Monoid, and "NO" otherwise. The following has been proved in [FGO12]:

Theorem 5 ([FGO12]).

- *If the Markov Monoid algorithm answers "YES" on input $\mathcal{A}$, then the probabilistic automaton $\mathcal{A}$ has value 1,*
- *The converse does not hold in general: there exists a probabilistic automaton that has value 1, such that the Markov Monoid algorithm answers "NO",*
- *The Markov Monoid algorithm can be implemented in* PSPACE.

4.2 Limit Operators for Prostochastic Words

We show in this subsection how to construct non-trivial prostochastic words. In particular, we want to define a limit operator that accounts for the stabilization operation from the Markov Monoid. 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 will define two limit operators: an operator ω_P, where P stands for "polynomial", and an operator ω_E, where E stands for "exponential". First, we analyze the automaton represented in Figure 2, which was introduced in [GO10]. As explained in [FGO12,FGKO14], if $x > \frac{1}{2}$, then we have $\lim_n P_{\mathcal{A}}((ba^n)^{2^n}) = 1$, but $\lim_n P_{\mathcal{A}}((ba^n)^n) < 1$. This exhibits two different behaviours; the first one shall be accounted for by the exponential prostochastic word $(\mathbf{ba}^{\omega_\mathbf{P}})^{\omega_\mathbf{E}}$, the second by the polynomial prostochastic word $(\mathbf{ba}^{\omega_\mathbf{P}})^{\omega_\mathbf{P}}$.

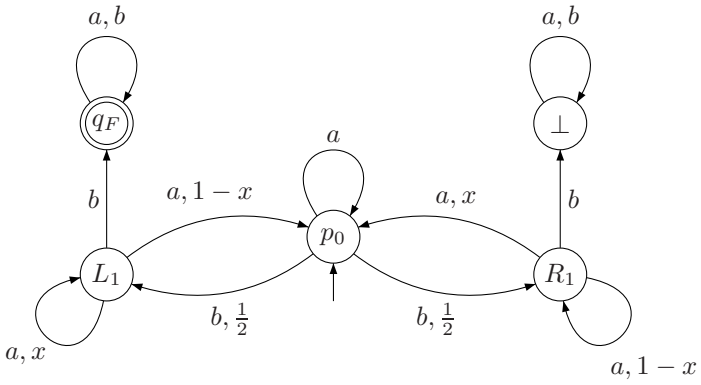

Fig. 2. Automaton accepting an exponential prostochastic word but no polynomial ones.

Informally speaking, this automaton consists of two symmetric parts, left and right. The left part leads to the accepting state, and the right part to the rejecting sink. To reach the accepting state with arbitrarily high probability, one needs to "tip the scales" to the left. Consider the following experiment, which consists in reading b and then a long sequence of a's. It results in the following situation: with high probability, the current state is p_0, with small probability it is L_1, and with even smaller probability it is R_1. To construct a sequence of words with arbitrarily high probability of being accepted, one has to play with this difference, and repeat the previous experiment many times. As shown by

precise calculations, what matters is that this experiment is repeated exponentially more than the length of the experiment, leading to the sequence of words $((ba^n)^{2^n})_{n\in\mathbb{N}}$.

We now turn to the definitions of ω_P and ω_E. Consider the two functions $f_P, f_E : \mathbb{N} \to \mathbb{N}$ defined as follows:

- $f_P(n) = k!$, where k is maximal such that $k! \leq n$,
- $f_E(n) = k!$, where k is maximal such that $k! \leq n^{\log(n)}$.

The function f_P grows linearly: roughly, $f_P(n) \sim n$, and the function f_E grows super-polynomially: roughly, $f_E(n) \sim n^{\log(n)}$. Both choices of n and $n^{\log(n)}$ are arbitrary; one could replace n by any polynomial, and $n^{\log(n)}$ by any function both super-polynomial and sub-exponential.

The functions f_P and f_E are *factorial-like*: for all $p \in \mathbb{N}$, there exists $k \in \mathbb{N}$, such that for all $n \geq k$, we have $p \mid f(n)$, *i.e.* p divides $f(n)$. The choice of factorial-like functions comes from the following classical result from Markov chain theory.

Lemma 8 (Powers of a Stochastic Matrix). *Let $M \in \mathcal{S}_{Q\times Q}(\mathbb{R})$ such that $[M]$ is idempotent, and $f : \mathbb{N} \to \mathbb{N}$ factorial-like.*

Then the sequence $(M^{f(n)})_{n\in\mathbb{N}}$ converges, denote M^∞ its limit. There exist two constants K and $C > 1$ such that

$$||M^{f(n)} - M^\infty|| \leq K \cdot C^{-f(n)} .$$

Furthermore, $[M^\infty] = [M]^\sharp$.

The two operators ω_P and ω_E take as input a sequence of finite words, and output a sequence of finite words. Formally, let $\mathbf{u}$ be a sequence of finite words, define:

$$\mathbf{u}^{\omega_P} = (u_n^{f_P(n\cdot|u_n|)})_{n\in\mathbb{N}} \qquad ; \qquad \mathbf{u}^{\omega_E} = (u_n^{f_E(n\cdot|u_n|)})_{n\in\mathbb{N}} .$$

It is not true in general that if $\mathbf{u}$ is Cauchy, then $\mathbf{u}^{\omega_P}$ is, nor is $\mathbf{u}^{\omega_E}$. In the next subsection, we will show that a sufficient condition is that $\mathbf{u}$ is fast.

4.3 Fast Sequences

This subsection introduces fast sequences, as the *key* technical tool for the proofs to follow. Indeed, we will prove that fast sequences are closed under concatenation, and that both operators ω_P and ω_E are defined for fast sequences. This will allow to define polynomial and exponential prostochastic words.

Definition 10 (Fast Sequences). *A sequence of finite words* $\mathbf{u}$ *is fast if it is Cauchy, and for every* $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$*, there exist a polynomial* P *and* $C > 1$ *such that*

$$||\phi(u_n) - \widehat{\phi}(\mathbf{u})|| \leq P(|u_n|) \cdot C^{-|u_n|} .$$

We start by closure under concatenation.

Lemma 9 (Concatenation and Fast Sequences). *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. Let $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$.

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

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

Lemma 10 (Limit Operators and Fast Sequences). *Let* $\mathbf{u}$ *be a fast sequence. Both sequences* $\mathbf{u}^{\omega_P}$ *and* $\mathbf{u}^{\omega_E}$ *are Cauchy. Moreover,* $\mathbf{u}^{\omega_P}$ *is fast.*

Proof. Let $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$, and $f \in \{f_P, f_E\}$.

Thanks to Lemma 8, the sequence $(\widehat{\phi}(\mathbf{u})^{f(n\cdot|u_n|)})_{n\in\mathbb{N}}$ converges, denote its limit by M^∞, there exists two constants K and $C_1 > 1$ such that $||\widehat{\phi}(\mathbf{u})^{f(n\cdot|u_n|)} - M^\infty|| \leq K \cdot C_1^{-f(n\cdot|u_n|)}$.

We proceed in two steps, using the following inequality:

$$\begin{aligned}
&||\phi(u_n^{f(n\cdot|u_n|)}) - M^\infty|| \\
&\quad \leq ||\phi(u_n)^{f(n\cdot|u_n|)} - \widehat{\phi}(\mathbf{u})^{f(n\cdot|u_n|)}|| + ||\widehat{\phi}(\mathbf{u})^{f(n\cdot|u_n|)} - M^\infty|| .
\end{aligned}$$

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

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

This gives:

$$
\begin{aligned}
&||\phi(u_n)^{f(n\cdot|u_n|)} - \widehat{\phi}(\mathbf{u})^{f(n\cdot|u_n|)}|| = \\
&||\sum_{k=0}^{f(n\cdot|u_n|)-1} \phi(u_n)^{f(n\cdot|u_n|)-k-1} \cdot (\phi(u_n) - \widehat{\phi}(\mathbf{u})) \cdot \widehat{\phi}^k(\mathbf{u})|| \\
&\leq \sum_{k=0}^{f(n\cdot|u_n|)-1} ||\phi(u_n)^{f(n\cdot|u_n|)-k-1}|| \cdot ||\phi(u_n) - \phi(\mathbf{u})|| \cdot ||\widehat{\phi}^k(\mathbf{u})|| \\
&\leq \sum_{k=0}^{f(n\cdot|u_n|)-1} ||\phi(u_n)||^{f(n\cdot|u_n|)-k-1} \cdot ||\phi(u_n) - \phi(\mathbf{u})|| \cdot ||\widehat{\phi}^k(\mathbf{u})|| \\
&= f(n\cdot|u_n|) \cdot ||\phi(u_n) - \widehat{\phi}(\mathbf{u})|| \,.
\end{aligned}
$$

Since $\mathbf{u}$ is fast, there exist a polynomial P and $C_2 > 1$ such that $||\phi(u_n) - \widehat{\phi}(\mathbf{u})|| \leq P(|u_n|) \cdot C_2^{-|u_n|}$. Altogether, we have

$$||\phi(u_n^{f(n\cdot|u_n|)}) - M^\infty|| \leq f(n\cdot|u_n|) \cdot P(|u_n|) \cdot C_2^{-|u_n|} + K \cdot C_1^{-f(n\cdot|u_n|)} \,.$$

It follows that $(\phi(u_n^{f(n\cdot|u_n|)}))_{n\in\mathbb{N}}$ converges, so both sequences $\mathbf{u}^{\omega_P}$ and $\mathbf{u}^{\omega_E}$ are Cauchy.

For $f = f_P$, we have $f(n\cdot|u_n|) \leq n\cdot|u_n|$, so there exist some polynomial Q and $C > 1$ such that $||\phi(u_n^{f(n\cdot|u_n|)}) - M^\infty|| \leq Q(|u_n|) \cdot C^{-|u_n|}$, implying that $\mathbf{u}^{\omega_P}$ is fast. ∎

We define an interpretation of ω-expressions into sequences of finite words:

- $\mathbf{a}$ is the constant sequence of the one-letter word a,
- $\mathbf{u_1} \cdot \mathbf{u_2}$ is the component-wise concatenation of $\mathbf{u}_1$ and $\mathbf{u}_2$,
- $\mathbf{u}^\omega$ is the sequence $\mathbf{u}^{\omega_P}$.

The sequences obtained as interpretation of ω-expressions are called *polynomial sequences*, and the prostochastic words defined by polynomial sequences are called *polynomial prostochastic words*.

A sequence is exponential if it is of the form $\mathbf{u}^{\omega_E}$, for some polynomial sequence $\mathbf{u}$. The prostochastic words defined by exponential sequences are called *exponential prostochastic words*.

Note that both notions are well defined thanks to Lemma 9 and 10.

4.4 A Characterization with Polynomial Prostochastic Words

The aim of this subsection is to prove that for given a probabilistic automaton $\mathcal{A}$, for every ω-expression u, the element $\langle u \rangle$ of the Markov Monoid of $\mathcal{A}$ is a value

1 witness if, and only if, the polynomial prostochastic word $\mathbf{u}$ is accepted by $\mathcal{A}$. This implies the following characterization 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 $\mathcal{A}$.

This is the first item of Theorem 1. It follows from the following proposition.

Proposition 1. *For all ω-expressions u, for every $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$, we have*

$$[\widehat{\phi}(\mathbf{u})] = \langle u \rangle \ .$$

Consequently, the element $\langle u \rangle$ of the Markov Monoid is a value 1 witness if, and only if, the regular prostochastic word $\mathbf{u}$ is accepted by $\mathcal{A}$.

We prove the first part of Proposition 1 by induction on the ω-expression u, which now essentially amounts to gather the results from the previous sections. The second part is a direct corollary of the first part.

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

The product case: let $w = u \cdot v$, and $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$.

We prove that $[\widehat{\phi}(\mathbf{w})] = \langle w \rangle$. Indeed, by definition $\widehat{\phi}(\mathbf{w}) = \widehat{\phi}(\mathbf{u}) \cdot \widehat{\phi}(\mathbf{v})$ and $\langle w \rangle = \langle u \rangle \cdot \langle v \rangle$, so the conclusion follows from the induction hypothesis.

The iteration case: let $v = u^\omega$, and $\phi : A \to \mathcal{S}_{Q\times Q}(\{0, \frac{1}{2}, 1\})$.

We prove that $[\widehat{\phi}(\mathbf{v})] = \langle v \rangle$. This follows from the definitions, the induction hypothesis and Lemma 8.

The proof of Proposition 1 is complete. It implies the first item of Theorem 1.

4.5 Undecidability for Exponential Prostochastic Words

The aim of this subsection is to show that undecidability is around the corner:

The following problem is undecidable:
given a probabilistic automaton $\mathcal{A}$,
determine whether there exists
an exponential prostochastic word accepted by $\mathcal{A}$.

This is the second item of Theorem 1.

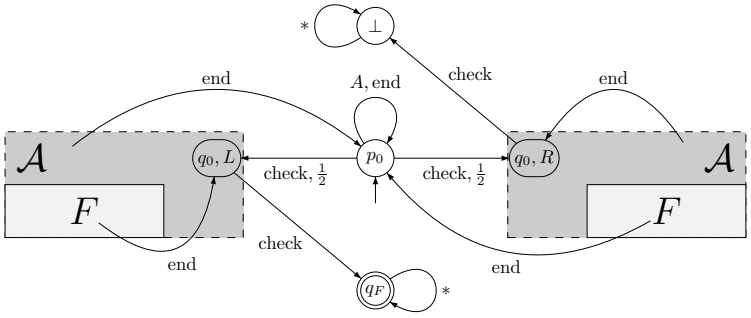

Fig. 3. Reduction.

Proof. We construct a reduction from the emptiness problem of probabilistic automata over finite words, proved to be undecidable in [Paz71]. Let $\mathcal{A}$ be a probabilistic automaton, we ask if there exists a finite word w such that $P_{\mathcal{A}}(w) > \frac{1}{2}$. We construct a probabilistic automaton $\mathcal{B}$ such that the following holds:

$$\begin{array}{c}\text{there exists a finite word } w \text{ such that } P_{\mathcal{A}}(w) > \frac{1}{2}\\ \text{if, and only if,}\\ \text{there exists an exponential prostochastic word accepted by } \mathcal{B}.\end{array}$$

The reduction is essentially as in [GO10], where they proved the undecidability of the value 1 problem. It is illustrated in Figure 3. In the original proof, it was enough to prove the existence of any prostochastic word accepted by $\mathcal{B}$. The challenge here is to improve the construction and the proof to show the existence of an exponential prostochastic word accepted by $\mathcal{B}$.

The automaton $\mathcal{B}$ is very similar to the one presented in Figure 2, except that the role of the letter a is now replaced by the simulation of a word in $\mathcal{A}$.

We fix the notations: the set of states of $\mathcal{A}$ is Q, its transition function is ϕ, without loss of generality we assume that it has a unique initial state q_0 (which has no ingoing transitions), and the set of final states is F.

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

final state is q_F. We define ϕ' as follows:

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

Assume that there exists a finite word w such that $P_{\mathcal{A}}(w) > \frac{1}{2}$, then we claim that $(\mathrm{check} \cdot (\mathbf{w} \cdot \mathrm{end})^{\omega_{\mathbf{P}}})^{\omega_{\mathbf{E}}}$ is accepted by $\mathcal{B}$. Denote $x = P_{\mathcal{A}}(w)$.

We have

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

and

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

We fix an integer N and analyze the action of reading $(\mathrm{check} \cdot (w \cdot \mathrm{end})^k)^N$: there are N "rounds", each of them corresponding to reading $\mathrm{check} \cdot (w \cdot \mathrm{end})^k$ from p_0. In a round, there are three outcomes: winning (that is, remaining in (q_0, L)) with probability $p_k = \frac{1}{2} \cdot x^k$, losing (that is, remaining in (q_0, R)) with probability $q_k = \frac{1}{2} \cdot (1 - x)^k$, 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:

$$\begin{aligned} P_{\mathcal{A}}((\mathrm{check} \cdot (w \cdot \mathrm{end})^k)^N) &= \textstyle\sum_{i=1}^{N-1} (1 - (p_k + q_k))^{i-1} \cdot p_k \\ &= p_k \cdot \tfrac{1 - (1 - (p_k + q_k))^{N-1}}{1 - (1 - (p_k + q_k))} \\ &= \tfrac{1}{1 + \frac{q_k}{p_k}} \cdot \left(1 - (1 - (p_k + q_k))^{N-1}\right) \end{aligned}$$

We now set $k = f_P(n \cdot (|w| + 1))$ and $N = f_E(n \cdot (1 + k \cdot (|w| + 1)))$. A simple calculation shows that the sequence $((1 - (p_k + q_k))^{N-1})_{n \in \mathbb{N}}$ converges

to 0 as n goes to infinity. Furthermore, $\frac{q_k}{p_k} = (\frac{1-x}{x})^k$, which converges to 0 as n goes to infinity since $x > \frac{1}{2}$. It follows that the acceptance probability converges to 1 as n goes to infinity. Consequently:

$$\lim_n P_{\mathcal{A}}((\mathrm{check} \cdot (w \cdot \mathrm{end})^k)^N) = 1 \ ,$$

i.e. $(\mathrm{check} \cdot (\mathbf{w} \cdot \mathrm{end})^{\omega_{\mathbf{P}}})^{\omega_{\mathbf{E}}}$ is accepted by $\mathcal{B}$.

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

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

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, so $\mathcal{B}$ accepts no prostochastic words thanks to Theorem 4. ∎

This concludes the optimality argument for the Markov Monoid algorithm, which consisted in first characterizing its computations using polynomial prostochastic words, and then showing that considering exponential prostochastic words leads to undecidability.

Conclusion and Perspectives

In this paper, we developed a profinite theory for probabilistic automata, called the prostochastic theory, and used it to formalize an optimality argument for the Markov Monoid algorithm. To the best of our knowledge, this is the first optimality argument for algorithms working on probabilistic automata.

This opens new perspectives. One of them is to further develop the prostochastic theory, for instance to better understand the class of fast prostochastic words, and another is to push our result further, using the prostochastic theory to construct an optimal algorithm for approximating the value.

References

[AGK+10] Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, editors. *Automata, Languages and Programming, 37th International Colloquium, ICALP 2010, Bordeaux, France, July 6-10, 2010, Proceedings, Part II*, volume 6199 of *Lecture Notes in Computer Science*. Springer, 2010.

[BBG12] Christel Baier, Nathalie Bertrand, and Marcus Größer. Probabilistic ω-automata. *Journal of the ACM*, 59(1):1, 2012.

[BMT77] Alberto Bertoni, Giancarlo Mauri, and Mauro Torelli. Some recursive unsolvable problems relating to isolated cutpoints in probabilistic automata. In Arto Salomaa and Magnus Steinby, editors, *Automata, Languages and Programming, Fourth Colloquium, University of Turku, Finland, July 18-22, 1977, Proceedings*, volume 52 of *Lecture Notes in Computer Science*, pages 87–94. Springer, 1977.

[CK97] Karel Culik and Jarkko Kari. *Digital images and formal languages*, pages 599–616. Springer-Verlag New York, Inc., 1997.

[CKV+11] Rohit Chadha, Vijay Anand Korthikanti, Mahesh Viswanathan, Gul Agha, and YoungMin Kwon. Model checking MDPs with a unique compact invariant set of distributions. In *QEST*, pages 121–130. IEEE Computer Society, 2011.

[CSV13] Rohit Chadha, A. Prasad Sistla, and Mahesh Viswanathan. Probabilistic automata with isolated cut-points. In Krishnendu Chatterjee and Jiri Sgall, editors, *MFCS*, volume 8087 of *Lecture Notes in Computer Science*, pages 254–265. Springer, 2013.

[CT12] Krishnendu Chatterjee and Mathieu Tracol. Decidable problems for probabilistic automata on infinite words. In *Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012, Dubrovnik, Croatia, June 25-28, 2012* [DBL12], pages 185–194.

[DBL12] *Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012, Dubrovnik, Croatia, June 25-28, 2012*. IEEE Computer Society, 2012.

[DEKM99] Richard Durbin, Sean R. Eddy, Anders Krogh, and Graeme Mitchison. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, July 1999.

[FGHO14] Nathanaël Fijalkow, Hugo Gimbert, Florian Horn, and Youssouf Oualhadj. Two recursively inseparable problems for probabilistic automata. In Erzsébet Csuhaj-Varjú, Martin Dietzfelbinger, and Zoltán Ésik, editors, *Mathematical Foundations of Computer Science 2014 - 39th International Symposium, MFCS 2014, Budapest, Hungary, August 25-29, 2014. Proceedings, Part I*, volume 8634 of *Lecture Notes in Computer Science*, pages 267–278. Springer, 2014.

[FGKO14] Nathanaël Fijalkow, Hugo Gimbert, Edon Kelmendi, and Youssouf Oualhadj. Deciding the value 1 problem for probabilistic leaktight automata. submitted, 2014.

[FGO12] Nathanaël Fijalkow, Hugo Gimbert, and Youssouf Oualhadj. Deciding the value 1 problem for probabilistic leaktight automata. In *Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012, Dubrovnik, Croatia, June 25-28, 2012* [DBL12], pages 295–304.

[GGP10] Mai Gehrke, Serge Grigorieff, and Jean-Éric Pin. A topological approach to recognition. In Abramsky et al. [AGK+10], pages 151–162.

[GO10] Hugo Gimbert and Youssouf Oualhadj. Probabilistic automata on finite words: Decidable and undecidable problems. In Abramsky et al. [AGK+10], pages 527–538.

[KMO+11] Stefan Kiefer, Andrzej S. Murawski, Joël Ouaknine, Björn Wachter, and James Worrell. Language equivalence for probabilistic automata. In Ganesh Gopalakrishnan and Shaz Qadeer, editors, *CAV*, volume 6806 of *Lecture Notes in Computer Science*, pages 526–540. Springer, 2011.

[KVAK10] Vijay Anand Korthikanti, Mahesh Viswanathan, Gul Agha, and YoungMin Kwon. Reasoning about MDPs as transformers of probability distributions. In *QEST*, pages 199–208. IEEE Computer Society, 2010.

[Moh97] Mehryar Mohri. Finite-state transducers in language and speech processing. *Computational Linguistics*, 23:269–311, June 1997.

[Paz71] Azaria Paz. *Introduction to probabilistic automata*. Academic Press, 1971.

[Pin09] Jean-Éric Pin. Profinite methods in automata theory. In Susanne Albers and Jean-Yves Marion, editors, *26th International Symposium on Theoretical Aspects of Computer Science, STACS 2009, February 26-28, 2009, Freiburg, Germany, Proceedings*, volume 3 of *LIPIcs*, pages 31–50. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2009.

[Rab63] Michael O. Rabin. Probabilistic automata. *Information and Control*, 6(3):230–245, 1963.

[Sch61] Marcel Paul Schützenberger. On the definition of a family of automata. *Information and Control*, 4(2-3):245–270, 1961.

[Tor11] Szymon Toruńczyk. *Languages of profinite words and the limitedness problem*. PhD thesis, University of Warsaw, 2011.