ELA

http://math.technion.ac.il/iic/ela

ON THE MARKOV CHAIN TREE THEOREM
IN THE MAX ALGEBRA[*]

BUKET BENEK GURSOY[†], STEVE KIRKLAND[‡], OLIVER MASON[†], AND SERGEĬ
SERGEEV[§]

Abstract. A max-algebraic analogue of the Markov Chain Tree Theorem is presented, and its connections with the classical Markov Chain Tree Theorem and the max-algebraic spectral theory are investigated.

Key words. Markov chains, Stochastic matrices, Directed spanning trees, Max algebra, Kleene star.

AMS subject classifications. 15A80, 60J10, 68R10, 15B51, 05C05.

1. Introduction. The Markov Chain Tree Theorem is a well-known result that relates the stationary distribution of an irreducible Markov chain with the weights of directed spanning trees in its associated digraph. Before recalling this result in Theorem 1.1, we introduce some necessary notation and terminology.

Let $D(A)$ denote the weighted directed graph associated with an irreducible matrix $A \in \mathbb{R}_+^{n \times n}$. $D(A)$ consists of the nodes $V = \{1, \ldots, n\}$ with a directed edge (i, j) from i to j of weight a_{ij} if and only if $a_{ij} > 0$. We say the edge $e = (i, j)$ is *outgoing* from i and write $t(e) = i$. A spanning subgraph $T = (V, E)$ of $D(A)$ is said to be an *i-tree* if the following conditions are satisfied:

(i) for every $j \neq i$ in $\{1, \ldots, n\}$, there is exactly one outgoing edge $e \in E$ whose beginning node is j;

(ii) there is no edge $e \in E$ whose beginning node is i;

(iii) the subgraph (V, E) contains no directed cycle.

ELA

Note that an i-tree is often referred to as a *rooted directed spanning tree (RST)* at $i \in \{1, \ldots, n\}$.

Given an i-tree T in $D(A)$, the *weight* of T is given by the product of the weights of the edges in T and is denoted by $\pi(T, A)$ or just by $\pi(T)$ when A is clear from the context. Let $\mathcal{T}_i$ denote the set of all i-trees of $D(A)$ for $1 \leq i \leq n$. The classical Markov Chain Tree Theorem, also known as the Freĭdlin-Wentzell formula [14, 23], can be stated as follows.

THEOREM 1.1. *Let $A \in \mathbb{R}_+^{n \times n}$ be an irreducible (row) stochastic matrix. Define $w \in \mathbb{R}_+^n$ by*

$$w_i = \sum_{T \in \mathcal{T}_i} \pi(T).$$

Then, $A^T w = w$. In particular, $\frac{w}{\sum_{i=1}^n w_i}$ is the unique stationary distribution of the Markov chain with transition matrix A.

This core result has appeared in a variety of different contexts [1, 2, 6, 14, 23, 24]. It was discovered by Shubert [22] in connection with flow-graph methods, and independently by Kohler and Vollmerhaus [15] motivated by problems in biological modelling. For another reference, which discusses its extension to general, not necessarily irreducible Markov chains, see Leighton and Rivest [16].

The primary contribution of this paper is to extend the Markov Chain Tree Theorem to the setting of the max algebra. We show this in two ways: first, we prove a max-algebraic version directly using combinatorial arguments; we then provide an alternative proof using dequantisation. We also relate the max version of the Markov Chain Tree Theorem with max-algebraic spectral theory. In keeping with Bapat [4], the max algebra $\mathbb{R}_+(\max)$ consists of the nonnegative real numbers equipped with the two operations $a \oplus b = \max(a, b)$ and $a \otimes b = ab$. These operations extend to nonnegative matrices and vectors in the standard way [3, 4, 7, 9].

2. Markov Chain Tree Theorem. In this section, we first show that Theorem 1.1 extends to the max algebra. We then provide an alternative proof of this result using dequantisation.

ELA

http://math.technion.ac.il/iic/ela

2.1. Markov Chain Tree Theorem in the max algebra. Let us first recall standard observations on graphs and spanning trees.

LEMMA 2.1. *Let D be a digraph and i be a node of D, to which every other node can be connected by a path. Then D contains an i-tree.*

COROLLARY 2.2. *If $A \in \mathbb{R}_+^{n \times n}$ is irreducible then for each node $i \in \{1, \ldots, n\}$ there exists an i-tree in $D(A)$ with nonzero weight.*

LEMMA 2.3. *Let D be a digraph, i be a node of D and T be an i-tree. Then for each node $j \neq i$ of D, there exists a unique directed path from j to i in T.*

We now consider an irreducible matrix A in $\mathbb{R}_+^{n \times n}$ which is row stochastic in a max-algebraic sense. Formally, we assume that $\max_{1 \leq j \leq n} a_{ij} = 1$ for $1 \leq i \leq n$ or using max-algebraic notation

$$A \otimes \mathbf{1} = \mathbf{1},$$

where $\mathbf{1}$ denotes the vector all of whose entries are equal to one. In a convenient abuse of notation, we refer to matrices satisfying the above condition as *max-stochastic*. We now present the main result of this section.

THEOREM 2.4. *Let $A \in \mathbb{R}_+^{n \times n}$ be an irreducible max-stochastic matrix. Define the vector w by*

$$w_i = \bigoplus_{T \in \mathcal{T}_i} \pi(T), \quad 1 \leq i \leq n. \tag{2.1}$$

Then

$$A^T \otimes w = w.$$

Proof. We first show that $A^T \otimes w \leq w$. To this end, let an arbitrary $i \in V = \{1, \ldots, n\}$ be given. Then, as $a_{ii} \leq 1$, it is immediate that $a_{ii} w_i \leq w_i$. Now consider $j \neq i$ such that $a_{ji} \neq 0$. Let T_j be a j-tree such that $\pi(T_j) = w_j$, let E_j be the set of edges of T_j and $(i, k) \in E_j$. Consider the set of edges formed by removing (i, k) from E_j and inserting (j, i) instead, and denote it by E_i. Consider the subgraph $T_i = (V, E_i)$. Note that there is exactly one outgoing edge from every $\ell \neq i$ and no outgoing edge from i. Further, T_i is acyclic as any cycle in T_i must contain the edge (j, i) (otherwise it would define a cycle in the original j-tree T_j); however there is no outgoing edge from i in T_i. It follows that the graph T_i is an i-tree. By construction and since all entries of a max-stochastic matrix are not greater than 1, we obtain that

$$w_i \geq \pi(T_i) = \pi(T_j) a_{ji} / a_{ik} \geq w_j a_{ji},$$

and since we were given an arbitrary i and took an arbitrary j such that $a_{ji} \neq 0$, it follows that $A^T \otimes w \leq w$.

ELA

http://math.technion.ac.il/iic/ela

To complete the proof, we show that $A^T \otimes w \geq w$. Let an arbitrary $i \in \{1, \ldots, n\}$ be given, and let $T_i = (V, E_i)$ be an i-tree such that $\pi(T_i) = w_i$. As A is a max-stochastic matrix by assumption, we know that $a_{ik} = 1$ for some k. If $k = i$ then $(A^T \otimes w)_i \geq a_{ii} w_i = w_i$. So let $k \neq i$. To show that $\bigoplus_{j=1}^{n} w_j a_{ji} \geq w_i$ we will construct a j-tree T_j such that $\pi(T_j) a_{ji} = \pi(T_i)$. Consider a path connecting k to i in T_i. By Lemma 2.3 this path is unique. Let j be the penultimate node on this path, meaning that $(j, i) \in E_i$. Removing the edge (j, i) from E_i and inserting the edge (i, k) we obtain the edge set E_j and the required j-tree $T_j = (V, E_j)$. Indeed, there is exactly one outgoing edge from each node other than j in T_j, and there is no outgoing edge from j. Furthermore, if there exists a cycle in T_j, it must contain the edge (i, k) as otherwise it would define a cycle in T_i. This would then imply that there exists a directed path in T_j from k to i, all of whose edges are also edges in T_i. This is impossible however, as the only such path in T_i contains the edge (j, i) which is not an edge in T_j. Therefore, T_j is indeed a j-tree, which satisfies $\pi(T_j) a_{ji} = \pi(T_i)$ by construction (as $a_{ik} = 1$). Hence, $\bigoplus_{j=1}^{n} w_j a_{ji} \geq w_i$ and $A^T \otimes w \geq w$, as i was arbitrary. The proof is complete. $\square$

The vector w defined in (2.1) will be called the maximal i-tree weight (ITW) vector of A.

2.2. Proof by dequantisation. In this subsection, we present an alternative proof of Theorem 2.4 using a procedure that can be seen as an instance of the Maslov dequantisation [17]. Note that the same procedure was used by Olsder and Roos [18] to derive max-algebraic analogues of the Cramer and Cayley-Hamilton formulae.

Consider the set of nonnegative numbers $\mathbb{R}_+$ equipped with the operations $a +_p b := (a^p + b^p)^{1/p}$ and $a \times_p b := ab$ for $1 \leq p < \infty$. This is a semiring isomorphic to the semiring of nonnegative numbers with the usual arithmetic, via the mapping $f(a) := a^{1/p}$. We denote by $\mathbb{R}_+(p)$ the semiring of nonnegative real numbers equipped with the operations $+_p$, $\times_p$ defined above. We say that $A \in \mathbb{R}_+^{n \times n}$ is p-stochastic if $a_{i1} +_p a_{i2} +_p \cdots +_p a_{in} = 1$ for $1 \leq i \leq n$.

The ITW vector of $A \in \mathbb{R}_+^{n \times n}$ defined as in Theorem 1.1 using the arithmetic of $\mathbb{R}_+(p)$ will be denoted by $w^{(p)}(A)$, and when defined in $\mathbb{R}_+(\max)$ (i.e., the maximal ITW vector), by $w^{\max}(A)$.

THEOREM 2.5. *Let* $A \in \mathbb{R}_+^{n \times n}$ *be max-stochastic. There exists an integer* P_0 *and a sequence* $A^{(p)}, p \geq P_0$ *in* $\mathbb{R}_+^{n \times n}$, *where each* $A^{(p)}$ *is* p-stochastic, *such that* $\lim_{p \to \infty} A^{(p)} = A$ *and* $\lim_{p \to \infty} w^{(p)}(A^{(p)}) = w^{\max}(A)$.

$$\boxed{\textbf{ELA}}$$

Proof. Let $B^\delta(A)$ denote the set of matrices C such that $|c_{ij} - a_{ij}| \le \delta$ for all i, j and such that $c_{ij} > 0$ if and only if $a_{ij} > 0$. We start by constructing a nondecreasing sequence of p-stochastic matrices $A^{(p)} \in B^\delta(A)$.

As A is max-stochastic, for each $i \in \{1, \ldots, n\}$ there are l_i entries $a_{ij} = 1$, where $0 < l_i \le n$. We denote the set of other entries in each row by $J_i := \{j \mid a_{ij} < 1\}$ for each i. Denoting $m_i = \max\{a_{ij} \mid j \in J_i\}$ choose P_0 so that $1 - m_i^p(n - l_i) \ge 0$ for all $p \ge P_0$. Then also $1 - \sum_{j \in J_i} a_{ij}^p \ge 0$ for all $p \ge P_0$. For $p \ge P_0$, define $A^{(p)}$ by

$$a_{ij}^{(p)} = \begin{cases} a_{ij}, & \text{if } a_{ij} < 1, \\ \delta_i^{(p)}, & \text{otherwise}, \end{cases}$$

where

$$\delta_i^{(p)} = \left(\frac{1 - \sum_{j \in J_i} a_{ij}^p}{l_i} \right)^{1/p}.$$

It is readily verified that $A^{(p)}$ is p-stochastic and that $a_{ij}^p \le a_{ij}$ for all i, j. We obtain

$$a_{ik} - a_{ik}^{(p)} \le 1 - \delta_i^{(p)} \le 1 - \left(\frac{1 - m_i^p(n - l_i)}{l_i} \right)^{1/p} \tag{2.2}$$

for all i and k. As $m_i < 1$ for all i, it follows that the right hand side of (2.2) converges to 0 as p tends to infinity. Hence, $A^{(p)}$ converges to A.

Next, note that

$$|w_i^{(p)}(A^{(p)}) - w_i^{\max}(A)| = |w_i^{(p)}(A^{(p)}) - w_i^{\max}(A^{(p)}) + w_i^{\max}(A^{(p)}) - w_i^{\max}(A)| \tag{2.3}$$
$$\le |w_i^{(p)}(A^{(p)}) - w_i^{\max}(A^{(p)})| + |w_i^{\max}(A^{(p)}) - w_i^{\max}(A)|.$$

Since $w_i^{\max}(A^{(p)}) = \max_{T \in \mathcal{T}_i} \pi(T, A^{(p)})$, we see that

$$w_i^{(p)}(A^{(p)}) - w_i^{\max}(A^{(p)}) \le (M_i^{1/p} - 1) \max_{T \in \mathcal{T}_i} \pi(T, A^{(p)}) \le M_i^{1/p} - 1, \tag{2.4}$$

where M_i is the number of i-trees in A (or $A^{(p)}$).

It is obvious from the definition of $A^{(p)}$ that $w_i^{\max}(A) \ge w_i^{\max}(A^{(p)})$. Let T' be an i-tree such that $w_i^{\max}(A) = \max_{T \in \mathcal{T}_i} \pi(T, A) = \pi(T', A)$. It follows that $w_i^{\max}(A^{(p)}) \ge \pi(T', A^{(p)})$. Then

$$|w_i^{\max}(A^{(p)}) - w_i^{\max}(A)| = w_i^{\max}(A) - w_i^{\max}(A^{(p)}) \le \pi(T', A) - \pi(T', A^{(p)}). \tag{2.5}$$

Let $E' = \{(i_1, j_1), (i_2, j_2), \ldots, (i_{n-1}, j_{n-1})\}$ be the edges in the i-tree T'. Then, it follows from (2.5) that

$$|w_i^{\max}(A^{(p)}) - w_i^{\max}(A)| \le a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_{n-1} j_{n-1}} - a_{i_1 j_1}^{(p)} a_{i_2 j_2}^{(p)} \cdots a_{i_{n-1} j_{n-1}}^{(p)} \tag{2.6}$$

$$\leq C(A) \max_{i,j}(a_{ij} - a_{ij}^{(p)}),$$

where $C(A)$ is a fixed constant that depends only on the entries of A.

Using (2.4) and (2.6) in (2.3), we obtain

$$|w_i^{(p)}(A^{(p)}) - w_i^{\max}(A)| \leq M_i^{1/p} - 1 + C(A) \max_{i,j}(a_{ij} - a_{ij}^{(p)}).$$

As we showed above that $\max(a_{ij} - a_{ij}^{(p)}) \to 0$ as $p \to \infty$ and $M_i^{1/p} \to 1$ as $p \to \infty$, the claim follows. $\square$

As each of the semirings $\mathbb{R}_+(p)$ is isomorphic to the nonnegative real numbers with the usual operations, it follows from the classical Markov Chain Tree Theorem (Theorem 1.1) that $(A^{(p)})^T \times_p w^{(p)} = w^{(p)}$ for all $p \geq P_0$. Passing to the limit and applying Theorem 2.5 yields another proof of Theorem 2.4.

We next present a numerical example to illustrate Theorem 2.4.

EXAMPLE 2.6.

$$A = \begin{bmatrix} 1 & 3/4 & 5/6 & 0 \\ 1/2 & 1 & 1/4 & 9/10 \\ 0 & 0 & 1 & 7/8 \\ 1/3 & 0 & 1 & 4/5 \end{bmatrix} \tag{2.7}$$
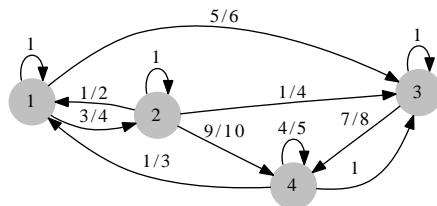

FIG. 2.1. $D(A)$ for (2.7).

Let T_i be an i-tree with maximal weight for $i = 1, 2, 3, 4$. Then,

- $T_1 : (3,4), (2,4), (4,1)$ $w_1 = \pi(T_1) = a_{34}a_{24}a_{41} = 21/80$;
- $T_2 : (3,4), (4,1), (1,2)$ $w_2 = \pi(T_2) = a_{34}a_{41}a_{12} = 7/32$;
- $T_3 : (1,3), (2,4), (4,3)$ $w_3 = \pi(T_3) = a_{13}a_{24}a_{43} = 3/4$;
- $T_4 : (2,4), (1,3), (3,4)$ $w_4 = \pi(T_4) = a_{24}a_{13}a_{34} = 21/32$.

Hence, $w = \begin{bmatrix} 21/80 \\ 7/32 \\ 3/4 \\ 21/32 \end{bmatrix}$ and $A^T \otimes w = w$.

ELA

http://math.technion.ac.il/iic/ela

As a final point for this section, we note the following consequence of Theorem 2.4 for nonnegative irreducible matrices that are not necessarily max-stochastic.

PROPOSITION 2.7. *Let $A \in \mathbb{R}_+^{n \times n}$ be irreducible and assume that for all $i \in \{1, \ldots, n\}$, there exists some j with $a_{ij} > 0$. Consider the diagonal matrix X given by*

$$x_{ii} = \max_{1 \leq j \leq n} a_{ij}.$$

Further let w be the maximal ITW vector for A. Then

$$A^T \otimes w = Xw.$$

Proof. Let $\hat{A} = X^{-1}A$. Then $\hat{A}$ is irreducible and max-stochastic. For $1 \leq i \leq n$, consider a spanning tree $\hat{T}$ in $D(\hat{A})$ rooted at i. It is clear that the weight of $\hat{T}$ takes the form

$$\hat{a}_{i_1 j_1} \cdots \hat{a}_{i_{n-1} j_{n-1}} = \frac{1}{x_{i_1 i_1} x_{i_2 i_2} \cdots x_{i_{n-1} i_{n-1}}} a_{i_1 j_1} \cdots a_{i_{n-1} j_{n-1}},$$

where $\{i_1, \ldots, i_{n-1}\} = \{1, \ldots, n\} \backslash \{i\}$. In fact, it is clear that there is a bijective correspondence between spanning trees T_i in $D(A)$ rooted at i and spanning trees $\hat{T}_i$ rooted at i in $D(\hat{A})$ with

$$\pi(\hat{T}_i) = \frac{x_{ii}}{x_{11} \cdots x_{nn}} \pi(T_i).$$

It follows that if we write $\hat{w}$ for the maximal ITW vector of $\hat{A}$, then

$$\hat{w} = \frac{X}{x_{11} \cdots x_{nn}} w. \tag{2.8}$$

As $\hat{A}$ is max-stochastic, we know from Theorem 2.4 that $\hat{A}^T \otimes \hat{w} = \hat{w}$. Noting that $\hat{A}^T = A^T X^{-1}$, we can use (2.8) to rewrite this as

$$\frac{1}{x_{11} \cdots x_{nn}} A^T \otimes w = \frac{X}{x_{11} \cdots x_{nn}} w.$$

The result follows immediately. ☐

The above result characterises the maximal ITW vector for a broader class of nonnegative matrices in max-algebraic terms. In particular, it applies to *symmetrical reciprocal* matrices, which arise in the *Analytic Hierarchy Process*: a widely used framework for multi-criteria decision making [20]. In the spirit of [5, 12, 13], it would be interesting to investigate the possibility of using the maximal ITW vector in this and related applications. We hope to report on this in more detail in future work.

ELA

http://math.technion.ac.il/iic/ela

3. Maximal ITW vectors and Kleene stars. We have seen that the maximal ITW vector w associated with the directed graph $D(A)$ is always a left max eigenvector of an irreducible max-stochastic matrix A. However, in contrast to the conventional algebra, the irreducibility of A is not sufficient to guarantee uniqueness (up to scalar multiple) of the max eigenvector. This naturally leads to the question of how to identify the maximal ITW vector using the tools of max spectral theory such as the power method or Kleene star. We next consider this question.

First, recall that for $A \in \mathbb{R}_+^{n \times n}$ with $\mu(A) \leq 1$ the series $I \oplus A \oplus A_\otimes^2 \oplus \cdots \oplus A_\otimes^n \oplus \cdots$ converges to a finite matrix called the Kleene star of A given by

$$A^* = I \oplus A \oplus A_\otimes^2 \oplus \cdots \oplus A_\otimes^{n-1},$$

where $\mu(A) \leq 1$ [3, 7, 8, 9, 19]. Here, $A_\otimes^k$ denotes the kth max-algebraic power of A and a_{ij}^* is the maximum weight of a path from i to j of any length in $D(A)$ (if $i \neq j$). In particular if A is irreducible, then A^* is positive [3, 9].

A cycle with the maximum cycle geometric mean is called a *critical cycle* [3, 4, 7, 8, 9, 19]. The set of nodes that lie on some critical cycle are said to be *critical nodes* and denoted by $N^C(A)$. The set of edges belonging to critical cycles are said to be *critical edges* and denoted by $E^C(A)$. The critical matrix of A [10, 11], A^C, is formed from the submatrix of A consisting of the rows and columns corresponding to critical nodes as follows. Set $a_{ij}^C = a_{ij}$ if (i, j) lies on a critical cycle and $a_{ij}^C = 0$ otherwise. Moreover, we use the notation $D^C(A)$ for the critical graph of A, the digraph which consists of all critical nodes and edges.

The following well-known result shows the connection of A^* with the max eigenvectors of A [3, 4, 11]. We adopt the notation $A_{i\cdot}^*$ for the ith row, and the notation $A_{\cdot i}^*$ for the ith column of the matrix A^*.

PROPOSITION 3.1. *Let $A \in \mathbb{R}_+^{n \times n}$ be an irreducible matrix with $\mu(A) = 1$. Assume that $D^C(A)$ has r strongly connected components. Then, the following are true.*

(i) *$\mu(A) = 1$ is the only max eigenvalue of A;*
(ii) *$A_{\cdot i}^*$ is a (right) max eigenvector associated with $\mu(A)$ for $i \in N^C(A)$;*
(iii) *For $i, j \in N^C(A)$ ($i \neq j$), $A_{\cdot i}^*$ and $A_{\cdot j}^*$ are scalar multiples of each other if they belong the same strongly connected component in $D^C(A)$.*

If one takes r columns of A^* from different strongly connected components of $D^C(A)$, then none of them can be expressed as a max-linear combination of the other columns. Moreover, any such set is strongly linear independent in the sense of [7]. For general (reducible) matrices, $\mu(A)$ is the largest eigenvalue of A.

A max-stochastic matrix has max eigenvalue 1, and $a_{ij} \leq 1$ for all i, j. This implies that $\mu(A) = 1$, and that $a_{ij} = 1$ for $(i, j) \in E^C(A)$. Such matrices are

said to be *visualised* [21]. Note that the max-stochastic matrices have an additional property: each node in $D(A)$ has an outgoing edge with weight 1. The spanning subgraph of $D(A)$ consisting of the edges of weight 1 defines the saturation digraph, denoted $\mathrm{Sat}(A)$.

Observe that for any matrix A with a positive eigenvector x, the matrix $B = X^{-1}AX$, where X is a diagonal matrix formed from x, is max-stochastic. An analogous property holds in nonnegative algebra, where it has many applications, and one can consider a generalisation to semifields (i.e., semirings with invertible multiplication). Thus, a max-stochastic matrix can be considered to be "eigenvector-visualised".

The Kleene star of a visualised matrix with $\mu(A) = 1$ (and hence of a max-stochastic one) has a very specific structure, as described, for example, in Proposition 4.1 of [21], which we now recall. Define $D^{C*}(A)$ to be the directed graph consisting of the nodes $\{1, \ldots, n\}$, all critical edges and the loops (i, i) for $1 \le i \le n$. We assume that $D^{C*}(A)$ has r' strongly connected components with node sets $N_1, \ldots, N_{r'}$.

Let $A_{\mu\nu}$ denote the submatrix of A formed from the rows with indices in N_μ and from the columns with indices in N_ν for $1 \le \mu, \nu \le r'$. Let A^{red} be the $r' \times r'$ matrix with entries $\alpha_{\mu\nu} = \max\{a_{ij} \mid i \in N_\mu, \, j \in N_\nu\}$, and let $E \in \mathbb{R}_+^{n \times n}$ be the $n \times n$ matrix with all entries equal to 1.

PROPOSITION 3.2 ([21], Proposition 4.1). *Let $A \in \mathbb{R}_+^{n \times n}$ be a visualised matrix, $\mu(A) = 1$ and r' be the number of strongly connected components of $D^{C*}(A)$. Further, let $A^{\mathrm{red}} = (\alpha_{\mu,\nu})$ be as defined above. Then*

1. *$\alpha_{\mu\mu} = 1$ for all $1 \le \mu \le r'$ and $\alpha_{\mu\nu} \le 1$ (resp. $\alpha_{\mu\nu} < 1$ for $\mu \ne \nu$), where $\mu, \nu \in \{1, \ldots, r'\}$;*
2. *for $1 \le \mu, \nu \le r'$, the corresponding submatrix of A^* is $A_{\mu\nu}^* = \alpha_{\mu\nu}^* E_{\mu\nu}$, where $\alpha_{\mu\nu}^*$ is the (μ, ν)-entry of $(A^{\mathrm{red}})^*$, and $E_{\mu\nu}$ is the (μ, ν)-submatrix of E.*

We proceed with the following preliminary result.

LEMMA 3.3. *Let $A \in \mathbb{R}_+^{n \times n}$ be an irreducible max-stochastic matrix. Then, for $1 \le j \le n$, $\min\limits_{1 \le i \le n} a_{ij}^* = \min\limits_{q \in N^C(A)} a_{qj}^*$.*

Proof. Let $j \in \{1, \ldots, n\}$ be given. It is immediate that

$$\min_{1 \le i \le n} a_{ij}^* \le \min_{q \in N^C(A)} a_{qj}^*. \tag{3.1}$$

To show the reverse inequality, consider some $l \notin N^C(A)$. We claim that there exists a path from l to some $k \in N^C(A)$ of weight 1. As A is max-stochastic, there exists at least one outgoing edge from l of weight 1, $a_{lk_1} = 1$. Moreover, as l is not critical, $k_1 \ne l$. If k_1 is critical, we are done. If not, then there exists $k_2 \notin \{l, k_1\}$

ELA

http://math.technion.ac.il/iic/ela

with $a_{k_1 k_2} = 1$. Continuing in this fashion, we must eventually arrive at some node $k = k_p$ which was already on the path. Hence, this node is on a critical cycle, and k is in $N^C(A)$. By construction, $l, k_1, \ldots, k_p = k$ is a path of weight 1, which we denote by P_1.

Next, note that a_{kj}^* is the maximal weight of a path P_2 between k and j. Concatenation of P_1 and P_2 yields the path $P_1 \circ P_2$ with weight a_{kj}^* connecting l to j. It follows that

$$a_{lj}^* \geq a_{kj}^* \geq \min_{q \in N^C(A)} a_{qj}^*.$$

As this must hold for any $l \notin N^C(A)$, we have that

$$\min_{1 \leq i \leq n} a_{ij}^* \geq \min_{q \in N^C(A)} a_{qj}^*. \tag{3.2}$$

Combining (3.2) and (3.1) yields the result. ☐

Recall that for a max-stochastic matrix A, each node has an outgoing edge with weight 1. The spanning subgraph of $D(A)$, which contains the edge (i,j) if and only if $a_{ij} = 1$ is known as the *saturation subgraph* $\text{Sat}(A)$ of A.

LEMMA 3.4. *Let $A \in \mathbb{R}_+^{n \times n}$ be an irreducible max-stochastic matrix. Assume that $D^C(A)$ is strongly connected. Let w be the maximal ITW vector of A. Then for all $i \in N^C(A)$, $w_i = 1$.*

Proof. Evidently $w_i \leq 1$ for all i, since it is obtained by multiplication of the entries of A, all not exceeding 1. Hence, it is enough to show that for a given critical node i, we can construct an i-tree in $\text{Sat}(A)$. In $\text{Sat}(A)$, each node is connected to a critical node. As $D^C(A)$ is strongly connected, the result follows from an application of Lemma 2.1. ☐

In the next result, we denote by y_C the critical subvector of y, i.e., the subvector corresponding to indices in $N^C(A)$.

THEOREM 3.5. *Let $A \in \mathbb{R}_+^{n \times n}$ be an irreducible max-stochastic matrix and w be the maximal ITW vector of A. Then, the following are true.*

(i) $w^T \leq \min_{i \in N^C(A)} A_{i \cdot}^*$;

(ii) *If $D^C(A)$ is strongly connected then $w^T = \min_{i \in N^C(A)} A_{i \cdot}^*$;*

(iii) *If $D^C(A)$ has no more than two components then $w_C^T = (\min_{i \in N^C(A)} A_{i \cdot}^*)_C$.*

Proof. (i): Consider a j-tree T $(1 \leq j \leq n)$, with weight w_j. There exists a path P in T from i to j for $1 \leq i \leq n$, $i \neq j$, with weight $w(P)$. Then,

$$w_j \leq w(P) \leq a_{ij}^* \text{ for all } i \in \{1, 2, \ldots, n\}.$$

Thus, $w_j \leq \min\limits_{1 \leq i \leq n} a_{ij}^*$, or equivalently (by Lemma 3.3), we have $w_j \leq \min\limits_{k \in N^C(A)} a_{kj}^*$.

(ii): In this case, all columns of the Kleene star with indices in $N^C(A)$ are scalar multiples of each other, and any eigenvector, including w, is a multiple of such a column. Let y be a column of the Kleene star with index in $N^C(A)$. By Proposition 3.2, all the components of y_C equal 1, and by Lemma 3.4 all the components of w_C equal 1. Hence, $y = w_C$.

(iii): Let $D^C(A)$ consist of two components, with sets of nodes N_1 and N_2 respectively. By Proposition 3.2, there exist α and β such that $A_{12}^* = \alpha E_{12}$, $A_{21}^* = \beta E_{21}$, $A_{11}^* = E_{11}$ and $A_{22}^* = E_{22}$. Hence, we need to show that $w_i = \beta$ when $i \in N_1$ and $w_i = \alpha$ when $i \in N_2$. We will give a proof only for $i \in N_1$, the other case being similar.

As we showed in part (i) that $w^T \leq \min\limits_{i \in N^C(A)} A_{i\cdot}^*$, it suffices to build an i-tree of weight β, for $i \in N_1$. Consider a path P of greatest weight connecting a node in N_2 to i; by Proposition 3.2 this weight is equal to β. Let k be the first node on P where it leaves N_2 and let l be the first node on P where it enters N_1. By optimality of P, only the subpath P' of P connecting k to l may have weight less than 1, and this weight is β. Using Lemma 2.1, construct a spanning tree directed to i in the first component of $D^C(A)$ (with node set N_1), and a spanning tree directed to k in the second component of $D^C(A)$ (with node set N_2). This makes an i-tree on the graph consisting of nodes and edges of $D^C(A)$ and P'. We need to complete this tree to an i-tree in $D(A)$ having the same weight. We can do this using the edges of $\mathrm{Sat}(A)$, since all remaining nodes of $D(A)$ can be connected by a path with edges in $\mathrm{Sat}(A)$ either to a node of $D^C(A)$ or to a node of P'. The resulting tree is directed to i and has weight β. $\square$

It follows immediately that if all nodes in $D(A)$ are critical and $D^C(A)$ has two strongly connected components, then the maximal ITW vector w is given by $\left(\min\limits_{i \in N^C(A)} A_{i\cdot}^*\right)^T$.

The following example illustrates that point (iii) in Theorem 3.5 does not necessarily hold when $D^C(A)$ consists of three components.

EXAMPLE 3.6. Consider the matrix A given in (2.7). There exist three strongly connected components in $D^C(A)$ and $N^C(A) = \{1, 2, 3\}$ such that $N_1^C(A) = \{1\}$, $N_2^C(A) = \{2\}$ and $N_3^C(A) = \{3\}$. The Kleene star of A is given by

$$A^* = \begin{bmatrix} 1 & 3/4 & 5/6 & 35/48 \\ 1/2 & 1 & 9/10 & 9/10 \\ 7/24 & 7/32 & 1 & 7/8 \\ 1/3 & 1/4 & 1 & 1 \end{bmatrix}.$$

ELA

The left max eigenvectors are $A_{1.}^*$, $A_{2.}^*$ and $A_{3.}^*$.

$$\text{Recall that } w = \begin{bmatrix} 21/80 \\ 7/32 \\ 3/4 \\ 21/32 \end{bmatrix}. \text{ Then, } w \le (\min_{i \in N^C(A)} A_{i.}^*)^T = \begin{bmatrix} 7/24 \\ 7/32 \\ 5/6 \\ 35/48 \end{bmatrix}.$$

However, $w_C^T \ne (\min_{i \in N^C(A)} A_{i.}^*)_C$ as there are three strongly connected components in $D^C(A)$.

Acknowledgment. We wish to thank the anonymous reviewer for their insightful and helpful suggestions. We would also like to thank the reviewer and editor for their prompt handling of the manuscript.

<div align="center">REFERENCES</div>

[1] D.J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM Journal on Discrete Mathematics*, 3:450–465, 1990.

[2] V. Anantharam and P. Tsoucas. A proof of the Markov chain tree theorem. *Statistics & Probability Letters*, 8:189–192, 1989.

[3] F. Baccelli, G. Cohen, G.J. Olsder, and J.-P. Quadrat. *Synchronization and Linearity. An Algebra for Discrete Event Systems*. Free Web Edition: http://www-roc.inria.fr/metalau/cohen/documents/BCOQ-book.pdf, 2001.

[4] R.B. Bapat. A max version of the Perron-Frobenius theorem. *Linear Algebra and its Applications*, 275/276:3–18, 1998.

[5] B. Benek Gursoy, O. Mason, and S. Sergeev. The Analytic Hierarchy Process, Max Algebra and Multi-objective Optimisation. *Linear Algebra and its Applications*, doi:10.1016/j.laa.2012.11.020, 2013.

[6] A. Broder. Generating random spanning trees. In *30th Annual Symposium on Foundations of Computer Science*, 442–447, 1989.

[7] P. Butkovič. *Max-Linear Systems: Theory and Algorithms*. Springer-Verlag London Ltd., London, 2010.

[8] B.A. Carré. An algebra for network routing problems. *Journal of the Institute of Mathematics and its Applications*, 7:273–294, 1971.

[9] R.A. Cuninghame-Green. *Minimax Algebra*. Lecture Notes in Economics and Mathematical Systems, 166, Springer-Verlag, Berlin-New York, 1979.

[10] L. Elsner and P. van den Driessche. On the power method in max algebra. *Linear Algebra and its Applications*, 302/303:17–32, 1999.

[11] L. Elsner and P. van den Driessche. Modifying the power method in max algebra. *Linear Algebra and its Applications*, 332/334:3–13, 2001.

[12] L. Elsner and P. van den Driessche. Max-algebra and pairwise comparison matrices. *Linear Algebra and its Applications*, 385:47–62, 2004.

[13] L. Elsner and P. van den Driessche. Max-algebra and pairwise comparison matrices II. *Linear Algebra and its Applications*, 432:927–935, 2010.

[14] M.I. Freĭdlin and A.D. Wentzell. *Perturbations of Stochastic Dynamic Systems*. Springer-Verlag, New York, 1984. (Translation of Russian edition, Nauka, Moscow, 1979.)

ELA

http://math.technion.ac.il/iic/ela

[15] H.-H. Kohler and E. Vollmerhaus. The frequency of cyclic processes in biological multistate systems. *Journal of Mathematical Biology*, 9:275–290, 1980.

[16] F.T. Leighton and R.L. Rivest. The Markov chain tree theorem. Technical Report MIT/LCS/TM-249, Massachusetts Institute of Technology, Laboratory for Computer Science, 1983.

[17] G.L. Litvinov and V.P. Maslov. The correspondence principle for idempotent calculus and some computer applications. In J. Gunawardena (editor), *Idempotency*, Cambridge University Press, 420–443, 1998.

[18] G.-J. Olsder and C. Roos. Cramer and Cayley-Hamilton in the max algebra. *Linear Algebra and its Applications*, 101:87–108, 1988.

[19] G.J. Olsder and J. van der Wounde. *Max Plus at Work. Modeling and Analysis of Synchronized Systems: A Course on Max-Plus Algebra and Its Applications*. Princeton University Press, Princeton, NJ, 2006.

[20] T.L. Saaty. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15:234–281, 1977.

[21] S. Sergeev, H. Schneider, and P. Butkovič. On visualization scaling, subeigenvectors and Kleene stars in max algebra. *Linear Algebra and its Applications*, 431:2395–2406, 2009.

[22] B.O. Shubert. A flow-graph formula for the stationary distribution of a Markov chain. *IEEE Transactions on Systems, Man, and Cybernetics*, 5:565–566, 1975.

[23] I. Sonin. The state reduction and related algorithms and their applications to the study of Markov chains, graph theory, and the optimal stopping problem. *Advances in Mathematics*, 145:159–188, 1999.

[24] J.R. Wicks. *An Algorithm to Compute the Stochastically Stable Distribution of a Perturbed Markov Matrix*. Ph.D. Thesis, Brown University, Providence, RI, USA, 2009.