INDUCIBILITY OF TOPOLOGICAL TREES

AUDACE A. V. DOSSOU-OLORY AND STEPHAN WAGNER

ABSTRACT. Trees without vertices of degree 2 are sometimes named topological trees. In this work, we bring forward the study of the inducibility of (rooted) topological trees with a given number of leaves. The inducibility of a topological tree S is the limit superior of the proportion of all subsets of leaves of T that induce a copy of S as the size of T grows to infinity. In particular, this relaxes the degree-restriction for the existing notion of the inducibility in d-ary trees. We discuss some of the properties of this generalised concept and investigate its connection with the degree-restricted inducibility. In addition, we prove that stars and binary caterpillars are the only topological trees that have an inducibility of 1. We also find an explicit lower bound on the limit inferior of the proportion of all subsets of leaves of T that induce either a star or a binary caterpillar as the size of T tends to infinity.

1. INTRODUCTION AND PRELIMINARY

The parameter *inducibility* of graphs first appears in the mathematical literature in a paper of Pippenger and Golumbic [1] in 1975. In plain language, Pippenger and Golumbic studied the maximum frequency of a simple graph occurring as a subgraph induced by vertices of another simple graph with a sufficiently large number of vertices. Over the past four decades, many authors have extensively discussed the inducibility in various settings and a number of papers has been published.

The inducibility of trees, however, was first initiated much more recently in [5] by Bubeck and Linial. They studied the concept in trees with a given number of vertices. Shortly afterwards, Czabarka, Székely and the second author of the current paper introduced in [3] the inducibility in rooted binary trees with a given number of leaves and also presented an application to random tanglegrams. The recent paper [2] proposed an extension of the inducibility of binary trees to d-ary trees for every $d \geq 2$.

The object of this paper is to continue the investigation of the inducibility of trees. In particular, it is natural to consider a variant of the concept that relaxes the degree-restriction on the vertices of the tree: one might be interested in knowing how large the number of appearances of a tree in another larger tree with a given number of leaves can be.

Before we can present our results, we need to give some formal definitions.

An unrooted tree that does not contain any subdivided edge, that is, no vertices of degree 2, is called a *topological* tree (see e.g. Bergeron et al. [7] or Allman and Rhodes [8]), or also *series-reduced* or *homeomorphically irreducible* tree. Based on this definition, we shall call any rooted tree in which every vertex has outdegree at least 2 a (rooted) topological tree.

When the tree consists of a single vertex, we shall treat it as both leaf and root. Figure 1 displays all the topological trees with fewer than five leaves: the counting sequence for the number of n-leaf topological trees starts

$$1, 1, 2, 5, 12, 33, 90, 261, 766, 2312, 7068, \ldots,$$
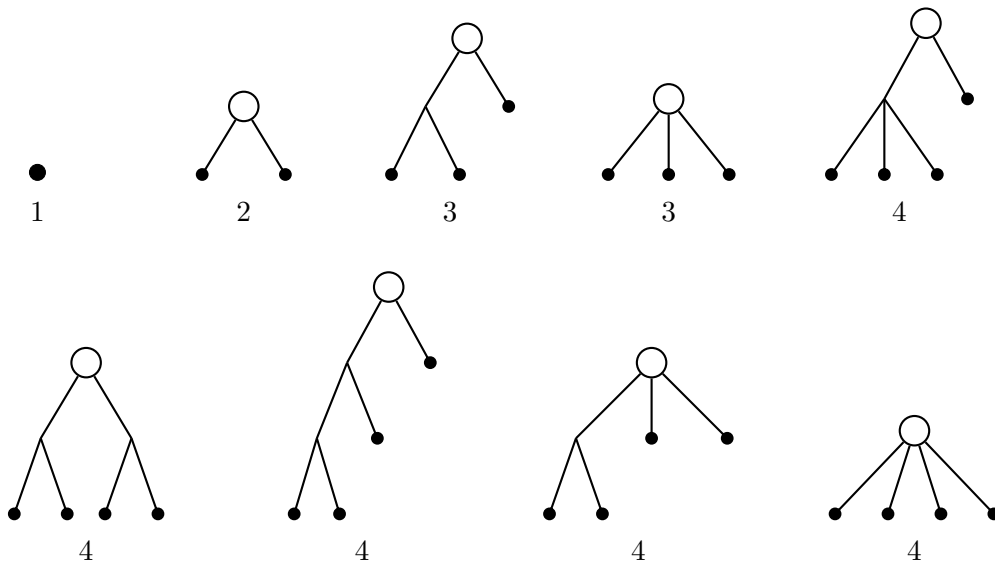
see A000669 in [6] for more information.

FIGURE 1. All the topological trees with fewer than five leaves.

Every subset L of leaves of a topological tree T induces another tree which is obtained by the following operation: first extract the smallest subtree of T that connects the leaves in L and then erase the vertices of outdegree 1 (if any). This operation is illustrated in Figure 2.

Such a subtree will be referred to as a *leaf-induced subtree* of T. It has a root in a natural way which we define to be the most recent common parent shared by the leaves in L. We shall write $|T|$ for the number of leaves of a topological tree T.

For two topological trees S, T, we denote by $c(S, T)$ the number of copies of S in T. Formally, $c(S, T)$ is the number of subsets of the leaf set of T that induce a tree isomorphic in the sense of rooted trees to S (i.e., the isomorphism maps the root of one tree to the root of the other tree). Let $\gamma(S, T) = c(S, T)/\binom{|T|}{|S|}$ be its normalised version, which lies
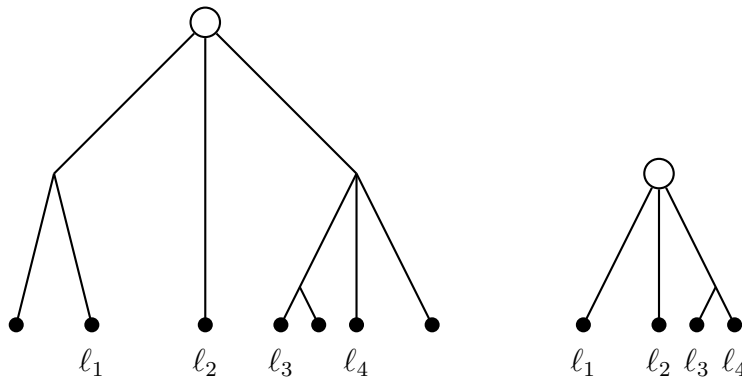
FIGURE 2. A ternary tree and the subtree induced by four leaves $\{\ell_1, \ell_2, \ell_3, \ell_4\}$.

between 0 and 1 by definition. For brevity, $\gamma(S, T)$ will be referred to as the *density* of S in T. We are specifically interested in the maximum of $\gamma(S, T)$ in the limit: the quantity

$$
(1) \qquad J(S) := \limsup_{\substack{|T| \to \infty \\ T \text{ topological tree}}} \gamma(S, T) = \limsup_{n \to \infty} \max_{\substack{|T| = n \\ T \text{ topological tree}}} \gamma(S, T),
$$

where the maximum runs over all topological trees, will be called the *inducibility of S* (in topological trees).

A topological tree with the property that every vertex has outdegree no more than $d \ (\geq 2)$ will be called a *d-ary tree*. In this note, we shall simply call a 2-ary tree a *binary tree*.

When the maximum of the density $\gamma(D, T)$ of a d-ary tree D in T is taken over all d-ary trees, we shall speak of the *inducibility in d-ary trees*. Strictly speaking, the *inducibility* $I_d(D)$ of a d-ary tree D in d-ary trees is

$$
I_d(D) = \lim_{n \to \infty} \max_{\substack{|T| = n \\ T \text{ d-ary tree}}} \gamma(D, T),
$$

as presented in a previous article [2], where the limit is shown to exist. We will prove that this is also the case for $J(S)$, see Theorem 3.

The d-ary trees that attain the maximum inducibility 1 are determined in [2]: for every $d \geq 2$, the maximal trees are so-called binary caterpillars (paths with one pendant edge dropped from all the vertices except for one of the endvertices). Things change, however, when the vertices of the tree are allowed to have any outdegree, see Theorem 4.

Clearly, for any given d-ary tree D, the inducibility $I_d(D)$ is non-decreasing with respect to d, and thus tends to a definite limit as $d \to \infty$. In the following section, we establish that $J(S) = \lim_{d \to \infty} I_d(S)$ for every topological tree S. We also prove a link between the maximum density of S in topological trees and the actual inducibility $J(S)$, and demonstrate that stars and binary caterpillars are the only topological trees with the maximum inducibility 1.

Furthermore, we find an explicit lower bound on the limit inferior of the proportion of all subsets of leaves of T that induce either a star or a binary caterpillar as the size of the topological tree T grows to infinity (see Proposition 6). In addition, we show that one can obtain a lower bound on the inducibility of any topological tree.

2. Auxiliary results

We recall the following fundamental results from [2]. We shall utilise them in various places in the note.

Theorem 1 ([2])**.** *For every fixed positive integer $d \geq 2$ and every d-ary tree D, the double inequality*

$$0 \leq \max_{\substack{|T|=n \\ T \ d\text{-ary tree}}} \gamma(D, T) - I_d(D) \leq |D|(-1 + |D|)n^{-1}$$

holds for all n.

Theorem 2 ([2])**.** *Let $d \geq 2$ be an arbitrary but fixed positive integer. Among d-ary trees, only binary caterpillars have $(d$-ary$)$ inducibility 1.*

3. Main results

As for the degree-restricted inducibility, we begin our investigation by providing an estimate on how much the general inducibility can differ from the maximum density $\gamma(S, T)$ in topological trees.

The result is an analogue of Theorem 1 for $J(S)$. It asserts that for every topological tree S, the maximum of $\gamma(S, T)$ tends to a definite limit as $|T| \to \infty$ and that the precise gap between the maximum density and the limit is of order at most $\mathcal{O}(|T|^{-1})$.

Theorem 3. *Let S be a topological tree. Then the double inequality*

$$0 \leq \max_{\substack{|T|=n \\ T \ topological \ tree}} \gamma(S, T) - J(S) \leq |S|(-1 + |S|)n^{-1}$$

is valid for all $n \geq |S|$. In particular, we have

$$J(S) = \lim_{n \to \infty} \max_{\substack{|T|=n \\ T \ topological \ tree}} \gamma(S, T).$$

Proof. We can follow the same averaging argument used in [2] to prove Theorem 1. Since we shall make heavy use of the intermediary results that appear in the proof, we present them for completeness.

Let S and T be two topological trees such that $|S| \leq |T|$. We write $L(T)$ for the set of leaves of T. For $l \in L(T)$, let us denote by $c_l(S, T)$ the number of l-containing subsets of leaves of T that induce a copy of S. Thus, since

$$\sum_{l \in L(T)} c_l(S, T) = |S| \cdot c(S, T),$$

we deduce that there exist leaves l_1 and l_2 of T (possibly $l_1 = l_2$) for which the double inequality

$$(2) \qquad c_{l_1}(S, T) \leq \frac{|S| \cdot c(S, T)}{|T|} \leq c_{l_2}(S, T)$$

is satisfied. The number of copies of S in T not involving the leaf l_1 is

$$c(S, T) - c_{l_1}(S, T) \geq \left(1 - \frac{|S|}{|T|}\right) c(S, T)$$

by relation (2). Let T^- be the topological tree that results when the leaf l_1 of T is removed and the unique vertex adjacent to l_1 is erased if it has outdegree 2 in T. We have

$$c(S, T^-) \geq \left(1 - \frac{|S|}{|T|}\right) c(S, T),$$

and dividing by $\binom{|T|-1}{|S|}$, we obtain

$$\gamma(S, T^-) \geq \gamma(S, T).$$

Since T is an arbitrary topological tree, this implies that

$$\max_{\substack{|T|=n-1 \\ T \text{ topological tree}}} \gamma(S, T) \geq \max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S, T)$$

for every $n \geq 2$. In particular, the assertion on the limit follows as the sequence

$$\left(\max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S, T) \right)_{n \geq 1}$$

is nonincreasing and bounded. Moreover, we get

$$J(S) \leq \max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S, T)$$

for all n.

Now, denote by T^+ the tree obtained after replacing the leaf l_2 of T by an internal vertex with two leaves l_2 and l_2' attached to it. So, $c(S, T)$ represents the number of copies of S in T^+ not involving l_2' (removing the leaf l_2' yields T) whereas the number of copies of S in T^+ involving l_2' is at least $c_{l_2}(S, T)$. Thus, it follows from relation (2) that

$$c(S, T^+) \geq \left(1 + \frac{|S|}{|T|}\right) c(S, T).$$

Dividing by $\binom{|T|+1}{|S|}$ yields

$$\gamma(S, T^+) \geq \left(1 - \frac{|S|(-1 + |S|)}{|T|(|T| + 1)}\right) \gamma(S, T),$$

and since T was assumed to be an arbitrary n-leaf topological tree, we have

$$\max_{\substack{|T|=n+1 \\ T \text{ topological tree}}} \gamma(S,T) \geq \left(1 - \frac{|S|(-1+|S|)}{n(n+1)}\right) \max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S,T)$$

for every n. After p iterations, we establish that

$$\max_{\substack{|T|=n+p \\ T \text{ topological tree}}} \gamma(S,T) \geq \left(\max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S,T)\right) \cdot \prod_{j=0}^{p-1} \left(1 - \frac{|S|(-1+|S|)}{(n+p-j)(n+p-j-1)}\right)$$

for all n, p with $p \geq 1$.

Letting $p \to \infty$ gives us the estimate

$$J(S) \geq \left(\max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S,T)\right) \cdot \left(1 - \sum_{i=0}^{\infty} \frac{|S|(-1+|S|)}{(n+i+1)(n+i)}\right)$$

$$= \left(\max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S,T)\right) \cdot \left(1 - \frac{|S|(-1+|S|)}{n}\right)$$

for every n, where in the first step we used the standard inequality (which is proved by a simple induction) $\prod_{i=1}^{m}(1-y_i) \geq 1 - \sum_{i=1}^{m} y_i$ (valid when $0 \leq y_i \leq 1$ for every i – see [9, p. 60]) giving us

$$\prod_{j=0}^{p-1} \left(1 - \frac{|S|(-1+|S|)}{(n+p-j)(n+p-j-1)}\right) \geq 1 - \sum_{j=0}^{p-1} \frac{|S|(-1+|S|)}{(n+p-j)(n+p-j-1)}$$

(recall that $n \geq |S|$). Putting things together, we obtain the desired double inequality and in particular, the asymptotic formula

$$\max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S,T) = J(S) + \mathcal{O}(n^{-1})$$

which is valid for all n. This completes the proof of the theorem. $\square$

The problem of determining extremal graph structures that maximise or minimise a given graph parameter is a topical subject within graph theory. In what follows, we address the question of characterising all the topological trees that have the maximal inducibility 1.

Unlike the degree-restricted inducibility $I_d(S)$, two families of topological trees are found to be maximal with respect to $J(S)$.

By a *binary caterpillar*, we mean a binary tree whose non-leaf vertices lie on a single path starting at the root. By a *star*, we mean a topological tree in which all leaves are children of the root.

An illustration of both classes of trees is given in Figure 3.

Theorem 4. *Both stars and binary caterpillars have the maximal inducibility 1. Moreover, every topological tree S that is not a star or a binary caterpillar satisfies $J(S) < 1$.*

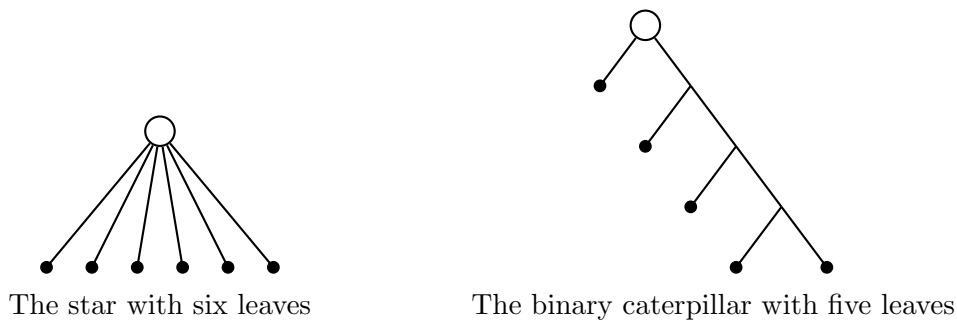The star with six leaves   The binary caterpillar with five leaves

FIGURE 3. A star and a binary caterpillar

We remark that Theorem 4 has a link to an observation made by Bubeck and Linial in [5]. In their context, the inducibility $\mathrm{ind}(R)$ of a tree R (not necessarily a topological tree) with k vertices is defined to be the limit superior of the proportion of R as a subtree among all k-vertex induced subtrees where the limit is taken over all sequences of trees whose number of vertices grows to infinity.

In the last section of their paper, Bubeck and Linial point out that stars and paths are the only trees R that have an inducibility $\mathrm{ind}(R)$ of 1.

Proof of Theorem 4. First off, note that we have both $J(C_k) = 1$ and $J(F_k^2) = 1$ for every k in view of the identity

$$c(C_k, C_n) = c(F_k^2, F_n^2) = \binom{n}{k},$$

i.e., every subset of k leaves of the star C_n induces C_k, and every subset of k leaves of the binary caterpillar F_n^2 induces F_k^2.

The rest of the proof is a refinement of the proof of Theorem 2. To begin, let us prove that for any fixed positive integers $k > 2$ and $n > (k-1)^{k-2}$, every n-leaf topological tree contains either a copy of the k-leaf star C_k or a copy of the k-leaf binary caterpillar F_k^2.

Note that a topological tree has C_k as a leaf-induced subtree if and only if it contains a vertex that has at least k children. Likewise, a topological tree contains F_k^2 as a leaf-induced subtree if and only if it has height at least $k-1$.

Fix $k > 2$, $n > (k-1)^{k-2}$, and consider a topological tree T with n leaves. If we suppose that T is a $(k-1)$-ary tree and T has height at most $k-2$, then T must have at most $(k-1)^{k-2}$ leaves because it is easy to see that for any fixed integers $h \geq 0$ and $m \geq 2$, an m-ary ($m \geq 2$) tree with height at most h can never have more than m^h leaves. This contradicts our assumption that $|T| = n > (k-1)^{k-2}$.

Therefore, for $k > 2$ and $n > (k-1)^{k-2}$, every n-leaf topological tree must contain either a copy of the star C_k or a copy of the binary caterpillar F_k^2.

As a next step, if S is neither $F_{|S|}^2$ nor $C_{|S|}$, then by the above discussion, we obtain $c(S, T) < \binom{|T|}{|S|}$ for every topological tree with n leaves as soon as $n > (k-1)^{k-2}$. Therefore,

fixing $n > (k-1)^{k-2}$, we get

$$\max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S,T) < 1\,.$$

Furthermore, we know from the proof of Theorem 3 that

$$J(S) \leq \max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S,T)$$

for every n: this finishes the proof of the theorem. $\qquad\square$

In particular, one obtains:

Corollary 5. *We have*

$$\liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \gamma(S,T) = 0$$

for every topological tree S with at least three leaves.

Proof. This is a consequence of the identities $J(C_k) = 1$ and $J(F_k^2) = 1$ because

$$\liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \gamma(S,T) \leq \liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \left(1 - \gamma(F_{|S|}^2, T)\right) = 1 - J(F_{|S|}^2)$$

for every topological tree S different from a binary caterpillar, and likewise

$$\liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \gamma(S,T) \leq \liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \left(1 - \gamma(C_{|S|}, T)\right) = 1 - J(C_{|S|})$$

for every topological tree S different from a star. $\qquad\square$

As one might expect (in view of the aforementioned analogy to the work [5]), the quantity

$$\liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \left(\gamma(C_k, T) + \gamma(F_k^2, T)\right)$$

is positive for every k. It seems arduous to determine its explicit value as a function of k. Nevertheless, we are able to say more about this quantity.

In fact, Bubeck and Linial [5] proved that in the context of induced subtrees of trees, the sum of the proportions of the k-vertex path and the k-vertex star is always greater than zero in the limit for every k. Specifically, they showed that the sum of the two proportions is bounded from below by an explicit constant that depends solely on k.

Proposition 6. *For every fixed positive integer $k \geq 2$, the inequalities*

$$\liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \left(\gamma(C_k, T) + \gamma(F_k^2, T)\right) \geq \frac{1}{\binom{1+(k-1)^{k-2}}{k}} \geq \frac{k!}{\left(1+(k-1)^{k-2}\right)^k}$$

are satisfied.

Proof. Let $k \geq 2$ be an arbitrary but fixed positive integer. In analogy to the proof of Theorem 3, we have

$$\sum_{l \in L(T)} \left(c_l(F_k^2, T) + c_l(C_k, T) \right) = k \left(c(F_k^2, T) + c(C_k, T) \right)$$

for every topological tree T. So there exists $l_1 \in L(T)$ such that the inequality

$$c_{l_1}(F_k^2, T) + c_{l_1}(C_k, T) \geq \frac{k}{|T|} \cdot \left(c(F_k^2, T) + c(C_k, T) \right)$$

holds. If we denote by T^- the topological tree obtained when removing the leaf l_1 of T as well as erasing the unique vertex adjacent to l_1 in T if it has outdegree 2 in T, then this implies that

$$c(F_k^2, T^-) + c(C_k, T^-) = \left(c(F_k^2, T) + c(C_k, T) \right) - \left(c_{l_1}(F_k^2, T) + c_{l_1}(C_k, T) \right)$$
$$\leq \left(1 - \frac{k}{|T|} \right) \cdot \left(c(F_k^2, T) + c(C_k, T) \right),$$

and dividing by $\binom{|T|-1}{k}$ yields

$$\gamma(F_k^2, T^-) + \gamma(C_k, T^-) \leq \gamma(F_k^2, T) + \gamma(C_k, T)$$

for every topological tree T with $n \geq 2$ leaves. Since T is arbitrary, it follows that

$$\min_{\substack{|T|=n-1 \\ T \text{ topological tree}}} \left(\gamma(F_k^2, T) + \gamma(C_k, T) \right) \leq \min_{\substack{|T|=n \\ T \text{ topological tree}}} \left(\gamma(F_k^2, T) + \gamma(C_k, T) \right),$$

so the sequence

$$\left(\min_{\substack{|T|=n \\ T \text{ topological tree}}} \left(\gamma(C_k, T) + \gamma(F_k^2, T) \right) \right)_{n \geq k}$$

is nondecreasing. Using this result, we derive that

$$\min_{\substack{|T|=n \\ T \text{ topological tree}}} \left(\gamma(C_k, T) + \gamma(F_k^2, T) \right) \geq \min_{\substack{|T|=1+(k-1)^{k-2} \\ T \text{ topological tree}}} \left(\gamma(C_k, T) + \gamma(F_k^2, T) \right)$$

for every $n > (k-1)^{k-2}$.

It follows that

$$\liminf_{\substack{|T| \to \infty \\ T \text{ topological tree}}} \left(\gamma(C_k, T) + \gamma(F_k^2, T) \right) \geq \min_{\substack{|T|=1+(k-1)^{k-2} \\ T \text{ topological tree}}} \left(\gamma(C_k, T) + \gamma(F_k^2, T) \right).$$

Furthermore, we know from the proof of Theorem 4 that

$$\gamma(C_k, T) + \gamma(F_k^2, T) \geq \frac{1}{\binom{|T|}{k}}$$

as soon as $|T| > (k-1)^{k-2}$. Hence, we obtain

$$\liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \big(\gamma(C_k, T) + \gamma(F_k^2, T)\big) \geq \frac{1}{\binom{1 + (k-1)^{k-2}}{k}} \, .$$

Consequently, this finishes the proof of the proposition. $\square$

We would very much welcome seeing a solution to the following question:

Question 1. What is the precise value of

$$\liminf_{\substack{|T|\to\infty \\ T \text{ topological tree}}} \big(\gamma(C_k, T) + \gamma(F_k^2, T)\big)$$

when $k > 3$?

Our next result gives a simple identity involving $J(S)$ and its d-ary counterpart $I_d(S)$. Specifically, although using equation (1) can be very difficult, we discover that to compute the inducibility $J(S)$, it is enough to properly understand the inducibility $I_d(S)$ as a single variable function of d.

Theorem 7. *We have*

$$J(S) = \lim_{d\to\infty} I_d(S)$$

for every topological tree S.

Proof. Let $\epsilon > 0$ be an arbitrary but fixed positive real number. We shall make use of the two estimates

$$(3) \qquad\qquad J(S) \leq \max_{\substack{|T|=n \\ T \text{ topological tree}}} \gamma(S, T)$$

and

$$I_d(S) \geq \left(1 - \frac{|S|(-1 + |S|)}{n}\right) \max_{\substack{|T|=n \\ T \; d-\text{ary tree}}} \gamma(S, T)$$

both valid for every n. They follow from the proof of Theorem 3 and Theorem 1, respectively.

For any positive integer $m_\epsilon > |S|(-1 + |S|)/\epsilon$, we have

$$\frac{|S|(-1 + |S|)}{m_\epsilon} \cdot \max_{\substack{|T|=m_\epsilon \\ T \; m_\epsilon-\text{ary tree}}} \gamma(S, T) < \epsilon \, ,$$

which implies

$$(4) \qquad \max_{\substack{|T|=m_\epsilon \\ T \text{ topological tree}}} \gamma(S, T) - \left(1 - \frac{|S|(-1 + |S|)}{m_\epsilon}\right) \max_{\substack{|T|=m_\epsilon \\ T \; m_\epsilon-\text{ary tree}}} \gamma(S, T) < \epsilon \, .$$

Employing the relation

$$I_{m_\epsilon}(S) \geq \left(1 - \frac{|S|(-1 + |S|)}{m_\epsilon}\right) \max_{\substack{|T| = m_\epsilon \\ T \ m_\epsilon-\text{ary tree}}} \gamma(S, T),$$

and invoking the fact that $\left(I_m(S)\right)_{m \geq 2}$ is a nondecreasing sequence of real numbers, inequality (4) implies

$$\left(\max_{\substack{|T| = m_\epsilon \\ T \text{ topological tree}}} \gamma(S, T)\right) - I_m(S) < \epsilon,$$

for every $m \geq m_\epsilon$.

Combining this with (3), we establish that

$$J(S) - I_m(S) < \epsilon$$

for every $m \geq m_\epsilon$. On the other hand, we have

$$J(S) \geq \sup_{d \geq 2} I_d(S) = \lim_{d \to \infty} I_d(S)$$

by definition of $J(S)$. Hence, we conclude that for every $\epsilon > 0$, there exists $\Delta \geq 2$ such that the inequality

$$|J(S) - I_d(S)| < \epsilon$$

holds for every $d > \Delta$. This completes the proof of the theorem. $\qquad \square$

Remark 1. For a topological tree T, let us denote by $\Delta(T)$ the maximum among the outdegrees of the vertices of T. Let $T_1^\bullet, T_2^\bullet, \ldots$ be a sequence of topological trees such that $|T_n^\bullet| \to \infty$ as $n \to \infty$ and

$$J(S) = \lim_{n \to \infty} \gamma(S, T_n^\bullet).$$

If $\max_{i \geq 1} \Delta(T_i^\bullet)$ exists, then the sequence $\left(I_d(S)\right)_{d \geq 2}$ becomes constant from the point $\Delta = \max_{i \geq 1} \Delta(T_i^\bullet)$ onwards, that is,

$$I_\Delta(S) = I_{1+\Delta}(S) = I_{2+\Delta}(S) = \cdots.$$

As already mentioned earlier, Theorem 7 can be very useful. The following corollary reveals one of its features:

Corollary 8. *We have*

$$J(B) = I_2(B)$$

for every binary tree B.

Proof. It is shown in [2] that $I_d(B) = I_2(B)$ for every d and binary tree B. Therefore, by Theorem 7, we obtain $J(B) = I_2(B)$. $\qquad \square$

For the next corollary, we first need to recall the definition of a complete d-ary tree.

The *complete d-ary tree* CD_h^d of height h is the d-ary tree defined recursively in the following way:

- CD_0^d is the single leaf;
- for $h > 0$, CD_h^d has d branches isomorphic to the tree CD_{h-1}^d.

Corollary 9. *There are infinitely many topological trees S with $J(S) \leq \epsilon$ for every $\epsilon > 0$.*

Proof. It is sufficient to consider binary trees. By a result that appears in [4], we know that the inducibility $I_2(CD_h^2)$ of the complete binary tree CD_h^2 of height h is at most $I_2(CD_2^2)^{2^{h-2}}$ for all $h \geq 2$. Therefore, since it is also proved in the same paper [4] (see also [3]) that $I_2(CD_2^2) = 3/7$, we deduce that

$$I_2(CD_h^2) \leq I_2(CD_2^2)^{2^{h-2}} = \left(\frac{3}{7}\right)^{2^{h-2}} \to 0 \text{ as } h \to \infty \,.$$

Thus, the assertion of the corollary follows from Corollary 8. $\qquad\square$

We close this section by collecting two general facts about the inducibility $J(S)$. We note that some results in [2] with respect to $I_d(S)$ hold analogously for $J(S)$ using our Theorem 7.

Note that $J(S) > 0$ for every topological tree S because by definition, we have $J(S) \geq I_{\Delta(S)}(S)$ where $\Delta(S)$ is the maximum outdegree among the vertices of S while it is proved in [2] that $I_{\Delta(S)}(S) > 0$. The following proposition provides a better lower bound that only depends on $|S|$.

Proposition 10. *We have*

$$J(S) \geq \frac{(-1 + |S|)!}{-1 + |S|^{-1+|S|}}$$

for every topological tree S with at least two leaves.

Proof. Consider a topological tree S. By a result in [2], we have

$$I_d(S) \geq \frac{(-1 + |S|)!}{-1 + |S|^{-1+|S|}}$$

for every $d \geq \Delta(S)$. Passing to the limit as $d \to \infty$, Theorem 7 gives us the desired statement. $\qquad\square$

For two topological trees S_1, S_2, denote by $\mathcal{F}(S_1; S_2)$ the unique topological tree which is constructed by appending the root of S_2 to every leaf of S_1. See Figure 4 for a diagram.

Theorem 11. *The tree $\mathcal{F}(S_1; S_2)$ satisfies*

$$J\big(\mathcal{F}(S_1; S_2)\big) \geq \frac{(|S_1| \cdot |S_2|)!}{(|S_2|!)^{|S_1|} \cdot |S_1|^{|S_1| \cdot |S_2|}} \cdot J(S_2)^{|S_1|}$$
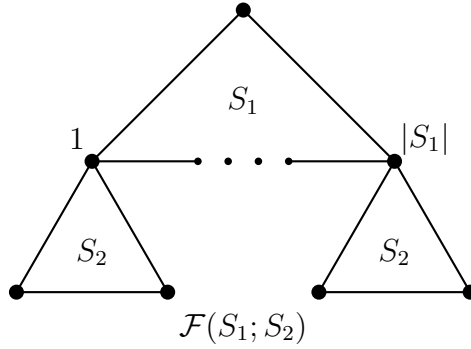
for every pair of topological trees S_1 and S_2.

FIGURE 4. A rough picture of the tree $\mathcal{F}(S_1; S_2)$ defined for Theorem 11.

Proof. Again, by a result in [2], we have

$$I_d\big(\mathcal{F}(S_1; S_2)\big) \geq \frac{(|S_1| \cdot |S_2|)!}{(|S_2|!)^{|S_1|} \cdot |S_1|^{|S_1| \cdot |S_2|}} \cdot I_d(S_2)^{|S_1|}$$

for every $d \geq \Delta\big(\mathcal{F}(S_1; S_2)\big)$. Taking the limit of both sides as $d \to \infty$, and invoking Theorem 7, we obtain the desired inequality of the theorem. □

It turns out that if one knows the inducibility of a topological tree S, then one can obtain a lower bound on the inducibility of any of the leaf-induced subtrees of S:

Theorem 12. *For any three topological trees R, S and T such that $|T| \geq |S| \geq |R|$, we have*

$$c(R, T) \geq \frac{c(R, S)}{\binom{-|R| + |T|}{-|R| + |S|}} \cdot c(S, T) \,.$$

In particular, we obtain

$$J(R) \geq J(S) \cdot \gamma(R, S) \,.$$

Proof. Given a topological tree S, we can count the number of appearances of any smaller topological tree R in a larger topological tree T by first counting the number of copies of S in T and then the number of copies of R in S. Clearly, $c(R, T)$ is overcounted in this way based on the observation that the intersection of the subsets of leaves of T that induce a copy of S may not be empty. So we would like to take this observation into account.

Assume $|T| \geq |S| \geq |R|$. Given a subset L of leaves of T that induces a copy of R, we can then choose any subset L' of $|S| - |R|$ leaves from the set of leaves of T without L so that $L \cup L'$ induces a subtree of T with $|S|$ leaves. Since the subtree induced by the subset $L \cup L'$ of leaves of T has $|S|$ leaves, we deduce that the quantity $c(S, T) \cdot c(R, S)$ is at most

$\binom{|T|-|R|}{|S|-|R|} \cdot c(R,T)$. Consequently, one obtains a simple lower bound on $c(R,T)$, namely

$$c(R,T) \geq \frac{c(S,T) \cdot c(R,S)}{\binom{|T|-|R|}{|S|-|R|}} \,.$$

As a next step, we take the density:

$$\gamma(R,T) \geq \gamma(R,S) \cdot \frac{\binom{|S|}{|R|} \cdot \binom{|T|}{|S|}}{\binom{|T|}{|R|} \cdot \binom{|T|-|R|}{|S|-|R|}} \cdot \gamma(S,T)$$
$$= \gamma(R,S) \cdot \gamma(S,T) \,.$$

Finally, in view of Theorem 3, we can take the limit to obtain the desired result.   □

Remark 2. Let $d \geq 2$ be a fixed positive integer. By the same argument as in the proof of Theorem 12, it is also seen that

$$I_d(D_1) \geq I_d(D_2) \cdot \gamma(D_1, D_2)$$

for any two d-ary tree D_1 and D_2 satisfying $|D_1| \leq |D_2|$.

4. Concluding comments

It would be interesting to answer the following question:

Question 2. Can we explicitly determine the inducibility $J(S)$ of any topological tree S other than a star or a binary tree? Note that in the degree-restricted context, the answer to this question is affirmative [4].

Finally, we conjecture that equality never holds in Proposition 10. It is also natural to formulate the following problem:

Question 3. Is there a sequence $(T_n)_{n \geq 1}$ of topological trees such that $|T_n| = n$ and $\lim_{n \to \infty} \gamma(S, T_n)$ exists and is positive for all topological trees S? In the degree-restricted context, the answer is positive [4].

References

[1] Nicholas Pippenger and Martin C. Golumbic. The inducibility of graphs. *J. Combinatorial Theory Ser. B*, 19(3):189–203, 1975.

[2] Éva Czabarka, Audace A. V. Dossou-Olory, László A. Székely, and Stephan Wagner. Inducibility of d-ary trees. Preprint, 2018. http://arxiv.org/abs/1802.03817

[3] Éva Czabarka, László A. Székely, and Stephan Wagner. Inducibility in binary trees and crossings in random tanglegrams. *SIAM J. Discrete Mathematics*, 31(3):1732–1750, 2017.

[4] Audace A. V. Dossou-Olory and Stephan Wagner. Further results on the inducibility of d-ary trees. In preparation.

[5] Sébastien Bubeck and Nati Linial. On the local profiles of trees. *J. Graph Theory*, 81(2):109–119, 2016.

[6] The On-Line Encyclopedia of Integer Sequences. Published electronically at https://oeis.org, 2018.

[7] F. Bergeron, G. Labelle, and P. Leroux. Combinatorial species and tree-like structures. *Cambridge Univ. Press*, 67, 1998.

[8] E. S. Allman and J. A. Rhodes. Mathematical models in biology: an introduction. *Cambridge Univ. Press*, 2004.

[9] Godfrey H. Hardy, John E. Littlewood, and George Pólya. *Inequalities*. Cambridge university press, 1952.

Audace A. V. Dossou-Olory and Stephan Wagner, Department of Mathematical Sciences, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

E-mail address: {audaced,swagner}@sun.ac.za