ON THE CONSTRUCTION AND COMPARISON
OF GRAPH IRREGULARITY INDICES

Tamás Réti, Edit Tóth–Laufer

Donát Bánki Faculty of Mechanical and Safety Engineering,
Óbuda University, Bécsi út 96/B, H-1034 Budapest, Hungary
e-mail: `reti.tamas@bgk.uni-obuda.hu`

ABSTRACT. Irregularity indices are generally used for quantitative characterization of topological structure of non-regular graphs. According to a widely accepted preconception, using a topological invariant (called a graph irregularity index) for that purpose, the results of graph irregularity classification should be consistent with our subjective judgements (intuitive feeling). In the case of structurally strongly similar graphs, it is difficult to select the proper irregularity index by which the irregularity ranking (ordering) of graphs can be performed in a consistent manner to our preliminary expectations. In this work we investigate various possibilities of constructing so-called composite irregularity indices obtained as a function of traditional irregularity indices and test their discriminatory performance. Moreover, it has been demonstrated in examples, that in some cases the subjective evaluation of graph irregularities leads to false conclusions. This phenomenon is called the paradox of quantitative graph irregularity characterization. From our study we have concluded that results of graph irregularity measuring depend not only on the choice of irregularity indices but it is influenced strongly on the preselected set of graphs to be investigated. Similar problems arise for the quantitative evaluation of information content, complexity or branching of molecular graphs.

1 Introduction, preliminary considerations

Throughout the paper, we consider only simple connected graphs. For a graph G with n vertices and m edges, $V(G)$ and $E(G)$ denote the set of vertices and edges, respectively. Let $d_i = d(u_i)$ be the degree of vertex u_i $(i = 1, 2, n)$, and denote by $(i, j) = (u_i, u_j)$ an edge of G connecting vertices u_i and u_j. Let $\Delta = \Delta(G)$ and $\delta = \delta(G)$ be the maximum and the minimum degrees, respectively, of vertices of G. We use the standard terminology in graph theory, for notations not defined here we refer the reader to [5, 19]. Let $A = A(G)$ be the adjacency matrix of a graph G. We denote by $\rho(G)$ the largest eigenvalue of $A(G)$ and call it the spectral radius of G.

For a connected graph G, denote by $\{m_{r,s} = m_{r,s}(G) : m_{r,s} > 0, 1 \leq s, r \leq \Delta\}$ the finite set of positive integers $m_{r,s}$ representing the numbers of edges in G with end-vertex degrees r and s. Similarly, for a graph G, denote by $\{n_r = n_r(G) : n_r > 0, 1 \leq r \leq \Delta\}$ the set of numbers n_r of vertices with degree r. For simplicity, numbers $m_{r,s}(G)$ are called the edge-parameters of G, and numbers $n_r(G)$ are called the vertex-parameters of G, respectively.

Connected graphs G_1 and G_2 are said to be edge-equivalent if for their corresponding edge-parameters sets $\{m_{r,s}(G_1) > 0\} = \{m_{r,s}(G_2) > 0\}$ holds. Analogously, graphs G_1 and G_2 are called vertex-equivalent if for their vertex-parameters sets $\{n_r(G_1) > 0\} = \{n_r(G_2) > 0\}$ is fulfilled. It is known that edge-equivalent graphs are also vertex-equivalent. A graph is called regular if all its vertices have the same degree. A graph which is not regular is called an irregular graph. A connected graph G is said to be bidegreed if its each vertex is of degree either Δ or $\delta (\Delta > \delta \geq 1)$.

A connected bidegreed bipartite graph $G(\Delta, \delta)$ is called semiregular if each vertex in the same part of bipartition has the same degree. A connected graph G is said to be harmonic (pseudo-regular) [6, 39] if there exists a positive constant $p(G)$ such that each vertex u of G has the same average neighbor degree number identical with $p(G)$. The spectral radius of a harmonic graph G is equal to $p(G)$. It is obvious that any connected R-regular graph G_R is a harmonic graph with $p(G_R) = \rho(G_R) = R$. A graph in which the maximum vertex degree is not larger than four is referred to as a chemical graph. In the literature there are several degree-based graph invariants of

the form

$$\sum_{(i,j)\in E} \varepsilon(d_i, d_j)$$

where ε is an appropriately selected non-negative function. These graph invariants (often referred to as topological indices) have proven to be useful in $QSAR/QSPR$ studies for characterizing quantitatively the structure of chemical graphs and predict molecular properties [11, 17, 18, 20, 22, 25, 29, 31, 33–35, 37, 38, 40]. Among them the most important are the first and second Zagreb indices $M_1(G)$ and $M_2(G)$ formulated as

$$M_1(G) = \sum_{i=1}^{n} d_i^2 \qquad \text{and} \qquad M_2(G) = \sum_{(i,j)\in E} d_i d_j \,.$$

Moreover the Randić index $Ra(G)$ given as

$$Ra(G) = \sum_{(i,j)\in E} 1/\sqrt{d_i d_j} \,.$$

By definition, a topological invariant $IT(G)$ is called an irregularity index of a graph G if $IT(G) \geq 0$ and $IT(G) = 0$ if and only if G is a regular graph. The majority of irregularity indices belong to the family of degree-based graph invariants, but there exist eigenvalue-based irregularity indices as well [1, 3, 4, 7, 8, 12, 15, 21, 24, 26, 27]. Widely used invariants of this type are the Albertson irregularity index [3],

$$AL(G) = \sum_{(i,j)\in E} |d_i - d_j|$$

and the Collatz–Sinogowitz irregularity index [8]

$$CS(G) = \rho(G) - \frac{2m}{n} \,.$$

Both have been extensively studied during the last decade [1, 12, 15, 24, 26]. As an example, consider the 7-vertex trees $T(3,2)$ and $T(2,3)$ depicted in Fig. 1

For both trees the corresponding Collatz–Sinogowitz irregularity index is identical: $CS(T(3,2)) = CS(T(2,3)) = 2 - 12/7 = 0.285714$, while the values of the Albertson index are strongly different: $AL(T(3,2)) = 6$ and $AL(T(2,3)) = 10$. Comparing $AL(G)$ and $CS(G)$ irregularity indices we can conclude that the practical application of $CS(G)$ index is more complicated, since the spectral radius cannot be directly deduced from the structure of graphs considered [21].
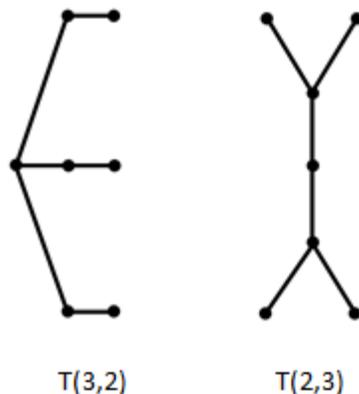
T(3,2)   T(2,3)

Figure 1: Seven-vertex trees having equal spectral radius 2.

In what follows, we concentrate our study on the construction and comparison of various irregularity indices devoted mainly to structural characterization of chemical graphs. Considering degree-based graph invariants, we evaluate their applicability for the construction of graph irregularity indices. Our investigations are extended to the critical evaluation of discriminatory power of the Collatz–Sinogowitz and the Albertson irregularity indices as well. Moreover, we study the relations between some traditional irregularity indices (Collatz–Sinogowitz irregularity index, Albertson irregularity index) and the widely used molecular descriptors applied primarily in mathematical chemistry. It will be also demonstrated that among traditional degree–based graph invariants, the Randić index and the Zagreb indices can be advantageously used for constructing various topological quantities for quantifying the irregularity of graphs.

2   Zagreb indices-based graph irregularity indices

The following Lemma serves as a useful tool to generate several types of Zagreb indices based irregularity indices.

Lemma 2.1. *[29] Let G be a connected graph with n vertices and m edges. Then $M_1(G) \geq 4m^2/n$ and $M_2(G) \geq 4m^3/n^2$ hold with equalities if and only if G is a regular graph. Based on Lemma 2.1, we can define the first and the second Zagreb irregularity indices $IRM_1(G)$ and $IRM_2(G)$ as follows [26]:*

$$IRM_1(G) = M_1(G) - \frac{4m^2}{n} \tag{1}$$

$$IRM_2(G) = M_2(G) - \frac{4m^3}{n^2} \tag{2}$$

Additionally, from Lemma 2.1 it is easy to deduce several simple irregularity indices formulated as

$$IRM_A(G) = \frac{IRM_1(G)}{2m} = \frac{M_1(G)}{2m} - \frac{2m}{n} \tag{3}$$

$$VAR(G) = \frac{M_1(G)}{n} - \left(\frac{2m}{n}\right)^2 = \left(\sqrt{\frac{M_1(G)}{n}} + \frac{2m}{n}\right)\left(\sqrt{\frac{M_1(G)}{n}} - \frac{2m}{n}\right) \tag{4}$$

$$\begin{aligned} IRM_B(G) &= \frac{IRM_2(G)}{m} = \frac{M_2(G)}{m} - \left(\frac{2m}{n}\right)^2 \\ &= \left(\sqrt{\frac{M_2(G)}{m}} + \frac{2m}{n}\right)\left(\sqrt{\frac{M_2(G)}{m}} - \frac{2m}{n}\right) \end{aligned}$$

$$IR_1(G) = \sqrt{\frac{M_1(G)}{n}} - \frac{2m}{n} \tag{5}$$

$$IR_2(G) = \sqrt{\frac{M_2(G)}{m}} - \frac{2m}{n} \tag{6}$$

Remark 2.2. *There is a broad class of connected graphs for which the identity $\rho(G) = \sqrt{M_2(G)/m}$ is fulfilled. These graphs are called Z_2 graphs, because they are defined on the basis of the second Zagreb index [2]. It is important to note that if G is an m-edge Z_2 graph, then $\rho^2(G) = M_2(G)/m$ is a positive number. The harmonic and bipartite semiregular graphs form subsets of Z_2 graphs [2]. From the previous considerations it follows that if G is a Z_2 graph then $IR_2(G) = CS(G)$ holds.*

3 Two simple methods for constructing irregularity indices

Various irregularity indices can be constructed as special cases of the general Randić index and the general zeroth-order Randić index published in Refs. [32, 41]. The general Randić index $R_\alpha(G)$ and the general zeroth-order Randić index $R_\alpha(G)$ are defined as

$$R_\alpha(G) = \sum_{(i,j)\in E} (d_i d_j)^\alpha \qquad\qquad Q_\alpha(G) = \sum_{i=1}^{n} d_i^\alpha$$

where α is a real number [41]. Depending on the particular choice of parameter α we obtain a set of degree-based topological indices which are widely used in $QSPR$ and $QSAR$ applications [41]. Consequently, $R_{-1/2}(G)$ and $R_1(G)$ are identical to the ordinary Randić index $Ra(G)$ and the second Zagreb index $M_2(G)$. Moreover, it follows that the invariant $Q_2(G)$ is equal to the first Zagreb index $M_1(G)$. In mathematical chemistry $Q_3(G)$ is referred to as the forgotten topological index denoted by $F(G)$ [17]. Additionally, $R_{1/2}(G)$ is called as reciprocal Randić index denoted by $RR(G)$ [23]. Starting with the general Randić index and the general zeroth-order Randić index, a family of different irregularity indices can be generated. To do this, consider the irregularity index of general type

$$IRZ(G, \alpha) = \sum_{(i,j)\in E} (d_i^\alpha - d_j^\alpha)^2 = \sum_{(i,j)\in E} (d_i^{2\alpha} + d_j^{2\alpha}) - 2 \sum_{(i,j)\in E} (d_i d_i)^\alpha \geq 0$$

where α is real number differing from zero. Using the transformation rule published in [14], the above inequality can be rewritten in the following alternative form:

$$\sum_{(i,j)\in E} (d_i^\alpha - d_j^\alpha)^2 = \sum_{i=1}^{n} d_i^{2\alpha+1} - 2 \sum_{(i,j)\in E} (d_i d_i)^\alpha \geq 0$$

with equality if and only if G is a regular graph. This implies that

$$IRZ(G, \alpha) = \sum_{i=1}^{n} d_i^{2\alpha+1} - 2 \sum_{(i,j)\in E} (d_i d_i)^\alpha = Q_{2\alpha+1}(G) - 2R_\alpha(G) \geq 0\,.$$

It should be noted that the above general inequality has been published in [32]. In particular cases, one obtains the following irregularity indices:

If $\alpha = 1$, then $Q_3(G) = F(G) = \sum_{i=1}^{n} d_i^3$, and this implies that

$$IRZ(G,1) = \sum_{i=1}^{n} d_i^3 - 2 \sum_{(i,j)\in E} (d_i d_j) = F(G) - 2M_2(G) \geq 0 \qquad (7)$$

If $\alpha = -1/2$, then $\sum_{i=1}^{n} d_i^0 = n$, and

$$IRZ(G,-1/2) = n - 2 \sum_{(i,j)\in E} 1/\sqrt{d_i d_j} = n - 2Ra(G) \geq 0 \,.$$

If $\alpha = 1/2$, then

$$IRZ(G,1/2) \; = \; \sum_{(i,j)\in E} (\sqrt{d_i} - \sqrt{d_j})^2$$

$$= \; \sum_{(i,j)\in E} (d_i + d_j) - 2 \sum_{(i,j)\in E} \sqrt{d_i d_j} = M_1(G) - 2RR(G) \geq 0 \,.$$

If $\alpha = 2$, then

$$IRZ(G,2) \; = \; \sum_{(i,j)\in E} (d_i^2 - d_j^2)^2 = \sum_{(i,j)\in E} (d_i^4 + d_i^4) - 2 \sum_{(i,j)\in E} (d_i d_j)^2$$

$$= \; \sum_{i=1}^{n} d_i^5 - 2R_2(G) \geq 0 \,.$$

If $\alpha = -1$, then

$$IRZ(G,-1) \; = \; \sum_{(i,j)\in E} (1/d_i - 1/d_j)^2 = \sum_{(i,j)\in E} (1/d_i^2 + 1/d_j^2) - 2 \sum_{(i,j)\in E} 1/(d_i d_j)$$

$$= \; \sum_{i=1}^{n} 1/d_i - 2 \sum_{(i,j)\in E} 1/(d_i d_j) = Q_{-1}(G) - 2R_{-1} \geq 0 \,.$$

Another simple technique for constructing irregularity indices is based on the following concept. Consider the irregularity index $IRV(G,\alpha)$ formulated as

$$IRV(G,\alpha) = \sum_{(i,j)\in E} \{(d_i d_j)^{\alpha/2} - \left(\frac{2m}{n}\right)^{\alpha}\}^2 \geq 0$$

where α is real number not equal to zero. It follows that

$$IRV(G,\alpha) = \sum_{(i,j)\in E} (d_i d_j)^{\alpha} + m \left(\frac{2m}{n}\right)^{2\alpha} - 2 \left(\frac{2m}{n}\right)^{\alpha} \sum_{(i,j)\in E} (d_i d_j)^{\alpha/2} \geq 0 \,.$$

For particular cases, we have

If $\alpha = 2$, then

$$
\begin{aligned}
IRV(G,2) &= \sum_{(i,j)\in E}(d_id_j)^2 + m\left(\frac{2m}{n}\right)^4 - 2\left(\frac{2m}{n}\right)^2\sum_{(i,j)\in E}(d_id_j) \\
&= R_2(G) + m\left(\frac{2m}{n}\right)^4 - 8\frac{m^2}{n^2}M_2(G) \geq 0 \, .
\end{aligned}
$$

If $\alpha = 1$, then

$$
\begin{aligned}
IRV(G,1) &= \sum_{(i,j)\in E}d_id_j + m\left(\frac{2m}{n}\right)^2 - 2\left(\frac{2m}{n}\right)\sum_{(i,j)\in E}\sqrt{d_id_j} \\
&= M_2(G) + 4\frac{m^3}{n^2} - \frac{4m}{n}RR(G) \geq 0 \, .
\end{aligned}
$$

If $\alpha = -1$, then

$$
\begin{aligned}
IRV(G,-1) &= \sum_{(i,j)\in E}1/d_id_j + m\left(\frac{2m}{n}\right)^{-2} - 2\left(\frac{2m}{n}\right)^{-1}\sum_{(i,j)\in E}1/\sqrt{d_id_j} \\
&= R_{-1}(G) + \frac{n^2}{4m} - \frac{n}{m}Ra(G) \geq 0 \, .
\end{aligned}
$$

4   Construction of composite irregularity indices

A topological graph invariant is called a composite index if it is generated as a function of two or more algebraically independent graph invariants. For example, the sum or the product of more irregularity indices results in a novel irregularity index. Generally, it is expected that a composite irregularity index has an increased discriminatory power comparing with the simple irregularity indices.

Surprisingly, in some cases the discriminating ability of a composite irregularity index obtained as a sum of 2 different irregularity indices is less than that of their components. As an example, consider the irregularity index $IRM_{2A}(G)$ obtained from formula (2):

$$
IRM_{2A}(G) = 2IRM_2(G) = 2M_2(G) - \frac{8m^3}{n^2}
$$

and define the corresponding composite irregularity index of the form

$$IRM_{Add}(G) = IRM_{2A}(G) + IRZ(G,1) = F(G) - \frac{8m^3}{n^2} \tag{8}$$

where $IRZ(G,1) = F(G) - 2M_2(G)$ according to equation (7).

It is clear that the values of $IRM_{Add}(G)$ index are equal for all graphs having the same vertex degree sequence. Namely, if graphs G_1 and G_2 are characterized by the same vertex degree sequence, then $IRM_{Add}(G_1) = IRM_{Add}(G_2)$ is fulfilled. This observation implies that for some graphs having an identical degree sequence but different second Zagreb indices, the irregularity indices $IRM_{2A}(G)$ and $IRZ(G,1)$) will be more discriminative than $IRM_{Add}(G)$.

By definition, a graph invariant $IT(G)$ is called a non-negative topological index (NN-index) if $IT(G) \geq 0$ holds for any connected graph G. It is easy to see that the irregularity indices form a particular subset of NN-indices, and the sum of an irregularity index and NN-indices results in an irregularity index as well. As an example, define the topological invariants $F_A(G)$ and $F_B(G)$ as follows:

$$F_A(G) = \rho(G) - \sqrt{\frac{M_1(G)}{n}} \tag{9}$$

$$F_B(G) = \rho(G) - \frac{2M_2(G)}{M_1(G)} \tag{10}$$

Fundamental properties of topological invariants $F_A(G)$ and $F_B(G)$ are characterized by the following lemma.

Lemma 4.1. *The topological invariants $F_A(G)$ and $F_B(G)$ are not irregularity indices. Both belong to the family of NN-indices.*

Proof. It is known that for $F_A(G)$ and $F_B(G)$ the following relations hold:

$F_A(G) \geq 0$ and $F_A(G) = 0$ if and only if G is regular or semiregular graph [39].

$F_B(G) \geq 0$ and $F_B(G) = 0$ if and only if G is a regular graph or a non-regular harmonic graph [40]. $\square$

Define the following composite irregularity indices as follows:

$$IR_{add,a}(G) = CS(G) + F_A(G) = 2\rho(G) - \frac{2m}{n} - \sqrt{\frac{M_1(G)}{n}}$$

$$IR_{add,b}(G) = CS(G) + F_B(G) = 2\rho(G) - \frac{2m}{n} - \frac{2M_2(G)}{M_1(G)}.$$

Consider now, the 7-vertex trees $T(3,2)$ and $T(2,3)$ depicted in Fig. 1 As we have observed, for these trees the Collatz–Sinogowitz irregularity index is identical, namely $CS(T(3,2)) = CS(T(2,3)) = 2 - 12/7 = 0.285714$. However, it is easy to see that for trees $T(3,2)$ and $T(2,3)$ the irregularity indices $IR_{add,a}(G)$ and $IR_{add,b}(G)$ are more selective (discriminative) than $CS(G)$ index. Because $M_1(T(3,2)) = M_2(T(3,2)) = M_2(T(2,3)) = 24$ and $M_1(T(2,3)) = 26$, we get the following results:

$$
\begin{aligned}
IR_{add,a}(T(3,2)) &= 2\rho(T(3,2)) - \frac{2m}{n} - \sqrt{\frac{M_1(T(3,2))}{n}} \\[1em]
&= 4 - \frac{12}{7} - \sqrt{\frac{24}{7}} = 0.4334074
\end{aligned}
$$

$$
\begin{aligned}
IR_{add,a}(T(2,3)) &= 2\rho(T(2,3)) - \frac{2m}{n} - \sqrt{\frac{M_1(T(2,3))}{n}} \\[1em]
&= 4 - \frac{12}{7} - \sqrt{\frac{26}{7}} = 0.358466
\end{aligned}
$$

$$
\begin{aligned}
IR_{add,b}(T(3,2)) &= 2\rho(T(3,2)) - \frac{2m}{n} - \frac{2M_2(T(3,2))}{M_1(T(3,2))} \\[1em]
&= CS(T(3,2)) = CS(T(2,3)) = 0.285714
\end{aligned}
$$

$$
\begin{aligned}
IR_{add,b}(T(2,3)) &= 2\rho(T(2,3)) - \frac{2m}{n} - \frac{2M_2(T(2,3))}{M_1(T(2,3))} \\[1em]
&= 4 - \frac{12}{7} - \frac{2 \cdot 24}{26} = 0.439560.
\end{aligned}
$$

It is worth noting that for tree $T(3,2)$ the corresponding irregularity index is $IR_{add,b}(T(3,2) = CS(T(3,2) = CS(T(2,3)$. The reason for this can be explained by the fact that $T(3,2)$ is a harmonic graph, for which the equality $\rho(T(3,2)) = 2M_2(T(3,2))/M_1(T(3,2)) = 2$ holds. As can be seen, when ranking the graph irregularity by indices $IR_{add,a}(G)$ and $IR_{add,b}(G)$ the results obtained are not consistent,

because the inequalities $IR_{add,a}(T(3,2)) > IR_{add,a}(T(2,3))$ and $IR_{add,b}(T(3,2)) < IR_{add,b}(T(2,3))$ hold.

An interesting observation is that a composite topological invariant can be a graph irregularity index in some particular cases, when their components are not irregularity indices. This phenomenon is demonstrated in the following example. Consider the topological invariants $F_A(G)$ and $F_B(G)$ represented by formulas (9) and (10). Previously, it has been verified that $F_A(G)$ and $F_B(G)$ are non-negative numbers, and they do not belong to the family of irregularity indices.

Lemma 4.2. *In the family of connected non-regular graphs there are no graphs which are semiregular and harmonic simultaneously.*

Proof. Let $G(\Delta, \delta)$ a semiregular graph. For any edge (i, j) of a bidegreed semiregular graph the equality $d_i = d(u_i) = \Delta$ and $d_j = d(u_j) = \delta$ holds. Denote the degree sum of vertices adjacent to a vertex u_i by $Q(u_i)$. For the corresponding average neighbor degree numbers we obtain the following identities: $\mu(u_i) = Q(u_i)/d(u_i) = \Delta\delta/\Delta = \delta$ and $\mu(u_j) = Q(u_j)/d(u_j) = \Delta\delta/\delta = \Delta$. Because $\mu(u_i) \neq \mu(u_j)$ holds for every edge (i, j), this implies that graph $G(\Delta, \delta)$ is not harmonic. From this it follows that if G is a harmonic graph, then G is not semiregular. $\square$

Proposition 4.3. *Let G be a connected graph, and consider the topological invariant $IR_{AB}(G)$ formulated as*

$$IR_{AB}(G) = F_A(G) + F_B(G) = 2\rho(G) - \sqrt{\frac{M_1(G)}{n}} - \frac{2M_2(G)}{M_1(G)}.$$

Then $IR_{AB}(G)$ is a graph irregularity index.

Proof. From Lemma 4.1 and Lemma 4.2 it follows that IR_{AB} is a non-negative number and $IR_{AB}(G) = 0$ if and only if G is a regular graph. Consequently, $IR_{AB}(G)$ is an irregularity index. $\square$

Remark 4.4. *Based on the concept outlined previously, it is possible to generate composite irregularity indices of multiplicative type as well. To do this, construct the composite graph invariant $IM(G)$ in the following form:*

$$IM(G) = (1 + F_A(G))(1 + F_B(G)) - 1\,.$$

It is easy to see that $IM(G) \geq 0$ and $IM(G) = 0$ if and only if G is a regular graph. Consequently, $IM(G)$ is an irregularity index.

5   Decomposition of irregularity indices

There are several ways to obtain an additive decomposition of a given irregularity index $IR(G)$. The decomposition of $IR(G)$ into K components can be formulated as

$$IR(G) = \sum_{k=1}^{K} p_k IR_k(G)$$

where $IR_k(G)$ represent a finite set of various irregularity indices and p_k are positive numbers. As an example, consider the irregularity index $IRM_{Add}(G)$ given by Eq.(8). Irregularity index $IRM_{Add}(G)$ can be decomposed as

$$IRM_{Add}(G) = (F(G) - 2M_2(G)) + \left(2M_2(G) - \frac{8m^3}{n^2}\right) = IRZ(G,1) + 2IRM_2(G)\,.$$

It is possible to perform a decomposition of an irregularity index into 2 graph invariants in such a way that only one component will be an irregularity index. Such decompositions of the Collatz–Sinogowitz irregularity index will be demonstrated in 3 examples.

Example 5.1. *Consider the topological invariant $F_A(G)$ and the irregularity index $IR_1(G)$ defined by Eq.(5). Using $F_A(G)$ and $IR_1(G)$ one obtains the following decomposition of $CS(G)$:*

$$CS(G) = \rho(G) - \frac{2m}{n} = \left(\rho(G) - \sqrt{\frac{M_1(G)}{n}}\right) + \left(\sqrt{\frac{M_1(G)}{n}} - \frac{2m}{n}\right) = F_A(G) + IR_1(G)\,.$$

It should be emphasized that $F_A(G)$ is an NN-index, because $F_A(G) = 0$ holds for regular and semiregular graphs.

Example 5.2. *As another similar example, consider the topological invariant*

$$\Phi(G) = \rho(G) - \sqrt{\frac{M_2(G)}{m}}$$

and the irregularity index $IR_2(G)$ defined by Eq.(6). It is important to note that $\Phi(G)$ can be a positive or a negative number, and $\Phi(G) = 0$ holds for any Z_2 graphs. This implies that $\Phi(G)$ is not an irregularity index. Using the above formula we get the following decomposition of the Collatz–Sinogowitz index:

$$CS(G) = \left(\rho(G) - \sqrt{\frac{M_2(G)}{m}}\right) + \left(\sqrt{\frac{M_2(G)}{m}} - \frac{2m}{n}\right) = \Phi(G) + IR_2(G).$$

Example 5.3. *An interesting decomposition of $CS(G)$ index can be obtained if we use the reciprocal Randić index $RR(G)$ for this purpose [23]. It is verified [16] that for an m-edge connected graph*

$$\rho(G) \geq \frac{1}{m} \sum_{(i,j)\in E} \sqrt{d_i d_j} = \frac{1}{m} RR(G)$$

where for example, equality holds for regular or semi-regular graphs. From the above inequality we can construct the topological invariant $\Psi(G)$ defined as

$$\Psi(G) = \rho(G) - \frac{1}{m} \sum_{(i,j)\in E} \sqrt{d_i d_j} = \rho(G) - \frac{RR(G)}{m}.$$

Moreover, it is known [23] that

$$\frac{RR(G)}{m} = \frac{1}{m} \sum_{(i,j)\in E} \sqrt{d_i dj} \geq \frac{2m}{n}$$

with equality if and only if G is regular. From the above inequality we can generate the irregularity index $IRS(G)$ formulated as

$$IRS(G) = \frac{RR(G)}{m} - \frac{2m}{n}.$$

Based on the previous considerations, the $CS(G)$ index can be decomposed into the sum of $\Psi(G)$ and $IRS(G)$ as follows

$$CS(G) = \left(\rho(G) - \frac{RR(G)}{m}\right) + \left(\frac{RR(G)}{m} - \frac{2m}{n}\right) = \Psi(G) + IRS(G).$$

In certain cases, the irregularity index $IRS(G)$ has a higher discriminatory performance than $CS(G)$ index. This observation is demonstrated in the following example. We have shown for trees $T(3, 2)$ and $T(2, 3)$ depicted in Fig. 1 that their $CS(G)$ indices are identical, namely $CS(T(3, 2)) = CS(T(2, 3)) = 2 - 12/7 = 0.2857143$. For these trees we have the following reciprocal Randić indices:

$$RR(T(3,2)) = \sum_{(i,j)\in E} \sqrt{d_i d_j} = 3\sqrt{2} + 3\sqrt{6} = 11.59111$$

$$RR(T(2,3)) = \sum_{(i,j)\in E} \sqrt{d_i d_j} = 4\sqrt{3} + 2\sqrt{6} = 11.82718 \,.$$

Consequently, for the corresponding irregularity indices one obtains

$$IRS(T(3,2)) = \frac{RR(T(3,2))}{m} - \frac{2m}{n} = \frac{1}{6}\left(3\sqrt{2} + 3\sqrt{6}\right) - \frac{12}{7} = 0.217566$$

and

$$IRS(T(2,3)) = \frac{RR(T(2,3))}{m} - \frac{2m}{n} = \frac{1}{6}\left(4\sqrt{3} + 2\sqrt{6}\right) - \frac{12}{7} = 0.256911 \,.$$

Comparing the computed values of $IRS(G)$ indices, the tree $T(3, 2)$ is judged to be more regular than $T(2, 3)$.

6 Problems of degeneracy

A well-known problem when performing topological graph characterizations is that there exist non-isomorphic graphs having the same value of topological indices. This phenomenon is called degeneracy [9, 10, 13, 30]. A graph invariant is referred to be degenerate if it has the same value for more than one graph [9].

Supposing that a set of graphs is given, the fundamental task to be solved is to select a highly discriminating irregularity index for ranking. If the irregularity index is not compatible with the preselected graph family, then the irregularity ordering of graphs by this index becomes impossible. The degeneracy problem is clearly demonstrated by the following example related to the Collatz–Sinogowitz irregularity index.

Let q be an arbitrary positive integer. It is easy to show that there exists an infinite family $\Omega(q)$ of connected graphs for which $CS(G) = q$ holds for any graph G

included in $\Omega(q)$. Such infinite graph families can be constructed as follows: Let C_r be a cycle of length r, where $r \geq 3$. Denote by C_r^k the graph obtained from C_r by attaching $k \geq 1$ pendent vertices to every vertex of C_r. Because C_r^k is a 2-walk $(2, k)$ linear graph [28, 36] for its spectral radius one obtains

$$\rho(C_r^k) = \frac{1}{2}\left(2 + \sqrt{4 + 4k}\right) = 1 + \sqrt{1 + k}\,.$$

If $(1 + k)$ is a perfect square then $\rho(C_r^k) = 1 + \sqrt{1 + k}$ will be a positive integer. The above identity is valid for any unicyclic graph C_r^3 with arbitrary $r \geq 3$. The smallest graph of this type is the graph C_3^3 depicted in Fig. 2.
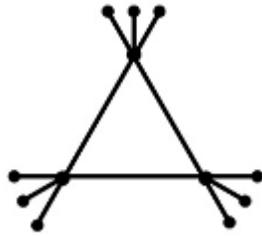

Figure 2: Unicyclic C_3^3 graph with a spectral radius 3.

Because graphs C_r^k are unicyclic, then $2m/n = 2$ holds for them. This implies that if $k = (q + 1)^2 - 1$, where q is a positive integer, then for the corresponding Collatz–Sinogowitz index we obtain

$$CS\left(C_r^{(q+1)^2-1}\right) = \rho\left(C_r^{(q+1)^2-1}\right) - \frac{2m}{n} = 1 + \sqrt{1 + (q + 1)^2 - 1} - 2 = q\,.$$

Remark 6.1. *For C_r^k graphs the vertex number is determined by parameters r and k, because $n = r(1 + k)$ holds. Moreover, the diameter of these graphs can be arbitrary large and depends only on parameter r. This implies that $CS(C_r^k)$ is independent of the graph diameters. Our investigations confirm the general observation formulated in [10]: A topological index can be unique only for a particular graph class, but it fails when it is applied to another class.*

In certain cases, the degeneracy can be avoided or efficiently decreased if instead of degree-based irregularity indices we use so-called combined irregularity indices which are constructed by the generalization of some traditional irregularity indices. This

concept is demonstrated in an example concerning the construction of two extended versions of the Albertson irregularity index.

For an edge $(i, j) = (u_i, u_j)$ of graph G, define the positive integers N_i and N_j where N_i is the number of vertices of G whose distance to vertex u_i is smaller than distance to vertex u_j, and analogously, N_j is the number of vertices of G whose distance to the vertex u_j is smaller than to u_i. Based on the previous considerations, an extended distance-based version of the Albertson irregularity index can be constructed. This extended (combined) irregularity index $IRA_f(G)$ is defined in the following general form:

$$IRA_f(G) = \sum_{(i,j)} |d_i - d_j| f(N_i, N_j).$$

In the above formula $f(x, y)$ is a non-negative symmetric function for which $f(x, y) = f(y, x) \geq 1$ holds and $f(x, y) = 1$ if and only if $x = y$. As an example, depending on the particular choice of function $f(x, y)$, two distance-based irregularity indices denoted by $IRAS(G)$ and $IRAQ(G)$ are constructed:

If $f(x, y) = (x^2 + y^2)/(2xy)$, then

$$IRAS(G) = \frac{1}{2} \sum_{(i,j)} |d_i - d_j| \left(\frac{N_i^2 + N_j^2}{N_i N_j} \right) \geq \sum_{(i,j)} |d_i - d_j| = AL(G), \qquad (11)$$

If $f(x, y) = 2(x^2 + y^2)/(x + y)^2$ then

$$IRAQ(G) = 2 \sum_{(i,j)} |d_i - d_j| \left(\frac{N_i^2 + N_j^2}{(N_i + N_j)^2} \right) \geq \sum_{(i,j)} |d_i - d_j| = AL(G). \qquad (12)$$

It is easy to see that in the above formulas equality holds, if for every edge (i, j) having different end-vertex degrees, $N_i = N_j$ is fulfilled.From this fact we can conclude that both irregularity indices can be considered as extended versions of the traditional degree-based Albertson irregularity index $AL(G)$.

7 Paradox of graph irregularity characterization

According to a general opinion, the concept of "graph irregularity" depends strongly on the subjective impression (feeling) about this notion [21]. Comparing two con-

nected graphs, based only on simple visual observations, in certain cases it is easy to judge that a given graph is more irregular than another. This concept forms the basis of the hypothesis that quantifying the notion of graph irregularity, by using an appropriately defined (constructed) irregularity index, the ordering of graphs according to their irregularity can be trustworthily performed. In what follows, we will demonstrate that there exist some particular classes of graphs where the concept mentioned is false, because it leads to serious inconsistency in the interpretations of computed results.
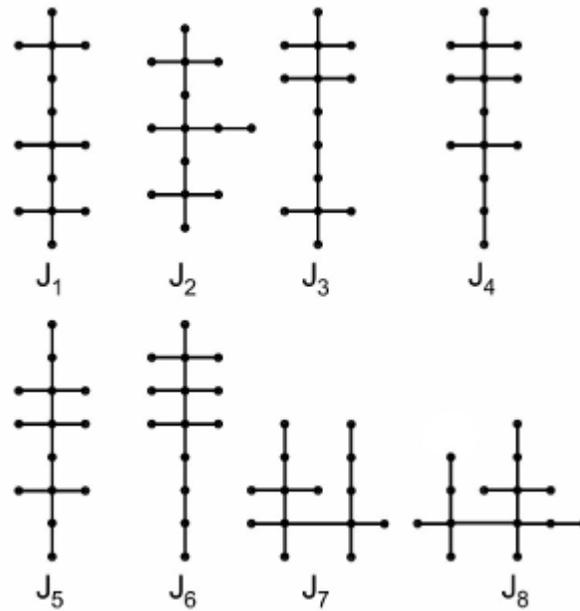

Figure 3: Eight non-isomorphic trees having identical vertex parameters.

In Fig. 3 eight non-isomorphic tree graphs J_i $(i = 1, 2, \ldots, 8)$ characterized by identical degree sequence $(n_1 = 8, n_2 = 3, n_4 = 3)$ are depicted. All of them have strongly similar structures. This observation implies that ranking the irregularity of these trees on the basis of direct visual impressions seems to be problematic, practically it is impossible.

For the structural characterization of graphs depicted in Fig.3 the irregularity indices $IRZ(G, 1)$ and $IRM_2(G)$ were selected and compared, where $IRZ(G, 1) = F(G) - 2M_2(G)$ and $IRM_2(G) = M_2(G) - 4m^3/n^2$, by definition. The correspondence between the computed values of these irregularity indices is illustrated in Fig. 4.

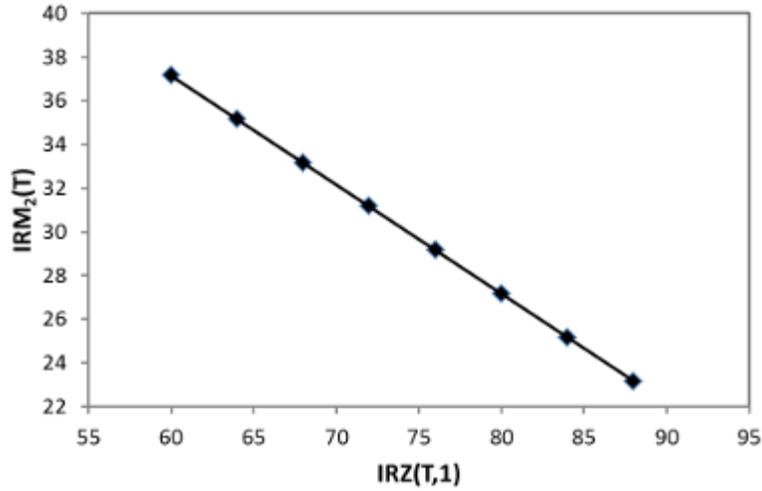As can be observed, for the irregularity classification of 8 trees we have completely

Figure 4: Relation between irregularity indices $IRZ(G,1)$ and $IRM_2(G)$ for 8 trees.

different results with $IRZ(G,1)$ and $IRM_2(G)$ indices. This is due to the fact that between them there is an exact linear relationship formulated as

$$IRM_2(G) = -\frac{IRZ(G,1)}{2} + \frac{1}{2}F(G) - \frac{4m^3}{n^2}.$$

It should be emphasized that the above formula is generally valid for all graphs having an identical degree sequence. Denoting by T an arbitrary 14-vertex tree characterized by the vertex degree sequence $(n_1 = 8, n_2 = 3, n_4 = 3)$, as a particular case, we get

$$IRM_2(T) = -\frac{IRZ(T,1)}{2} + 67.153265.$$

On the basis of these observations it can be concluded that there exist several graphs for which the irregularity ranking performed by $IRM_2(G)$ and $IRZ(G,1)$ cannot be compatible with preliminary subjective perception. Our findings demonstrate clearly that in some cases the subjective concepts concerning the visual-type graph irregularity characterization (ordering) are based on unrealistic, speculative considerations. This strange phenomenon called "the graph irregularity paradox" can be interpreted as follows: There exist some particular classes of connected graphs for which the result of evaluation (ranking) of graph irregularity by using various irregularity indices is determined primarily on their definitions, consequently, in such cases the results of classification depend only to a limited extent on the true topological structure of graphs considered.

8 Conclusions, final remarks

The result of graph irregularity classification using various irregularity indices is determined i) by the type of the selected irregularity index and ii) the specific structural properties of the graph family to be investigated. We have demonstrated that sometimes the application of two different irregularity indices can lead to inconsistent results whose interpretation is problematic.
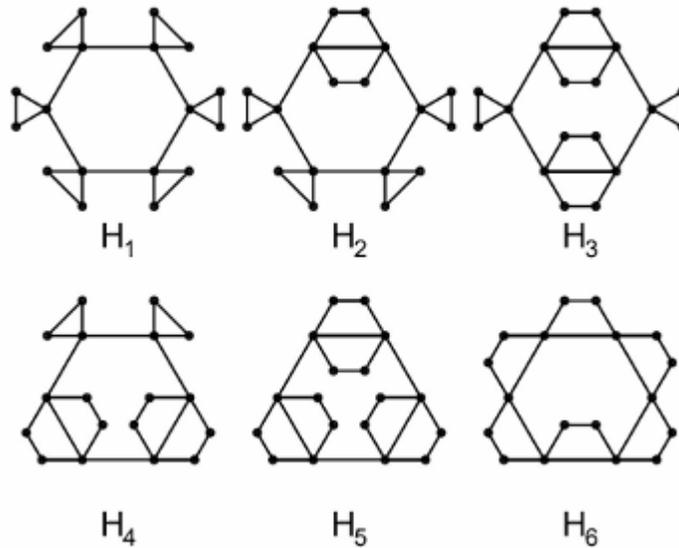

Figure 5: Non-isomorphic harmonic graph having spectral radius 3.

It should be mentioned that there exist non-isomorphic graphs whose irregularity classification (ranking) cannot be performed using traditional degree-based irregularity indices. In such cases, there is no preconception helping us to decide on the correct irregularity ranking of graphs. As an example, consider the 6 non-isomorphic harmonic graphs depicted in Fig. 5.

They are bidegreed graphs with $\Delta = 4$ and $\delta = 2$, additionally, all of them are characterized by the same set of edge parameters $(m_{22} = 6, m_{24} = 12, m_{44} = 6)$. Due to this fact, all degree-based irregularity indices will be equal for them. For example, for the corresponding Albertson irregularity indices we got $AL(H_i) = 24$ for $i = 1, 2, \ldots, 6$.

Moreover, because these graphs are also harmonic, they have identical spectral radius equal to 3. This implies that these graphs cannot be distinguished by the

Collatz–Sinogowitz index. For unicyclic graphs C_r^k we have demonstrated that among these graphs the discrimination is impossible by means of $CS(G)$ index, because there exist infinitely many C_r^k graphs characterized by the same $CS(G)$ value.

However, it seems likely [9,21] that for a given family of structurally strongly similar non-isomorphic graphs it is possible to construct such distance and/or eigenvalue-based irregularity indices by which a better discrimination can be achieved than by using the traditional degree-based irregularity indices. Our investigations have confirmed that the distance-based irregularity indices $IRAS(G)$ and $IRAQ(G)$ have an increased discriminatory performance, consequently both of them can be successfully applied to the structural characterization of graphs depicted in Fig. 5.

Starting with the complex $IRAS(G)$ index for 6 graphs depicted in Fig. 5, we obtained the following results of irregularity ordering:

$$
\begin{aligned}
IRAS(H_1) &= 192,75 > IRAS(H_2) = 157,07 > IRAS(H_3) = IRAS(H_4) \\
&= 129,25 > IRAS(H_5) = IRAS(H_6) = 97,50 \,.
\end{aligned}
$$

A completely similar result of ordering has been obtained by the $IRAQ(G)$ irregularity index:

$$
\begin{aligned}
IRAQ(H_1) &= 42,685 > IRAQ(H_2) = 40,957 > IRAQ(H_3) = IRAQ(H_4) \\
&= 39,907 > IRAQ(H_5) = IRAQ(H_6) = 38,519 \,.
\end{aligned}
$$

As we can conclude, in both cases the bipartite graphs H_5 and H_6 having equal irregularity indices are judged to be the "most regular" graphs.

REFERENCES

[1] H. Abdo, N. Cohen, D. Dimitrov, Graphs with maximal irregularity, *Filomat* **28** (2014) 1315–1322.

[2] H. Abdo, D. Dimitrov, T. Réti, D. Stevanović, Estimation of the spectral radius of graph by the second Zagreb index, *MATCH Commun. Math. Comput. Chem.* **72** (2014) 741–751.

[3] M. O. Albertson, The irregularity of a graph, *Ars Comb.* **46** (1997) 2019–2025.

[4] F. K. Bell, A note on the irregularity of a graph, *Lin. Algebra Appl.* **161** (1992) 45–54.

[5] J. A. Bondy, U. S. R. Murty, *Graph Theory with Applications*, Esevier, New York, 1976.

[6] B. Borovićanin, S. Grünewald, I. Gutman, M. Petrović, Harmonic graphs with a small number of cycles, *Discr. Math.* **265** (2003) 31–44.

[7] G. Chartrand, P. Erdős, O. Oellermann, How to define an irregular graph, *College Math. J.* **19** (1988) 36–42.

[8] L. Collatz, U. Sinogowitz, Spektren endlicher Grafen, *Abh. Math. Sem. Univ. Hamburg* **21** (1957) 63–77.

[9] M. Dehmer, M. Grabner, B. Furtula, Structural discrimination of networks by using distance, degree and eigenvalue–based measures, *PLoS ONE* **7** (2012) e38564.

[10] M. Dehmer, M. Grabner, K. Varmuza, Information indices with high discriminative power for graphs, *PLoS ONE* **7** (2012) e31214.

[11] H. Deng, A unified approach to the extremal Zagreb indices for trees, unicyclic graphs and bicyclic graphs, *MATCH Commun. Math. Comput. Chem.* **57** (2007) 597–616.

[12] D. Dimitrov, T. Réti, Graphs with equal irregularity indices, *Acta Polytech. Hung.* **11** (2014) 41–57.

[13] T. Došlić, T. Réti, Novel degree–based molecular descriptors with increased discriminating power, *Acta Polytech. Hung.* **4** (2012) 17–30.

[14] T. Došlić, T. Réti, D. Vukičević, On the vertex degree indices of connected graphs, *Chem. Phys. Lett.* **512** (2011) 283–286.

[15] E. Estrada, Randić index, irregularity and complex biomolecular networks, *Acta Chim. Slov.* **57** (2010) 597–603.

[16] O. Favaron, M. Mahéo, J. F. Saclé, Some eigenvalue properties in graphs (conjectures of Graffiti-II), *Dicr. Math.* **111** (1993) 197–220.

[17] B. Furtula, I. Gutman, A forgotten topological index, *J. Math. Chem.* **53** (2015) 1184–1190.

[18] B. Furtula, I. Gutman, M. Dehmer, On the structure-sensitivity of degree–based topological indices, *Appl. Math. Comput.* **219** (2013) 8973–8978.

[19] C. Godsil, G. Royle, *Algebraic Graph Theory*, Springer, New York, 2001.

[20] I. Gutman, Degree–based topological indices, *Croat. Chem. Acta,* **86** (2013) 351–361.

[21] I. Gutman, Irregularity of molecular graphs, *Kragujevac J. Sci.* **38** (2016) 99–109.

[22] I. Gutman, K. C. Das, The first Zagreb index 30 years after, *MATCH Commun. Math. Comput. Chem.* **50** (2004) 83–92.

[23] I. Gutman, B. Furtula, C. Elphick, Three new/old vertex–degree–based topological indices, *MATCH Commun. Math. Comput. Chem.* **72** (2014) 617–632.

[24] I. Gutman, P. Hansen, H. Mélot, Variable neighborhood search for extremal graphs. 10. Comparison of irregularity indices for chemical trees, *J. Chem. Inf. Model.* **45** (2005) 222–230.

[25] I. Gutman, J. Tošović, Testing the quality of molecular structure descriptors. Vertex–degree–based topological indices, *J. Serb. Chem. Soc.* **78** (2013) 805–810.

[26] A. Hamzeh, T. Réti, An analogue of Zagreb index inequality obtained from graph irregularity measures, *MATCH Commun. Math. Comput. Chem.* **72** (2014) 669–683.

[27] B. Horoldagva, L. Buyantogtokh, S. Dorjsembe, I. Gutman, Maximum size of maximally irregular graphs, *MATCH Commun. Math. Comput. Chem.* **76** (2016) 81–98.

[28] Y. Hou, F. Tian, Unicyclic graphs with exactly two main eigenvalues, *Appl. Math. Lett.* **19** (2006) 1143–1147.

[29] A. Ilić, D. Stevanović, On comparing Zagreb indices, *MATCH Commun. Math. Comput. Chem.* **62** (2009) 681–687.

[30] E. V. Konstantinova, The discrimination ability of some topological and information distance indices for graphs of unbranched hexagonal systems, *J. Chem. Inf. Comput. Sci.* **36** (1996) 54–57.

[31] B. Liu, Z. You, A survey on comparing Zagreb indices, *MATCH Commun. Math. Comput. Chem.* **65** (2011) 581–589.

[32] M. Lu, H. Liu, F. Tian, The connectivity index, *MATCH Commun. Math. Comput. Chem.* **51** (2004) 149–154.

[33] R. Nasiri, G. H. Fath–Tabar, The second minimum of the irregularity of graphs, *El. Notes Discr. Math.* **45** (2014) 133–140.

[34] S. Nikolić, G. Kovačević, A. Miličević, N. Trinajstić, The Zagreb indices 30 years after, *Croat. Chem. Acta* **76** (2003) 113–124.

[35] M. Randić, The connectivity index 25 years after, *J. Mol. Graphics Model.* **20** (2001) 19–35.

[36]] L. Shi, On graphs with given main eigenvalues, *Appl. Math. Lett.* **22** (2009) 1870–1877.

[37] R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors* WileyVCH, Weinheim, 2009.

[38] N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, 1992.

[39] A. Yu, M. Lu, F. Tian, On the spectral radius of graphs, *Lin. Algebra Appl.* **387** (2004) 41–49.

[40] B. Zhou, Zagreb indices, *MATCH Commun. Math. Comput. Chem.* **52** (2004) 113–118.

[41] B. Zhou, D. Vukičević, On general Randić and general zeroth–order Randić index, *MATCH Commun. Math. Comput. Chem.* **62** (2009) 189–196.