Geometry and algorithms for upper triangular tropical matrix identities

DOI:
10.1016/j.jalgebra.2019.03.024

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):
Johnson, M., & Tran, N. (2019). Geometry and algorithms for upper triangular tropical matrix identities. Journal of Algebra. Advance online publication. https://doi.org/10.1016/j.jalgebra.2019.03.024

Published in:
Journal of Algebra

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
GEOMETRY AND ALGORITHMS FOR UPPER TRIANGULAR TROPICAL MATRIX IDENTITIES

MARIANNE JOHNSON AND NGOC MAI TRAN

Abstract. We provide geometric methods and algorithms to verify, construct and enumerate pairs of words (of specified length over a fixed $m$-letter alphabet) that form identities in the semigroup $\mathcal{UT}_n(T)$ of $n \times n$ upper triangular tropical matrices. In the case $n = 2$ these identities are precisely those satisfied by the bicyclic monoid, whilst in the case $n = 3$ they form a subset of the identities which hold in the plactic monoid of rank 3. To each word we associate a signature sequence of lattice polytopes, and show that two words form an identity for $\mathcal{UT}_n(T)$ if and only if their signatures are equal. Our algorithms are thus based on polyhedral computations and achieve optimal time complexity in some cases. For $n = m = 2$ we prove a Structural Theorem, which allows us to quickly enumerate the pairs of words of fixed length which form identities for $\mathcal{UT}_2(T)$. This allows us to recover a short proof of Adjan’s theorem on minimal length identities for the bicyclic monoid, and to construct minimal length identities for $\mathcal{UT}_3(T)$, providing counterexamples to a conjecture of Izhakian in this case. We conclude with six conjectures at the intersection of semigroup theory, probability and combinatorics, obtained through analysing the outputs of our algorithms.

Key words: Tropical polynomials; bicyclic monoid; semigroup identities; Dyck paths; tropical matrices.

1. Introduction

Consider arithmetic over the tropical semifield $T = (\mathbb{R} \cup \{-\infty\}, \oplus, \odot)$, where $x \oplus y = \max(x, y)$ and $x \odot y = x + y$. For each positive integer $n$, the set of $n \times n$ upper triangular matrices

$$\mathcal{UT}_n(T) = \{ M \in T^{n \times n} : M_{ij} = -\infty \text{ if } i < j \}$$

is a semigroup under matrix multiplication. We write $w \sim_n v$ to denote that a pair of words $w, v$ over an alphabet $\Sigma$ is a semigroup identity for $\mathcal{UT}_n(T)$, meaning that each morphism $\varphi$ from the free semigroup on $\Sigma$ to $\mathcal{UT}_n(T)$ satisfies $\varphi(w) = \varphi(v)$. The study of semigroup identities for $\mathcal{UT}_n(T)$ has attracted much attention in recent literature [5, 8, 9, 10, 14, 16, 19], with interesting connections between the equational theory of these monoids and the bicyclic [6] and plactic [4, 7] monoids. Typical results in this direction are as follows. The equivalence relation determined by $\sim_n$ is a nonidentical relation [8] which is a refinement of the relation determined by identities for the bicyclic monoid [10]; in the case $n = 2$ this refinement is trivial [6]; in the

Date: March 23, 2019.

The authors would like to thank the Mittag-Leffler Institute and the organisers of the Spring 2018 program on Tropical Geometry, Amoebas and Polytopes for providing them with the opportunity to interact.
case $n = 3$, the relation $\sim_3$ is also a refinement of the corresponding relation for the plactic monoid of rank three \cite{4,7}. With such results one can deduce, for example, that words in the same $\sim_n$ equivalence class must contain the same number of each letter. However, little insight into the structure of $\sim_n$ equivalence classes has been gained through analyzing known identities of $\mathcal{UT}_n(\mathbb{T})$. In part, this is because existing methods for constructing these identities require delicate arguments, developed with the primary focus of proving the existence of nontrivial identities \cite{8,10,14,19}.

In this work, we give new geometric methods and algorithms to verify, construct and enumerate identities in $\mathcal{UT}_n(\mathbb{T})$ amongst words with a given number of occurrences of each letter from a fixed alphabet. In general, verifying whether two given words form an identity in a given semigroup is a major computational challenge. In the case of $\mathcal{UT}_n(\mathbb{T})$, the naïve algorithm would be to multiply out the matrices symbolically, and then test for functional equality of the resulting pairs of tropical polynomials corresponding to each entry. Both of these operations are costly when the words are long. Our approach is based upon a geometric interpretation of the characterisation of $\mathcal{UT}_n(\mathbb{T})$ identities given in \cite{6}, which involves more polynomial pairs but in fewer variables and monomials. A key observation is that the polynomials which appear in \cite{6} are tropical polynomials with trivial coefficients, and thus they are functionally equal if and only if their Newton polytopes are equal. This translates the problem to verifying equality of lattice polytopes, which can be efficiently computed and, more importantly, adds geometric insights that allow us to deduce structural information about these identities.

Our strongest structural result applies to $\mathcal{UT}_2(\mathbb{T})$ identities over a two letter alphabet. Let $W(\ell_a, \ell_b)$ be the set of words with $\ell_a$ occurrences of $a$ and $\ell_b$ occurrences of $b$. This set is naturally in bijection with Northeast staircase paths from $(0,0)$ to $(\ell_a, \ell_b)$. In particular, $W(\ell_a, \ell_b)$ is a distributive lattice with respect to the natural partial order $\preceq$, in which $w \preceq v$ if and only if the path of $w$ does not rise above the path of $v$. The following Structural Theorem shows that $\mathcal{UT}_2(\mathbb{T})$ identities (or equivalently, identities for the bicyclic monoid \cite{6}) respect the distributive lattice structure of $W(\ell_a, \ell_b)$. In particular, it implies that the $\sim_2$ equivalence class of $w \in W(\ell_a, \ell_b)$ is completely specified by the unique minimal and maximal words in this class. We give an explicit and efficient construction for these words in Theorem 4.9.

**Theorem 1.1** (Structural Theorem). Let $w,u,v \in W(\ell_a, \ell_b)$.

(i) If $w \preceq u \preceq v$ and $w \sim_2 v$, then $w \sim_2 u$.
(ii) If $w \sim_2 v$, then $w \sim_2 w \lor v \sim_2 w \land v$.

In general, for any finite alphabet, we provide algorithms to solve the problems listed below.

- **CheckPair**: given a pair of words $w,v$ over alphabet $\Sigma$ and a positive integer $n$, decide if $w \sim_n v$.
- **ListWord**: given a word $w$ over alphabet $\Sigma$ and a positive integer $n$, compute the equivalence class of $w$ in $\mathcal{UT}_n(\mathbb{T})$. 


• ListAll: given an alphabet $\Sigma$, a vector $(\ell_1, \ldots, \ell_{|\Sigma|})$ of non-negative integers, and a positive integer $n$, list all $UT_n(T)$ equivalence classes of words with $\ell_i$ occurrences of the $i$th letter in $\Sigma$.

The worst-case time complexity of our various algorithms is summarised in the following theorem. In each instance $n$ is constant (indicating which semigroup $UT_n(T)$ we consider), and the complexity of our algorithms is given in terms of:

• the size, $m$, of the alphabet of the given word(s),
• the length, $\ell$, and content $(\ell_1, \ldots, \ell_m)$ of the given word(s),
• the size, $N_d(w)$, of the $\sim_d$ equivalence class containing the given word $w$ for each $d < n$, and
• the total number $T_n(\ell_1, \ldots, \ell_m)$ of $\sim_n$ equivalence classes of words with the same content as the given word.

**Theorem 1.2.** The worst-case time complexity of our algorithms (assuming a random access model of computation with unit time arithmetic operations, including comparison of numbers) is as follows:

| Task    | $m = 2, n = 2$ | $m = 2, n > 2$ | $m > 2$ |
|---------|----------------|----------------|---------|
| CheckPair | $\ell$ | $2^{n-1} \ell^2$ | $m^{n-1} \left( \frac{\ell}{n} \right)^2$ |
| ListWord | $\ell$ | $\sum_{d=2}^{n-1} \begin{pmatrix} \ell \\ d \end{pmatrix}^2 N_d(w)$ | $m^{n-1} \left( \frac{\ell}{(\ell_1, \ldots, \ell_m)} (n-1) \right)^2$ |
| ListAll | $T_2(\ell_1, \ell_2)\ell^2$ | $2^{n-1} \left( \frac{\ell}{\ell_1} \right)^2 (n-1)^2$ | $m^{n-1} \left( \frac{\ell}{(\ell_1, \ldots, \ell_m)} (n-1) \right)^2$ |

Furthermore, the above time complexity for CheckPair and ListWord is best possible for the case $m = n = 2$.

To the best of our knowledge, the only other algorithmic approaches to studying $UT_n(T)$ arise in connection with problem CheckPair. In [6] it is established that for each fixed $n$, problem CheckPair has polynomial time complexity where the degree of the polynomial depends upon $n$; in the case where $m = n = 2$ it is also shown that this can be achieved in linear time assuming a random access model of computation with unit time arithmetic operations, including comparison of numbers. Moreover, the connection to the bicyclic monoid established in [6] allows one to identify algorithmic approaches to identity checking for the bicyclic monoid with those for $UT_2(T)$. In particular we mention that Shleifer [17] enumerated all identities of the bicyclic monoid of length at most 13, by means of a computer program, whilst Pastijn [15] gave an algorithm for identity verification in the bicyclic monoid via linear programming (cf. Section 5 for further comparison and discussion). In this paper we streamline the approach taken in [6], and use geometric insights to create efficient methods to enumerate $UT_n(T)$ identities.

Our algorithms are efficient enough to find the shortest identities that hold in $UT_3(T)$ through exhaustive search (noting that it suffices to consider words over a two letter alphabet). Without our results, constructing shortest identities is a prohibitive computation. For instance, to prove that there is no $UT_3(T)$ identity of length 21, after grouping words by pairs that have the same number of each letter, there are still $\sum_{i=1}^{10} \begin{pmatrix} 21 \\ i \end{pmatrix} > 1.34 \times 10^{11}$ pairs of words to consider. In comparison, our algorithm takes less than 30 minutes on a conventional laptop to complete this task.
Theorem 1.3. The shortest $UT_3(T)$ identities have length 22. Up to exchanging the roles of $a$ and $b$ and reversing the words, these are:

\[
\begin{align*}
abbaabba & \ab baabbbbbaba \sim_3 \ab baabbbab ba baaabbbbbaba \\
abbaabaab & \ab babaababab ba baaabababa \\
abababaab & \ab baabababaab ba baaabababa \\
abbaabaab & \ab abbaabababaab ba baaabababa \\
abababaab & \ab baabababaab ba baaabababa \\
abbaabaab & \ab abbaabababaab ba baaabababa \\
abbaabaab & \ab abbaabababaab ba baaabababa \\
abbaababa & \ab baaababaabbaab ba abaababaabba \\
abbaababa & \ab abbaabababaab ba baaabababa \\
abbaababa & \ab abbaabababaab ba baaabababa \\
abbaababa & \ab abbaabababaab ba baaabababa \\
abbaababa & \ab abbaabababaab ba baaabababa \\
abbaababa & \ab abbaabababaab ba baaabababa \\
\end{align*}
\]

Theorem 1.3 disproves a conjecture of Izhakian [8, Conjecture 6.1], which in the case $n = 3$ predicts that a certain identity of length 26 is shortest. Moreover, since $\sim_3$ is a refinement of the relation defined by identities of $P_3$, the plactic monoid of rank three [4, 7], Theorem 1.3 represents the shortest identities currently known to hold in $P_3$ (these will be shortest if the refinement turns out to be trivial). We note that Taylor [19] has disproved [8, Conjecture 6.1] in the case $n = 4$ by means of a carefully constructed example. The analogue of Theorem 1.3 for $n = 2$ corresponds to Adjan’s result on minimal identities for the bicyclic monoid [4]. For ease of comparison, we restate Adjan’s result and provide a new geometric proof in Section 4.

Finally, our algorithms provide a wealth of data and insights, which lead to interesting results and conjectures regarding the statistics of $UT_n(T)$ equivalence classes. We present six such conjectures in our paper. Conjectures 6.2 and 6.3 state that weaker versions of the Structural Theorem hold for $n \geq 3$. Conjectures 6.5, 6.6 and 6.8 concern probabilistic constructions for long $UT_n(T)$ identities on two-letter alphabets. Conjecture 6.9 gives a candidate for the largest $\sim_2$ class amongst words of fixed length $\ell$. We present further details, data and reasoning to support these conjectures in Section 6.

Our paper is organised as follows. We review essential definitions and results in Section 2, refining a result of [9] to show that $\sim_n$ can be characterised in terms of certain tropical polynomials. In Section 3 we show how to compute the support sets of these polynomials, and use this to provide appealing geometric interpretations of several properties of the equational theory of the bicyclic monoid. In Section 4 we further specialise to the two-letter case, where we collect our major results. This includes a geometric proof of Adjan’s theorem, proof of the Structural Theorem, and an explicit construction of the minimal and maximal elements of each $UT_2(T)$ equivalence class (Theorem 4.9). In Section 5 we give the algorithms and prove their complexity. We discuss the conjectures in conjunction with the results obtained through the algorithms in Section 6. Section 7 concludes
the paper. Documented code to reproduce all our experimental results is available at https://github.com/princengoc/tropicalsemigroup.

Notations. We write \( \mathbb{N}_{\geq 0} \) to denote the set of non-negative integers, and \( \mathbb{N} \) to denote the set of positive integers. If \( \Sigma \) is a finite alphabet, then \( \Sigma^* \) (resp. \( \Sigma^+ \)) will denote the free monoid (resp. free semigroup) on \( \Sigma \), that is, the set of finite words (resp. finite nonempty words) over \( \Sigma \) under the operation of concatenation. For a word \( w \in \Sigma^+ \), we write \( w_i \) for the \( i \)th letter of this word, \( |w|_a \) for the number of occurrences of the letter \( a \) in \( w \) for each \( a \in \Sigma \), and \( |w| = \sum_{a \in \Sigma} |w|_a \) for the length of the word. Fixing a total order on \( \Sigma \), the sequence \( (|w|_a : a \in \Sigma) \) is called the content of the word, denoted \( c(w) \). For each \( d \in \mathbb{N}_{\geq 0} \), let \( \Sigma^d = \{ w \in \Sigma^* : |w| = d \} \). For each \( n \in \mathbb{N} \) and each pair of words \( w, v \in \Sigma^+ \), write \( w \sim_n v \) if \( w = v \) is a semigroup identity for \( \mathcal{UT}_n(T) \). If \( w \) is the unique element in its \( \sim_n \)-class, we say that \( w \) is an isoterm for \( \mathcal{UT}_n(T) \). We write \( w \leftrightarrow u \) to denote that \( u \) is obtained from \( w \) by a single adjacent letter swap, that is \( w = w' xy w'' \) and \( u = w' yx w'' \) where \( w', w'' \in \Sigma^* \) and \( x, y \in \Sigma \) with \( x \neq y \). For \( X, Y \subseteq \mathbb{R}^N \), we write \( X \simeq Y \) if there exists an invertible affine linear transformation \( f \) such that \( f(X) = Y \). For a finite set of points \( X \), let \( \text{conv}(X) \) denote the convex hull. Write \( \mathcal{V}(P) \) for the set of vertices of a polytope \( P \). For presentation of algorithms, \( \leftarrow \) means variable assignment, \( \oplus \) means list addition. For a dictionary \( L \), we write \( \text{keys}(L) \) to denote the list of its keys. If \( u \) is not a key in a dictionary \( L \), then \( L[u] \) is defined to be the empty list.

2. Background

In this section we recall the connection between tropical polynomials and identities of the semigroup \( \mathcal{UT}_n(T) \). To each word \( w \in \Sigma^+ \) we then associate a signature, which is a sequence of lattice polytopes. The main result of this section is Theorem 2.4, which states that \( w \sim_n v \) if and only if the first \( \sum_{j=1}^{n-1} |\Sigma|^j \) polytopes in their signatures are equal. This reduces identity verification in \( \mathcal{UT}_n(T) \) to checking equality of a sequence of lattice polytope pairs, a key idea of our paper.

Let \( I \) be a finite subset of \( \mathbb{N}_{\geq 0}^m \). The tropical polynomial with support \( I \) and coefficients \( (c_a \in \mathbb{R} : a \in I) \) in formal variables \( x_1, \ldots, x_m \) is the formal algebraic expression \( f = \bigoplus_{a \in I} (c_a \odot x_1^{a_1} \odot \cdots \odot x_m^{a_m}) \). We say that \( f \) has trivial coefficients if \( c_a = 0 \) for all \( a \in I \). The Newton polytope of \( f \), denoted \( \text{Newt}(f) \), is the convex hull of \( I \). The polynomial expression \( f \) defines a piece-wise linear function from \( \mathbb{R}^m \) to \( T \), which by an abuse of notation we shall also denote by \( f \), with

\[
\begin{align*}
f(x) &= \max_{a \in I} \left( c_a + \sum_{i=1}^m a_i x_i \right) \quad \text{for all } x \in \mathbb{R}^m,
\end{align*}
\]

where we regard \( f(x) = -\infty \) for all \( x \in \mathbb{R}^m \) if \( I = \emptyset \). Different polynomial expressions may define the same piece-wise linear function [13 §1]. In general, it is difficult to check if two tropical polynomials define the same function [12]. For formal tropical polynomials with trivial coefficients, the following lemma reduces this task to a convex hull computation [13 §1].
Lemma 2.1. Two formal tropical polynomials $f, g$ each having trivial coefficients define the same function if and only if $\text{Newt}(f) = \text{Newt}(g)$.

For each formal tropical polynomial $f$ with trivial coefficients there is a canonical reduced expression $[f]$ formed by taking the sum of monomials corresponding to the vertices of $\text{Newt}(f)$ defining the same function; we shall write simply $[f] = [g]$ to denote that two formal tropical polynomials $f$ and $g$ with trivial coefficients define the same function.

Fix an alphabet $\Sigma$ and $\ell \in \mathbb{N}$ and consider a word $w \in \Sigma^\ell$. For each $s \in \Sigma$ let $x(s)$ be the upper triangular matrix whose $(i,j)$th entry for $i \geq j$ is the variable $x(s, i, j)$ and associate to $w$ the matrix $x(w) := x(w_1) \odot \cdots \odot x(w_\ell)$ obtained by multiplying out the corresponding matrices tropically. Note that each entry of $x(w)$ on or above the main diagonal is a tropical polynomial with trivial coefficients. It then follows immediately from the definitions that for $w, v \in \Sigma^+$,

$$w \sim_n v \iff [x(w)_{jk}] = [x(v)_{jk}] \text{ for each } 1 \leq j \leq k \leq n. \quad (1)$$

In this formulation, verifying semigroup identities in $\mathcal{UT}_n(T)$ is equivalent to verifying functional equality of $\binom{n+1}{2}$ pairs of tropical polynomials in $|\Sigma|\binom{n+1}{2}$ variables. This problem has two difficulties. First, obtaining these polynomials through symbolic matrix multiplication is costly. Second, to apply Lemma 2.1 one must compute the corresponding Newton polytopes, which live in high dimension due to the large number of variables.

We now derive an alternative characterisation, Theorem 2.4. It builds on [6, Theorem 5.2], which utilises the well-known connection between tropical matrix multiplication and optimal paths in a weighted directed graph on $n$ nodes [3,13]. The characterisation of Theorem 2.4 also translates $w \sim_n v$ into functional equality of a system of tropical polynomials with trivial coefficients. Though the system is larger than that in [1], there are fewer variables, and the monomials involved capture combinatorial properties of the two words, meaning that the Newton polytopes are of lower dimension and can be constructed directly from the two words. This leads to significant gain in computation and understanding.

Introduce variables $x(s, i)$ for each letter $s \in \Sigma$ and each $i \in \{1, \ldots, n-1\}$. Fix a word $w \in \Sigma^+$. Let $d \in \mathbb{N}$ and for each word $u \in \Sigma^d$, define the formal tropical polynomial $g_u^w$ associated with $w$ as follows

$$g_u^w := \bigoplus_{\pi \in I, s \in \Sigma} x(s, k)^{N_s^w(\pi_{k-1}, \pi_k)}, \quad (2)$$

in variables $x(s, k), s \in \Sigma, 1 \leq k \leq |u|$, where $I$ is the set

$$\{(\pi_0, \ldots, \pi_{|u|+1}) : 0 = \pi_0 < \cdots < \pi_{|u|+1} = |w| + 1, \ w_{\pi_k} = u_k, 1 \leq k \leq |u|\}, \quad (3)$$

and $N_s^w(\pi_{k-1}, \pi_k)$ denotes the number of occurrences of $s$ lying strictly between $w_{\pi_{k-1}}$ and $w_{\pi_k}$. Recall that $u$ is said to be a scattered subword of $w$ if $w$ has a factorisation of the form

$$w = i(1) \ u_1 i(2) \cdots i(d) \ u_d i(d+1)$$

where $i(j) \in \Sigma^*$. The set $I$ thus records the positions of the letters $u_1, \ldots, u_d$ within $w$ for each such factorisation, whilst the exponents $N_s^w(\pi_{k-1}, \pi_k)$ for
s ∈ Σ record the content of the factors i(k) for k = 1, . . . , d. Notice that, by
definition, g_u^w(x) = −∞ for all x ∈ R^{|Σ|d} if and only if I = ∅.

**Definition 2.2** (Signature of a word). Let w ∈ Σ^+. For each d ∈ N, the d-
signature of w is the lexicographically ordered sequence of Newton polytopes
(Newt(g_u^w) : u ∈ Σ^d). The signature of w in UT_n(T) is the sequence of its
d-signatures for 1 ≤ d ≤ n − 1.

**Example 2.3.** Let Σ = {a, b, c} and w = acbaacbcb. Then
\[ g_u^w = x(a, 1)^0 x(b, 1)^0 x(c, 1)^0 + x(a, 1)^1 x(b, 1)^1 x(c, 1)^1 + x(a, 1)^2 x(b, 1)^1 x(c, 1)^1 \]
\[ g_v^w = x(a, 1)^1 x(b, 1)^0 x(c, 1)^0 + x(a, 1)^3 x(b, 1)^1 x(c, 1)^1 + x(a, 1)^2 x(b, 1)^2 x(c, 1)^2 \]
\[ g_c^w = x(a, 1)^1 x(b, 1)^0 x(c, 1)^1 + x(a, 1)^3 x(b, 1)^1 x(c, 1)^2 + x(a, 1)^3 x(b, 1)^2 x(c, 1)^3. \]
The degree 1 signature of w is (Newt(g_u^w), Newt(g_v^w), Newt(g_c^w)), where
\[ \text{Newt}(g_u^w) = \text{conv} \{(0, 0, 0), (1, 1, 1), (2, 1, 1)\} \]
\[ \text{Newt}(g_v^w) = \text{conv} \{(1, 0, 0), (3, 1, 1), (3, 2, 2)\} \]
\[ \text{Newt}(g_c^w) = \text{conv} \{(1, 0, 1), (3, 1, 2), (3, 2, 3)\}. \]

These polytopes are illustrated in Figure 1.

**Theorem 2.4.** Let w, v ∈ Σ^+, n ≥ 2. Then
\[ w \sim_u v \Leftrightarrow |g_u^w| = |g_u^v| \text{ for all } u ∈ Σ^d, 1 ≤ d ≤ n − 1 \tag{4} \]
or equivalently, w ∼_n v if and only their signatures in UT_n(T) are equal.

The right-hand condition of (4) requires \( \sum_{d=1}^{n-1} |Σ|^d \) pairs of tropical
polynomial functions in (n − 1)|Σ| variables to be equal. Theorem 2.4 is a
simplification of [6] Theorem 5.2, which provides a similar condition requiring
equality of \( \sum_{d=1}^{n-1} \binom{n-1}{d} |Σ|^d \) pairs of tropical polynomials in n|Σ| variables.
Indeed, our proof proceeds by showing that the two theorems are equivalent.
For ease of reference, let us recall [6] Theorem 5.2.

**Theorem 2.5.** [6] Theorem 5.2 For 0 ≤ d ≤ n − 1, let \( P_d^n \) denote the set of all
increasing sequences of length d + 1 in \{1, . . . , n\}. Recalling the definition
of the set I from [3], given u ∈ Σ^d, ρ ∈ P_d^n, and w ∈ Σ^+, define
\[ f_{u, ρ}^w := \bigoplus_{\pi ∈ I} x(s, ρ_k)^{N_w(π_k−1:π_k)}, \text{ in variables } x(s, ρ_i), s ∈ Σ, 1 ≤ i ≤ d+1. \]
For $w, v \in \Sigma^+$, $w \sim_n v \Leftrightarrow [f^w_{u,\rho}] = [f^v_{u,\rho}]$ for all $u \in \Sigma^d, \rho \in \mathcal{P}_d^n, 0 \leq d \leq n - 1$.

**Proof of Theorem 2.4.** For each positive integer $d$ let $\tau_d = (1, \ldots, d + 1)$. If $w \sim_n v$, then it follows from Theorem 2.5 that for each $u \in \Sigma^d, 1 \leq d \leq n - 1$ we have in particular $[f^w_{u,\tau_d}] = [f^v_{u,\tau_d}]$. Evaluating at $x(s, d + 1) = 0$ for all $s \in \Sigma$ then yields $[g^w_{s,\tau_d}] = [g^v_{s,\tau_d}]$. For the converse, first notice that if $[g^w_{s,\tau_d}] = [g^v_{s,\tau_d}]$, then setting $x(a, 1) = 1$ and $x(s, 1) = 0$ for all other $s \in \Sigma$ yields $|w|_a - 1 = |v|_a - 1$. Thus for each $n \geq 2$, the right hand side of (4) implies that $w$ and $v$ have the same content. By [6, Lemma 2.5] it will therefore suffice to show that the right hand side of (4) implies that $\varphi(w) = \varphi(v)$ for all morphisms $\varphi : \Sigma^+ \rightarrow \mathcal{UT}_n(T)$ such that for all $s \in \Sigma$ and $1 \leq i \leq n - 1$, $\varphi(s)_{i,i} = y(s, i) \in \mathbb{R}$ and $\varphi(s)_{n,n} = 0$. In this case it is easy to see that for each $1 \leq j < i \leq n$ we have $\varphi(w)_{i,j} = -\infty = \varphi(v)_{i,j}$ and $\varphi(w)_{j,i} = \bigodot y(s, j)^{|w|_s} = \bigodot y(s, j)^{|v|_s} = \varphi(v)_{j,i}$. It remains to show that all entries above the diagonal agree; we shall proceed by induction on $n$. Write $[[g^w_{s,\tau_d}]]$ to denote evaluation of the function corresponding to the tropical polynomial $g^w_{s,\tau_d}$ at $x(s, i) = y(s, i)$ for all $s \in \Sigma$ and $1 \leq i \leq n - 1$.

For $n = 2$, it is then easy to see that

$$\varphi(w)_{1,2} = \bigodot_{a \in \Sigma} \varphi(a)_{1,2} \odot [[g^w_{a,\tau_2}]] = \bigodot_{a \in \Sigma} \varphi(a)_{1,2} \odot [[g^v_{a,\tau_2}]] = \varphi(v)_{1,2},$$

as required. Now suppose that (4) holds for all $2 \leq n \leq k$. If $[g^w_{s,\tau_d}] = [g^v_{s,\tau_d}]$ for all $u \in \Sigma^d, 1 \leq d \leq k - 1$, then by our inductive hypothesis we have $w \sim_k v$ which by Theorem 2.5 yields $[f^w_{u,\rho}] = [f^v_{u,\rho}]$ for all $u \in \Sigma^d, \rho \in \mathcal{P}_d^k, 0 \leq d \leq k - 1$. By applying simple changes of variables of the form $x(s, \rho_i) \mapsto x(s, \rho'_i)$, where $\rho$ and $\rho'$ are increasing sequences of the same length over possibly different sets, we may further deduce that $[f^w_{u,\rho}] = [f^v_{u,\rho}]$ for all $u \in \Sigma^d, \rho \in \mathcal{P}_d^{k+1}, 0 \leq d \leq k - 1$. Thus if $\varphi : \Sigma^+ \rightarrow \mathcal{UT}_{k+1}(T)$ is any morphism and $1 \leq j - i < k$, arguing as in [6, Lemma 5.1] immediately yields

$$\varphi(w)_{i,j} = \bigodot_{d=1}^{j-i} \bigodot_{u \in \Sigma^d} \bigodot_{\rho \in \mathcal{P}_{d+1}^{k+1}, \rho_{d+1} = i, \rho_d = j} \varphi(u_t)_{\rho_{d-1},\rho_0} \odot [[f^w_{u,\rho}]] = \bigodot_{d=1}^{j-i} \bigodot_{u \in \Sigma^d} \bigodot_{\rho \in \mathcal{P}_{d+1}^{k+1}, \rho_{d+1} = i, \rho_d = j} \varphi(u_t)_{\rho_{d-1},\rho_0} \odot [[f^v_{u,\rho}]] = \varphi(v)_{i,j}.$$

The remaining entries in position $(1, k + 1)$ can be similarly expressed and it is easy to see that these will differ only if $[f^w_{u,\tau_k}]$ and $[f^v_{u,\tau_k}]$ differ for some word $u$ of length $k$. Since we may assume that $\varphi(s)_{k+1,k+1} = 0$ for all $s \in \Sigma$, it follows that a difference occurs in position $(1, k + 1)$ only if the functions corresponding to $f^w_{u,\tau_k}$ and $f^v_{u,\tau_k}$ differ when evaluated at $x(s, k + 1) = 0$ for all $s \in \Sigma$. But it is straightforward to check that such an evaluation yields the functions corresponding to polynomials $g^w_{\tau_k}$ and $g^v_{\tau_k}$ respectively, which agree by assumption. This completes the proof of (4). The second
statement now follows by application of Lemma 2.1 since each polynomial \( g_i^w \) has trivial coefficients.

\[ \square \]

3. Computing support sets

3.1. Staircase paths, lattice polytopes and degree 1 supports. We begin by showing that the degree-one signature of a word \( w \) can be easily read off from an associated staircase path. From Theorem 2.4 this gives a simple way to construct and verify identities in \( UT_2(T) \). Throughout this section let \( \Sigma = \{a_1, \ldots, a_m\} \) and to simplify notation let us write \( g_i^w \) for the formal tropical polynomial obtained from \( g_{ai}^w \) by the variable substitution \( x(a_k, 1) \leftrightarrow x_k \). Let \( \leq \) denote the partial order on \( \mathbb{N}^m_0 \) given by \( p \leq q \) if \( p_i \leq q_i \) for all \( i = 1, \ldots, m \). Recall that for \( w \in \Sigma^+ \), we write \( c(w) = (|w|_a_1, \ldots, |w|_a_m) \in \mathbb{N}^m_0 \) to denote the content of \( w \). For \( c \in \mathbb{N}^m_0 \) let \( W(c) = \{ w : c(w) = c \} \) be the set of words with prescribed content \( c \).

**Definition 3.1.** The path \( \gamma^w \) of a word \( w \in W(\ell_1, \ldots, \ell_m) \) is the set of lattice points of the staircase walk from \((0, \ldots, 0) \in \mathbb{N}^m_0 \) to \((\ell_1, \ldots, \ell_m) \in \mathbb{N}^m_0 \), formed by reading the word \( w \) from left to right and moving by the \( i \)th unit vector \( e_i := (0, \ldots, 0, 1, 0 \ldots 0) \) when one reads the letter \( a_i \), for \( i = 1, \ldots, m \). For each \( i = 1, \ldots, m \), the \( a_i \)-height of \( \gamma^w \) is the set \( \gamma^{w,a_i} = \{ \gamma^w_j : 0 \leq j \leq \ell_i - 1 \} \), where \( \gamma^w_j := \max\{ p \in \gamma^w : p_i = j \} \). (5)

**Lemma 3.2.** For \( i = 1, \ldots, m \), the support of \( g_i^w \) is \( \gamma^{w,a_i} \subseteq \mathbb{N}^m_0 \).

**Proof.** Since \( \gamma^w \) is a finite increasing sequence of points in \( \mathbb{N}^m_0 \) with respect to the order \( \leq \), every nonempty subset of \( \gamma^w \) has a maximal element with respect to this order. Thus, the \( a_i \)-height of \( \gamma^w \) is well-defined. For \( w \in \Sigma^\ell \), let \( w(j) = w_1 \cdots w_j \) denote the length \( j \) prefix of \( w \). By definition,

\[ g_i^w = \bigoplus_{h:w_h=a_i, k=1}^m x_k^{c(w(h-1))} \]

where \( c(w(h-1)) \in \mathbb{N}^m_0 \) records the content of the length \( h - 1 \) prefix of \( w \). Suppose that \( a_i \) occurs in positions \( h_0, \ldots, h_{|w|_a_i-1} \) of the word \( w \). Then the support of \( g_i^w \) is \( \{ c(w(h_j - 1)) : 0 \leq j \leq |w|_a_i - 1 \} \). Each such point clearly lies on the path \( \gamma^w \) (since it records the content of a prefix of \( w \)) and by definition the \( h \)th coordinate of \( c(w(h_j - 1)) \) is equal to \( j \) (since this is the number of occurrences of \( a_i \) before the \( (j + 1) \)th occurrence of \( a_i \)). Moreover, since \( w = w(h_j - 1)a_iw' \) for some \( w' \in \Sigma^\ell \) it follows that each point \( p \in \gamma^w \) that strictly exceeds \( c(w(h_j - 1)) \) in some coordinate is the content of a prefix of length greater than or equal to \( h_j \) and hence \( p_i > j \). Thus \( c(w(h_j - 1)) = \max\{ p \in \gamma^w : p_i = j \} \), as required.

**Corollary 3.3.** Let \( w, v \in \Sigma^+ \). Then \( w \sim_2 v \) if and only if \( \text{conv}(\gamma^{w,a_i}) = \text{conv}(\gamma^{v,a_i}) \) for all \( i = 1, \ldots, m \).

**Example 3.4.** Figure 2 illustrates two minimal length identities of the bicyclic monoid \( \mathbb{B} \). The left plot shows the paths \( \gamma^w \) (black) and \( \gamma^v \) (red) in the identity \( w := abab abba \sim_2 abba ba abba =: v \).
It is easy to see that \( \text{conv}(\gamma_{w,a}) = \text{conv}(\gamma_{v,a}) \) and \( \text{conv}(\gamma_{w,b}) = \text{conv}(\gamma_{v,b}) \) (the polygons shown in blue and green). Similarly, the right plot shows the paths \(\gamma_w\) (black) and \(\gamma_v\) (red) of the words in the identity
\[
w = abba \text{ ab} baab \sim_2 \text{ ab} ba \text{ baab} =: v.
\]

**Example 3.5.** The degree-one signature of \( w = acbaacbcb \), along with its path is illustrated in Figure 1. Note that each point of \(\gamma_w\), save the final one, is a vertex of one of the polytopes in its signature. So the path, and hence the word, is uniquely determined from this signature, showing that \( w \) is an isoterm for \( UT_2(T) \).

### 3.2. Degree \( n \) support sets.

Fix an alphabet \( \Sigma = \{a_1, \ldots, a_m\} \) and a word \( w \in \Sigma^+ \). For each scattered subword \( u \) of \( w \) with \( |u| = d \), let \( \gamma_u \subseteq \mathbb{N}_{\geq 0}^d \) denote the nonempty support set of the polynomial \( g_w^u \). Proposition 3.7 shows how these support sets can be recursively computed from the staircase path of \( w \). We begin with an example to illustrate the idea.

**Example 3.6.** Let \( \Sigma = \{a, b\} \) and consider the support set \( \gamma_{ab} \) of the polynomial \( g_{ab}^w \) in variables \( x(a,1), x(b,1), x(a,2), x(b,2) \). To compute this, we must consider each instance of \( ab \) as a scattered subword of \( w \). Suppose that for some \( 0 \leq i \leq |w|_a - 1 \) and \( 0 \leq j \leq |w|_b - 1 \), the \((i+1)\)th \( a \) occurs before the \((j+1)\)th \( b \). Then before the \((i+1)\)th \( a \) there are \( \gamma_{i}^{w,a} \) many \( a \)'s and \( b \)'s. Between the \((i+1)\)th \( a \) and the \((j+1)\)th \( b \) there are \( \gamma_{j}^{w,b} - \gamma_{i}^{w,a} - e_1 \) many \( a \)'s and \( b \)'s. Thus if \((i,i',j,j') \in \gamma_{ab} \) then we have \((i,i') = \gamma_{i}^{w,a} \), and \((j',j) = \gamma_{j}^{w,b} - \gamma_{i}^{w,a} - e_1 \). It follows easily from this observation that
\[
\gamma_{ab} \simeq (\gamma_{w,a} \times \gamma_{w,b}) \cap C_a,
\]
where the inequalities specifying the polyhedron
\[
C_a = \{(i,i',j,j) \in \mathbb{R}_{\geq 0}^4 : i < j' \text{ and } i' \leq j\}
\]
ensure that the \((j+1)\)th \( b \) is occurring after the \((i+1)\)th \( a \).

**Proposition 3.7.** Let \( w \in \Sigma^+ \) and let \( u \) be a scattered subword of \( w \) of length \( d \). Then for each \( a_j \in \Sigma \)
\[
\gamma_{u,a_j} \simeq (\gamma_{u} \times \gamma_{a_j}) \cap C_u,
\]
where \( C_u = \{y \in \mathbb{R}^{m(d+1)} : \sum_{k=0}^{d-1} y_{mk+r} + |u|_{a_r} \leq y_{md+r}, 1 \leq r \leq m \} \).
Proof. Consider the invertible affine linear map $\Pi : \mathbb{R}^{m(d+1)} \to \mathbb{R}^{m(d+1)}$ defined by

$$\Pi(p) = \begin{pmatrix} I_{md} & \cdots & I_m \\ I_m & \cdots & I_m \\ 0 & \cdots & I_m \end{pmatrix} p + \sum_{r=1}^m |u|_{a_r} e_{md+r},$$

where $e_i$ denotes the $i$th unit vector of $\mathbb{R}^{m(d+1)}$, and $I_k$ denotes the $k \times k$ identity matrix. We claim that the image of $\gamma^{ua_j}$ under $\Pi$ is $(\gamma^u \times \gamma^{a_j}) \cap C_u$. Let $p \in \gamma^{ua_j}$. Then, by definition, there is a factorisation of $w$ of the form:

$$w = i(0) u_1 i(1) u_2 \cdots i(d-1) u_d i(d) a_j w',$$

in which $i(0), \ldots, i(d), w' \in \Sigma^*$ and

$$p_{km+r} = |i(k)|_{a_r} \text{ for } 0 \leq k \leq d, \ 1 \leq r \leq m.$$

Then

$$\Pi(p) = (p_1, \ldots, p_{md}) \sum_{k=0}^d p_{km+1} + |u|_{a_1}, \ldots, \sum_{k=0}^d p_{km+m} + |u|_{a_m}).$$

Since the first $md$ coordinates describe the first $d$ intermediary factors in our factorisation of $w$, it is immediate that restricting to these coordinates yields a point of $\gamma^u$. Noting that the final $m$ coordinates describe the content of the prefix $v := i(0) u_1 i(1) u_2 \cdots i(d-1) u_d i(d)$ of $w = va_j w'$, it follows that restricting to these coordinates gives a point of $\gamma^{a_j}$. Finally, $0 \leq p_{md+r}$ for each $r = 1, \ldots, m$, so $\Pi(p)$ satisfies the inequalities specified by $C_u$. This shows that $\Pi(\gamma^{ua_j}) \subseteq (\gamma^u \times \gamma^{a_j}) \cap C_u$. For the reverse inclusion, suppose that $y \in (\gamma^u \times \gamma^{a_j}) \cap C_u$. By definition of $\gamma^u$, there is a factorisation of $w$ of the form:

$$w = i(0) u_1 i(1) u_2 \cdots i(d-1) u_d w',$$

in which $i(0), \ldots, i(d-1), w' \in \Sigma^*$ and

$$y_{km+r} = |v|_{a_r} \text{ for } 0 \leq k \leq d-1, 1 \leq r \leq m.$$

Similarly, by the definition of $\gamma^{a_j}$, there is another factorisation of $w$ of the form $w = va_j w''$ in which $v, w'' \in \Sigma^*$ and

$$y_{md+r} = |v|_{a_r} \text{ for } 1 \leq r \leq m.$$

The inequalities of $C_u$ then allow us to deduce that $i(0) u_1 i(1) u_2 \cdots i(d-1) u_d$ must be a prefix of $v$, so that

$$w = i(0) u_1 i(1) u_2 \cdots i(d-1) u_d i(d) a_j w'',$$

for some $i(d) \in \Sigma^*$. Taking $p$ to be the point of $\gamma^{ua_j}$ corresponding to this factorisation then yields $\Pi(p) = y$, so $\Pi(\gamma^{ua_j}) \supseteq (\gamma^u \times \gamma^{a_j}) \cap C_u$. Thus we have equality between these two sets as claimed.

\[\square\]

Remark 3.8. It follows from the previous result that every support set can be built from the degree 1 supports by taking set products, intersecting with appropriate polyhedra and applying affine linear transformations. Notice however that one cannot build the degree $d$ signature of a word directly from the degree 1 signature polytopes in this way, since although taking convex hulls commutes with affine transformations and set products, it need not commute with taking intersections and so it may happen that

$$\text{conv}(\gamma^u \times \gamma^a \cap C_u) \neq \text{conv}(\gamma^u \times \gamma^a) \cap C_u.$$
Therefore, Proposition 3.7 gives a clear geometric interpretation as to why \( w \sim_n v \) need not imply \( w \sim_{n+1} v \).

3.3. Properties of \( UT_n(T) \) identities. Reversal of words yields an involution on \( \sim_n \) equivalence classes.

**Lemma 3.9.** Let \( w, v \in \Sigma^+ \), and let \( \text{rev}(w), \text{rev}(v) \) denote the words obtained from \( w \) and \( v \) by reading from right-to-left. Then \( w \sim_n v \) if and only if \( \text{rev}(w) \sim_n \text{rev}(v) \).

**Proof.** Let \( w \sim_n v \) and consider the involution \( R : UT_n(T) \to UT_n(T) \) obtained by reflecting along the anti-diagonal. It is straightforward to check that \( R \) is an anti-automorphism of \( UT_n(T) \), meaning that for all \( A, B \in UT_n(T) \) one has \( R(AB) = R(B)R(A) \). Let \( \varphi : \Sigma^+ \to UT_n(T) \) be a morphism. Consider the morphism \( \psi : \Sigma^+ \to UT_n(T) \) constructed from \( \varphi \) by defining \( \psi(a) = R(\varphi(a)) \) for all \( a \in \Sigma \). Since \( R \) is an involution, we have

\[
\varphi(w_1 \cdots w_t) = \varphi(w_t) \cdots \varphi(w_1) = R(\psi(w_t)) \cdots R(\psi(w_1)) = R(\psi(w_1) \cdots \psi(w_t)).
\]

Since \( \psi \) is a morphism and \( w \sim_n v \), the two expressions are equal. \( \square \)

**Remark 3.10.** In the case where \( n = 2 \), the previous lemma can also be read off from the degree 1 signature. For each \( i = 1, \ldots, m \), define the map \( T_i^w : \mathbb{R}^m \to \mathbb{R}^m \) via \( p \mapsto -p + c(w) - e_i \). It is straightforward to check that \( T(\gamma_{w,a}) = \gamma_{\text{rev}(w),a} \), and that \( T \) is affine linear and invertible. Thus \( \text{conv}(\gamma_{w,a}) = \text{conv}(\gamma_{\text{rev}(w),a}) \) if and only if \( \text{conv}(\gamma_{\text{rev}(w),a}) = \text{conv}(\gamma_{\text{rev}(v),a}) \), and the conclusion follows.

The following lemma shows how some properties of identities for \( UT_n(T) \) manifest themselves in terms of geometric properties of the degree 1 signature of the words.

**Lemma 3.11.** Let \( i \in \{1, \ldots, m\} \), \( n \geq 2 \), and \( w, v \in \Sigma^+ \).

(i) The minimal and maximal elements of \( \gamma_{w,a} \) are vertices of \( \text{conv}(\gamma_{w,a}) \).

In particular, \( w \sim_n v \) implies that \( c(w) = c(v) \).

(ii) If \( w = w' a_i^k w'' \) where \( w' \) does not contain \( a_i \) and \( w'' \) does not start with \( a_i \), then \( \text{conv}(\gamma_{w,a}) \) has vertices \( c(w') \) and \( c(w') + (k-1)e_i \). In particular, \( w \sim_n v \) implies \( v = v' a_i^k v'' \), where \( c(v') = c(w') \) and \( v'' \) does not start with \( a_i \).

(iii) If \( w = w' a_i^k w'' \) where \( w'' \) does not contain \( a_i \) and \( w' \) does not end with \( a_i \), then \( \text{conv}(\gamma_{w,a}) \) has vertices \( c(w') \) and \( c(w') + (k-1)e_i \). In particular, \( w \sim_n v \) implies \( v = v' a_i^k v'' \), where \( c(v'') = c(w'') \) and \( v' \) does not end with \( a_i \).

(iv) If \( w = a_i^{k_{i_1}} \cdots a_i^{k_{i_t}} w'' \) where the indices \( i_1, \ldots, i_t \) are distinct elements of \( \{1, \ldots, m\} \), \( k_1, \ldots, k_t \in \mathbb{N} \), and \( w'' \in \Sigma^* \setminus a_i \Sigma^* \), then each \( v \) in the \( \sim_n \)-class of \( w \) has the form \( v = a_i^{k_{i_1}} \cdots a_i^{k_{i_t}} v'' \) where \( v'' \in \Sigma^* \setminus a_i \Sigma^* \) with \( c(v') = c(w'') \) and the first letter of \( v'' \) is equal to the first letter of \( w'' \).

**Proof.** Noting that \( w \sim_n v \) implies \( w \sim_2 v \), it suffices to prove each implication in the case \( n = 2 \).
(i) Since \( \gamma^{w,a_i} \) is a finite increasing sequence of points in \( \mathbb{N}_{\geq 0}^m \) with respect to the natural partial order \( \preceq \), it is immediate that the minimal and maximal elements of this set must be vertices of the convex hull \( \text{conv}(\gamma^{w,a_i}) \). Suppose then that \( w \sim_2 v \). Corollary 3.3 then implies that the maximal element of \( \bigcup_{j=1}^m \gamma^{w,a_j} \) must be equal to the maximal element of \( \bigcup_{j=1}^m \gamma^{v,a_j} \), and that this common point lies in \( \gamma^{w,a_i} \cap \gamma^{v,a_i} \) for some \( i \). Since the maximal element of \( \gamma^{w,a_i} \) is \( c(w) - e_i \), the result follows.

(ii) Noting that \( \gamma^{w,a_i} \) is contained in the affine cone \( c(w) + \mathbb{R}_{\geq 0}^m \), we see that the segment of points \( c(wa_1^j) \) with \( 0 \leq j \leq k-1 \) are the unique elements of \( \gamma^{w,a_i} \) lying on the extremal ray \( c(w') + \mathbb{R}_{\geq 0}e_i \). Thus the two extremal points of this segment, namely \( c(w') \) and \( c(w' a^{k-1}) \), must be vertices of \( \text{conv}(\gamma^{w,a_i}) \). If \( w \sim_2 v \), then these two points must also be vertices of \( \text{conv}(\gamma^{v,a_i}) \), from which we deduce that \( v = v' a^k v'' \) where \( c(v') = c(w') \) and the first letter of \( v'' \) is not equal to \( a_i \).

(iii) If \( w'' \) does not contain \( a_i \), then the polytope \( \text{conv}(\gamma^{w,a_i}) \) has vertices \( c(\text{rev}(w)) \), \( c(\text{rev}(w'')) + (k-1)e_i \), and applying the transformation from Remark 3.10 then yields that the polytope \( \text{conv}(\gamma^{w,a_i}) \) has vertices:
\[
c(w'), c(w') + (k-1)e_i.
\]

(iv) By repeated application of part (ii) one may deduce that \( v = w'v'' \) where \( w' = a_{i_1}^{k_1} \cdots a_{i_t}^{k_t} \) and where the first letter of \( v'' \) is not equal to \( a_i \).

Now let \( a_i \) (respectively, \( a_j \)) denote the first letter of \( w'' \) (respectively, \( v'' \)). It remains to show that \( j = i \). Suppose first that \( j \not\in \{i_1, \ldots, i_{t-1}\} \). Since \( j \neq i \), reasoning as above yields that \( c(w') \) must be a vertex of \( \text{conv}(\gamma^{w,a_j}) \), and so in particular \( c(w') \in \gamma^{w,a_j} \). Since the sets \( \gamma^{w,a_1}, \ldots, \gamma^{w,a_m} \) are disjoint and \( c(w') \in \gamma^{w,a_i} \) we obtain \( j = i \). Likewise, if \( i \not\in \{i_1, \ldots, i_{t-1}\} \), we obtain \( i = j \).

Now suppose that \( i, j \in \{i_1, \ldots, i_{t-1}\} \) and suppose for contradiction that \( j \neq i \). Without loss of generality we may assume that \( i = i_r \) and \( j = i_s \) where \( 1 \leq r < s < t \). Since \( UT_n(\mathbb{T}) \) contains an identity element, it is clear that if \( w \sim_2 v \), then \( (w \setminus \Delta) \sim_n (v \setminus \Delta) \), where \( (w \setminus \Delta) \) and \( (v \setminus \Delta) \) are the words obtained from \( w \) and \( v \) by deleting all occurrences of the letters from any proper subset \( \Delta \) of \( \Sigma \). Taking \( \Delta = \Sigma \setminus \{i, j\} \) we find that \( c(w \setminus \Delta) = a^{k_r} a^{k_s} a^{k_t} (w'' \setminus \Delta) \) and \( c(v \setminus \Delta) = a^{k_r} a^{k_s} a^m (v'' \setminus \Delta) \). Applying the reasoning of the first paragraph to these two words then yields a contradiction.

Say that \( a^k \) is a block of \( w \in \Sigma^+ \) if \( w = w'a^k w'' \) where \( w' \) does not end with \( a_i \) and \( w'' \) does not begin with \( a_i \). Lemma 3.11 then tells us that any word with very few blocks of each letter must be an isoterm.

**Corollary 3.12.** Let \( w \in \Sigma^+ \) be a word in which each letter \( a_i \in \Sigma \) occurs in at most two distinct blocks. Then \( w \) is an isoterm for \( UT_2(\mathbb{T}) \).

**Proof.** By Lemma 3.11(ii), the first block of each letter fixes a pair of points through which every path \( \gamma^v \) with \( v \sim_2 w \) must pass. Likewise, by Lemma 3.11(iii), the last block of each letter fixes a pair of points through which every path \( \gamma^v \) with \( v \sim_2 w \) must pass. If each letter occurs in at most two distinct blocks, then it follows that each block of \( w \) is either the first block of that letter or the last block of that letter (possibly both), and it is straightforward to verify that the points described above uniquely determine the word.

Corollary 3.13. Let $|\Sigma| = 2$. If the total number of blocks of $w \in \Sigma^+$ (counted with multiplicity) is no more than five, then $w$ is an isoterm for $\text{UT}_2(\Sigma)$.

Proof. If $w$ is a word on two letters with fewer than five blocks, then it is clear that each block is either the first or last block of a letter, and the reasoning of the previous corollary applies. Suppose then that $w$ has five blocks, say $w = a^ib^ja^kb^la^l$, and that $w \sim_2 v$. Parts (ii) and (iii) of Lemma 3.11 yield that $v = a^ib^jv''$ and $v = v'b^ka^l$ for some $v', v'' \in \Sigma^*$. Since $v$ and $w$ are required to have the same content we see that $v = w$. □

More generally, one can use similar reasoning to show that any word over an $m$-letter alphabet containing no more than $m + 3$ blocks must be an isoterm. For the sake of brevity we omit the proof.

Say that an identity system $E \subseteq W(c) \times W(c)$ is left 1-hereditary if for each pair $(v, w) \in E$, whenever $v = v'zv''$ and $w = w'zw''$ where $v'$ and $w'$ do not contain any occurrences of $z \in \Sigma$, then $(v', w') \in E$. Dually, say that $E$ is right 1-hereditary if for each pair $(v, w) \in E$, we have that whenever $v = v'zv''$ and $w = w'zw''$ where $v''$ and $w''$ do not contain any occurrences of $z \in \Sigma$, then $(v'', w'') \in E$. The equational theory of the bicyclic monoid is known to be both left and right 1-hereditary [15, 16]. Since $\text{UT}_2(\Sigma)$ identities are exactly the identities of the bicyclic monoid [6, Theorem 4.1], we obtain a geometric proof of this statement by consideration of the degree 1 signatures of the words. The following easy lemma (depicted in Figure 3) captures the essence of this property in terms of the polytopes involved.

Lemma 3.14. Let $\epsilon \in \mathbb{R}^m_0$, with $c_j \neq 0$ for some $j$, and let $B_\epsilon = \{p \in \mathbb{R}^m_0 : p_j = 0, p_i \leq c_j$ for all $i \neq j\}$. If $P$ and $Q$ are finite sets with $P \subseteq B_\epsilon$ and $Q \subseteq c + \mathbb{R}^m_0$, then every vertex of $\text{conv}(P)$ is a vertex of $\text{conv}(P \cup Q)$.

Proof. It is easily verified that $\text{conv}(P \cup Q) \cap B_\epsilon = \text{conv}(P)$. Suppose $p$ is a vertex of $\text{conv}(P)$, but not of $\text{conv}(P \cup Q)$. Then $p$ lies on a line segment strictly between $q$ and $r$ for some $q, r \in \text{conv}(P \cup Q)$. Taking one of $q, r$ outside the box $B_\epsilon$ contradicts that $p$ lies in this box, whilst taking both $q, r$ in $B_\epsilon$, and hence in $\text{conv}(P)$, contradicts that $p$ is a vertex of $P$. □

Proposition 3.15. [15, Lemma 2] (cf. [15, 4.16]) The equational theory of the bicyclic monoid is both left and right 1-hereditary.

Proof. Suppose that $v \sim_2 w$, where $v = v'a_jv''$ and $w = w'a_jw''$ and where $v'$ and $w'$ do not contain any occurrences of $a_j \in \Sigma$. Then we have $\text{conv}(\gamma_{w',a_j}) = \text{conv}(\gamma_{v',a_j})$ for all $i$. The fact that the $j$-polytopes agree tells us that $c(w'a_j) = c(v'a_j)$; let $c$ denote this common content. For each $i$ such that $c_i \neq 0$, the points of $\gamma_{w',a_i}$ (respectively, $\gamma_{v',a_i}$) each lie in either the box $B_\epsilon$ or the cone $c + \mathbb{R}^m_0$. The points lying in the box are precisely the points of $\gamma_{w',a_i}$ (respectively, $\gamma_{v',a_i}$). Applying Lemma 3.14 first with $P = \gamma_{w',a_i}$ and the set difference $Q = \gamma_{w',a_i} \setminus \gamma_{w',a_i}$, and second with $P = \gamma_{v',a_i}$ and the set difference $Q = \gamma_{v',a_i} \setminus \gamma_{v',a_i}$, then yields that $\text{conv}(\gamma_{w',a_i}) = \text{conv}(\gamma_{v',a_i})$. Thus $v' \sim_2 w'$. □
4. Identities for the bicyclic monoid on two letter alphabets

This section collects our main results on $UT_2(T)$ identities over a two-letter alphabet. In particular, we prove Theorem 1.1, providing an algorithm to compute the minimal and maximal elements of each equivalence class guaranteed by this theorem (cf. Theorem 4.9), and give an alternative proof of the minimality of Adjan’s identity for the bicyclic monoid.

Let $\Sigma = \{a, b\}$. The degree-one signature of a word $w \in W(\ell_a, \ell_b)$ is the pair of polygons $\text{Newt}(g_w^a), \text{Newt}(g_w^b)$, which we denote as $A(w)$ and $B(w)$, respectively. To further simplify notation, we write $\alpha_w$ for $\gamma_w^a$ and $\beta_w$ for $\gamma_w^b$ defined in (5) and denote the unique element of $\alpha_w$ with first coordinate $0 \leq i \leq \ell_a - 1$ by $(i, \alpha_w^i)$, and the unique element of $\beta_w$ with second coordinate $0 \leq i \leq \ell_b - 1$ by $(\beta_w^i, i)$. The dual of $w$ is the word $\tilde{w} \in W(\ell_b, \ell_a)$ obtained by swapping each occurrence of the letter $a$ in $w$ by the letter $b$, and vice versa. Note that $(i, \alpha_w^i) = (i, \beta_{\tilde{w}}^i)$ for all $0 \leq i \leq \ell_a - 1$.

There is a one-to-one correspondence between a word and its path; for words on two letters, fixing the length and the $a$-height $\alpha_w$ also uniquely determines $w$.

Lemma 4.1. Let $\ell_a, \ell_b \in \mathbb{N}$ and $\alpha : \{0, 1 \ldots, \ell_a - 1\} \rightarrow \mathbb{N}_{\geq 0}$ a non-decreasing map $i \mapsto \alpha_i$. If $\ell_b \geq \alpha_{\ell_a - 1}$, then

$$w = b^{\alpha_0} a b^{(\alpha_1 - \alpha_0)} a b^{(\alpha_2 - \alpha_1)} a \ldots b^{(\alpha_{\ell_a - 1} - \alpha_{\ell_a - 2})} a b^{(\ell_b - \alpha_{\ell_a - 1})}$$

(6)

is the unique word in $W(\ell_a, \ell_b)$ such that $\alpha_w = \{(i, \alpha_i) : 0 \leq i \leq \ell_a - 1\}$.

Equation (6) elucidates the interpretation of $\alpha_i^w$ for each $i$. Applying the same argument to the dual word $\tilde{w}$, one sees that the word length $\ell$ together with the $b$-height $\beta_w$ also uniquely determines $w$. Lemma 1.1 and Corollary 3.3 therefore converts the study of $UT_2(T)$ identities over a two letter alphabet to the study of non-decreasing functions $\alpha : \{0, 1 \ldots, \ell_a - 1\} \rightarrow \{0, 1, \ldots, \ell_b\}$ for each fixed pair of non-negative integers $(\ell_a, \ell_b)$.

Fix $\ell_a, \ell_b \in \mathbb{N}$. For staircase walks $\gamma, \gamma'$ that both start from $(0, 0)$ and end at $(\ell_a, \ell_b)$, define $\gamma \leq \gamma'$ if the walk $\gamma'$ does not go above the walk $\gamma$, and $\gamma < \gamma'$ if additionally the two paths are not equal. Define $w \preceq w'$ if
\[ \gamma^w \preceq \gamma^{w'} \] This turns \( W(\ell_a, \ell_b) \) into a distributive lattice. The following lemma follows straightforwardly from the definitions.

**Lemma 4.2.** For \( w, v \in W(\ell_a, \ell_b) \), the following are equivalent:

- \( (k, \alpha^v_k) \preceq (k, \alpha^v_k) \) for all \( k \in \{0, \ldots, \ell_a - 1\} \).
- \( (\beta^v_k, k) \succeq (\beta^v_k, k) \) for all \( k \in \{0, \ldots, \ell_b - 1\} \).

In particular, the \( a \)-height and \( b \)-height of \( w \land v \) and \( w \lor v \) are given by

\[
\begin{align*}
\alpha^w_{i \land u} &= \min(\alpha^v_i, \alpha^u_i), & \beta^w_{i \land u} &= \max(\beta^v_i, \beta^u_i), \\
\alpha^w_{i \lor u} &= \max(\alpha^v_i, \alpha^u_i), & \beta^w_{i \lor u} &= \min(\beta^v_i, \beta^u_i).
\end{align*}
\]

**Proof of Theorem 1.1** Let \( w, u, v \in W(\ell_a, \ell_b) \) and suppose that \( w \sim_2 v \).

Thus we may write \( A := A(w) = A(v) \) and \( B := B(w) = B(v) \).

(i) Suppose \( w \preceq u \preceq v \). By definition, we have

\[
A = \text{conv}\{ (i, \alpha^w_i) : 0 \leq i \leq \ell_a - 1 \} = \text{conv}\{ (i, \alpha^u_i) : 0 \leq i \leq \ell_a - 1 \}, \quad \text{and} \quad A(u) = \text{conv}\{ (i, \alpha^v_i) : 0 \leq i \leq \ell_a - 1 \}.
\]

Since \( w \preceq u \preceq v \), we have \( \alpha^w_i \preceq \alpha^u_i \preceq \alpha^v_i \) for all \( i \). Thus each point \( (i, \alpha^u_i) \) lies on a (vertical) line segment contained within \( A \), giving \( A(u) \subseteq A \). If \( p \) is a vertex of \( A \), then clearly \( p, p + (1, 0) \in \gamma^w \cap \gamma^u \), from which it follows that \( p, p + (1, 0) \in \gamma^v \), giving \( A \subseteq A(u) \). Thus \( A(u) = A \). Likewise,

\[
B = \text{conv}\{ (\beta^w_i, i) : 0 \leq i \leq \ell_b - 1 \} = \text{conv}\{ (\beta^v_i, i) : 0 \leq i \leq \ell_b - 1 \}, \quad \text{and} \quad B(u) = \text{conv}\{ (\beta^u_i, i) : 0 \leq i \leq \ell_b - 1 \}.
\]

Since \( w \preceq u \preceq v \), we have \( \beta^w_i \preceq \beta^u_i \preceq \beta^v_i \) for all \( i \). Thus each point \( (\beta^u_i, i) \) lies on a (horizontal) line segment contained within \( B \), and hence \( B(u) \subseteq B \).

(ii) By definition \( \{ (i, \alpha^w_{i \land u}), (i, \alpha^v_{i \land v}) \} = \{ (i, \alpha^u_i), (i, \alpha^v_i) \} \), and so one immediately has that \( A(w \land v) \), \( A(w \lor v) \subseteq A \). Conversely, if \( (i, j) \) is a vertex of \( A \), then this point must belong to both \( \alpha^w \) and \( \alpha^v \), giving \( \alpha^w = \alpha^v = j \), and hence \( \alpha^w_{i \land u} = \alpha^v_{i \land v} = j \), from which it follows that \( A \subseteq A(w \land v), A(w \lor v) \). The same reasoning shows that \( B = B(w \land v) = B(w \lor v) \), whence the conclusion follows from Corollary 3.3. \( \square \)

**Corollary 4.3.** Each \( \sim_2 \)-equivalence class of \( \{a, b\}^+ \) is an interval of the lattice \( W(\ell_a, \ell_b) \) with respect to the order \( \preceq \) for some \( \ell_a, \ell_b \in \mathbb{N} \). If \( w \) is the unique minimal element in its equivalence class, then \( \tilde{w} \) is the unique maximal element in its class.

**Corollary 4.4.** Suppose \( w, v \in W(\ell_a, \ell_b) \). Then \( w \sim_2 v \) if and only if there exists a non-negative integer \( k \) and a sequence of words \( u(i) \), for \( i = 0, \ldots, k \) such that:

(i) \( w = u(0) \sim_2 u(1) \sim_2 \cdots \sim_2 u(k) = v \); and

(ii) \( u(i) \leftrightarrow u(i + 1) \) for all \( i \).

In particular, \( w \) is an iseterm if and only if \( w \neq w' \) for all \( w' \leftrightarrow w \).

For general alphabets, \( \sim_2 \)-equivalence classes need not be connected through adjacent swaps as the following example illustrates.
Example 4.5. Consider the words
\[ w = abc 
\] \[ v = abc 
\]
The first block of \( v \) is a difference in the first block of \( w \), and \( u \) is a common three points of these two sets as vertices. Let \( i \) either \( \ell \) be co-linear. Thus we must have \( \ell \) assume that \( \ell \) is a vertex of \( A \). By Corollary 4.4, there exists a word \( \ell \), with \( \ell \) connected to any other word of its equivalence class of \( \sim \) reached from \( \ell \) cannot have been obtained via a swap involving \( a \) and \( c \) only. We see that the equivalence class of \( \ell \) is not connected through adjacent swaps. (In fact, one finds that \( \ell \) is not \( \sim \) related to any of its neighbours and hence not connected to any other word of its \( \sim \)-class by adjacent swaps.)

Theorem 1.1 gives a short geometric proof of Adjan’s result that the identities in Example 3.4 are minimal length identities for the bicyclic monoid. We begin this proof with the following lemma.

Lemma 4.6. Let \( v \in W(\ell_a, \ell_b) \) and suppose that \( v \) is not an isoterm for \( UT_2(\mathbb{T}) \). Then \( \ell_a \geq 4 \), and if \( \ell_a = 4 \), then \( \ell_b > 6 \).

Proof. By Corollary 4.4 there exists a word \( u \) with \( v \leftrightarrow u \) and \( v \sim u \). In other words, \( \alpha^v \) and \( \alpha^u \) differ by one in exactly one coordinate, say, \( \alpha^v_i = \alpha^u_i + 1 \) for some \( 0 \leq i \leq \ell_a - 1 \), where neither \( (i, \alpha^v_i) \) nor \( (i, \alpha^v_i + 1) \) is a vertex of \( A(v) \). By Lemma 3.11, \( i \neq 0 \) and \( i \neq \ell_a - 1 \). Thus \( \ell_a \geq 3 \). If \( \ell_a = 3 \), this implies \( A(v) = \text{conv}\{(0, \alpha^v_0), (2, \alpha^v_2)\} = \text{conv}\{(0, \alpha^u_0), (2, \alpha^u_2), (1, \alpha^u_1), (1, \alpha^u_1 + 1)\} \), providing an immediate contradiction, since the four points above cannot be co-linear. Thus we must have \( \ell_a \geq 4 \). Suppose then that \( \ell_a = 4 \). Then either \( i = 1 \) or \( i = 2 \). By considering the reverse words if necessary, we can assume that \( i = 1 \) and hence \( A(v) = \text{conv}\{(0, \alpha^v_0), (1, \alpha^v_1), (1, \alpha^v_1 + 1), (2, \alpha^v_2), (3, \alpha^v_3)\} \)

Since these points cannot be co-linear, \( A(v) \) must be the triangle with the common three points of these two sets as vertices. Let \( e = \alpha^v_0 - \alpha^u_0 \) and \( f = \alpha^v_2 - \alpha^u_2 \). Note that \( e, f \in \mathbb{N}_{\geq 0} \) as \( \alpha^v \) is a staircase path. In fact \( e, \alpha^v_0 > 0 \), since \( \alpha^v_0 \geq \alpha^v_1 + 1 \geq \alpha^v_1 \geq \alpha^u_0 \), whilst taking \( \alpha^v_0 = 0 \) yields a difference in the first block of \( b \)'s of the two words \( v \) and \( u \) which, by Lemma 3.11 would contradict \( B(v) = B(u) \). We consider two cases.

(I) Suppose that the point \( (2, \alpha^v_2) \) lies above the line segment \( \text{conv}\{(0, \alpha^v_0), (3, \alpha^v_3)\} \).

It follows from (7) that \( e \) and \( f \) must satisfy \( |e + f| + 1 \leq |e| \),

which in turn implies \( e \geq 6 + 2f \). Then \( \ell_b \geq \alpha^v_0 + e + f \geq \alpha^v_0 + 6 > 6 \).
(II) Suppose that the point \((2, \alpha_2^v)\) lies below the line segment \(\text{conv}\{(0, \alpha_0^v), (3, \alpha_3^v)\} \).

It follows from \((7)\) that \(e\) and \(f\) must satisfy

\[
\left\lfloor \frac{e}{2} \right\rfloor + 1 \leq \left\lfloor \frac{e + f}{3} \right\rfloor.
\]

Since \(e \neq 0\), we see that \(\ell_b \geq \alpha_0^v + e + f \geq \alpha_0^v + 3\left\lfloor \frac{e}{3} \right\rfloor + 3 \geq \alpha_0^v + 6 > 6.\)

\[\square\]

**Theorem 4.7.** Let \(w \in W(\ell_a, \ell_b)\). If \(\ell_a + \ell_b \leq 9\), then \(w\) is an isoterm for the bicyclic monoid. Moreover, the only identities for the bicyclic monoid involving words of length 10 are given by

\[
\begin{align*}
\text{abba } &\text{abba } \sim_2 \text{abba ba abba, baab ab baab } \sim_2 \text{baab ba baab}, \\
\text{abba } &\text{baab } \sim_2 \text{abba ba baab, baab ab abba } \sim_2 \text{baab ab abba}.
\end{align*}
\]

**Proof.** Suppose \(w, v \in W(\ell_a, \ell_b)\) with \(w \sim_2 v\). By Lemma 4.6, \(\min(\ell_a, \ell_b) \geq 4\) and \(\min(\ell_a, \ell_b) = 4\) implies \(\max(\ell_a, \ell_b) > 6\). Thus \(\ell_a + \ell_b \geq 10\). Suppose \(\ell_a + \ell_b = 10\). By Lemma 4.6, we must have \(\ell_a = \ell_b = 5\). Enumerating all possible configurations of such words (using Algorithm 8 from Section 5, for example) gives us the desired identities. Figure 2 in Example 3.4 visualises two of these identities. \[\square\]

We now give an explicit construction of the minimal and maximal elements (with respect to \(\leq\)) of a given \(\sim_2\) equivalence class based on the signature of the class. As we shall discuss in Section 5, this construction gives a significant speedup to various algorithms for two-letter identities in \(\mathcal{UT}_{\mathcal{N}}(\mathbb{T})\). Due to Corollary 4.3, it is sufficient to construct the maximum element.

First we introduce some notations. For each point \(p \in \mathbb{N}_0^2\) let us write \(x(p)\) and \(y(p)\) to denote the first and second coordinates of \(p\), respectively. For \(p \preceq q\) we write \([p, q]\) to denote the set of all points \(r \in \mathbb{N}_0^2\) such that \(p \preceq r \preceq q\). Similarly, we write \((p, q)\), \([p, q]\) and \((p, q)\) to denote the corresponding intervals of \(\mathbb{N}_0^2\) with respect to the partial order \(\preceq\) which exclude one or both of the end points.

Let \((A, B)\) be the signature of some word in \(W(\ell_a, \ell_b)\), and let \(\mathcal{V}\) be the monotonically increasing sequence

\[
(0, 0) = p_1 \prec p_2 \prec \cdots \prec p_M \prec p_{M+1} = (\ell_a, \ell_b)
\]

consisting of the vertices of \(A\) and the vertices of \(B\), together with the point \((\ell_a, \ell_b)\). Define the label function \(\text{label} : \mathcal{V} \to \{a, b, \emptyset\}\), where \(\text{label}(p_i) = a\) if \(p_i\) is a vertex of \(A\); \(\text{label}(p_i) = b\) if \(p_i\) is a vertex of \(B\), and \(\text{label}(p_i) = \emptyset\) if \(p_i = (\ell_a, \ell_b)\). If \(\gamma\) is a staircase path with signature \((A, B)\), then \(\gamma\) must go through all points in \(\mathcal{V}\). In particular,

\[
\gamma = \bigcup_{i=1}^{M} \gamma(i),
\]

where \(\gamma(i)\) is a staircase path from \(p_i\) to \(p_{i+1}\). We shall construct the maximal path \(\gamma\) by constructing \(\gamma(i)\) for each \(i\). This is done through iterative calls to the procedure \text{MaxSegment} (see Algorithm 1), each of which constructs a segment of \(\gamma(i)\). Figures 4 and 5 illustrate the idea. For proof convenience,
the function MaxPath returns both the path $\gamma(i)$ and an extension of the function label to the points of $\gamma(i)$.

**Algorithm 1 Maxpath**

**Input:** points $p_i, p_{i+1} \in V$, label($p_i$), label($p_{i+1}$)  
**Output:** $\gamma(i) \subset \mathbb{N}_2^2$ and a function $\text{label}_i : \gamma(i) \rightarrow \{a, b, \emptyset\}$

1. Set $p \leftarrow p_i$, label$_i(p) \leftarrow \text{label}(p_i)$, $\gamma(i) \leftarrow \{p\}$
2. while $p \neq p_{i+1}$ do  
3. $\gamma(i) \leftarrow \gamma(i) \cup (p, p')$
4. for $q \in (p, p')$ do label$_i(q) \leftarrow \text{label}_i(p)$
5. $p \leftarrow p'$, label$_i(p) \leftarrow \text{label}_i(p')$
6. return ($\gamma(i), \text{label}_i$)

**procedure MaxSegment($p, p_{i+1}, \text{label}(p), \text{label}(p_{i+1})$)**

8. if $y(p) = y(p_{i+1})$ or $x(p) = x(p_{i+1})$ then

9. $p' \leftarrow p_{i+1}$, label$_i(p') \leftarrow \text{label}(p_{i+1})$

11. if label($p$) = $b$ then

12. $Y \leftarrow \{k : y(p) < k \leq y(p_{i+1}) \land (x(p), k) \in A, (x(p), k-1) \in B\}$

13. $y \leftarrow \text{max} Y$

14. $p' \leftarrow (x(p), y)$, label$_i(p') \leftarrow a$

15. if label($p$) = $a$ then

16. $X \leftarrow \{k : x(p) < k \leq x(p_{i+1}) \land (k, y(p)) \in B, (k-1, y(p)) \in A, (k, y(p) + r) \in A \text{ for some } r > 0\}$

17. $x \leftarrow \text{min} X$

18. $p' \leftarrow (x, y(p))$, label$_i(p') \leftarrow b$

19. return ($p', \text{label}(p')$)

**Definition 4.8.** For $i = 1, \ldots, M$, let ($\gamma(i), \text{label}_i$) denote the output of the Maxpath algorithm with input $(p_i, p_{i+1})$. Let $\gamma := \bigcup_{i=1}^M \gamma(i)$, and define label : $\gamma \rightarrow \{a, b, \emptyset\}$ by label($p$) = label$_i(p)$ if $p \in \gamma(i)$.

**Theorem 4.9.** Let $w \in W(\ell_a, \ell_b)$ be a word with signature $(A, B)$. The path of the unique maximal word in the $\sim_2$-class of $w$ is given by $\gamma$.

**Proof.** Let $\gamma$ be any path with signature $(A, B)$, and let $\gamma(i)$ be its subpath from $p_i$ to $p_{i+1}$. Let $\alpha$ and $\beta$ denote the respective $a$- and $b$-heights of $\gamma$. For each $i = 1, \ldots, M - 1$, we shall show by induction on the number of iterations of MaxSegment that after each iteration with input $p, p_{i+1}$:

(i) Point $p'$ is well-defined and $[p, p']$ is either a North or East segment.

(ii) For each point $q \in (p, p') \setminus V$

\[
\begin{align*}
\text{label}(q) = a & \Rightarrow \alpha(x(q)) \leq y(q) \text{ and } q \in A, \quad \text{and} \quad (8) \\
\text{label}(q) = b & \Rightarrow \beta(y(q)) \geq x(q) \text{ and } q \in B. \quad (9)
\end{align*}
\]

Once this statement has been established, the remaining steps of the proof proceed as follows. Condition (i) implies that Maxpath terminates in finitely many steps; thus $\gamma(i)$ is well-defined. By construction, $\gamma(i)$ is a union of North and East segments, and so a staircase path from $p_i$ to $p_{i+1}$. Thus,
\( \gamma = \bigcup_{i=1}^{m} \gamma(i) \) is a staircase path from \((0,0)\) to \((\ell_a, \ell_b)\). Note that \texttt{Maxpath} does not change the labels of points in \( \mathcal{V} \), and thus the \texttt{label} function returned by \texttt{Maxpath} is well-defined and extends the definition on the vertices \( \mathcal{V} \) to the entire of \( \gamma \). Let 
\[
\overline{\alpha} = \{ q \in \gamma : \text{label}(q) = a \}, \quad \overline{\beta} = \{ q \in \gamma : \text{label}(q) = b \}.
\]
Since \( \gamma \) is a staircase path, by construction of \texttt{label}, \( \overline{\alpha} \) is the \( a \)-height of \( \gamma \), and \( \overline{\beta} \) is the \( b \)-height of \( \gamma \). Therefore, condition (ii) implies that \( \gamma \preceq \gamma \).

Furthermore, (8) implies \( \alpha \subseteq A \), so \( A(\gamma) \subseteq A \). By definition of \texttt{label} on points in \( \mathcal{V} \), all vertices of \( A \) are in \( \overline{\alpha} \), so \( A(\gamma) \supseteq A \). Thus \( A = A(\gamma) \). A similar argument applied to \( B \) and \( \overline{\beta} \) implies \( B = B(\gamma) \). So \( \gamma \) has signature \((A, B)\), and it dominates any path \( \gamma \) with signature \((A, B)\). Thus it is the desired maximal path.

Now we prove (i) and (ii) by induction on the number of iterations of \texttt{MaxSegment}. Consider the first iteration, where \((p_i, p_{i+1}, \text{label}(p_i), \text{label}(p_{i+1}))\) is the input to \texttt{MaxSegment}. Let \( p = p_i \). There are three cases:

(I) If \( p \) shares a coordinate with \( p_{i+1} \): Then \texttt{MaxSegment} has well-defined output \( p_{i+1} \), and (i) clearly holds. If \texttt{label}(\( p \)) = \( a \), then \( p \) is a vertex of \( A \), and so \([p, p_{i+1}]\) must be an East segment. Thus \( y(p) = y(p_{i+1}) \) and it follows from the fact that \( \gamma \) is a staircase path that for any \( q \in (p, p_{i+1}) \), we have \( \alpha(i, x(q)) = y(q) = y(p) \), giving (8). If \texttt{label}(\( p \)) = \( b \), then by the same reasoning, \([p, p_{i+1}]\) is a North segment, and \( \beta(i, y(q)) = x(q) = x(p) \) for all \( q \in (p, p_{i+1}) \), giving (7).

(II) If \( p \) does not share a coordinate with \( p_{i+1} \) and \texttt{label}(\( p \)) = \( b \): The set \( Y \) defined in line 12 is nonempty, since \( \alpha(x(p)) \in Y \). Thus the output \( p' \) is well-defined, and \([p, p']\) is a North segment. This proves (i). For (ii), note that \( p' \in A \) is the unique point in this segment with label \( a \) and since we chose \( p' \) to have maximal \( y \) coordinate amongst points in \( Y \), we have

Figure 4. Algorithm \texttt{Maxpath} constructs a staircase path from \((0,0)\) to \((\ell_a, \ell_b)\), as a union of staircase paths between the consecutive vertex points of polytopes \( A \) and \( B \).
Again, we have three cases.

(III) If \( p \) does not share a coordinate with \( p_{i+1} \) and \( \text{label}(p) = a \): The set \( X \) defined in line 16 is nonempty, since \( \beta_{y(p)} \in X \). Thus the output \( p' \) is well-defined, and \([p, p']\) is an East segment. This proves (i). For (ii), note that \( p' \in B \) is the unique point in this segment with label \( b \) and since we chose \( p' \) to have minimal \( x \) coordinate amongst points in \( X \), we have \( \beta_{y(p')} = \beta_{y(p)} \geq x(p') \), giving (9). By construction, \( p \not\equiv (0, 1) \in A \), so \( q \in A \) for each \( q \in (p, p') \) by convexity. If (8) did not hold for some \( q \in (p, p') \), then we would have \( \alpha_{x(q)} > y(q) = y(p) \). But this contradicts that \( \gamma \) is a Northeast staircase path passing through the points \((x(q), \alpha_{x(q)})\) and \((\beta_{y(p)}, y(p))\), since by assumption the former lies directly North of \( q \), whilst the latter lies directly to the East of \( q \), since \( x(q) < x(p') \leq \beta_{y(p)} \).

Now suppose the induction hypothesis holds for the first \( j \) iterations of \text{MaxSegment}. Consider the \((j + 1)\)th iteration. Let \((p, \text{label}(p))\) be the output of \text{MaxSegment} in the \( j \)th iteration. If \( p = p_{i+1} \), then we are done. Otherwise \( p \) is not a vertex. Run \text{MaxSegment} with input \((p, p_{i+1}, \text{label}(p), \text{label}(p_{i+1}))\).

Again, we have three cases.

(A) At \( p \) with \( \text{label}(p) = b \), move North as much as possible to point \( p' \) such that \( p' \in A \) and \( p' < (0, 1) \in B \). Label \( p' \) as \( a \), and all other points between \( p \) and \( p' \) as \( b \).

(B) At \( p \) with \( \text{label}(p) = a \), move East as little possible to the first point \( p' \) such that \( p' \in B \), \( p' < (1, 0) \in A \), and \( p' + (0, r) \in A \) for some \( r > 0 \). Label \( p' \) as \( b \) and all other points between \( p \) and \( p' \) as \( a \).

(C) At the boundary follow the unique line segment to point \( p_{i+1} \). Give all points between \( p \) and \( p_{i+1} \) the same label as that of \( p \).

(D) In the proof of Theorem 4.9 the function \text{label}, constructed together with points on the original path \( \gamma \) (in red) are used to show that each call to \text{MaxSegment} has well-defined output.

**Figure 5.** Procedure \text{MaxSegment} in pictures.
(I) If \( p \) shares a coordinate with \( p_{i+1} \): Then \textbf{MaxSegment} outputs \( p_{i+1} \), and (i) holds. Let \( q \) be a point in \((p, p_{i+1})\).

If \( \text{label}(p) = a \), then \([p, p_{i+1}]\) is an East segment and \( \text{label}(q) = a \). Thus \( y(q) = y(p_{i+1}) \geq \alpha_{x(q)} \). Since \( p_{i+1} \) is a vertex point lying on the path \( \gamma \), one easily verifies that

\[
R := \{p_{i+1}, p_{i+1} - (1, 0), p_{i+1} + (0, 1)\} \cap A \neq \emptyset.
\]

Noting that \( q \) lies in the convex hull of \( p, r \) and \((x(q), \alpha_{x(q)})\) for each \( r \in R \), we obtain \( q \in A \). This proves (8) for case (I).

If \( \text{label}(p) = b \), then \([p, p_{i+1}]\) is a North segment, \( \text{label}(q) = b \), and \( x(q) = x(p_{i+1}) \). As \( \gamma(i) \) is a staircase path from \( p \) to \( p_{i+1} \), \( \beta_{y(p)} \leq \beta_{y(q)} \leq x(p_{i+1}) \). Since the induction hypothesis holds at the non-vertex point \( p \), we have \( \beta_{y(p)} \geq x(p) = x(p_{i+1}) \), giving \( x(q) = x(p_{i+1}) = \beta_{y(q)} \). Since \( p_{i+1} \) is a vertex point lying on the path \( \gamma \), one easily verifies

\[
R := \{p_{i+1}, p_{i+1} + (1, 0), p_{i+1} - (0, 1)\} \cap B \neq \emptyset.
\]

Noting that \( q \) lies in the convex hull of \( p, r \) and \((\beta_{y(q)}, y(q))\) for each \( r \in R \), we obtain \( q \in B \). This proves (9) for case (I).

(II) If \( p \) does not share a coordinate with \( p_{i+1} \) and \( \text{label}(p) = b \): Then \( p \) is the last point of an East move by \textbf{MaxSegment}. Thus \( p \) arises by taking the minimum of a set of the form \( X \) from line 16 (for the previous value of \( p \)). This implies that \( y(p) + 1 \) lies in the set \( Y \) from line 12 (for the current value of \( p \)). Thus \( q' \) is well defined and (i) holds. Using the inductive hypothesis we have \( \beta_{y(p)} \geq x(p) \) and hence \( \alpha_{x(p)} \leq y(p) \). The reasoning of the base step, case (II) then applies to show that (ii) holds.

(III) If \( p \) does not share a coordinate with \( p_{i+1} \) and \( \text{label}(p) = a \): Then \( p \) is the last point of a North move by \textbf{MaxSegment}. We now argue that \( \beta_{y(p)} \in X \). By the inductive hypothesis, \( \alpha_{x(p)} \leq y(p) \). Therefore, \( \beta_{y(p)} \geq x(p) \). This last inequality must be strict, or else \( p \) belongs to \( \beta \), which would contradict the maximality of the \( y \) coordinate of \( p \) obtained in the previous step. Let \( q' = (\beta_{y(p)}, y(p)) \) be the point on \( \gamma \) at height \( y(p) \). It follows from the definition of \( \beta \) that \( q' \in B \), \( \beta_{y(p)} \leq x(p_{i+1}) \), and that the path \( \gamma \) continues to the North immediately after \( q' \). Since \( q' \triangleq p_{i+1} \) and \( p_{i+1} \) is a vertex of either \( A \) or \( B \), there must be at least one more East step in \( \gamma \) after its North segment from \( q' \). Thus \( q' + (0, r) \in A \) for some \( r > 0 \). It remains to show that \( q' - (1, 0) \in A \). Let \( q^- \in \alpha \) be the point with \( x \)-coordinate \( x(q') - 1 \), and \( q^+ \in \alpha \) be the point with \( x \)-coordinate \( x(q') \). That is, \( q^- \) is the point of \( \gamma \) at which the last East move before \( q' \) occurred and let \( q^+ \in \alpha \) be the point of \( \gamma \) at which the first East move after \( q' \) occurs. It is immediate that \( q' - (1, 0) \) lies in the convex hull of \( q^-, q^+ \) and \( p \), which are all points in \( A \), giving \( q' - (1, 0) \in A \). The reasoning of the base step, case (III) then applies to show that (ii) holds.

This concludes the induction step, and thus establishes (i) and (ii), as desired.

\( \square \)

Dually, one may define the path of the minimal word via Corollary 4.3 and \textbf{Maxpath}. We denote this minimal path by \( \gamma' \).

\textbf{Corollary 4.10.} The cardinality of the \( \sim_2 \) equivalence class containing \( w \in W(\ell_a, \ell_b) \) is equal to the number of lattice paths bounded by \( \gamma_1 \) and \( \gamma_2 \).
Example 4.11. Consider the word
\[ w = baabbaababaababaabaababaaba. \]
Figure 6(A) shows \( A(w) \) and its vertices in blue, and \( B(w) \) and its vertices in orange. The boxes between the points \( p_i \) and \( p_{i+1} \) are shown in grey. Figure 6(B) and 6(C) show the paths \( \gamma \) and \( \overline{\gamma} \), respectively, in black. Figure 6(D) shows \( \gamma \) in black and \( \overline{\gamma} \) in red. The minimal and maximal words are
\[ \underline{w} = baabababababababaababaaba, \quad \overline{w} = baabababababaababaababaaba, \]
and one can easily see that the equivalence class of \( w \) has size \( 2^5 = 32 \).

Example 4.12 (The Generalized Catalan family). For integers \( r,k \geq 2 \), define the word \( w(r,k) \in W(r+k+1,r+k+1) \) by
\[ w(r,k) = (ab)^k(ab)r(a^kb). \]
We call the \( \sim_2 \)-equivalence class of \( w(r,k) \) the generalised Catalan family \( C(r,k) \). Notice that in this case, \( A := A(w(r,k)) \) is a trapezium with vertices
\[ V(A) = \{(0,0), (1,k), (r+1,r+k), (r+k,r+k+1)\}, \]
while \( B := B(w(r,k)) \) is another trapezium with vertices
\[ V(B) = \{(1,0), (1,k-1), (r+1,r+k-1), (r+k+1,r+k)\}. \]
The pair of lines constructed by extending the parallel edges of \( B \) are the shift by \( (1,0) \) of the corresponding lines constructed from the parallel edges of \( A \). The path of each word in \( C(r,k) \) is thus constrained to pass through
each point of \( V(A) \cup V(B) \cup \{(2,k),(r+k+1,r+k+1)\} \), whilst remaining within the union of the two trapezia. By Theorem 4.9, \( w(r,k) \) is the maximal element of this equivalence class. The minimal word is given by \((ab^k)a^rb^r(a^k b)\) when \( r < k \) and \((ab^k)a^k(ba)^{r-k}b^k(a^k b)\) when \( r \geq k \). Figure 7 illustrates the paths of the two words and the corresponding signature for the family \( C(5,4) \).

**Figure 7.** The degree-one signature of \( w(5,4) \) from Example 4.12 with maximal and minimal words shown. The size of the equivalence class is the number of Dyck paths of length 8 with height at most 4.

The size of the equivalence class \( C(r,k) \) is equal to the number of staircase paths from \((0,0)\) to \((r,r)\) weakly bounded by the lines \( y = x \) and \( y = x - k \), or in other words, the number of Dyck paths of length \( 2r \) with height at most \( k \). A formula for the size of this equivalence class is therefore provided by [11, Theorem 10.3.4]:

\[
|C(r,k)| = 2^{2r+2} \frac{k+1}{k+2} \sum_{j=1}^{\frac{k+1}{2}} \cos^{2r} \left( \frac{j\pi}{k+2} \right) \sin^2 \left( \frac{j\pi}{k+2} \right)
\]

For fixed \( k \) and \( r \to \infty \), the first order term in \( |C(r,k)| \) is

\[
|C(r,k)| \sim O \left( 4^r \cos^{2r} \left( \frac{\pi}{k+2} \right) \right) \sim O \left( 2^\ell \cos^\ell \left( \frac{\pi}{k+2} \right) \right),
\]

where \( \ell = 2(r+k+1) \) is the word length. That is, this equivalence class grows exponentially fast in the word length. Here are some asymptotics for low values of \( k \).

| \( k \) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| \( |C(r,k)| \) | \( 2^r \) | 2.61^r | 3^r | 3.24^r | 3.41^r | 3.53^r | 3.61^r | 3.68^r | 3.73^r |

5. Algorithms

In this section we present and discuss our algorithms. We analyse their time complexity, assuming a random access model of computation in which arithmetic operations and comparison takes unit time.
5.1. Algorithms for the general case. The Signature algorithm computes the degree-$d$ signature of a word. It underpins the first step in all algorithms on a general alphabet. To compute the degree $d$ signature of a word from definition, we need to compute the support of the polynomials $g_u^w$ defined in (2) for all $u \in \Sigma^d$. This means for each $u \in \Sigma^d$, one needs to list the ways that $u$ can appear as a scattered subword of $w$, and in each case, for each letter $s \in \Sigma$, count the number of occurrences of $s$ within each intermediary factor. To improve efficiency, instead of computing $g_u^w$ for one $u$ at a time, the Signature algorithm computes the monomial $\odot_{s\in\Sigma} \odot_{k=1}^{d} x(s,k)^{N_v^\pi(s_k-1,n_k)}$ for each $\pi$ in (3), then uses a lookup table indexed by elements of $\Sigma^d$ to associate this monomial with the correct polynomial $g_u^w$.

Lemma 5.1. Signature$(w,d)$ outputs the degree-$d$ signature of $w$, and has time complexity $O(m^d \binom{d}{\ell}^2)$, where $\ell = |w|$. When $m = 2$ and $d = 1$, it has time complexity $O(\ell)$.

Proof. From (2), it is straightforward to verify that $g_u^w = L[w]$ if and only if $u$ appears as a scattered subword of $w$. This proves the algorithm’s correctness. Now we analyse its time complexity. The calculation in Lines 2 and 3 requires $m\ell$ counts, and thus has time complexity $O(m\ell)$. Lines 10-13 are executed within three loops, and thus has time complexity $O(md\binom{d}{\ell})$. The last loop computes the convex hull of the support of $g_u^w$ for each $u \in \Sigma^d$. There are $md$ such polytopes to compute in the signature (some of which can be empty). Each polytope is the convex hull of at most $O(d^2)$ points, so in the worst case the time complexity of computing their convex hull is $O\left(d^2 \binom{d}{\ell}^2\right)$. Thus, overall, the time complexity of the algorithm is $O(md\binom{d}{\ell}) + O(m^2 \binom{d}{\ell}^2) = O(m^d \binom{d}{\ell}^2)$. For $m = 2$ and $d = 1$, the time complexity of computing the support of each polynomial is $O(\ell)$. As these points are naturally sorted, the two polygons in the signature can be computed in linear time using Andrew’s monotone chain algorithm [2]. Thus, the case $m = 2, d = 1$ has time complexity $O(\ell)$. \hfill \Box

Proof of Theorem 1.2 for $m > 2$. Fix $n$. Problem CheckPair is solved by Algorithm 3, which calls Signature$(w,d)$ for $d = 1, \ldots, n - 1$. Thus it has time complexity of the worst-case, which is that of Signature$(w,n-1)$. Problem ListAll is solved by Algorithm 4 which first computes the UT$n(T)$ signatures for all words in $W(\ell_{a_1}, a_i \in \Sigma)$, and stores these as keys in a hash table. There are $\binom{\ell}{\ell_1,\ldots,\ell_m}$ such words, and by Lemma 5.1 computing the UT$n(T)$ signature in each case has complexity $O(m^{n-1} \binom{\ell}{n-1}^2)$. Therefore, the time complexity is $O(m^{n-1} \binom{\ell}{\ell_1,\ldots,\ell_m} \binom{\ell}{n-1}^2)$. Finally, ListWord is solved by Algorithm 5 which finds a specific entry from the lookup table created from Algorithm 4 and thus has the same time complexity. \hfill \Box

5.2. Two-letter alphabets. For a two-letter alphabet and $n = 2$, there is a natural speedup of the equivalence class enumeration algorithm. In particular, thanks to the Structural Theorem, to represent an equivalence class it is sufficient to compute the minimal and maximal element, which
Algorithm 2 Signature

Input: Word $w$ on alphabet $\Sigma = \{a_1, \ldots, a_m\}$, positive integer $d$

Output: Degree-$d$ signature of $w$

1: $L \leftarrow$ empty dictionary
2: $c(0, s) \leftarrow 0$ for all $1 \leq s \leq m$
3: for $1 \leq i \leq |w|$ do
4: \quad $c(i, s) \leftarrow$ number of occurrences of letter $a_s$ before $i$th letter of $w$
5: for $\pi \in [\binom{d}{j}]$ do
6: \quad $u \leftarrow (w_{\pi_1}, \ldots, w_{\pi_d})$, $u_0 \leftarrow 0$, $h \leftarrow \emptyset \in \mathbb{R}^{md}$
7: \quad for $1 \leq s \leq m$ do
8: \quad \quad if $u_{k-1} = a_s$ then
9: \quad \quad \quad $N_s^w(\pi_{k-1}, \pi_k) \leftarrow c(\pi_k, s) - c(\pi_{k-1}, s) - 1$
10: \quad \quad else
11: \quad \quad \quad $N_s^w(\pi_{k-1}, \pi_k) \leftarrow c(\pi_k, s) - c(\pi_{k-1}, s)$
12: \quad \quad \quad $h \leftarrow h + N_s^w(\pi_{k-1}, \pi_k)e_{(k-1)m+s}$
13: \quad if $u \notin \text{keys}(L)$ then $L[u] \leftarrow \{h\}$ else $L[u] \leftarrow L[u] \cup \{h\}$
14: for $u \in \Sigma^d$ in lexicographic order do
15: \quad $P_u^w \leftarrow$ convex hull of $L[u]$
16: return $(P_u^w, u \in \Sigma^d)$

Algorithm 3 Identity

Input: words $w, w'$, integer $n \geq 2$

Output: True iff $w \sim_n w'$

1: for $d = 1, \ldots, n-1$ do
2: \quad if $\text{Signature}(w, d) \neq \text{Signature}(w', d)$ then return False
3: return True

Algorithm 4 ListIdentitiesGeneral

Input: $n \geq 2, c \in \mathbb{N}^{[\Sigma]}$

Output: All $UT_n(T)$ identities for words $w \in W(c)$.

1: $L \leftarrow$ empty dictionary
2: for $w \in W(c)$ do
3: \quad $V^w \leftarrow (\text{Signature}(w, 1), \ldots, \text{Signature}(w, n-1))$
4: \quad $L[V^w] \leftarrow w$
5: return $L$

Algorithm 5 EquivalenceClassGeneral

Input: $n \geq 2$, $\Sigma$ an alphabet of size $m$, $w \in \Sigma^+$

Output: Equivalence class of $w$ in $UT_n(T)$

1: $c = \binom{|w|}{a_i}, a_i \in \Sigma$
2: $V^w \leftarrow (\text{Signature}(w, 1), \ldots, \text{Signature}(w, n-1))$
3: $L = \text{ListIdentitiesGeneral}(n, c)$
4: return $L[V^w]$
can be done very efficiently. This leads to an algorithm with optimal time
complexity for equivalence class computation for $\mathcal{UT}_2(\mathbb{T})$, and a very efficient
algorithm for identity enumeration. As $\mathcal{UT}_n(\mathbb{T})$ identities must also be
$\mathcal{UT}_{n-1}(\mathbb{T})$ identities, this also gives a shortcut for these problems for $\mathcal{UT}_n(\mathbb{T})$
over two-letter alphabets.

Algorithm 6 Minmax

**Input:** word $w \in W(\ell_a, \ell_b)$

**Output:** minimal and maximal words $(w, \overline{w})$ of the equivalence class of $w$
in $\mathcal{UT}_2(\mathbb{T})$

1: $\overline{w} \leftarrow \text{Maxword}(w)$
2: $w \leftarrow \text{Maxword}(\overline{w})$
3: **return** $(w, \overline{w})$.

4: **procedure** Maxword$(w)$
5: $(A, B) \leftarrow \text{Signature}(w, 1)$
6: $V \leftarrow \text{vertex}(A) \cup \text{vertex}(B) \cup (\ell_a, \ell_b)$, sorted in increasing $\preceq$ and labelled.
7: for $p_i, p_{i+1} \in V$ do
8: $\gamma(i) \leftarrow \text{MaxPath}(p_i, p_{i+1}, \text{label}(p_i), \text{label}(p_{i+1}))$
9: Compute $w$ from $\gamma$ via (6)
10: **return** $w$.

Algorithm 7 EquivalenceClass

**Input:** word $w \in W(\ell_a, \ell_b)$, $n \geq 3$

**Output:** equivalence class of $w$ in $\mathcal{UT}_{n'}(\mathbb{T})$ for all $2 \leq n' \leq n$

1: $(w, \overline{w}) \leftarrow \text{Minmax}(w)$
2: $E_1 \leftarrow \{v \in W(\ell_a, \ell_b) : w \preceq v \preceq \overline{w}\}$.
3: for $n'$ from 3 to $n$ do
4: $T \leftarrow \text{Signature}(w, n'-1)$.
5: $E_{n'-1} \leftarrow \emptyset$
6: for $v \in E_{n'-2}$ do
7: $S(v) \leftarrow \text{Signature}(v, n'-1)$
8: if $S(v) = T$ then $E_{n'-1} + v$
9: **return** $E_n$.

Lemma 5.2. Algorithm 6 has time complexity $O(\ell)$.

**Proof.** Step 1 of Procedure Maxword has complexity $O(\ell)$ as noted in Lemma 5.1. Since the vertices of $A$ and $B$ are already sorted, merging them to a labelled and sorted list $V$ requires $O(\ell)$ operations. Computation of $\gamma$ requires one to compute $O(\ell)$ coordinate slices of the lattice polytopes $A(w)$ and $B(w)$, and thus has complexity $O(\ell)$. Finally, reading $w$ from $\gamma$ requires $O(\ell)$ operations. \hfill $\square$

**Proof of Theorem 1.2 for $m = 2$.** Lemma 5.1 implies that for $n = 2$, Algorithm 2 has time complexity $O(\ell)$. Now consider the problem ListWord. Fix $n \geq 3$. For each $d = 2, \ldots, n-1$ Algorithm 7 computes the $d$-signature of
Algorithm 8 ListIdentities2

Input: \( \ell_a, \ell_b \in \mathbb{N} \)

Output: All \( \mathcal{UT}_2(\mathbb{T}) \) identities for all words in \( W(\ell_a, \ell_b) \)

1: \( L \leftarrow \) empty dictionary
2: \( \text{wlist} = [a^{\ell_a}b^{\ell_b}] \)
3: while \( \text{wlist} \neq \emptyset \) do
4: \( \text{wlist}' = \emptyset \)
5: for \( w \in \text{wlist} \) do
6: \( (w, \overline{w}) \leftarrow \text{MinMax}(w) \)
7: if \( w \notin \text{keys}(L) \) then
8: \( L[w] \leftarrow (w, \overline{w}) \)
9: \( \text{wlist}' \leftarrow \text{wlist}' \cup \{w' : w' \leftrightarrow w, w' > w\} \).
10: \( \text{wlist} \leftarrow \text{wlist}' \)
11: return \( L \)

Each word in the \( \sim_d \) class of \( w \). Thus its complexity is \( \sum_{d=2}^{n-1} O(2^d(\ell_d)^2)N_d(w) \).

For \( n = 2 \), the equivalence class is efficiently represented by the minimal and maximal elements, which can be computed in \( O(\ell) \) operations by Lemma 5.2. Finally, consider problem ListAll. For \( n = 2 \), each equivalence class is represented by its min-max pair \((w, \overline{w})\) in Algorithm 8. For each equivalence class, we make at most \( \ell \) calls to MinMax, which by Lemma 5.2 has complexity \( O(\ell) \). Thus each equivalence class requires at most \( O(\ell^2) \) computations.

So the overall time complexity is \( O(T_2(\ell_1, \ll_2)\ell^2) \). For \( n \geq 3 \), the worst-case complexity we give is that of the general-alphabet algorithm. Finally, to see that the algorithms for \( \mathcal{UT}_2(\mathbb{T}) \) have optimal time complexity, note that any algorithm would need to read in the word. Reading a word of length \( \ell \) already takes \( O(\ell) \) operations, and thus the first two algorithms for \( \mathcal{UT}_2(\mathbb{T}) \) have optimal time complexity.

Proof of Theorem 1.3. We use Algorithm 4 to list all \( \mathcal{UT}_3(\mathbb{T}) \) identities of lengths 21 and 22. The algorithm verifies that there are no \( \mathcal{UT}_3(\mathbb{T}) \) identities of length 21, and returns the identities in Theorem 1.3 as the only identities of length 22.

Remark 5.3. Pastijn [15] gave an algorithm for identity verification in the bicyclic monoid (and hence in \( \mathcal{UT}_2(\mathbb{T}) \) by [6]) via linear programming. Given two words \( w, v \), Pastijn’s algorithm essentially checks for equality of the polygons \( A(w) \) and \( A(v) \) by recursively checking if each point \( p \in A(w) \) can be separated from \( A(v) \) by a hyperplane. In the case of a two letter alphabet, this naïve convex hull computation has time complexity \( O(\ell^2) \), where \( \ell \) is the length of the word. In comparison, our algorithm attains the optimal time complexity of \( O(\ell) \) as shown in Theorem 1.2.

6. Statistics on \( \mathcal{UT}_2(\mathbb{T}) \) and \( \mathcal{UT}_3(\mathbb{T}) \): data and conjectures

In this section we discuss six conjectures obtained by analyzing the data obtained from our algorithms. We hope that these conjectures will fuel new developments at the intersection of semigroup theory, combinatorics and probability.
6.1. Structural theorem for $UT_n(\mathbb{T})$. On a two-letter alphabet, the Structural Theorem leads to an efficient algorithm to compute the minimal and maximal elements of an equivalence class in $UT_2(\mathbb{T})$, and thus resulted in the dramatic speedup for various algorithms for $UT_2(\mathbb{T})$. Unfortunately, neither parts of the Structural Theorem hold for $UT_3(\mathbb{T})$.

**Lemma 6.1.** There exists words $w, w'$ over a two letter alphabet such that $w \lor w' \sim_3 w \land w'$ but $w \not\sim_3 w'$. There also exists another pair of words $w, w'$ over a two letter alphabet such that $w \land w' \sim_3 w \sim_3 w'$ but $w \not\sim_3 w \lor w'$.

**Proof.** For the first pair, let $u = baaaabaaaaababbbabbaa, v = babbabababababaa$, and define

$$w = u ba b ab v, \quad w' = u ab b ba v.$$  

Then $w \land w' = u ba b ba v$ and $w \lor w' = u ab b ab v$. For the second pair, let $u = abaaaabbbbaabbaabbbaa, v = abbbaababababababbb$, and define

$$w = u ab ba v, \quad w' = u ba ab v.$$  

Then $w \land w' = u ba ba v, w \lor w' = u ab ab v$. Figure 3(a) and (b) show the paths and $UT_2(\mathbb{T})$ signatures of $w \lor w', w \land w'$ for the first and second pair, respectively. It is immediate from this figure that in both examples, all four words $w, w', w \lor w'$ and $w \land w'$ are equivalent in $UT_2(\mathbb{T})$. However, in the first pair, $w$ is an isoterm in $UT_3(\mathbb{T})$, while the other three are equivalent.

In the second example, all but $w \lor w'$ are equivalent in $UT_3(\mathbb{T})$. $\square$

In the first example considered in the proof of the previous lemma, the three words lying in the same $\sim_3$ class are connected by adjacent swaps, however, unlike the situation for $UT_2(\mathbb{T})$, the order of these swaps is important. Based on further experiments, we conjecture that the property of being connected via adjacent swaps holds for $\sim_n$ classes more generally. That is, while the Structural Theorem does not hold for $n \geq 3$, the following conjecture states that Corollary 4.4 generalises.

**Conjecture 6.2.** Let $w, v \in \Sigma^*$ with $w \neq v$ and $|\Sigma| = 2$, and let $n \geq 2$. Then $w \sim_n v$, if and only if there exists a positive integer $k$ and a sequence of words $u(i)$, for $i = 0, \ldots, k$ such that

$$w = u(0) \leftrightarrow u(1) \leftrightarrow \cdots \leftrightarrow u(k) = v \text{ and } u(0) \sim_n u(1) \sim_n \cdots \sim_n u(k).$$

Conjecture 6.2 is not true for alphabets of size greater than two, see Example 4.5. On the other hand, it holds trivially for the existing families of $UT_n(\mathbb{T})$ identities constructed in [8, 19], as these are of the form $w \sim_n v$ for some $w \leftrightarrow v$.

One can formulate a weaker version of Conjecture 6.2 as follows. Say that a word is locally isolated in $UT_n(\mathbb{T})$ if $w' \leftrightarrow w$ implies $w' \not\sim_n w$, that is, $w$ does not form $UT_n(\mathbb{T})$ identities with any of its neighbours. By Corollary 4.4, a word over a two letter alphabet is an isoterm for $UT_2(\mathbb{T})$ if and only if it is locally isolated. We conjecture that this holds for all $UT_n(\mathbb{T})$.

**Conjecture 6.3.** A word over a two letter alphabet is an isoterm for $UT_n(\mathbb{T})$ if and only if it is locally isolated in $UT_n(\mathbb{T})$.  

Verifying whether a word is locally isolated in $\mathcal{UT}_n(T)$ can be done efficiently, since this amounts to comparing with each of its neighbours, whose number is bounded above by the length of the word. This allows us to estimate the fraction of words in $W(\ell_a, \ell_b)$ that are locally isolated in $\mathcal{UT}_n(T)$ through random sampling. Since an isoterm for $\mathcal{UT}_n(T)$ must be locally isolated, this gives numerical upper bounds for the fraction of $\mathcal{UT}_n(T)$ isoterms in $W(\ell_a, \ell_b)$. If Conjecture 6.3 holds, then this upper bound is tight.

Example 6.4. We estimate the fraction of locally isolated $\mathcal{UT}_3(T)$ words in $W(\ell_a, \ell_a)$ for $15 \leq \ell_a \leq 40$. For each value of $\ell_a$, we draw 50,000 independent samples from $W(\ell_a, \ell_a)$, and count the number of locally isolated $\mathcal{UT}_3(T)$ words. We plot the estimators together with the 95% confidence intervals in Figure 9. Note that the fraction of locally isolated words sharply drops with the word length. Conjecture 6.3 states that this fraction is in fact an unbiased estimate of the true fraction of $\mathcal{UT}_3(T)$ isoterms.

Existing $\mathcal{UT}_n(T)$ identities in the literature have rather intricate and symmetrical constructions [4, 7, 19]. In contrast, our experimental findings suggest that for long words, it is extremely easy to randomly generate identities. For instance, if one picks a random word from $W(40, 40)$, and just checks amongst its neighbours, Figure 9 suggests that the probability that one finds a $\mathcal{UT}_3(T)$ identity in this way is more than 85%.

Conjecture 6.5. Fix $r \in (0, 1)$ and $n \geq 2$. Let $i(n, \ell, r)$ be the fraction of $\mathcal{UT}_n(T)$ isoterms in $W(\ell r, \ell (1-r))$. Then $i(n, \ell, r)$ is a monotone decreasing function in $\ell$. Furthermore, $i(n, \ell, r) \to 0$ as $\ell \to \infty$.

6.2. Fraction of isoterms. What fraction of the identities which hold in $\mathcal{UT}_2(T)$ also hold in $\mathcal{UT}_3(T)$? For short words this fraction is small: for words with length less than 21, for example, it is zero. However, we claim that for long words, a randomly chosen $\mathcal{UT}_2(T)$ identity has a high chance of being a $\mathcal{UT}_3(T)$ identity. Specifically, consider the setup of Conjecture 6.5. Fix $r \in (0, 1)$, Consider words in $W(\ell r, \ell (1-r))$. Rescale the polygon $\text{conv}(a)$ so that it fits inside a rectangle with side lengths $r \times (1-r)$. For
most words, this implies the distance between adjacent points in the rescaled polygon are of order $O(\ell^{-1})$. As $\ell \to \infty$, this has the same effect as finely discretising the polytope $\text{conv}(\alpha)$. For large $\ell$, one would thus expect

$$\text{conv}(\alpha^w \times \alpha^w \cap C_a) = \text{conv}(\alpha^w \times \alpha^w) \cap C_a$$

for most words $w \in W(\ell r, \ell (1 - r))$, where $C_a$ is the polyhedron defined in Proposition 3.7. A similar argument applies to the other polytopes in the
degree-two signature. In view of the discussion following Lemma 3.7, this means for large $\ell$, most $UT_2(\mathbb{T})$ identities are also $UT_3(\mathbb{T})$ identities. This should especially be true for identities coming from a single adjacent swap.

**Conjecture 6.6.** Fix $r \in (0, 1)$ and consider words in $W(\ell r, \ell (1 - r))$. Let $w$ be a word chosen uniformly at random in $W(\ell r, \ell (1 - r))$, let $u$ be a word chosen uniformly at random from the $\sim 2$-class of $w$, and let $v$ be a word chosen uniformly at random from amongst the immediate neighbours of $w$ in this class. Then $P(w \sim 3 u)$ and $P(w \sim 3 v)$ are monotone increasing in $\ell$, and they tend to 1 as $\ell \to \infty$.

**Example 6.7.** Fix $r = 0.5$. For each $\ell = 200, 400, \ldots, 1200$, we draw 50 random words from $W(\ell/2, \ell/2)$. For each word $w$, we compute the ratio of $UT_3(\mathbb{T})$-class neighbours to $UT_2(\mathbb{T})$-class neighbours, that is,

$$\frac{|\{w' : w' \leftrightarrow w, w' \sim 3 w\}|}{|\{w' : w' \leftrightarrow w, w' \sim 2 w\}|}.$$

Figure 10 compares the distribution of these fractions for different $\ell$. In support of Conjecture 6.6, this fraction increases as $\ell$ increases.

Say that an equivalence class is a twin if it has size 2. Figure 11 below shows the distribution of isoterms, twin, and other classes amongst the $\sim 2$ classes of words of length 22, as the content of the word varies. The distribution is symmetric as one can swap $a$’s and $b$’s. While the distribution looks approximately binomial, a quick analysis reveals that it is much flatter: it has heavier tail, and the mode is not as high. Note that the distribution of number of words of length $\ell$ as a function of $\ell_a$ is indeed $\text{Binomial}(\ell, 1/2)$. Thus, should the distribution in Figure 11 be approximately binomial, one should find that the ratio of equivalence classes to total number of words be close to constant. We plot this ratio in Figure 12 (black line). One finds that this fraction drops sharply as $\ell_a$ approaches $n/2$, see Figure 12. A similar phenomenon is observed for the ratio of isoterms relative to the number of equivalence classes (cf. Figure 12, blue line). Based on this figure, we conjecture that the fraction of isoterms amongst equivalence classes is still asymptotically zero for a fixed letter ratio, as the word lengths tend to infinity. In other words, for long words, it is extremely difficult to construct isoterms through random sampling.

**Conjecture 6.8.** Fix $r \in (0, 1)$. Let $e(m, r)$ be the ratio of number of $UT_2(\mathbb{T})$ isoterms to the number of $UT_2(\mathbb{T})$ equivalence classes in $W(\ell r, \ell (1 - r))$. Then $e(m, r)$ is a monotone non-increasing function in $\ell$. Furthermore, $e(m, r) \to 0$ as $\ell \to \infty$.

We conclude with a conjecture of a combinatorial nature, concerning the size of the largest equivalence class amongst words with fixed number of letters. This conjecture has been verified numerically through exhaustive enumeration up to words of length 24.

**Conjecture 6.9.** For each fixed $\ell_a \geq 5$, the largest $UT_2(\mathbb{T})$ equivalence class in $W(\ell_a, \ell_a)$ belongs to the generalised Catalan family defined in Example 4.12.
Figure 10. Fraction of $UT_3(\mathbb{T})$-class neighbours to $UT_2(\mathbb{T})$-class neighbours of a randomly chosen word with equal number of $a$’s and $b$’s, for different word lengths.

7. Summary

This work gives new geometric methods and algorithms to verify, construct and enumerate identities which hold in the semigroup of upper-triangular $n \times n$ tropical matrices $UT_n(\mathbb{T})$. Each word over a fixed alphabet $\Sigma$ has an associated signature, which is a sequence of polytopes, and two words over the same alphabet form an identity in $UT_n(\mathbb{T})$ if and only if the first $\sum_{d=1}^{n-1} |\Sigma|^d$ terms of their signature agree. This translates problems concerning semigroup identities in $UT_n(\mathbb{T})$ to questions about polytopes. With this insight, we obtained the Structural Theorem for $UT_2(\mathbb{T})$ identities, a short proof of Adjan’s Theorem for the bicyclic monoid and explicit construction of the shortest $UT_3(\mathbb{T})$ identities. Furthermore, we show that the signature of a word can be efficiently computed. This yields algorithms to verify and enumerate $UT_n(\mathbb{T})$ identities. We implement these algorithms, and produce six interesting conjectures and a wealth of data that support them. These conjectures are at the intersection of combinatorics, semigroup
Figure 11. Number of isoterms, twins and non-twin classes of identities in $UT_2(T)$ for words with prescribed number of a’s and of length 22.

Figure 12. The ratio of $UT_2(T)$ classes to words in black, and $UT_2(T)$ isoterms to classes in blue, as a function of $\ell_a$, for binary words of length 22.

theory and probability. Our results and conjectures call for new research direction at this interface.
GEOMETRY OF UPPER TRIANGULAR TROPICAL MATRIX IDENTITIES

References

[1] S. I. Adjan. Defining relations and algorithmic problems for groups and semigroups. Proceedings of the Steklov Institute of Mathematics, No. 85 (1966). Translated from the Russian by M. Greendlinger. American Mathematical Society, Providence, R.I., 1966.

[2] A. M. Andrew. Another efficient algorithm for convex hulls in two dimensions. Information processing letters, 9(5):216–219, 1979.

[3] P. Butkovič. Max-linear systems: theory and algorithms. Springer Monographs in Mathematics. Springer-Verlag London, Ltd., London, 2010.

[4] A. J. Cain, G. Klein, L. Kubat, A. Malheiro, and J. Okniński. A note on identities in plactic monoids and monoids of upper-triangular tropical matrices. arXiv:1705.04596.

[5] Y. Chen, X. Hu, Y. Luo, and O. Sapir. The finite basis problem for the monoid of two-by-two upper triangular tropical matrices. Bull. Aust. Math. Soc., 94(1):54–64, 2016.

[6] L. Daviaud, M. Johnson, and M. Kambites. Identities in upper triangular tropical matrix semigroups and the bicyclic monoid. J. Algebra, 501:503–525, 2018.

[7] Z. Izhakian. Tropical plactic algebra, the cloaktic monoid, and semigroup representations. arXiv:1701.05156.

[8] Z. Izhakian. Semigroup identities in the monoid of triangular tropical matrices. Semigroup Forum, 88(1):145–161, 2014.

[9] Z. Izhakian. Erratum to: Semigroup identities in the monoid of triangular tropical matrices [ MR3164156]. Semigroup Forum, 92(3):733, 2016.

[10] Z. Izhakian and S. W. Margolis. Semigroup identities in the monoid of two-by-two tropical matrices. Semigroup Forum, 80(2):191–218, 2010.

[11] C. Krattenthaler. Lattice path enumeration. In Handbook of enumerative combinatorics, Discrete Math. Appl. (Boca Raton), pages 589–678. CRC Press, Boca Raton, FL, 2015.

[12] B. Lin and N. M. Tran. Linear and rational factorization of tropical polynomials. arXiv preprint arXiv:1707.03332, 2017.

[13] D. Maclagan and B. Sturmfels. Introduction to tropical geometry, volume 161. American Mathematical Soc., 2015.

[14] J. Okniński. Identities of the semigroup of upper triangular tropical matrices. Comm. Algebra, 43(10):4422–4426, 2015.

[15] F. Pastijn. Polyhedral convex cones and the equational theory of the bicyclic semigroup. J. Aust. Math. Soc., 81(1):63–96, 2006.

[16] Y. Shitov. A semigroup identity for tropical 3×3 matrices. ArXiv e-prints, June 2014.

[17] F. G. Shleifer. Looking for identities on a bicyclic semigroup with computer assistance. Semigroup Forum, 41(2):173–179, 1990.

[18] L. M. Shneerson. On the axiomatic rank of varieties generated by a semigroup or monoid with one defining relation. Semigroup Forum, 39(1):17–38, 1989.

[19] M. Taylor. On upper triangular tropical matrix semigroups, tropical matrix identities and T-modules. Ph. D. Thesis, University of Manchester, 2017.