Community Heroes and Sleeping Members
Interdependency of the Tenets of Energy Justice

Astola, Mandi; Laes, Erik; Bombaerts, Gunter; Ryszawska, Bozena; Rozwadowska, Magdalena; Szymanski, Piotr; Ruess, Anja; Nyborg, Sophie; Hansen, Meiken

DOI
10.1007/s11948-022-00384-3

Publication date
2022

Document Version
Final published version

Published in
Science and Engineering Ethics

Applied Mathematics and Computation 437 (2023) 127533

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

The effect of graph operations on the degree-based entropy[☆]

Yanni Dong[a,b,c], Hajo Broersma[c,*], Changwu Song[d], Pengfei Wan[e], Shenggui Zhang[a,b]

[a] School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, PR China
[b] Xi'an-Budapest Joint Research Center for Combinatorics, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, PR China
[c] Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, PO Box 217, AE Enschede 7500, The Netherlands
[d] Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, PO Box 5, AA Delft 2600, The Netherlands
[e] School of Mathematics and Statistics, Yulin University, Yulin, Shaanxi 719000, PR China

ARTICLE INFO

Article history:
Received 3 May 2022
Revised 31 August 2022
Accepted 3 September 2022

Keywords:
Graph entropy
Graph operation
Degree
Tree
Unicyclic graph

ABSTRACT

The degree-based entropy $I_d(G)$ of a graph G on $m > 0$ edges is obtained from the well-known Shannon entropy $-\sum_{i=1}^{n} p(x_i)\log p(x_i)$ in information theory by replacing the probabilities $p(x_i)$ by the fractions $\frac{d_G(v_i)}{2m}$, where $\{v_1, v_2, \ldots, v_n\}$ is the vertex set of G, and $d_G(v_i)$ is the degree of v_i. We continue earlier work on $I_d(G)$. Our main results deal with the effect of a number of graph operations on the value of $I_d(G)$. We also illustrate the relevance of these results by applying some of these operations to prove a number of extremal results for the degree-based entropy of trees and unicyclic graphs.

1. Introduction

We start this section with some background and motivation underpinning our research and informing the reader who is less experienced with graph theory.

1.1. Background and motivation

Many of the graph concepts related to our current work originate not only from applications of graph theory in chemistry, but also share common ground with the concept of entropy in information theory, as we will see.

A graph is a mathematical object consisting of a finite set of vertices which can be interpreted as the abstraction of the atoms of a molecule, and a set of edges which can be interpreted as the abstraction of the bonds between pairs of atoms in that molecule. In this graph model, each edge represents one bond between one pair of atoms, and multiple bonds between the same pair of atoms are represented by multiple edges between the corresponding pair of vertices in the graph. If such

☆ This research was supported by the National Natural Science Foundation of China (Nos. 12071370, 12131013 and U1803263) and the China Scholarship Council (No. 202006290070).

* Corresponding author.
E-mail address: h.j.broersma@utwente.nl (H. Broersma).

https://doi.org/10.1016/j.amc.2022.127533

multiple bonds exist, then an incidence function is used to associate each edge of the graph with its two end vertices. So, the graph is specified by its vertex set, its edge set, and – if necessary – an incidence function.

To study the structural properties of graphs, one of the first concepts that appeared rather naturally in the graph theory literature is the concept of the degree of a vertex, i.e., the number of edges that have this vertex as one of their end vertices. This vertex degree represents the valency of the corresponding atom in the molecule. Many other and more sophisticated graph concepts that are based on the vertex degrees, the powers of these degrees, the sums or reciprocal sums of these powers, or the distribution of the degrees have appeared over the last decennia. Within the application area of chemistry these graph invariants are usually grouped under the general umbrella term of degree-based topological indices. We refer the interested reader to the survey paper due to Gutman [16] for more information and for a critical comparison of a large number of these indices.

In this paper, we focus on a degree-based graph invariant which is inspired by the seminal work of Shannon [20] in information theory. He defined what is now known as the Shannon entropy of a discrete random variable X as

$$-\sum_{i=1}^{n} p(x_i)\log_b p(x_i),$$

where the possible outcomes x_i of X occur with probability $p(x_i)$ for $i = 1, 2, \ldots, n$, and where b denotes the base of the logarithm. The most commonly used value for b is 2, but other used values for b are 10 and e, the base of the natural logarithm. Although the base of the logarithms in the above expression could be any reasonable constant, throughout the paper we use the base 2. We take the liberty to omit the subscript b or 2, and we use the common convention that $0\log 0 = 0$. Instead of using probabilities in the above formula, one could be tempted to use any type of fractions that together add up to one. In fact, this is the general idea behind a number of graph invariants that have been studied in the context of the Shannon entropy. To be more particular, in the next section we will demonstrate how vertex degrees have been applied in a rather natural way by Cao et al. in [4] to define a graph invariant which reveals a remarkable resemblance with the above expression, as is shown in Definition 1 below.

1.2. Degree-based graph entropy

Recalling that the degree $d_G(v)$ of a vertex v in a graph G is the number of edges having this vertex as an end vertex, it is obvious that each edge contributes 2 to the sum of the degrees taken over all vertices of G. Hence, the sum of the vertex degrees is equal to twice the number of edges of G, a folklore result that goes back to Euler [14], who proved this result in 1736. This result implies that in a graph G with $m > 0$ edges, the fraction $\frac{d_G(v)}{2m}$ is between 0 and 1 for every vertex v of G, and that the sum of these fractions taken over all vertices of G is equal to 1. In the light of the above discussion, from the graph theoretic point of view it makes sense to replace the probabilities in the above formula of the Shannon entropy by these fractions. This is the basic idea behind the following definition of the degree-based entropy which we adopted from Cao et al. [4].

Definition 1. Let G be a graph with vertex set $V = \{v_1, v_2, \ldots, v_n\}$ and $m > 0$ edges. Let $d_G(v_i)$ denote the degree of vertex v_i in G. The degree-based (graph) entropy of G, denoted by $I_d(G)$, is defined as

$$I_d(G) = -\sum_{i=1}^{n} \frac{d_G(v_i)}{2m}\log\frac{d_G(v_i)}{2m}. \quad (1)$$

For later reference, we also define a function

$$h_d(G) = \sum_{i=1}^{n} d_G(v_i)\log d_G(v_i). \quad (2)$$

Straightforward calculations show that $I_d(G) = \log(2m) - \frac{1}{2m}h_d(G)$. This function $h_d(G)$ comes in handy if we want to compare $I_d(G)$ and $I_d(G')$ for two graphs G and G' with the same number of edges, or if we want to obtain the minimum or maximum of $I_d(G)$, where G ranges over all members of a class of graphs with m edges.

In fact, there have appeared earlier studies on graph entropy notions which are based on replacing the probabilities in the Shannon entropy by graph invariants. In the next section, we give a short overview for further background and information.

1.3. More background

As with the above degree-based entropy of (1) in Definition 1, the origin of all graph entropy measures dates back to the seminal work on information entropy due to Shannon. With the introduction of this concept in his 1948 paper [20] he laid the foundation for the development of information theory. The first graph entropy measure which is based on the Shannon entropy is usually attributed to Rashevsky. To measure the information content of an organism, already back in 1955 Rashevsky [19] defined a graph entropy based on certain probability distributions associated with the automorphisms of graphs. Since then graph entropies have been applied to structural graph measures in quantitative graph theory [10,13].

Many different invariant-based entropies have meanwhile been proposed to measure the complexity of graphs. We refer the interested reader to the two survey papers [11,21] and the two books [2,9] for general information and a wealth of pointers to related literature on graph entropies, and the four papers [4–6,8] for some of the more recent examples. In an influential and much cited paper of 2008 [7], it was Dehmer who introduced a general concept of graph entropies, thereby unifying several different existing entropy-based measures. The degree-based graph entropy defined in [4] and (1) of Definition 1 above is a special case of this more general concept.

One of the fundamental and first natural problems in studying any (new) graph invariant is determining its minimum and maximum values, and characterizing the extremal graphs attaining these values. For such problems, restrictions to special graph classes are also often considered. This holds in particular for graph entropies and more generally for topological indices of graphs. Especially in application areas like chemistry, the class of trees and more sophisticated special graph classes are motivated by the atomic structure of hydrocarbon molecules or their carbon atom skeleton.

Recently, several groups of researchers have studied the extremal properties of degree-based entropies. Cao et al. [4] were the first to study the extremal values of entropies based on degree powers for certain families of graphs, and in more detail the entropy defined in (1) of Definition 1. More recently, Ghalavand et al. [15] established the first maximum and minimum values of the entropy in (1) for families of trees and unicyclic graphs, by applying majorization techniques. We come back to this in Section 3, where we extend and apply one of their fundamental lemmas. In a very recent paper, Yan [22] investigated the extremal properties of this entropy for general graphs. The maximum value of this entropy was determined for the class of bipartite graphs in [12], by characterizing the degree sequences of the extremal graphs.

1.4. Our contributions

Our main results in the next section deal with the effect of certain graph operations on the value of the degree-based entropy from (1). It has been demonstrated in many papers that graph operations can form an effective and valuable tool in determining extremal values of several topological indices. Examples of their benefit in obtaining these values have been illustrated with respect to the eccentric connectivity coindex [1], the first and second Zagreb indices [17], and the hyper-Wiener index [18]. For the degree-based graph entropy, results in this direction are generally lacking. Our contributions are motivated by the above observations.

The rest of this paper is organized as follows. In Section 2, we present our results involving the effect of graph operations on the value of the degree-based entropy from (1). We also determine several extremal values of this entropy for trees and unicyclic graphs with some given parameters. In Section 3, we present some preliminary results that will be used in our later proofs. Most of the proofs of our results are postponed to Section 4.

2. Main results

In the next two subsections, we present our main results. We start with the results expressing the effect of certain graph operations on the value of the degree-based entropy. This is followed by a subsection in which we present our extremal results.

2.1. The effect of graph operations

We use standard graph-theoretic terminology and notation, as can be found in the textbook of Bondy and Murty [3]. Whenever we use the term graph, we mean a finite and undirected graph in which we allow multiple edges but no loops. Let $G = (V, E)$ be a graph. We call G an n-vertex graph if $|V| = n$. In the presence of multiple edges, we use an incidence function $\psi_G : E \to V \times V$ that defines the ends $\psi_G(e) = \{u, v\}$ of each edge $e \in E$ incident with two distinct vertices u and v of G. We use the shorthand uv or vu for $\{u, v\}$. If G is a simple graph, we can avoid the use of ψ_G, and use uv or vu to indicate the unique edge e with ends u and v. For a subset $S \subseteq E$, we use $G - S$ to denote the graph $(V, E \setminus S)$ (for which we restrict ψ_G to $E \setminus S$ if necessary). Similarly, we use $G + F$ to denote the graph obtained from G by adding a set F of new edges incident with pairs of distinct vertices of G (possibly creating multiple edges and defining or extending ψ_G in the obvious way). If $S = \{e\}$ or $F = \{e\}$, we use $G - e$ and $G + e$ as shorthand for $G - \{e\}$ and $G + \{e\}$, respectively. Similarly, we use $G - e + f$ as shorthand for $(G - e) + f$.

The following four results deal with the effect of edge additions and edge deletions on the value of the degree-based entropy $I_d(G)$ of (1) in Definition 1.

Theorem 1. *Let u, v, w and x be four vertices of a graph G. Set $G' = G + e$ and $G'' = G + f$, in which $\psi_{G'}(e) = uv$ and $\psi_{G''}(f) = wx$. If $d_G(u) \geq d_G(w)$ and $d_G(v) \geq d_G(x)$, then $I_d(G') \leq I_d(G'')$, with equality holding in the latter inequality if and only if $d_G(u) = d_G(w)$ and $d_G(v) = d_G(x)$.*

The following known result is an easy consequence of Theorem 1.

Corollary 1 [4]**.** *Let u, v and w be three vertices of a simple graph G. Suppose that u and v are adjacent, and w and v are not adjacent, and set $G' = G - uv + wv$. If $d_G(u) - d_G(w) \geq 2$, then $I_d(G) < I_d(G')$.*

In fact, since Theorem 1 holds for multigraphs, we can deduce the slightly stronger statement in which we assume e (with $\psi_G(e) = uv$) is one of possibly more than one edges joining u and v, and f (with $\psi_{G'}(f) = wv$) is a new edge of G' joining the possibly already adjacent vertices w and v of G. We also immediately obtain the following result involving the deletion and addition of pendant edges.

Corollary 2. *Let u, v and w be three vertices of a graph G. Suppose that $d_G(v) = 1$, u and v are adjacent, and w and v are not adjacent. Set $G' = G - e + f$, in which $\psi_G(e) = uv$ and $\psi_{G'}(f) = wv$. Then $I_d(G) \le I_d(G')$ if and only if $d_G(u) > d_G(w)$.*

In our next result, we compare the effect of adding one edge between two different pairs of vertices with the same degree sum.

Theorem 2. *Let s be a positive integer, and let u, v, w and x be four vertices of a graph G. Suppose that $d_G(u) \ge d_G(v) \ge 1$, $d_G(w) \ge d_G(x) \ge 1$, and $d_G(u) + d_G(v) = d_G(w) + d_G(x) = s$. Set $G' = G + e$ and $G'' = G + f$, in which $\psi_{G'}(e) = uv$ and $\psi_{G''}(f) = wx$. If $d_G(u) - d_G(v) \le d_G(w) - d_G(x)$, then $I_d(G') \le I_d(G'')$, with equality holding in the latter inequality if and only if $d_G(u) = d_G(w)$.*

In our next results, we consider a number of more global operations on a graph, the first of which is the so-called k-blow up of G, denoted by $G^{(k)}$. This is the graph obtained by replacing every vertex v of G with $k > 0$ distinct copies, and joining every copy of u to every copy of v in $G^{(k)}$ with ℓ_{uv} edges if and only if there are ℓ_{uv} edges joining u and v in G. We deduce the following expression for the degree-based entropy of $G^{(k)}$.

Theorem 3. *Let G be a graph with at least one edge, and let $k \ge 1$ be an integer. Then $I_d(G^{(k)}) = I_d(G) + \log k$.*

In the next result, we consider the graph $G/\{x, y\}$ obtained from a graph G by identifying two distinct nonadjacent vertices x and y, i.e., replacing x and y by a single new vertex z and making z incident to all edges that were incident to x or y (possibly creating multiple edges). The following result shows that the degree-based entropy decreases if two distinct nonadjacent vertices are identified.

Theorem 4. *Let G be a graph with at least one edge. Suppose that x and y are two distinct nonadjacent vertices of G. Set $G' = G/\{x, y\}$. Then $I_d(G) > I_d(G')$ if and only if $d_G(x) > 0$ and $d_G(y) > 0$.*

We next consider the identification of two vertices $x \in V(G)$ and $z \in V(H)$ from disjoint graphs G and H, resulting in the graph denoted as $GxHz$. We observe the following effect of the degree of x on the degree-based entropy of $GxHz$.

Theorem 5. *Let G and H be two disjoint graphs. Suppose that x and y are two vertices of G, and z is a non-isolated vertex of H. If $d_G(x) \ge d_G(y)$, then $I_d(GxHz) \le I_d(GyHz)$, with equality holding in the latter inequality if and only if $d_G(x) = d_G(y)$.*

Our final result of this subsection deals with the weak product (also known as tensor product or Kronecker product) $G \times G'$ of two disjoint graphs G and G'. This is the graph with vertex set $V(G) \times V(G')$, in which every pair of edges $f \in E(G)$ with $\psi_G(f) = uv$ and $f' \in E(G')$ with $\psi_{G'}(f') = u'v'$ produces an edge $e \in E(G \times G')$ with $\psi_{G\times G'}(e) = (u, u')(v, v')$ (in other words, $(u, u')(v, v')$ has multiplicity ab if uv has multiplicity a and $u'v'$ has multiplicity b).

We deduce the following nice relationship between the degree-based entropies of $G \times G'$, G and G'.

Theorem 6. *Let G and G' be two graphs with at least one edge. Then $I_d(G \times G') = I_d(G) + I_d(G')$.*

In the next subsection, we present our extremal results.

2.2. Extremal results

Before we can present our first result in this subsection, we need some additional notation.

As usual, P_n denotes the path on n vertices. For an integer $k \ge 1$, let $T_n(n_0, n_1, \ldots, n_k)$ be the n-vertex tree which is obtained from the path $P_{k+1} = v_0 v_1 \cdots v_k$ by attaching n_i pendant vertices to the vertex v_i for $i = 1, \ldots, k-1$, so with $n - 2 = \sum_{i=1}^{k-1}(n_i + 1)$. Now we let $\mathcal{T}^*_{n,k}$ (see Fig. 1) denote the set of all n-vertex trees $T_n(n_0, n_1, \ldots, n_k)$ with $n_1 = \cdots = n_{i-1} = n_{i+1} = \cdots = n_{k-1} = 0$ and $n_i = n - 1 - k$ for $i = 1, 2, \ldots, k-1$.

All the trees of Fig. 1 have the same degree-based entropy, and they appear naturally in the following extremal result.

Theorem 7. *Let T be an n-vertex tree with diameter $k \ge 1$. If $I_d(T)$ attains the minimum value among all n-vertex trees with diameter k, then $T \in \mathcal{T}^*_{n,k}$.*

We can prove an analogous result for unicyclic graphs (possibly a tree with one double edge). Let C_n denote the cycle on $n \ge 2$ vertices (where C_2 corresponds to a double edge). Let $C_n(n_1, n_2, \ldots, n_k)$ be the n-vertex unicyclic graph obtained from the cycle $C_k = u_1 u_2 \cdots u_k u_1$ by attaching n_i pendant neighbors to the vertex u_i for $i = 1, 2, \ldots, k$. Figure 2 shows the unicyclic graph $C_n(n_1, n_2, \ldots, n_k)$ for $n_1 = n - k > 0$ and $n_2 = \cdots = n_k = 0$ (i.e., $C_n(n-k, \underbrace{0, \ldots, 0}_{k-1})$).

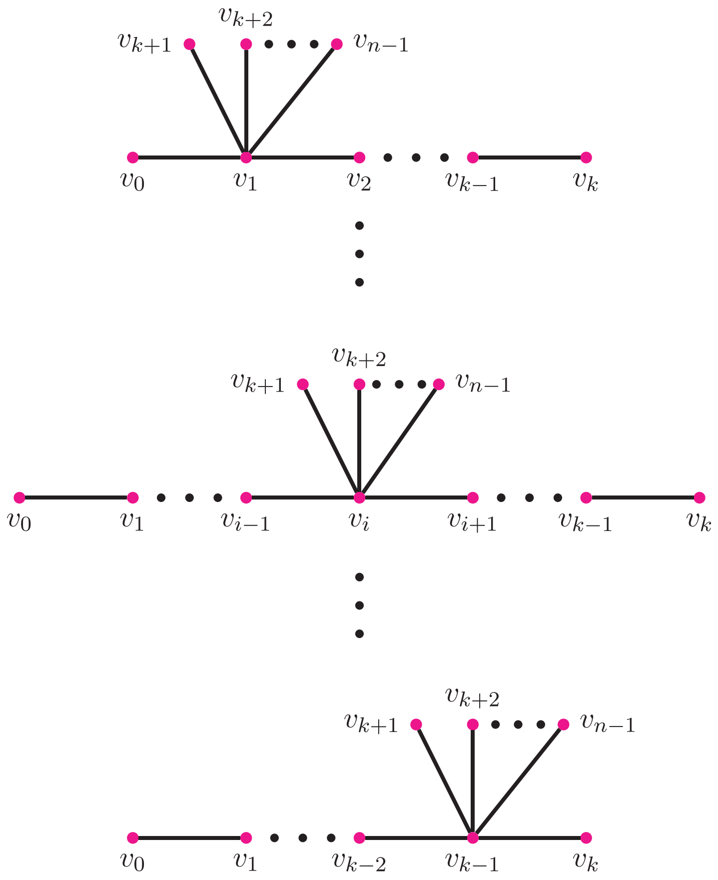

Fig. 1. The trees in $\mathcal{T}^*_{n,k}$.

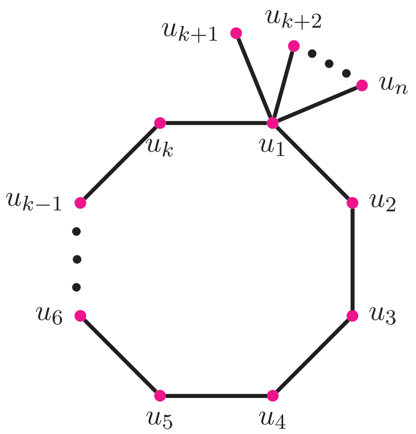

Fig. 2. The unicyclic graph $C_n(n_1, n_2, \ldots, n_k)$ for $n_1 = n - k > 0$ and $n_2 = \cdots = n_k = 0$.

Theorem 8. *Let C be an n-vertex unicyclic graph containing a cycle on $k \geq 2$ vertices. If $I_d(C)$ attains the minimum value among all n-vertex unicyclic graphs containing a cycle on k vertices, then $C \cong C_n(n-k, \underbrace{0, \ldots, 0}_{k-1})$.*

We continue with extremal results for specific subclasses of trees and unicyclic graphs, but first need some additional terminology and notation.

We say that a graph G admits a (p, q)-bipartition if $V(G) = V_1 \cup V_2$ for disjoint sets V_1 and V_2 with $|V_1| = p > 0$ and $|V_2| = q > 0$, and each edge of G has ends in V_1 and V_2.

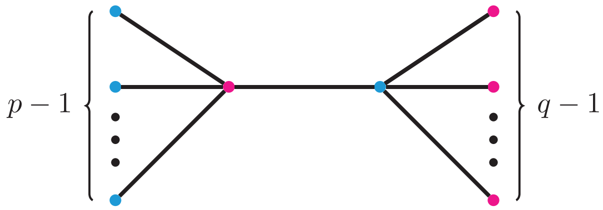

Fig. 3. The tree $S^*(p,q)$.

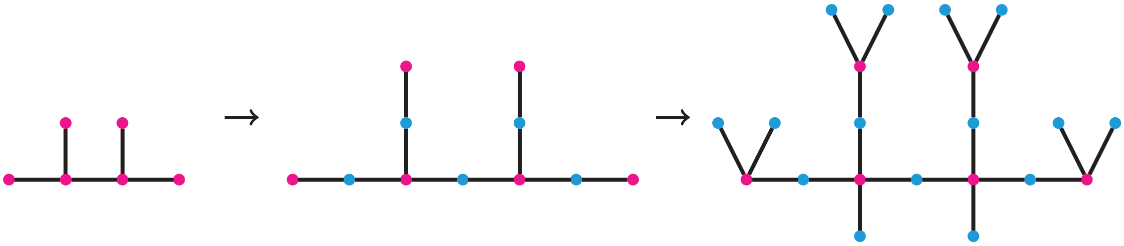

Fig. 4. The construction of a tree in $\mathcal{T}^*(15,6)$.

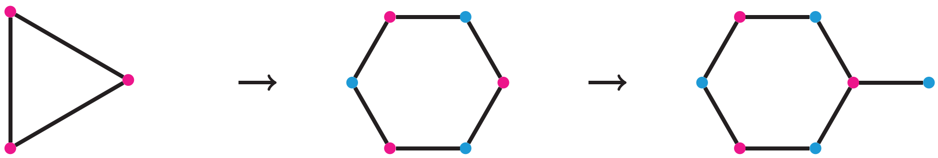

Fig. 5. The construction of one member of $\mathcal{C}^*(4,3)$.

Let $\mathcal{T}(p,q)$ denote the set of all trees admitting a (p,q)-bipartition. Let $S^*(p,q)$ be the member of $\mathcal{T}(p,q)$ obtained by attaching $p-1$ and $q-1$ pendant vertices to the two vertices of a P_2, respectively (as indicated in Fig. 3).

We consider another specific subclass of $\mathcal{T}(p,q)$. Let $\mathcal{T}^*(p,q)$ denote the set of all trees that can be obtained from a q-vertex tree T with $\Delta(T) \le \left\lceil \frac{p+q-1}{q} \right\rceil$ in the following way. First subdivide every edge of T, i.e., replace each edge $e=uv$ by a path $ux_{uv}v$ for a newly added vertex x_{uv}. The new tree clearly admits a $(q-1,q)$-partition with $V_2=V(T)$ and V_1 consisting of the newly added vertices. Next attach $p-q+1$ pendant vertices to the vertices of V_2 in such a way that the maximum degree of the vertices in V_2 exceeds their minimum degree by at most 1. The construction of one member of $\mathcal{T}^*(15,6)$ is illustrated in Fig. 4. Clearly, by construction every tree in $\mathcal{T}^*(p,q)$ has a (p,q)-bipartition.

The next result determines the minimum value of $I_d(T)$ among all trees $T \in \mathcal{T}(p,q)$ and characterizes the unique extremal tree.

Theorem 9. *Let p and q be integers with $p \ge q \ge 1$. Then $I_d(T)$ attains the minimum value among all trees in $\mathcal{T}(p,q)$ if and only if $T \cong S^*(p,q)$.*

The following result determines the maximum value of $I_d(T)$ among all trees $T \in \mathcal{T}(p,q)$ and characterizes all the extremal trees.

Theorem 10. *Let p and q be integers with $p \ge q \ge 1$. Then $I_d(T)$ attains the maximum value among all trees in $\mathcal{T}(p,q)$ if and only if $T \in \mathcal{T}^*(p,q)$.*

We finish this section with the counterparts of the above tree results for unicyclic graphs. For this, we let $\mathcal{C}(p,q)$ denote the set of all unicyclic graphs (possibly a tree with one double edge) admitting a (p,q)-bipartition. Obviously, every member of $\mathcal{C}(p,q)$ has a unique cycle on an even number of vertices. We define a subclass $\mathcal{C}^*(p,q)$ of $\mathcal{C}(p,q)$ in a similar way as we did for trees. Let $\mathcal{C}^*(p,q)$ consist of all unicyclic graphs that can be obtained from a q-vertex unicyclic graph C with $\Delta(C) \le \left\lceil \frac{p+q}{q} \right\rceil$ in the following way. First subdivide every edge of C to obtain a unicyclic graph which admits a (q,q)-bipartition with $V_2=V(C)$ and V_1 consisting of the newly added vertices. Next attach $p-q$ pendant vertices to the vertices of V_2 in such a way that the maximum degree of the vertices in V_2 exceeds their minimum degree by at most 1. The construction of one member of $\mathcal{C}^*(4,3)$ is illustrated in Fig. 5. Clearly, by construction every unicyclic graph in $\mathcal{C}^*(p,q)$ has a (p,q)-bipartition.

The next result determines the minimum value of $I_d(C)$ among all unicyclic graphs $C \in \mathcal{C}(p,q)$ and identifies the unique extremal graph.

Theorem 11. *Let p and q be integers with $p \geq q \geq 1$. Then $I_d(C)$ attains the minimum value among all unicyclic graphs in $\mathcal{C}(p, q)$ if and only if $C \cong C_n(p-1, q-1)$.*

Note that $C_n(p-1, q-1)$ in the above statement is a special case of the previously defined class $C_n(n_1, n_2, \ldots, n_k)$, and that $C_n(p-1, q-1)$ can be obtained from the tree in Fig. 3 by replacing the middle edge by a double edge.

Our final result determines the maximum value of $I_d(C)$ among all unicyclic graphs $C \in \mathcal{C}(p, q)$ and characterizes all the extremal graphs.

Theorem 12. *Let p and q be integers with $p \geq q \geq 1$. Then $I_d(C)$ attains the maximum value among all unicyclic graphs in $\mathcal{C}(p, q)$ if and only if $C \in \mathcal{C}^*(p, q)$.*

All proofs of our main results are postponed to Section 4. In the next section, we introduce some additional terminology and tools, together with several auxiliary results that we will use in our proofs.

3. Preliminaries

In this section, we give some additional terminology and notation, and we state and prove a number of lemmas which will be used in the proofs of our results.

Let G be a graph, and let S be a nonempty subset of $V(G)$. If $|S| = a_1 + a_2 + \cdots + a_{t+1}$ and the number of vertices with degree d_i in S is a_i for $i = 1, 2, \ldots, t+1$ and $d_1 > d_2 > \cdots > d_{t+1} = 0$, then by $D(S) = [d_1^{a_1}, d_2^{a_2}, \ldots, d_t^{a_t}]$ we denote the degree sequence of S. We use $D(G)$ instead of $D(V(G))$. If there exists a graph G with degree sequence $D = D(G)$, then D is called graphic, and G is called a realization of D. We use $\Delta(G)$ (resp., $\delta(G)$) to denote the maximum (resp., minimum) degree of the graph G. By S_n we denote the star on n vertices. We say v is the center of the star S_n if $d_{S_n}(v) = n-1$.

In the context of our research, majorization is a useful relationship between two non-increasing integer (degree) sequences $A = [a_1, a_2, \ldots, a_n]$ and $B = [b_1, b_2, \ldots, b_n]$. We say that A majorizes B, denoted by $A \succeq B$, if for all $k \in \{1, 2, \ldots, n-1\}$:

$$\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i, \quad \text{and} \quad \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i.$$

If at least one of the above inequalities is strict, then we say the majorization is strict. We use $A \succ B$ to express that A strictly majorizes B.

Ghalavand et al. [15] proved the following result for simple graphs, but it is straightforward to extend their proof to graphs. We omit the proof.

Lemma 1. *Let G and G′ be two graphs with the same number of vertices and edges. If $D(G) \succeq D(G')$, then $I_d(G) \leq I_d(G')$, with equality holding in the inequality if and only if $D(G) = D(G')$.*

The following extremal result shows that $I_d(S_n)$ attains the minimum value among all n-vertex trees.

Theorem 13 [4]**.** *Let T be an n-vertex tree. Then $I_d(T) \geq I_d(S_n)$, with equality holding in the inequality if and only if $T \cong S_n$.*

Let G be a fixed graph with at least one edge, and let T be a randomly chosen n-vertex tree. Denote by $GuTw$ the graph obtained from G and T by identifying a fixed non-isolated vertex $u \in V(G)$ and a randomly chosen vertex $w \in V(T)$. Let v be the center of the star S_n. For our proofs of Theorems 7 and 8, the following two lemmas are key ingredients.

The first result shows that among all n-vertex trees, $I_d(GuS_nv)$ attains the minimum value.

Lemma 2. *Let T be an n-vertex tree with $w \in V(T)$, and let v be the center of the star S_n. Suppose that G is a fixed graph with a fixed non-isolated vertex $u \in V(G)$. Then $I_d(GuTw) \geq I_d(GuS_nv)$, with equality holding in the inequality if and only if $GuTw \cong GuS_nv$.*

Proof. Suppose that $I_d(GuTw) \leq I_d(GuS_nv)$ and $GuTw$ is not isomorphic to GuS_nv. So we have $d_T(w) < n-1$. This implies that $d_{GuTw}(u) = d_G(u) + d_T(w) < d_G(u) + n - 1 = d_{GuS_nv}(u)$. By Theorem 13, and recalling (2) and the remarks we made there, we have $h_d(T) \leq h_d(S_n)$. Since $d_G(u) \geq 1$, the function $g(t) = (t + d_G(u))\log(t + d_G(u)) - t\log t$ strictly increases as t increases for $t > 0$. This implies

$$\begin{aligned} h_d(GuS_nv) - h_d(GuTw) =& h_d(S_n) - (n-1)\log(n-1) + d_{GuS_nv}(u)\log d_{GuS_nv}(u) \\ & - h_d(T) + d_T(w)\log d_T(w) - d_{GuTw}(u)\log d_{GuTw}(u) \\ =& h_d(S_n) - h_d(T) \\ & + (d_G(u) + n - 1)\log(d_G(u) + n - 1) - (n-1)\log(n-1) \\ & - (d_G(u) + d_T(w))\log(d_G(u) + d_T(w)) + d_T(w)\log d_T(w) \\ =& h_d(S_n) - h_d(T) + g(n-1) - g(d_T(w)) \\ >& 0, \end{aligned}$$

a contradiction. □

For our next lemma, let $\mathcal{T}_{n,k}$ (resp. $\mathcal{C}_{n,k}$) denote the set of all trees $T_n(n_0, n_1, \ldots, n_k)$ (resp. all unicyclic graphs $C_n(n_1, n_2, \ldots, n_k)$). We consider the extremal results for $\mathcal{T}_{n,k}$ and $\mathcal{C}_{n,k}$.

Lemma 3. *Let $\mathcal{T}_{n,k}$ and $\mathcal{C}_{n,k}$ be defined as above. Then*

(a) *$I_d(T)$ attains the minimum value among all trees in $\mathcal{T}_{n,k}$ if and only if $T \in \mathcal{T}^*_{n,k}$ for $n > k \geq 1$;*
(b) *$I_d(C)$ attains the minimum value among all unicyclic graphs in $\mathcal{C}_{n,k}$ if and only if $C \cong C_n(n-k, \underbrace{0, \ldots, 0}_{k-1})$ for $n \geq k \geq 2$.*

Proof. We only prove (a) because (b) can be proved similarly.

Suppose that $T \notin \mathcal{T}^*_{n,k}$ and $I_d(T)$ attains the minimum value among all trees in $\mathcal{T}_{n,k}$. Let $P_{k+1} = v_0 v_1 \cdots v_k$ be the diametrical path of T. This implies that there exist two distinct vertices v_i and v_j with $d_T(v_i) \geq 3$ and $d_T(v_j) \geq 3$. Without loss of generality, we may assume that $d_T(v_i) \leq d_T(v_j)$. Let $v \notin V(P_{k+1})$ be a neighbor of v_i. Set $T' = T - v_i v + v_j v$. By Corollary 2, we have $I_d(T') < I_d(T)$, a contradiction. □

We need the next two lemmas for our proof of Theorem 11. Let n, p, q and k be four integers with $p+q=n$, $p \geq q \geq 1$ and $k \geq 2$. The following lemma shows that, among all possible values for $n_1, n_2, \ldots, n_k$ with $C_n(n_1, n_2, \ldots, n_k) \in \mathcal{C}(p,q)$, the minimum value is attained by $I_d(C_n(p - \frac{k}{2}, q - \frac{k}{2}, \underbrace{0, \ldots, 0}_{k-2}))$.

Lemma 4. *Let n, p, q and k be four integers with $p+q=n$, $p \geq q \geq 1$ and $k \geq 2$. If $C_n(n_1, n_2, \ldots, n_k) \in \mathcal{C}(p,q)$, then $I_d(C_n(n_1, n_2, \ldots, n_k)) \geq I_d(C_n(p - \frac{k}{2}, q - \frac{k}{2}, \underbrace{0, \ldots, 0}_{k-2}))$.*

Proof. Let $C_k = u_1 u_2 \cdots u_k u_1$ be the cycle of $C_n(n_1, n_2, \ldots, n_k)$ and let (V_1, V_2) correspond to a (p,q)-bipartition of $C_n(n_1, n_2, \ldots, n_k)$. Suppose that $u_{2j} \in V_1$ and $u_{2j-1} \in V_2$ for $j = 1, 2, \ldots, \frac{k}{2}$. This implies that

$$\sum_{j=1}^{\frac{k}{2}} \left(d_{C_n(n_1,n_2,\ldots,n_k)}(u_{2j-1}) - 2 \right) = p - \frac{k}{2}$$

and

$$\sum_{j=1}^{\frac{k}{2}} \left(d_{C_n(n_1,n_2,\ldots,n_k)}(u_{2j}) - 2 \right) = q - \frac{k}{2}.$$

It follows that

$$2 \leq d_{C_n(n_1,n_2,\ldots,n_k)}(u_{2j}) \leq p - \frac{k}{2} + 2$$

and

$$2 \leq d_{C_n(n_1,n_2,\ldots,n_k)}(u_{2j-1}) \leq q - \frac{k}{2} + 2$$

for $j = 1, 2, \ldots, \frac{k}{2}$, and

$$d_{C_n(n_1,n_2,\ldots,n_k)}(u) = 1$$

for $u \in V(C_n(n_1, n_2, \ldots, n_k)) \setminus V(C_k)$. We have $D(C_n(p - \frac{k}{2}, q - \frac{k}{2}, \underbrace{0, \ldots, 0}_{k-2})) = [p - \frac{k}{2} + 2, q - \frac{k}{2} + 2, 2^{k-2}, 1^{n-k}]$. This implies $D(C_n(p - \frac{k}{2}, q - \frac{k}{2}, \underbrace{0, \ldots, 0}_{k-2})) \succeq D(C_n(n_1, n_2, \ldots, n_k))$. By Lemma 1, we have

$$I_d(C_n(p - \frac{k}{2}, q - \frac{k}{2}, \underbrace{0, \ldots, 0}_{k-2})) \leq I_d(C_n(n_1, n_2, \ldots, n_k)).$$

□

We state and prove one more lemma to complete this section. This lemma shows the effect of shortening the cycle of $C_n(s, t, \underbrace{0, \ldots, 0}_{k-2})$ on the value of the degree-based entropy.

Lemma 5. *Let n, s, t and k be four integers with $s+t+k=n$, $n \geq k \geq 4$ and $s, t \geq 0$. Then*

$$I_d(C_n(s+1, t+1, \underbrace{0, \ldots, 0}_{k-4})) < I_d(C_n(s, t, \underbrace{0, \ldots, 0}_{k-2})).$$

Proof. Without loss of generality, we assume that $s \geq t$. It follows from

$$D(C_n(s,t,\underbrace{0,\ldots,0}_{k-2})) = [s+2, t+2, 2^{k-2}, 1^{n-k}]$$

and

$$D(C_n(s+1,t+1,\underbrace{0,\ldots,0}_{k-4})) = [s+3, t+3, 2^{k-4}, 1^{n-k+2}]$$

that $D(C_n(s+1,t+1,\underbrace{0,\ldots,0}_{k-4})) \succ D(C_n(s,t,\underbrace{0,\ldots,0}_{k-2}))$. By Lemma 1, we have $I_d(C_n(s+1,t+1,\underbrace{0,\ldots,0}_{k-4})) < I_d(C_n(s,t,\underbrace{0,\ldots,0}_{k-2}))$. □

4. Proofs

In this final section, we gathered all the missing proofs of the statements in earlier sections.

Proof of Theorem 1. Without loss of generality, we assume that $d_G(u) \geq d_G(v)$. If $d_G(u) = d_G(w)$ or $d_G(v) = d_G(x)$, then we have $D(G') \succeq D(G'')$. Using Lemma 1, we get $I_d(G') \leq I_d(G'')$, with equality holding in this inequality if and only if $d_G(u) = d_G(w)$ and $d_G(v) = d_G(x)$. We consider $d_G(u) > d_G(w)$ and $d_G(v) > d_G(x)$ in the following. It follows that $D(G') \succ D(G'')$ if $d_G(u) = d_G(v)$. Then, using Lemma 1 again, we have $I_d(G') < I_d(G'')$. We only prove the case $d_G(u) > d_G(w) > d_G(v) > d_G(x)$, since the other cases $d_G(u) > d_G(v) = d_G(w) > d_G(x)$, $d_G(u) > d_G(v) > d_G(w) > d_G(x)$, $d_G(u) > d_G(v) > d_G(w) = d_G(x)$, and $d_G(u) > d_G(v) > d_G(x) > d_G(w)$ can be proved similarly. Let us relabel the vertices of the graph G as $v_1, v_2, \ldots, v_n$ such that $d_G(v_1) \geq d_G(v_2) \geq \cdots \geq d_G(v_n)$. Suppose that $v_i = u$, $v_j = w$, $v_s = v$ and $v_r = x$. We may assume $i = \min\{l | d_G(v_l) = d_G(u)\}$, $j = \min\{l | d_G(v_l) = d_G(w)\}$, $s = \min\{l | d_G(v_l) = d_G(v)\}$, $r = \min\{l | d_G(v_l) = d_G(x)\}$ and $i < j < s < r$.

For each $k \in \{1, 2, \ldots, i-1\}$, we have

$$\sum_{t=1}^{k} d_{G'}(v_t) = \sum_{t=1}^{k} d_{G''}(v_t);$$

for each $k \in \{i, i+1, \ldots, j-1\}$, we have

$$\sum_{t=1}^{k} d_{G'}(v_t) > \sum_{t=1}^{k} d_{G''}(v_t);$$

for each $k \in \{j, j+1, \ldots, s-1\}$, we have

$$\sum_{t=1}^{k} d_{G'}(v_t) = \sum_{t=1}^{k} d_{G''}(v_t);$$

for each $k \in \{s, s+1, \ldots, r-1\}$, we have

$$\sum_{t=1}^{k} d_{G'}(v_t) > \sum_{t=1}^{k} d_{G''}(v_t);$$

for each $k \in \{r, r+1, \ldots, n\}$, we have

$$\sum_{t=1}^{k} d_{G'}(v_t) = \sum_{t=1}^{k} d_{G''}(v_t).$$

Therefore, $D(G') \succ D(G'')$. Using Lemma 1, we conclude that $I_d(G') < I_d(G'')$. □

Proof of Corollary 1. Set $H = G - uv$. We have $G = H + uv$ and $G' = H + wv$. Since $d_G(u) - d_G(w) \geq 2$, $d_H(u) = d_G(u) - 1 > d_G(w) = d_H(w)$. By Theorem 1, we have $I_d(G) < I_d(G')$. □

Proof of Corollary 2. Set $H = G - e$. We have $G = H + e$ and $G' = H + f$.

Suppose that $d_G(u) \leq d_G(w)$. We have $d_H(u) = d_G(u) - 1 < d_G(w) = d_H(w)$. By Theorem 1, we have $I_d(G) > I_d(G')$, a contradiction.

Hence, $d_G(u) > d_G(w)$ and $d_H(u) = d_G(u) - 1 \geq d_G(w) = d_H(w)$. By Theorem 1, we have $I_d(G) \leq I_d(G')$. □

Proof of Theorem 2. Since $d_G(u) + d_G(v) = d_G(w) + d_G(x) = s$ and $d_G(u) - d_G(v) \leq d_G(w) - d_G(x)$, we have $d_G(u) \geq \frac{s}{2}$, $d_G(w) \geq \frac{s}{2}$ and $d_G(w) \geq d_G(u)$. If $d_G(u) = d_G(w)$, then $I_d(G') = I_d(G'')$. We consider the case $d_G(w) > d_G(u)$ in the following.

Let $g(t) = t\log\frac{t+1}{t} + \log(t+1) + (s-t)\log\frac{s-t+1}{s-t} + \log(s-t+1)$ for $\frac{s}{2} \le t \le s-1$. By calculating the first-order derivative, we obtain

$$g'(\frac{s}{2}) = 0$$

and

$$g'(t) = \log\frac{t+1}{t} - \log\frac{s-t+1}{s-t} < 0$$

for $\frac{s}{2} < t \le s-1$. This implies $g(t)$ strictly decreases as t increases for $\frac{s}{2} \le t \le s-1$. Let m be the number of edges of G. Because $d_G(u) + d_G(v) = d_G(w) + d_G(x) = s$, we have

$$\begin{aligned}
I_d(G') =& \log(2m+2) - \frac{1}{2m+2}h_d(G) + \frac{1}{2m+2}(d_G(u)\log d_G(u) \\
&+ d_G(v)\log d_G(v) - (d_G(u)+1)\log(d_G(u)+1) - (d_G(v)+1)\log(d_G(v)+1)) \\
=& \log(2m+2) - \frac{1}{2m+2}h_d(G) - \frac{1}{2m+2}\Big(d_G(u)\log\frac{d_G(u)+1}{d_G(u)} \\
&+ \log(d_G(u)+1) + d_G(v)\log\frac{d_G(v)+1}{d_G(v)} + \log(d_G(v)+1)\Big) \\
=& \log(2m+2) - \frac{1}{2m+2}h_d(G) - \frac{1}{2m+2}\Big(d_G(u)\log\frac{d_G(u)+1}{d_G(u)} \\
&+ \log(d_G(u)+1) + (s-d_G(u))\log\frac{s-d_G(u)+1}{s-d_G(u)} + \log(s-d_G(u)+1)\Big) \\
=& \log(2m+2) - \frac{1}{2m+2}h_d(G) - \frac{1}{2m+2}g(d_G(u)).
\end{aligned}$$

By similar calculations, $I_d(G'') = \log(2m+2) - \frac{1}{2m+2}h_d(G) - \frac{1}{2m+2}g(d_G(w))$.

Since $d_G(x) \ge 1$, $s-1 \ge d_G(w)$. This implies $\frac{s}{2} \le d_G(u) < d_G(w) \le s-1$. So we have $g(d_G(u)) > g(d_G(w))$. Thus $I_d(G') < I_d(G'')$. □

Proof of Theorem 3. Let $\{v_1, v_2, \ldots, v_n\}$ denote the vertex set of G. We use $v_{i1}, v_{i2}, \ldots, v_{ik}$ to denote k copies of v_i in the blow-up graph $G^{(k)}$. Let m be the number of edges of G. By definition, we have $d_{G^{(k)}}(v_{ij}) = kd_G(v_i)$ for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, k$. Thus $\sum_{j=1}^{k}\sum_{i=1}^{n} d_{G^{(k)}}(v_{ij}) = \sum_{j=1}^{k}\sum_{i=1}^{n} kd_G(v_i) = \sum_{j=1}^{k} 2km = 2k^2m$. So we have

$$\begin{aligned}
I_d(G^{(k)}) &= \log(2k^2m) - \frac{1}{2k^2m}\sum_{i=1}^{n}\sum_{j=1}^{k} d_{G^{(k)}}(v_{i,j})\log d_{G^{(k)}}(v_{i,j}) \\
&= \log(2k^2m) - \frac{1}{2k^2m}\sum_{i=1}^{n}(k^2 d_G(v_i))\log(kd_G(v_i)) \\
&= \log(2k^2m) - \frac{k^2}{2k^2m}(\log k\sum_{i=1}^{n} d_G(v_i) + \sum_{i=1}^{n} d_G(v_i)\log d_G(v_i)) \\
&= \log(2m) - \frac{1}{2m}h_d(G) + \log k \\
&= I_d(G) + \log k.
\end{aligned}$$

□

Proof of Theorem 4. Let m be the number of edges of G. Identifying x and y of G, we use a vertex z to replace these vertices. This implies $d_{G'}(z) = d_G(x) + d_G(y)$.

Suppose that $d_G(x) = 0$ or $d_G(y) = 0$. It follows from $d_G(x) = 0$ (resp. $d_G(y) = 0$) that $d_{G'}(z) = d_G(y)$ (resp. $d_{G'}(z) = d_G(x)$). Because the case with $d_G(y) = 0$ can be proved similarly, we only consider the case that $d_G(x) = 0$. We have

$$\begin{aligned}
I_d(G') - I_d(G) &= \frac{1}{2m}(d_G(x)\log d_G(x) + d_G(y)\log d_G(y) - d_{G'}(z)\log d_{G'}(z)) \\
&= \frac{1}{2m}(0\log 0 + d_G(y)\log d_G(y) - d_G(y)\log d_G(y)) \\
&= 0,
\end{aligned}$$

a contradiction.

This contradiction implies $d_G(x) > 0$ and $d_G(y) > 0$, and $\frac{d_G(x)^{d_G(x)} + d_G(y)^{d_G(y)}}{(d_G(x)+d_G(y))^{(d_G(x)+d_G(y))}} < 1$. So we have

$$I_d(G') - I_d(G) = \frac{1}{2m}(d_G(x)\log d_G(x) + d_G(y)\log d_G(y) - d_{G'}(z)\log d_{G'}(z))$$

$$
\begin{aligned}
&= \frac{1}{2m}(d_G(x)\log d_G(x) + d_G(y)\log d_G(y) \\
&\quad - (d_G(x)+d_G(y))\log(d_G(x)+d_G(y))) \\
&= \frac{1}{2m}\log\frac{d_G(x)^{d_G(x)} + d_G(y)^{d_G(y)}}{(d_G(x)+d_G(y))^{(d_G(x)+d_G(y))}} \\
&< 0.
\end{aligned}
$$

□

Proof of Theorem 5. Let G and H be two graphs satisfying the hypothesis of the theorem. Since z is a non-isolated vertex of H, we have $d_H(z) \geq 1$. If $d_G(x) > d_G(y)$, then

$$\frac{d_G(y)^{d_G(y)}(d_G(x)+d_H(z))^{d_G(x)}}{d_G(x)^{d_G(x)}(d_G(y)+d_H(z))^{d_G(y)}} > 1$$

and

$$\frac{d_G(x)+d_H(z)}{d_G(y)+d_H(z)} > 1.$$

This implies

$$
\begin{aligned}
h_d(GxHz) - h_d(GyHz) &= (d_G(x)+d_H(z))\log(d_G(x)+d_H(z)) + d_G(y)\log d_G(y) \\
&\quad - (d_G(y)+d_H(z))\log(d_G(y)+d_H(z)) - d_G(x)\log d_G(x) \\
&= d_G(x)\log\frac{d_G(x)+d_H(z)}{d_G(x)} + d_G(y)\log\frac{d_G(y)}{d_G(y)+d_H(z)} \\
&\quad + d_G(z)\log\frac{d_G(x)+d_H(z)}{d_G(y)+d_H(z)} \\
&= \log\frac{d_G(y)^{d_G(y)}(d_G(x)+d_H(z))^{d_G(x)}}{d_G(x)^{d_G(x)}(d_G(y)+d_H(z))^{d_G(y)}} \\
&\quad + d_G(z)\log\frac{d_G(x)+d_H(z)}{d_G(y)+d_H(z)} \\
&> 0.
\end{aligned}
$$

So we have $I_d(GxHz) < I_d(GyHz)$.

If $d_G(x) = d_G(y)$, then $D(GxHz) = D(GyHz)$. Thus $I_d(GxHz) = I_d(GyHz)$. □

Proof of Theorem 6. Let m (resp. m') be the number of edges of G (resp. G'). By definition, $d_{G\times G'}((u,u')) = d_G(u)d_{G'}(u')$ for $u \in V(G)$ and $u' \in V(G')$. So we have

$$
\begin{aligned}
&\sum_{\substack{u\in V(G)\\ u'\in V(G')}} d_{G\times G'}((u,u')) \\
&= \sum_{u\in V(G)}\sum_{u'\in V(G')} d_G(u)d_{G'}(u') \\
&= 2m'\sum_{u\in V(G)} d_G(u) \\
&= 4mm'.
\end{aligned}
$$

This implies

$$
\begin{aligned}
I_d(G\times G') &= \log(4mm') - \frac{1}{4mm'}\sum_{\substack{u\in V(G)\\ u'\in V(G')}} d_{G\times G'}((u,u'))\log d_{G\times G'}((u,u')) \\
&= \log(4mm') - \frac{1}{4mm'}\sum_{\substack{u\in V(G)\\ u'\in V(G')}} d_G(u)d_{G'}(u')\log(d_G(u)d_{G'}(u')) \\
&= \log(4mm') - \frac{1}{4mm'}\sum_{\substack{u\in V(G)\\ u'\in V(G')}} (d_G(u)d_{G'}(u')\log d_G(u) + d_G(u)d_{G'}(u')\log d_{G'}(u')) \\
&= \log(4mm') - \frac{2m'\sum_{u\in V(G)} d_G(u)\log d_G(u) + 2m\sum_{u'\in V(G')} d_{G'}(u')\log d_{G'}(u')}{4mm'}
\end{aligned}
$$

$$= \log(2m) - \frac{1}{2m}h_d(G) + \log(2m') - \frac{1}{2m'}h_d(G')$$
$$= I_d(G) + I_d(G').$$

□

Proof of Theorem 7. Let $P_{k+1} = v_0v_1\cdots v_k$ be a diametrical path of T, and let T^i be the component of $T - E(P_{k+1})$ containing v_i for $i = 0, 1, \ldots, k$. Let H be the component of $T - E(T^i)$ containing v_i. By Lemma 2, we have $I_d(Hv_iT^iv_i) \geq I_d(Hv_iS_{|V(T^i)|}v)$ in which v is the center of $S_{|V(T^i)|}$. This implies that $T \in \mathcal{T}_{n,k}$. And by Lemma 3 (a), we have $T \in \mathcal{T}^*_{n,k}$. □

Proof of Theorem 8. Let $C_k = u_1u_2\cdots u_ku_1$ be the cycle of C, and let T^i be the component of $C - E(C_k)$ containing u_i for $i = 1, 2, \ldots, k$. Let H be the component of $C - E(T^i)$ containing u_i. By Lemma 2, we have $I_d(Hu_iT^iu_i) \geq I_d(Hu_iS_{|V(T^i)|}v)$ in which v is the center of $S_{|V(T^i)|}$. This implies that $C \in \mathcal{C}_{n,k}$. And by Lemma 3 (b), we have $C \cong C_n(n-k, \underbrace{0, \ldots, 0}_{k-1})$. □

Proof of Theorem 9. Let $P_{t+1} = v_0v_1\cdots v_t$ be a diametrical path of T. Suppose that the diameter of T is at least 4 (i.e, $t \geq 4$). Let A be the set of neighbors of v_3 excluding v_2 of T. Let H_1 (resp. H_2) be the component of $T - A$ (resp. $T - v_2v_3$) containing v_3. Then T can be obtained from H_1 and H_2 by identifying $v_3 \in V(H_1)$ and $v_3 \in V(H_2)$. Let T' be the tree obtained from H_1 and H_2 by identifying $v_1 \in V(H_1)$ and $v_3 \in V(H_2)$. We may partition the vertex set of T' and T in the same way. Thus T' has a (p, q)-bipartition. Clearly, $d_{H_1}(v_1) > d_{H_1}(v_3)$. By Theorem 5, we have $I_d(T') < I_d(T)$, a contradiction. Therefore, the diameter of T is at most 3, that is, $T \cong S^*(p, q)$. □

Proof of Theorem 9. Let V_1 and V_2 be two subsets of $V(T)$ satisfying $|V_1| = p$ and $|V_2| = q$, and such that each edge of T joins a vertex in V_1 and a vertex in V_2. Let $r = (p+q-1) - q\lfloor \frac{p+q-1}{q} \rfloor$. We state a claim.

Claim 1. *$D(V_1) = [2^{q-1}, 1^{p-q+1}]$ and $D(V_2) = [(\lceil \frac{p+q-1}{q} \rceil)^r, (\lfloor \frac{p+q-1}{q} \rfloor)^{q-r}]$.*

Proof. We only prove $D(V_1) = [2^{q-1}, 1^{p-q+1}]$, since $D(V_2) = [(\lceil \frac{p+q-1}{q} \rceil)^r, (\lfloor \frac{p+q-1}{q} \rfloor)^{q-r}]$ can be proved similarly. Suppose that $D(V_1) = [2^{q-1}, 1^{p-q+1}]$ does not hold. This implies there exist two vertices $u \in V_1$ and $v \in V_1$ satisfying $d_T(u) - d_T(v) \geq 2$. Let P be the path from u to v, and let $w \notin V(P)$ be a neighbor of u. Let A be the set of neighbors of u excluding w. Let H_1 (resp. H_2) be the component of $T - A$ (resp. $T - uw$) containing u. Then T can be obtained from H_1 and H_2 by identifying $u \in V(H_1)$ and $u \in V(H_2)$. Let T' be the tree obtained from H_1 and H_2 by identifying $u \in V(H_1)$ and $v \in V(H_2)$. We may partition the vertex set of T and T' in the same way. This implies that T' has a (p, q)-bipartition. Clearly, $d_{H_2}(u) > d_{H_2}(v)$. By Theorem 5, we have $I_d(T) < I_d(T')$, a contradiction. Since all graphs in $\mathcal{T}^*(p, q)$ are realizations of $D(V_1) = [2^{q-1}, 1^{p-q+1}]$ and $D(V_2) = [(\lceil \frac{p+q-1}{q} \rceil)^r, (\lfloor \frac{p+q-1}{q} \rfloor)^{q-r}]$, this pair of degree sequences is graphic. □

Using Claim 1, to prove $T \in \mathcal{T}^*(p, q)$, it suffices to show that realizations of $D(V_1) = [2^{q-1}, 1^{p-q+1}]$ and $D(V_2) = [(\lceil \frac{p+q-1}{q} \rceil)^r, (\lfloor \frac{p+q-1}{q} \rfloor)^{q-r}]$ are in $\mathcal{T}^*(p, q)$. Let T'' be a tree obtained from T by deleting pendant vertices in V_1, and identifying vertices with degree 2 in V_1 with one of their neighbors (avoiding loops). It follows that T'' is a q-vertex tree with maximum degree at most $\lceil \frac{p+q-1}{q} \rceil$, and T can be obtained by subdividing every edge of T'' and attaching the pendant vertices to the original vertices. Thus $T \in \mathcal{T}^*(p, q)$. □

Proof of Theorem 11. Let V_1 and V_2 be two subsets of $V(C)$ satisfying $|V_1| = p$ and $|V_2| = q$, and such that each edge of C joins a vertex in V_1 and a vertex in V_2. Let $C_k = u_1u_2\cdots u_ku_1$ be the cycle of C. If $k = n = 2$, then $C \cong C_2$. Suppose that $C \cong C_n$ (i.e., $k = n$) for $n \geq 4$. If $C_n \in \mathcal{C}(p, q)$ and $n \geq 4$, then $C_n(1, 1, \underbrace{0, \ldots, 0}_{n-4}) \in \mathcal{C}(p, q)$. Then, by Lemma 5, $I_d(C_n) > I_d(C_n(1, 1, \underbrace{0, \ldots, 0}_{n-4}))$ for $n \geq 4$, which leads to a contradiction. So we have $k < n$ for $n \geq 4$. This implies there is a vertex u_i of degree at least 3.

Let A be the set of neighbors of u_i of C excluding u_{i-1} and u_{i+1}, where the addition is taken modulo k. Let H_1 (resp. T^1) be the component of $C - A$ (resp. $C - \{e_1, e_2\}$) containing u_i in which $\psi_C(e_1) = u_{i-1}u_i$ and $\psi_C(e_2) = u_iu_{i+1}$. Then C can be obtained from H_1 and T^1 by identifying $u_i \in V(H_1)$ and $u_i \in V(T^1)$. We next prove that T^1 is a star.

Suppose that T^1 is not a star. This implies that T^1 has a (p', q')-bipartition with $p', q' \geq 2$. Let T' be the tree obtained by attaching $p'-1$ and $q'-1$ pendant vertices to the two vertices u and w of a P_2, respectively. By Theorem 9, we have $I_d(T') \leq I_d(T^1)$ (i.e., $h_d(T') \geq h_d(T^1)$). Let C' be the unicyclic graph obtained from H_1 and T' by identifying either $u_i \in V(H_1)$ and $u \in V(T')$, or $u_i \in V(H_1)$ and $w \in V(T')$, such that C' has a (p, q)-bipartition. Without loss of generality, we assume that C' is obtained from H_1 and T' by identifying $u_i \in V(H_1)$ and $u \in V(T')$. It is easy to check that $d_{T^1}(u_i) \leq d_{T'}(u)$. Thus $(d_{T^1}(u_i)+2)\log(d_{T^1}(u_i)+2) - d_{T^1}(u_i)\log d_{T^1}(u_i) - (d_{T'}(u)+2)\log(d_{T'}(u)+2) + d_{T'}(u)\log d_{T'}(u) \leq 0$. So we have

$$\begin{aligned} h_d(C') - h_d(C) &= h_d(H_1) + h_d(T') - 2\log 2 - d_{T'}(u)\log d_{T'}(u) + (d_{T'}(u)+2)\log(d_{T'}(u)+2) \\ &\quad - h_d(H_1) - h_d(T^1) + 2\log 2 + d_{T^1}(u_i)\log d_{T^1}(u_i) - (d_{T^1}(u_i)+2)\log(d_{T^1}(u_i)+2) \\ &= h_d(T') - h_d(T^1) + (d_{T^1}(u_i)\log d_{T^1}(u_i) - (d_{T^1}(u_i)+2)\log(d_{T^1}(u_i)+2)) \end{aligned}$$

$$- (d_{T'}(u)\log d_{T'}(u) - (d_{T'}(u)+2)\log(d_{T'}(u)+2)) \geq 0.$$

This implies $I_d(C') \leq I_d(C)$. Let B be the set of neighbors of w in T' excluding u of T'. Since $p', q' \geq 2$, we have $B \neq \emptyset$. Let H_2 (resp. T^2) be the component of $C' - B$ (resp. $C' - f$) containing w in which $\psi_{C'}(f) = uw$. Then C' can be obtained from H_2 and T^2 by identifying $w \in V(H_2)$ and $w \in V(T^2)$. Let C'' be the unicyclic graph obtained from H_2 and T^2 by identifying $u_{i+1} \in V(H_2)$ and $w \in V(T^2)$. Clearly, we may partition the vertices of C' and C'' in the same way, that is, C'' also has a (p,q)-bipartition. It is easy to check that $d_{H_2}(u_{i+1}) > d_{H_2}(w)$, where the addition is taken modulo k. By Theorem 5, we have $I_d(C'') < I_d(C') \leq I_d(C)$, a contradiction. There exist some integers $n_1, n_2, \ldots, n_k$ and k such that $C \cong C_n(n_1, n_2, \ldots, n_k)$. By Lemmas 4 and 5, we have $C \cong C_n(p-1, q-1)$. □

Proof of Theorem 12. Let V_1 and V_2 be two subsets of $V(C)$ satisfying $|V_1| = p$ and $|V_2| = q$, and such that each edge of C joins a vertex in V_1 and a vertex in V_2. The statement is trivial for $q = 1$. So we only consider $q \geq 2$ in the following. Let $r = (p+q) - q\lfloor \frac{p+q}{q} \rfloor$. We state a claim.

Claim 2. *$D(V_1) = [2^q, 1^{p-q}]$ and $D(V_2) = [(\lceil \frac{p+q}{q} \rceil)^r, (\lfloor \frac{p+q}{q} \rfloor)^{q-r}]$.* □

Proof. We only prove $D(V_1) = [2^q, 1^{p-q}]$, since $D(V_2) = [(\lceil \frac{p+q}{q} \rceil)^r, (\lfloor \frac{p+q}{q} \rfloor)^{q-r}]$ can be proved similarly. Suppose that $D(V_1) = [2^q, 1^{p-q}]$ does not hold. This implies there exist two vertices $u \in V_1$ and $v \in V_1$ satisfying $d_C(u) - d_C(v) \geq 2$. Let $C_k = u_1u_2\cdots u_ku_1$ be the cycle of C. Let T^i be the component of $C - E(C_k)$ containing u_i for $i = 1, 2, \ldots, k$. We distinguish three cases.

Case 1. $u \in V(C_k)$ and $v \in V(C_k)$.

We have $d_C(u) \geq 4$. Let $w \notin V(C_k)$ be a neighbor of u. Let A be the set of neighbors of u excluding w. Let H_1 (resp. H_2) be the component of $C - A$ (resp. $C - e$) containing u in which $\psi_C(e) = uw$. Then C can be obtained from H_1 and H_2 by identifying $u \in V(H_1)$ and $u \in V(H_2)$. Let C' be the unicyclic graph obtained from H_1 and H_2 by identifying $u \in V(H_1)$ and $v \in V(H_2)$. We may partition the vertex set of C' and C in the same way. This implies that C' has a (p,q)-bipartition. Clearly, $d_{H_2}(u) > d_{H_2}(v)$. By Theorem 5, we have $I_d(C) < I_d(C')$, a contradiction.

Case 2. $u \in V(C_k)$ and $v \notin V(C_k)$.

Let $w \in V(C_k)$ be the neighbor of u different from v. Let $C' = C - e + f$ in which $\psi_C(e) = uw$ and $\psi_C(f) = wv$. We may partition the vertex set of C' and C in the same way. Thus C' has a (p,q)-bipartition. Since $d_C(u) - d_C(v) \geq 2$,

$$\begin{aligned}
h_d(C) - h_d(C') &= d_C(u)\log d_C(u) + d_C(v)\log d_C(v) \\
&\quad - (d_C(u)-1)\log(d_C(u)-1) - (d_C(v)+1)\log(d_C(v)+1) \\
&= (d_C(u)\log d(u) - (d_C(u)-1)\log(d_C(u)-1)) \\
&\quad - ((d_C(v)+1)\log(d_C(v)+1) - d_C(v)\log d_C(v)) \\
&= (\log \xi_1 + \frac{1}{\ln 2}) - \left(\log \xi_2 + \frac{1}{\ln 2}\right) \\
&> 0,
\end{aligned}$$

where $\xi_1 \in (d_C(u)-1, d_C(u))$ and $\xi_2 \in (d_C(v), d_C(v)+1)$. Thus $I_d(C) < I_d(C')$, a contradiction.

Case 3. $u \notin V(C_k)$.

Suppose that $u \in V(T^i)$. If $v \in V(T^i)$, we denote the path from u to v by P; otherwise, we denote the path from u to u_i by P. Let $w \notin V(P)$ be a neighbor of u. Let A be the set of neighbors of u excluding w. Let H_1 (resp. H_2) be the component of $C - A$ (resp. $C - e$) containing u in which $\psi_C(e) = uw$. Then C can be obtained from H_1 and H_2 by identifying $u \in V(H_1)$ and $u \in V(H_2)$. Let C' be the unicyclic graph obtained from H_1 and H_2 by identifying $u \in V(H_1)$ and $v \in V(H_2)$. We may partition the vertex set of C and C' in the same way. Thus C' has a (p,q)-bipartition. Clearly, $d_{H_2}(u) > d_{H_2}(v)$. By Theorem 5, we have $I_d(C) < I_d(C')$, a contradiction.

Since all graphs in $\mathcal{C}^*(p,q)$ are realizations of $D(V_1) = [2^q, 1^{p-q}]$ and $D(V_2) = [(\lceil \frac{p+q}{q} \rceil)^r, (\lfloor \frac{p+q}{q} \rfloor)^{q-r}]$, this pair of degree sequences is graphic.

Using Claim 2, to prove $C \in \mathcal{C}^*(p,q)$, it suffices to show that realizations of $D(V_1) = [2^q, 1^{p-q}]$ and $D(V_2) = [(\lceil \frac{p+q}{q} \rceil)^r, (\lfloor \frac{p+q}{q} \rfloor)^{q-r}]$ are in $\mathcal{C}^*(p,q)$. Let C'' be a unicyclic graph obtained by deleting all pendant vertices in V_1, and identifying the vertices of degree 2 in V_1 with one of their neighbors (avoiding loops). It is easy to check that C'' is a q-vertex unicyclic graph with maximum degree at most $\lceil \frac{p+q}{q} \rceil$, and C can be obtained by subdividing every edge of C'' and attaching the pendant vertices to the original vertices. Thus $C \in \mathcal{C}^*(p,q)$. □

References

[1] M. Azari, Eccentric connectivity coindex under graph operations, J. Appl. Math. Comput. 62 (2020) 23–35.
[2] D. Bonchev, D.H. Rouvray, Complexity in Chemistry, Biology and Ecology, Springer, New York, 2005.

[3] J.A. Bondy, U.S.R. Murty, Graph Theory, Springer, New York, 2008.
[4] S. Cao, M. Dehmer, Y. Shi, Extremality of degree-based graph entropies, Inf. Sci. 218 (2014) 22–33.
[5] Z. Chen, M. Dehmer, Y. Shi, A note on distance-based graph entropies, Entropy 16 (2014) 5416–5427.
[6] K.C. Das, Y. Shi, Some properties on entropies of graphs, MATCH Commun. Math. Comput. Chem. 78 (2017) 259–272.
[7] M. Dehmer, Information processing in complex networks: graph entropy and information functionals, Appl. Math. Comput. 201 (2008) 82–94.
[8] M. Dehmer, Z. Chen, A. Mowshowitz, H. Jodlbauer, F. Emmert-Streib, Y. Shi, S. Tripathi, C. Xia, On the degeneracy of the Randić entropy and related graph measures, Inf. Sci. 501 (2019) 680–687.
[9] M. Dehmer, F. Emmert-Streib, Z. Chen, X. Li, Y. Shi, Mathematical Foundations and Applications of Graph Entropy, Wiley, Weinheim, 2016.
[10] M. Dehmer, F. Emmert-Streib, M. Grabner, A computational approach to construct a multivariate complete graph invariant, Inf. Sci. 260 (2014) 200–208.
[11] M. Dehmer, A. Mowshowitz, A history of graph entropy measures, Inf. Sci. 181 (2011) 57–78.
[12] Y. Dong, S. Qiao, B. Chen, P. Wan, S. Zhang, Maximum values of degree-based entropies of bipartite graphs, Appl. Math. Comput. 401 (2021) 126094.
[13] F. Emmert-Streib, M. Dehmer, Y. Shi, Fifty years of graph matching, network alignment and comparison, Inf. Sci. 346–347 (2016) 180–197.
[14] L. Euler, Solutio problematis ad geometriam situs pertinentis, in: N.L. Biggs, E.K. Lloyd, R.J. Wilson (Eds.), Commentarii Academiae Scientiarum Imperialis Petropolitanae, vol. 8, Oxford University Press, 1976, pp. 128–140. Graph Theory 1736–1936
[15] A. Ghalavand, M. Eliasi, A.R. Ashrafi, First degree-based entropy of graphs, J. Appl. Math. Comput. 59 (2019) 37–46.
[16] I. Gutman, Degree-based topological indices, Croat. Chem. Acta 86 (2013) 351–361.
[17] M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, The first and second Zagreb indices of some graph operations, Discrete Appl. Math. 4 (2009) 804–811.
[18] M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, The hyper-wiener index of graph operations, Comput. Math. Appl. 56 (2008) 1402–1407.
[19] N. Rashevsky, Life, information theory, and topology, Bull. Math. Biophys. 17 (1955) 229–235.
[20] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 379–423.
[21] G. Simonyi, Graph entropy: a survey, in: Combinatorial optimization (New Brunswick, NJ, 1992–1993), in: DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Vol. 20, Amer. Math. Soc., Providence, RI, 1995, pp. 399–441.
[22] J. Yan, Topological structure of extremal graphs on the first degree-based graph entropies, MATCH Commun. Math, Comput. Chem. 85 (2021) 275–284.