Polynomially Ambiguous Probabilistic Automata on Restricted Languages

Paul C. Bell
Department of Computer Science, Byrom Street, Liverpool John Moores University,
Liverpool, L3-3AF, UK
p.c.bell@ljmu.ac.uk

Abstract

We consider the computability and complexity of decision questions for Probabilistic Finite Automata (PFA) with sub-exponential ambiguity. We show that the emptiness problem for non-strict cut-points of polynomially ambiguous PFA remains undecidable even when the input word is over a bounded language and all PFA transition matrices are commutative. In doing so, we introduce a new technique based upon the Turakainen construction of a PFA from a Weighted Finite Automata which can be used to generate PFA of lower dimensions and of subexponential ambiguity. We also study freeness/injectivity problems for polynomially ambiguous PFA and study the border of decidability and tractability for various cases.

2012 ACM Subject Classification Theory of computation → Quantitative automata; Theory of computation → Probabilistic computation

Keywords and phrases Probabilistic finite automata, ambiguity, undecidability, bounded language, emptiness

Digital Object Identifier 10.4230/LIPIcs.ICALP.2019.105

Category Track B: Automata, Logic, Semantics, and Theory of Programming

Acknowledgements We thank the referees for their careful reading of this manuscript and their helpful improvements.

1 Introduction

Probabilistic finite automata (PFA) are a simple yet expressive model of computation, obtained by extending nondeterministic finite automata so that transitions from each state (and for each input letter) form probability distributions. As input letters are read from some alphabet Σ, the automaton transitions among states according to these probabilities. The probability of accepting a word $w \in \Sigma^*$ is given by the probability of the automaton being in one of its final states, denoted $f_{\mathcal{P}}(w) = \mathbf{x}^T M_{w_1} M_{w_2} \cdots M_{w_k} \mathbf{y}$, where $\mathbf{x}$ represents the initial state, $\mathbf{y}$ represents the final state and each M_{w_i} is a row stochastic matrix representing the transition probabilities for letter $w_i \in \Sigma$.

The PFA model has been studied extensively over the years, ever since its introduction by Rabin [27]; for example see [10] for a survey of 416 research papers related to PFA in the eleven years since their introduction to just 1974. They have been used to study Arthur-Merlin games [2], space bounded interactive proofs [15], quantum complexity theory [33], the joint spectral radius and semigroup boundedness [8], Markov decision processes and planning questions [9], and text and speech processing [24] among many others.

There are a variety of interesting questions that one may ask about PFA. A central question is the emptiness problem for cut-point languages; given some probability $\lambda \in [0, 1]$, does there exist a finite input word whose probability of acceptance is greater than λ (i.e. does there exist $w \in \Sigma^*$ such that $f_{\mathcal{P}}(w) > \lambda$, see Section 2.2). This problem is known to be undecidable [26], even for a fixed number of dimensions and for two input matrices [7, 19].

46th International Colloquium on Automata, Languages, and Programming (ICALP 2019).
Editors: Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi;
Article No. 105; pp. 105:1–105:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

A second natural question is the *freeness problem* (or *injectivity problem*) for PFA, studied in [3] – given a PFA $\mathcal{P}$ over alphabet Σ determine whether the acceptance function $f_{\mathcal{P}}(w)$ is injective (i.e. do there exist two distinct words with the same acceptance probability).

When studying the frontiers of decidability of a problem, there are two competing objectives, namely, determine the most general version of the problem which is decidable, and the most restricted specialization which is undecidable; the latter being the focus of this paper.

Various classes of restrictions may be studied for PFA depending upon the structure of the PFA or on possible input words. Some restrictions relate to the number of states of the automaton, the alphabet size and whether one defined the PFA over the algebraic real numbers or the rationals. One may also study PFA with finite, polynomial or exponential ambiguity (in terms of the underlying NFA), PFA defined for restricted input words (for example those coming from regular, bounded or letter monotonic languages), PFA with isolated thresholds (a probability threshold is isolated if it cannot be approached arbitrarily closely) and PFA where all matrices commute, for which cut-point languages and non-free languages generated by such automata necessarily become commutative.

The cut-point emptiness problem for PFA is known to be undecidable for rational matrices [26], even over a binary alphabet when the PFA has dimension 46 in [7]; later improved to dimension 25 [19]. The authors of [6] show that the problem of determining if a threshold is isolated (resp. if a PFA has any isolated threshold) is undecidable and this was shown to hold even for PFA with 420 (resp. 2354) states over a binary alphabet [7].

A natural restriction on PFA was studied in [4], where possible input words of the PFA are restricted to be from some letter monotonic language of the form $\mathcal{L} = a_1^* a_2^* \cdots a_k^*$ with each $a_i \in \Sigma$ (analogous to a 1.5 way PFA, whose read head may "stay put" on an input word letter but never moves left), then the problem remains undecidable. In other words, does there exist $w \in \mathcal{L}$ such that $f_{\mathcal{P}}(w) > \lambda$? This restriction is inspired by the well-known property that many language-theoretic problems become decidable or tractable when restricted to bounded languages, and especially letter-monotonic languages [13]. Nevertheless, the emptiness problem for PFA on letter-monotonic languages was shown to be undecidable for high (but finite) dimensional matrices over the rationals via an encoding of Hilbert's tenth problem on the solvability of Diophantine equations and the utilization of Turakainen's method to transform weighted integer automata to a PFA [4].

The authors of [17] recently studied decision problems for PFA of various degrees of ambiguity in order to map the frontier of decidability for restricted classes of PFA. The degree of ambiguity of a PFA is defined as the maximum number of accepting runs over all possible words and can be used to give various classifications of ambiguity including finite, polynomial and exponential ambiguity. The ambiguity of a PFA is a property of the underlying NFA and is independent of the transition probabilities in so much as we only need care if the probability is zero or positive. The degree of ambiguity of automata is a well-known and well-studied property in automata theory [31]. The authors of [17] show that the emptiness problem for PFA remains undecidable even for polynomially ambiguous automata (quadratic ambiguity), before going on to show **PSPACE**-hardness results for finitely ambiguous PFA and that emptiness is in **NP** for the class of k-ambiguous PFA for every $k > 0$. The emptiness problem for PFA was later shown to also be undecidable even for linearly ambiguous automata in [16].

1.1 Our Contributions

In this paper, we show that the emptiness problem is undecidable even for polynomially ambiguous PFA defined over letter monotonic languages when all matrices are rational and commutative. This combination of restrictions on the PFA significantly increases the difficulty of proving undecidability. The study of PFA over letter monotonic languages is a particularly interesting intermediate model, lying somewhere between single letter alphabets, for which we have decidability results, and PFA defined with multi-letter alphabets, for which most decision problems are undecidable.

▶ **Theorem 1.** *The emptiness problem for polynomially ambiguous probabilistic finite automata on letter monotonic languages is undecidable for non-strict cut-points, even when all matrices are commutative.*

We note a few difficulties with proving this result. Firstly, Post's correspondence problem, whose variants are often used for showing undecidability results in such settings, is actually decidable over letter monotonic languages [18][1]. Secondly, although other reductions of undecidable computational problems to matrices are possible, the standard technique of Turakainen (shown in [30]) to modify such matrices to stochastic matrices introduces exponential ambiguity (indeed all such matrices are strictly positive, and thus we might think of such matrices as being *maximally exponentially ambiguous*)[2]. Finally, we note that matrix problems for commutative matrices are often decidable; indeed there are polynomial time algorithms for solving the orbit problem [22, 14] and the vector reachability problem for commutative matrices [1]. Since the matrices commute, it is the Parikh vector of letters of the input word which is important.

We use a reduction of Hilbert's tenth problem and various new encoding techniques to avoid the use of Turakainen's method for converting from weighted to probabilistic automata, so as to retain polynomial ambiguity. We then move on to the freeness/injectivity problem to show the following two results.

▶ **Theorem 2.** *The injectivity problem for linearly ambiguous four state probabilistic finite automata is undecidable.*

▶ **Theorem 3.** *The injectivity problem for linearly ambiguous three-state probabilistic finite automata over letter-monotonic languages is NP-hard.*

These results are proven via an encoding of the mixed modification PCP and our new encoding technique and the injectivity problem for three state PFA over letter monotonic languages is **NP**-hard via an encoding of the variant of the subset sum problem and a novel encoding technique. We conclude with some open problems.

2 Preliminaries

2.1 Linear Algebra

Given $A = (a_{ij}) \in \mathbb{F}^{m\times m}$ and $B \in \mathbb{F}^{n\times n}$, we define the direct sum $A \oplus B$ and Kronecker product $A \otimes B$ of A and B by:

$$A \oplus B = \left[\begin{array}{c|c} A & \mathbf{0}_{m,n} \\ \hline \mathbf{0}_{n,m} & B \end{array}\right], \quad A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mm}B \end{bmatrix},$$

[1] Although it is undecidable in general (i.e. not over a letter monotonic language) with an alphabet with at least five letters [25].

[2] This is due to an essential step of the Turakainen procedure that adds a positive constant offset to each element of every generator matrix, thus making all matrices strictly positive [30].

where $\mathbf{0}_{i,j}$ denotes the zero matrix of dimension $i \times j$. Note that neither $\oplus$ nor $\otimes$ are commutative in general. Given a finite set of matrices $\mathcal{G} = \{G_1, G_2, \ldots, G_m\} \subseteq \mathbb{F}^{n\times n}$, $\langle\mathcal{G}\rangle$ is the semigroup generated by $\mathcal{G}$. We will use the following notations:

$$\bigoplus_{j=1}^{m} G_j = G_1 \oplus G_2 \oplus \cdots \oplus G_m, \qquad \bigotimes_{j=1}^{m} G_j = G_1 \otimes G_2 \otimes \cdots \otimes G_m$$

Given a single matrix $G \in \mathbb{F}^{n\times n}$, we inductively define $G^{\otimes k} = G \otimes G^{\otimes(k-1)} \in \mathbb{F}^{n^k \times n^k}$ for $k > 0$ with $G^{\otimes 0} = 1$ as the k-fold Kronecker power of G. Similarly, $G^{\oplus k} = G \oplus G^{\oplus(k-1)} \in \mathbb{F}^{n^k \times n^k}$ for $k > 0$ with $G^{\oplus 0}$ being a zero dimensional matrix. The rationalle for the base cases is that $G \otimes G^{\otimes 0} = 1 \otimes G = G$ and that $G \oplus G^{\oplus 0} = G$ as expected.

The following properties of $\oplus$ and $\otimes$ are well known, see [20] for proofs.

▶ **Lemma 4.** *Let $A, B, C, D \in \mathbb{F}^{n\times n}$. We note that:*

- *Associativity - $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ and $(A \oplus B) \oplus C = A \oplus (B \oplus C)$, thus $A \otimes B \otimes C$ and $A \oplus B \oplus C$ are unambiguous.*
- *Mixed product properties: $(A\otimes B)(C\otimes D) = (AC\otimes BD)$ and $(A\oplus B)(C\oplus D) = (AC\oplus BD)$.*
- *If A and B are stochastic matrices, then so are $A \oplus B$ and $A \otimes B$.*

It is trivial to prove that if $A, B \in \mathbb{F}^{n\times n}$ are both upper-triangular then so are $A \oplus B$ and $A \otimes B$. This follows directly from the definition of the Kronecker sum and product.

2.2 Probabilistic Finite Automata (PFA)

Probabilistic Finite Automata (PFA) with n states over an alphabet Σ are defined as $\mathcal{P} = (\mathbf{x}, \{M_a | a \in \Sigma\}, \mathbf{y})$ where $\mathbf{x} \in \mathbb{R}^n$ is the initial probability distribution; $\mathbf{y} \in \{0,1\}^n$ is the final state vector and each $M_a \in \mathbb{R}^{n\times n}$ is a (row) stochastic matrix. For a word $w = w_1 w_2 \cdots w_k \in \Sigma^*$, we define the acceptance probability $f_{\mathcal{P}} : \Sigma^* \to \mathbb{R}$ of $\mathcal{P}$ as:

$$f_{\mathcal{P}}(w) = \mathbf{x}^T M_{w_1} M_{w_2} \cdots M_{w_k} \mathbf{y},$$

which denotes the acceptance probability of w.

For any $\lambda \in [0,1]$ and PFA $\mathcal{A}$ over alphabet Σ, we define a cut-point language to be: $L_{\geq\lambda}(\mathcal{A}) = \{w \in \Sigma^* | f_{\mathcal{A}}(w) \geq \lambda\}$, and a strict cut-point language $L_{\geq\lambda}(\mathcal{A})$ by replacing $\geq$ with $>$. The (strict) emptiness problem for a cut-point language is to determine if $L_{\geq\lambda}(A) = \emptyset$ (resp. $L_{>\lambda}(A) = \emptyset$).

Let $\Sigma_\ell = \{x_1, x_2, \ldots, x_\ell\}$ be an ℓ-letter alphabet for some $\ell > 0$. A language $L \subseteq \Sigma_\ell^*$ is called a *bounded language* if and only if there exist words $w_1, w_2, \ldots, w_m \in \Sigma_\ell^+$ such that $L \subseteq w_1^* w_2^* \cdots w_m^*$. A language L is called *letter-monotonic* if there exists letters $u_1, u_2, \ldots, u_m \in \Sigma_\ell$ such that $L \subseteq u_1^* u_2^* \cdots u_m^*$. One thus sees that letter monotonic languages are more restricted than bounded languages. We will be interested in PFA which are defined over a bounded language or a letter monotonic language $\mathcal{L}$, whereby all input words necessarily come from $\mathcal{L}$. In this case a cut-point language for a PFA $\mathcal{P}$ over bounded/letter monotonic language $\mathcal{L}$ and a probability $\lambda \in [0,1]$ is defined as $L_{\geq\lambda}(\mathcal{A}) = \{w \in L | f_{\mathcal{A}}(w) \geq \lambda\}$; similarly for nonstrict cut point languages. We may then ask similar emptiness questions for such languages, as before.

We also study the *freeness/injectivity problem* for PFA. Given a PFA $\mathcal{P}$ over alphabet Σ determine whether the acceptance function $f_{\mathcal{P}}(w)$ is injective (i.e. do there exist two distinct words with the same acceptance probability). Such problems can readily be studied when the input words are necessarily derived from a bounded or letter-monotonic language.

2.3 PFA Ambiguity

The degree of ambiguity of a finite automaton is a structural parameter, roughly indicating the number of accepting runs for a given input word [31]. We here define only those notions required for our later proofs, see [31] for full details of these notions and a thorough discussion.

Let $w \in \Sigma^*$ be an input word of an NFA $\mathcal{N} = (Q, \Sigma, \delta, Q_I, Q_F)$. For each $(p, w, q) \in Q \times \Sigma^* \times Q$, let $\mathrm{da}_{\mathcal{N}}(p, w, q)$ be defined as the number of all paths for w in $\mathcal{N}$ leading from state p to state q. The degree of ambiguity of w in $\mathcal{N}$, denoted $\mathrm{da}_{\mathcal{N}}(w)$, is defined as the number of all *accepting paths* for w. The degree of ambiguity of $\mathcal{N}$, denoted $\mathrm{da}(\mathcal{N})$ is the supremum of the set $\{\mathrm{da}_{\mathcal{N}}(w) | w \in \Sigma^*\}$. $\mathcal{N}$ is called infinitely ambiguous if $\mathrm{da}(\mathcal{N}) = \infty$, finitely ambiguous if $\mathrm{da}(\mathcal{N}) < \infty$, and unambiguous if $\mathrm{da}(\mathcal{N}) \leq 1$. The degree of growth of the ambiguity of $\mathcal{N}$, denoted $\deg(M)$ is defined as the minimum degree of a univariate polynomial h with positive integral coefficients such that for all $w \in \Sigma^*$, $\mathrm{da}_{\mathcal{N}}(w) \leq h(|w|)$ if such a polynomial exists, or infinity otherwise.

The above notions relate to NFA. We may derive an analogous notation of ambiguity for PFA by considering an embedding of a PFA $\mathcal{P}$ to an NFA $\mathcal{N}$ with the property that for each letter $a \in \Sigma$, if the probability of transitioning from a state i to state j is nonzero under $\mathcal{P}$, then there is an edge from state i to j under $\mathcal{N}$ for letter a. The degree of (growth of) ambiguity of $\mathcal{P}$ is then defined as the degree of (growth of) ambiguity of $\mathcal{N}$.

We may use the following notions to determine the degree of ambiguity of a given NFA (and thus a PFA) $\mathcal{A}$ as is shown in the theorem which follows. A state $q \in Q$ is called *useful* if there exists an accepting path which visits q.

EDA. There is a useful state $q \in Q$ such that, for some word $v \in \Sigma^*$, $da_{\mathcal{A}}(q, v, q) \geq 2$.

IDA$_d$. There are useful states $r_1, s_1, \ldots, r_d, s_d \in Q$ and words $v_1, u_2, v_2, \ldots, u_d, v_d \in \Sigma^*$ such that for all $1 \leq \lambda \leq d$, r_λ and s_λ are distinct and $(r_\lambda, v_\lambda, r_\lambda), (r_\lambda, v_\lambda, s_\lambda), (s_\lambda, v_\lambda, s_\lambda) \in \delta$ and for all $2 \leq \lambda \leq d$, $(s_{\lambda-1}, u_\lambda, r_\lambda) \in \delta$.

▶ **Theorem 5** ([21, 28, 31])**.** *An NFA (or PFA) $\mathcal{A}$ having the EDA property is equivalent to it being exponentially ambiguous. For any $d \in \mathbb{N}$, an NFA (or PFA) $\mathcal{A}$ having property IDA$_d$ is equivalent to $\deg(\mathcal{A}) \geq d$.*

Clearly, if $\mathcal{N}$ agrees with IDA$_d$ for some $d > 0$, then it also agrees with IDA$_1, \ldots,$ IDA$_{d-1}$. One must be careful with these notions of ambiguity when considering NFA/PFA $\mathcal{A}$, where inputs are necessarily from a bounded language $\mathcal{L}$. In such cases, the above criteria do not suffice to determine the ambiguity of $\mathcal{A}$, since the number of paths must be determined not over Σ^*, but over all paths from $\mathcal{L}$. Of course, the degree of ambiguity of $\mathcal{A}$ cannot *increase* by restricting to a bounded input language, but it may decrease.

As an example, if an NFA has property EDA, then there exists three words w_1, w_2 and w_3 such that $w_1w_2w_3$ is an accepting word and $da_{\mathcal{A}}(q, w_2, q) \geq 2$, thus $w_1w_2w_3$ has at least two distinct accepting runs. However, this implies that $da_{\mathcal{A}}(w_1w_2^kw_3) \geq 2^k$ and thus $w_1w_2^kw_3$ has at least 2^k accepting runs. Now, if we are given some bounded language $\mathcal{L}$ such that $w_1w_2w_3 \in \mathcal{L}$ and $da_{\mathcal{A}}(q, w_2, q) \geq 2$ then the same implication is not possible, unless $w_2 \in \Sigma$ is a single letter, otherwise there is no guarantee that $w_1w_2^kw_3 \in \mathcal{L}$. Nevertheless, in the results of this paper we will use the standard definitions of ambiguity since the distinction is not relevant in our results as will become clear.

We note the following trivial lemma, which will be useful later.

▶ **Lemma 6.** *Probabilistic finite automata defined over upper-triangular matrices are polynomially ambiguous.*

Proof. This lemma is immediate from Theorem 5 and property (EDA), since a PFA defined over upper-triangular matrices clearly does not have property (EDA). This is since a transition matrix (for a letter "a") which is upper-triangular only defines transitions of the form $\delta(i,a) = j$ where $i \leq j$ and thus the states entered for any run are monotonically nondecreasing. ◀

2.4 Reducible Undecidable Problems

We will require the following undecidable problems for proving later results. The first is a variant of the famous *Post's Correspondence Problem* (PCP).

▶ **Problem 7** (Mixed Modification PCP (MMPCP)). *Given a binary alphabet Σ_2, a finite set of letters $\Sigma = \{s_1, s_2, \ldots, s_{|\Sigma|}\}$, and a pair of homomorphisms $h, g : \Sigma^* \to \Sigma_2^*$, the MMPCP asks to decide whether there exists a word $w = x_1 \ldots x_k \in \Sigma^+, x_i \in \Sigma$ such that*

$$h_1(x_1)h_2(x_2)\ldots h_k(x_k) = g_1(x_1)g_2(x_2)\ldots g_k(x_k),$$

where $h_i, g_i \in \{h, g\}$, and there exists at least one j such that $h_j \neq g_j$.

▶ **Theorem 8** ([12]). *The Mixed Modification PCP is undecidable for $|\Sigma| \geq 9$.*

A second useful undecidable problem is *Hilbert's tenth problem*: Let $P(n_1, n_2, \ldots, n_k)$ be an integer polynomial with k variables - determine if there exists a procedure to find if there exist $x_1, x_2, \ldots, x_k \in \mathbb{Z}$ such that: $P(x_1, x_2, \ldots, x_k) = 0$. It is well known that this may be reduced to a problem in formal power series. It was shown in [29, p.73] that the above problem can be reduced to that of determining for a $\mathbb{Z}$-rational formal power series $S \in \mathbb{Z}\langle\langle A \rangle\rangle$, whether there exists any word $w \in A^*$ such that $(S, w) = 0$. The undecidability of this problem was shown in 1970 by Y. Matiyasevich (building upon work of Davis, Putman, Robinson and others). For more details, see the excellent reference [23]. We may, without loss of generality, restrict the variables to be natural numbers [23, p.6].

3 Cut-point languages for polynomially ambiguous PFA over letter monotonic languages

It was proven in [4] that the emptiness problem is undecidable for probabilistic finite automata even when input words are given over a letter-monotonic language, i.e., given a letter-monotonic language $\mathcal{L}$, it is undecidable to determine if $\{w \in \mathcal{L} | f_{\mathcal{P}}(w) \Delta \lambda\}$ is empty for $\Delta \in \{\leq, <, >, \geq\}$. The constructed PFA $\mathcal{P}$ of [4] has exponential ambiguity, due to the well-known Turakainen conversion of arbitrary integer matrices into stochastic matrices. Here, we show that the emptiness problem for PFA over letter-monotonic languages can also be achieved even when all matrices have polynomial ambiguity by a modified Turakainen procedure.

The following property of the Kronecker product will also be required for the proof of Theorem 1.

▶ **Lemma 9.** *Let $A_1, \ldots, A_\ell \in \mathbb{F}^{n \times n}$. Then for any index sequence $(i_1, j_1), \ldots, (i_\ell, j_\ell)$ with each $(i_s, j_s) \in [1, n] \times [1, n]$ then there exists $1 \leq i, j \leq n^\ell$ such that*

$$\prod_{m=1}^{\ell} (A_m)_{i_m, j_m} = \left(\bigotimes_{m=1}^{\ell} A_m \right)_{i,j}$$

Proof. The proof proceeds by induction. For the base case when $\ell = 1$, we just set $(i, j) = (i_1, j_1)$ and we are done. Assume then that the result holds for some $\ell - 1$, then for sequence $(i_1, j_1), (i_2, j_2), \ldots, (i_{\ell-1}, j_{\ell-1})$ there exists $1 \le i', j' \le n^{\ell-1}$ such that:

$$\prod_{m=1}^{\ell-1} (A_m)_{i_m,j_m} = \left(\bigotimes_{m=1}^{\ell-1} A_m\right)_{i',j'}$$

By the definition of Kronecker product,

$$\left(\left(\bigotimes_{m=1}^{\ell-1} A_m\right) \otimes A_\ell\right)_{ni'+i_\ell, nj'+j_\ell} = \prod_{m=1}^{\ell-1} (A_m)_{i_m,j_m} \times (A_\ell)_{i_\ell,j_\ell}$$

as required. ◀

Note that we can of course work out the particular value of i and j, but in general the formula for i, j does not have a nice form when $\ell > 2$, and anyway will not be necessary for us, so we settle for an existential proof of such i and j.

3.1 Proof of Theorem 1

Proof. We will construct a polynomially ambiguous probabilistic finite automaton $\mathcal{P}$, a cut-point $\lambda \in [0, 1]$ and a letter monotonic language $\mathcal{L}$.

We begin by encoding an instance of Hilbert's tenth problem into a set of integer matrices. Let $P(x_1, x_2, \ldots, x_t) = 0$ be a Diophantine equation. Homogenenization of polynomials is a well known technique, as is used for example in the study of Gröbner bases [11], which allows us to convert such a Diophantine equation to $P^h(x_0, x_1, x_2, \ldots, x_t) = 0$ with a new dummy variable x_0 such that P^h is a homogeneous polynomial (each term having the same degree d) and for which $P^h(x_0, x_1, \ldots, x_t) = P(x_1, x_2, \ldots, x_t)$ when $x_0 = 1$. We thus assume a homogeneous Diophantine equation $P^h(x_0, x_1, \ldots, x_t) = 0$ with implied constraint $x_0 = 1$ which will be dealt with later. Furthermore, we assume that P^h gives nonnegative values, which may be assumed by redefining $P^h = (P^h)^2$, which clearly does not affect whether a zero exists for such a polynomial.

Notice that given $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, then $A^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$. We will generalise this property to a set of $t+1$ matrices $A_0, A_1, \ldots, A_t \in \mathbb{Z}^{(t+3)\times(t+3)}$ so that given any tuple $(x_0, x_1, x_2, \ldots, x_t)$, then x_i appears as an element on the superdiagonal of $A_0^{x_0} A_1^{x_1} \cdots A_t^{x_t}$ for each $0 \le i \le t$. We will also have the property that each A_i has the same row sum of 2 for every row, which will be useful when we later convert to stochastic matrices.

We define each matrix A_i for $0 \le i \le t+1$ in the following way:

$$A_i = \begin{pmatrix} 1 & \delta_{0,i} & 0 & \cdots & 0 & 0 & 1-\delta_{0,i} \\ 0 & 1 & \delta_{1,i} & \cdots & 0 & 0 & 1-\delta_{1,i} \\ 0 & 0 & 1 & \cdots & 0 & 0 & 1-\delta_{2,i} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \delta_{t,i} & 1-\delta_{t,i} \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 2 \end{pmatrix} \in \mathbb{N}^{(t+3)\times(t+3)}, \tag{1}$$

where $0 \le j \ne i \le t$ and $\delta_{\ell,i} \in \{0, 1\}$ is the Kronecker delta (thus $\delta_{i,i} = 1$ and $\delta_{\ell,i} = 0$ for $\ell \ne i$). We also denote $J = A_{t+1}$, noting that this is the matrix (1) when all $\delta_{\ell,i}$ have the value 0. Notice then that every row sum of A_i and J is 2. This structure is retained under matrix powers and it is easy to see that:

$$A_i^k = \begin{pmatrix} 1 & k\delta_{0,i} & 0 & \cdots & 0 & 0 & 2^k - k\delta_{0,i} - 1 \\ 0 & 1 & k\delta_{1,i} & \cdots & 0 & 0 & 2^k - k\delta_{1,i} - 1 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 2^k - k\delta_{2,i} - 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & k\delta_{t,i} & 2^k - k\delta_{t,i} - 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 2^k - 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 2^k \end{pmatrix} \in \mathbb{N}^{(t+3)\times(t+3)} \tag{2}$$

All row sums of A_i^k are 2^k and exactly one element of the superdiagonal is equal to k, with all other elements on the superdiagonal (excluding that on row $t+2$) zero. Taking powers of A_i will allow us to choose any positive value of variable x_i. Note that J^k has the same form as the matrix of (2) with all $\delta_{\ell,i} = 0$ and acts as a kind of identity matrix, (in its upperleft block) while retaining the 2^k row sum. Indeed, one sees that for all $0 \leq i, j \leq t+1$, then $A_iA_j = A_jA_i$, i.e. these matrices commute (as does J since $J = A_{t+1}$). We now show how to compute terms of P^h.

We may write $P^h(x_0, x_1, \ldots, x_t) = \sum_{j=1}^r T_j(x_0, x_1, \ldots, x_t)$, where T_j denotes the j'th term of P^h, with P^h having r terms. Since P^h is a homogeneous polynomial, each term has the same degree d. We may thus write each term as:

$$T_j(x_0, x_1, \ldots, x_t) \quad = \quad c_jR_j(x_0, x_1, \ldots, x_t), \tag{3}$$

with $c_j \in \mathbb{Z}$ and $R_j(x_0, x_1, \ldots, x_t) = \prod_{\ell=0}^t x_\ell^{r_{j,\ell}}$ with $r_{j,\ell} \geq 0$ and $\sum_{\ell=0}^t r_{j,\ell} = d$. For convenience, we define a d-dimensional vector $s_j = \bigotimes_{\ell=0}^t \ell^{\otimes r_{j,\ell}} \in [0,t]^d$. For example, if $t = 4, d = 8$ and $T_j(x_0, x_1, x_2, x_3, x_4) = 6x_1^2x_3^5x_4$, then $R_j(x_0, x_1, x_2, x_3, x_4) = x_0^0x_1^2x_2^0x_3^5x_4^1$ and thus $s_j = (1, 1, 3, 3, 3, 3, 3, 4)^T \in [0, 4]^8$. By $s_j[i]$ we denote the i'th element of vector s_j.

We now define $t+1$ matrices corresponding to term T_j:

$$X_{j,i} = \bigotimes_{\ell=0}^{i-1} J^{\otimes r_{j,\ell}} \otimes A_i^{\otimes r_{j,i}} \otimes \bigotimes_{\ell=i+1}^{d} J^{\otimes r_{j,\ell}},$$

where $0 \leq i \leq t$. The dimension of such matrices is $(t+3)^d \times (t+3)^d$ since each submatrix has dimension $(t+3) \times (t+3)$ and we take the d-fold Kronecker product. Similarly, we see that the row sum of each $X_{j,i}$ is 2^d since the row sum of each A_i and J is 2 and we take a d-fold Kronecker product. Clearly then, by the mixed product property (see Lemma 4):

$$X_{j,i}^k = \bigotimes_{\ell=0}^{i-1} (J^k)^{\otimes r_{j,\ell}} \otimes (A_i^k)^{\otimes r_{j,i}} \otimes \bigotimes_{\ell=i+1}^{d} (J^k)^{\otimes r_{j,\ell}},$$

for any $k \geq 0$. In the example when $r_{j,0} = 0$, $r_{j,1} = 2$, $r_{j,2} = 0$, $r_{j,3} = 5$ and $r_{j,4} = 1$, then $X_{j,3} = J^{\otimes 0} \otimes J^{\otimes 2} \otimes J^{\otimes 0} \otimes A_i^{\otimes 5} \otimes J^{\otimes 1} = J^{\otimes 2} \otimes A_i^{\otimes 5} \otimes J$. We then see that $X_{j,3}^k = (J^k)^{\otimes 2} \otimes (A_3^k)^{\otimes 5} \otimes J^k$.

Now, we see that:

$$\begin{aligned} X_{j,0}^{x_0}X_{j,1}^{x_1} \cdots X_{j,t}^{x_t} \quad &= \quad \prod_{i=0}^{t} \left(\bigotimes_{\ell=0}^{i-1} (J^{x_i})^{\otimes r_{j,\ell}} \otimes (A_i^{x_i})^{\otimes r_{j,i}} \otimes \bigotimes_{\ell=i+1}^{d} (J^{x_i})^{\otimes r_{j,\ell}} \right) && (4) \\ &= \quad \bigotimes_{\ell=0}^{d} \left(D_{\ell,0}^{x_0} D_{\ell,1}^{x_1} \cdots D_{\ell,t}^{x_t} \right), && (5) \end{aligned}$$

where $D_{\ell,i} \in \{J, A_i\}$ for $0 \le i \le t$. The derivation of Eqn (5) from Eqn (4) follows by the mixed product property of the Kronecker product (Lemma 4). For each product $D_{\ell,0}^{x_0}D_{\ell,1}^{x_1}\cdots D_{\ell,t}^{x_t}$, we see that $D_{\ell,s_j[\ell]} = A_{s_j[\ell]}$ and $D_{\ell,j} = J$ for all $0 \le j \le d$ with $j \ne s_j[\ell]$. As discussed earlier, matrices A_i and J commute, for any $0 \le i \le t$ and thus we may rewrite (5) as:

$$X_{j,0}^{x_0}X_{j,1}^{x_1}\cdots X_{j,t}^{x_t} = \bigotimes_{\ell=0}^{d}\left(A_{s_j[\ell]}^{x_{s_j[\ell]}} \otimes J^{\overline{x_{s_j[\ell]}}}\right), \text{ where } \overline{x_{s_j[\ell]}} = \sum_{\substack{0\le q\le d\\ q\ne s_j[l]}} x_q \tag{6}$$

By Lemma 9, we see that some element of $X_{j,0}^{x_0}X_{j,1}^{x_1}\cdots X_{j,t}^{k_t}$ is thus equal to $R_j(x_0, x_1, \ldots, x_t)$ as required, since there is an element on the superdiagonal of $A_{s_j[\ell]}^{x_{s_j[\ell]}}$ equal to $x_{s_j[\ell]}$ for each $0 \le \ell \le d$. Let us assume that $R_j(x_0, x_1, \ldots, x_t)$ appears at row i_1 and column i_2. Now, we may define a vector $u'_j = c_j e_{i_1}$ and $v'_j = e_{i_2}$ where c_j is the coefficient of term T_j as in Eqn (3) and $e_{i_1}, e_{i_2} \in \mathbb{Z}^{(t+3)^d}$ are basis vectors. We may now see that

$$(u'_j)^T X_{j,0}^{x_0}X_{j,1}^{x_1}\cdots X_{j,t}^{x_t} v'_j = c_j R_j(x_0, x_1, \ldots, x_t) = T_j(x_0, x_1, \ldots, x_t) \tag{7}$$

In order to derive the sum of the r such terms $\sum_{j=1}^{r} T_j(x_0, x_1, \ldots, x_t)$, we will utilise the *direct sum*. For $0 \le \ell \le d$, we define Y'_ℓ by:

$$Y'_\ell = \bigoplus_{j=1}^{r} X_{j,\ell} \in \mathbb{N}^{r(t+3)^d \times r(t+3)^d}$$

We shall now modify each Y'_ℓ so that they are row stochastic. We recall that the row sum of each A_ℓ and J is 2. Therefore, the row sum of each $X_{j,\ell}$ is 2^d, since $X_{j,\ell}$ is a d-fold Kronecker product of A_i and J matrices. Then the row sum of each Y'_ℓ is also 2^d since direct sums do not modify the row sum. We thus see that $Y_\ell = \frac{1}{2^d}Y'_\ell$ is row stochastic.

We now consider the coefficients of each term. We previously multiplied each initial vector u_j by c_j and we may consider taking the Kronecker sum of each u_j before normalising the resulting vector (normalising according to L^1 norm). We face an issue however, since some coefficients c_j may be negative and thus the resulting vector is not stochastic (it must be nonnegative). Fortunately we may modify a technique utilised by Bertoni [5] to solve this issue. Given a PFA for which $u^T X v = \lambda \in [0, 1]$, then by defining $v' = \mathbf{1} - v$ where $\mathbf{1}$ is the all-one vector of appropriate dimension (i.e. swapping between final and non final states), then $u^T X v = 1 - \lambda \in [0, 1]$.

Now, since each $X_{j,\ell}$ has a row sum of 2^d and u'_j is of unit length (L^1 norm), then Eqn. (7) can be adapted to the following:

$$\begin{aligned}(u'_j)^T X_{j,0}^{x_0}X_{j,1}^{x_1}\cdots X_{j,t}^{x_t}(\mathbf{1} - v'_j) &= 2^{d(x_0+x_1+\ldots+x_t)} - c_j R_j(x_0, x_1, \ldots, x_t) \\ &= 2^{d(x_0+x_1+\ldots+x_t)} - T_j(x_0, x_1, \ldots, x_t) \end{aligned} \tag{8}$$

Let us assume, without loss of generality, that we have arranged the terms of P^h such that those terms with a positive coefficient (positive terms) appear first, followed by those with a negative coefficient (negative terms). Since we have r terms in P^h, there exists some $1 \le r' < r$ such that we have r' postive and $r - r'$ negative terms. Let us define $u_j = |c_j| e_{i_1}$, which is similar to u'_j defined previously, but using the absolute value of the corresponding coefficient.

We define $v = \bigoplus_{j=1}^{r'} v_j \oplus \bigoplus_{j=r'+1}^{r} (\mathbf{1} - v_j) \in \{0,1\}^{r(t+3)^d}$ as the final vector, so that we take the Kronecker sum of all final vectors, but we swap final and non-final states for the negative terms.

We now define the initial vector u, which must be a probability distribution. Let $g = \sum_{j=1}^{r} |c_j|$ be the sum of absolute values of coefficients and define $u = \frac{1}{g} \bigoplus_{j=1}^{r} u_j \in [0,1]^{r(t+3)^d}$. Note that u is stochastic (a probability distribution).

We now see that:

$$\begin{aligned} & u^T Y_0 Y_1^{a_1} \cdots Y_t^{a_t} v && (9) \\ = \; & \frac{\sum_{j=1}^{r'} u_j \left(\bigotimes_{\ell=0}^{d} A_{s_j[\ell]}^{x_{s_j[\ell]}} \otimes J^{\overline{x_{s_j[\ell]}}} \right) v_j + \sum_{j=r'+1}^{r} u_j \left(\bigotimes_{\ell=0}^{d} A_{s_j[\ell]}^{x_{s_j[\ell]}} \otimes J^{\overline{x_{s_j[\ell]}}} \right) (\mathbf{1} - v_j)}{g 2^{d(1+a_1+\cdots+a_t)}} \end{aligned}$$

Here we used the definition of matrices Y_i and Eqn. (6) to rewrite the expressions for $X_{j,0} \cdots X_{j,t}$. Notice that the power of Y_0 is set at 1, since that constraint is required by the conversion from a standard Diophantine polynomial to a homogeneous one as explained previously. Now, using Eqn. (7) and Eqn. (8), we can rewrite Eqn. (9) as:

$$\frac{\sum_{j=1}^{r'} T_j(x_0, \ldots, x_t) + \sum_{j=r'+1}^{r} \left(2^{d(1+a_1+\ldots+a_t)} - |T_j(x_0, \ldots, x_t)| \right)}{g 2^{d(1+a_1+\cdots+a_t)}} \tag{10}$$

$$= \frac{(r - r' + 1)}{g} + \frac{\sum_{j=1}^{r'} T_j(x_0, \ldots, x_t) + \sum_{j=r'}^{r} T_j(x_0, \ldots, x_t)}{g 2^{d(1+a_1+\cdots+a_t)}} \tag{11}$$

$$= \frac{(r - r' + 1)}{g} + \frac{P^h(x_0, x_1, \ldots, x_t)}{g 2^{d(1+a_1+\cdots+a_t)}} \tag{12}$$

We therefore define $\mathcal{P} = (u, \{Y_a | a \in \Sigma_t\}, v)$ and $\Sigma_t = \{0, 1, \ldots, t\}$ as our PFA, with letter monotonic language $\mathcal{L} = 01^*2^* \cdots t^*$ and $\lambda = \frac{r-r'+1}{g} \in [0,1] \cap \mathbb{Q}$ as the cut-point. There exists some word $w = 01^{x_1}2^{x_2} \cdots t^{x_t} \in \mathcal{L}$ such that $f_{\mathcal{P}}(w) \leq \lambda$ if and only if $P^h(1, x_1, x_2, \ldots, x_t) = 0$. Therefore the strict emptiness problem for $\mathcal{P}$ is undecidable on letter monotonic languages. Since $\mathcal{P}$ is upper-triangular, then it is polynomially ambiguous. We note the surprising fact that all generator matrices are in fact *commutative* (each $X_{j,i}$ is commutative and direct sums do not affect commutativity), which leads to the undecidability of non-strict cut-points for polynomially ambiguous PFA defined over commutative matrices. In this case, the order of the input word in irrelevant, only the Parikh vector of alphabet letters is important. In fact we may redefine $u = uY_0$ and $\mathcal{L} = 1^*2^* \cdots t^*$ to remove Y_0 and all constraints on $\mathcal{L}$. ◀

4 Injectivity problems for polynomially ambiguous PFA

We now study the injectivity of acceptance probabilities of polynomially ambiguous PFA. The next result begins with a proof technique from [4], where the undecidability of the injectivity problem (called the freeness problem in [4], although we here rename it injectivity) was shown for exponentially ambiguous PFA over five states. We show that the injectivity problem remains undecidable even when the PFA is polynomially ambiguous and over four states by using our new encoding technique (avoiding the Turakainen procedure which increases the matrix dimensions by two and generates an exponentially ambiguous PFA).

4.1 Proof of Theorem 2

Proof. Let $\Sigma = \{x_1, x_2, \ldots, x_{n-2}\}$ and $\Delta = \{x_{n-1}, x_n\}$ be distinct alphabets and $h, g : \Sigma^* \to \Delta^*$ be an instance of the mixed modification PCP. The naming convention will become apparent below. We define two injective mappings $\alpha, \beta : (\Sigma \cup \Delta)^* \to \mathbb{Q}$ by:

$$\begin{aligned} \alpha(x_{i_1} x_{i_2} \cdots x_{i_m}) &= \Sigma_{j=1}^{m} i_j (n+1)^{j-1}, \\ \beta(x_{i_1} x_{i_2} \cdots x_{i_m}) &= \Sigma_{j=1}^{m} i_j (n+1)^{-j}, \end{aligned}$$

and $\alpha(\varepsilon) = \beta(\varepsilon) = 0$. Thus α represents $x_{i_1}x_{i_2}\cdots x_{i_m}$ as a reverse $(n+1)$-adic number and β represents $x_{i_1}x_{i_2}\cdots x_{i_m}$ as a fractional number $(0.x_{i_1}x_{i_2}\cdots x_{i_m})_{(n+1)}$ (e.g. if $n = 9$, then $x_1x_2x_3$ is represented as $\alpha(x_1x_2x_3) = 321_{10}$ and $\beta(x_1x_2x_3) = 0.123_{10}$, where subscript 10 denotes base 10). Note that $\forall w \in (\Sigma \cup \Delta)^*, \alpha(w) \in \mathbb{N}$ and $\beta(w) \in [0,1) \cap \mathbb{Q}$. It is not difficult to see that $\forall w_1, w_2 \in (\Sigma \cup \Delta)^*, (n+1)^{|w_1|}\alpha(w_2) + \alpha(w_1) = \alpha(w_1w_2)$ and $(n+1)^{-|w_1|}\beta(w_2) + \beta(w_1) = \beta(w_1w_2)$.

Define $\gamma'' : (\Sigma \cup \Delta)^* \times (\Sigma \cup \Delta)^* \to \mathbb{Q}^{3\times 3}$ by

$$\gamma''(u,v) = \begin{pmatrix} (n+1)^{|u|} & 0 & \alpha(u) \\ 0 & (n+1)^{-|v|} & \beta(v) \\ 0 & 0 & 1 \end{pmatrix}.$$

It is easy to verify that $\gamma''(u_1,v_1)\gamma''(u_2,v_2) = \gamma''(u_1u_2, v_1v_2)$, i.e., γ'' is a homomorphism.

Let $\mathcal{G}'' = \{\gamma''(x_i, g(x_i)), \gamma''(x_i, h(x_i)) | x_i \in \Sigma, 1 \le i \le n-2\}$, $\mathcal{S}'' = \langle \mathcal{G}'' \rangle$, $\rho'' = (1,1,0)^T$ and $\tau'' = (0,0,1)^T$. Assume that there exist $M_1 = G_{i_1}G_{i_2}\cdots G_{i_t} \in \langle \mathcal{G}'' \rangle$ and $M_2 = G_{j_1}G_{j_2}\cdots G_{j_{t'}} \in \langle \mathcal{G}'' \rangle$ such that $t \neq t'$ or else at least one $G_{i_p} \neq G_{j_p}$ where $1 \le p \le t$ and $\lambda = \rho''^T M_1 \tau'' = \rho''^T M_2 \tau''$. We see that:

$$\lambda = \rho''^T M_1 \tau'' = \alpha(x_{i_1}x_{i_2}\cdots x_{i_t}) + \beta(f_1(x_{i_1})f_2(x_{i_2})\cdots f_t(x_{i_t})),$$
$$\lambda = \rho''^T M_2 \tau'' = \alpha(x_{j_1}x_{j_2}\cdots x_{j_{t'}}) + \beta(f_1'(x_{j_1})f_2'(x_{j_2})\cdots f_{t'}'(x_{j_{t'}})),$$

where each $f_i, f_i' \in \{g,h\}$. Since $\alpha(w) \in \mathbb{N}$ and $\beta(w) \in (0,1) \cap \mathbb{Q}$, $\forall w \in (\Sigma \cup \Delta)^*$, injectivity of α and β implies that if $\rho''^T M_1 \tau'' = \rho''^T M_2 \tau''$, then $t = t'$ and $i_k = j_k$ for $1 \le k \le t$. Furthermore, if $\rho^T M_1 \tau = \rho^T M_2 \tau$, we have that $\beta(f_1(x_{i_1})f_2(x_{i_2})\cdots f_t(x_{i_t})) = \beta(f_1'(x_{i_1})f_2'(x_{i_2})\cdots f_t'(x_{i_t}))$ and since at least one $f_p \neq f_p'$ for $1 \le p \le t$ by our above assumption, then this corresponds to a correct solution to the MMPCP instance (h,g). On the other hand, if there does not exist a solution to (h,g), then $\beta(f_1(x_{i_1})f_2(x_{i_2})\cdots f_t(x_{i_t})) \neq \beta(f_1'(x_{i_1})f_2'(x_{i_2})\cdots f_t'(x_{i_t}))$, and injectivity of β implies that $\rho''^T M_1 \tau'' \neq \rho''^T M_2 \tau''$.

We now use our new technique to encode such matrices and vectors to a linearly ambiguous four state PFA. We first define a mapping $\gamma' : (\Sigma \cup \Delta)^* \times (\Sigma \cup \Delta)^* \to \mathbb{N}^{3\times 3}$ to make all matrices be nonnegative integral:

$$\gamma'(u,v) \quad = \quad (n+1)^{|v|}\gamma''(u,v) = \begin{pmatrix} (n+1)^{|u|+|v|} & 0 & (n+1)^{|v|}\alpha(u) \\ 0 & 1 & (n+1)^{|v|}\beta(v) \\ 0 & 0 & (n+1)^{|v|} \end{pmatrix} \in \mathbb{N}^{3\times 3}$$

We next define the following morphism $\gamma : (\Sigma \cup \Delta)^* \times (\Sigma \cup \Delta)^* \to \mathbb{Q}^{4\times 4}$ to make all such matrices be row stochastic:

$$\gamma(u,v) = (n+1)^{-k} \begin{pmatrix} (n+1)^{|u|+|v|} & 0 & (n+1)^{|v|}\alpha(u) & \delta_1 \\ 0 & 1 & (n+1)^{|v|}\beta(v) & \delta_2 \\ 0 & 0 & (n+1)^{|v|} & \delta_3 \\ 0 & 0 & 0 & \delta_4 \end{pmatrix},$$

where $\delta_j \in \mathbb{N}$ are chosen so that the row sum of each row of $\gamma(u,v)$ is $(n+1)^k$ for some k. Any sufficiently large k can be used so long as each row has the same sum $(n+1)^k$ and thus $\gamma(u,v)$ becomes row stochastic. We use the same k value for all matrices of $\mathcal{G}$ which we define as $\mathcal{G} = \{\gamma(x_i, g(x_i)), \gamma(x_i, h(x_i)) | x_i \in \Sigma, 1 \le i \le n-2\}$, so that $\mathcal{S} = \langle \mathcal{G} \rangle$, and finally $\rho = (1,1,0,0)^T$ and $\tau = (0,0,1,0)^T$ are the initial and final state vectors respectively.

Assume that there exist $M_1 = G_{i_1} \cdots G_{i_t} \in \langle \mathcal{G} \rangle$ and $M_2 = G_{j_1} \cdots G_{j_{t'}} \in \langle \mathcal{G} \rangle$ such that $t \neq t'$ or else at least one $G_{i_p} \neq G_{j_p}$ for $1 \leq p \leq t$ and $\lambda = \rho^T M_1 \tau = \rho^T M_2 \tau$. We see that:

$$\lambda = \rho^T M_1 \tau = (n+1)^{-kt} \left(\alpha(x_{i_1} x_{i_2} \cdots x_{i_t}) + \beta(f_1(x_{i_1}) f_2(x_{i_2}) \cdots f_t(x_{i_t})) \right),$$
$$\lambda = \rho^T M_2 \tau = (n+1)^{-kt'} \left(\alpha(x_{j_1} x_{j_2} \cdots x_{j_{t'}}) + \beta(f'_1(x_{j_1}) f'_2(x_{j_2}) \cdots f'_{t'}(x_{j_{t'}})) \right),$$

where each $f_i, f'_i \in \{g, h\}$. If $t = t'$, then the same argument as previously shows that $i_k = j_k$ for $1 \leq k \leq t$. If $t \neq t'$, assume without loss of generality that $t' < t$. In this case we see that:

$$(n+1)^{-kt''} \left(\alpha(x_{i_1} \cdots x_{i_t}) + \beta(f_1(x_{i_1}) \cdots f_t(x_{i_t})) \right) = \alpha(x_{j_1} \cdots x_{j_{t'}}) + \beta(f'_1(x_{j_1}) \cdots f'_{t'}(x_{j_{t'}})),$$

where $t'' = t - t'$. This is a contradiction however since the number of nonzero digits (where a digit is understood base $(n+1)$ here) in the left hand side of this expression is exactly $2t$, and the number of digits in the right expression is $2t' < 2t$. Note that the multiplication by $(n+1)^{-kt''}$ does not alter the number of nonzero digits, it is only a right shift of all digits, kt'' times. Thus, since the left and right sides have a different number of nonzero digits they cannot be equal and thus $t = t'$ as required. ◀

4.2 Proof of Theorem 3

Proof. We use a reduction from the equal subset sum problem, defined thus: given a set of positive integers $S = \{x_1, x_2, \ldots, x_k\} \subseteq \mathbb{N}$, do there exist two disjoint nonempty subsets $S_1, S_2 \subseteq S$ such that $\sum_{\ell \in S_1} \ell = \sum_{m \in S_2} m$? This problem is known to be NP-complete [32]. Note that although there is a requirement that the sets S_1 and S_2 be disjoint, this is not crucial so long as $S_1 \neq S_2$ (since if some element x_j is in both S_1, S_2, then the equality also holds when x_j is removed from both sets). We may therefore require that $S_1 \neq S_2$, with both nonempty such that the sum of elements of each set is identical. We define the set of matrices $M = \{A_i, B_i | 1 \leq i \leq k\} \subseteq \mathbb{Q}^{3 \times 3}$ in the following way:

$$A_i = \frac{1}{x_i + 1} \begin{pmatrix} 1 & x_i & 0 \\ 0 & 1 & x_i \\ 0 & 0 & x_i + 1 \end{pmatrix}, \quad B_i = \frac{1}{x_i + 1} \begin{pmatrix} 1 & 0 & x_i \\ 0 & 1 & x_i \\ 0 & 0 & x_i + 1 \end{pmatrix}$$

Note that A_i and B_i are thus row stochastic. Let $u = (1, 0, 0)^T$ be the initial probability distribution, $v = (0, 1, 0)^T$ be the final state vector and let $\mathcal{P} = (u, \{A_i, B_i\}, v)$ be our PFA. Define letter monotonic language $\mathcal{L} = (a_1|b_1)(a_2|b_2) \cdots (a_k|b_k) \subseteq a_1^* b_1^* a_2^* b_2^* \cdots a_k^* b_k^*$ and define a morphism $\varphi : \{a_i, b_i | 1 \leq i \leq k\}^* \to \{A_i, B_i | 1 \leq i \leq k\}^*$ in the natural way (e.g. the morphism induced by $\varphi(a_i) = A_i$ and $\varphi(b_i) = B_i$). Now, for a word $w = w_1 w_2 \cdots w_k \in \mathcal{L}$, note that $w_j \in \{a_j, b_j\}$ for $1 \leq j \leq k$. Define that $\mathfrak{v}(a_i) = x_i$ and $\mathfrak{v}(b_i) = 0$. In this case, we see that (due to the structure of A_i and B_i)

$$u^T \varphi(w_1 w_2 \cdots w_k) v = \frac{1}{\sum_{j=1}^{k} (x_j + 1)} \sum_{\ell=1}^{k} \mathfrak{v}(w_\ell)$$

Note of course that the factor $\frac{1}{\sum_{j=1}^{k}(x_j+1)}$ is the same for any $w \in \mathcal{L}$.

Assume then that there exists two words $\alpha, \beta \in \mathcal{L}$ with $\alpha \neq \beta$ such that $u^T \varphi(\alpha) v = u^T \varphi(\beta) v$ (i.e. assume that $\mathcal{P}$ is not free). Then $\sum_{\ell=1}^{k} \mathfrak{v}(\alpha_\ell) = \sum_{i \in S_1}^{k} x_i = \sum_{i \in S_2}^{k} x_i = \sum_{\ell=1}^{k} \mathfrak{v}(\beta_\ell)$, where $S_1 = \{x_i ; |\alpha|_{a_i} > 0\}$ and $S_2 = \{x_i ; |\beta|_{a_i} > 0\}$. This is true if and only if the instance S of the equal subset sum problem has a solution as required (note that only the empty set has a sum of zero which has unique representation $b_1 \cdots b_k$). Since A_i and B_i are upper-triangular, with initial state 1 and final state 2, then $\mathcal{P}$ is linearly ambiguous. ◀

5 Conclusion

There are a variety of open problems remaining. For example, does Theorem 1 still hold for quadratic ambiguity, when taken alongside the other constraints (letter monotonic language and commutative matrices). Another direction is to improve the complexity lower bound of Theorem 3 to show it is either PSPACE-hard, EXPSPACE-hard or undecidable, under the same constraints as in the theorem statement.

References

1 L. Babai, R. Beals, J-Y. Cai, G. Ivanyos, and E. M. Luks. Multiplicative equations over commuting matrices. In *Proc. of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA 96, 1996.

2 L. Babai and S. Moran. Arthur–Merlin games: a randomized proof system, and a hierarchy of complexity classes. *Journal of Computer and System Sciences*, 36:254–276, 1988.

3 P. C. Bell, S. Chen, and L. M. Jackson. Scalar ambiguity and freeness in matrix semigroups over bounded languages. In *Language and Automata Theory and Applications*, volume LNCS 9618, pages 493–505, 2016.

4 P. C. Bell, V. Halava, and M. Hirvensalo. Decision Problems for Probabilistic Finite Automata on Bounded Languages. *Fundamenta Informaticae*, 123(1):1–14, 2012.

5 A. Bertoni. The solution of problems relative to probabilistic automata in the frame of the formal language theory. *GI Jahrestagung*, pages 107–112, 1974.

6 A. Bertoni, G. Mauri, and M. Torelli. Some recursively unsolvable problems relating to isolated cutpoints in probabilistic automata. In *Automata, Languages and Programming*, volume 52, pages 87–94, 1977.

7 V. Blondel and V. Canterini. Undecidable problems for probabilistic automata of fixed dimension. *Theory of Computing Systems*, 36:231–245, 2003.

8 V. Blondel and J. Tsitsiklis. The boundedness of all products of a pair of matrices is undecidable. *Systems and Control Letters, Elsevier*, 41:2:135–140, 2000.

9 V. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36:1249–1274, 2000.

10 R. G. Bukharaev. Probabilistic automata. *Journal of Mathematical Sciences*, 13(3):359–386, 1980.

11 J. Buresh-Oppenheim, M. Clegg, R. Impagliazzo, and T. Pitassi. Homogenization and the polynomial calculus. *Computational complexity*, 11(3-4):91–108, 2002.

12 J. Cassaigne, J. Karhumäki, and T. Harju. On the Decidability of the Freeness of Matrix Semigroups. *International Journal of Algebra and Computation*, 9(3-4):295–305, 1999.

13 É. Charlier and J. Honkala. The freeness problem over matrix semigroups and bounded languages. *Information and Computation*, 237:243–256, 2014.

14 V. Chonev, J. Ouaknine, and J. Worrell. On the Complexity of the Orbit Problem. *Journal of the ACM*, 63(3):1–18, 2016.

15 A. Condon and R. J. Lipton. On the complexity of space bounded interactive proofs. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 462–467, 1989.

16 L. Daviaud, M. Jurdzinski, R. Lazic, F. Mazowiecki, G. A. Pérez, and J. Worrell. When is containment decidable for probabilistic automata? In *International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 121:1–121:14, 2018.

17 N. Fijalkow, C. Riveros, and J. Worrell. Probabilistic automata of bounded ambiguity. In *28th International Conference on Concurrency Theory (CONCUR)*, pages 19:1–19:14, 2017.

18 V. Halava, J. Kari, and Y. Matiyasevich. On post correspondence problem for letter monotonic languages. *Theoretical Computer Science*, 410:30–32, 2009.

19 M. Hirvensalo. Improved undecidability results on the emptiness problem of probabilistic and quantum cut-point languages. *SOFSEM 2007: Theory and Practice of Computer Science, Lecture Notes in Computer Science*, 4362:309–319, 2007.

20 R. A. Horn and C. R. Johnson. *Topics in matrix analysis.* Cambridge University Press, 1991.

21 O. Ibarra and B. Ravikumar. On sparseness, ambiguity and other decision problems for acceptors and transducers. In *Proc. STACS 1986*, volume 210, pages 171–179, 1986.

22 R. Kannan and R. J. Lipton. Polynomial-time algorithm for the orbit problem. *Journal of the ACM*, 33(4):808–821, 1986.

23 Yu. Matiyasevich. *Hilbert's Tenth Problem.* MIT Press, 1993.

24 M. Mohri, F. Pereira, and M. Riley. Weighted finite-state transducers in speech recognition. *Computer Speech & Language*, 16(1):69–88, 2002.

25 T. Neary. Undecidability in Binary Tag Systems and the Post Correspondence Problem for Five Pairs of Words. In *STACS15*, pages 649–661, 2015.

26 A. Paz. *Introduction to Probabilistic Automata.* Academic Press, 1971.

27 M. O. Rabin. Probabilistic automata. *Information and Control*, 6:230–245, 1963.

28 C. Reutenauer. *Propertiétés arithmétiques et topologiques de séries rationnelles en variables non commutatives.* Thèse troisième cycle, Université Paris VI, 1977.

29 A. Salomaa and M. Soittola. *Automata-Theoretic Aspects of Formal Power Series.* Springer-Verlag, 1978.

30 P. Turakainen. Generalized automata and stochastic languages. *Proceedings of the American Mathematical Society*, 21:303–309, 1969.

31 A. Weber and H. Seidl. On the degree of ambiguity of finite automata. *Theoretical Computer Science*, 88(2):325–349, 1991.

32 H. J. Woeginger and Z. Yu. On the equal-subset-sum problem. *Information Processing Letters*, 42(6):299–302, 1992.

33 A. Yakaryilmaz and A. C. Say. Unbounded-error quantum computation with small space bounds. *Information and Computation*, 209(6):873–892, 2011.