Semigroup identities of tropical matrices through matrix ranks

Izhakian, Zur; Merlet, Glenn

Publication date:
2018

Document Version
Early version, also known as pre-print

Link to publication

Citation for published version (APA):
Izhakian, Z., & Merlet, G. (2018). *Semigroup identities of tropical matrices through matrix ranks*. ArXiv. http://arxiv.org/abs/1806.11028v1

SEMIGROUP IDENTITIES OF TROPICAL MATRICES THROUGH MATRIX RANKS

ZUR IZHAKIAN AND GLENN MERLET

Abstract. We prove the conjecture that, for any n, the monoid of all $n \times n$ tropical matrices satisfies nontrivial semigroup identities. To this end, we prove that the factor rank of a large enough power of a tropical matrix does not exceed the tropical rank of the original matrix.

Introduction

Tropical matrices are matrices over the max-plus semiring [28], that is $\mathbb{T} := \mathbb{R} \cup \{-\infty\}$ equipped with the operations of maximum as addition and summation as multiplication:

$$a \vee b := \max\{a, b\}, \qquad a + b := \mathrm{sum}\{a, b\}.$$

This semiring is additively idempotent, i.e., $a \vee a = a$ for every $a \in \mathbb{T}$, in which $\mathbb{0} := -\infty$ is the zero element and $\mathbb{1} := 0$ is the multiplicative identity. More generally, one may consider $\mathbb{T}$ as an ordered semiring whose addition is determined as maximum, e.g., a semiring obtained from an ordered monoid $(\mathcal{S}, \cdot\ )$ by setting the addition to be maximum and $\cdot$ as multiplication. $\mathcal{M}_n(\mathbb{T})$ denotes the monoid of all $n \times n$ square matrices with entries in $\mathbb{T}$, and induced multiplication. These matrices correspond uniquely to weighted digraphs (see [4, 24] for recent expositions), which play a central role in algebraic methods, applications to combinatorics, semigroup representations, automata theory, and many other methodologies.

Any finitely generated semigroup of tropical matrices has polynomial growth [3, 33]; thus the free semigroup on two generators is not isomorphic to a tropical matrix sub-semigroup. Growth rate of groups is an important subject of study in combinatorial and geometric group theory, delivered to semigroup theory as well, involving semigroup identities [34]. While Gromovs theory [12] implies that every finitely generated group having polynomial growth satisfies a nontrivial semigroup identity (since it is virtually nilpotent), Shneerson has given examples which show that this does not hold for semigroups [32].

Tropical matrices enable natural linear representations of semigroups; therefore, the question whether tropical matrices satisfy nontrivial semigroup identities arises immediately [19]. If they do satisfy identities, then any faithfully represented semigroup inherits these identities, and complicated computations are saved [16]. As well, these identities define varieties of tropically represented semigroups [29, Ch. VII], where matrix view may provide a classification (or bases) for these varieties. Birkhoff HSP Theorem states that varieties are the only classes of semigroup stable under homomorphisms, submonoids, and products. In addition, by the one-to-one correspondence, matrix identities are carried over to labeled weighted digraphs, with multiplication replaced by walk composition, and are interpreted as the impossibility of word separation in automata theory [8]. (See Section 1.4 for details.)

Semigroup identities have been found for certain submonoids of tropical matrices, including triangular matrices, and for arbitrary 2×2 and 3×3 matrices [14, 17, 19, 31, 27]. In this paper we prove the existence of identities for all $n \times n$ tropical matrices, for any n, to wit:

Theorem 3.7. *The monoid $\mathcal{M}_n(\mathbb{T})$ satisfies a nontrivial semigroup identity for every $n \in \mathbb{N}$. The length of this identity grows with n as $e^{Cn^2 + o(n^2)}$, for some $C \leqslant 1/2 + \ln(2)$.*

This theorem further supports the insight that, in many senses, the behavior of tropical matrices is similar to that of matrices over a field [2, 13, 18, 20, 21, 22, 23], and has immediate consequences in semigroup representations.

Corollary 3.9. *Any semigroup which is faithfully represented by $\mathcal{M}_n(\mathbb{T})$ satisfies a nontrivial identity.*

Date: June 29, 2018.

2010 *Mathematics Subject Classification.* Primary: 20M05, 20M07, 20M30, 47D03; Secondary: 16R10, 68Q70, 14T05.

Key words and phrases. Tropical (max-plus) matrices, idempotent semirings, semigroup identities, ranks of matrices, semigroup varieties, semigroup representations, word separation.

Our semigroup identities arise from an idea of Y. Shitov [31], resulting in Lemma 3.1, which paves the way to constructing identities for matrices by induction on their size. The further step towards this aim is detecting new relations for those matrices which cannot be factorized to a product of matrices of smaller size, said to have factor rank n. Unfortunately, Shitov was only able to deal with matrices having maximal determinantal rank, and thus to conclude the existence of identities only for 3×3 matrices. (See Definition 1.3 for various notions of rank and [1] for an extensive survey.)

To prove Theorem 3.7, we rely on tropical rank and give a generalization of the first author's result [17] to obtain identities for matrices of maximal rank (Theorem 3.4), based on identities of triangular matrices ([14, Theorem 4.10] or [27]). Since tropical rank is the smallest among other notions of ranks [1], especially smaller than determinantal rank, this is not enough to construct identities for $\mathcal{M}_n(\mathbb{T})$, and additional ingredient is needed. Specifying a new relationship between tropical and factor rank is then a crucial obstacle, confronted in this paper. We introduce two results of similar flavor.

Proposition 2.6. *Let $A \in \mathcal{M}_n(\mathbb{T})$ and $\overline{n} = \operatorname{lcm}(1, \dots, n)$. If $\operatorname{rk}_{\mathrm{tr}}(A^{\overline{n}}) < n$, then $\operatorname{rk}_{\mathrm{fc}}(A^{t\overline{n}}) < n$ for any $t \geqslant 3n-2$.*

Theorem 2.20. $\operatorname{rk}_{\mathrm{fc}}(A^t) \leqslant \operatorname{rk}_{\mathrm{tr}}(A)$ *for any $A \in \mathcal{M}_n(\mathbb{T})$ and $t \geqslant (n-1)^2 + 1$.*

The proof of the latter is based on the so-called weak CSR expansion – a method developed by T. Nowak, S. Sergeev, and the second author in [26]. The former is proven in the same spirit, but the simplification derived from the power $\overline{n}$ allows for a self-contained exposition of graph theoretic arguments.

These results are interesting for their own sake, as they introduce new relationships between different notions of rank, concerning also their tendency to unite for large powers. Indeed, in their earlier paper [11] the authors have shown that, taking powers of a matrix, at the limit all notions of rank coincide. This limit is reached for irreducible matrices, but the exponent can be arbitrary large.

The paper is organized as follows. Section 1 recalls the relevant setup and results to be used in the paper. Section 2 introduces the relationships between the factor rank of a matrix power and its original tropical rank. Section 3 applies these relationships to prove the existence of semigroup identifies for $\mathcal{M}_n(\mathbb{T})$.

1. Preliminaries

As the paper combines several areas of study, we provide the relevant background.

1.1. **Semigroup identities.** Given an **alphabet** $\mathcal{A}$, i.e., a finite set of **letters**, the free monoid of finite sequences generated by $\mathcal{A}$ is denoted by $\mathcal{A}^*$. The elements of $\mathcal{A}^*$ are termed **words**, its identity element is the empty word, denoted by e. The **length** of a word w, denoted by $\ell(w)$, is the number of its letters. We write $\#_a(w)$ for the number of occurrences of a letter a in w. Both $\ell(w)$ and $\#_a(w)$ are nonnegative integers. The free semigroup $\mathcal{A}^+$ is obtained from $\mathcal{A}^*$ by excluding the empty word.

A (nontrivial) **semigroup identity** is a formal equality $u = v$, written as pair $\langle u, v\rangle$, where u and v are two different words in $\mathcal{A}^+$, cf. [34]. For a monoid identity one allows u and v to be the empty word as well, i.e., $u, v \in \mathcal{A}^*$. The **length** of $\langle u, v\rangle$ is defined to be $\max\{\ell(u), \ell(v)\}$. An identity $\langle u, v\rangle$ is said to be an **n-letter identity**, if u and v involve at most n different letters from $\mathcal{A}$.

A semigroup $\mathcal{S} := (\mathcal{S}, \cdot\,)$ **satisfies a semigroup identity** $\langle u, v\rangle$, if

$$\phi(u) = \phi(v) \quad \text{for every semigroup homomorphism } \phi : \mathcal{A}^+ \longrightarrow \mathcal{S}. \tag{1.1}$$

The set of all semigroup identities satisfied by $\mathcal{S}$ is denoted by $\operatorname{Id}(\mathcal{S})$. Note that even if $\mathcal{S}$ is a monoid or a group, u, v are still taken to be elements of the free semigroup $\mathcal{A}^+$. With this setting, $\langle ab, e\rangle$ is not a legal semigroup identity, but it is a monoid identity.

Theorem 1.1 ([14, Theorem 3.10])**.** *A semigroup that satisfies an n-letter identity, $n \geqslant 2$, also satisfies a 2-letter identity of the same length.*

In this view, regarding existence of semigroup identities, one may restrict to a 2-letter alphabet. Therefore, in the sequel, **we always assume that** $\mathcal{A} = \{a, b\}$.

Notation 1.2. *Given a word $w \in \mathcal{A}^+$ and elements $s', s'' \in \mathcal{S}$, we write $w\,[\![s', s'']\!]$ for the evaluation of w in $\mathcal{S}$, obtained by substituting $a \mapsto s'$, $b \mapsto s''$. Similarly, we write $\langle u, v\rangle\,[\![s', s'']\!]$ for the pair of evaluations $u\,[\![s', s'']\!]$ and $v\,[\![s', s'']\!]$ in $\mathcal{S}$ of the words u and v.*

In the certain case that $\mathcal{S} = \mathcal{A}^+$, to indicate that for $u, v \in \mathcal{A}^+$ the evaluation $w\,[\![u, v]\!]$ is again a word in $\mathcal{A}^+$, we use the particular notation $w\,[u, v\,]$. Similarly, we write $\langle u, v\rangle\,[u, v\,]$ for $\langle u, v\rangle\,[\![u, v]\!]$.

With these notations, condition (1.1) reads as

$$\langle u, v\rangle \in \mathrm{Id}(\mathcal{S}) \quad \text{iff} \quad u\,[\![s', s'']\!] = v\,[\![s', s'']\!] \text{ for every } s', s'' \in \mathcal{S}.$$

Note also that, if $\langle u, v\rangle \in \mathrm{Id}(\mathcal{S})$, then $\langle u, v\rangle [w_1, w_2] \in \mathrm{Id}(\mathcal{S})$ for any $w_1, w_2 \in \mathcal{A}^+$.

1.2. **Tropical matrices.** Tropical matrices are matrices with entries in $\mathbb{T} := \mathbb{R} \cup \{-\infty\}$, whose multiplication is induced from the semiring operations of $\mathbb{T}$ as in the familiar matrix construction. The set of all $n \times n$ tropical matrices form the multiplicative monoid $\mathcal{M}_n := \mathcal{M}_n(\mathbb{T})$. The identity of $\mathcal{M}_n$, denoted by I, is the matrix with $\mathbb{1} := 0$ on the main diagonal and whose off-diagonal entries are all $\mathbb{0} := -\infty$. Formally, for any nonzero matrix $A \in \mathcal{M}_n$ we set $A^0 := I$. A matrix $A \in \mathcal{M}_n$ with entries $A_{i,j}$ is written as $A = (A_{i,j})$, where $i, j = 1, \dots, n$. We denote by $\mathcal{U}_n := \mathcal{U}_n(\mathbb{T})$ the submonoid of all (upper) tropical triangular matrices in $\mathcal{M}_n$. We write $\mathcal{M}_{m,n} := \mathcal{M}_{m,n}(\mathbb{T})$ for the set of all $m \times n$ tropical matrices. A **permutation matrix** is an $n \times n$ matrix $P_\pi = (P_{i,j})$, with π a permutation over $\{1, \dots, n\}$, such that $P_{i,\pi(i)} = \mathbb{1}$ for each $i = 1, \dots, n$ and $P_{i,j} = \mathbb{0}$ for all $j \neq \pi(i)$.

Definition 1.3. *Given a tropical matrix $A \in \mathcal{M}_n$.*

*(i) The **permanent** of A is defined as:*

$$\mathrm{per}(A) = \bigvee_{\pi \in \mathrm{S}_n} \sum_i A_{i,\pi(i)},$$

*where S_n denotes the set of all the permutations over $\{1, \dots, n\}$. The **weight** of a permutation $\pi \in \mathrm{S}_n$ is $\omega(\pi) = \sum_i A_{i,\pi(i)}$, so that $\mathrm{per}(A) = \bigvee_{\pi \in \mathrm{S}_n} \omega(\pi)$.*

*(ii) A is called **nonsingular**, if there exists a unique permutation $\tau_A \in \mathrm{S}_n$ that reaches $\mathrm{per}(A)$; that is, $\mathrm{per}(A) = \omega(\tau_A) = \sum_i A_{i,\tau_A(i)}$. Otherwise, A is said to be **singular**.*

*(iii) The **tropical rank** of A, denoted $\mathrm{rk}_{\mathrm{tr}}(A)$, is the largest k for which A has a $k \times k$ nonsingular submatrix. Equivalently, $\mathrm{rk}_{\mathrm{tr}}(A)$ is the maximal number of independent columns (or rows) of A for an adequate notion of independence* [20].

*(iv) The **factor rank** (also called Schein/Barvinok rank) of A, denoted $\mathrm{rk}_{\mathrm{fc}}(A)$, is the smallest k for which A can be written as $A = BC$ with $B \in \mathcal{M}_{n,k}$ and $C \in \mathcal{M}_{k,n}$. Equivalently, $\mathrm{rk}_{\mathrm{fc}}(A)$ is the minimal number of vectors whose tropical span contains the span of the columns (or rows) of A, or the minimal number of rank-one matrices A_i needed to write A additively as $A = \bigvee_i A_i$, cf.* [1].

*(v) The **trace** $\mathrm{tr}(A) = \sum_i A_{i,i}$ is the usual trace taken with respect to summation, although it corresponds to the tropical product of diagonal entries in $\mathbb{T}$.*

By definition of $\mathrm{rk}_{\mathrm{tr}}$, a matrix $A \in \mathcal{M}_n$ is nonsingular iff $\mathrm{rk}_{\mathrm{tr}}(A) = n$. From the last characterization of $\mathrm{rk}_{\mathrm{fc}}$ it readily follows that this rank is subadditive:

$$\mathrm{rk}_{\mathrm{fc}}(A \vee B) \leqslant \mathrm{rk}_{\mathrm{fc}}(A) + \mathrm{rk}_{\mathrm{fc}}(B). \tag{1.2}$$

As known, the above notions of rank do not coincide [2, §8]. Nevertheless, the inequality

$$\mathrm{rk}_{\mathrm{tr}}(A) \leqslant \mathrm{rk}_{\mathrm{fc}}(A) \tag{1.3}$$

holds for every $A \in \mathcal{M}_n$ [9, Theorem 1.4].

It is easily seen that $\mathrm{per}(A) \geqslant \mathrm{tr}(A)$ and $\mathrm{tr}(AB) \geqslant \mathrm{tr}(A) + \mathrm{tr}(B)$ for any $A, B \in \mathcal{M}_n$. Furthermore, for products of matrices, we have the following.

Theorem 1.4 ([13, Theorem 2.6], [21, Theorem 3.5], [25, Proposition 3.4])**.** *Any $A, B \in \mathcal{M}_n$ satisfy*

$$\mathrm{per}(AB) \geqslant \mathrm{per}(A) + \mathrm{per}(B).$$

If AB is nonsingular, then A and B are nonsingular, $\mathrm{per}(AB) = \mathrm{per}(A) + \mathrm{per}(B)$, and $\tau_{AB} = \tau_B \circ \tau_A$.

1.3. **Digraphs and automata.** Any matrix $(A_{i,j}) \in \mathcal{M}_n$ is uniquely associated with the **weighted digraph** $\mathrm{G}(A) := (\mathcal{V}, \mathcal{E})$ over the set of node $\mathcal{V} := \{1, \dots, n\}$ with a directed arc $\varepsilon_{i,j} := (i, j) \in \mathcal{E}$ of **weight** $A_{i,j}$ from i to j for every $A_{i,j} \neq \mathbb{0}$. With this one-to-one correspondence, we say that $\mathrm{G}(A)$ is the graph of the matrix A, and conversely that B is the matrix of the weighted digraph G', if $\mathrm{G}' = \mathrm{G}(B)$.

A **walk** γ on $\mathrm{G}(A)$ is a sequence of arcs $\varepsilon_{i_1,j_1}, \dots, \varepsilon_{i_m,j_m}$, with $j_k = i_{k+1}$ for every $k = 1, \dots, m-1$. We write $\gamma := \gamma_{i,j}$ to indicate that γ is a walk from $i = i_1$ to $j = j_m$. The **length** of a walk γ, denoted by $\ell(\gamma)$, is the number of its arcs. Formally, we may consider also walks of length 0, one on each node. The **weight** of γ, denoted by $\omega(\gamma)$, is the sum of weights of its arcs, counting repeated arcs.

We write $\gamma_{i,j} \circ \gamma_{j,h}$ for the composition of the walk $\gamma_{i,j}$ from i to j with the walk $\gamma_{j,h}$ from j to h. Similarly, $(\rho)^k$ denotes for the composition $\rho \circ \cdots \circ \rho$ of a loop ρ repeated k times. A walk $\gamma = \varepsilon_{i_1,i_2}, \dots, \varepsilon_{i_m,j_m}$ may also be viewed as the sequence of nodes $(i_1, i_2, \dots, i_m, j_m)$. For convenience, we use this point of view as well, depending of context, and write $i \in \gamma$ to indicate that the node i appears in γ.

A walk γ is **simple** (or elementary) if it has no repeated nodes, i.e., a node appears in γ at most once, except possibly as first and last node. A (simple) walk that starts and ends at the same node is called a **(simple) cycle**. An arc $\rho_i := \varepsilon_{i,i}$ is called a **loop**. A **1-cyclic walk** is a walk which contains simple cycles of length at most 1.

A digraph G is called **strongly connected**, if there is a walk from i to j for any nodes i, j. Maximal strongly connected subgraphs of G are called **strongly connected components** (**s.c.c.'s**). When there are no arcs between different s.c.c.'s, G is said to be **completely reducible**.

The **cyclicity** $\mathrm{cyc}(G)$ of a strongly connected digraph G is the greatest common divisor of the lengths of its cycles. If G is not strongly connected, then its cyclicity $\mathrm{cyc}(G)$ is the least common multiple of the cyclicities of its s.c.c.'s. It is well-known that the lengths of all walks on G which start at a same node i and end at a same node j are congruent modulo $\mathrm{cyc}(G)$.

Remark 1.5. *A permutation $\pi \in \mathrm{S}_n$ uniquely corresponds to a disjoint union Θ of simple cycles $\theta_1, \dots, \theta_m$ that cover all the nodes of $\mathrm{G}(A)$. We say that θ_t is a cycle of π, and write $\omega(\Theta) = \sum_t \omega(\theta_t)$, so that $\omega(\Theta) = \omega(\pi)$. The permanent of A is the highest weight $\omega(\Theta)$ over all such Θ. Accordingly, a matrix $A \in \mathcal{M}_n$ is nonsingular, if $\mathrm{G}(A)$ has a unique covering Θ of highest weight by simple cycles.*

The **spectral radius** of a matrix $(A_{i,j}) \in \mathcal{M}_n$ is the value

$$\lambda(A) = \bigvee_{j \leqslant n} \bigvee_{i_1,\dots,i_j} \frac{A_{i_1 i_2} + A_{i_2 i_3} + \cdots + A_{i_j i_1}}{j}, \tag{1.4}$$

that is, the maximal mean weight of (simple) cycles in $\mathrm{G}(A)$. A simple cycle of $\mathrm{G}(A)$ is called **critical**, if its mean weight equals $\lambda(A)$. A node of $\mathrm{G}(A)$ is said to be a **critical node**, if it belongs to some critical cycle. The **critical graph** of A, denoted by $\mathrm{G}^{\mathrm{cr}}(A)$, is the union of all the critical cycles of $\mathrm{G}(A)$, over the node set $\mathcal{V}$. If $\mathrm{G}(A)$ is acyclic, then $\lambda(A) = \mathbb{0}$ and $\mathrm{G}^{\mathrm{cr}}(A)$ has no arcs. We may also view $\mathrm{G}^{\mathrm{cr}}(A)$ as a subgraph of $\mathrm{G}(A)$ over a subset of nodes in $\mathcal{V}$.

The **kleene star** of A is the matrix

$$A^\star = \bigvee_{k \in \mathbb{N}} A^k.$$

(Some entries might be $+\infty$, unless A is normalized by $\lambda(A) \leqslant \mathbb{1}$.)

Remark 1.6. *Taking a power A^t of a matrix $A \in \mathcal{M}_n$, it is easily verified that the entry $(A^t)_{i,j}$ is the highest weight of walks from i to j of length t on $\mathrm{G}(A)$. Thus, the (i,j)-entry of $A^\star$ is the supremum of the weights of all walks from i to j on $\mathrm{G}(A)$. Obviously, the supremums are reached and $A^\star \in \mathcal{M}_n(\mathbb{T})$ if the weight of the cycles are nonpositive, that is if $\lambda(A) \leqslant \mathbb{1}$.*

While powers of a single matrix correspond to walks on $\mathrm{G}(A)$, to deal with products of matrices, $\mathrm{G}(A)$ needs a generalization. We restrict to products of two matrices, which suffices our purpose.

Definition 1.7. *The **labeled-weighted digraph** $\mathrm{G}(A, B)$, written **lw-digraph**, of matrices $A, B \in \mathcal{M}_n$ is the digraph over the nodes $\mathcal{V} := \{1, \dots, n\}$ with a directed arc $\varepsilon_{i,j}$ from i to j labeled a of weight $A_{i,j}$ for every $A_{i,j} \neq \mathbb{0}$ and a directed arc from i to j labeled b of weight $B_{i,j}$ for every $B_{i,j} \neq \mathbb{0}$. A walk $\gamma = \varepsilon_{i_1,j_1}, \dots, \varepsilon_{i_m,j_m}$ on $\mathrm{G}(A, B)$ is labeled by the sequence of arcs' labels along γ, from ε_{i_1,j_1} to ε_{i_m,j_m}, which is a word in $\{a, b\}^+$. (In particular, every walk labeled by w has length $\ell(w)$.) The weight $\omega(\gamma)$ of γ is the sum of its arcs' weights.*

Note that $\mathrm{G}(A, B)$ may have parallel arcs, but with different labels, and that $\mathrm{G}(A, A) = \mathrm{G}(A)$. With this definition, we have the following proposition.

Proposition 1.8. *Given a word $w \in \{a, b\}^+$ of length $\ell(w)$ and matrices $A, B \in \mathcal{M}_n$, the (i,j)-entry of the matrix $w[\![A, B]\!]$ is the maximum over the weights of all walks $\delta_{i,j}$ on $\mathrm{G}(A, B)$ from i to j labeled by w.*

lw-digraphs are the core of **weighted automata** – a widely studied extension of standard (i.e., boolean) automata. (See [10] for an overview on automata theory.)

Remark 1.9. *In automata theory, nodes are called states, arcs are called transitions, and one consider set of walks (also called runs) from an initial to a final state. The initial and final state might also have weights. Thus, a weighted automaton is defined by* $(\mathrm{G}(A,B), \mathbf{r}, \mathbf{c})$*, where* $\mathbf{r}$ *is a row vector and* $\mathbf{c}$ *is a column vector. The weight of a word* w *is the sum (here max) of weights of walks labeled by* w*, given by* $\mathbf{r}\,(w\,[\![A,B]\!])\,\mathbf{c}$.

1.4. **Word separation.** Due to Remark 1.9, existence of a nontrivial semigroup identity for $\mathcal{M}_n$ can be understood as the impossibility to **separate** two words by weighted automata. Recall that a standard automaton is said to separate a pair of words (u, v) if it accepts u but not v. Determining the size of the smallest automaton that separates a pair of words is an old open problem in automata theory. See [8] for a survey on the subject. Let $\mathrm{sep}(m)$ be the smallest size of automata, in terms of state number n, necessary to separate all pairs of words of length m (or at most m, since separating words of different length is easier). The best known upper bound is $\mathrm{sep}(m) = \mathrm{O}\left(m^{2/5}\ln^{3/5}(m)\right)$ [30, Theorem 3].

Since there are finitely many automata having $n = \mathrm{sep}(m)$ states, and 2^m words of length m in $\{a,b\}^*$, obviously, there is a pair of words that cannot be separated by such an automaton. This simple argument on cardinality gives words of length $2^{2n^2+\mathrm{o}(n)}$. This means that there exists a nontrivial semigroup identity of length $2^{2n^2+\mathrm{o}(n)}$, satisfied by the monoid $\mathcal{M}_n$ of $n\times n$ boolean matrices. Analyzing powers of boolean matrices, shorter identities of order $e^{n+\mathrm{o}(n)}$ are obtained, so that $\mathrm{sep}(m) \geqslant \ln(m) + \mathrm{o}(\ln(m))$. To the best of our knowledge, this is the best lower bound.

A weighted automaton is said to separate two words, if it assigns these words with different weights. As there are infinitely many weighted automata having a given number of states, it is not obvious that not all pairs of words can be separated by automata of a given size. Denote by $\mathrm{sep}_{\mathbf{S}}(m)$ the smallest size of weighted automata, having weights in the semigroup $\mathbf{S}$, necessary to separate all pairs of words of length m (or at most m). It follows from Remark 1.9 that $\mathrm{sep}_{\mathbf{S}}(m) > n$ iff there exists a semigroup identity for $\mathcal{M}_n(\mathbf{S})$. Theorem 3.7 implies that $\mathrm{sep}_{\mathbb{T}}(m) \geqslant c\ln^{1/2}(m)$ for some $c > 0$. As far as we know, there is no better upper bound than the one for boolean matrices.

2. Ranks of large powers of a matrix

In this section we assume that A is a matrix in $\mathcal{M}_n$, and set $\overline{n} = \mathrm{lcm}(1,\dots,n)$.

2.1. Direct approach.

Lemma 2.1. *If* $\mathrm{rk}_{\mathrm{tr}}(A^{\overline{n}}) = n$*, then* $\mathrm{per}(A^{\overline{n}}) = \mathrm{tr}(A^{\overline{n}})$.

Proof. Follows from Theorem 1.4. (See also [17, Lemma 2.8] or [31, Corollary 4].) ☐

We start with an easy lemma that links weights of permutations to weights of simple cycles.

Lemma 2.2. *Given a permutation* $\tau \in \mathrm{S}_n$*, let* μ_i *be the average weight of the unique (simple) cycle of* τ *that contains the node* i*. Suppose*

$$\omega(\theta) < \sum_{i\in\theta}\mu_i, \tag{2.1}$$

for every simple cycle θ *of* $\mathrm{G}(A)$ *which is not a cycle of* τ*, then* A *is nonsingular and* $\tau = \tau_A$.

Proof. A permutation $\pi \in \mathrm{S}_n$ corresponds to a disjoint union of simple cycles $\theta_1,\dots,\theta_m$ (cf. Remark 1.5). If θ_t is also a cycle of τ, then $\mu_i = \frac{\omega(\theta_t)}{\ell(\theta_t)}$ for every $i \in \theta_t$, by definition of μ_i. Therefore $\omega(\theta_t) = \sum_{i\in\theta_t}\mu_i$. Using (2.1), we see that

$$\omega(\pi) = \sum_{t=1}^{m}\omega(\theta_t) \leqslant \sum_{t=1}^{m}\sum_{i\in\theta_t}\mu_i = \sum_{i=1}^{n}\mu_i = \omega(\tau),$$

where equality can only be reached if all cycles of π are cycles of τ; that is, if $\pi = \tau$. This implies that τ is maximally unique; hence, A is nonsingular and $\tau = \tau_A$. ☐

Remark 2.3. *The critical nodes of* $\mathrm{G}^{\mathrm{cr}}(A^{\overline{n}})$ *and* $\mathrm{G}^{\mathrm{cr}}(A)$ *are the same, while* $\mathrm{G}^{\mathrm{cr}}(A^{\overline{n}})$ *also has a loop at each node belonging to a critical cycle of* A.

More generally:

Lemma 2.4 ([24, Lemma 3.6])**.** *The matrix of* $\mathrm{G}^{\mathrm{cr}}(A^t)$ *is the* t*'th power of the matrix of* $\mathrm{G}^{\mathrm{cr}}(A)$*, for any* $t \geqslant 1$*. Moreover,* $\lambda(A^t) = t\lambda(A)$.

From this simple lemma, in the spirit of [26], we deduce:

Lemma 2.5. *Suppose that $B = A^{\overline{n}}$ for some $A \in \mathcal{M}_n$, and that $t \geqslant 2n-2$. For every $(B^t)_{i,j} \neq \mathbb{0}$ there exists a walk $\gamma_{i,j}$ from i to j on $\mathrm{G}(B)$, with weight $(B^t)_{i,j}$ and length t, of the form*

$$\gamma_{i,j} = \gamma_{i,h} \circ \rho^s \circ \gamma_{h,j}, \tag{2.2}$$

where ρ is a loop at node h, and $\gamma_{i,h}$, $\gamma_{h,j}$ are simple walks, possibly empty.[1]

Proof. Proof by induction on the matrix size n. The case of $n = 1$ is trivial.

Assume $n > 1$. Since $B^t = A^{t\overline{n}}$, for $(B^t)_{i,j} \neq \mathbb{0}$ there is a walk γ_0 on $\mathrm{G}(A)$ from i to j of length $t\overline{n}$ and weight $(B^t)_{i,j}$. Note that $\mathrm{G}(A)$ and $\mathrm{G}(B)$ have the same critical nodes (Remark 2.3), so we often say critical node without specific details. There are two cases.

Case I: γ_0 passes through a critical node c.
Let θ be a simple cycle, possibly a loop, on $\mathrm{G}^{\mathrm{cr}}(A)$ which contains c. We insert θ repeated $\overline{n}/\ell(\theta)$ times in γ_0 to obtain a walk γ_1 on $\mathrm{G}(A)$. This walk has length $(t+1)\overline{n}$ and weight $B_{i,j} + \overline{n}\lambda(A)$, where $\lambda(A)$ is the spectral radius of $\mathrm{G}(A)$ defined by (1.4).

The sequence of nodes positioned at $0, \overline{n}, 2\overline{n}, \dots, (t+1)\overline{n}$ in γ_1 determines a walk γ_2 on $\mathrm{G}(B)$ that passes through a critical node c' (not necessarily c). This walk has length $t+1$ and weight at least $B_{i,j} + \overline{n}\lambda(A) = B_{i,j} + \lambda(B)$.

- If c' appears twice (or more) in γ_2, then the closed subwalk from the first occurrence of c' in γ_2 to last occurrence can be replaced by loops on c'.
- If a node k appears twice (or more) on the same side of the occurrence of c' in γ_2, then the closed subwalk from the first occurrence of k to last occurrence can be replaced by loops on c'.

These exchanges do not decrease the weight of γ_2, since the loop at c' is critical. They provide a new walk γ_3 which can be decomposed as in (2.2), but has length $t+1$ and weight at least $(B^t)_{i,j} + \lambda(B) = (B^t)_{i,j} + B_{c,c}$. Since $t+1 \geqslant 2(n-1)+1$, while simple walks have length at most $n-1$, γ_3 has at least one loop ρ which can be removed to get the desired walk γ_4 of the right length. The weight of γ_4 is proved to be at least $(B^t)_{i,j}$, but, clearly, it cannot be strictly greater.

Case II: γ_0 does not pass through any critical node.
Then, γ_0 is a walk on the graph $\mathrm{G}(\widetilde{A})$, where $\widetilde{A}$ is the matrix obtained by deleting from A the rows and columns corresponding to critical nodes. Thus, $(B^t)_{i,j} = (\widetilde{A}^{t\overline{n}})_{i,j}$ and, by the induction hypothesis, $(B^t)_{i,j}$ is the weight of a walk $\gamma_{i,j}$ of length t on $\mathrm{G}(\widetilde{A}^{\overline{n}})$ of the form (2.2). By definition of $\widetilde{A}$ and Remark 1.6 the weights of the arcs of $\mathrm{G}(\widetilde{A}^{\overline{n}})$ are at most the weights of the corresponding arcs of $\mathrm{G}(B)$, so that the weight of $\gamma_{i,j}$ as a walk on $\mathrm{G}(B)$ is at least $(B^t)_{i,j}$. As it cannot be strictly greater, $\gamma_{i,j}$ has weight $(B^t)_{i,j}$. $\square$

Proposition 2.6. *If $\mathrm{rk}_{\mathrm{tr}}(A^{\overline{n}}) < n$, then $\mathrm{rk}_{\mathrm{fc}}(A^{t\overline{n}}) < n$ for any $t \geqslant 3n-2$.*

Proof. Set $B = A^{\overline{n}}$ and $t \geqslant 3n-2$. Assume that $\mathrm{rk}_{\mathrm{tr}}(B) < n$. Since B is singular, by Lemma 2.2 applied to the identity permutation, there is a simple cycle θ on $\mathrm{G}(B)$, which is not a loop, whose weight is at least the sum of weights of loops at its nodes. Let c be a node of θ whose loop has minimal weight, and let h_0 be the node proceeding c in θ. We prove that for any i,j:

$$(B^t)_{i,j} \leqslant \bigvee_{h \neq c} \left((B^n)_{i,h} + (B^{t-n})_{h,j} \right). \tag{2.3}$$

If $(B^t)_{i,j} = \mathbb{0}$, then this inequality holds trivially. Otherwise, let $\gamma_{i,j} = \gamma_{i,h} \circ \rho^s \circ \gamma_{h,j}$ be a walk given by Lemma 2.5, where h is a node of $\gamma_{i,j}$. Since $\gamma_{i,h}$ and $\gamma_{h,j}$ are simple walks, they have length at most $n-1$, so that all nodes of $\gamma_{i,j}$ at positions n to $t-n+1$ are the same, namely h.

- When $h \neq c$, (2.3) follows from

$$(B^t)_{i,j} = \omega(\gamma_{i,j}) = \omega\left(\gamma_{i,h} \circ \rho^{n-\ell(\gamma_{i,h})}\right) + \omega\left(\rho^{t-n-\ell(\gamma_{h,j})} \circ \gamma_{h,j}\right) \leqslant (B^n)_{i,h} + (B^{t-n})_{h,j}.$$

[1] Note that $\gamma_{i,j}$ needs not be 1-cyclic.

– If $h = c$, then ρ is the loop at c. We have $n-1+\ell(\gamma_{h,j})+\ell(\theta) \leqslant 3n-2 \leqslant t$ and, by definition of c, $\omega(\theta) \geqslant \ell(\theta)\omega(\rho)$, so that

$$\begin{aligned}(B^t)_{i,j} = \omega(\gamma_{i,j}) &= \omega\left(\gamma_{i,h} \circ \rho^{n-1-\ell(\gamma_{i,h})}\right) + \omega\left(\rho^{\ell(\theta)}\right) + \omega\left(\rho^{t-(n-1)-\ell(\gamma_{h,j})-\ell(\theta)} \circ \gamma_{h,j}\right) \\ &\leqslant \omega\left(\gamma_{i,h} \circ \rho^{n-1-\ell(\gamma_{i,h})}\right) + \omega\left(\theta\right) + \omega\left(\rho^{t+1-n-\ell(\gamma_{h,j})-\ell(\theta)} \circ \gamma_{h,j}\right) \\ &= \omega\left(\gamma_{i,h} \circ \rho^{n-1-\ell(\gamma_{i,h})} \circ \theta \circ \rho^{t+1-n-\ell(\gamma_{h,j})-\ell(\theta)} \circ \gamma_{h,j}\right) \\ &\leqslant (B^n)_{i,h_0} + (B^{t-n})_{h_0,j}.\end{aligned}$$

Thus, inequality (2.3) holds in all cases. Since the reverse inequality always holds, B^t is the tropical sum of the $n-1$ matrices $\left((B^n)_{i,h} + (B^{t-n})_{h,j}\right)_{i,j}$ with $h \neq c$. Each of these matrices has rank 1, as it is the tropical product of a row of B^n by a column of B^{t-n}. Therefore, Definition 1.3.(iv) of factor rank implies $\mathrm{rk}_{\mathrm{fc}}(B^t) < n$. ☐

2.2. **CSR approach.** To prove Theorem 2.20 below we use the so-called weak CSR expansion of powers, developed by T. Nowak, S. Sergeev and the second author [26]. We first recall the relevant setup and results.

Definition 2.7. *For a completely reducible subgraph H of $\mathrm{G}^{\mathrm{cr}}(A)$, $A \in \mathcal{M}_n$, we set*

$$M_H = \left(\left(-\lambda(A) + A\right)^{\mathrm{cyc}(H)}\right)^{\star},$$

and define the matrices $C = C_H, S = S_H, R = R_H$ in $\mathcal{M}_n$ as follows

$$C_{i,j} = \begin{cases} M_{i,j} & \text{if } j \in H, \\ \mathbb{0} & \text{otherwise,}\end{cases} \qquad S_{i,j} = \begin{cases} A_{i,j} & \text{if } (i,j) \in H, \\ \mathbb{0} & \text{otherwise,}\end{cases} \qquad R_{i,j} = \begin{cases} M_{i,j} & \text{if } i \in H, \\ \mathbb{0} & \text{otherwise.}\end{cases} \tag{2.4}$$

*The matrices C_H, S_H and R_H are named the **CSR terms** of A with respect to H.*

This CSR expansion provides a useful tool for analyzing tropical matrices, especially their powers. For this purpose, we are interested in products $C_H(S_H)^t R_H$ with $t \in \mathbb{N}$, whose interpretation in terms of walks on $\mathrm{G}(A)$ is given by Theorem 2.10 below.

Remark 2.8. *If $\mathrm{G}(A)$ is acyclic, then $\lambda(A) = \mathbb{0}$ and the matrices M_H, C_H, S_H, R_H are not defined by* (2.4)*. In this case, $\mathrm{G}^{\mathrm{cr}}(A)$ has no arcs, and we formally set these matrices to be zero matrix, which is consistent with Theorems 2.9 and 2.10 below.*

Alternatively to (2.4)*, when $\lambda(A) \neq \mathbb{0}$, the matrices C_H and R_H can be extracted respectively from the columns and the rows of M_H which are indexed by the nodes of H, while S_H can be obtained from the square submatrix indexed by the critical nodes of $\mathrm{G}^{\mathrm{cr}}(A)$. The products $C_H(S_H)^t R_H$ obtained with this approach are the same as those obtained via* (2.4)*. Note that M_H, C_H and R_H remain unchanged when (tropically) multiplying A by any $\alpha \in \mathbb{R}$, but S_H is multiplied by α.*

The matrix $B[A]$ is defined by[2]

$$(B[A])_{i,j} = \begin{cases} \mathbb{0} & \text{if } i \text{ or } j \text{ is a critical node in } \mathrm{G}^{\mathrm{cr}}(A), \\ A_{i,j} & \text{else.}\end{cases} \tag{2.5}$$

In graph view, the digraph $\mathrm{G}(B[A])$ is the subgraph of $\mathrm{G}(A)$ induced by the set of non-critical nodes, i.e., the digraph obtained from $\mathrm{G}(A)$ by omitting all arcs incident to critical nodes, in particular all arcs of $\mathrm{G}^{\mathrm{cr}}(A)$. Therefore, if $\mathrm{G}(A)$ is acyclic, then $B[A] = A$.

Theorem 2.9 ([26, Theorem 4.1])**.** *Given $A \in \mathcal{M}_n$, let C_H, S_H, R_H be the CSR terms* (2.4) *of A for $H = \mathrm{G}^{\mathrm{cr}}(A)$, and let $B[A]$ be the matrix* (2.5)*. Then*

$$A^t = C_H(S_H)^t R_H \vee (B[A])^t, \quad \text{for any } t \geqslant (n-1)^2 + 1. \tag{2.6}$$

Note that $\mathrm{G}^{\mathrm{cr}}(A)$ is a completely reducible subgraph of $\mathrm{G}(A)$, unless $\mathrm{G}(A)$ is acyclic. In the latter case, all matrices in (2.6) are zero, so the equation holds obviously.

A main approach for proving CSR results, for instance Theorem 2.9, is the interpretation of a product $C_H(S_H)^t R_H$ in terms of walks on $\mathrm{G}(A)$, based on the following notations ($\mathcal{N}$ denotes a node subset):

- $\mathcal{W}^{\,t}(i \to j)$ is the set of all walks from i to j of length t,

[2]It is called the **Nachtigall matrix subordinate to** A in [26], denoted there by B_N.

- $\mathcal{W}^{\,t}(i \xrightarrow{k} j) = \bigcup_{t_1+t_2=t} \{\gamma_{i,k} \circ \gamma_{k,j} \mid \gamma_{i,k} \in \mathcal{W}^{\,t_1}(i \to k),\ \gamma_{k,j} \in \mathcal{W}^{\,t_2}(k \to j)\}$,
- $\mathcal{W}^{\,t}(i \xrightarrow{\mathcal{N}} j) = \bigcup_{k \in \mathcal{N}} \mathcal{W}^{\,t}(i \xrightarrow{k} j)$,
- $\mathcal{W}^{*}(i \xrightarrow{\mathcal{N}} j) = \bigcup_{t \geqslant 0} \mathcal{W}^{\,t}(i \xrightarrow{\mathcal{N}} j)$,
- $\mathcal{W}^{\,t,p}(i \xrightarrow{\mathcal{N}} j) = \{\gamma \in \mathcal{W}^{*}(i \xrightarrow{\mathcal{N}} j) \,|\, \ell(\gamma) = t \mod p\}$, with $p \in \mathbb{N}$.

Theorem 2.10 ([26, Theorem 6.1])**.** *Let $A \in \mathcal{M}_n$ be a matrix with $\lambda(A) = \mathbb{1}$, and let C_H, S_H, R_H be the CSR terms of A for H a completely reducible subgraph of $\mathrm{G}^{\mathrm{cr}}(A)$. Let $p \in \mathbb{N}$ be a multiple of $\mathrm{cyc}(H)$, and let $\mathcal{N}$ be a subset of nodes of H that contains at least one node from each s.c.c. of H. Then, for every $i, j = 1, \dots, n$ and $t \in \mathbb{N}$:*

$$\big(C_H (S_H)^t R_H\big)_{i,j} = \max\big\{w(\gamma) \mid \gamma \in \mathcal{W}^{\,t,p}(i \xrightarrow{\mathcal{N}} j)\big\}. \tag{2.7}$$

The theorem has the following corollaries

Corollary 2.11 ([26, Corollary 6.2])**.** *$C_H(S_H)^t R_H$ depends only on the set of s.c.c.'s of $\mathrm{G}^{\mathrm{cr}}(A)$ intersecting H – a completely reducible subgraph of $\mathrm{G}^{\mathrm{cr}}(A)$.*

Corollary 2.12 ([26, Corollary 6.3])**.** *If $H_1, \dots, H_q$ are the s.c.c.'s of H, then*

$$C_H (S_H)^t R_H = \bigvee_{\xi=1}^{q} C_{H_\xi} (S_{H_\xi})^t R_{H_\xi}. \tag{2.8}$$

Definition 2.13. *Let H be a subgraph of $\mathrm{G}(A)$, and let $p \in \mathbb{N}$. The **cycle removal threshold** $\mathrm{T}^{p}_{\mathrm{cr}}(H)$ (resp. **strict cycle removal threshold** $\widetilde{\mathrm{T}}^{p}_{\mathrm{cr}}(H)$) of H is the smallest $T \in \mathbb{N} \cup \{0\}$ for which the following holds: for each walk $\gamma \in \mathcal{W}^{*}(i \xrightarrow{H} j)$ of length $\geqslant T$ there is a walk $\delta \in \mathcal{W}^{*}(i \xrightarrow{H} j)$ obtained from γ by removing cycles (resp. at least one cycle), and possibly inserting cycles from H, such that $\ell(\delta) \leqslant T$ and $\ell(\delta) = \ell(\gamma) \mod p$.*

Proposition 2.14 ([26, Proposition 9.5])**.** *Given a subgraph H of $\mathrm{G}(A)$ with m nodes, then*

$$\mathrm{T}^{p}_{\mathrm{cr}}(H) \leqslant pn + n - m - 1, \quad \text{for any } p \in \mathbb{N}.$$

Corollary 2.15. *For a simple cycle θ of $\mathrm{G}(A)$ with $\ell(\theta) \leqslant n - 1$ the following holds:*

$$\mathrm{T}^{\ell(\theta)}_{\mathrm{cr}}(\theta) \leqslant \ell(\theta) n + n - \ell(\theta) - 1 = \ell(\theta)(n-2) + n + \ell(\theta) - 1 \leqslant (n-1)^2 + 1 + (\ell(\theta) - 1).$$

Corollary 2.16. *For a node i of $\mathrm{G}(A)$ and $p \leqslant n$ the following holds:*

$$\widetilde{\mathrm{T}}^{p}_{\mathrm{cr}}(\{i\}) \leqslant pn + n - 2 + 1 \leqslant n^2 + (p - 1).$$

The next proposition allows to deal with Hamiltonian cycles.

Proposition 2.17 ([26, Proposition 9.4])**.** *For a simple cycle θ of length n in $\mathrm{G}(A)$ the following holds:*

$$\widetilde{\mathrm{T}}^{n}_{\mathrm{cr}}(\theta) \leqslant n^2 - n + 1 = (n-1)^2 + 1 + (\ell(\theta) - 1).$$

Remark 2.18. *The above bounds on $\mathrm{T}^{p}_{\mathrm{cr}}$ are applied to produce from a given walk $\gamma_1 \in \mathcal{W}^{\,t}(i \xrightarrow{H} j)$ a new walk $\gamma_2 \in \mathcal{W}^{\,t,p}(i \xrightarrow{H} j)$ with $\ell(\gamma_2) \leqslant T$, and the bound $\mathrm{T}^{p}_{\mathrm{cr}}(H) \leqslant T$, by omitting cycles from γ_1 and possibly inserting cycles of H.*

The next lemma (included implicitly in [26]) completes Theorem 2.10 for matrices $A \in \mathcal{M}_n$ that are not normalized, i.e., have $\lambda(A) \neq \mathbb{1}$.

Lemma 2.19. *Given $A \in \mathcal{M}_n$, $t \in \mathbb{N}$, and indices $i, j = 1, \dots, n$.*

(i) *$(C_H (S_H)^t R_H)_{i,j} \geqslant \omega(\gamma_{i,j})$ for any completely reducible subgraph H of $\mathrm{G}^{\mathrm{cr}}(A)$ and any walk $\gamma_{i,j} \in \mathcal{W}^{\,t}(i \xrightarrow{H} j)$;*

(ii) *Assume θ is a simple critical cycle. If $t \geqslant (n-1)^2 + 1$, then there exists $\gamma_{i,j} \in \mathcal{W}^{\,t}(i \xrightarrow{\theta} j)$ such that $(C_\theta (S_\theta)^t R_\theta)_{i,j} = \omega(\gamma_{i,j})$.*

Proof. Both in (i) and (ii), since $(C_H (S_H)^t R_H)_{i,j}$ and $\omega(\gamma_{i,j})$ decreases by $t\lambda(A)$ when replacing A by $(-\lambda(A)) + A$, we may assume that $\lambda(A) = \mathbb{1}$.

(i): Follows immediately from Theorem 2.10.

(ii): Take $\gamma_1 \in \mathcal{W}^{t,g}(i \xrightarrow{\mathcal{N}} j)$ having weight $(C_\theta(S_\theta)^t R_\theta)_{i,j}$, cf. Theorem 2.10. If $\ell(\theta) = n$, apply Proposition 2.17, otherwise use Corollary 2.15, to get a walk γ_2 such that $\ell(\gamma_2) = t \mod \ell(\theta)$ and $\ell(\gamma_2) \leqslant t+\ell(\theta)-1$, so $\ell(\gamma_2) \leqslant t$. If needed, insert additional copies of θ to get a walk γ_3 of length exactly t. Since $\lambda(A) = \mathbb{1}$, $\omega(\gamma_2) \geqslant \omega(\gamma_1)$. Since θ is critical, $\omega(\gamma_3) = \omega(\gamma_2)$. Thus, $\omega(\gamma_3) \geqslant \omega(\gamma_1) = (C_\theta(S_\theta)^t R_\theta)_{i,j}$. The reverse inequality is given by (i), so that $\gamma_{i,j} = \gamma_3$ has the desired properties. □

We are now ready to prove the main result of this section.

Theorem 2.20. $\mathrm{rk}_{\mathrm{fc}}(A^t) \leqslant \mathrm{rk}_{\mathrm{tr}}(A)$ *for any* $t \geqslant (n-1)^2 + 1$.

Proof. Fix $t \geqslant (n-1)^2+1$, and apply Theorem 2.9 recursively to get A^t as the sum of $C_{H_\xi}(S_{H_\xi})^t R_{H_\xi}$ defined by successive matrices subordinate to A. Explicitly, we start with $A_1 = A$ and define inductively $A_{\xi+1} = B[A_\xi]$. At each step, we set $C_\xi = C_{H_\xi}$, $S_\xi = S_{H_\xi}$, $R_\xi = R_{H_\xi}$ to be the CSR terms of A_ξ with respect to $H_\xi = \mathrm{G}^{\mathrm{cr}}(A_\xi)$.

By definition (2.5) of $B[A_\xi]$ we get a sequence of nested digraphs

$$\mathrm{G}(A) = \mathrm{G}(A_1) \supseteq \mathrm{G}(A_2) \supseteq \mathrm{G}(A_3) \supseteq \cdots, \tag{2.9}$$

such that for any $\xi > \varsigma$, $\mathrm{G}(A_\xi)$ is the subgraph of $\mathrm{G}(A_\varsigma)$ obtained by removing all the arcs of $\mathrm{G}(A_\varsigma)$ that are incident to some node that is critical for some A_ζ, where $\xi > \zeta \geqslant \varsigma$. Thus, $\mathrm{G}(A_\xi)$ can be viewed as a digraph on a subset of nodes of $\mathrm{G}(A)$, i.e., as an induced subgraph.

Since $\mathrm{G}(A_1)$ has finitely many nodes and $\mathrm{G}^{\mathrm{cr}}(A_\xi)$ and $\mathrm{G}^{\mathrm{cr}}(A_\varsigma)$ are arc-disjoint for any $\xi \neq \varsigma$, the sequence (2.9) stabilizes after finitely many steps, when $\mathrm{G}(A_\xi)$ is acyclic. Therefore, (2.9) restricts to matrices $A_1, \dots, A_q$ with strict inclusions, where $A_q^t = \mathbb{0}$, since $\mathrm{G}(A_q)$ is acyclic. Applying Theorem 2.9 recursively, we obtain

$$A^t = \bigvee_{\xi=1}^{q-1} C_\xi (S_\xi)^t R_\xi.$$

If $\xi < q$, then H_ξ is not acyclic; hence H_ξ is completely reducible. Let Θ_ξ be a collection of simple cycles of H_ξ that contains one cycle from each s.c.c. of H_ξ, each of them having minimal length. By Corollaries 2.11 and 2.12, we have

$$C_\xi(S_\xi)^t R_\xi = \bigvee_{\theta \in \Theta_\xi} C_\theta (S_\theta)^t R_\theta,$$

where C_θ, S_θ, R_θ are CSR terms of A_ξ with respect to $\theta \in \Theta_\xi$. Namely, the collection $\Theta = \bigcup_\xi \Theta_\xi$ of node-disjoint simple cycles gives

$$A^t = \bigvee_{\theta \in \Theta} C_\theta (S_\theta)^t R_\theta, \tag{2.10}$$

where $C_\theta, S_\theta, R_\theta$ are the CSR terms of the unique A_ξ such that θ is a simple cycle of $\mathrm{G}^{\mathrm{cr}}(A_\xi)$. The factor rank is subadditive, cf. (1.2), and thus (2.10) implies

$$\mathrm{rk}_{\mathrm{fc}}(A^t) \leqslant \sum_{\theta \in \Theta} \mathrm{rk}_{\mathrm{fc}}\left(C_\theta (S_\theta)^t R_\theta\right) \leqslant \sum_{\theta \in \Theta} \ell(\theta). \tag{2.11}$$

(Later we show that some terms can be omitted to get $\sum_{\theta \in \Theta} \ell(\theta) \leqslant \mathrm{rk}_{\mathrm{tr}}(A)$.)

Let Φ be a subcollection of Θ for which (2.10) holds as well. **Assume that**

$$\sum_{\theta \in \Phi} \ell(\theta) > \mathrm{rk}_{\mathrm{tr}}(A). \tag{2.12}$$

Denote by $\mathcal{N}$ the set of nodes of all $\theta \in \Phi$, and by Q the principal minor of A indexed by the nodes in $\mathcal{N}$. Since the $\theta \in \Phi$ are node-disjoint simple cycles, $\mathcal{N}$ has strictly more than $\mathrm{rk}_{\mathrm{tr}}(A)$ elements; hence Q is singular. By Lemma 2.2, applied to the permutation of $\mathcal{N}$ whose cycles are the $\theta \in \Phi$ (cf. Remark 1.5), $\mathrm{G}(Q)$ has a simple cycle $\theta \notin \Phi$ such that

$$\omega(\theta) \geqslant \sum_{i \in \theta} \sum_{H_\xi \ni i} \lambda(A_\xi). \tag{2.13}$$

(The right part runs over all nodes $i \in \theta$ and for each i accumulates the spectral radius $\lambda(A_\xi)$ for the unique H_ξ containing i.)

Let l and m be respectively the smallest and largest ξ such that $\theta \cap H_\xi \neq \varnothing$. Note that θ is a simple cycle belonging $\mathrm{G}(A_l)$, since all its node occur in H_ξ for some $\xi \geqslant l$. Assume first that $l = m$. Then, $\omega(\theta) \geqslant \ell(\theta)\lambda(A_l)$ by (2.13), implying that θ is a simple cycle of H_l. Since each simple cycle in $\Phi \cap \Theta_l$

belongs to a different *s.c.c.* of H_l, all nodes of θ appear in the same simple cycle $\tilde{\theta}$ in $\Phi \cap \Theta_l$. Since $\theta \notin \Phi$, there is an arc of θ that does not belong to $\tilde{\theta}$. Starting with this arc and going back along the arcs of $\tilde{\theta}$, we build a cycle of H_l shorter than $\tilde{\theta}$. This contradicts the minimality of the length of $\tilde{\theta}$.

We are left with the case where $l < m$. Let θ_ξ, with $\xi = l, m$, be a simple cycle in $\mathrm{G}^{\mathrm{cr}}(A_\xi)$ that belongs to Φ such that $\theta \cap \theta_\xi \neq \varnothing$, and let k_ξ be a node of this nonempty intersection. It remains to show that

$$C_{\theta_m}(S_{\theta_m})^t R_{\theta_m} \leqslant C_{\theta_l}(S_{\theta_l})^t R_{\theta_l}. \tag{2.14}$$

Fix indices i and j, for which $(C_{\theta_m}(S_{\theta_m})^t R_{\theta_m})_{i,j} \neq \mathbb{0}$. By Lemma 2.19, there is a walk $\gamma_1 \in \mathcal{W}^*(i \xrightarrow{\theta_m} j)$ on $\mathrm{G}(A_m)$ such that $\omega(\gamma_1) = (C_{\theta_m}(S_{\theta_m})^t R_{\theta_m})_{i,j}$. In particular, γ_1 intersects θ_m, and θ_m intersects θ at some k_m. Insert θ_m into γ_1 to get a walk $\gamma_2 \in \mathcal{W}^{\,t+\ell(\theta_m)}(i \xrightarrow{k_m} j)$. Note that γ_2 lives on $\mathrm{G}(A_m)$, so it visits at most $n-1$ different nodes, since all arcs incident to H_l do not belong to $\mathrm{G}(A_m)$.

By Corollary 2.16, applied to $p = \ell(\theta)$ and $i = k_m$, there is another walk $\gamma_3 \in \mathcal{W}^{\,t,\ell(\theta_m)}(i \xrightarrow{k_m} j)$ on $\mathrm{G}(A_m)$, of length at most $(n-1)^2 + \ell(\theta_m) - 1$ (cf. Remark 2.18). But $\ell(\gamma_3) = t \mod \ell(\theta_m)$, and thus $\ell(\gamma_3) \leqslant (n-1)^2 < t$. Inserting copies of θ_m into γ_3 at k_m , we get a walk $\gamma_4 \in \mathcal{W}^{\,t}(i \xrightarrow{k_m} j)$. Namely, γ_4 is obtained from γ_1 by adding a copy of θ and copies of θ_m, and removing cycles having average weight at most $\lambda(A_m)$, which is the average weight of θ_m. Therefore $\omega(\gamma_4) \geqslant \omega(\gamma_1) = (C_{\theta_m}(S_{\theta_m})^t R_{\theta_m})_{i,j}$.

Now we reduce γ_4 (cf. Remark 2.18), and then insert copies of θ at k_m to produce a new walk $\gamma_6 \in \mathcal{W}^{\,t}(i \xrightarrow{\theta_l} j)$ for which

$$\omega(\gamma_6) \geqslant \omega(\gamma_4) \geqslant \big(C_{\theta_m}(S_{\theta_m})^t R_{\theta_m}\big)_{i,j}\,. \tag{2.15}$$

Note that γ_4 is a walk on $\mathrm{G}(A_m)$, so it visits $\tilde{n} \leqslant n-1$ different nodes, at most $n - \ell(\theta)$ of which do not belong to θ, as γ_4 is also a walk on $\mathrm{G}(A)$.

- When $\tilde{n} < n-1$, we apply Proposition 2.14 with $p = \ell(\theta)$ to get a walk $\gamma_5 \in \mathcal{W}^{\,t,\ell(\theta)}(i \xrightarrow{\theta} j)$ of length at most $(\tilde{n}-1)\,\ell(\theta) + n - 1 \leqslant (n-1)^2 < t$. Then, we insert at least one copy of θ to get a walk γ_6 of length t. Since the average weight of θ is larger than $\lambda(A_m)$ by (2.13), and thus larger than the average weight of each cycle of γ_4, inequality (2.15) holds.
- The equality $\tilde{n} = n-1$ implies that θ_l is a loop at k_l. In this case, we apply Proposition 2.14 with $p = 1$ and get $\gamma_5 \in \mathcal{W}^*(i \xrightarrow{k_m} j)$ of length at most $n-1+n-1 = 2n-2$. Then we insert θ once and enough copies of the loop θ_l to get a walk γ_6 of length t. Since θ_l is a critical cycle for A_l, its average weight is $\lambda(A_l)$, while $\lambda(A_l) \geqslant \lambda(A_m)$; thus (2.15) holds.

Finally (2.14) follows from (2.15) by Lemma 2.19 (*i*), and we have proved that the sum $\sum_{\theta \in \Phi} \ell(\theta)$ is not minimal, as long as this sum is strictly larger than $\mathrm{rk}_{\mathrm{tr}}(A)$. Thus, the inequality $\mathrm{rk}_{\mathrm{fc}}(A^t) \leqslant \mathrm{rk}_{\mathrm{tr}}(A)$ follows from (2.11), applied to a minimal subcollection Φ that satisfies (2.10). □

3. Semigroup identities of tropical matrices

The following auxiliary results lead to Theorems 3.6 and 3.7. We begin with an idea of Y. Shitov [31], implemented in the following lemma.

Lemma 3.1. *Let $A, B, C \in \mathcal{M}_n$ such that $A = PQ$, where $P \in \mathcal{M}_{n\times k}$, $Q \in \mathcal{M}_{k\times n}$, $k < n$, and let $w \in \{a,b\}^+$. Then $(wa)\,[\![AB, AC\,]\!] = P\big(w\,[\![QBP, QCP\,]\!]\big)Q$.*

Proof. Straightforward by induction on the length of the word w. □

To deal with matrices that cannot be factorized as above, we use Theorem 3.4 which extends a result from [17]. To this ends additional results are needed, based on the following conditions: A pair of matrices $A, B \in \mathcal{M}_n$ and a word $w \in \{a,b\}^+$ satisfy (PR) if:

$$\mathrm{per}(A) = \mathrm{tr}(A), \quad \mathrm{per}(B) = \mathrm{tr}(B), \quad \text{and } \mathrm{rk}_{\mathrm{tr}}(w\,[\![A, B\,]\!]) = n. \tag{PR}$$

Lemma 3.2. *Assume that* (PR) *holds for $A, B \in \mathcal{M}_n$, $w \in \{a,b\}^+$, and write $w = w_1 w_2 \cdots w_{\ell(w)}$ as a sequence of letters. For each index $i = 1, \dots, n$ we have*

$$\big(w\,[\![A, B\,]\!]\big)_{i,i} = \sum_{t=1}^{\ell(w)} \big(w_t\,[\![A, B\,]\!]\big)_{i,i} = \#_a(w) A_{i,i} + \#_b(w) B_{i,i}, \tag{3.1}$$

i.e., the i'th diagonal entry of $w\,[\![A, B\,]\!]$ is $\#_a(w) A_{i,i} + \#_b(w) B_{i,i}$.

Proof. Applying Theorem 1.4, as (PR) holds, we have

$$\mathrm{per}\big(w\,[\![A,B\,]\!]\,\big) = \sum_{t=1}^{\ell(w)} \mathrm{per}\big(w_t\,[\![A,B\,]\!]\,\big) = \sum_{t=1}^{\ell(w)} \mathrm{tr}\,\big(w_t\,[\![A,B\,]\!]\,\big) \leqslant \mathrm{tr}\,\big(w\,[\![A,B\,]\!]\,\big),$$

implying the equality

$$\mathrm{tr}\,\big(w\,[\![A,B\,]\!]\,\big) = \sum_{t=1}^{\ell(w)} \mathrm{tr}(w_t\,[\![A,B\,]\!]), \tag{3.2}$$

since $\mathrm{per}\big(w\,[\![A,B\,]\!]\,\big) \geqslant \mathrm{tr}\,\big(w\,[\![A,B\,]\!]\,\big)$. Proposition 1.8 obviously implies

$$\big(w\,[\![A,B\,]\!]\,\big)_{i,i} \geqslant \sum_{t=1}^{\ell(w)} \big(w_t\,[\![A,B\,]\!]\,\big)_{i,i} = \#_a(w)A_{i,i} + \#_b(w)B_{i,i}\ , \tag{3.3}$$

since the right hand side corresponds to the weight of the walk from i to itself, composed of loops only. On the other hand, since $w\,[\![A,B\,]\!]$ is nonsingular, we have

$$\begin{aligned} \sum_{i=1}^{n} \big(w\,[\![A,B\,]\!]\,\big)_{i,i} &= \mathrm{tr}(w\,[\![A,B\,]\!]) \overset{(3.2)}{=} \sum_{t=1}^{\ell(w)} \mathrm{tr}(w_t\,[\![A,B\,]\!]) \\ &= \#_a(w)\,\mathrm{tr}(A) + \#_b(w)\,\mathrm{tr}(B) = \sum_{i=1}^{n} \big(\#_a(w)A_{i,i} + \#_b(w)B_{i,i}\big), \end{aligned}$$

so that the inequality in (3.3) cannot be strict for any i. Hence, (3.1) holds. □

Lemma 3.3. *Assume that* (PR) *holds for* $A,B \in \mathcal{M}_n$, $w \in \{a,b\}^+$. *For each entry* $W_{i,j} \neq \mathbb{0}$ *of* $W = w\,[\![A,B\,]\!]$ *there is a 1-cyclic walk* $\gamma_{i,j}$ *on* $\mathrm{G}(A,B)$, *labeled by* w *of weight* $W_{i,j}$.

Proof. An (i,j)-entry of $w\,[\![A,B\,]\!]$ corresponds to the weight of a walk $\gamma_0 := \gamma_{i,j}$ from i to j on $\mathrm{G}(A,B)$ labeled by w, by Proposition 1.8. Assume γ_0 is not 1-cyclic, which means that γ_0 returns to a node h which it has already left. Let γ_1 be the subwalk γ_1 of γ_0 which starts at the first occurrence h and ends at the last occurrence of h. Let v be the factor of w labeling γ_1. Since $v\,[\![A,B\,]\!]$ is a factor of $w\,[\![A,B\,]\!]$, it follows from Proposition 1.4 that $\mathrm{rk}_{\mathrm{tr}}(v\,[\![A,B\,]\!]) = n$, and A,B,v satisfy (PR). Hence, by Lemma 3.2 the walk γ_2 labeled by v that stays at h has weight at least as that of γ_1. Then γ_1 can be replaced by γ_2 in γ_0 to obtain a walk γ_3 that does not return to h after leaving h and whose weight is at least as that of γ_0. Repeating this process sequentially for each recurrent node, we receive a 1-cyclic walk γ with weight at least as that of γ_0. Since γ cannot have a strictly larger weight, we are done. □

We are now ready to prove:

Theorem 3.4. *Suppose that* $\langle u,v\rangle \in \mathrm{Id}(\mathcal{U}_n)$, *with* $u,v \in \{a,b\}^+$, *and that* $A,B \in \mathcal{M}_n$ *satisfy*

$$\mathrm{per}(A) = \mathrm{tr}(A) \quad \textit{and} \quad \mathrm{per}(B) = \mathrm{tr}(B); \tag{3.4}$$

$$\mathrm{rk}_{\mathrm{tr}}(u\,[\![A,B\,]\!]) = \mathrm{rk}_{\mathrm{tr}}(v\,[\![A,B\,]\!]) = n. \tag{3.5}$$

Then, $u\,[\![A,B\,]\!] = v\,[\![A,B\,]\!]$.

Proof. We prove that the following inequality holds for any entry (i,j):

$$(u\,[\![A,B\,]\!])_{i,j} \leqslant (v\,[\![A,B\,]\!])_{i,j}\,. \tag{3.6}$$

The case of $(u\,[\![A,B\,]\!])_{i,j} = \mathbb{0}$ is obvious. Otherwise, Lemma 3.3 gives a 1-cyclic walk $\gamma_{i,j}$, i.e., $\gamma_{i,j}$ never returns to a node which it has already left. Thus, the nodes of $\mathrm{G}(A,B)$ can be permuted, say by $\pi \in S_n$, in a way that $\gamma_{i,j}$ has only arcs that go forward. Let $P := P_\pi$ be the matrix associated to π, and let T_A and T_B be the upper triangular matrices obtained respectively from $P^{-1}AP$ and $P^{-1}BP$ by setting all entries below the diagonal to $\mathbb{0}$. Then the matrix $P^{-1}AP$ satisfies

$$\big(u\,\big[\![P^{-1}AP, P^{-1}BP\,\big]\!]\,\big)_{\pi(i),\pi(j)} = \big(u\,[\![T_A,T_B\,]\!]\,\big)_{\pi(i),\pi(j)},$$

and we can compute

$$\begin{aligned} (u\,[\![A,B\,]\!])_{i,j} &= \big(u\,\big[\![P^{-1}AP, P^{-1}BP\big]\!]\big)_{\pi(i),\pi(j)} \\ &= \big(u\,[\![T_A,T_B\,]\!]\,\big)_{\pi(i),\pi(j)} = \big(v\,[\![T_A,T_B\,]\!]\,\big)_{\pi(i),\pi(j)} \\ &\leqslant \big(v\,\big[\![P^{-1}AP, P^{-1}BP\big]\!]\big)_{\pi(i),\pi(j)} = (v\,[\![A,B\,]\!])_{i,j}\,. \end{aligned}$$

Thus (3.6) holds for each entry (i,j). The reverse inequality holds by symmetry, so that $u\,[\![A,B]\!] = v\,[\![A,B]\!]$. □

To apply Theorem 3.3, matrices which satisfy (3.4) should be detected; this is done by Lemma 2.1.

Remark 3.5. *Lemma 2.1, and consequently Theorem 3.4, also hold if the maximality condition of tropical rank is replaced by maximality of the so-called determinantal rank, which is larger. As well, modifying the notion of nonsingularity accordingly, Theorem 1.4 holds, cf.* [31, Theorem 2]. *Nevertheless, Theorem 2.6 holds for tropical rank, which suffices our needs.*

We can finally prove our main result:

Theorem 3.6. *Given $n \in \mathbb{N}$, let $\overline{n} = \mathrm{lcm}(1,\dots,n)$. For any $t \geqslant (n-1)^2+1$ and every $\langle u,v\rangle \in \mathrm{Id}(\mathcal{M}_{n-1})$, where $u,v,p,\hat{p},q,\hat{q},r,\hat{r} \in \{a,b\}^+$, the following hold:*

(i) If $\langle q,r\rangle \in \mathrm{Id}(\mathcal{U}_n)$, then

$$\langle ua,va\rangle\left[\left((qr)^t\right)\left[a^{\overline{n}},b^{\overline{n}}\right],\left((qr)^t r\right)\left[a^{\overline{n}},b^{\overline{n}}\right]\right] \in \mathrm{Id}(\mathcal{M}_n); \tag{3.7}$$

(ii) If $(p\hat{q}p, p\hat{r}p) \in \mathrm{Id}(\mathcal{U}_n)$, then

$$\langle ua,va\rangle\left[(w\hat{q}p)\left[a^{\overline{n}},b^{\overline{n}}\right],(w\hat{r}p)\left[a^{\overline{n}},b^{\overline{n}}\right]\right] \in \mathrm{Id}(\mathcal{M}_n) \tag{3.8}$$

with $w = (p\hat{q}p\hat{r}p)^t$.

Proof. (i): Let $A,B \in \mathcal{M}_n$, and let

$$X = \left((qr)^t\right)[\![A^{\overline{n}},B^{\overline{n}}]\!], \quad Y = \left((qr)^t r\right)[\![A^{\overline{n}},B^{\overline{n}}]\!] = XR, \quad \text{with } R = r\,[\![A^{\overline{n}},B^{\overline{n}}]\!],$$

be matrices in $\mathcal{M}_n$.

- If $\mathrm{rk}_{\mathrm{fc}}(X) < n$, then $X = PQ$ for some matrices $P \in \mathcal{M}_{n,n-1}$ and $Q \in \mathcal{M}_{n-1,n}$. (Add columns and rows of $\mathbb{0}$, if $\mathrm{rk}_{\mathrm{fc}}(X) < n-1$.) Hence $QP, QRP \in \mathcal{M}_{n-1}$, and using Lemma 3.1 we obtain

  $$(ua)\,[\![X,Y]\!] = P\big(u\,[\![QP,QRP]\!]\big)Q = P\big(v\,[\![QP,QRP]\!]\big)Q = (va)\,[\![X,Y]\!],$$

  since $\langle u,v\rangle \in \mathrm{Id}(\mathcal{M}_{n-1})$ by assumption. Therefore,

  $$(ua)\,[\![X,Y]\!] = (va)\,[\![X,Y]\!]. \tag{3.9}$$

- If $\mathrm{rk}_{\mathrm{fc}}(X) = n$, then $\mathrm{rk}_{\mathrm{tr}}\left((qr)\,[\![A^{\overline{n}},B^{\overline{n}}]\!]\right) = n$ by Theorem 2.20, implying that

  $$\mathrm{rk}_{\mathrm{tr}}\left(q\,[\![A^{\overline{n}},B^{\overline{n}}]\!]\right) = \mathrm{rk}_{\mathrm{tr}}\left(r\,[\![A^{\overline{n}},B^{\overline{n}}]\!]\right) = n$$

  by Proposition 1.4. Then $q\,[\![A^{\overline{n}},B^{\overline{n}}]\!] = r\,[\![A^{\overline{n}},B^{\overline{n}}]\!]$ by Theorem 3.4, since $\langle q,r\rangle \in \mathrm{Id}(\mathcal{U}_n)$. Thus X and Y are both powers of $q\,[\![A^{\overline{n}},B^{\overline{n}}]\!]$, and hence commute. This implies (3.9), since $\#_a(u) = \#_a(v)$ and $\#_b(u) = \#_b(v)$.

Therefore, (3.9) holds for any $A,B \in \mathcal{M}_n$, which means that (3.7) holds true.

(ii): The proof of (3.8) follows along the same lines of (i). □

Replacing Theorem 2.20 in the proof by Proposition 2.6, a similar result is obtained, but with longer identities, in which t is exchanged by $\overline{n}t$ and $t \geqslant (n-1)^2+1$ by $t \geqslant 3n-2$. Consequentially, by this change, only a subset of identities is produced.

Theorem 3.7. *The monoid $\mathcal{M}_n$ satisfies a nontrivial semigroup identity for every $n \in \mathbb{N}$. The length of this identity grows with n as $e^{Cn^2+o(n^2)}$ for some $C \leqslant 1/2+\ln(2)$.*

Proof. The case of $n=1$ is trivial, while [19, Theorem 3.9] proves the case of $n=2$. For tropical triangular matrices there exists an identity $\langle q,r\rangle \in \mathrm{Id}(\mathcal{U}_n)$ by [14, Theorem 4.10], or [27, Theorem 0.1]. The proof then easily follows from Theorem 3.6 by induction. To bound the length, note that $\langle q,r\rangle \in \mathrm{Id}(\mathcal{U}_n)$ given by [14, Theorem 4.10] has length $\ell(q) = \ell(r) = 2^{n+o(n)}$, while $\overline{n} = e^{n+o(n)}$ – a fact that follows from the Prime Number Theorem. □

Remark 3.8. *Decreasing the length of $\langle q,r\rangle \in \mathrm{Id}(\mathcal{U}_n)$ would lead to a better bound on C. Yet, with this method whose formulas includes $\overline{n}$, such bound cannot be lower than $1/2$.*

We immediately conclude the following:

Corollary 3.9. *Any semigroup which is faithfully represented by $\mathcal{M}_n$ satisfies a nontrivial identity.*

Example 3.10. *Set* $p = a^2b^2a[ab, ba]$*. Then* $\langle pabp, pbap\rangle \in \mathrm{Id}(\mathcal{U}_3)$ *by* [15]*, while* $\mathcal{M}_2$ *satisfies an identity* $\langle u, v\rangle$ *of length 17 by* [7]*. Thus, by Theorem 3.6.(ii),* $\mathcal{M}_3$ *satisfies an identity of length 19,656, while Theorem 3.6.(i) gives a length 24,816. In* [31] *Shitov pointed out that a matrix* $A \in \mathcal{M}_3$ *has either determinantal rank* 3 *or factor rank at most* 2*. Consequently, for* $\mathcal{M}_3$*, Remark 3.5 allows to omit exponent* t *in Theorem 3.6, which reduces the identity length to 4,968 and 5,808, respectively.* [31] *provides identities of length 1,795,308.*

References

[1] M. Akian, R. Bapat, S. Gaubert. Max-plus algebra, In: L. Hogben, R. Brualdi, A. Greenbaum, R. Mathias (eds.) *Handbook of Linear Algebra*. Chapman and Hall, London, 2006.

[2] M. Akian, S. Gaubert, A. Guterman. Linear independence over tropical semirings and beyond. In *Tropical and idempotent mathematics*, vol. 495 of *Contemp. Math.*, pp 1–38. Amer. Math. Soc., Providence, RI, 2009.

[3] F. d'Alessandro, E. Pasku. A combinatorial property for semigroups of matrices, *Semigroup Forum*, 67(1):22–30, 2003.

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

[5] Y. Chen, X. Hu, Y. Luo, O. Sapir. The finite basis problem for the monoid of 2×2 upper triangular tropical matrices, *Bulletin Australian Math. Society*, 94(1):54–64, 2016.

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

[7] L. Daviaud, M. Johnson. The shortest identities for max-plus automata with two states, In: K. G. Larsen, H. L. Bodlaender, J.-F. Raskin (eds.) *MFCS 2017*, Art. 48:1–13, 2017.

[8] E. Demaine, S. Eisenstat, J. Shallit, D. Wilson. Remarks on separating words, *DCFS 2011*, LNCS 6008, 147–157, 2011.

[9] M. Develin, F. Santos, B. Sturmfels. On the rank of a tropical matrix. In *Combinatorial and computational geometry*, vol. 52 of *Math. Sci. Res. Inst. Publ.*, pp 213–242, Cambridge Univ. Press, Cambridge, 2005.

[10] M. Droste, W. Kuich, W. Vogler (eds.), *Handbook of weighted automata*, Springer-Verlag, 2009.

[11] P. Guillon, Z. Izhakian, J. Mairesse, G. Merlet. The ultimate rank of tropical matrices. *J. Alg.*, 437:222–248, 2015.

[12] M. Gromov. Groups of polynomial growth and expanding maps. *Publ. Math IHES*, 53:53–73, 1985.

[13] Z. Izhakian. Tropical arithmetic and matrix algebra. *Commun. Alg.*, 37(4):1445–1468, 2009.

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

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

[16] Z. Izhakian. Tropical plactic algebra, the cloaktic monoid, and semigroup representations, arXiv:1701.05156, 2017.

[17] Z. Izhakian. Semigroup identities of tropical matrix semigroups of maximal rank, *Semigroup Forum*, 92(3):712–732, 2016.

[18] Z. Izhakain, M. Johnson, M. Kambites. Tropical matrix groups, *Semigroup Forum*, 96(1):178–196, 2018.

[19] Z. Izhakian, S. Margolis. Semigroup identities in the monoid of 2-by-2 tropical matrices. *Semigroup Furom*, 80(2):191–218, 2010.

[20] Z. Izhakian, L. Rowen. The tropical rank of a tropical matrix. *Commun. Alg.*, 37(11):3912–3927, 2009.

[21] Z. Izhakian, L. Rowen. Supertropical matrix algebra. *Israel J. Math.*, 182(1):383–424, 2011.

[22] Z. Izhakian, L. Rowen. Supertropical matrix algebra II: Solving tropical equations. *Israel J. Math.*, 186(1):69–97, 2011.

[23] Z. Izhakian, L. Rowen. Supertropical matrix algebra III: Powers of matrices and generalized eigenspaces. *J. Alg.*, 341(1):125–149, 2011.

[24] B. Heidergott, G. J. Olsder, J. van der Woude. *Max Plus at Work: Modeling and analysis of synchronized systems: A course on max-plus algebra and its applications*, Princeton University Press, 2006

[25] G. Merlet. Semigroup of matrices acting on the max-plus projective space. *Linear Alg. Appl.*, 432(8):1923–1935, 2010.

[26] G. Merlet, T. Nowak, S. Sergeev. Weak CSR expansions and transience bounds in max-plus algebra. *Linear Alg. Appl.*, 461:163–199, 2014.

[27] J.Okninski. Identities of the semigroup of upper triangular tropical matrices, *Commun. Alg.*, 43(10):4422–4426, 2015.

[28] J.-E. Pin. Tropical semirings. *Cambridge Univ. Press, Cambridge*, 11:50–69, 1998.

[29] J.-E. Pin. *Mathematical foundations of automata theory*, ver. Nov. 30, 2016.

[30] J.-M. Robson. Separating strings with small automata, *Inform. Process. Lett.*, 30:209–214, 1989.

[31] Y. Shitov. A semigroup identity for tropical 3x3 matrices, arXiv:1406.2601, 2014.

[32] L. Shneerson. Identities in finitely generated semigroups of polynomial growth. *J. Alg.,* 154(1):67–85, 1993.

[33] I. Simon. Recognizable sets with multiplicities in the tropical semiring, in MFCS 88, M. Chytil, L. Janiga, V. Koubek (eds.), *Lecture Notes in Computer Science*, Springer, 324:107–120, 1988.

[34] L. N. Shevrin, M. V. Volkov. Identities of semigroups, *Izv. VUZ. Matematika* (Russian), 11:3–47, 1985; *Soviet Math. Iz. Vuz. 29* (English translation), 11:1–64, 1985.

Institute of Mathematics, University of Aberdeen, AB24 3UE, Aberdeen, UK.
E-mail address: zzur@abdn.ac.uk

Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
E-mail address: glenn.merlet@univ-amu.fr