UTILITAS MATHEMATICA 124 (2025) 3–24

Utilitas Mathematica
www.combinatorialpress.com/um

Line digraphs of polytrees

Sergiy Kozerenko✉, Bohdan-Yarema Dekhtiar

ABSTRACT

We characterize line digraphs of polytrees, including several of their well-known subclasses. For a given undirected tree, we characterize its orientations with weak line digraphs, and count the exact number. Furthermore, we find the minimum, maximum, and average sizes of these line digraphs. We provide an explicit formula for the number of weak components in line digraphs of polytrees in terms of the inner sources and sinks. Additionally, we count the average number of weak components in them among all orientations of a fixed tree. Finally, we propose an algorithm for finding weak components in line digraphs of polytrees.

Keywords: line digraph, polytree, in-tree, out-tree

2020 Mathematics Subject Classification: 05C76, 05C20, 05C05.

1. Introduction

For any collection of sets $\mathcal{F}$ of arbitrary nature one can associate the corresponding intersection graph, which is an undirected graph having $\mathcal{F}$ as its vertex set and in which two sets $A, B \in \mathcal{F}$ are adjacent provided $A \cap B \neq \emptyset$. This construction is "universal" in a sense that any graph is isomorphic to the intersection graph for some family $\mathcal{F}$. However, if we restrict ourselves to some particular collections of sets $\mathcal{F}$, the problem of characterizing resulting intersection graphs becomes much more interesting. Among such well-known "geometric" set families are intervals on the real line [10] (interval graphs), chords on a circle [12] (circle graphs), curves on a plane [3] (string graphs) and many other similar constructions.

Graph theory itself also suggests several natural families of sets to consider for the

✉ Corresponding author.
E-mail address: kozerenkosergiy@ukr.net (S. Kozerenko).

Received 11 Jul 2025; Revised 15 Aug 2025; Accepted 18 Sep 2025; Published 26 Sep 2025.

DOI: 10.61091/um124-01

intersection graphs. We only mention the families of edges in a graph [2] (line graphs), cliques [11] (clique graphs), blocks [7] (block graphs) etc. Thus, the line graph $L(G)$ of a graph $G = (V(G), E(G))$ has a vertex set $E(G)$ with two edges being adjacent if they share a common vertex. A well-known Whitney isomorphism theorem [13] states that, apart from one trivial exception, the line graph $L(G)$ characterizes its parent graph G uniquely (up to isomorphism). Thus, the transition to the line graph does not lose any information. Line graphs admit many nice characterizations [8, Theorem 8.4] (as do several its subclasses, for example, the line graphs of complete graphs, trees and bipartite graphs).

Moving from undirected graphs to their directed versions (shortly, digraphs), we can modify the construction of the intersection graph as follows. Let $\mathcal{F}$ be a collection of ordered pairs of sets. The corresponding intersection digraph has elements from $\mathcal{F}$ as vertices with two of them (X, Y) and (Z, T) being adjacent provided $X \cap T \neq \emptyset$ (see [4]).

For a digraph $D = (V(D), A(D))$, its line digraph $L(D)$ is then defined as the intersection digraph for the family of pairs $\mathcal{F} = \{(\{v\}, \{u\}) : (u, v) \in A(D)\}$. In other words, the vertices of $L(D)$ can be identified with the arcs in D with the existence of an arc $\alpha \to \beta$ in $L(D)$ provided the head of α is the tail of β. Similarly to the line graphs, line digraphs also admit nice characterizations [1] (the main criterion will be presented in Theorem 2.4).

In this work, we consider the properties of line digraphs of polytrees (which are orientations of undirected trees). The paper is organized as follows.

In Section 2, we briefly give all the necessary preliminary information about undirected graphs, digraphs, and line digraphs.

In Section 3.1, we present criteria for line digraphs of polytrees (Theorem 3.1) and several of their subclasses, including in-trees (out-trees) and the orientations of paths and stars (Propositions 3.5–3.7). We also characterize polytrees which are line digraphs themselves (Proposition 3.3).

Section 3.2 is devoted to studying the polytrees with weak line digraphs. In particular, at first we characterize such polytrees (Proposition 3.8). Then, for a given tree, we find an explicit number of its orientations having weak line digraphs (Theorem 3.11). After finding the exact number of arcs in line digraphs of out-trees and in-trees (Proposition 3.12), we calculate the minimum size of $L(T)$ among those orientations T of a given tree X which have weak $L(T)$'s (Theorem 3.13). Similarly, after proving that any orientation T of a given tree X which maximizes the size of $L(T)$ necessarily has a weak $L(T)$ (Proposition 3.14), we explicitly calculate the maximum size of such a line digraph in Theorem 3.16. Also, using the first Zagreb index, we calculate the average size of line digraphs among all orientations of a given tree.

In Section 3.3, we calculate the number of weak components in line digraphs of polytrees in terms of the inner sources and sinks (Proposition 3.19), and as a corollary, provide the average number of weak components among orientations of a given tree. We then proceed by presenting an algorithm for finding weak components in line digraphs of polytrees and justify its correctness in Theorem 3.24.

We note that some of the results from this paper were announced at International

Conference of Young Mathematicians [5] in 2023.

2. Preliminaries

2.1. Undirected graphs

An *undirected graph* or just a *graph* is an ordered pair $G = (V, E)$, where $V = V(G)$ is the set of its *vertices* and $E = E(G)$ is the set of its *edges*. Hence, our graphs are *simple*, i.e. they do not have loops nor multiple edges. Also, all graphs in this paper are assumed to be finite. For convenience, we denote an edge $\{u, v\}$ as uv.

For a graph G, its *line graph* $L(G)$ is an intersection graph on the collection of $E(G)$, i.e. $V(L(G)) = E(G)$ with two edges $e_1, e_2 \in E(G)$ being adjacent in $L(G)$ provided they share a vertex.

A graph is called *connected* if there is a path between any pair of its vertices. On the vertex set $V(G)$ of a connected graph G, one can define a *metric* d_G with $d_G(u, v)$ equal to the length of the shortest path between u and v. A *metric interval* $[u, v]_G$ is the set of all vertices w such that there is a shortest path between u and v containing w. A set A of vertices in a graph is called *convex* if $[u, v]_G \subset A$ for all $u, v \in A$. For any set A of vertices of G, its *convex hull* $\mathrm{Conv}_G(A)$ is defined as the minimal (by inclusion) convex set that contains A. A set of vertices $A \subset V(G)$ is called *Chebyshev* provided for any vertex $u \in V(G)$ there exists a unique vertex $v \in A$ with $d_G(u, v) = d_G(u, A) := \min\{d_G(u, a) : a \in A\}$. We denote such a vertex v by $\mathrm{pr}_A(u)$.

A connected graph without cycles is called a *tree*. A path P_n is a connected graph whose vertices have degrees at most two. A star $K_{1,n-1}$ is a complete bipartite graph with one of the parts consisting of a single vertex, called the *center* of the star. In a tree, a vertex with a degree of one is called a *leaf*. We use $\mathrm{Leaf}(X)$ to denote the set of all leaf vertices in X.

2.2. Directed graphs

A *directed graph*, often referred to as a *digraph*, is an ordered pair $D = (V, A)$, where $V = V(D)$ is the set of its *vertices* and $A = A(D) \subset V \times V$ is the set of its *arcs*. Thus, our digraphs can have *loops* (arcs of the form (u, u)). For an arc $\alpha = (u, v)$ in a digraph D, the vertex u is called the *tail* of α and v is called the *head* of α. Consequently, α is called an *out-arc* from the vertex u and an *in-arc* into the vertex v.

The *out-degree* $d^+_D(u)$ of a vertex $u \in V(D)$ in a digraph D is the number of out-arcs from u. A vertex u is called a *sink* if $d^+_D(u) = 0$. Similarly, the *in-degree* $d^-_D(u)$ of a vertex $u \in V(D)$ in a digraph D is the number of in-arcs into u. A vertex u is called a *source* if $d^-_D(u) = 0$.

If $\alpha = (u, v)$ and $\beta = (v, w)$ are two arcs in a digraph D, then we say that α is *adjacent to* β and β is *adjacent from* α. And both α and β are *adjacent with* each other. Similarly, in this case, u is *adjacent to* v, v is *adjacent from* u, and both are *adjacent with* each other. Finally, we say that u is *incident to* α, v is *incident from* α, and both vertices are *incident with* α.

A *walk* is a finite sequence of vertices in which every next vertex is adjacent from the previous one. A *dipath* is a walk with different vertices. If there exists a dipath beginning at vertex u and ending at vertex v, then u is *connected to* v, v is *connected (reachable) from* u, and both are *connected with* each other. Similarly, we define connectedness for the first and last arcs of a dipath. A *directed cycle* is a walk in which exactly two vertices coincide - the first and the last one. The *length* of a directed cycle is the number of different vertices in it. Directed cycle of length 1 is called a *loop*.

For a digraph D, its *converse digraph* is the digraph D^{co} with the vertex set $V(D^{co}) = V(D)$ and the arc set $A(D^{co}) = \{(a,b) : (b,a) \in A(D)\}$.

For a digraph, D by $[D]$ we denote the corresponding undirected graph which is obtained from D, i.e. $V([D]) = V(D)$ and $E([D]) = \{uv : u \neq v$ and $(u,v) \in A(D)$ or $(v,u) \in A(D)\}$. A digraph is called *weakly connected* or *weak* if $[D]$ is connected. A *weak component* of a digraph is its maximal weak subgraph. A digraph D is *empty* if $A(D) = \emptyset$.

2.3. *Line digraphs*

Given a digraph D, its *line digraph* is a digraph $L(D)$ with the vertex set $V(L(D)) = A(D)$ and the arc set

$$A(L(D)) = \{(\alpha,\beta) \in A(D) \times A(D) : \alpha = (a,b), \beta = (b,c) \text{ for } a,b,c \in V(D)\}.$$

An example of a digraph and its line digraph is given in Figure 1.

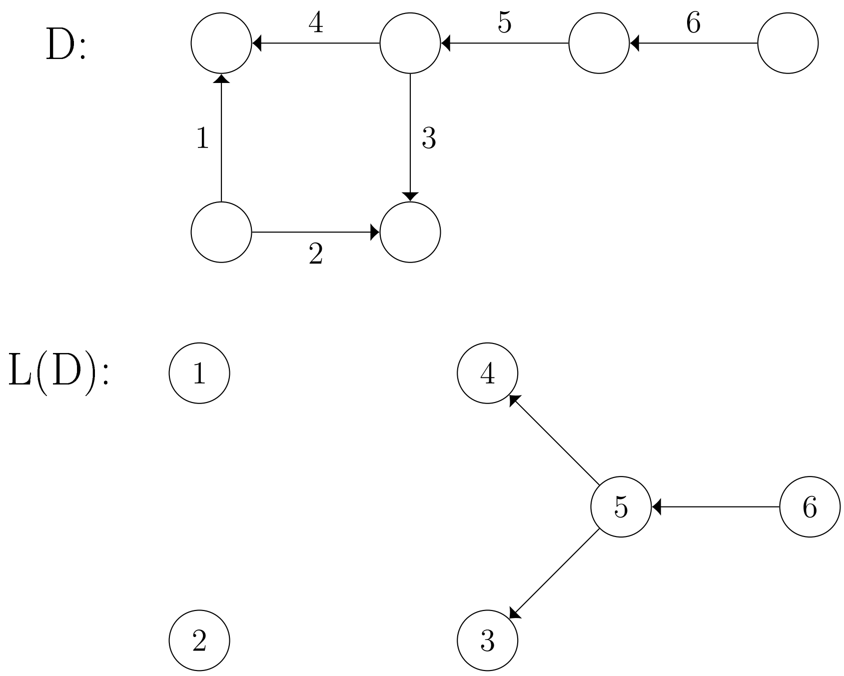

Fig. 1. A digraph D and its line digraph $L(D)$

The next simple result about the size of a line digraph will be frequently used without reference.

Theorem 2.1. [9] *For any digraph D we have*

$$|A(L(D))| = \sum_{u \in V(D)} d_D^+(u) \cdot d_D^-(u).$$

Let G be a graph. An *orientation of G* is any digraph D without directed cycles of length ≤ 2 and $[D] = G$. A *polytree, polypath, polystar, polycycle* is an orientation of a tree, path, star, cycle, respectively. An *in-tree (out-tree)* is a polytree in which all edges are oriented towards (from) some sink (source). Note that the converse digraph of an in-tree is an out-tree and vice versa.

A *bipartite* digraph is an orientation of a bipartite graph in which all arcs are oriented from one part (of the corresponding bipartition) to another. A *two-port* digraph is a bipartite digraph D with $[D]$ being a complete bipartite graph. We shall denote it as $K(A, B)$, where $A, B \subset V(D)$ are the corresponding parts. The next characterization is straightforward.

Proposition 2.2. *A graph G is bipartite if and only if there exists its orientation D with $L(D)$ being an empty digraph.*

Proof. *Necessity.* Let A and B be the parts of G. Construct the corresponding bipartite digraph D with every edge oriented from A to B. Then D contains no dipath of length greater than one, thus $L(D)$ is empty.

Sufficiency. Since $L(D)$ has no arcs, every vertex of D is either a source or a sink. Let A be the set of sources of D, and B the set of its sinks. Then every arc's tail belongs to A, and every arc's head belongs to B. Therefore, by neglecting the orientation of arcs we get a bipartite graph G with parts equal to A and B. □

It is easy to see that $L(D)$ is a directed cycle if and only if D is a directed cycle itself. The next result provides a characterization of line digraphs of other polycycles.

Proposition 2.3. *For a digraph D there exists a polycycle but not a directed cycle C such that $L(C) \simeq D$ if and only if D has an even number of weak components each being a dipath.*

Proof. *Necessity.* As a polycycle which is not a directed cycle, C consists of $2k$ dipaths $W_1, W_2, \ldots, W_{2k}$, where W_i and $W_{(i+1) \bmod 2k}$ share the starting vertex or the endvertex. Clearly, $L(W_1), L(W_2), \ldots, L(W_{2k})$ are weak components in $L(C)$ and each $L(W_i)$ is a dipath of length $|W_i| - 1$ for all $i \in [1, 2k]$.

Sufficiency. Let $P_1, \ldots, P_{2k}$ be the weak components of D. Hence, every P_i is a dipath. Identify the endvertices in P_1 and P_2, identify starting vertices in P_2 and P_3, and so on alternately. Then, since $2k$ is even, P_{2k} and P_1 will share the starting vertex, and we will have built a polycycle C consisting of $P_1, P_2, \ldots, P_{2k}$. It is easy to see that $L(C) = D$. □

Given a set S, its *improper partition* is a collection of (possibly empty) subsets $S_i \subset S$ such that $S = \sqcup S_i$.

The main characterization of line digraphs is due to Harary and Norman [9].

Theorem 2.4. [9] *A digraph D is a line digraph if and only if there exist two improper partitions A_i and B_i of $V(D)$ such that $A(D) = \bigcup_i K(A_i, B_i)$.*

3. Main results

3.1. Characterizations of line digraphs of polytrees and their subclasses

Our first result is a characterization of line digraphs of polytrees.

Theorem 3.1. *Let D be a line digraph. Then there exists a polytree T with $L(T) \simeq D$ if and only if every polycycle in D is a bipartite subgraph.*

Proof. *Necessity.* Let C' be a polycycle in D. Clearly, C' can not be a directed cycle. Thus C' consists of $2k$ dipaths $W'_1, W'_2, \dots, W'_{2k}$ (see Proposition 2.3) with dipaths W'_i and W'_{i+1} being connected at a single vertex, which is either their common source or their common sink, depending on the parity of i (we shall write $i+1$ instead of $(i+1) \bmod 2k$ for simplicity). Then $L^{-1}(C')$, which is a subgraph of T, consists of $2k$ dipaths $W_1, W_2, \dots, W_{2k}$, with $|W_i| = |W'_i| + 1$. Also, W_i and W_{i+1} share an arc - the very arc that generates the common vertex of W'_i and W'_{i+1}. All the arcs in $L^{-1}(C')$, save for the shared ones, form a polycycle C whose dipaths are of length $|W_i| - 2 = |W'_i| - 1$ (the first and last arcs of each dipath W_i are shared with other dipaths and thus not included). But there is a single case when $|W'_i| = 1$ for all $i \in [1, 2k]$ and C becomes a single vertex. Since T is a polytree, it contains no polycycles, so C is indeed a vertex and $|W'_i| = 1$ for all i. But that means that C' is a bipartite subgraph, with every source adjacent to both of its neighbors and every sink incident from its neighbors.

Sufficiency. As D is a line digraph, there exist two improper partitions A_i and B_i such that $A(D) = \bigcup_i K_i$, where $K_i := K(A_i, B_i)$. Let $E = L^{-1}(D)$. Each K_i corresponds to a vertex of E. Every other vertex of E is a source or a sink; construct E in such a way that all those vertices are of degree one, therefore not a being part of a polycycle.

First we shall prove that in such a case, E has no polycycles. To the contrary, let C be a polycycle in E, $a_1 - a_2 - \dots - a_n - a_1$ be its vertices, and α_i be the arc between a_i and a_{i+1}. As previously shown, each vertex a_i is associated with a two-port subgraph K_i of D. Every arc α_i corresponds to a vertex α'_i of D, which is common for K_i and K_{i+1}. So α'_{i-1} and α'_i are distinct vertices that belong to K_i. If α'_{i-1} and α'_i are from different ports, then there is an arc between them. Otherwise, if they both are from port A_i, there exists a polypath $\alpha'_{i-1} - \beta - \alpha'_i$, where β is any vertex from port B_i (similarly, if they are both from B_i). In every case, there is a polypath between α'_{i-1} and α'_i. Hence, the vertices $\alpha'_1, \alpha'_2, \dots, \alpha'_n$ together with some intermediate vertices form a polycycle C' of D. Note that C' is not a bipartite subgraph, as every weak bipartite subgraph of D is either K_i or a subgraph of K_i. We have a contradiction, which implies E has no polycycles.

Assume E consists of k weak components $E_1, E_2, \dots, E_k$. Each of them has at least one sink; let a_i be a sink that belongs to E_i. By merging $a_1, a_2, \dots, a_k$ into one vertex, we obtain a digraph T with $L(T) = L(E) = D$. Clearly, T is weak and has no polycycles, thus it is a polytree. □

Example 3.2. Consider a digraph D with the vertex set $V(D) = \{1, \dots, 7\}$ and the arc set $A(D) = \{(1,3),(2,3),(3,6),(3,7),(4,6),(4,7),(5,6),(5,7)\}$ (see Figure 2). Then D is a line digraph with improper partitions being $A_1 = \{1,2\}$, $B_1 = \{3\}$, $A_2 = \{3,4,5\}$, $B_2 = \{6,7\}$, $A_3 = \{6,7\}$, $B_3 = \{1,2,4,5\}$. Also, D contains three polycycles ($4-6-3-7-4$, $4-6-5-7-4$, and $3-6-5-7-3$), all of which are bipartite subgraphs. We conclude that D is a line digraph of a polytree T (depicted in Figure 2).

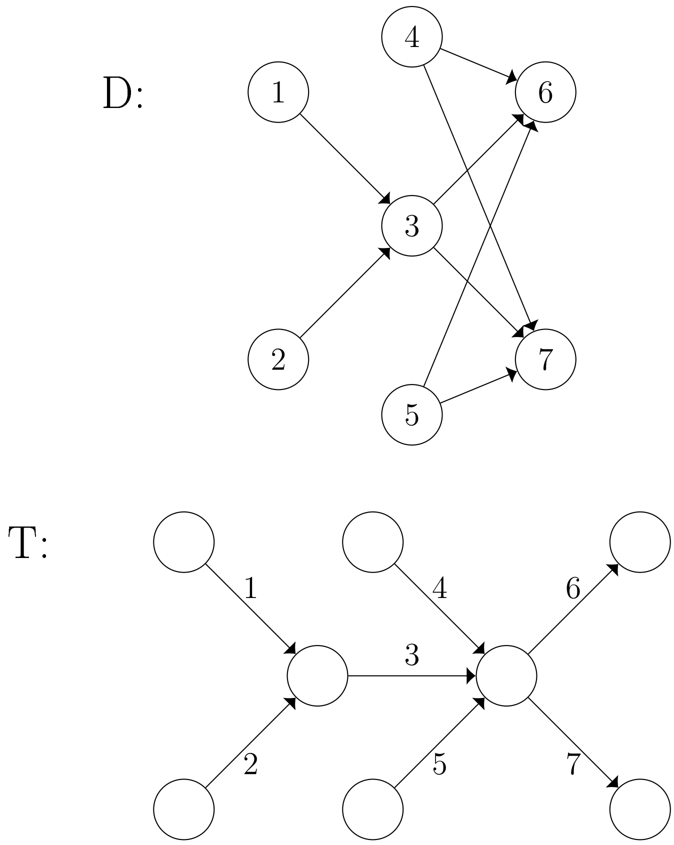

Fig. 2. A digraph D which is the line digraph of a polytree T

Next we characterize polytrees which are line digraphs themselves.

Proposition 3.3. *Let T be a polytree. Then T is a line digraph if and only if for all arcs $(a,b) \in A(T)$ either $d^+_T(a) = 1$ or $d^-_T(b) = 1$. Moreover, this is also equivalent to the existence of a polytree T' such that $L(T') \simeq T$.*

Proof. *Necessity* Let A_i and B_i be two improper partitions as in Theorem 2.4, where i ranges from 1 to n. Let A_k be the subset containing a and B_k the subset containing b. Suppose $d^+_T(a) > 1$ and $d^-_T(b) > 1$. Then there exists a vertex d adjacent from a and a vertex c adjacent to b. The vertices c and d are distinct, for otherwise T would contain a polycycle $a-c-b-a$. Since all the arcs incident from a are contained within $K(A_k, B_k)$, it holds $d \in B_k$. Similarly, $c \in A_k$. Thus T contains a polycycle $a-b-c-d-a$, a contradiction.

Sufficiency Let (a,b) be an arc such that $d_T^+(a) = 1$. Then we put $A_i := N_T^-(b)$ (obviously, $a \in A_i$) and $B_i := \{b\}$, so that $K(A_i, B_i)$ is the set of all arcs ending at b. The case when $d_T^-(b) = 1$ is covered similarly, that is, $A_i := N_T^+(a)$ and $B_i := \{b\}$. If simultaneously $d_T^+(a) = 1$ and $d_T^-(b) = 1$, the two processes yield the same result. It is easy to see that we thus constructed two improper partitions A_i and B_i of $V(T)$ such that $A(T) = \bigcup_i K(A_i, B_i)$. Then Theorem 2.4 implies that T is a line digraph. The existence of T' follows from Theorem 3.1. □

Remark 3.4. Proposition 3.3 immediately implies that a polytree is a line digraph is and only if it does not contain a "forbidden" subtree depicted in Figure 3.

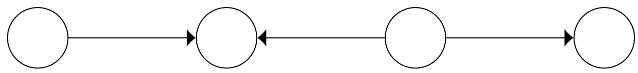

Fig. 3. The "forbidden" subtree for polytrees which are line digraphs (the middle arc violates the condition of Proposition 3.3)

We now consider several specific classes of polytrees and characterize their line digraphs.

Proposition 3.5. *Let D be a digraph. Then $D \simeq L(T)$ for an in-tree (out-tree) T if and only if every weak component of D is an in-tree (out-tree).*

Proof. We shall offer proof for in-trees, considering the proof for out-trees sufficiently similar to be omitted.

Necessity. Let E be a weak component of D. Additionally, let s be the unique sink of T. Since s is reachable from at least one vertex of T (in fact, it is reachable from every vertex), there is an in-arc α to s associated with a vertex α' of E. This in-arc α is unique, as otherwise the connectedness of E would imply the existence of a vertex in T from which we could reach s by two different polypaths. Therefore, α' is the only sink of E, and so it can be reached from every vertex. Consequently, E is an in-tree.

Sufficiency. Let again E be a weak component of D. As it is an in-tree, there is no vertex of E with out-degree greater than one, as otherwise there would be two different dipaths from it to s, s being the sink of E. It follows from Proposition 3.3 that there exists a polytree F such that $L(F) = E$. If α is the arc of F associated with s, then it is reachable from every other arc of F. Thus vertex b of F, which is incident from α, is reachable from every other vertex of F. Hence, F is an in-tree. Now let $F_1, F_2, \ldots, F_n$ be in-trees for which $L(F_i) \simeq E_i$ for all $i \in [1, n]$, E_i being the weak components of D. Let all F_i's have a common sink b, and let it be the only common vertex of different F_i

(therefore there will be no common arcs). This way we obtain a digraph T, which is still an in-tree. Also, $L(T) \simeq D$, which completes the proof. □

Proposition 3.6. *Let D be a digraph. Then $D \simeq L(T)$ for a polypath T if and only if every weak component of D is a dipath.*

Proof. *Necessity.* By definition, a polypath is given by a sequence of dipaths where each two consecutive paths share a common vertex (alternately starting and ending vertices). After taking the line digraph, every dipath of length k will become a separate weak component, which by itself is a dipath of length $k-1$. That completes the proof.

Sufficiency. Let $D_1, D_2, \dots, D_n$ be the weak components of D, each of them being a dipath of length $k_1, k_2, \dots, k_n$, respectively. Let T_i, $i \in [1, n]$ be digraphs such that $D_i = L(T_i)$ for all $i \in [1, n]$. Then every T_i is a dipath of length $k_i + 1$. Let a digraph T be a sequence of $T_1, T_2, \dots, T_n$ with T_1 and T_2 sharing an endvertex, T_2 and T_3 sharing the starting vertex, and so on alternately. Then T is a polypath with $L(T) = D$. □

Proposition 3.7. *Let D be a digraph. Then $D \simeq L(S)$ for a polystar S if and only if D is either a two-port digraph or an empty digraph.*

Proof. *Necessity.* If the center p of S is either the common source or the common sink of all its arcs, then D is empty. Otherwise, let $\alpha_1, \alpha_2, \dots, \alpha_n$ be arcs ending at p, and $\beta_1, \beta_2, \dots, \beta_m$ be arcs starting at p. Let $a_1, \dots, a_n$, $b_1, \dots, b_m$ be corresponding vertices of $D = L(S)$. Then D is a two-port digraph, its ports being $A = \{a_1, a_2, \dots, a_n\}$ and $B = \{b_1, b_2, \dots, b_m\}$. So D is either a two-port or an empty digraph.

Sufficiency. If D is a two-port digraph, then S is a vertex together with arcs incident to and from it (which is a polystar). If D is empty with n vertices, then $D = L(S)$ for a polystar S with n arcs, all of them adjacent to (or all adjacent from) the center. In any case, D is a line digraph of a polystar. □

3.2. Polytrees with weak line digraphs

We begin by giving a characterization of polytrees with weak line digraphs.

Proposition 3.8. *Let T be a polytree. Then $L(T)$ is weak if and only if every source and every sink of T is a leaf in $[T]$.*

Proof. *Necessity.* Let s be the source in T of out-degree $m \geq 2$ (the proof is similar for a sink). As T is a polytree, there is no vertex to which we could get from s by two different polypaths; thus each of the neighbors of s belongs to a different induced subgraph, all the subgraphs sharing exactly one vertex, s. In $L(T)$ these subgraphs will become different weak components.

Sufficiency. Divide the arcs of T into two non-empty subsets $A \sqcup B = A(T)$. As T is weak, there is a vertex a and two arcs $\alpha \in A, \beta \in B$ such that a is incident with both α

and β. Since a is not a source nor a sink, we may presume that a is incident from α to β (if, for instance, α and β are both in-arcs at a, then, as a a is not a sink, there exists $\gamma \in B$ such that γ is an out-arc at a (if $\gamma \in A$, swap A and B); in that case $\beta := \gamma$). Let A' be the set of vertices of $L(T)$ generated by the arcs from A, and let B' be the set of vertices generated by the arcs from B. Then $A' \sqcup B' = V(L(T))$. It is easy to see that there is an arc between A' and B' – this is the arc between vertices which correspond to α and β. Since every such pair of sets A' and B' can be obtained in a similar way, this means that for every such partition of vertices of $L(T)$ into two sets, there is an arc between vertices from different sets. This implies that $L(T)$ is weak. □

Proposition 3.8 asserts that for an (undirected) tree X any its orientation T with weak $L(T)$ gives a non-trivial partition of $\mathrm{Leaf}(X)$ into two subsets (the sets of sources and sinks of T). The next result shows the converse is also true, i.e. that for any non-trivial partition of the set $\mathrm{Leaf}(X)$ there is a corresponding orientation T of X.

Proposition 3.9. *Let X be a tree with $|V(X)| \geq 2$ and $\mathrm{Leaf}(X) = A \sqcup B$ be a partition of its leaf vertices. Then there is an orientation T of X having A as the set of its sources and B as the set of its sinks if and only if A and B both are nonempty sets.*

Proof. *Necessity.* It is clear that every acyclic digraph (in particular, a polytree) with at least two vertices has a source as well as a sink.

Sufficiency. Assume that $A, B \neq \emptyset$. Consider the convex hull $C = \mathrm{Conv}_X(B)$ of B. Since $B \neq \emptyset$, then trivially $C \neq \emptyset$ as well. Further, since $A \neq \emptyset$, we can fix a vertex $x_0 \in A$. Since C is convex, it is also connected, which in case of a tree means C has to be Chebyshev. Thus we can construct an orientation T of X as follows: orient the subtree $X[C]$ as an out-tree from the root $\mathrm{pr}_C(x_0)$ and orient each edge $uv \in E(X)\backslash E(C)$ as $u \to v$ if $v \in [u, \mathrm{pr}_C(u)]_X$ (i.e. we orient the edge uv "towards" C). Since $T[C]$ is an out-tree rooted at a non-leaf vertex $\mathrm{pr}_C(x_0)$ in X, it is clear that the elements of $\mathrm{Leaf}(X[C])$ are exactly the sinks in $T[C]$. Hence, the inclusion $B \subset \mathrm{Leaf}(X[C]) \cap \mathrm{Leaf}(X)$ implies that the elements of B are sinks in T as well.

Now consider a vertex $u \in A$ and the corresponding leaf edge $uv \in E(X)$ (clearly $uv \notin E(C)$). By construction of T, we have $(u, v) \in A(T)$ which implies that u is a source in T. Therefore, the elements of A are the sources in T.

Now we only need to show that there are no other sources or sinks in T. Indeed, since $T[C]$ is an out-tree, any non-leaf vertex in $X[C]$ is not a source nor a sink in T. Also, the minimality of C implies $\mathrm{Leaf}(X[C]) = B$, thus no element of $C\backslash B$ can be a source or a sink. Finally, assume that $y \in V(X)\backslash(C \cup A)$. Then y has two neighbors $u, v \in N_X(y)$ with $y \in [u, \mathrm{pr}_C(u)]_X$ and $v \in [y, \mathrm{pr}_C(y)]_X$. By construction, $(u, y), (y, v) \in A(T)$ implying that y also is not a source nor a sink in T. □

Example 3.10. Consider a tree X with the vertex set $V(X) = \{1, \ldots, 8\}$ and the edge set $E(X) = \{12, 23, 26, 34, 37, 38, 45, 56\}$. Consider a partition of $\mathrm{Leaf}(X) = \{1, 5, 6, 7, 8\}$ into two sets $A = \{5, 6, 8\}$ and $B = \{1, 7\}$. We have that $C := \mathrm{Conv}_X(B) = \{1, 2, 3, 7\}$.

Fix a vertex $x_0 = 8 \in A$ and orient X as in proof of Proposition 3.9: the subtree $X[C]$ is oriented as an out-tree from the root $\mathrm{pr}_C(x_0) = 3$ and other edges are oriented towards $X[C]$ (see Figure 4). It is clear that the sources of this orientation are exactly the vertices from A and the sinks are the vertices from B.

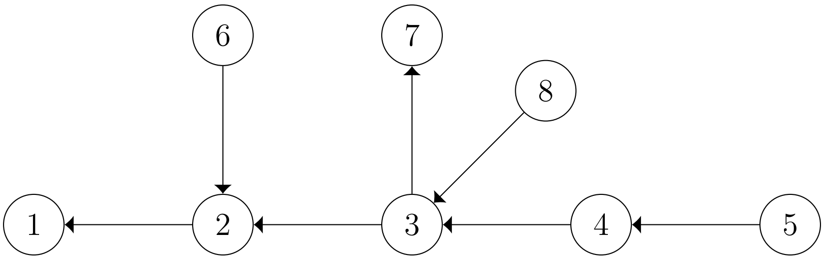

Fig. 4. The orientation of the tree X from Example 3.10

In fact, for a given tree X, we can explicitly calculate the number of orientations of X having weak line digraphs. Note that this number depends only on the degree sequence of X and not on its structure.

Theorem 3.11. *Let X be a tree with $n \geq 3$ vertices. Then the number of orientations of T having weak line digraphs equals*

$$2 \cdot \prod_{u \in V(X) \setminus \mathrm{Leaf}(X)} (2^{d_X(u)-1} - 1).$$

Proof. Let us call an orientation T of X *good* if $L(T)$ is weak.

We use induction on the number of non-leaf vertices in X. First, note that since $n \geq 3$, X has at least one such vertex. If $|V(X) \setminus \mathrm{Leaf}(X)| = 1$, then $X \simeq K_{1,n-1}$ is a star. And among 2^{n-1} orientations of X exactly two orientations have the center of X as a source or a sink. By Proposition 3.8, they yield a disconnected line digraph. Hence, in case of a star, there are $2^{n-1} - 2 = 2(2^{(n-1)-1} - 1)$ good orientations T of X.

Further, assume $|V(X) \setminus \mathrm{Leaf}(X)| \geq 2$. Then there exists a vertex $v_0 \in \mathrm{Leaf}(X \setminus \mathrm{Leaf}(X))$. Consider the tree $X' = X \setminus (N_X(v_0) \cap \mathrm{Leaf}(X))$ which is obtained from X by deleting all $d_X(v_0) - 1$ leaf neighbors of v_0. Since v_0 is a leaf vertex in X', the tree X' has fewer non-leaf vertices than X.

By induction assumption, the number of good orientations of X' equals

$$2 \cdot \prod_{u \in V(X') \setminus \mathrm{Leaf}(X')} (2^{d_{X'}(u)-1} - 1) = 2 \cdot \prod_{u \in V(X) \setminus (\mathrm{Leaf}(X) \cup \{v_0\})} (2^{d_X(u)-1} - 1).$$

Proposition 3.8 immediately asserts that each good orientation T of X induces the corresponding good orientation $T' = T \setminus \{v_0 x : x \in N_X(v_0) \cap \mathrm{Leaf}(X)\}$ of X'. Hence, in order to calculate the number of good orientations of X, we need to calculate how many of them induce the same good orientation T' of X'.

Therefore, let us fix a good orientation T' of X'. Overall there are $2^{d_X(v_0)-1}$ orientations T of X having T' as a suborientation. Among them only one orientation has v_0 as a source or a sink (depending on the orientation in T' of the unique edge incident to v_0 in X'). Thus, for a given T' we have exactly $2^{d_X(v_0)-1} - 1$ corresponding good orientations T. This implies that the number of good orientations of X equals

$$2 \cdot \prod_{u \in V(X) \setminus (\mathrm{Leaf}(X) \cup \{v_0\})} (2^{d_X(u)-1} - 1) \cdot (2^{d_X(v_0)-1} - 1) = 2 \cdot \prod_{u \in V(X) \setminus \mathrm{Leaf}(X)} (2^{d_X(u)-1} - 1).$$

□

Now we turn our attention to the orientations of trees which are extremal with respect to the number of arcs in their line digraphs. But at first, we explicitly calculate the number of arcs in $L(T)$ for out-trees and in-trees T.

Proposition 3.12. *If T is an n-vertex out-tree or an in-tree rooted at u, then $|A(L(T))| = n - 1 - d_{[T]}(u)$.*

Proof. As $|A(D)| = |A(D^{co})|$ and $(L(D))^{co} = L(D^{co})$ for any digraph D, we only consider the case of out-trees T. Thus let T be an out-tree rooted at u. Then $d_T^+(u) = d_{[T]}(u)$ and $d_T^-(u) = 0$. Further, for all $x \in V(T) \setminus \{u\}$ it holds $d_T^+(x) = d_{[T]}(x) - 1$ and $d_T^-(x) = 1$. Therefore,

$$\begin{aligned}
|A(L(T))| &= \sum_{x \in V(T)} d_T^+(x) \cdot d_T^-(x) \\
&= d_T^+(u) \cdot d_T^-(u) + \sum_{x \in V(T) \setminus \{u\}} d_T^+(x) \cdot d_T^-(x) \\
&= \sum_{x \in V(T) \setminus \{u\}} (d_{[T]}(x) - 1) = \sum_{x \in V(T) \setminus \{u\}} d_{[T]}(x) - (n-1) \\
&= (2(n-1) - d_{[T]}(u)) - n + 1 = n - 1 - d_{[T]}(u).
\end{aligned}$$

□

From Proposition 2.2 it follows that any tree X admits an orientation T with $|A(L(T))| = 0$. However, if we restrict ourselves to the orientations with weak line digraphs, we obtain the following result.

Theorem 3.13. *Let X be an n-vertex tree, $n \geq 2$. Then the minimum number of arcs in $L(T)$ among orientations T of X with weak $L(T)$'s equals $n - 2$.*

Proof. Consider an orientation T of X which is an out-tree rooted at some leaf vertex $u \in \mathrm{Leaf}(X)$. By Proposition 3.12, $|A(L(T))| = n - 1 - d_{[T]}(u) = n - 1 - 1 = n - 2$. Also, by Proposition 3.8, $L(T)$ is weak. Hence, the minimum number of arcs in $L(T)$ among orientations T of X with weak $L(T)$'s is at most $n - 2$. We need to show that any such orientation T has at least $n - 2$ arcs in $L(T)$. To do so, we use induction on $n \geq 2$. If $n = 2$, then $X = P_2$ and for any orientation T of X it clearly holds $|A(L(T))| = 0 = 2 - 2$.

Now let $n \geq 3$. At first, assume that X is a star centered at u. Then $|A(L(T))|= d_T^+(u) \cdot d_T^-(u)$. It is clear, that this number is minimized only if $\min\{d_T^+(u), d_T^-(u)\} = 1$. In this case, $|A(L(T))|= n-2$ as $d_T^+(u)+d_T^-(u) = |E(X)|= n-1$. Further, assume that X is not a star. This means that $\mathrm{Leaf}(X-\mathrm{Leaf}(X)) \neq \emptyset$. Fix a vertex $v \in \mathrm{Leaf}(X-\mathrm{Leaf}(X))$ and consider a tree $X' = X - (N_X(v) \cap \mathrm{Leaf}(X))$. It is clear that $|V(X')|\geq 2$. Thus by induction assumption, the number of arcs in $L(T')$ among orientations T' of X' with weak $L(T')$'s is at least $|V(X')|-2 = n - (d_X(v) - 1) - 2 = n - 1 - d_X(v)$. Since $v \in \mathrm{Leaf}(X - \mathrm{Leaf}(X))$, there is a unique vertex $w \in V(X')$ with $e = vw \in E(X')$. If e is oriented in T' as $w \to v$, then every edge vx, $x \in N_X(v) \cap \mathrm{Leaf}(X)$ must be oriented in T as $v \to x$ (since $L(T)$ is weak). Similarly, if e is oriented in T' as $v \to w$, then each edge vx, $x \in N_X(v) \cap \mathrm{Leaf}(X)$ is oriented in T as $x \to v$. In both cases, we obtain

$$|A(L(T))|\geq |A(L(T'))|+|N_X(v) \cap \mathrm{Leaf}(X)|\geq n-1-d_X(v)+(d_X(v)-1) = n-2.$$

□

Now we consider orientations of trees with maximum number of arcs in their line digraphs. At first, we show that any such an orientation T must have a weak $L(T)$.

Proposition 3.14. *Let T be an orientation of a tree X which has a maximum number of arcs in $L(T)$ among all orientations of X. Then $L(T)$ is weak.*

Proof. To the contrary, assume that T is such an orientation of X, but with a non-weak line digraph $L(T)$. By Proposition 3.8, there is a non-leaf vertex $u \in V(X)\backslash\mathrm{Leaf}(X)$ which is a source or a sink in T. Without loss of generality, assume that u is a source in T. Fix a vertex $x \in N_X(u)$ and consider the forest $X - \{ux\}$. Let X' be the connected component in $X - \{ux\}$ which contains x. Consider the new orientation T' of X which coincides with T on edges from $E(X)\backslash(E(X') \cup \{ux\})$ and is opposite on edges from $E(X') \cup \{ux\}$. In other words, we take T and conversely reorient the edges from $E(X') \cup \{ux\}$. Then for all $v \in V(X)\backslash(V(X') \cup \{u\})$ we have $d_{T'}^+(v) = d_T^+(v)$ and $d_{T'}^-(v) = d_T^-(v)$. Also, for all $v \in V(X')$ it holds $d_{T'}^+(v) = d_T^-(v)$ and $d_{T'}^-(v) = d_T^+(v)$. Moreover, $d_{T'}^+(u) = d_X(u) - 1 = d_T^+(u) - 1$ and $d_{T'}^-(u) = 1 = d_T^-(u) + 1$. Hence,

$$\begin{aligned}
|A(L(T'))|&=\sum_{v\in V(T')} d_{T'}^+(v) \cdot d_{T'}^-(v) = \sum_{v\in V(X)\backslash(V(X')\cup\{u\})} d_{T'}^+(v) \cdot d_{T'}^-(v) \\
&\quad+\sum_{v\in V(X')} d_{T'}^+(v) \cdot d_{T'}^-(v) + d_{T'}^+(u) \cdot d_{T'}^-(u) \\
&=\sum_{v\in V(X)\backslash(V(X')\cup\{u\})} d_T^+(v) \cdot d_T^-(v) + \sum_{v\in V(X')} d_T^-(v) \cdot d_T^+(v) + d_X(u) - 1 \\
&=\sum_{v\in V(T)\backslash\{u\}} d_T^+(v) \cdot d_T^-(v) + d_X(u) - 1 = \sum_{v\in V(T)} d_T^+(v) \cdot d_T^-(v) + d_X(u) - 1 \\
&=|A(L(T))|+d_X(u) - 1 \geq |A(L(T))|+1.
\end{aligned}$$

Thus $L(T')$ has more arcs than $L(T)$. The obtained contradiction proves the proposition. □

Example 3.15. Consider a tree X with $V(X) = \{1, \ldots, 7\}$, $E(X) = \{12, 23, 34, 37, 45, 46\}$ and its orientation T depicted in Figure 5.

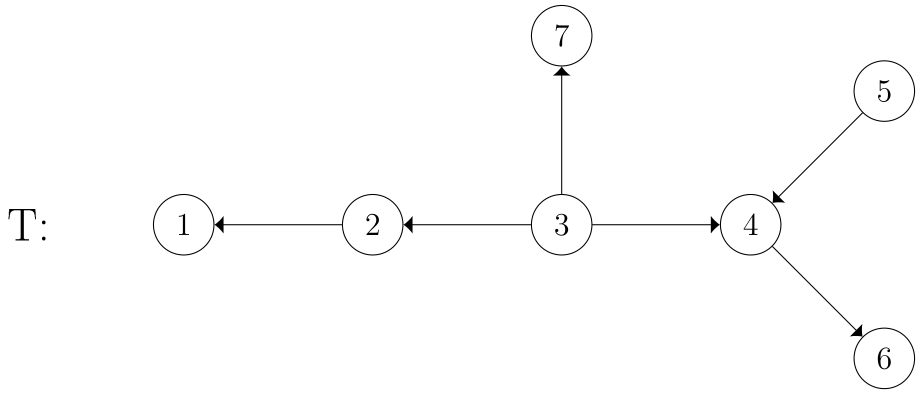

Fig. 5. An orientation T of a tree X from Example 3.15

Then the line digraph $L(T)$ has 3 arcs. Since the vertex 3 is a source in T, we can reorient X to obtain a polytree T' with more arcs in $L(T')$. Indeed, exploiting the idea in proof of Proposition 3.14, we fix an arbitrary edge incident to 3, say 34. Then we consider the forest $X - \{ux\}$ and its connected component X' which contains the vertex 4. Clearly, $X' = X[\{4, 5, 6\}]$. Let us reorient the edges from $E(X') \cup \{34\} = \{34, 45, 46\}$ in the opposite way to obtain the new orientation T' of X (see Figure 6). We have $|A(T')| = 5 > 3 = |A(T)|$.

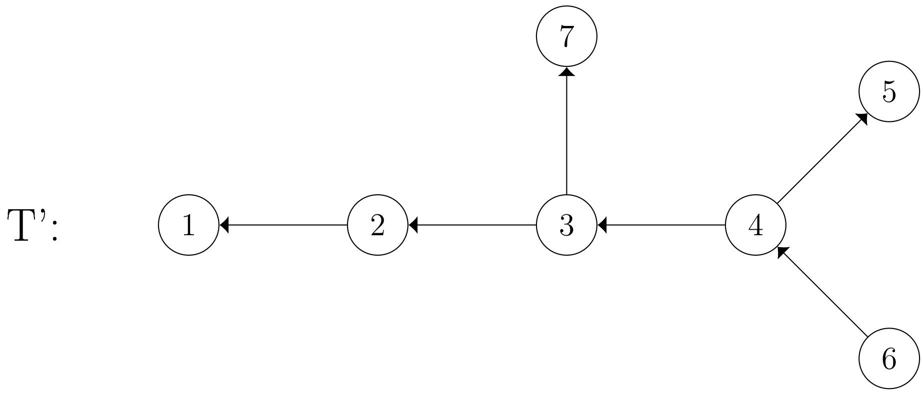

Fig. 6. The new orientation T' of X having $|A(T')| > |A(T)|$ (Example 3.15)

It turns out that we can explicitly calculate the maximum number of arcs in $L(T)$'s among orientations T of a given tree X in terms of the degrees of vertices in X.

Theorem 3.16. *Let X be a tree. Then the maximum number of arcs in $L(T)$ among* *orientations T of X equals*

$$\sum_{u\in V(X)} \left\lfloor \frac{d_X(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_X(u)}{2} \right\rceil .$$

Proof. Let T be an orientation of X. Since for every vertex $u \in V(X)$ we have $d_T^+(u) + d_T^-(u) = d_X(u)$, then $d_T^+(u) \cdot d_T^-(u) \le \left\lfloor \frac{d_X(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_X(u)}{2} \right\rceil$. Therefore,

$$|A(L(T))| = \sum_{u\in V(T)} d_T^+(u) \cdot d_T^-(u) \le \sum_{u\in V(X)} \left\lfloor \frac{d_X(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_X(u)}{2} \right\rceil .$$

To show that there exists an orientation which attains this bound, we again use the same induction on $|V(X)| \ge 1$ as in proof of Theorem 3.13. The cases $|V(X)| = 1$ and $|V(X)| = 2$ are trivial. Now assume that $|V(X)| \ge 3$. If X is a star, then divide $\mathrm{Leaf}(X) = A \sqcup B$ into two almost equal parts A and B with $|A| = \left\lfloor \frac{n-1}{2} \right\rfloor$, $B = \left\lceil \frac{n-1}{2} \right\rceil$. Consider the orientation T of X with the vertices from A being sources and the vertices from B being sinks in it. Clearly,

$$|A(L(T))| = |A| \cdot |B| = \left\lfloor \frac{n-1}{2} \right\rfloor \cdot \left\lceil \frac{n-1}{2} \right\rceil = \sum_{u\in V(X)} \left\lfloor \frac{d_X(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_X(u)}{2} \right\rceil .$$

Now suppose that X is not a star. Then there is a vertex $v \in \mathrm{Leaf}(X - \mathrm{Leaf}(X))$. Consider the tree $X' = X - (N_X(v) \cap \mathrm{Leaf}(X))$. It is clear that $|V(X')| \ge 2$. By induction assumption, there exists an orientation T' of X' with $|A(L(T'))| = \sum_{u\in V(X')} \left\lfloor \frac{d_{X'}(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_{X'}(u)}{2} \right\rceil$. We want to construct the desired orientation T of X as an extension of T', thus we only need to properly orient the edges vx for $x \in N_X(v) \cap \mathrm{Leaf}(X)$. To do this, divide the set $N_X(v) \cap \mathrm{Leaf}(X) = A \sqcup B$ into two almost equal parts A and B with $|A| = \left\lfloor \frac{d_X(v)-1}{2} \right\rfloor$ and $|B| = \left\lceil \frac{d_X(v)-1}{2} \right\rceil$. Since $v \in \mathrm{Leaf}(X - \mathrm{Leaf}(X))$, there is a unique vertex $w \in V(X')$ with $vw \in E(X')$. If the edge vw is oriented in T' as $w \to v$, then for every $a \in A$ orient each edge va as $a \to v$ and for every $b \in B$ orient each edge vb as $v \to b$ (see Figure 7). The obtained orientation T of X yields

$$\begin{aligned}
|A(L(T))| &= |A(L(T'))| + d_T^+(v) \cdot d_T^-(v) = |A(L(T'))| + (|A|+1) \cdot |B| \\
&= |A(L(T'))| + \left\lceil \frac{d_X(v)}{2} \right\rceil \cdot \left\lfloor \frac{d_X(v)}{2} \right\rfloor \\
&= \sum_{u\in V(X')} \left\lfloor \frac{d_{X'}(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_{X'}(u)}{2} \right\rceil + \left\lceil \frac{d_X(v)}{2} \right\rceil \cdot \left\lfloor \frac{d_X(v)}{2} \right\rfloor \\
&= \sum_{u\in V(X')\setminus\{v\}} \left\lfloor \frac{d_X(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_X(u)}{2} \right\rceil + \left\lceil \frac{d_X(v)}{2} \right\rceil \cdot \left\lfloor \frac{d_X(v)}{2} \right\rfloor \\
&= \sum_{u\in V(X)} \left\lfloor \frac{d_X(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_X(u)}{2} \right\rceil .
\end{aligned}$$

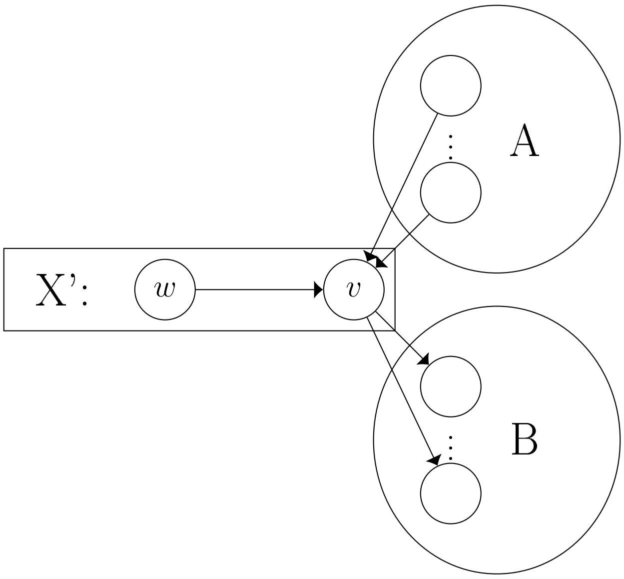

Fig. 7. An orientation T in case there is an arc $w \to v$ in T'

Similarly, if the edge vw is oriented in T' as $v \to w$, then for every $a \in A$ orient each edge va as $v \to a$ and for every $b \in B$ orient each edge vb as $b \to v$. The obtained orientation T of X also yields the desired equality $|A(L(T))| = \sum\limits_{u \in V(X)} \left\lfloor \frac{d_X(u)}{2} \right\rfloor \cdot \left\lceil \frac{d_X(u)}{2} \right\rceil$. □

It turns out that it is possible to explicitly calculate the average number of arcs in line graphs $L(T)$ among orientations T on a given tree X. To present this result, we need to recall the definition of one well-known topological index. Namely, for a graph G, the number $M_1(G) = \sum\limits_{x \in V(G)} d_G^2(x)$ is called the *first Zagreb index* of G.

Proposition 3.17. *Let X be a tree with n vertices. Then the average number of arcs in $L(T)$ among orientations T of X equals*

$$\frac{M_1(X)}{4} - \frac{n-1}{2}.$$

Proof. We use the standard counting argument. Denote by $\mathcal{T}(X)$ the class of all orientations of X (i.e., polytrees T with $[T] = X$). It is clear that $|\mathcal{T}(X)| = 2^{n-1}$. In order to calculate the sum $\sum_{T \in \mathcal{T}(X)} |A(L(T))|$, for a fixed possible arc (e_1, e_2), we calculate the number of trees $T \in \mathcal{T}(X)$ with $(e_1, e_2) \in A(L(T))$. Indeed, to be an arc in some $L(T)$, the pair (e_1, e_2) must consist of two adjacent edges e_1, e_2 in X. It is clear that the number of such pairs equals the number of 2-paths on X. Every such path is encoded by its middle vertex, say $x \in V(X)$, and an ordered pair (u, v) of its neighbors $u, v \in N_X(x)$. Hence, the number of such pairs equals

$$\sum_{x \in V(X)} d_X(x)(d_X(x) - 1) = M_1(X) - 2(n-1).$$

Further, given an arc (e_1, e_2), we have exactly $2^{n-1-2} = 2^{n-3}$ trees $T \in \mathcal{T}(X)$ with $(e_1, e_2) \in A(L(T))$. Therefore, the average number of arcs in $L(T)$ among orientations T on X equals

$$\frac{(M_1(X) - 2(n-1)) \cdot 2^{n-3}}{2^{n-1}} = \frac{M_1(X)}{4} - \frac{n-1}{2}.$$

□

Corollary 3.18. *Let X be a tree with n vertices. Then the average number of arcs in $L(T)$ among orientations T of X is at least $\frac{n-2}{2}$ and at most $\frac{(n-1)(n-2)}{4}$.*

Proof. These bounds follow from the corresponding bounds on $M_1(X)$ for trees. Namely, by [6], for an n-vertex tree X it holds that $M_1(P_n) \leq M_1(X) \leq M_1(K_{1,n-1})$. Given that $M_1(P_n) = 4n - 6$ and $M_1(K_{1,n-1}) = n(n-1)$, the result immediately follows from Proposition 3.17. □

3.3. The number of weak components in the line graph of a polytree and an algorithm for finding them all

Denote by $\mathrm{Si}(D)$ and $\mathrm{So}(D)$ the sets of sinks and sources in a digraph D, respectively.

Proposition 3.19. *Let T be a polytree with $n \geq 2$. Then the number of weak components in $L(T)$ equals*

$$\sum_{u \in \mathrm{Si}(T) \cup \mathrm{So}(T)} (d_{[T]}(u) - 1) + 1.$$

Proof. Denote the number of weak components in a digraph D as $k(D)$. We use induction on $|\mathrm{Si}(T) \cup \mathrm{So}(T)| \geq |\mathrm{Leaf}([T])|$. If $|\mathrm{Si}(T) \cup \mathrm{So}(T)| = |\mathrm{Leaf}([T])|$, then by Proposition 3.8, $L(T)$ is weak implying that $k(L(T)) = 1 = \sum_{u \in \mathrm{Si}(T) \cup \mathrm{So}(T)} (d_{[T]}(u) - 1) + 1$.

Now assume that T has a sink or a source which is a non-leaf vertex in $[T]$. Without loss of generality, assume that $x \in \mathrm{So}(T) \backslash \mathrm{Leaf}([T])$. Delete x from T and add new vertices x_u with new arcs $x_u \to u$ for all out-neighbors $u \in N_T^+(x)$. Denote the obtained digraph by T'. It is clear that T' has exactly $d = d_T^+(x)$ weak components $T_1, \ldots, T_d$, each being a polytree with fewer sinks and sources than T. By induction assumption, $k(L(T_i)) = \sum_{u \in \mathrm{Si}(T_i) \cup \mathrm{So}(T_i)} (d_{[T_i]}(u) - 1) + 1$. Since x is a source in T, it is pretty straightforward to see that $k(L(T)) = \sum_{i=1}^{d} k(L(T_i))$. Also,

$$\mathrm{Si}(T) \cup \mathrm{So}(T) = \left(\{x\} \cup \left(\bigcup_{i=1}^{d} (\mathrm{Si}(T_i) \cup \mathrm{So}(T_i)) \right) \right) \backslash \{x_u : u \in N_T^+(x)\}.$$

Therefore,

$$k(L(T)) = \sum_{i=1}^{d} k(L(T_i)) = \sum_{i=1}^{d} \left(\sum_{u \in \mathrm{Si}(T_i) \cup \mathrm{So}(T_i)} (d_{[T_i]}(u) - 1) + 1 \right)$$

$$= \sum_{u \in \bigcup_{i=1}^{d} (\mathrm{Si}(T_i) \cup \mathrm{So}(T_i))} (d_{[T_i]}(u) - 1) + d$$

$$= \sum_{u \in \left(\bigcup_{i=1}^{d} (\mathrm{Si}(T_i) \cup \mathrm{So}(T_i))\right) \setminus \{x_u : u \in N_T^+(x)\}} (d_{[T_i]}(u) - 1) + d$$

$$= \sum_{u \in (\mathrm{Si}(T) \cup \mathrm{So}(T)) \setminus \{x\}} (d_{[T]}(u) - 1) + (d-1) + 1 = \sum_{u \in \mathrm{Si}(T) \cup \mathrm{So}(T)} (d_{[T]}(u) - 1) + 1.$$

□

Corollary 3.20. *Let X be a tree. Then the average number of weak components in $L(T)$ among orientations T of X equals*

$$\sum_{u \in V(X)} \frac{d_X(u) - 1}{2^{d_X(u)-1}} + 1.$$

Proof. Let $|V(X)| = n$. Recall that by $\mathcal{T}(X)$ we denote the class of all orientations of X. By Proposition 3.19,

$$\sum_{T \in \mathcal{T}(X)} k(L(T)) = \sum_{T \in \mathcal{T}(X)} \left(\sum_{u \in \mathrm{Si}(T) \cup \mathrm{So}(T)} (d_X(u) - 1) + 1 \right)$$
$$= \sum_{T \in \mathcal{T}(X)} \sum_{u \in \mathrm{Si}(T) \cup \mathrm{So}(T)} (d_X(u) - 1) + 2^{n-1}.$$

For every vertex $u \in V(X)$ there are $2 \cdot 2^{n-1-d_X(u)} = 2^{n-d_X(u)}$ orientations T of X in which u is a sink or a source. Therefore,

$$\sum_{T \in \mathcal{T}(X)} k(L(T)) = \sum_{u \in V(X)} (d_X(u) - 1) 2^{n-d_X(u)} + 2^{n-1}.$$

This means that the average number of weak components in $L(T)$ among orientations T of X equals

$$\frac{1}{|\mathcal{T}(X)|} \sum_{T \in \mathcal{T}(X)} k(L(T)) = \frac{1}{2^{n-1}} \left(\sum_{u \in V(X)} (d_X(u) - 1) 2^{n-d_X(u)} + 2^{n-1} \right)$$
$$= \sum_{u \in V(X)} \frac{d_X(u) - 1}{2^{d_X(u)-1}} + 1.$$

□

Also, note that Corollary 3.20 implies that for a given tree X among its orientations T, the line digraph $L(T)$ on average has at most $\frac{1}{2}(|V(X)| - |\mathrm{Leaf}(X)|) + 1$ weak components.

In what follows, we are going to provide an algorithm that for a given polytree T constructs its induced subgraphs that become weak components in $L(T)$.

Algorithm 1 Construct induced subgraphs of a polytree T which become weak components in $L(T)$

Require: Polytree T

Ensure: Set $\mathcal{C}$ of induced subgraphs of T which are mapped to different weak components when applying line digraph operator

```
1: C ← ∅.
2: A_out ← the set of out-arcs from sources of T.
3: while A_out ≠ ∅ do
4:     (a, b) ← an out-arc from A_out.
5:     S ← T[{a, b}].
6:     S_prev ← null digraph.
7:     while S_prev ≠ S do
8:         S_prev ← S.
9:         S ← T[{α ∈ A(T) | α is connected with some arc from S}].
10:    end while
11:    C ← C ∪ {S}.
12:    A_out ← A_out \ A(S).
13: end while
```

Remark 3.21. After every iteration of the **while** cycle (lines 7–10) we either have added new arcs to S, or $S = S_{prev}$ and we terminate. Since the total number of arcs in T is finite, this process cannot continue indefinitely.

As for the main **while** cycle (lines 3–13), every iteration of it begins with extracting an arc from A_{out} and ends with deleting some arcs from A_{out}, including the one extracted. Thus the algorithm performs at most as many iterations as there are elements of A_{out}, which proves its termination.

Example 3.22. Consider the polytree T from Figure 8. Let us apply Algorithm 1 to it.

Initially A_{out} contains four arcs: a, d, h, i. We shall pick arc a. Hence, $S = T[a]$ as we enter the inner **while** cycle. On each iteration of it, we add all the arcs connected with any of the arcs already in S, obtaining: $A(S) = \{a\} \to \{a, b, c, e, f\} \to \{a, b, c, d, e, f, g, h\}$. There is no arc we can add to that, so the inner **while** cycle terminates and we add the newly built subgraph to $\mathcal{C}$.

Now A_{out} contains a single arc i. The inner cycle shall go $A(S) = \{i\} \to \{i, j, k\}$. After that we add S to $\mathcal{C}$. This way we divided T into two subgraphs which are mapped to different weak components of $L(T)$.

Lemma 3.23. *Every subgraph $C_1 \in \mathcal{C}$ generated by Algorithm 1 is weak.*

Proof. As shown in remark above, C_1 is generated by accumulating arcs through consecutive iterations of the inner **while** cycle. Line 9 of the algorithm implies that for all $k > 1$ an arc generated on iteration k is connected to some arc generated on step $k - 1$. Given that on step 1 the cycle C_1 is connected, the fact that after terminating the inner

while cycle C_1 is still weak follows easily by induction. □

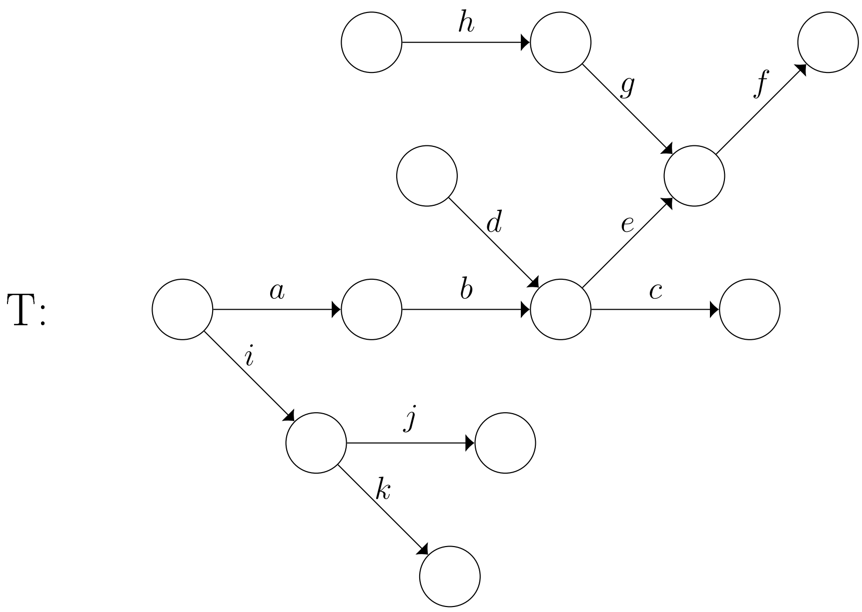

Fig. 8. A polytree T from Example 3.22

The next theorem justifies Algorithm 1.

Theorem 3.24. *Let $C_1, C_2, \ldots, C_n$ be subgraphs obtained by applying Algorithm 1 to the polytree T. Then the next statements hold:*

1. *$L(C_k)$ is a weak digraph for all $k \in [1, n]$.*
2. *$L(C_k \cup \{\alpha\})$ is not weak, for all $k \in [1, n]$ and for every $\alpha \in A(T) \backslash A(C_k)$.*
3. *$A(C_1) \sqcup A(C_2) \sqcup \ldots \sqcup A(C_n) = A(T)$.*

Proof. 1. Statement *1* can be proved by induction similar to the one used in Lemma 3.23. On the first iteration, C_k consists of a single arc, therefore clearly $L(C_k)$ is weak. Let then α be an arc added to C_k on iteration $i \geq 2$. There exists an arc β already in C_k such that α is connected with β by a dipath P. But $L(P)$ is also a dipath, α and β being its end-vertices. This means that they belong to the same weak component in $L(C_k)$. However, by induction assumption there has only been one weak component so far; thus $L(C_k)$ is weak.

2. If $\alpha \notin A(C_k)$, then α isn't connected with any arc from C_k (if it were connected with an arc β, it would be generated on the next iteration after β and therefore belong to C_k), and so the vertex of $L(T)$ associated with it will lie in different weak component than $L(C_k)$. This completes the proof of statement *2*.

3. Finally, we shall prove that all sets $A(C_i)$ are pairwise disjoint. Let $A(C_1) \cap A(C_2) = S$. Lemma 3.23 implies that C_1 is weak, thus there exists $a \in V(C_1)$ such that a is incident with arcs $\alpha \in S$ and $\beta \in C_1 \backslash S$. From statement *1* together with Proposition 3.8 we have that a is not a source nor a sink, hence without loss of generality, we assume a is incident from α to β. This implies that there is a dipath containing both α and β, namely, $\alpha - \beta$

is a dipath. Thus, since α is in C_2 ($S \subset A(C_2)$), $\beta \in C_2$ as well – it will be generated on the next iteration after α. A contradiction, which implies $A(C_1) \cap A(C_2) = \emptyset$.

To prove statement *3* we still need to show that for all arcs $\alpha \in A(T)$ there is $k \in \mathbb{N}$ with $\alpha \in A(C_k)$. If α is an out-arc from a source, then the algorithm doesn't terminate until it is marked as visited, that is, it belongs to some C_k. Otherwise, we assert that α is connected by a dipath to some out-arc from a source: as α is not an out-arc from a source, there is an arc β_1 which is adjacent to α. Now we have that either β_1 is an out-arc from a source, or there is an arc β_2 adjacent to β_1. The process terminates when we reach an arc β_m that is an out-arc from a source (it cannot continue indefinitely, since the total number of arcs in T is finite). Thus α will be generated no later than on the next iteration after β_m. $\square$

Acknowledgments

The authors express their gratitude to Y.-L. Dekhtiar, who carefully read the draft version of the paper and made suggestions that helped improve the clarity of its presentation.

References

[1] J. S. Bagga and L. W. Beineke. A survey of line digraphs and generalizations. *Discrete Mathematics Letters*, 6:68–83, 2021. 10.47443/dml.2021.s109.

[2] L. W. Beineke. Characterizations of derived graphs. *Journal of Combinatorial Theory*, 9:129–135, 1970. https://doi.org/10.1016/S0021-9800(70)80019-9.

[3] S. Benzer. On the topology of the genetic fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, 45(11):1607–1620, 1959. https://doi.org/10.1073/pnas.45.11.1607.

[4] S. Das, M. Sen, A. B. Roy, and D. B. West. Interval digraphs: an analogue of interval graphs. *Journal Graph Theory*, 13(2):189–202, 1989. https://doi.org/10.1002/jgt.3190130206.

[5] B.-Y. Dekhtiar and S. Kozerenko. Line digraphs of polytrees and their weak components. In *International Conference of Young Mathematicians*, Kyiv, Ukraine, June 2023.

[6] I. Gutman. Multiplicative zagreb indices of trees. *Bulletin of the International Mathematical Virtual Institute*, 1:13–19, 2011.

[7] F. Harary. A characterization of block graphs. *Canadian Mathematical Bulletin*, 6:1–6, 1963. https://doi.org/10.4153/CMB-1963-001-x.

[8] F. Harary. *Graph Theory*. Addison-Wesley, Reading, Mass., 1969, page 274.

[9] F. Harary and R. Z. Norman. Some properties of line digraphs. *Rendiconti del Circolo Matematico di Palermo*, 9:161–168, 1960. https://doi.org/10.1007/BF02854581.

[10] C. G. Lekkerkerker and J. C. Bohland. Representation of a finite graph by a set of intervals on the real line. *Fundamenta Mathematicae*, 51:45–64, 1962.

[11] F. S. Roberts and J. H. Spencer. A characterization of clique graphs. *Journal of Combinatorial Theory, Series B*, 10:102–108, 1971. https://doi.org/10.1016/0095-8956(71)90070-0.

[12] J. Spinrad. Recognition of circle graphs. *Journal of Algorithms*, 16:264–282, 1994. https://doi.org/10.1006/jagm.1994.1012.

[13] H. Whitney. Congruent graphs and the connectivity of graphs. *American Journal of Mathematics*, 54(1):150–168, 1932. https://doi.org/10.2307/2371086.

Sergiy Kozerenko
Graph Theory and Network Analysis Laboratory, Kyiv School of Economics
Mykoly Shpaka str. 3, 03113 Kyiv, Ukraine

Bohdan-Yarema Dekhtiar
Graph Theory and Network Analysis Laboratory, Kyiv School of Economics
Mykoly Shpaka str. 3, 03113 Kyiv, Ukraine