Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 34(9): 5993–6010.
DOI: 10.3934/era.2026265
Received: 30 April 2026
Revised: 16 June 2026
Accepted: 29 June 2026
Published: 13 July 2026

Research article

On the difference between Sombor index and harmonic-arithmetic index and its applications

Zhenhua Su[1,2,*]

[1] School of Mathematics and Computational Sciences, Huaihua University, Huaihua, Hunan, China
[2] Key Laboratory of Intelligent Control Technology for Wuling-Mountain Ecological Agriculture in Hunan Province, Huaihua, Hunan, China

* **Correspondence:** Email: szh820@163.com.

Abstract: The Sombor index (SO) and harmonic-arithmetic index (HA) stand as two well-established topological indices constructed based on vertex degrees. The difference between the SO index and HA index of a graph G is introduced as

$$(SO-HA)(G)=\sum_{uv\in E(G)}\sqrt{d(u)^2+d(v)^2}-\sum_{uv\in E(G)}\frac{4d(u)d(v)}{(d(u)+d(v))^2},$$

where $d(u)$ stands for the degree of vertex $u\in V(G)$. In the present work, we determined the trees, unicyclic graphs, and bicyclic graphs that achieved extremal $SO-HA$ values, characterized the chemical graphs with extremal $SO-HA$ values, and obtained the trees with fixed maximum degree minimizing the $SO-HA$ index. Furthermore, by performing QSPR analysis on octane isomers, we observed that the $SO-HA$ index possessed favorable physicochemical predictability.

Keywords: Sombor index; harmonic-arithmetic index; difference; QSPR analysis

1. Introduction

All graphs discussed herein are simple, undirected, and connected. Let G denote such a graph with vertex set $V(G)$ and edge set $E(G)$. For a vertex $u\in V(G)$, $d(u)$ represents the degree of u, i.e., the number of vertices adjacent to u. We use Δ to denote the maximum degree of graph G. An acyclic graph G is referred to as a tree, denoted by T. Let $\mathcal{T}_{n,\Delta}$ denote the family of all trees on n vertices with maximum degree Δ. Undefined notions and terminologies conform to those presented in the graph theory monograph [1].

As a vital subfield of mathematical chemistry, chemical graph theory characterizes molecular architectures and represents chemical data in a mathematical form. Within this field, vertices denote atoms, and edges between them represent chemical bonds, which constitutes a molecular graph. The invariants of molecular graphs in graph theory are defined as topological indices. In 1947, Wiener [2] presented the first topological index, termed the Wiener index, and verified its close correlation with the boiling points of paraffins. Topological indices have become one of the most commonly used tools to facilitate chemists' analysis and prediction of the physicochemical properties of compounds.

For a graph G, its degree-based index is given by

$$TI_f = TI_f(G) = \sum_{uv\in E(G)} f(d(u), d(v)),$$

with $f(x, y) = f(y, x)$ being a real-valued function.

Gutman [3] presented the Sombor index in 2021, which is rooted in geometric distance. It is expressed as

$$SO(G) = \sum_{uv\in E(G)} \sqrt{d(u)^2 + d(v)^2}.$$

That is, by setting $f(x, y) = \sqrt{x^2 + y^2}$ in the expression of TI_f, then $TI_f(G) = SO(G)$. Since its introduction, the SO index has been extensively investigated owing to its great significance. Research topics cover its extremal problems [4, 5], chemical applications [6, 7], and correlations between the SO index and other topological indices [8, 9]. For more comprehensive surveys on the SO index, readers may refer to the review papers [10, 11].

In 2023, Albalahi et al. [12] proposed a new degree-based topological index, termed the harmonic-arithmetic index (HA), which is defined by considering the ratio of the harmonic mean to the arithmetic mean of $d(u)$ and $d(v)$, that is,

$$HA(G) = \sum_{uv\in E(G)} \frac{4d(u)d(v)}{(d(u) + d(v))^2}.$$

Similarly, the problems related to the HA index have attracted considerable attention from researchers [13–15]. In particular, several inequalities relating the HA index to various degree-based indices were established by Ali et al. [16].

Recently, the difference between two well-established topological indices has attracted growing research attention. Aarthi et al. [17] derived analytical results concerning the difference between the atom-bond sum-connectivity index and the Randić index. To date, no further progress has been made on related problems concerning index differences.

Inspired by the aforementioned literature, we mainly investigate the extremal properties of the difference index formed by the SO index and the HA index, as well as its relevant predictive performance for chemical properties. For a graph G, the difference of $SO - HA$ is defined by

$$(SO - HA)(G) = \sum_{uv\in E(G)} \sqrt{d(u)^2 + d(v)^2} - \sum_{uv\in E(G)} \frac{4d(u)d(v)}{(d(u) + d(v))^2}.$$

In the present work, we determine the extremal values of the $SO - HA$ index for trees, unicyclic graphs, and bicyclic graphs in Section 2. Section 3 derives the extremal bounds for chemical graphs. Section 4 presents the minimum $SO - HA$ index for trees with fixed maximum degree. The Quantitative Structure-Property Relationships (QSPR) analysis on octane isomers is carried out in Section 5, which shows the excellent predictive performance of the $SO - HA$ index. Section 6 outlines future research directions.

2. Extremal graphs for SO-HA among trees, unicyclic, and bicyclic graphs

Lemma 2.1. *(Gutman [3]) Let T be a nontrivial tree with order $n \geq 3$. Then,*

$$2\sqrt{2}(n-3) + 2\sqrt{5} \leq SO(T) \leq (n-1)\sqrt{n^2-2n+2}.$$

The lower bound is attained iff $T \cong P_n$, and the upper bound iff $T \cong S_n$.

Lemma 2.2. *(Albalahi et al. [12]) Let T be a nontrivial tree with order $n \geq 3$. Then,*

$$4(1-\frac{1}{n})^2 \leq HA(T) \leq n - \frac{11}{9}.$$

The lower bound is attained iff $T \cong S_n$, and the upper bound iff $T \cong P_n$.

From the above two lemmas, we immediately obtain the following theorem.

Theorem 2.3. *Let T be a nontrivial tree with order $n \geq 3$. Then,*

$$(2\sqrt{2}-1)n + 2\sqrt{5} - 6\sqrt{2} + \frac{11}{9} \leq (SO-HA)(T) \leq (n-1)\sqrt{n^2-2n+2} - 4(1-\frac{1}{n})^2.$$

The lower bound is attained iff $T \cong P_n$, and the upper bound iff $T \cong S_n$.

Proof. By Lemmas 2.1 and 2.2, P_n is the tree that minimizes both $SO(T)$ and $-HA(T)$. Thus, P_n minimizes $(SO-HA)(T)$, and

$$(SO-HA)(T) \geq 2\sqrt{2}(n-3) + 2\sqrt{5} - n + \frac{11}{9} = 2\sqrt{2}-1) + 2\sqrt{5} - 6\sqrt{2} + \frac{11}{9}.$$

Similarly, by Lemmas 2.1 and 2.2, S_n is the tree that maximizes both $SO(T)$ and $-HA(T)$. Hence,

$$(SO-HA)(T) \leq (n-1)\sqrt{n^2-2n+2} - 4(1-\frac{1}{n})^2.$$

Combining the above arguments, the theorem follows.

Lemma 2.4. *(Cruz and Rada [4], Das [5]) Let G be a unicyclic graph with order $n \geq 4$. Then,*

$$2\sqrt{2}n \leq SO(G) \leq (n-3)\sqrt{n^2-2n+2} + 2\sqrt{n^2-2n+5} + 2\sqrt{2}.$$

Equality holds for the lower bound iff $G \cong C_n$, and for the upper bound iff $G \cong S'_n$.

Lemma 2.5. *(Ding et al. [14]) Let G be a unicyclic graph with order $n \geq 4$. Then,*

$$\frac{4(n-1)(n-3)}{n^2} + \frac{16(n-1)}{(n+1)^2} + 1 \leq HA(G) \leq n.$$

Equality holds for the lower bound iff $G \cong S'_n$, and for the upper bound iff $G \cong C_n$.

From the above two lemmas, we immediately obtain the following result.

Theorem 2.6. *Let G be a unicyclic graph with order $n \geq 4$. Then,*

$$(2\sqrt{2}-1)n \leq (SO-HA)(G) \leq g(n),$$

where $g(n) = (n-3)\sqrt{n^2-2n+2} + 2\sqrt{n^2-2n+5} + 2\sqrt{2} - \frac{4(n-1)(n-3)}{n^2} - \frac{16(n-1)}{(n+1)^2} - 1$. *Equality holds for the lower bound iff* $G \cong C_n$, *and for the upper bound iff* $G \cong S'_n$.

Proof. By Lemmas 2.4 and 2.5, C_n is the unicyclic graph that minimizes $SO(G)$ and maximizes $HA(G)$, which implies that C_n minimizes $(SO-HA)(G)$. Similarly, S'_n is the unicyclic graph that maximizes $SO(G)$ and minimizes $HA(G)$, so S'_n maximizes $(SO-HA)(G)$. Combining the above arguments, the theorem follows.

Let $P = \{(x,y) \in \mathbb{N}\times\mathbb{N} : 1 \leq x \leq y \leq n-1\} - \{(1,1),(2,2),(2,3),(3,3)\}$, let $P' = \{(x,y) \in P : x+y \leq n+2\}$, and define

$$g(x,y) = f(x,y) + 6(f(3,3)-f(2,3))(\frac{x+y}{xy}-1) + f(3,3) - 2f(2,3). \tag{2.1}$$

Lemma 2.7. *(Su and Deng [18]) Let G denote a bicyclic graph. Then,*

$$TI_f(G) = (2f(2,3)-f(3,3))n + 5f(3,3) - 4f(2,3) + g(2,2)m_{2,2} + \sum_{(x,y)\in P'} g(x,y)m_{x,y}, \tag{2.2}$$

where $g(x,y)$ is defined as in Eq (2.1), and $m_{x,y}$ denotes the number of edges in G satisfying $(d(u),d(v)) = (x,y)$.

Lemma 2.8. *Let* $f(x,y) = \frac{4xy}{(x+y)^2}$. *Then*

$$g(x,y) = \frac{4xy}{(x+y)^2} + \frac{6}{25}(\frac{x+y}{xy}-1) - \frac{23}{25} < g(2,2)$$

holds for all $\{(x,y) \in \mathbb{N}\times\mathbb{N} : 1 \leq x \leq y \leq n-1\} - \{(1,1),(1,2),(2,1),(2,2)\}$.

Proof. Since $g(2,2) = \frac{2}{25}$, it suffices to show that

$$\frac{4xy}{(x+y)^2} + \frac{6}{25}(\frac{x+y}{xy}-1) < 1$$

to establish the lemma. Let

$$F(x,y) = \frac{4xy}{(x+y)^2} + \frac{6}{25}(\frac{x+y}{xy}-1).$$

Without loss of generality, let $1 \leq x \leq y$.

Case 1. $x = y$. Then, $F(x,y) = F(x,x) = 1 + \frac{6}{25}(\frac{2}{x}-1)$, which is strictly decreasing. Since $x \neq 1, 2$, we deduce that $x \geq 3$, and thus

$$F(x,y) \leq F(3,3) = 1 + \frac{6}{25}(\frac{2}{3}-1) < 1.$$

Case 2. $1 \leq x < y$. If $x = 1$, then $F(x,y) = F(1,y) = \frac{4y}{(1+y)^2} + \frac{6}{25y}$, which is strictly decreasing. It follows that

$$F(1,y) \leq F(1,3) = \frac{12}{16} + \frac{6}{75} = \frac{83}{100} < 1.$$

If $x \geq 2$, then $y \geq x + 1$. Note that

$$\frac{4xy}{(x+y)^2} < 1, \quad \frac{x+y}{xy} - 1 = \frac{1}{y} + \frac{1}{x} - 1 < 0.$$

Hence,

$$F(x,y) = \frac{4xy}{(x+y)^2} + \frac{6}{25}\left(\frac{x+y}{xy} - 1\right) < 1.$$

This completes the proof of the lemma.

Before stating the next theorem, we introduce three special classes of bicyclic graphs with n vertices, denoted as B_1–B_3, as illustrated in Figure 1. Specifically,

B_1 is constructed by identifying a single common vertex between two cycles C_1 and C_2;

B_2 is formed by sharing a common path between cycles C_1 and C_2;

B_3 consists of two cycles C_1 and C_2 linked by an independent connecting path.

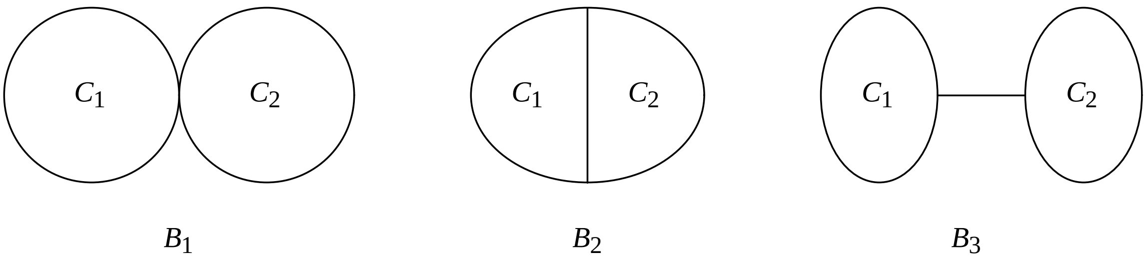

Figure 1. Bicyclic graphs of B_1–B_3.

Theorem 2.9. *Let G be a bicyclic graph with order $n \geq 5$. Then,*

$$HA(G) \leq n - 3 + 4 \cdot \frac{24}{25},$$

and the upper bound is achieved iff $G \cong B_2$ or $G \cong B_3$.

Proof. Since G is a bicyclic graph, $m_{2,2} \leq n-3$, and the unique graph satisfying $m_{2,2} = n-3$ is B_1. Consequently,

$$HA(B_1) = (n-3)f(2,2) + 4f(2,4) = n - 3 + 4 \cdot \frac{8}{9},$$

$$HA(B_2) = HA(B_3) = (n-4)f(2,2) + f(3,3) + 4f(2,3) = n - 3 + 4 \cdot \frac{24}{25}.$$

It is clear that $HA(B_1) < HA(B_2) = HA(B_3)$. Thus, we only need to consider the case $m_{2,2} \leq n-4$ in the following.

On the one hand, setting $f(x,y) = \frac{4xy}{(x+y)^2}$ in Eq (2.2) of Lemma 2.7, we have

$$HA(G) = \frac{23}{25}n + \frac{29}{25} + \frac{2}{25}m_{2,2} + \sum_{(x,y)\in P'} g(x,y)m_{x,y}. \tag{2.3}$$

On the other hand, in a bicyclic graph G, every edge of type $(1,2)$ is connected to an edge of type $(2,y)$ via a path, which implies $m_{1,2} \leq m_{2,y}$. Furthermore, we obtain

$$g(1,2) + g(2,y) \leq g(1,2) + g(2,3) = \frac{4}{45} < 2 \cdot \frac{2}{25} = 2g(2,2),$$

Combining Lemma 2.8 and Eq (2.3), we deduce that $HA(G)$ attains its maximum value when $m_{2,2} = n-4$ and $m_{x,y} = 0$ for all $(x,y) \in P'$. In this case,

$$HA(G) = \frac{23}{25}n + \frac{29}{25} + \frac{2}{25}(n-4) = n - 3 + 4 \cdot \frac{24}{25},$$

where equality holds iff $G \cong B_2$ or $G \cong B_3$.

Lemma 2.10. *(Cruz and Rada [4]) Let G be a bicyclic graph with order $n \geq 5$. Then*

$$SO(G) \geq 2\sqrt{2}n - 5\sqrt{2} + 4\sqrt{13},$$

and the lower bound is achieved iff $G \cong B_2$ or $G \cong B_3$.

Hence, by an analogous argument to Theorem 2.6, the result follows directly from Theorem 2.9 and Lemma 2.10.

Theorem 2.11. *Let G denote a bicyclic graph of order $n \geq 5$. Then,*

$$(SO - HA)(G) \geq (2\sqrt{2} - 1)n + 4\sqrt{13} - 5\sqrt{2} - \frac{21}{25},$$

the lower bound is attained if and only if $G \cong B_2$ or $G \cong B_3$.

3. Extremal chemical graphs for SO-HA

A graph G is called a chemical graph if the degree of every vertex v satisfies $1 \leq d(v) \leq 4$. In this section, we derive the sharp upper and lower bounds for the $SO - HA$ index, and start with the following lemma.

Lemma 3.1. *Let G be a connected chemical graph. For any edge $e = uv \in E(G)$ with $d(u), d(v) \geq 2$, we have*

$$(SO - HA)(G - e) < (SO - HA)(G).$$

Proof. Denote the neighborhood of u by $N_G(u) = \{u_1 = v, u_2, \ldots, u_r\}$ and the neighborhood of v by $N_G(v) = \{v_1 = u, v_2, \ldots, v_s\}$. Then $d(u) = r$ and $d(v) = s$, where $2 \leq r, s \leq 4$. Consequently, we have

$$\begin{aligned}
&(SO - HA)(G) - (SO - HA)(G - e) \\
&= \sum_{i=2}^{r}\Big(\sqrt{r^2 + d(u_i)^2} - \frac{4rd(u_i)}{(r + d(u_i))^2} - \sqrt{(r-1)^2 + d(u_i)^2} + \frac{4(r-1)d(u_i)}{(r - 1 + d(u_i))^2}\Big) \\
&+ \sum_{i=2}^{s}\Big(\sqrt{s^2 + d(v_i)^2} - \frac{4sd(v_i)}{(s + d(v_i))^2} - \sqrt{(s-1)^2 + d(v_i)^2} + \frac{4(s-1)d(v_i)}{(s - 1 + d(v_i))^2}\Big) \\
&+ \sqrt{r^2 + s^2} - \frac{4rs}{(r+s)^2}.
\end{aligned}$$

Define the function $g(r, d(u_i)) = \sqrt{r^2 + d(u_i)^2} - \frac{4rd(u_i)}{(r+d(u_i))^2}$. Under the constraints $2 \leq r \leq 4$ and $1 \leq d(u_i) \leq 4$, we compute the partial derivative of g with respect to r as follows

$$\begin{aligned}\frac{\partial g}{\partial r} &= \frac{r}{\sqrt{r^2 + d(u_i)^2}} + \frac{4d(u_i)(r - d(u_i))}{(r + d(u_i))^3} \\ &\geq \frac{2}{\sqrt{2^2 + 4^2}} + \frac{4 \cdot 4(2-4)}{(2+4)^3} = \frac{1}{\sqrt{5}} - \frac{4}{27} > 0.\end{aligned}$$

This implies that $g(r, d(u_i))$ is strictly increasing with respect to r. Accordingly,

$$\sum_{i=2}^{r}\left(\sqrt{r^2 + d(u_i)^2} - \frac{4rd(u_i)}{(r + d(u_i))^2} - \sqrt{(r-1)^2 + d(u_i)^2} + \frac{4(r-1)d(u_i)}{(r - 1 + d(u_i))^2}\right) > 0.$$

By the same argument,

$$\sum_{i=2}^{s}\left(\sqrt{s^2 + d(v_i)^2} - \frac{4sd(v_i)}{(s + d(v_i))^2} - \sqrt{(s-1)^2 + d(v_i)^2} + \frac{4(s-1)d(v_i)}{(s - 1 + d(v_i))^2}\right) > 0.$$

Furthermore, since $2 \leq r, s \leq 4$, we obtain

$$\sqrt{r^2 + s^2} \geq 2\sqrt{2} > 1 \geq \frac{4rs}{(r+s)^2}.$$

Combining the above inequalities yields $(SO - HA)(G) > (SO - HA)(G - e)$.

Theorem 3.2. *Let G denote a chemical graph with n vertices. Then,*

$$(2\sqrt{2} - 1)n + 2\sqrt{5} - 6\sqrt{2} + \frac{11}{9} \leq (SO - HA)(G) \leq 2n(4\sqrt{2} - 1).$$

The lower bound is attained iff $G \cong P_n$, and the upper bound iff G is 4-regular.

Proof. On one hand, according to Table 1, the inequality

$$F(x, y) = \sqrt{x^2 + y^2} - \frac{4xy}{(x+y)^2} > 0$$

holds for all integers x, y with $1 \leq x, y \leq 4$. Combining this with Lemma 3.1, we obtain $(SO-HA)(G-e) < (SO - HA)(G)$. It follows that the graph attaining the minimum value of $SO - HA$ must be a tree. By Theorem 2.3, we have

$$(SO - HA)(G) \geq (2\sqrt{2} - 1)n + 2\sqrt{5} - 6\sqrt{2} + \frac{11}{9},$$

with equality holding iff $G \cong P_n$.

Table 1. Values of $F(x, y) = \sqrt{x^2 + y^2} - \frac{4xy}{(x+y)^2}$.

$F(1,2) = \sqrt{5} - \frac{8}{9} = 1.3472$	$F(1,3) = \sqrt{10} - \frac{12}{16} = 2.4123$	$F(1,4) = \sqrt{17} - \frac{16}{25} = 3.4831$
$F(2,2) = \sqrt{8} - 1 = 1.8284$	$F(2,3) = \sqrt{13} - \frac{24}{25} = 2.6456$	$F(2,4) = \sqrt{20} - \frac{32}{36} = 3.5832$
$F(3,3) = \sqrt{18} - 1 = 3.2426$	$F(3,4) = \sqrt{25} - \frac{48}{49} = 4.0204$	$F(4,4) = \sqrt{32} - 1 = 4.6569$

On the other hand, as G is a chemical graph, it follows that

$$2m = \sum_{i=1}^{n} d(v_i) \le 4n,$$

that is, $m \le 2n$. From Table 1, $F(x, y) \le F(4, 4)$ for $1 \le x, y \le 4$. Therefore,

$$\begin{aligned}(SO - HA)(G) &= \sum_{v_i v_j \in E(G)} \sqrt{d(v_i)^2 + d(v_j)^2} - \frac{4d(v_i)d(v_j)}{(d(v_i) + d(v_j))^2} \\ &\le (\sqrt{32} - 1)m \le (4\sqrt{2} - 1) \cdot 2n.\end{aligned}$$

Equality holds iff $d(v_i) = d(v_j) = 4$ for every edge $v_i v_j \in E(G)$, i.e., G is 4-regular.

Combining the above arguments, the theorem follows.

4. Minimal SO-HA among trees with fixed maximum degree

Lemma 4.1. *Let $f_1(x, y) = \sqrt{x^2 + y^2}$ and $f_2(x, y) = \frac{4xy}{(x+y)^2}$. Then,*
(i) $f_1(x, y)$ is strictly increasing in x;
(ii) $f_2(x, y)$ is strictly increasing in $x < y$;
(iii) $2f_2(x, 2) - f_2(x - 1, 2)$ is strictly decreasing in $x \ge 3$.

Proof. (i) Since $\frac{\partial f_1(x,y)}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} > 0$, the first statement is readily obtained.

(ii) Note that $\frac{\partial f_2(x,y)}{\partial x} = \frac{4y(y-x)}{(x+y)^3}$. It follows that $\frac{\partial f_2(x,y)}{\partial x} > 0$ for $x < y$, and thus the second statement is valid.

(iii) Since

$$2f_2(x, 2) - f_2(x - 1, 2) = 8(\frac{2x}{(x + 2)^2} - \frac{(x - 1)}{(x + 1)^2}) = 8h(x),$$

and

$$h'(x) = \frac{-x^4 + x^3 - 18x - 20}{(x + 1)^3(x + 2)^3} < 0,$$

the third statement holds.

Lemma 4.2. *Let $g(x) = \sqrt{x^2 + 4} - \sqrt{x^2 + 1} - \frac{8x}{(x+2)^2} + \frac{4x}{(x+1)^2}$ with $x \ge 3$, then $g(x) < \sqrt{13} - \sqrt{10}$.*

Proof. Note that

$$4x^3 + 16x^2 + 16x < 8x^3 + 16x^2 + 8x,$$

which yields

$$4x(x + 2)^2 < 8x(x + 1)^2,$$

or, equivalently,

$$\frac{-8x}{(x + 2)^2} + \frac{4x}{(x + 1)^2} < 0.$$

On the other hand, since

$$g_1(x) = \sqrt{x^2 + 4} - \sqrt{x^2 + 1} = \frac{3}{\sqrt{x^2 + 4} + \sqrt{x^2 + 1}},$$

$g_1(x)$ is strictly decreasing in x. Combining the above two facts, we obtain

$$g(x) < g_1(x) + 0 \leq g_1(3) = \sqrt{13} - \sqrt{10}.$$

This concludes the proof of the lemma.

Lemma 4.3. *Let $T \in \mathcal{T}_{n,\Delta}$. If T contains a vertex of degree ≥ 3 distinct from the maximum degree vertex, then there exists $T' \in \mathcal{T}_{n,\Delta}$ such that*
(i) $SO(T) > SO(T')$;
(ii) $HA(T) < HA(T')$.

Proof. Let $d_T(u) = \Delta$ and $d_T(v) = r \geq 3$, with $N_T(v) = \{v_1, v_2, \ldots, v_r\}$. Suppose that the paths $P_1 = vv_1v_1^2 \cdots v_1^p$ and $P_2 = vv_2v_2^2 \cdots v_2^q$ are attached to v via v_1 and v_2, respectively. Then, T is depicted in Figure 2, where $u \in T_0$. Let $d_T(v_r) = s$. Then either $s < r$ or $s \geq r$.

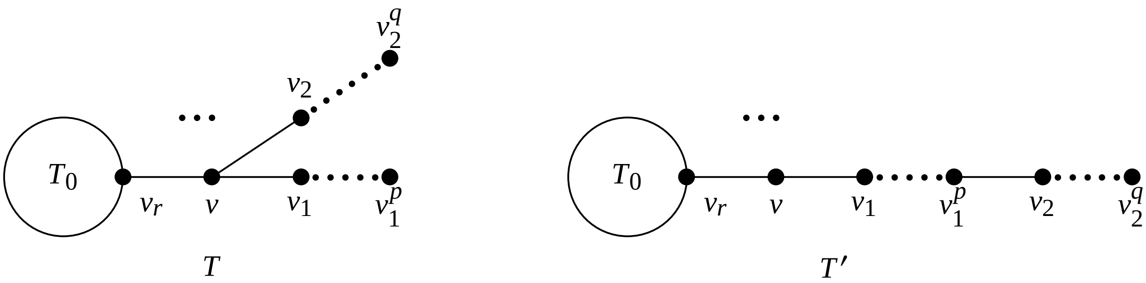

Figure 2. Trees T and T'.

Transformation 1. $T' = T - vv_2 + v_2v_1^p$.
Case 1. If $p = q = 1$. By Transformation 1, we have

$$\begin{aligned} TI_f(T) - TI_f(T') = &\sum_{i=3}^{r-1} f(r, d_T(v_i)) - \sum_{i=3}^{r-1} f(r-1, d_T(v_i)) + f(r,s) - f(r-1,s) \\ &+ 2f(r,1) - f(r-1,2) - f(2,1). \end{aligned}$$

(i) $f(x,y) = \sqrt{x^2+y^2}$. Using Lemma 4.1(i) together with $r \geq 3$, we get

$$SO(T) - SO(T') > f(r,1) - f(r-1,2) = \sqrt{r^2+4r+5} - \sqrt{r^2+2r+5} > 0.$$

Thus $SO(T) > SO(T')$.
(ii) $f(x,y) = \frac{4xy}{(x+y)^2}$. If $r > s$, then by Lemma 4.1(ii) and $r \geq 3$,

$$HA(T) - HA(T') < 2f(r,1) - f(r-1,2) - f(2,1) = \frac{8}{(r+1)^2} - \frac{8}{9} < 0.$$

If $r \leq s$, note that $max\{f(r,s) - f(r-1,s)\} = f(3,9) - f(2,9) \approx 0.1550$. By Lemma 4.1(ii),

$$\begin{aligned} HA(T) - HA(T') &< f(r,s) - f(r-1,s) + 2f(r,1) - f(r-1,2) - f(2,1) \\ &= 0.1550 + 2 \cdot \frac{3}{4} - 1 - \frac{8}{9} \approx -0.2339 < 0. \end{aligned}$$

Hence, $HA(T) < HA(T')$.

Case 2. $p \geq 2$ and $q = 1$ (the case $p = 1$ and $q \geq 2$ can be proved similarly). By Transformation 1,

$$TI_f(T) - TI_f(T') = \sum_{i=3}^{r-1} f(r, d_T(v_i)) - \sum_{i=3}^{r-1} f(r-1, d_T(v_i)) + f(r,s) - f(r-1,s) \\ + f(r,2) - f(r-1,2) + f(r,1) - f(2,2).$$

(i) $f(x,y) = \sqrt{x^2+y^2}$. By Lemma 4.1(i), we have

$$SO(T) - SO(T') > f(r,1) - f(2,2) = \sqrt{10} - \sqrt{8} > 0.$$

Thus, $SO(T) > SO(T')$.

(ii) $f(x,y) = \frac{4xy}{(x+y)^2}$. If $r > s$, then by Lemma 4.1(ii), we deduce

$$HA(T) - HA(T') < f(r,1) - f(2,2) = \frac{4r}{(r+1)^2} - 1 < 0.$$

If $r \leq s$, then $Max\{f(r,s) - f(r-1,s)\} \approx 0.1550$. By Lemma 4.1(ii), we have

$$HA(T) - HA(T') < f(r,s) - f(r-1,s) + f(r,1) - f(2,2) \\ \leq 0.1550 + \frac{3}{4} - 1 \approx -0.095 < 0.$$

Hence, $HA(T) < HA(T')$.

Case 3. $p \geq 2$ and $q \geq 2$. By Transformation 1, we have

$$TI_f(T) - TI_f(T') = \sum_{i=3}^{r-1} f(r, d_T(v_i)) - \sum_{i=3}^{r-1} f(r-1, d_T(v_i)) + f(r,s) - f(r-1,s) \\ + 2f(r,2) - f(r-1,2) + f(2,1) - 2f(2,2).$$

(i) $f(x,y) = \sqrt{x^2+y^2}$. By Lemma 4.1(i), we have

$$SO(T) - SO(T') > f(r,2) + f(2,1) - 2f(2,2) \geq \sqrt{13} + \sqrt{5} - 2\sqrt{8} \approx 0.1848 > 0.$$

Thus, $SO(T) > SO(T')$.

(ii) $f(x,y) = \frac{4xy}{(x+y)^2}$. If $r > s$, then by Lemma 4.1(ii), we have

$$HA(T) - HA(T') < f(r,2) + f(2,1) - 2f(2,2) < f(2,1) - f(2,2) = -\frac{1}{9} < 0.$$

If $r \leq s$, then $Max\{f(r,s) - f(r-1,s)\} \approx 0.1550$. Moreover, by Lemma 4.1(iii), $2f(r,2) - f(r-1,2) \leq 2f(3,2) - f(2,2) = 0.92$. Thus, we get

$$HA(T) - HA(T') < f(r,s) - f(r-1,s) + 2f(r,2) - f(r-1,2) + f(2,1) - 2f(2,2) \\ \leq 0.1550 + 0.92 + \frac{8}{9} - 2 \approx -0.0361 < 0.$$

Hence, $HA(T) < HA(T')$.

Combining all three cases, the lemma follows.

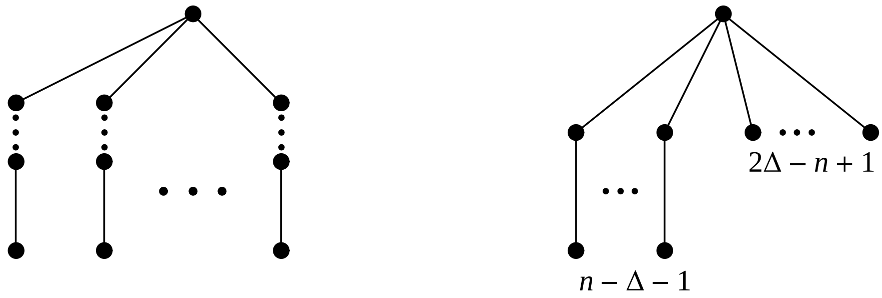

Figure 3. Trees of $T^1_{n,\Delta}$ and $T^2_{n,\Delta}$.

Before stating the following results, we first introduce the two families of trees depicted in Figure 3.

Let $T^1_{n,\Delta}$ denote the starlike tree where each pendant path has length at least 2.

Let $T^2_{n,\Delta}$ denote the tree formed by attaching $(n-\Delta-1)$ pendant edges to each noncentral vertex of the star S_Δ.

It is straightforward to verify that $T^1_{n,\Delta}$ and $T^2_{n,\Delta}$ belong to $\mathcal{T}_{n,\Delta}$.

Theorem 4.4. *Let $T \in \mathcal{T}_{n,\Delta}$ with $3 \le \Delta \le n-2$ and $n \ge 7$.*

(i) For $3 \le \Delta \le \lfloor \frac{n-1}{2} \rfloor$, then

$$(SO-HA)(T) \ge \Delta(\sqrt{\Delta^2+4} - \frac{8\Delta}{(\Delta+2)^2}) + \Delta(\sqrt{5} - \frac{8}{9}) + (n-2\Delta-1)(2\sqrt{2}-1),$$

with equality holding iff $T \cong T^1_{n,\Delta}$.

(ii) For $\lfloor \frac{n-1}{2} \rfloor < \Delta \le n-2$, then

$$\begin{aligned}(SO-HA)(T) \ge\ & (n-\Delta-1)(\sqrt{\Delta^2+4} - \frac{8\Delta}{(\Delta+2)^2}) + (n-\Delta-1)(\sqrt{5} - \frac{8}{9}) \\ & + (2\Delta-n+1)(\sqrt{\Delta^2+1} - \frac{4\Delta}{(\Delta+1)^2}),\end{aligned}$$

with equality holding iff $T \cong T^2_{n,\Delta}$.

Proof. Let $T \in \mathcal{T}^{min}_{n,\Delta}$. If T contains a vertex of degree > 2 other than the maximum degree vertex, then by Lemma 4.3, there exists a tree $T' \in \mathcal{T}_{n,\Delta}$ satisfying $SO(T') < SO(T)$ and $-HA(T') < -HA(T)$. Therefore,

$$(SO-HA)(T') < (SO-HA)(T),$$

which contradicts $T \in \mathcal{T}^{min}_{n,\Delta}$. Hence, T is a starlike tree, i.e., $T = S(P_1, P_2, \dots, P_\Delta)$, where $|P_i| = k_i$ and $k_1 \ge k_2 \ge \cdots \ge k_\Delta$.

(i) For $3 \le \Delta \le \lfloor \frac{n-1}{2} \rfloor$, we claim that $k_\Delta \ge 2$. Otherwise, $k_\Delta = 1$. Since $\lfloor \frac{n-1}{2} \rfloor \ge \Delta$, we have $n-1 \ge 2\Delta$, and thus $k_1 \ge 3$. Without loss of generality, let $P_1 = vv_1v_1^2\cdots v_1^{k_1}$ and $P_\Delta = vv_\Delta$, and tree T is depicted in Figure 4.

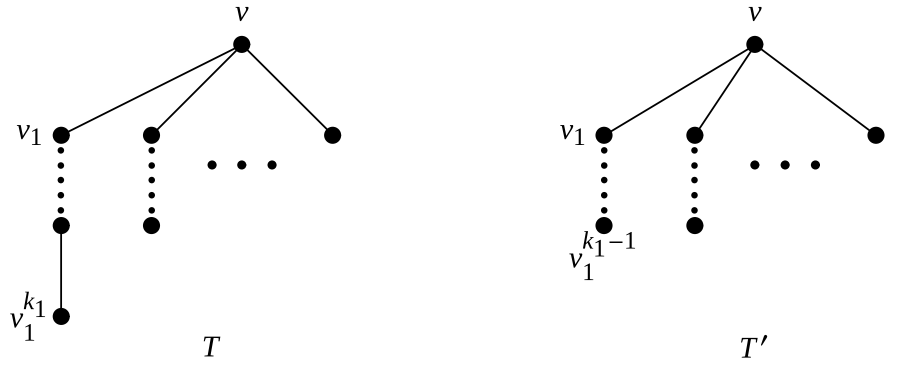

Figure 4. Trees T and T' in Transformation 2.

Transformation 2. $T' = T - v_1^{k_1-1}v_1^{k_1} + v_1^{k_1-1}v_\Delta$.

Let $f_1(x, y) = \sqrt{x^2 + y^2}$ and $f_2(x, y) = \frac{4xy}{(x+y)^2}$. By Lemma 4.2,

$$(f_1(2, \Delta) - f_2(2, \Delta)) - (f_1(1, \Delta) - f_2(1, \Delta)) < \sqrt{13} - \sqrt{10}.$$

Thus, we have

$$\begin{aligned}&(SO - HA)(T') - (SO - HA)(T)\\ &= (f_1(2, \Delta) - f_2(2, \Delta)) - (f_1(1, \Delta) - f_2(1, \Delta)) + (f_1(2, 1) - f_2(2, 1)) - (f_1(2, 2) - f_2(2, 2))\\ &< \sqrt{13} - \sqrt{10} + \sqrt{5} - \frac{8}{9} - \sqrt{8} + 1 \approx -0.0380 < 0.\end{aligned}$$

This contradicts $T \in \mathcal{T}_{n,\Delta}^{min}$. Hence, $k_\Delta \geq 2$, and we obtain

$$\begin{aligned}(SO - HA)(T) &\geq \Delta(f_1(2, \Delta) - f_2(2, \Delta)) + \Delta(f_1(1, 2) - f_2(1, 2)) + (n - 2\Delta - 1)(f_1(2, 2) - f_2(2, 2)),\\ &= \Delta(\sqrt{\Delta^2 + 4} - \frac{8\Delta}{(\Delta + 2)^2}) + \Delta(\sqrt{5} - \frac{8}{9}) + (n - 2\Delta - 1)(2\sqrt{2} - 1),\end{aligned}$$

with equality holding iff $T \cong T_{n,\Delta}^1$.

(ii) For $\lfloor \frac{n-1}{2} \rfloor < \Delta \leq n - 2$, we claim that $k_1 \leq 2$. Otherwise, $k_1 \geq 3$. Since $\lfloor \frac{n-1}{2} \rfloor < \Delta$, we have $n - 1 < 2\Delta$, and thus $k_\Delta = 1$. Let $P_1 = vv_1v_1^2 \cdots v_1^{k_1}$ and $P_\Delta = vv_\Delta$. Similarly to (i), applying Transformation 2, and by Lemma 4.2, we deduce $TI_f(T') < TI_f(T)$, which contradicts $T \in \mathcal{T}_{n,\Delta}^{min}$. Hence, $k_1 \leq 2$, and we have

$$\begin{aligned}(SO - HA)(T) \geq\ &(n - \Delta - 1)(\sqrt{\Delta^2 + 4} - \frac{8\Delta}{(\Delta + 2)^2}) + (n - \Delta - 1)(\sqrt{5} - \frac{8}{9})\\ &+ (2\Delta - n + 1)(\sqrt{\Delta^2 + 1} - \frac{4\Delta}{(\Delta + 1)^2}),\end{aligned}$$

with equality holding iff $T \cong T_{n,\Delta}^2$.

Combining (i) and (ii), the proof of the theorem is complete.

5. Physicochemical connection for SO-HA

Chemical applicability constitutes the most critical feature of a topological index, defined as its ability to accurately forecast at least one physicochemical property or biological activity among a set of molecules. In the present section, we focus primarily on exploring the application of the $SO-HA$ index and its correlation with the physicochemical properties of chemical compounds.

Since octane isomers possess 18 distinct molecular structures and well-known physicochemical properties, the QSPR analysis of octane isomers provides the most practical approach. We select five physicochemical properties as research objects: enthalpy of formation (HFORM), entropy (S), enthalpy of vaporization (HVAP), acentric factor (AF), and standard enthalpy of vaporization (DHVAP). All measured values of the octane isomers are presented in Table 2, and the raw data are retrieved from http://www.molecular-descriptors.eu/dataset/dataset.htm.

Table 2. Physicochemical properties of octane isomers.

Molecule	AF	$DHVAP$	S	$HVAP$	$HFORM$
octane	0.3979	9.915	111.67	73.19	-49.82
2-methyl-heptane	0.3779	9.484	109.84	70.30	-51.50
3-methyl-heptane	0.3710	9.521	111.26	71.30	-50.82
4-methyl-heptane	0.3715	9.483	109.32	70.91	-50.69
3-ethyl-hexane	0.3625	9.476	109.43	71.70	-50.40
2,2-dimethyl-hexane	0.3394	8.915	103.42	67.70	-53.71
2,3-dimethyl-hexane	0.3482	9.272	108.02	70.20	-51.13
2,4-dimethyl-hexane	0.3442	9.029	106.98	68.50	-52.44
2,5-dimethyl-hexane	0.3568	9.051	105.72	68.60	-53.21
3,3-dimethyl-hexane	0.3226	8.973	104.74	68.50	-52.61
3,4-dimethyl-hexane	0.3403	9.316	106.59	70.20	-50.91
2-methyl-3-ethyl-pentane	0.3324	9.209	106.06	69.70	-50.48
3-methyl-3-ethyl-pentane	0.3069	9.081	101.48	69.30	-51.38
2,2,3-trimethyl-pentane	0.3008	8.826	101.31	67.30	-52.61
2,2,4-trimethyl-pentane	0.3054	8.402	104.09	64.87	-53.57
2,3,3-trimethyl-pentane	0.2932	8.897	102.06	68.10	-51.73
2,3,4-trimethyl-pentane	0.3174	9.014	102.39	68.37	-51.97
2,2,3,3-tetramethylbutane	0.2553	8.410	93.06	66.20	-53.99

Next, we investigate the correlation between the $SO-HA$ index and physicochemical properties. For comparison, we also compute the SO index and the HA index, so that the prediction performances of the individual SO, HA, and combined $SO-HA$ indices can be evaluated. Based on the structures of octane isomers, the values of the SO, HA, and $SO-HA$ indices are calculated and presented in Table 3.

Table 3. Calculated values of the three indices of octane isomers.

Molecule	SO	HA	$SO-HA$
octane	18.6143	6.7778	11.8365
2-methyl-heptane	20.6515	6.3489	14.3026
3-methyl-heptane	20.5024	6.4478	14.0546
4-methyl-heptane	20.5024	6.4478	14.0546
3-ethyl-hexane	20.3533	6.5467	13.8066
2,2-dimethyl-hexane	24.7344	5.6978	19.0366
2,3-dimethyl-hexane	22.3995	6.0989	16.3006
2,4-dimethyl-hexane	22.5396	6.0189	16.5207
2,5-dimethyl-hexane	22.6886	5.9200	16.7686
3,3-dimethyl-hexane	24.4910	5.8356	18.6554
3,4-dimethyl-hexane	22.2504	6.1978	16.0527
2-methyl-3-ethyl-pentane	22.2504	6.1978	16.0527
3-methyl-3-ethyl-pentane	24.2477	5.9733	18.2744
2,2,3-trimethyl-pentane	26.3732	5.4985	20.8747
2,2,4-trimethyl-pentane	26.7716	5.2689	21.5027
2,3,3-trimethyl-pentane	26.2790	5.5374	20.7416
2,3,4-trimethyl-pentane	24.2967	5.7500	18.5467
2,2,3,3-tetramethylbutane	30.3955	4.8400	25.5555

We use the linear regression model

$$P = A \times TI + B$$

to investigate the correlation between indices and physicochemical properties, where TI denotes an index, P denotes a physicochemical property, and constants A and B are fitting coefficients. Using the values in Tables 2 and 3, we establish the following five linear regression models with the aid of Origin.

$$AF = -0.0117 \times SO + 0.6093, \quad AF = 0.0667 \times HA - 0.062,$$

$$AF = -0.01 \times (SO - HA) + 0.5096, \quad DHVAP = -0.125 \times SO + 12.046,$$

$$DHVAP = 0.763 \times HA + 4.5738, \quad DHVAP = -0.1078 \times (SO - HA) + 11,$$

$$S = -1.4727 \times SO + 139.8, \quad S = 8.4586 \times HA + 54.942,$$

$$S = -1.2586 \times (SO - HA) + 127.29, \quad HFORM = -0.3421 \times SO - 43.842,$$

$$HFORM = 2.2044 \times HA - 64.985, \quad HFORM = -0.2973 \times (SO - HA) - 46.662,$$

$$HVAP = -0.6303 \times SO + 83.883, \quad HVAP = 3.9339 \times HA + 45.69,$$

$$HVAP = -0.5451 \times (SO - HA) + 78.641.$$

By comparing the computed values of the three indices in Table 3 with the physicochemical properties in Table 2, and performing regression analysis using Origin, we obtain the correlation coefficients and statistical variables for each index and property, as shown in Tables 4–8. Here, we show R–correlation coefficient; R^2–coefficient of determination; $RMSE$–root mean squared error; F–ratio of two variables; P–probability value.

Table 4. Statistical parameters of the linear regression model for AF.

Indices	R	R^2	Adjusted-R^2	$RMSE$	$F(>2.5)$	$P(\leq 0.05)$
SO	-0.9594	0.9205	0.9155	0.0103	164.4989	0.0000
HA	0.9155	0.8382	0.8281	0.0147	82.8665	0.0000
$SO-HA$	-0.9547	0.9114	0.9058	0.0109	164.4989	0.0000

Table 5. Statistical parameters of the linear regression model for $DHVAP$.

Indices	R	R^2	Adjusted-R^2	$RMSE$	$F(>2.5)$	$P(\leq 0.05)$
SO	-0.9470	0.8967	0.8903	0.1270	138.9497	0.0000
HA	0.9690	0.9389	0.9351	0.0976	246.0206	0.0000
$SO-HA$	-0.9517	0.9056	0.8997	0.1214	153.56504	0.0000

Table 6. Statistical parameters of the linear regression model for S.

Indices	R	R^2	Adjusted-R^2	$RMSE$	$F(>2.5)$	$P(\leq 0.05)$
SO	-0.9465	0.8959	0.8894	1.5026	137.6671	0.0000
HA	0.9115	0.8308	0.8203	1.9153	78.5766	0.0000
$SO-HA$	-0.9430	0.8892	0.8823	1.5498	128.4523	0.0000

Table 7. Statistical parameters of the linear regression model for $HVAP$.

Indices	R	R^2	Adjusted-R^2	$RMSE$	$F(>2.5)$	$P(\leq 0.05)$
SO	-0.9032	0.8157	0.8042	0.8966	70.8203	0.0000
HA	0.9451	0.8933	0.8866	0.6823	133.9322	0.0000
$SO-HA$	-0.9107	0.8293	0.8186	0.8630	77.7260	0.0000

Table 8. Statistical parameters of the linear regression model for $HFORM$.

Indices	R	R^2	Adjusted-R^2	$RMSE$	$F(>2.5)$	$P(\leq 0.05)$
SO	-0.7937	0.6300	0.6069	0.7847	27.2442	0.0000
HA	0.8575	0.7353	0.7188	0.6637	44.4442	0.0000
$SO-HA$	-0.8042	0.6467	0.6246	0.7668	29.2867	0.0000

In statistical analysis, a result is considered significant if the F-value is greater than 2.5, or if the p-value is at most 0.05. Accordingly, all results in Tables 4–8 are statistically significant. Meanwhile, Randić and Trinajstic suggested that an index is meaningful if $R \geq 0.8$ with at least one property of the benchmark dataset. Clearly, the $|R|$-values of the SO, HA, and $SO - HA$ indices in Tables 4–8 are all greater than 0.8, implying that all three indices are chemically meaningful.

In general, a larger $|R|$-value indicates stronger predictive power. From Tables 4–8, the predictive powers of the three indices for the properties AF, S, DHVAP, HVAP, and HFORM are ordered as follows:

(i) For physicochemical properties of AF and S:

$$SO > (SO - HA) > HA;$$

(ii) For physicochemical properties of $DHVAP$, $HVAP$, and $HFORM$:

$$SO < (SO - HA) < HA.$$

Consequently, we observed that the SO index exhibits the best predictive ability for AF and S. In contrast, the HA index performs best in predicting $DHVAP$, $HVAP$, and $HFORM$. Nevertheless, the $SO - HA$ index shows favorable and more stable predictive performance for all five physicochemical properties.

6. Concluding remarks

This paper mainly investigates the extremal values of the $SO - HA$ index for trees and unicyclic graphs, as well as chemical graphs. We also characterize trees with fixed maximum degree that minimize this index. For practical application, QSPR analysis of octane isomers demonstrates that the $SO - HA$ index has excellent predictive performance for chemical properties.

This work advances the research on degree-based topological indices and their difference operations. Several directions are suggested for future studies. First, one may investigate the $SO - HA$ index on more general graphs such as tricyclic graphs. Second, further comparisons between the $SO - HA$ index and other classical topological indices can be conducted, and its applications in predicting the physicochemical properties and biological activities of organic molecules can be extended. Moreover, exploring the difference indices derived from other topological indices is also a meaningful research topic.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the Department of Education of Hunan Province (25A0571).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. J. A. Bondy, U. S. R. Murty, Graph theory with applications, *Elsevier,* New York, 1976.
2. H. Wiener, Structural determination of paraffin boiling points, *J. Amer. Chem. Soc.*, **69** (1947), 17–20.
3. I. Gutman, Geometric approach to degree-based topological indices: Sombor indices, *MATCH Commun. Math. Comput. Chem.*, **86** (2021), 11–16.
4. R. Cruz, J. Rada, Extremal values of the Sombor index in unicyclic and bicyclic graphs, *J. Math. Chem.*, **59** (2021), 1098–1116. https://doi.org/10.1007/s10910-021-01232-8
5. K. C. Das, Open problems on Sombor index of unicyclic and bicyclic graphs, *Appl. Math. Comput.*, **473** (2024), 128647. https://doi.org/10.1016/j.amc.2024.128647
6. I. Redžepović, Chemical applicability of Sombor indices, *J. Serb. Chem. Soc.*, **86** (2021), 445–457. https://doi.org/10.2298/JSC201215006R
7. H. Liu, L. You, Y. Huang, Ordering chemical graphs by Sombor indices and its applications, *MATCH Commun. Math. Comput. Chem.*, **87** (2022), 5–22. https://doi.org/10.46793/match.87-1.005L
8. S. Filipovski, Relations between Sombor index and some degree-based topological indices, *Iran. J. Math. Chem.*, **12** (2021), 19–26. https://doi.org/10.22052/IJMC.2021.240385.1533
9. Z. Wang, Y. Mao, Y. Li, B. Furtula, On relations between Sombor and other degree-based indices, *J. Appl. Math. Comput.*, **68** (2022), 1–17. https://doi.org/10.1007/s12190-021-01516-x
10. H. Liu, I. Gutman, L. You, Y. Huang, Sombor index: Review of extremal results and bounds, *J. Math. Chem.*, **66** (2022), 771–798. https://doi.org/10.1007/s10910-022-01333-y
11. B. A. Ravindrana, Sombor index on trees-A review, *MATCH Commun. Math. Comput. Chem.*, **94** (2025), 291–305. https://doi.org/10.46793/match.94-2.291B
12. A. M. Albalahi, A. Ali, A. M. Alanazi, A. A. Bhatti, A. E. Hamza, Harmonic–arithmetic index of (molecular) trees, preprint, arXiv:2302.11099.
13. K. Ramesh, S. Loganathan, Harmonic-Arithmetic index: Trees with maximum degrees and comparative analysis of antidrugs, *Symmetry*, **17** (2025), 167. https://doi.org/10.3390/sym17020167
14. G. Ding, L. Zhong, X. Wang, Harmonic-Arithmetic index of unicyclic graphs, *Iran J. Math. Chem.*, **16** (2025), 93–108. https://doi.org/10.22052/IJMC.2024.255486.1909
15. R. Kalpana, L. Shobana, Harmonic-Arithmetic index: Lower bound for *n*-order trees with fixed number of pendant vertices and monogenic semigroup for graph operations, *Int. J. Math. Math. Sci.*, **2025** (2025), 5576652. https://doi.org/10.1155/ijmm/5576652
16. A. Ali, E. Milovanović, S. Stankov, M. Matejić, I. Milovanović, Inequalities involving the harmonic-arithmetic index, *Afr. Mat.*, **35** (2024), 46. https://doi.org/10.1007/s13370-024-01183-8
17. K. Aarthi, S. Elumalai, S. Balachandran, S. Mondal, On difference between atom-bond sum-connectivity index and Randić index of graphs, *J. Appl. Math. Comput.*, **71** (2025), 2727–2748. https://doi.org/10.1007/s12190-024-02339-2

18. Z. Su, H. Deng, A note on the extremal graphs with respect to Vertex-degree-based topological indices for c-cyclic graphs, *MATCH Commun. Math. Comput. Chem.*, **95** (2026), 163–178. https://doi.org/10.46793/match.95-1.12725