MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **79** (2018) 685-716

ISSN 0340 - 6253

Mathematical Aspects of Balaban Index

Martin Knor[a], Riste Škrekovski[b,d,e], Aleksandra Tepeh[b,c]

[a]*Slovak University of Technology in Bratislava, Faculty of Civil Engineering, Department of Mathematics, Radlinského 11, 813 68, Bratislava, Slovakia,*
knor@math.sk

[b]*Faculty of Information Studies, Ljubljanska 8000 Novo Mesto, Slovenia,*
[d]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana, Slovenia,*
[e]*FAMNIT, University of Primorska, 6000 Koper, Slovenia,*
skrekovski@gmail.com

[c]*Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova ulica 17, 2000 Maribor, Slovenia,*
aleksandra.tepeh@gmail.com

(Received August 25, 2017)

Abstract

Balaban index is defined as $J(G) = \frac{m}{m-n+2} \sum \frac{1}{\sqrt{w(u) \cdot w(v)}}$, where the sum is taken over all edges of a connected graph G, n and m are the cardinalities of the vertex and the edge set of G, respectively, and $w(u)$ (resp. $w(v)$) denotes the sum of distances from u (resp. v) to all the other vertices of G. In the paper we summarize known results, clarify some ambiguities in the literature, and expose problems and conjectures on this molecular descriptor with attractive properties. In parallel, we discuss a related sum-Balaban index.

1 Introduction

A molecular graph is a connected undirected graph corresponding to structural formula of a chemical compound, so that vertices of the graph correspond to atoms of the molecule and edges of the graph correspond to the bonds between these atoms. Molecular graphs have fundamental applications in chemoinformatics, quantitative structure-property relationships (QSPR), quantitative structure-activity relationships (QSAR), virtual screening of chemical libraries, and computational drug design. QSPR, QSAR and virtual screening are based on the structure-property principle, which states that the physicochemical and biological properties of chemical compounds can be predicted from their chemical structure. One of the simplest methods that have been devised for correlating structures with biological activities or physical-chemical properties involve molecular descriptors called topological indices.

Since physical properties or bioactivities are expressed in numbers whereas chemical structures are discrete graphs, in order to associate graphs with numbers one has to rely on graph-theoretical invariants such as local vertex invariants, e.g. vertex degree, distance sum, etc. Hundreds of topological indices have been introduced so far. With respect to the invariant which plays a crucial role in the definition, we can divide topological indices into three types: degree-based indices, distance-based indices and spectrum-based indices. Degree-based indices include (general) Randić index, Zagreb index, connective eccentricity index, etc. Distance-based indices include Wiener index, Wiener polarity index, Szeged index, Kirchhoff index, ABC index, Harary index and so on. Eigenvalues of graphs, various graph energies, Estrada index etc. belong to spectrum-based indices. There are also topological indices whose definition is based on both degrees and distances such as degree distance, Gutman index, graph entropies. For more details about molecular descriptors see [50], and for further recent topics and open problems in chemical graph theory an interested reader is referred to [3, 34, 36].

Balaban index, the main subject of this paper, is a distance-based topological index. It was introduced by Alexandru T. Balaban over 30 years ago [7, 8]. To present its definition we need the following (standard) notation: for a graph G, by $V(G)$ and $E(G)$ we denote the vertex and edge sets of a graph G, respectively. We set $n = |V(G)|$ and $m = |E(G)|$. For vertices, $u, v \in V(G)$, we use $d_G(u, v)$ to denote the distance from u to v in G, and for $x \in V(G)$, the *transmission* of x (also known as the *status* or simply the *distance* of a vertex) is defined as $w(x) = \sum_{y \in V(G)} d_G(x, y)$.

Balaban index $J(G)$ of a graph G is defined as

$$J(G) = \frac{m}{m-n+2} \sum_{uv \in E(G)} \frac{1}{\sqrt{w(u) \cdot w(v)}},$$

where the sum is taken over all edges uv of G. The denominator $m - n + 2$ in the definition is used in order to have better comparability between acyclic and cyclic graphs with the same number of vertices. Recall that the *cyclomatic number* μ of G, which is the minimum number of edges that must be removed from G in order to transform it to an acyclic graph, is defined by $\mu = m - n + 1$. Thus Balaban index (often also referred to as J index) is sometimes given as

$$J(G) = \frac{m}{\mu + 1} \sum_{uv \in E(G)} \frac{1}{\sqrt{w(u) \cdot w(v)}}.$$

Balaban index was originally named the "average distance-sum connectivity index". Namely, it is based on a Randić type formula, today called the *Randić index* [45], and known also as the *connectivity index* $R(G)$, defined by

$$R(G) = \sum_{uv \in E(G)} \frac{1}{\sqrt{\deg(u) \cdot \deg(v)}},$$

where $\deg(u)$ denotes the degree of u in G. Note that in the definition of Balaban index, vertex degrees are replaced by transmissions.

Another related topological index, the so called *sum-Balaban index* $\mathrm{SJ}(G)$, was introduced in 2010 by Balaban et al. [12] and independently also by Deng [16]. As indicated by the name itself, in computing sum-Balaban index we sum up the transmissions instead of multiplying them, i.e., for a connected graph G:

$$\mathrm{SJ}(G) = \frac{m}{m - n + 2} \sum_{uv \in E(G)} \frac{1}{\sqrt{w(u) + w(v)}}.$$

The *Wiener index* of a graph G, denoted by $W(G)$, is the sum of distances between all (unordered) pairs of vertices of G

$$W(G) = \sum_{\{u,v\} \subseteq V(G)} d(u, v).$$

In [9], Balaban index is compared with Wiener index regarding the alkanes, and it was observed that Balaban index reduces the degeneracy of the later index and provides much higher discriminating ability. Therefore Balaban index is also called a "sharpened Wiener index". Note that both descriptors assume that the graphs under consideration are connected.

While the descriptive properties of Balaban index were widely discussed, its mathematical properties are less studied, which could be due to the fact that Balaban index is more difficult to handle theoretically than numerically. This also might be the reason why quite some mathematically wrong arguments appeared in the literature on this topic. The aim of this paper is to collect the existing knowledge, expose correct results and show directions in which this useful molecular descriptor with attractive properties, as well as its derived measure, sum-Balaban index, could be explored in more detail.

2 Balaban index of trees

From mathematical aspect, one direction to studying properties of each topological index is to determine the extremal values of the index among a given class of graphs. Sun [49], and also Dong and Guo [20] independently considered trees with given number of vertices. Their results hold, however Deng [15] corrected mistakes in their proofs. In the following theorem P_n and S_n denote a path and a star, respectively, on n vertices.

Theorem 1 *If T is a tree on $n \geq 2$ vertices, then*

$$(n-1)\sum_{i=1}^{n-1}\frac{1}{\sqrt{w_i w_{i+1}}} = J(P_n) \leq J(T) \leq J(S_n) = \sqrt{\frac{(n-1)^3}{2n-3}},$$

where $w_i = \frac{(n-i+1)(n-i)}{2} + \frac{(i-1)i}{2}$. Moreover, the lower bound is attained if and only if T is P_n, and the upper bound if and only if T is S_n.

The proof of Deng is based on so called *path-sliding* and *edge-lifting* transformations, which enabled him to characterize also trees of given order with the second maximal (minimal, respectively) Balaban index [15]. Both transformations increase Balaban index and were introduced already in [19].

Theorem 2 *Let G_1 and G_2 be two graphs with n_1 and n_2 vertices, respectively, $n_1, n_2 \geq 2$. If G is the graph obtained from G_1 and G_2 by adding an edge between a vertex u^* of G_1 and a vertex v^* of G_2, G' is the graph obtained by identifying u^* of G_1 to v^* of G_2 and adding a pendant edge to $u^*(v^*)$, then G' is called the* edge-lifting *transformation of G and we have $J(G) < J(G')$.*

Theorem 3 *Let G_0 be a graph with $n_0 \geq 2$ vertices, and $P = v_1 v_2 \ldots v_r$ a path of length $r-1 \geq 2$. If G (resp. G') is the graph obtained by identifying a vertex v^* of G_0 to v_{k-1} (resp. v_k) in P, $2 \leq k \leq \left\lfloor \frac{r-1}{2} \right\rfloor$, then G' is called the* path-sliding *transformation of G and we have $J(G) < J(G')$.*

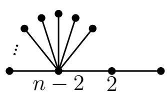

Figure 1. A double star $D_{n-2,2}$.

A double star $D_{a,b}$ is a tree consisting of $a+b$ vertices, two of which have degrees a and b, while the remaining ones have degree 1 (by symmetry, we may assume that $a \geq b$). Deng proved that $J(D_{n-x,x})$, as a (continuous) function of x, is convex. To explain how this result can be generalized (which was done in [38]), we need the notion of a discrete convex function. This concept can be introduced in several different ways (an interested reader should consult [43]), but generally a (discrete) function f is *strictly convex* if for every $x_0 < x_1 < x_2$ from the domain of f it holds

$$f(x_1) < \frac{x_2 - x_1}{x_2 - x_0} f(x_0) + \frac{x_1 - x_0}{x_2 - x_0} f(x_2).$$

However, if the domain is the set of integers greater than or equal to b, then the above property is equivalent to $2f(x_0+1) < f(x_0) + f(x_0+2)$ for all $x_0 \geq b$.

Theorem 4 *Let G be a graph with two distinct vertices u^* and v^*. Let $a \geq 2$ and $0 \leq x \leq a$. Attach x pendant edges to u^*, attach $a-x$ pendant edges to v^*, and denote the resulting graph by G_x. Then $J(G_x)$, as a (discrete) function of x, is strictly convex.*

We remark that the fact that S_n and $D_{n-2,2}$ (see Figure 1) are trees of order n with the largest and second largest, respectively, Balaban index (as originally proved by Deng) is a direct consequence of Theorem 4. Moreover, this theorem enabled us to characterize trees of order n with the third, fourth, $\ldots$ and seventh maximum value of Balaban index. The results are summarized in Table 1, and in what follows we explain the notation in the table. By T_i we denote a tree which has the i-th greatest value of Balaban index. A *caterpillar* $H_{a_1,a_2,\ldots,a_{d-1}}$ is a tree consisting of a diametric path of length d (i.e., with $d+1$ vertices) and a couple of pendant edges, such that the degrees of vertices of the diametric path are $1, a_1, a_2, \ldots, a_{d-1}, 1$ (due to symmetry, we may assume that $a_1 \geq a_{d-1}$ in $H_{a_1,a_2,\ldots,a_{d-1}}$). See Figure 2 for an example.

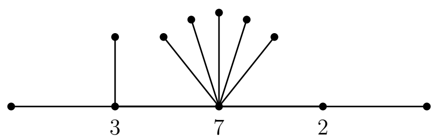

Figure 2. A caterpillar $H_{3,7,2}$.

Let $n \geq 7$. Then by R_n we denote the graph obtained from a star on $n-3$ vertices by subdividing three distinct edges, see Figure 3 for R_{11}.

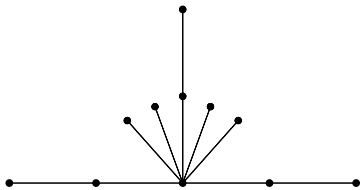

Figure 3. The graph R_{11}.

<table>
<tr><th>n</th><th>T_1</th><th>T_2</th><th>T_3</th><th>T_4</th><th>T_5</th><th>T_6</th><th>T_7</th></tr>
<tr><td>$2-3$</td><td rowspan="11">S_n</td><td></td><td rowspan="2"></td><td rowspan="3"></td><td rowspan="3"></td><td rowspan="3"></td><td rowspan="4"></td></tr>
<tr><td>4</td><td rowspan="10">$D_{n-2,2}$</td></tr>
<tr><td>5</td><td>P_5</td></tr>
<tr><td>6</td><td rowspan="8">$D_{n-3,3}$</td><td rowspan="2">$H_{2,n-3,2}$</td><td rowspan="2">$H_{n-3,2,2}$</td><td>P_6</td></tr>
<tr><td>7</td><td>$H_{3,3,2}$</td><td>R_7</td></tr>
<tr><td>8</td><td>$D_{4,4}$</td><td>$H_{2,5,2}$</td><td rowspan="2">$H_{n-3,2,2}$</td><td rowspan="2">$H_{3,n-4,2}$</td></tr>
<tr><td>9</td><td rowspan="5">$H_{2,n-3,2}$</td><td rowspan="4">$D_{n-4,4}$</td></tr>
<tr><td>10</td><td>$D_{5,5}$</td><td>$H_{n-3,2,2}$</td></tr>
<tr><td>11</td><td rowspan="2">$H_{n-3,2,2}$</td><td>$D_{6,5}$</td></tr>
<tr><td>$12-17$</td><td rowspan="2">$H_{3,n-4,2}$</td></tr>
<tr><td>$18-\infty$</td><td>$H_{n-3,2,2}$</td><td>$D_{n-4,4}$</td></tr>
</table>

Table 1. First seven trees with maximal values of Balaban index.

It can be observed that for every k, $J(T_1) - J(T_k)$ is bounded by a constant depending on k but not on n, see [38]. Moreover, we believe the following holds.

Conjecture 5 *For every k, $\lim_{n\to\infty} \big(J(T_1) - J(T_k)\big)$ is a constant.*

As mentioned earlier, Deng [15] characterized trees of given order with the second minimum Balaban index. Let T_0 be a tree obtained by attaching a pendant vertex v_n to the vertex v_2 of the path $v_1 v_2 \dots v_{n-1}$, see Figure 4.

Theorem 6 *Let T be a tree with $n \geq 4$ vertices. If T is not a path, then*

$$J(T) \geq J(T_0)$$

with equality if and only if T is T_0.

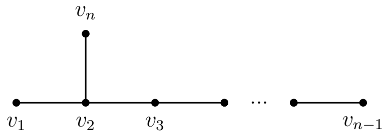

Figure 4. A tree T_0.

We believe that using Theorems 2, 3 and 4 further ranking of trees with small Balaban index can be obtained.

Problem 7 *Find trees of order n, with the third, fourth, etc. minimum Balaban index.*

Regarding the sum-Balaban index, the basic observation on trees was given by Deng [16], and Xing et al. [52]. They showed that the minimum and maximum value of sum-Balaban index among all trees of order n is, as in the case of Balaban index, attained precisely for P_n and S_n, respectively. In [52] also trees with the second-largest, and third-largest (as well as the second-smallest, and third-smallest) sum-Balaban index among the n-vertex trees for $n \geq 6$ were determined. Their proof is based on specific transformations, which increase sum-Balaban index. In [40] we presented an alternative proof as well as additional tools which enabled us to give further ranking up to seventh maximum sum-Balaban index. In fact we obtained almost the same table of results as for Balaban index. The only differences are that T_5 and T_6 are interchanged for $18 \leq n \leq 22$, and also T_6 and T_7 are interchanged for $n = 11$.

2.1 Trees satisfying certain conditions

In their study of trees with extreme Balaban index, Dong and Guo [20] discovered that *greedy trees* play a crucial role. Greedy trees were used already in [51] and are (assuming that degrees of the non-leaf vertices are given) achieved by the following 'greedy algorithm':

(1) Label a vertex with the largest degree as v (the root);

(2) Label the neighbors of v as $v_1, v_2, \ldots$, assign the largest degrees available to them so that $d(v_1) \geq d(v_2) \geq \cdots$;

(3) Label the neighbors of v_1 (except v) as $v_{11}, v_{12}, \ldots$ They take all the largest degrees available so that $\deg(v_{11}) \geq \deg(v_{12}) \geq \cdots$, then do the same for $v_2, v_3, \ldots$;

(4) Repeat (3) for all the newly labeled vertices, always start with the neighbors of the labeled vertex with largest degree whose neighbors are not labeled yet.

Theorem 8 *Given a degree sequence $\pi = (d_1, d_2, \ldots, d_n)$, the greedy tree maximizes the Balaban index among all trees with degree sequence π.*

Note that extremal graphs play the opposite roles in the cases of Balaban (sum-Balaban) and Wiener index (which is not surprising if we observe the role of distances in the definitions). Namely, a star maximizes Balaban (sum-Balaban) index and it minimizes Wiener index, while we have it vice-versa for a path. In this context we mention that it is known that the greedy tree also minimizes the Wiener index among all trees with given degree sequence, [56]. It would be interesting to explore whether greedy trees are the solution of the following problem as well.

Problem 9 *Find trees that maximize the sum-Balaban index among all n-vertex trees with given degree sequence.*

The problem of finding trees that maximize Wiener index among trees of given degree sequence is still open, see [36]. There is also no literature on trees that minimize Balaban (sum-Balaban) index among trees with given degree sequence.

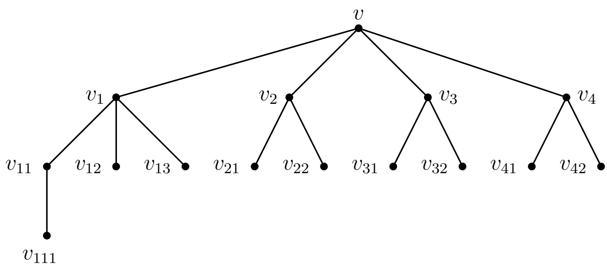

Figure 5. A greedy tree with degree sequence $(4, 4, 3, 3, 3, 2, 1, 1, \ldots, 1)$.

Recall that a partition of n is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ where the λ_i are weakly decreasing and $\sum_{i=1}^{l} \lambda_i = n$. Suppose $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_l)$ are partitions of n. Then λ *dominates* μ, if $\lambda_1 + \lambda_2 + \cdots + \lambda_i \geq \mu_1 + \mu_2 + \cdots + \mu_i$ for all $1 \leq i \leq l$ and $\sum_{i=1}^{l} \lambda_i = \sum_{i=1}^{l} \mu_i$. For the degree sequence $(d_1, \ldots, d_n)$ of any tree with n vertices, we assume that $d_1 \geq \cdots \geq d_n$. Then it can be seen as a partition of $2(n-1)$. A degree sequence is called a *dominating degree sequence* in a tree set S, if it dominates the degree sequence of any tree in S. Dong and Guo [20] noted that for greedy trees T and T' with degree sequences π and π', respectively, where π dominates π' it holds $J(T) \geq J(T')$ (and the equality holds if and only if T' is isomorphic to T). Using this observation and Theorem 8 they obtained the following.

Theorem 10 *Let S be a set of some trees with n vertices, and let T be a greedy tree in S with a dominating degree sequence. Then T has the maximum Balaban index in S.*

The above theorem enabled the authors to derive a series of theorems on certain families of n-vertex trees. These include a characterization of a tree with maximum Balaban index among all n-vertex trees with given maximum degree, among *starlike trees* (trees with just one branching vertex; i.e. vertex of degree at least three) on n vertices and given number of pendant vertices, as well as among all *chemical trees* (i.e. trees in which the maximum vertex degree is 4) of order n with k pendant vertices. See [20] for details. The extremal graphs which attain the maximum sum-Balaban index among trees with given number of vertices and maximum degree, are determined in [54].

Let $T(n; n_1, \ldots, n_k)$ denote a starlike tree on n vertices with the branching vertex u, such that k components of $T - u$ are paths of lengths $n_1 - 1, \ldots, n_k - 1$. With $S(n, k)$ we denote a starlike tree $T(n; n_1, \ldots, n_k)$ where n_i equals to $\lfloor \frac{n-1}{k} \rfloor$ or $\lceil \frac{n-1}{k} \rceil$ for $1 \leq i \leq k$. It seems that if T is a starlike tree $T(n; n_1, \ldots, n_k)$, then $\mathrm{SJ}(T) \leq \mathrm{SJ}(S(n, k))$, and equality holds if and only if T is $S(n, k)$. The following problem may have the same solutions as for Balaban index.

Problem 11 *Characterize trees that maximize sum-Balaban index among all n-vertex chemical trees with k pendant vertices.*

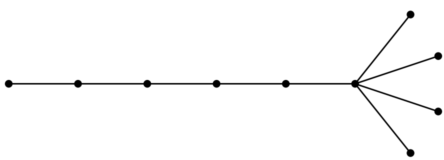

Figure 6. A comet $C(10, 5)$.

Beside the above mentioned results regarding the extremal trees in the sense of maximum values, Dong and Guo also explored the minimum value of Balaban index among

n-vertex trees with the maximum degree Δ. Let $C(n, \Delta)$ denote a *comet*, i.e., a tree obtained from a star with $\Delta - 1$ leaves and a path on $n - \Delta + 1$ vertices by identifying the central vertex of the star with an end-vertex of the path, see Figure 6 for $C(10, 5)$.

Theorem 12 *Let T be a tree with n vertices and the maximum degree Δ. Then $J(T) \geq J(C(n, \Delta))$ and the equality holds if and only if T is $C(n, \Delta)$.*

We believe that comets minimize also sum-Balaban index in the family of n-vertex trees with given maximum degree, i.e., for a tree T with n vertices and the maximum degree Δ, we have $\mathrm{SJ}(T) \geq \mathrm{SJ}(C(n, \Delta))$.

In earlier paper Dong and Guo [19] considered also n-vertex trees with given diameter (i.e. the maximum distance between any two vertices) d. In Theorem 19 they claim that the upper bound for this class of graphs is achieved for trees constructed from P_{d+1} and S_{n-d} by identifying a central vertex of the path with the central vertex of the star (recall that a central vertex in a graph is a vertex with minimum eccentricity). Our computer experiments show that the result is probably correct, however the proof in [19] is not correct as it indirectly (through Lemma 7) uses Lemma 6, which was disproved by Deng [15]. Finding a correct proof is a possible task for the future.

In [53] the authors found a tree with the maximum sum-Balaban index among all trees with n vertices and diameter d. In addition, they gave a new proof of the result that the star S_n is the graph which has the maximum sum-Balaban index among all trees with n vertices.

3 Balaban index of graphs with given order

Before we consider graphs with given order and their Balaban and sum-Balaban indices, let us mention the following general property of the two indices from [16].

Proposition 13 *Let G be a connected graph with order $n \geq 3$. Then $\mathrm{SJ}(G) \geq J(G)$ with equality if and only if $G = K_3$.*

3.1 Upper bound

The study of extremal values of Balaban index in the class of connected graphs with n vertices was initiated by Dong and Guo in [19]. They claimed that for such graphs the upper bound is attained by complete graphs, i.e.,

$$J(G) \leq J(K_n) = \frac{n^3 - n^2}{2(n^2 - 3n + 4)}$$

for any connected n-vertex graph G. But this is not true in general. The authors themselves observed a mistake and later stated in [20], that for a connected graph G with n vertices

$$J(G) \leq J(S_n) = \sqrt{\frac{(n-1)^3}{2n-3}},$$

and the equality holds only for S_n. However, it was brought to our attention that two years later (seemingly unaware of the paper by Dong and Guo), Aouchiche et al. [4] posed a conjecture, which we state here as a theorem.

Theorem 14 *For any connected graph G on $n \geq 2$ vertices, we have*

$$J(G) \leq \begin{cases} J(K_n), & \text{if } n \leq 7 \\ J(S_n), & \text{if } n \geq 8. \end{cases}$$

It turned out that in their proof from [20], Dong and Guo neglect cases with small n. A complete reasoning for the above theorem is given in [39]. We only mention that we used the following general result from [39].

Theorem 15 *Let G be a connected graph on n vertices with $\mu \geq 1$. Then*

(1) *$J(G)$ is maximum if and only of G is the complete graph K_n;*

(2) *$\mathrm{SJ}(G)$ is maximum if and only of G is the complete graph K_n.*

Using the above result, an analogous theorem was proved also for sum-Balaban index.

Theorem 16 *For any connected graph G on $n \geq 2$ vertices, we have*

$$\mathrm{SJ}(G) \leq \begin{cases} \mathrm{SJ}(K_n), & \text{if } n \leq 5 \\ \mathrm{SJ}(S_n), & \text{if } n \geq 6. \end{cases}$$

In [4] the authors also posed a conjecture about the second maximum Balaban index, which should be attained either by a complete graph with one edge missing or a comet.

Conjecture 17 *For any connected graph G on $n \geq 2$ vertices, such that G is different from S_n and K_n, we have*

$$J(G) \leq \begin{cases} J(K_n - e), & \text{if } n \leq 9, \\ J(C(n, n-2)), & \text{if } n \geq 10. \end{cases}$$

We believe that an analogous result holds for sum-Balaban index.

Conjecture 18 *For any connected graph G on $n \geq 2$ vertices, such that G is different from S_n and K_n, we have*

$$\mathrm{SJ}(G) \leq \begin{cases} \mathrm{SJ}(K_n - e), & \text{if } n \leq 5, \\ \mathrm{SJ}(C(n, n-2)), & \text{if } n \geq 6. \end{cases}$$

It was observed in [39] that for n big enough, the Balaban index of a double star always exceeds the Balaban index of the complete graph on n vertices.

Theorem 19 *Let a and b be positive integers such that $a, b \geq 2$, $a + b = n$ and $n \geq 70$. Then $J(D_{a,b}) > J(K_n)$.*

The above theorem implies the following.

Corollary 20 *For every k there exists n_0 such that for every $n \geq n_0$ the first k graphs of order n with the biggest value of Balaban index are trees.*

Similar conclusions were derived for sum-Balaban index.

Theorem 21 *Let a and b be positive integers such that $a, b \geq 2$, $a + b = n$ and $n \geq 8$. Then $\mathrm{SJ}(D_{a,b}) > \mathrm{SJ}(K_n)$.*

Corollary 22 *For every k there exists n_0 such that for every $n \geq n_0$ the first k graphs of order n with the biggest value of sum-Balaban index are trees.*

In [20] the authors point out the importance of the coefficient $\frac{|E(G)|}{\mu+1}$ in the definition of Balaban index. This coefficient was neglected in an inaccurate lemma (Lemma 3 in [19]), which stated that removing an arbitrary edge from a graph results in strictly smaller Balaban index (this statement led to the false conclusion that the upper bound is attained by K_n). Since this does not hold in general (as one can see by comparing Balaban indices of a star S_n, and a graph S_n^+ obtained from the star S_n by adding an edge [4]), Dong and Guo posed the following problem.

Problem 23 *Under what conditions $J(G - e) \leq J(G)$ (resp. $J(G - e) \geq J(G)$) for an edge e of a graph G?*

3.2 Lower bound

Using the AutoGraphiX software, Aouchiche, Caporossi and Hansen [4] showed that among all connected graphs on n vertices, the path on n vertices is not a graph for which the lower bound is attained, as claimed in [19]. For instance, C_5 has smaller Balaban index than P_5, and the same is true for two other graphs on 5 vertices, whose Balaban index is even smaler than that of C_5. Thus the following problem regarding Balaban index from Dong and Guo [19, 20] remains open, as well as an analogous problem for sum-Balaban index.

Problem 24 *Among n-vertex graphs, find those with the minimum Balaban (sum-Balaban) index.*

However, some general properties and potential candidates for extremal graphs were presented in [32] and [33].

Theorem 25 *Let G be a graph on $n \geq 4$ vertices. Then*

$$J(G) \geq \frac{4}{n-1} \qquad \textit{and} \qquad \mathrm{SJ}(G) \geq 2\sqrt{\frac{n}{n-1}}\,.$$

In [35] a class of graphs H_n of order n is constructed, for which $J(H_n) \leq \frac{32}{n}$. Hence the minimum value of Balaban index is of order $\Theta(n^{-1})$ and it tends to zero. Using more involved argument, we have proved the following lower bound in [32], which is for large n roughly twice the bound of Theorem 25. Similar result for sum-Balaban index is from [33].

Theorem 26 *Let G be a graph on n vertices, where n is big enough. Then*

$$J(G) \geq \frac{8}{n} + o(n^{-1}) \qquad \textit{and} \qquad \mathrm{SJ}(G) \geq 4 + o(1)\,.$$

By the results and arguments from [32], one would expect that a graph with the minimum Balaban index will have $\Theta(n)$ edges and vertices v with big value of $w(v)$, and analogous properties are expected also for sum-Balaban index. For small values of n the extremal graphs for both indices were found in [32, 33] and it was observed that they are either *dumbbell graphs* (i.e. graphs obtained from a path and two complete graphs, which are attached to the end-vertices of the path) or graphs similar to dumbbell graphs, see Figure 7. Motivated by this we studied the Balaban and sum-Balaban index of dumbbell graphs and graphs alike.

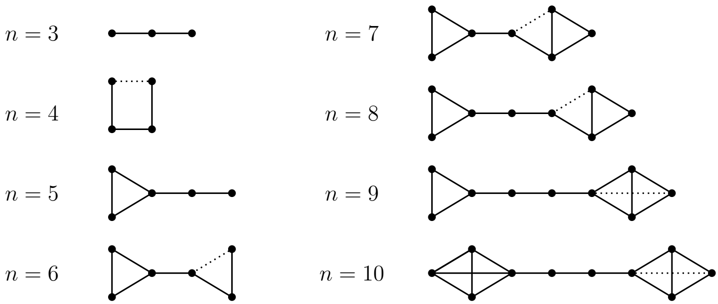

Figure 7. Graphs with the smallest value of sum-Balaban index for $n \in \{3, 4, \ldots, 10\}$. Some of the graphs contain a dotted edge. By removing such an edge we obtain the graph with the minimum value of Balaban index.

3.2.1 Bounds for balanced dumbbell graphs

First we define dumbbell graphs more precisely. Let K_a and $K'_{a'}$ be two disjoint complete graphs on a and a' vertices, respectively, and let P_b be a path on b vertices $(v_0, v_1, \ldots, v_{b-1})$ disjoint from the cliques. The *dumbbell graph* $D_{a,b,a'}$ is obtained from $K_a \cup P_b \cup K'_{a'}$ by joining all vertices of K_a with v_0 and all vertices of $K'_{a'}$ with v_{b-1}. Thus, $D_{a,b,a'}$ has $a + b + a'$ vertices. In the literature, it is often assumed that $a = a'$, here we call such graphs *balanced dumbbell graphs*. In what follows, we always assume $a \leq a'$.

Considering small values of n (up to 200), our computer tests show that among dumbbell graphs $D_{a,b,a'}$ on n vertices, the minimum value of Balaban (sum-Balaban) index is achieved for those with

$$a' = a \quad \text{or} \quad a' = a + 1. \tag{1}$$

We strongly believe this is true in general, and henceforth, we state it as a conjecture, see [32] and [33].

Conjecture 27 *Among all dumbbell graphs $D_{a,b,a'}$ on n vertices, the minimum value of Balaban (sum-Balaban) index is achieved for those with $a' = a$ or $a' = a + 1$.*

The only exception seems to be the case $n = 13$ in which the lowest sum-Balaban index among all dumbbell graphs is attained by $D_{2,7,4}$.

In the rest of this section we discuss the sizes of a and b for the optimal dumbbell graphs. When dealing with large graphs, there is not much difference between the cases $a' = a$ and $a' = a + 1$, so for the sake of simplicity, we restrict ourselves to balanced dumbbell graphs. We denote such dumbbell graphs by $D^*_{a,b}$. Thus, $D^*_{a,b}$ stands for $D_{a,b,a}$ and it has $2a+b$ vertices. In [32] we proved the following statement.

Theorem 28 *Let $D^*_{a,b}$ be a balanced dumbbell graph on n vertices, where n is big enough, with the smallest possible value of Balaban index. Then a and b are asymptotically equal to $\sqrt[4]{\pi/2}\sqrt{n}$ and n, respectively. That is, $a = \sqrt[4]{\pi/2}\sqrt{n} + o(\sqrt{n})$ and $b = n - o(n)$.*

In [33] we observed the same for sum-Balaban index, except that the constant standing by $\sqrt{n}$ is slightly different. We got $a = \sqrt[4]{\sqrt{2}\log\left(1+\sqrt{2}\right)}\sqrt{n} + o(\sqrt{n})$.

It is to be noted that π appears in Theorem 28 naturally, since the extremal balanced dumbbell graphs contain a very long path, and it is known that $\lim_{n\to\infty} J(P_n) = \pi$, see [11] (and Section 8 for more on asymptotic values of Balaban index). Theorem 28 and analogous observation for sum-Balaban index yield the following consequences.

Corollary 29 *Let D be a balanced dumbbell graph on n vertices, where n is big enough, with the minimum value of Balaban index. Then*

$$J(D) \sim \frac{1}{n}\left[\pi + 2\sqrt{2\pi} + 2\right] \doteq \frac{10.15}{n}.$$

In the corollary below the constant Q equals $\sqrt{2}\ln(1+\sqrt{2}) \doteq 1.24650$.

Corollary 30 *Let D be a balanced dumbbell graph on n vertices, where n is big enough, with the minimum value of sum-Balaban index. Then*

$$\mathrm{SJ}(D) \sim \frac{1}{\sqrt{2}}\sqrt{r-1-Q+\sqrt{(r-1-Q)^2-4Q}} \doteq 4.47934.$$

Comparing the above corollaries with the lower bound presented in Theorem 26, we see that the asymptotic value of Balaban (resp. sum-Balaban) index for optimum balanced dumbbell graph is only about 1.27 (resp. 1.12) times higher than our lower bound. Our expectation is that the optimal balanced dumbbell graph is not much different from the optimal graph. Namely, we have the following conjecture.

Conjecture 31 *Dumbbell graphs asymptotically attain the minimum value of Balaban (sum-Balaban) index among graphs on n vertices.*

3.2.2 Bounds for dumbbell-like graphs

Dumbbell-like graphs are obtained from dumbbell graphs by removing or attaching some edges from or to the cliques. More precisely, a *dumbbell-like graph*, $D^{\ell}_{a,b,a'}$, is obtained from the dumbbell graph $D_{a,b,a'}$ by either inserting ℓ edges between v_1 and K_a if $\ell > 0$, or by removing $-\ell$ edges between v_{b-1} and $K'_{a'}$ if $\ell < 0$. Note that we assume $a \leq a'$, so we always add edges to the smaller clique and remove them from the bigger one. We have the following conjecture which is supported by our computer experiments.

Conjecture 32 *Dumbbell-like graphs attain the minimum value of Balaban (sum-Balaban) index among graphs on n vertices.*

3.3 Bounds in terms of various parameters

An upper bound on the Balaban index of a connected graph in terms of its order n, size m and radius r (i.e., the minimum eccentricity of any vertex in a graph) is given by Aouchiche et al. [4]. Here we add the bound for sum-Balaban index as well.

Theorem 33 *Let G be a graph on n vertices and m edges with radius r. Then*

$$J(G) \leq \frac{2m^2}{(m-n+2)(r(r-1)+2(n-1))} \quad \textit{and}$$

$$\mathrm{SJ}(G) \leq \frac{m^2}{(m-n+2)\sqrt{r(r-1)+2(n-1)}}.$$

The bounds are best possible as shown by the complete graph K_n.

Proof. Let u be an arbitrary vertex in $V(G)$. By $\mathrm{ecc}(u)$ we denote its eccentricity. Then

$$w(u) \geq 1+2+\cdots+\mathrm{ecc}(u)+(n-\mathrm{ecc}(u)-1) \geq \frac{\mathrm{ecc}(u)(\mathrm{ecc}(u)-1)}{2}+n-1 \geq \frac{r(r-1)}{2}+n-1.$$

From this the upper bound for sum-Balaban index easily follows. For the complete graph K_n we have $r = 1$ and $m = \frac{n(n-1)}{2}$. When we substitute these values into the derived upper bound, we get exactly the sum-Balaban index of K_n. ■

Graphs with given number of vertices and diameter were first considered by Dong and Guo in [19], where they stated a result regarding the upper bound. However, the authors themselves observed a mistake in their proof and posed the problem of characterizing graphs with the maximum Balaban index among graphs with n vertices and diameter d, [20]. It seems that the for $d \geq 2$, the graph constructed from P_{d+1} and S_{n-d} by identifying a central vertex of the path with the central vertex of the star, is the answer. If $d = 1$, the answer is of course the complete graph.

The same problem would be interesting also for sum-Balaban index.

Problem 34 *Characterize graphs with the maximum sum-Balaban index among graphs with n vertices and diameter d.*

The following result from [4] improves the bound $J(G) \leq \frac{nm}{2(m-n+2)}$, given in [57].

Theorem 35 *Let G be a graph on n vertices and m edges with maximum degree Δ. Then*

$$J(G) \leq \frac{m^2}{(m-n+2)\Delta},$$

with equality if and only if G is the complete graph K_n.

A year earlier Xing et al. [52] proved this kind of bound for sum-Balaban index.

Theorem 36 *Let G be a connected graph with $n \geq 2$ vertices, m edges and maximum degree Δ. Then*

$$\mathrm{SJ}(G) \leq \frac{m}{2(m-n+2)}\sqrt{\frac{nm\Delta}{2n-2-\Delta}} \leq \frac{m}{2(m-n+2)}\sqrt{nm},$$

with the first equality if and only if G is a regular graph with diameter at most two, and with the second equality if and only if G is the complete graph.

Dong and Guo posed the problem of finding the lower bound for Balaban index [20]. Independently at the same time Ghorbani [27] showed that $J(G) \geq \frac{m}{(m-n+2)d}$, however, this bound is really rough, so the following problem is still interesting.

Problem 37 *Characterize graphs with the minimum Balaban (sum-Balaban) index among graphs with n vertices and diameter d.*

The problems regarding the extremal graphs with respect to the minimum value seem to be quite challenging. However, we think that (for Balaban as well as for sum-Balaban index) for sufficiently large n they will have structure similar to dumbbell-like graphs $D^{\ell}_{a,b,a'}$, whose diameter equals $b+1$.

Aouchiche et al. [4] obtained the lower bound for graphs where beside the number of vertices and diameter also the number of edges is prescribed [1]. To their result we add an analogous inequality for sum-Balaban index, which can easily be derived using the estimate of a vertex transmission from [4].

[1]There is a typographical error in Theorem 2 in [4], where the opposite inequality should be used.

Theorem 38 *Let G be a graph on n vertices with m edges and diameter d. Then*

$$J(G) \leq \frac{2m^2}{(m-n+2)(2nd-d(d+1))} \quad \textit{and}$$

$$\mathrm{SJ}(G) \leq \frac{m^2}{(m-n+2)\sqrt{2nd-d(d+1)}}.$$

The equalities hold if and only if $d=1$ and G is the complete graph K_n.

For more bounds in terms of parameters like spectral radius, Wiener index, clique number, minimum degree of a graph, etc. see [4, 52, 57].

4 Unicyclic graphs

Let G be a unicyclic graph (i.e., a graph containing exactly one cycle) on n vertices. Then $m=n$, $\mu(G)=1$, and thus $J(G) = \frac{n}{2}\sum_{uv\in E(G)} \frac{1}{\sqrt{w(u)\cdot w(v)}}$.

In [4] Aouchiche et al. stated a conjecture about the bounds on Balaban index for unicyclic graphs. You and Dong [21], and independently Deng and Chang [17] proved the conjecture for the upper bound. Let S_n^+ denote the graph, obtained from the star S_n by adding an edge between two nonadjacent vertices of the star (see the left hand side graph on Figure 8 for S_6^+).

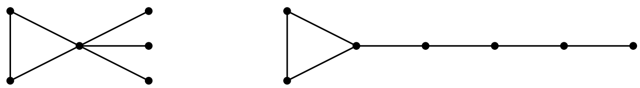

Figure 8. The graphs S_6^+ and $L_{7,3}$.

Theorem 39 *Let G be a connected unicyclic graph on $n \geq 4$ vertices. Then*

$$J(G) \leq \frac{n}{2}\left(\frac{1}{2n-4} + \frac{2}{\sqrt{(2n-4)(n-1)}} + \frac{n-3}{\sqrt{(2n-3)(n-1)}}\right),$$

and the equality holds if and only if G is S_n^+.

A star with an extra edge is also the extremal graph in the case of sum-Balaban index, [21].

Theorem 40 *Let G be a connected unicyclic graph on $n \geq 4$ vertices. Then*

$$\mathrm{SJ}(G) \leq \frac{n}{2}\left(\frac{1}{\sqrt{4n-8}} + \frac{2}{\sqrt{3n-5}} + \frac{n-3}{\sqrt{3n-4}}\right),$$

and the equality holds if and only if G is S_n^+.

Recently Fang et al. [24] characterized unicyclic n-vertex graphs with the second largest Balaban (sum-Balaban) index.

Recall that the *girth* $g = g(G)$ is the length of a shortest cycle in G. Aouchiche et al. [4] give the upper bound for the Balaban index of unicyclic graphs with given girth.

Theorem 41 *Let G be a unicyclic graph on $n \geq 3$ vertices with girth g. Then*

$$J(G) \leq \begin{cases} \frac{2n^2}{g^2}, & \text{if } n \text{ is even,} \\ \frac{2n^2}{g^2-1}, & \text{if } n \text{ is odd,} \end{cases}$$

with equality if and only if G is the cycle C_n.

One can quickly check that an analogous result holds for sum-Balaban index.

Theorem 42 *Let G be a unicyclic graph on $n \geq 3$ vertices with girth g. Then*

$$\mathrm{SJ}(G) \leq \begin{cases} \frac{\sqrt{2}n^2}{2g}, & \text{if } n \text{ is even,} \\ \frac{\sqrt{2}n^2}{2\sqrt{g^2-1}}, & \text{if } n \text{ is odd,} \end{cases}$$

with equality if and only if G is the cycle C_n.

Proof. One can check that for a cycle on at least 3 vertices we have $\mathrm{SJ}(C_n) = \frac{\sqrt{2}n}{2}$ if n is even, and $\mathrm{SJ}(C_n) = \frac{\sqrt{2}n^2}{2\sqrt{n^2-1}}$, if n is odd. Thus the equality holds if G is a cycle. Now assume G is not a cycle, and let C be the unique cycle in G, and $v \in V(G)$. If $v \in V(C)$, we obtain $w(v) = \sum_{w \in C} d(v,w) + \sum_{w \notin C} d(v,w) > \sum_{w \in C} d(v,w)$, and if $v \notin V(C)$ and u is the closest vertex to v that belongs to C, we have $w(v) > \sum_{w \in C} d(v,w) > \sum_{w \in C} d(u,w)$ (note that $d(v,w) > d(u,w)$ for every $w \in V(C)$). This means that the transmission of any vertex in G is bigger than $\frac{g^2}{4}$ if n is even, and it is bigger than $\frac{g^2-1}{4}$ if n is odd. From this observation the bound for $\mathrm{SJ}(G)$ in each of the cases can be derived. ■

Since for unicyclic graphs $m = n$, using Theorem 33, the authors derived the upper bound in terms of radius, [4].

Theorem 43 *Let G be a unicyclic graph on $n \geq 3$ vertices with radius r. Then*

$$J(G) \leq \frac{n^2}{r(r-1) + 2(n-1)}.$$

The conjecture from [4] regarding the lower bound remains unsolved. A lollipop $L_{n,g}$ is obtained from a cycle C_g and a path P_{n-g} by adding an edge between a vertex from the cycle and an endvertex from the path (see the right hand side graph of Figure 8 for $L_{7,3}$). Our computer experiments support the conjecture that lollipops are the extremal graphs for both Balaban and sum-Balaban indices.

Conjecture 44 *Let G be a connected unicyclic graph on $n \geq 5$ vertices. Then*

$$J(G) \geq J(L_{n,3}) \quad \textit{and} \quad \mathrm{SJ}(G) \geq \mathrm{SJ}(L_{n,3}),$$

and equalities hold if and only if G is $L_{n,3}$.

It would also be interesting to find a characterization of connected unicyclic graph with the minimum (maximum) Balaban (sum-Balaban) index among all unicyclic graphs of order n and girth g with k pendant vertices.

5 Bicyclic graphs

A connected graph on n vertices is said to be *bicyclic* if it contains exactly $n+1$ edges or, equivalently, it contains exactly two independent cycles. Also note that $\mu = 2$ for such graphs.

A *double lollipop* $L_{n,g1,g2}$, with $n \geq g_1+g_2$ and $g_1, g_2 \geq 3$, is the bicyclic graph obtained from two cycles C_{g_1} and C_{g_2} and a path $P_{n-g_1-g_2}$ by adding an edge between a vertex from the cycle C_{g_1} and an endpoint of the path and another edge between a vertex from the cycle C_{g_2} and the other endpoint of the path. Denote by S_n^{++} the bicyclic graph obtained from the star S_n by adding two edges with a common vertex. See Figure 9 for S_7^{++} and $L_{14,6,5}$.

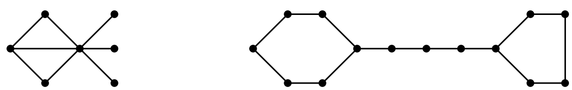

Figure 9. The graphs S_7^{++} and $L_{14,6,5}$.

After numerical experiments using AutoGraphiX, Aouchiche et al. [4] conjectured that double lollipops $L_{n,3,3}$ attain the minimum value of Balaban index. Our computer investigations indicate that the same is true for sum-Balaban index.

Conjecture 45 *Let G be a bicyclic graph on $n \geq 5$ vertices. Then*

$$J(G) \geq J(L_{n,3,3}) \quad \textit{and} \quad \mathrm{SJ}(G) \geq \mathrm{SJ}(L_{n,3,3}).$$

The bounds are attained only for the double lollipop $L_{n,3,3}$.

As we have seen so far, finding the lower bound for Balaban (sum-Balaban) index is a challenging problem in case of many graph families. On the other hand, the authors of [4] gave the conjecture on the upper bound, which turned out to be true. This was confirmed by Deng and Chang [17] in the same year, although it seems that they were not aware of the conjecture in [4]. It is surprising that three years later in the same journal the same topic (beside sum-Balaban index) is studied by Chen et al. [14] and also Fang et al. [22], who exposed some flaws in [14] and rediscovered the result of Deng and Chang (one can find this result as Theorem 3.2 in [22], where it is clear from the context that bicyclic graphs are under consideration, and not unicyclic graphs as stated by a mistake).

Theorem 46 *The graph S_n^{++} has the largest Balaban (sum-Balaban) index among all n-vertex bicyclic graphs.*

6 Graphs with high connectivity

The Balaban index of k-connected and k-edge-connected graphs of size n was first studied in [19]. However, due to already mentioned false belief that $J(G-e) < J(G)$ for any edge e of a graph G, the upper bound given by Dong and Guo in [19] does not hold as observed by the authors themselves in [20] and independently by Aouchiche et al. [4]. Thus Dong and Guo posed a problem of characterizing graphs with the maximum (minimum) Balaban index among k-connected (k-edge-connected) graphs with n vertices. Although the case of the minumum Balaban index may be hard to solve, Theorems 14 and 15 yield the following corollaries.

Corollary 47 *Let G be a graph with the maximum value of Balaban index in the class of k-connected (k-edge-connected) graphs of order n. Then we have:*

(1) *if $k = 1$ and, $n = 2$ or $n \geq 8$, then G is the star S_n;*

(2) *if $k = 1$ and $n \leq 7$, or $k \geq 2$, then G is the complete graph K_n.*

Analogously, by Theorems 15 and 16 we have:

Corollary 48 *Let G be a graph with the maximum value of sum-Balaban index in the class of k-connected (k-edge-connected) graphs of order n. Then we have:*

(1) *if $k = 1$ and, $n = 2$ or $n \geq 6$, then G is the star S_n;*

(2) *if* $k = 1$ *and* $n \leq 5$*, or* $k \geq 2$*, then* G *is the complete graph* K_n*.*

The problem for the minimum from [19, 20] remains open.

Problem 49 *Characterize graphs with the minimum Balaban (sum-Balaban) index among k-connected (k-edge-connected) graphs with n vertices.*

7 Balaban index of regular graphs

Regular graphs, more precisely their subclass, vertex transitive graphs, were first considered by Ghorbani in [27]. He gave a formula for Balaban index of a k-regular vertex transitive graph with n vertices and m edges. After a minor correction this formula reads as follows[2] :

$$J(G) = \frac{mn^2k}{4(m-n+2)W(G)}.$$

General r-regular graphs with $r \geq 3$ were studied in [35], where an upper bound for Balaban index for these graphs was obtained.

Theorem 50 *Let* G *be an* r*-regular graph on* n *vertices with* $r \geq 3$*. Then*

$$J(G) \leq \frac{r^2(r-1)^2}{2(r-2)^2 \lfloor \log_{r-1} \frac{(r-2)n+2}{r} \rfloor}.$$

The bound itself can be improved, but we wanted to keep it in as elegant form as possible, since the real value of this result can be seen from its corollary.

Corollary 51 *Let* $r \geq 3$*. For* r*-regular graphs* G *on* n *vertices it holds*

$$\lim_{n \to \infty} J(G) = 0.$$

Namely, Balaban index of regular graphs which are really big in the number of vertices, is close to 0. The number of such graphs is enormously large, and we conclude that the Balaban index does not distinguish them well. There are many (cubic) graphs interesting from the chemists point of view, for which the above observation apply; we devote to them Section 10. From this perspective also a class of cubic multigraphs, so called annulenes, is interesting. Hence, one could extend the study of Balaban index to multigraphs, see [11], where two alternative definitions for Balaban index of multigraphs are discussed. The following problem is interesting as well.

---

[2] We advise an interested reader to be alert when reading the paper as it contains several flaws, starting with incorrect computations of Balaban index for stars and complete graphs, to not precisely written definition of Randić index, etc.

Conjecture 52 *For graphs G on n vertices, such that the degree of every vertex lies in the interval $[a,b]$, $a \geq 3$, it holds*

$$\lim_{n\to\infty} J(G) = 0.$$

Regarding sum-Balaban index, the upper bound was given by Lei and Yang [42].

Theorem 53 *Let G be an r-regular graph on n vertices with $r \geq 3$. Then*

$$\mathrm{SJ}(G) \leq \frac{r^2(r-1)\sqrt{n}}{2(r-2)^{\frac{3}{2}}\sqrt{2\lfloor \log_{r-1}\frac{(r-2)n+2}{r}\rfloor}}.$$

A natural problem is also the following.

Problem 54 *For given r and n, find a graph on n vertices with maximum degree r which has the maximum value of Balaban (sum-Balaban) index.*

7.1 Cubic graphs with small value of Balaban index

Let n be even and $n \geq 10$. If $4 \nmid n$, then L_n is obtained from $(n-10)/4$ copies of $K_4 - e$ joined into a path by edges connecting the vertices of degree 2, to which at the ends we attach two pendant blocks, each on 5 vertices, see Figure 10 for L_{18}. On the other hand, if $4 \mid n$, then L_n is obtained from $(n-12)/4$ copies of $K_4 - e$, joined into a path-like manner by edges connecting the vertices of degree 2, to which ends we attach two pendant blocks, one on 5 vertices and the other on 7 vertices, see Figure 11 for L_{20}.

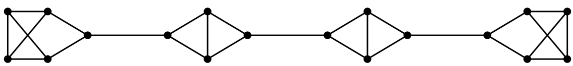

Figure 10. The graph L_{18}.

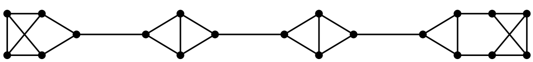

Figure 11. The graph L_{20}.

In [35] the following conjecture about L_n was proposed.

Conjecture 55 *Among n-vertex cubic graphs, L_n has the smallest Balaban (sum-Balaban) index.*

It would also be interesting to consider extremal graphs among n-vertex r-regular graphs, where $r > 3$. If r is odd we expect that extremal graphs are similar to L_n, but the case when r is even seems to be more challenging.

Problem 56 *Let $r \geq 4$. Characterize n-vertex r-regular graphs with the smallest Balaban (sum-Balaban) index.*

8 Asymptotic values

Unlike most other topological indices, owing to the factor $\frac{m}{m-n+2}$, Balaban index does not increase with the increasing number of vertices and cycles. It is thus interesting to explore the behaviour of Balaban index for various infinite classes of graphs.

Using the results from Section 3, we can observe that there are classes of graphs G_n and H_n such that $\lim_{n\to\infty} J(G_n) = \infty$ (take the star on n vertices for G_n) and $\lim_{n\to\infty} J(H_n) = 0$ (see Corollary 51). In [11] Balaban et al. discussed various infinite families of acyclic and cyclic graphs, and found many examples when J tends to a constant finite value. Interestingly, as already mentioned, J has the asymptotic value π for a path on n vertices, when n tends to infinity. They have found several examples of classes of trees whose Balaban index tends to the value 2π. Moreover, they obtained several general formulae for the asymptotic value for families of graphs with specific structure, and they observed how one can tell from the structure that the asymptotic value is a rational multiple of π.

In [37] it was proved that for every positive real number r there exists a sequence of graphs $\{G^r_{n_i}\}_{i=1}^{\infty}$, where $|V(G^r_{n_i})| = n_i$ and $\{n_i\}_{i=1}^{\infty}$ is increasing, such that $\lim_{n_i\to\infty} J(G^r_{n_i})$ $= r$. In fact, the sequence of corresponding graphs $\{G^r_{n_i}\}_{i=1}^{\infty}$ is very simple. Let $Q_{a,b}$ be a graph obtained from a clique K_a and a path P_b by joining one vertex of the clique with an endvertex of the path (see Figure 12 for $Q_{6,3}$). Then $|V(Q_{a,b})| = a + b$.

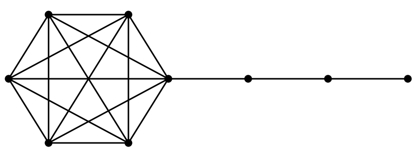

Figure 12. The graph $Q_{6,3}$.

Theorem 57 *Let $r \in \mathbb{R}$, $r > 0$, and let $\{b_a\}_{a=1}^{\infty}$ be a sequence of integers such that $\lim_{a\to\infty} b_a/a = 1/\sqrt{r}$. Then $\lim_{a\to\infty} J(Q_{a,b_a}) = r$.*

To fulfill the assumptions in Theorem 57 it suffices to choose $b_a = \lfloor a/\sqrt{r} \rfloor$ for every $a \in \mathbb{N}$. Consequently, every positive number is an accumulation point for Balaban index of a class of graphs. However, the problem still remains open for specific classes of graphs, such as the chemical ones:

Problem 58 *Is it true that for every positive real number r there exists a sequence of graphs $\{G^r_{n_i}\}_{i=1}^{\infty}$, where $|V(G^r_{n_i})| = n_i$, $\{n_i\}_{i=1}^{\infty}$ is increasing and G^r_n has maximum degree at most 4, such that*

$$\lim_{n_i\to\infty} J(G^r_{n_i}) = r\,?$$

In [39] accumulation points of sum-Balaban index were considered. Recall that the constant Q equals $\sqrt{2}\ln(1+\sqrt{2}) \doteq 1.24650$, and $1 + Q + 2\sqrt{Q} \doteq 4.47934$.

Theorem 59 *Let $r \geq 1 + Q + 2\sqrt{Q}$. Further, let $\{D_{a_i,b_i}\}_{i=1}^{\infty}$ be a sequence of balanced dumbbell graphs on $n_i = 2a_i + b_i$ vertices such that $n_i \to \infty$ and*

$$\lim_{i\to\infty} \frac{a_i}{\sqrt{n_i}} = \frac{1}{\sqrt{2}}\sqrt{r - 1 - Q + \sqrt{(r-1-Q)^2 - 4Q}}.$$

Then $\lim_{i\to\infty} \mathrm{SJ}(D_{a_i,b_i}) = r$.

Although we have a conjecture that for graphs G on large number of vertices $\mathrm{SJ}(G) \geq 1+Q+2\sqrt{Q}$ (see Corollary 30 and Conjecture 27), it is proved only that $\mathrm{SJ}(G) \geq 4+o(1)$ (see Theorem 26). Hence, if our conjecture is false, then the problem of accumulation points of sum-Balaban index for values in interval $[4, 4.47934)$ remains open.

9 Balaban index vs. Randić index

In the class of trees, the star S_n maximizes the Balaban index (see Theorem 1) and minimizes the Randić index [13]. Hence, for every tree T we have $\frac{J(T)}{R(T)} \leq \frac{n-1}{\sqrt{2n-3}}$, with equality if and only if T is the star S_n. This observation was pointed out by Aouchiche et al. [4], who proposed to study an extension of this bound to the class of all connected graphs. Based on their computer experiments for $n \geq 5$ they proposed the conjecture, which turns out to be true, see the following result from [39].

Theorem 60 *For any connected graph G on $n \geq 2$ vertices, we have*

$$\frac{J(G)}{R(G)} \leq \begin{cases} \frac{n^2-n}{n^2-3n+4}, & \text{if } n \leq 4 \\ \frac{n-1}{\sqrt{2n-3}}, & \text{if } n \geq 5, \end{cases}$$

with equality if and only if G is K_n for $n \leq 4$, and for $n \geq 5$ equality holds only if G is S_n.

In the same paper a similar observation was done for the class of unicyclic graphs. For this class Gao and Lu [26] proved that S_n^+ has the minimum Randić index, but on the other hand it has the maximum Balaban index (see Theorem 39).

Theorem 61 *For any connected unicyclic graph G on $n \geq 4$ vertices, we have*

$$\frac{J(G)}{R(G)} \leq \frac{J(S_n^+)}{R(S_n^+)}$$

with equality if and only if G is S_n^+.

10 Fullerenes and nanotubes

Balaban index was often and successfully used in QSAR/QSPR modeling [18, 50]. Some recent uses can be found in [10,31,44,46,48]. In Section 2 we have already considered trees, but there are other families of graphs interesting from chemists point of view. Dendrimers, which are repetitively branched molecules, are an instance of such families. Balaban index of some infinite classes of dendrimers is computed in [5] and [47]. Regular dendrimers were considered in [28]. More recent studies are devoted to fullerene and nanotubical graphs.

10.1 Fullerene graphs

Fullerenes [41] are polyhedral molecules made of carbon atoms arranged in pentagonal and hexagonal faces, and their corresponding graphs, fullerene graphs, are 3-connected, cubic planar graphs with only pentagonal and hexagonal faces. By Corollary 51, if $\mathcal{G}$ is the class of fullerenes, then

$$\lim_{n\to\infty} \{J(G);\, G \in \mathcal{G} \text{ and } |V(G)| = n\} = 0.$$

We remark that the upper bound given in Theorem 50 is very rough. For instance, if G is the well-known Buckminster fullerene, then the bound in the mentioned theorem with

$r = 3$ gives $J(G) \leq \frac{36}{2\lfloor \log_2 62/3 \rfloor} = 4.5$, while $J(G) = 0.91$. Nevertheless, in [35] we give a better upper bound for the Balaban index of fullerene graphs, which tends to 0 for $n \to \infty$ much faster.

Theorem 62 *Let G be a fullerene graph on $n \geq 60$ vertices. Then* $J(G) \leq \frac{25}{\sqrt{n}}$.

Sum-Balaban index of fullerene graphs was considered in [42].

Theorem 63 *Let G be a fullerene graph on $n \geq 60$ vertices. Then* $\mathrm{SJ}(G) \leq 9\sqrt[4]{n}$.

10.2 Nanotubical structures

Nanotubical graphs are obtained by wrapping a hexagonal grid into a tube so that hexagons with coodrinates (x, y) and $(x + k, y + l)$ are identified, and then possibly by closing the tube with patches, also called caps, see [2]. In practice, the ratio

$$\text{length of the cylindrical part} : \text{circumference of the cylindrical part}$$

can be of order 100 000 000:1. It is a well known fact that in a nanotubical fullerene of type (k, l) on n vertices, the circumference of the cylindrical part is $(k+l)$ and the diameter of the cylindrical part is approximately $n/(k + l)$, since when n is large enough comparing to $k + l$, the caps are negligible small [1]. This encourages us to assume that nanotubical fullerenes of type (k, l) on n vertices satisfy

$$k + l \in o(n)\,.$$

In [2], Balaban and sum-Balaban indices of infinite open nanotubes were considered. The leading term depends on the circumference of the cylindrical part of the nanotubical graph, but not on its specific type.

Theorem 64 *Let G be a nanotubical graph (open or not) of type (k, l) on n vertices. Then*

$$J(G) \sim \frac{9\pi(k + l)}{2n} \quad \textit{and} \quad \mathrm{SJ}(G) \sim \frac{9\sqrt{2}}{2}\sqrt{k + l} \cdot \log(1 + \sqrt{2}).$$

Exact formulae for special kinds of nanotubical structures were determined in [23, 29, 30, 55].

11 Conclusion

The paper is a state-of-the-art presentation on mathematical properties of Balaban and sum-Balaban indices. We exposed and corrected several flaws in the literature, included some new observations, and what is more important, with the presentation of many open problems in the field we would like to encourage further studies of structural properties of Balaban and sum-Balaban indices, relations between the two indices as well as with other topological indices.

Acknowledgments: The first author acknowledges partial support by Slovak research grants VEGA 1/0026/16, VEGA 1/0142/17, APVV-0136-12 and APVV-15-0220. All authors were partially supported by Slovenian research agency ARRS, program no. P1–00383, project no. L1–4292, and National Scholarship Programme of the Slovak Republic SAIA.

References

[1] V. Andova, D. Blenkuš, T. Došlić, F. Kardoš, R. Škrekovski, On diameter of nanotubical fullerene graphs, *MATCH Commun. Math. Comput. Chem.* **73** (2015) 529–542.

[2] V. Andova, M. Knor, R. Škrekovski, Distance based indices in nanotubical structures: Balaban index and sum–Balaban index, manuscript, 2017.

[3] V. Andova, F. Kardoš, R. Škrekovski, Mathematical aspects of fullerenes, *Ars Math. Contemp.* **11** (2016) 353–379.

[4] M. Aouchiche, G. Caporossi, P. Hansen, Refutations, results and conjectures about the Balaban index, *Int. J. Chem. Model.* **5** (2013) 189–202.

[5] A. R. Ashrafi, H. Shabani, M. V. Diudea, Balaban index of dendrimers, *MATCH Commun. Math. Comput. Chem.* **69** (2013) 151–158.

[6] A. T. Balaban, Chemical graphs. VIII. Valence isomerism and general cubic graphs, *Rev. Roum. Chim.* **15** (1970) 463–485.

[7] A. T. Balaban, Highly discriminating distance based numerical descriptor, *Chem. Phys. Lett.* **89** (1982) 399–404.

[8] A. T. Balaban, Topological indices based on topological distances in molecular graphs, *Pure Appl. Chem.* **55** (1983) 199–206.

[9] A. T. Balaban, A comparison between various topological indices, particularly between the index J and Wiener index W, in: D. H. Rouvray, R. B. King (Eds.), *Topology in Chemistry Discrete Mathematics of Molecules*, Horwood, Chichester, 2002, pp. 89–112.

[10] T. S. Balaban, A. T. Balaban, D. Bonchev, A topological approach to predicting properties of infinite polymers. Part VI. Rational formulas for the normalized Wiener index and a comparison with index J , *J. Mol. Struct. (Theochem)* **535** (2001) 81–92.

[11] A. T. Balaban, N. Ionescu–Pallas, T. S. Balaban, Asymptotic values of topological indices J and J' (average distance sum connectivities) for infinite cyclic and acyclic graphs, *MATCH Commun. Math. Comput. Chem.* **17** (1985) 121–146.

[12] A. T. Balaban, P. V. Khadikar, S. Aziz, Comparison of topological indices based on iterated 'sum' versus 'product' operations, *Iranian J. Math. Chem.* **1** (2010) 43–67.

[13] B. Bollobás, P. Erdős, Graphs of extremal weights, *Ars Comb.* **50** (1998) 225–233.

[14] Z. Chen, M. Dehmer, Y. Shi, H. Yang, Sharp upper bounds for the Balaban index of bicyclic graphs, *MATCH Commun. Math. Comput. Chem.* **75** (2016) 105–128.

[15] H. Deng, On the Balaban index of trees, *MATCH Commun. Math. Comput. Chem.* **66** (2011) 253–260.

[16] H. Deng, On the sum-Balaban index, *MATCH Commun. Math. Comput. Chem.* **66** (2011) 273–284.

[17] B. Deng, A. Chang, Maximal Balaban Index of Graphs, *MATCH Commun. Math. Comput. Chem.* 70 (2013) 259–286.

[18] J. Devillers, A. T. Balaban (Eds.), *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, 1999.

[19] H. Dong, X. Guo, Character of graphs with extremal Balaban index, *MATCH Commun. Math. Comput. Chem.* **63** (2010) 799–812.

[20] H. Dong, X. Guo, Character of trees with extreme Balaban index, *MATCH Commun. Math. Comput. Chem.* **66** (2011) 261–272.

[21] X. Dong, L. You, The maximum Balaban index (sum-Balaban Index) of unicyclic graphs, *J. Math. Res. Appl.* **34** (2014) 392–402.

[22] Z. Chen, Y. Gao, Y. Shao, W. Gao, G. Jing, Z. Li, Maximum Balaban index and sum–Balaban index of bicyclic graphs, *MATCH Commun. Math. Comput. Chem.* **75** (2016) 129–156.

[23] M. Eliasi, B. Taeri, Balaban index of zigzag polyhex nanotorus, *J. Comp. Theor. Nanosci.* **4** (2007) 1174–1178.

[24] W. Fang, Y. Gao, K. Fan, Z. Li, The second largest Balaban index (sum-Balaban index) of unicyclic graphs, *J. Math. Res. Appl.* **37** (2017) 391–403.

[25] P. W. Fowler, Resistance distances in fullerene graphs, *Croat. Chem. Acta* **75** (2002) 401–408.

[26] J. Gao, M. Lu, On the Randić index of unicyclic graphs, *MATCH Commun. Math. Comput. Chem.* **53** (2005) 377–384.

[27] M. Ghorbani, Remarks on the Balaban index, *Serdica J. Comput.* **7** (2013) 25–34.

[28] A. Heydari, Balaban index of regular dendrimers, *Optoel. Adv. Materials Rapid Commun.* **11** (2011) 1260–1262.

[29] A. Iranmanesh, Y. Alizadeh, Balaban and Randić indices of IPR C_{80} fullerene isomers, zigzag nanotubes and graphene, *Int. J. Nanosci. Nanotechn.* **7** (2011) 28–34.

[30] A. Iranmanesh, A. R. Ashrafi, Balaban index of an armchair polyhex, TUC4C8 (R) and TUC4C8 (S) nanotorus, *J. Comput. Theor. Nanosci.* **4** (2007) 514–517.

[31] M. Jalali–Heravi, M. Asadollahi–Baboli, QSAR analysis of platelet–derived growth inhibitors using GA-ANN and shuffling crossvalidation, *QSAR Comb. Sci.* **27** (2008) 750–757.

[32] M. Knor, J. Kranjc, R. Škrekovski, A. Tepeh, A search for the minimum value of Balaban index, *Appl. Math. Comput.* **286** (2016) 301–310.

[33] M. Knor, J. Kranjc, R. Škrekovski, A. Tepeh, On the minimum value of sum-Balaban index, *Appl. Math. Comput.* **303** (2017) 203–210.

[34] M. Knor, R. Škrekovski, Wiener index of line graphs, in M. Dehmer, F. Emmert–Streib (Eds.), *Quantitative Graph Theory: Mathematical Foundations and Applications*, CRC Press, 2014, 279–301.

[35] M. Knor, R. Škrekovski, A. Tepeh, Balaban index of cubic graphs, *MATCH Commun. Math. Comput. Chem.* **73** (2015) 519–528.

[36] M. Knor, R. Škrekovski, A. Tepeh, Mathematical aspects of Wiener index, *Ars Math. Contemp.* **11** (2016) 327–352.

[37] M. Knor, R. Škrekovski, A. Tepeh, A note on accumulation points of Balaban index, *MATCH Commun. Math. Comput. Chem.* **11** (2017) 163–168.

[38] M. Knor, R. Škrekovski, A. Tepeh, Convexity result and trees with large Balaban index, submitted (2016).

[39] M. Knor, R. Škrekovski, A. Tepeh, Some remarks on Balaban and sum-Balaban index, submitted, 2016.

[40] M. Knor, R. Škrekovski, A. Tepeh, Trees with large sum-Balaban index, submitted, 2016.

[41] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, R. E. Smalley, C60: buckminsterfullerene, *Nature* **318** (1985) 162–163.

[42] H. Lei, H. Yang, Bounds for the sum-Balaban index and (revised) Szeged index of regular graphs, *Appl. Math. Comput.* **268** (2015) 1259–1266.

[43] K. Murota, Discrete Convex Analysis, SIAM, Philadelphia, 2003.

[44] A. Panaye, J. P. Doucet, J. Devillers, N. Marchand–Geneste, J. M. Porcher, Decision trees versus support vector machine for classification of androgen receptor ligands, *SAR QSAR Env. Res.* **19** (2008) 129–151.

[45] M. Randić, On characterization of molecular branching, *J. Am. Chem. Soc.* **97** (1975) 6609–6615.

[46] V. Rastija, M. Medrić–Sarić, QSAR Study of antioxidant activity of wine polyphenols, *Eur. J. Med. Chem.* **44** (2009) 400–408.

[47] H. Shabani, A. R. Ashrafi, M. V. Diudea, Balaban index of an infinite class of dendrimers, *Croat. Chem. Acta* **83** (2010) 439–442.

[48] W. Shu, X. Bo, Z. Zheng, S. Wang, A novel representation of RNA secondary structure based on element–contact graphs, *BMC Bioinformatics* **9** (2008) 188–196.

[49] L. Sun, Bounds on the Balaban index of trees, *MATCH Commun. Math. Comput. Chem.* **63** (2010) 813–818.

[50] R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors*, Wiley–VCH, Weinheim, 2000.

[51] H. Wang, The extremed values of the Wiener index of a tree with given degree sequence, *Discr. Appl. Math.* **156** (2008) 2647–2654.

[52] R. Xing, B. Zhou, A. Graovac, On sum–Balaban index, *Ars Comb.* **104** (2012) 211–223.

[53] L. You, H. Han, The maximum Balaban index (sum-Balaban index) of trees with given diameter, *Ars Comb.* **112** (2013) 115–128.

[54] Z. You, H. Han, The maximum sum-Balaban index of tree graph with given vertices and maximum degree, *Adv. Appl. Math.* **2** (2013) 147–151.

[55] H. Yousefi–Azari, A. R. Ashrafi, Topological indices of nanotubes, nanotori and nanostars, *Digest J. Nanomater. Biostruct.* **3** (2008) 251–255.

[56] X. D. Zhang, Q. Y. Xiang, L. Q. Xu, R. Y. Pan, The Wiener index of trees with given degree sequences, *MATCH Commun. Math. Comput. Chem.* **60** (2008) 623–644.

[57] B. Zhou, N. Trinajstić, Bounds on the Balaban index, *Croat. Chem. Acta.* **81** (2008) 319–323.