Hanyuan Deng, Slavko Radenković,
and Ivan Gutman

THE ESTRADA INDEX

Abstract. If λ_i, $i = 1, 2, \ldots, n$, are the eigenvalues of the graph G, then the Estrada index EE of G is the sum of the terms e^{λ_i}. This graph invariant appeared for the first time in year 2000, in a paper by Ernesto Estrada, dealing with the folding of protein molecules. Since then a remarkable variety of other chemical and non-chemical applications of EE were communicated.

The mathematical studies of the Estrada index started only a few years ago. Until now a number of lower and upper bounds were obtained, and the problem of extremal EE for trees solved. Also, approximations and correlations for EE were put forward, valid for chemically interesting molecular graphs.

In this paper the relevant results on the Estrada index are surveyed.

Mathematics Subject Classification (2000): 05C50; 05C90; 92E10

Keywords: Estrada index; spectrum (of graph); chemistry

Contents

1. Introduction: the Estrada index and its various applications

Let G be a graph without loops and multiple edges. Let n and m be, respectively, the number of vertices and edges of G. Such a graph will be referred to as an (n, m)-graph.

The eigenvalues of the adjacency matrix of G are said to be [1] the eigenvalues of G and to form the spectrum of G. A graph of order n has n (not necessarily distinct, but necessarily real-valued) eigenvalues; we denote these by $\lambda_1, \lambda_2, \dots, \lambda_n$, and assume to be labelled in a non-increasing manner: $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. The basic properties of graph eigenvalues can be found in the book [1].

A graph-spectrum-based invariant, recently put forward by Estrada is defined as

$$EE = EE(G) = \sum_{i=1}^{n} e^{\lambda_i}. \tag{1}$$

We proposed [2] to call it the *Estrada index*, a name that in the meantime has been commonly accepted.

Although invented in year 2000 [3], the Estrada index has already found a remarkable variety of applications. Initially it was used to quantify the degree of folding of long-chain molecules, especially proteins [3, 4, 5]; for this purpose the EE-values of pertinently constructed weighted graphs were employed. Another, fully unrelated, application of EE (this time of simple graphs, like those studied in the present paper) was put forward by Estrada and Rodríguez-Velázquez [6, 7]. They showed that EE provides a measure of the centrality of complex (communication, social, metabolic, etc) networks; these ideas were recently further elaborated and extended [8]. In [9] a connection between EE and the concept of extended atomic branching was established, which was an attempt to apply EE in quantum chemistry. Another such application, this time in statistical thermodynamics, was proposed by Estrada and Hatano [10] and later further extended in [11]. Recently, Carbó-Dorca [12] endeavored to find connections between EE and the Shannon entropy.

The proposed biochemical [3, 4, 5], physico–chemical [9, 10], network–theoretical [6, 7, 8], and infomation–theoretical [12] applications of the Estrada index are nowadays widely accepted and used by other members of the scientific community; see, for example [13–20]. In addition, this graph invariant is worth attention of mathematicians. Indeed, in the last few years quite a few mathematicians became interested in the Estrada index and communicated mathematical results on EE in mathematical journals. In what follows we briefly survey the most significant of these results.

2. Elementary properties of the Estrada index

Directly from the definition of the Estrada index, Eq. (1) we conclude the following [7, 21].

1° Denoting by $M_k = M_k(G) = \sum_{i=1}^{n} (\lambda_i)^k$ the k-th spectral moment of the graph G, and bearing in mind the power–series expansion of e^x, we have

$$EE(G) = \sum_{k=0}^{\infty} \frac{M_k(G)}{k!}. \tag{2}$$

At this point one should recall [1] that $M_k(G)$ is equal to the number of self-returning walks of length k of the graph G. The first few spectral moments of an (n, m)-graph satisfy the following relations [1]:

$$M_0 = n; \quad M_1 = 0; \quad M_2 = 2m; \quad M_3 = 6t$$

where t is the number of triangles.

2° As a direct consequence of (2), for any graph G of order n, different from the complete graph K_n and from its (edgeless) complement $\bar{K}_n$,

$$EE(\bar{K}_n) < EE(G) < EE(K_n).$$

3° If G is a graph on n vertices, then $EE(G) \geqslant n$; equality holds if and only if $G \cong \bar{K}_n$ [22].

4° The eigenvalues of a bipartite graph satisfy the pairing property [1]: $\lambda_{n-i+1} = -\lambda_i$, $i = 1, 2, \ldots, n$. Therefore, if the graph G is bipartite, and if n_0 is nullity (= the multiplicity of its eigenvalue zero), then

$$EE(G) = n_0 + 2 \sum_{+} \cosh(\lambda_i) \tag{3}$$

where cosh stands for the hyperbolic cosine [$\cosh(x) = (e^x + e^{-x})/2$], whereas $\sum_+$ denotes summation over all positive eigenvalues of the corresponding graph.

5° If $\mathbf{A}(G)$ is the adjacency matrix of the graph G, then $EE(G) = \operatorname{tr} e^{\mathbf{A}(G)}$, with tr standing for the trace of the respective matrix.

3. Bounds for the Estrada index

Numerous lower and upper bounds for the Estrada index have been communicated. In what follows we first state the simplest and earliest such bounds (as Theorem 3.1), and provide them with complete proofs. The other bounds will only be stated, and their proofs skipped.

Theorem 3.1. [2] *Let G be an (n,m)-graph. Then the Estrada index of G is bounded as*

$$\sqrt{n^2+4m} \leqslant EE(G) \leqslant n-1+e^{\sqrt{2m}}. \tag{4}$$

Equality on both sides of (4) *is attained if and only if $G \cong \bar{K}_n$.*

Proof of the lower bound (4). From the definition of the Estrada index, Eq. (1), we get

$$EE^2 = \sum_{i=1}^{n} e^{2\lambda_i} + 2\sum_{i<j} e^{\lambda_i}\, e^{\lambda_j}. \tag{5}$$

In view of the inequality between the arithmetic and geometric means,

$$\begin{aligned} 2\sum_{i<j} e^{\lambda_i}\, e^{\lambda_j} &\geqslant n(n-1)\left(\prod_{i<j} e^{\lambda_i}\, e^{\lambda_j}\right)^{2/[n(n-1)]} \\ &= n(n-1)\left[\left(\prod_{i=1}^{n} e^{\lambda_i}\right)^{n-1}\right]^{2/[n(n-1)]} \\ &= n(n-1)\left(e^{M_1}\right)^{2/n} = n(n-1). \end{aligned} \tag{6}$$

By means of a power–series expansion, and bearing in mind the properties of M_0, M_1, and M_2, we get

$$\sum_{i=1}^{n} e^{2\lambda_i} = \sum_{i=1}^{n}\sum_{k\geqslant 0} \frac{(2\lambda_i)^k}{k!} = n+4m+\sum_{i=1}^{n}\sum_{k\geqslant 3}\frac{(2\lambda_i)^k}{k!}. \tag{7}$$

Because we are aiming at an (as good as possible) lower bound, it may look plausible to replace $\sum_{k\geqslant 3}\frac{(2\lambda_i)^k}{k!}$ by $8\sum_{k\geqslant 3}\frac{(\lambda_i)^k}{k!}$. However, instead of $8=2^3$ we shall use a multiplier $\gamma \in [0,8]$, so as to arrive at:

$$\begin{aligned} \sum_{i=1}^{n} e^{2\lambda_i} &\geqslant n+4m+\gamma\sum_{i=1}^{n}\sum_{k\geqslant 3}\frac{(\lambda_i)^k}{k!} \\ &= n+4m-\gamma n-\gamma m+\gamma\sum_{i=1}^{n}\sum_{k\geqslant 0}\frac{(\lambda_i)^k}{k!} \end{aligned}$$

i.e.,

$$\sum_{i=1}^{n} e^{2\lambda_i} \geqslant (1-\gamma)n+(4-\gamma)m+\gamma EE. \tag{8}$$

By substituting (6) and (8) back into (5), and solving for EE we obtain

$$EE \geqslant \frac{\gamma}{2} + \sqrt{\left(n - \frac{\gamma}{2}\right)^2 + (4-\gamma)m}. \tag{9}$$

It is elementary to show that for $n \geqslant 2$ and $m \geqslant 1$ the function

$$f(x) := \frac{x}{2} + \sqrt{\left(n - \frac{x}{2}\right)^2 + (4-x)m}$$

monotonically decreases in the interval $[0, 8]$. Consequently, the best lower bound for EE is attained not for $\gamma = 8$, but for $\gamma = 0$.

Setting $\gamma = 0$ into (9) we arrive at the first half of Theorem 3.1.

Remark. If in Eq. (7) we utilize also the properties of the third spectral moment, we get

$$\sum_{i=1}^{n} e^{2\lambda_i} = n + 4m + 8t + \sum_{i=1}^{n} \sum_{k \geqslant 4} \frac{(2\lambda_i)^k}{k!}$$

which, in a fully analogous manner, results in

$$EE \geqslant \sqrt{n^2 + 4m + 8t}. \tag{10}$$

Proof of the upper bound (4). Starting with Eq. (2) we get

$$\begin{aligned}
EE &= n + \sum_{i=1}^{n} \sum_{k \geqslant 1} \frac{(\lambda_i)^k}{k!} \leqslant n + \sum_{i=1}^{n} \sum_{k \geqslant 1} \frac{|\lambda_i|^k}{k!} \\
&= n + \sum_{k \geqslant 1} \frac{1}{k!} \sum_{i=1}^{n} \left[(\lambda_i)^2\right]^{k/2} \leqslant n + \sum_{k \geqslant 1} \frac{1}{k!} \left[\sum_{i=1}^{n} (\lambda_i)^2\right]^{k/2} \\
&= n + \sum_{k \geqslant 1} \frac{1}{k!} (2m)^{k/2} = n - 1 + \sum_{k \geqslant 0} \frac{(\sqrt{2m})^k}{k!}
\end{aligned}$$

which directly leads to the right–hand side inequality in (4).

From the derivation of (4) it is evident that equality will be attained if and only if the graph G has no non-zero eigenvalues. This, in turn, happens only in the case of the edgeless graph $\bar{K}_n$ [1].

By this the proof of Theorem 3.1 is completed. □

Recently Zhou [23] arrived at the following generalizations of Theorem 3.1:

Theorem 3.2. [23] *If G is a graph on n vertices and k_0 is an integer, $k_0 \geqslant 2$, then*

$$EE(G) \geqslant \sqrt{n^2 + \sum_{k=2}^{k_0} \frac{2^k\, M_k(G)}{k!}} \tag{11}$$

with equality if and only if $G \cong \bar{K}_n$.

For $k_0 = 2$ and $k_0 = 3$, the right–hand side of (11) reduces to the lower bounds (4) and (10), respectively.

Theorem 3.3. [23] *Let G be an (n, m)-graph and k_0 same as in Theorem 3.2. Then*

$$EE(G) \leqslant n - 1 - \sqrt{2m} + e^{\sqrt{2m}} + \sum_{k=2}^{k_0} \frac{M_k - \left(\sqrt{2m}\right)^k}{k!}$$

with equality if and only if $G \cong \bar{K}_n$.

Note that for $k_0 = 2$, Theorem 3.3 yields $EE \leqslant n - 1 + e^{\sqrt{2m}} - \sqrt{2mm}$, which is better than the right-hand side of (4).

If graph parameters other than n and m are included into consideration, then further bounds for the Estrada index could be deduced.

Theorem 3.4. [23] *Let G be a graph on n vertices, and d_i, $i = 1, 2, \dots, n$, the degrees of its vertices. Let $D = \sum_{i=1}^{n} (d_i)^2$. Then*

$$EE(G) \geqslant e^{\sqrt{D/n}} + (n - 1)\, e^{-\frac{1}{n-1}\sqrt{D/n}}$$

with equality if and only if either $G \cong K_n$ or $G \cong \bar{K}_n$.

Theorem 3.5. [23] *Let λ_1 be the greatest eigenvalue of an (n, m)-graph G. Let k_0 be the same as in Theorems 3.2 and 3.3. Then*

$$EE(G) \leqslant n - 2 - \lambda_1 - \sqrt{2m - (\lambda_1)^2} + e^{\sqrt{2m-(\lambda_1)^2}} + \sum_{k=2}^{k_0} \frac{M_k - (\lambda_1)^k - \left(\sqrt{2m - (\lambda_1)^2}\right)^k}{k!}$$

with equality if and only if $G \cong \bar{K}_n$.

The special cases of Theorem 3.5 for $k_0 = 2$ and $k_0 = 3$ read:

$$EE \leqslant n - 2 - \lambda_1 - \sqrt{2m - (\lambda_1)^2} + e^{\lambda_1} + e^{\sqrt{2m-(\lambda_1)^2}} \quad \text{and}$$

$$EE \leqslant n - 2 - \lambda_1 - \sqrt{2m - (\lambda_1)^2} + e^{\lambda_1} + e^{\sqrt{2m-(\lambda_1)^2}} + t - \frac{(\lambda_1)^3}{6} - \frac{\left(\sqrt{2m - (\lambda_1)^2}\right)^3}{6}$$

respectively.

Theorem 3.6. [24] *If G is an (n, m)-graph either without isolated vertices or having the property $2m/n \geqslant 1$, then $EE(G) \geqslant n \cosh\left(\sqrt{2m/n}\right)$ with equality if and only if G is a regular graph of degree 1.*

Recall that $2m/n$ is equal to the average vertex degree. Thus, if G is connected, then necessarily $2m/n \geqslant 1$, and the 2-vertex complete graph (K_2) is the only graph for which equality holds.

Theorem 3.7. [24] *If G is an (n,m)-graph, such that $2m/n < 1$, then*

$$EE(G) \geqslant n - 2m + 2m\cosh(1).$$

Equality holds if and only if G consists of $n-2m$ isolated vertices and m copies of K_2.

Theorem 3.8. [24, 25] *If G is an (n,m)-graph with at least one edge, and if n_0 is its nullity, then*

$$EE(G) \geqslant n_0 + (n-n_0)\cosh\left(\sqrt{\frac{2m}{n-n_0}}\right).$$

Equality holds if and only if $n-n_0$ is even, and if G consists of copies of complete bipartite graphs K_{r_i,s_i}, $i = 1,\dots,(n-n_0)/2$, such that all products $r_i \cdot s_i$ are mutually equal.

Theorem 3.8 should be compared with inequality (3). It was first proven for bipartite graphs [25] and eventually extended to all graphs. The same result was later communicated also in [23].

If the graph G is regular of degree r, then its greatest eigenvalue is equal to r. If, in addition, G is bipartite, then its smallest eigenvalue is equal to $-r$ [1]. Bearing these facts in mind, some of the above bounds could have been simplified [2]:

Theorem 3.9. [2] *Let G be a regular graph of degree r and of order n. Then*

$$e^r + \sqrt{n + 2nr - (2r^2+2r+1) + (n-1)(n-2)\,e^{-2r/(n-1)}} \leqslant EE(G) \leqslant n - 2 + e^r + e^{\sqrt{r(n-r)}}.$$

The lower bound is improved by including into the consideration also the third spectral moment:

$$EE(G) \geqslant e^r + \sqrt{n + 2nr - (2r^2+2r+1) + (n-1)(n-2)\,e^{-2r/(n-1)} - \tfrac{4}{3}\,(r^3-6t)}.$$

Theorem 3.10. [2] *Let G be a bipartite regular graph of degree r and of order n. Then*

$$2\cosh(r) + \sqrt{(n-2)^2 + 2nr - 4\,r^2} \leqslant EE(G) \leqslant n - 4 + 2\cosh(r) + 2\cosh\left(\sqrt{nr/2 - r^2}\right).$$

4. Estrada indices of some graphs

For graphs whose spectra are known [1], by Eq. (1) one gets explicit expressions for their Estrada index. In particular:

$$EE(K_n) = e^{n-1} + (n-1)\,e^{-1}$$

$$EE(K_{a,n}) = a + b - 2 + 2\cosh\left(\sqrt{ab}\right).$$

If S_n is the n-vertex star, then $EE(S_n) = n - 2 + 2\cosh\left(\sqrt{n-1}\right)$. If Q_n is the hypercube on 2^n vertices, then $EE(Q_n) = [2\cosh(1)]^n$ [22].

The $(n+1)$-vertex wheel W_{n+1} is obtained by joining a new vertex to each vertex of the n-vertex cycle C_n. Then $EE(W_{n+1}) = EE(C_n) - e^2 + 2e\cosh\left(\sqrt{n-1}\right)$ [22].

The Estrada index of the cycle C_n can be approximated as $EE(C_n) \approx n\, I_0$, [26] where

$$I_0 = \frac{1}{\pi}\int_0^\pi e^{2\cos x}\, dx = \sum_{k=0}^\infty \frac{1}{(k!)^2} = 2.27958530\cdots.$$

In an analogous manner [26, 27]

$$EE(P_n) \approx (n+1)\, I_0 - \cosh(2)$$
$$EE(Z_n) \approx (n+2)\, I_0$$
$$EE(ZZ_n) \approx (n+1)\, I_0 + 2 + \cosh(2)$$

where P_n is the n-vertex path, Z_n is the $(n+2)$-vertex tree obtained by attaching two pendent vertices to a terminal vertex of P_n, whereas ZZ_n is the $(n+4)$-vertex tree obtained by attaching two pendent vertices to each of the two terminal vertices of P_n.

For positive integers n and m, the tree $P_{n,m}$ on $(m+1)n$ vertices is obtained by attaching m pendent vertices to each vertex of P_n. Then [27]

$$EE(P_{n,m}) \approx (m-1)n + 2(n+1)\, J_m$$

where

$$J_m = \frac{1}{\pi}\int_0^\pi e^{\cos x}\, \cosh\left(\sqrt{m+\cos^2 x}\,\right) dx.$$

Approximations for the Estrada index of Bethe and double-Bethe trees were reported in [28]. Expressions and approximate expressions for EE of several other graphs can be found in [22].

In [29] the following approximate expression for the Estrada index of an (n,m)-graph was deduced using a Monte Carlo technique:

$$n\left(\sqrt{6m/n}\right)^{-1} \sinh\left(\sqrt{6m/n}\right)$$

where sinh stands for the hyperbolic sine [$\sinh(x) = (e^x - e^{-x})/2$]. In [29] also some more complicated approximations for EE of (n,m)-graphs were proposed.

4.1. Estrada index of line graphs.

Theorem 4.1. [30] *If G is an r-regular graph with n vertices and $m = rn/2$ edges, and $L(G)$ is its line graph, then $EE(L(G)) = e^{r-2}EE(G) + (m-n)e^{-2}$.*

By Theorem 4.1, if G is a connected r-regular graph, then $EE(L(G)) = EE(G)$ holds if and only if $r = 1, 2$, i.e., if and only if either $G \cong K_2$ or $G \cong C_n$ [22]. To see this, suppose that $EE(L(G)) = EE(G)$ and $r \geqslant 3$. Then $m > n$ and $EE(G) = (n-m)/[e^2\,(e^{r-2}-1)]$. This would imply that $EE(G) < 0$, a contradiction.

The k-th iterated line graph $L^k(G)$ of a graph G is defined recursively by $L^k(G) = L(L^{k-1}(G))$ where $L^0(G) \equiv G$ and $L^1(G) \equiv L(G)$.

Theorem 4.2. [30] *If G is an r-regular graph with n vertices, and $k \geqslant 1$, then*

$$EE(L^k(G)) = a_k(r)\, EE(G) + b_k(r)\, n$$

where $a_k(r)$ and $b_k(r)$ are functions depending solely on the variable r and parameter k.

In [30] it was shown that $a_k(r) = e^{(r-2)(2^k-1)}$, which implies $a_k(r) = O\big(e^{(r-2)2^k}\big)$. An explicit expression for $b_k(r)$ could not be determined, but it was established [30] that $b_k(r)$ has the same asymptotic behavior as $a_k(r)$, viz., $b_k(r) = O\big(e^{(r-2)2^k}\big)$.

4.2. Estrada index of some graph products. Let G and H be two graphs with disjoint vertex sets. The join $G+H$ of G and H is the graph obtained by connecting all vertices of G with all vertices of H. If $G_1, G_2, \dots, G_n$ are graphs with mutually disjoint vertex sets, then we denote $G_1 + G_2 + \cdots + G_n$ by $\sum_{i=1}^n G_i$. In the case that $G_1 = G_2 = \cdots = G_n = G$, we denote $\sum_{i=1}^n G_i$ by $n\,G$.

Theorem 4.3. [22] *Let G and H be r- and s-regular graphs with p and q vertices, respectively. Then*

$$EE(G+H) = EE(G) + EE(H) - (e^r + e^s) + 2e^{(r+s)/2} \cosh\left(\tfrac{1}{2}\sqrt{(r-s)^2 + 4pq}\right).$$

Corollary 4.4. [22] *If G is an r-regular n-vertex graph then*

$$EE(2G) = 2EE(G) - 2e^r + 2e^r \cosh(n)$$
$$EE(3G) = 3EE(G) - 3e^r + 2e^r \cosh(n) + 2e^{(2r+n)/2} \cosh(3n/2) - e^{r+n}.$$

The Cartesian product $G \times H$ of graphs G and H has the vertex set $V(G \times H) = V(G) \times V(H)$ and $(a,x)(b,y)$ is an edge of $G \times H$ if $a = b$ and $xy \in E(H)$, or $ab \in E(G)$ and $x = y$. If $G_1, G_2, \dots, G_n$ are graphs with mutually disjoint vertex sets, then we denote $G_1 \times G_2 \times \cdots \times G_n$ by $\prod_{i=1}^n G_i$. In the case that $G_1 = G_2 = \cdots = G_n = G$, we denote $\prod_{i=1}^n G_i$ by G^n.

Theorem 4.5. [22] *$EE(G \times H) = EE(G)\, EE(H)$. More generally,*

$$EE\left(\prod_{i=1}^r G_i\right) = \prod_{i=1}^r EE(G_i).$$

In particular, $EE(G^r) = EE(G)^r$.

5. Graphs with extremal Estrada indices

In [2] de la Peña, Gutman and Rada put forward two conjectures:

Conjecture A. Among n-vertex trees, the path P_n has the minimum and the star S_n the maximum Estrada index, i.e., $EE(P_n) < EE(T_n) < EE(S_n)$, where T_n is any n-vertex tree different from S_n and P_n.

Conjecture B. Among connected graphs of order n, the path P_n has the minimum Estrada index.

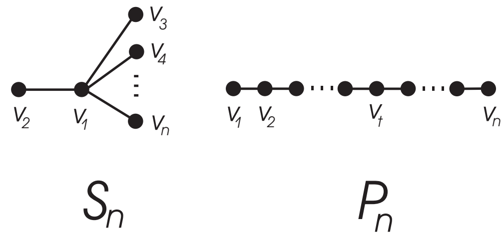

FIGURE 1. The star S_n and the path P_n, and the labelling of their vertices.

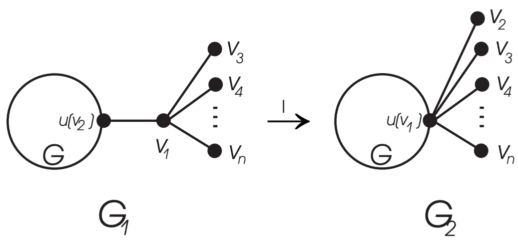

FIGURE 2. Transformation I.

In what follows we first state some transformations of graphs and establish the respective change in the spectral moments, and then provide a complete proof of these conjectures.

Lemma 5.1. [31] *Let S_n be the n-vertex star with vertices $v_1, v_2, \ldots, v_n$, and center v_1, as shown in Figure 1. Then there is an injection ξ_1 from $W_{2k}(v_2)$ to $W_{2k}(v_1)$, and ξ_1 is not surjective for $n \geqslant 3$ and $k \geqslant 1$, where $W_{2k}(v_1)$ and $W_{2k}(v_2)$ are the sets of self-returning walks of length $2k$ of v_1 and v_2 in S_n, respectively.*

Proof. Let $\xi_1 : W_{2k}(v_2) \to W_{2k}(v_1)$, $\forall w \in W_{2k}(v_2)$, if $w = v_2 v_1 v_{i_1} \ldots v_{i_{2k-3}} v_1 v_2$, then $\xi_1(w) = v_1 v_2 v_1 v_{i_1} \ldots v_{i_{2k-3}} v_1$.

Obviously, ξ_1 is injective. However, there is no $w \in W_{2k}(v_2)$ such that

$$\xi_1(w) = v_1 v_3 v_1 v_3 v_1 \ldots v_3 v_1 \in W_{2k}(v_1)$$

and ξ_1 is not surjective for $n \geqslant 3$ and $k \geqslant 1$. □

Lemma 5.2. [31] *Consider the Transformation I shown in Figure 2. Let u be a non-isolated vertex of a simple graph G. Let G_1 and G_2 be the graphs obtained from G by, respectively, identifying a leaf v_2 and the center v_1 of the n-vertex star S_n with the vertex u, cf. Figure 2. Then $M_{2k}(G_1) < M_{2k}(G_2)$ for $n \geqslant 3$ and $k \geqslant 2$.*

Proof. Let $W_{2k}(G)$ denote the set of self-returning walks of length $2k$ of G. Then $W_{2k}(G_i) = W_{2k}(G) \cup W_{2k}(S_n) \cup A_i$ is a partition, where A_i is the set of self-returning walks of length $2k$ of G_i, each of them containing both at least one edge in $E(G)$ and at least one edge in $E(S_n)$, $i = 1, 2$. So, $M_{2k}(G_i) = |W_{2k}(G)| + |W_{2k}(S_n)| + |A_i| = M_{2k}(G) + M_{2k}(S_n) + |A_i|$. Obviously, it is enough to show that $|A_1| < |A_2|$.

Let $\eta_1 : A_1 \to A_2$, $\forall w \in A_1$, $\eta_1(w) = (w - w \cap S_n) \cup \xi_1(w \cap S_n)$, i.e., $\eta_1(w)$ is the self-returning walk of length $2k$ in A_2 obtained from w by replacing its every maximal (v_2, v_2)-section in S_n (which is a self-returning walk of v_2 in S_n) with its image under the map ξ_1.

By Lemma 5.1, ξ_1 is injective. It is easily shown that η_1 is also injective. However, there is no $w \in A_1$ such that $\eta_1(w) \in A_2$ and $\eta_1(w)$ does not pass the edge v_1v_2 in G_2. So, η_1 is not surjective. Consequently, $|A_1| < |A_2|$ and $M_{2k}(G_1) < M_{2k}(G_2)$. □

Lemma 5.3. [31] *Let $P_n = v_1v_2 \dots v_n$ be the n-vertex path, depicted in Figure 1. Then there is an injection ξ_2 from $W'_{2k}(v_1)$ to $W'_{2k}(v_t)$, and ξ_2 is not a surjection for $n \geqslant 3$, $1 < t < n$ and $k \geqslant 1$, where $W'_{2k}(v_1)$ and $W'_{2k}(v_t)$ are the sets of self-returning walks of length $2k$ of v_1 and v_t in P_n, respectively.*

Proof. First, let $f : \{v_1, v_2, \dots, v_t\} \to \{v_1, v_2, \dots, v_t\}$, $f(v_i) = v_{t-i+1}$ for $i = 1, 2, \dots, t$. Then we can induce a bijection by f from the set of self-returning walks of length $2k$ of v_1 in the sub-path $P_t = v_1v_2 \dots v_t$ and the set of self-returning walks of length $2k$ of v_t in P_t.

Secondly, let $\xi_2 : W'_{2k}(v_1) \to W'_{2k}(v_t)$, $\forall w \in W'_{2k}(v_1)$.

(i) If w is a walk of $P_t = v_1v_2 \dots v_t$, i.e., w does not pass the edge v_tv_{t+1}, then $\xi_2(w) = f(w)$.

(ii) If w passes the edge v_tv_{t+1}, we can decompose w into $w = w_1 \cup w_2 \cup w_3$, where w_1 is the first (v_1, v_t)-section of w, w_3 is the last (v_t, v_1)-section of w, and the rest w_2 is the internal maximal (v_t, v_t)-section of w, i.e., w is a self-returning walk of v_1, first passing the walk w_1 from v_1 to v_t, next passing the walk w_2 from v_t to v_t, and last passing the walk w_3 from v_t to v_1; then $\xi_2(w) = w_1^{-1} \cup w_3^{-1} \cup w_2$, that is, $\xi_2(w)$ is a self-returning walk v_t, first passing the reverse of w_1 from v_t to v_1, next passing the reverse of w_3 from v_1 to v_t, and last passing the walk w_2 from v_t to v_t.

Obviously, ξ_2 is injective. And ξ_2 is not surjective since there is no $w \in W'_{2k}(v_1)$ such that $\xi_2(w)$ is a self-returning walk not passing the edge v_tv_{t-1} in P_n of length $2k$ of v_t. □

Lemma 5.4. [31] *Let u be a non-isolated vertex of a simple graph H. If H_1 and H_2 are the graphs obtained from H by identifying, respectively, an end vertex v_1 and an internal vertex v_t of the n-vertex path P_n to u, cf. Figure 3, then $M_{2k}(H_1) < M_{2k}(H_2)$ for $n \geqslant 3$ and $k \geqslant 2$.*

Proof. Let B_i be the set of self-returning walks of length $2k$ of H_i, each of them containing both at least one edge in $E(H)$ and at least one edge in $E(P_n)$, $i = 1, 2$. Similarly to the proof of Lemma 5.2, it is enough to show that $|B_1| < |B_2|$.

Let $\eta_2 : B_1 \to B_2$, $\forall w \in B_1$, $\eta_2(w) = (w - w \cap P_n) \cup \xi_2(w \cap P_n)$, i.e., $\eta_2(w)$ is the self-returning walk of length $2k$ in B_2 obtained from w by replacing its every section in P_n (which is a self-returning walk of v_1 in P_n) with its image under the map ξ_2.

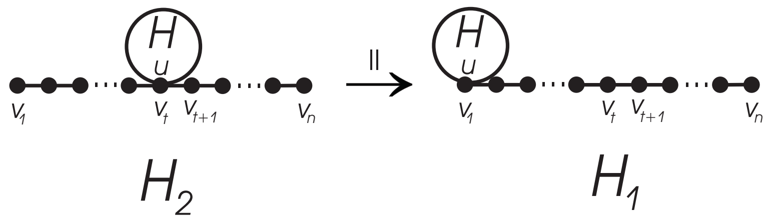

Figure 3. Transformation II.

By Lemma 5.3, ξ_2 is injective. It follows that η_2 is also injective. But, η_2 is not surjective since there is no $w \in B_1$ with $\eta_2(w) \in B_2$ not passing the edges $v_t v_{t-1}$ in H_2. So, $|B_1| < |B_2|$. □

Theorem 5.5. [31] *If T_n is a n-vertex tree different from S_n and P_n, then*

$$EE(P_n) < EE(T_n) < EE(S_n). \tag{12}$$

Proof. Repeating Transformation I, as shown in Figure 2, any n-vertex tree T can be changed into the n-vertex star S_n. By Lemma 5.2, we have $M_{2k}(T) < M_{2k}(S_n)$ for $k \geqslant 2$. This implies

$$EE(T) = \sum_{k\geqslant 0} \frac{M_{2k}(T)}{(2k)!} < \sum_{k\geqslant 0} \frac{M_{2k}(S_n)}{(2k)!} = EE(S_n).$$

On the other hand, repeating Transformation II, as shown in Figure 3, any n-vertex tree T can be changed into the n-vertex path P_n. By Lemma 5.4, we have $M_{2k}(T) > M_{2k}(P_n)$ for $k \geqslant 2$. Consequently,

$$EE(T) = \sum_{k\geqslant 0} \frac{M_{2k}(T)}{(2k)!} > \sum_{k\geqslant 0} \frac{M_{2k}(P_n)}{(2k)!} = EE(P_n).$$

So the inequalities (12) hold. □

Theorem 5.5 shows that the path P_n and the star S_n have the minimum and the maximum Estrada indices among n-vertex trees, i.e., Conjecture A is true.

Zhao and Jia [32] have determined also the trees with the second and the third greatest Estrada index. In fact, they proved:

Theorem 5.6. [32] *Let $S_n^1 \cong S_n$ be the n-vertex star, cf. Figure 1, and let the n-vertex trees S_n^i, $i = 2,3,4,5,6$, be those shown in Figure 4. Let T_1 and T_2 be n-vertex trees, such that $T_1 \notin \{S_n^i \mid i = 1,2,3,4,5,6\}$ and $T_2 \notin \{S_n^i \mid i = 1,2,3\}$. Then for $n \geqslant 6$,*

$$EE(S_n^1) > EE(S_n^2) > EE(S_n^3) > EE(S_n^5) > EE(S_n^6) > EE(T_1)$$

and

$$EE(S_n^1) > EE(S_n^2) > EE(S_n^3) > EE(T_2).$$

Consequently, among n-vertex trees, the first three trees with the greatest Estrada indices are S_n, S_n^2 and S_n^3, respectively.

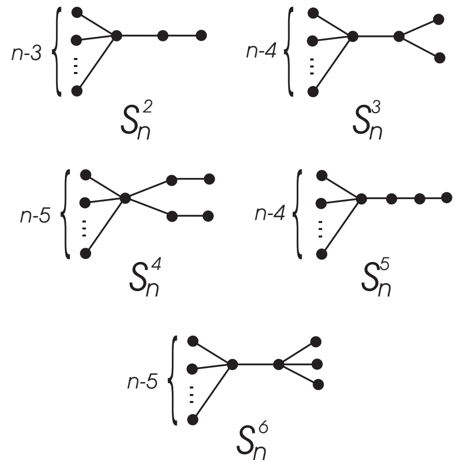

FIGURE 4. The graphs S_n^i, $i = 2, 3, 4, 5, 6$, having the second, third, fourth, fifth, and sixth greatest Estrada indices among n-vertex trees [32, 33].

Recently it was demonstrated [33] that $EE(S_n^3) > EE(S_n^4) > EE(S_n^5)$, from which follows:

Theorem 5.7. [33] *Among n-vertex trees, $n \geqslant 6$, the first six trees with the greatest Estrada indices are S_n, S_n^2, S_n^3, S_n^4, S_n^5, S_n^6, respectively, cf. Figure 4.*

Let G be a connected graph of order n and let e be an edge of G. The graph $G' = G - e$ is obtained from G by deleting the edge e. Obviously, any self-returning walk of length k of G' is also a self-returning walk of length k of G. Thus,

$$M_k(G') \leqslant M_k(G) \quad \text{and} \quad EE(G') \leqslant EE(G).$$

In particular, if T is a spanning tree of G, then

$$M_k(T) \leqslant M_k(G) \quad \text{and} \quad EE(T) \leqslant EE(G).$$

From Theorem 5.5 it follows that $EE(P_n) \leqslant EE(G)$. So, we have:

Theorem 5.8. [31] *If G is a simple connected graph of order n different from the complete graph K_n and the path P_n, then*

$$EE(P_n) < EE(G) < EE(K_n).$$

Theorem 5.7 shows that the path P_n and the complete graph K_n have the minimum and the maximum Estrada indices among connected graphs of order n, i.e., Conjecture B is true.

6. Estrada indices of molecular graphs

In view of the chemical origin of the Estrada index, it is natural than molecular graphs [34], especially acyclic and benzenoid, were among the first whose structure–dependence was systematically examined.

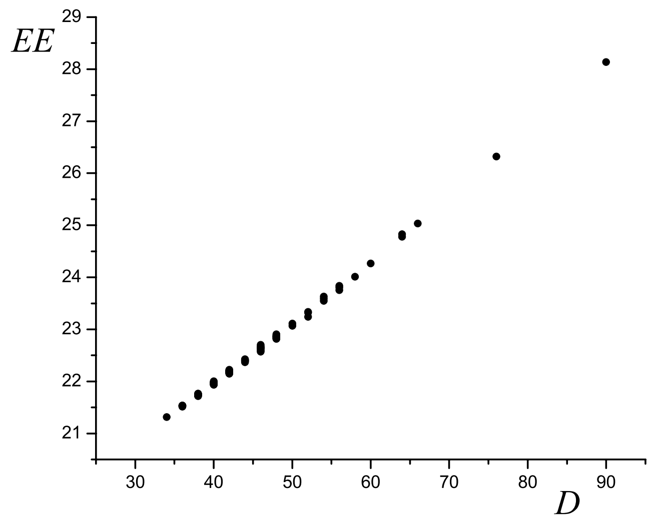

FIGURE 5. Correlation between the Estrada indices and the parameter D (= sum of squares of vertex degrees) for the 106 trees on 10 vertices.

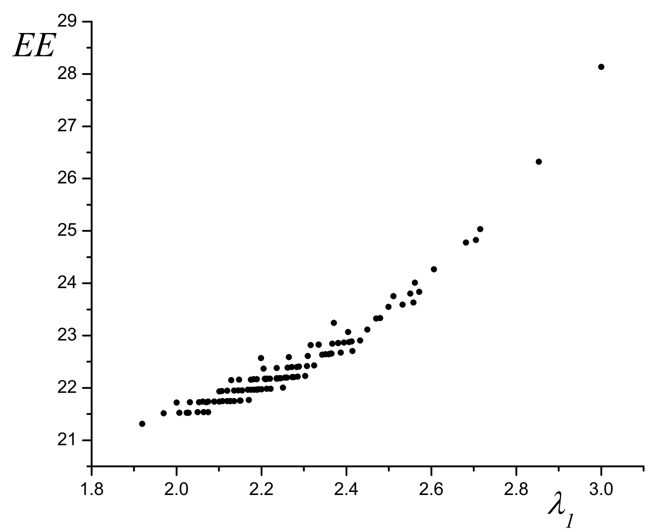

FIGURE 6. Correlation between the Estrada index (EE) and the greatest graph eigenvalue λ_1 for the 106 trees on 10 vertices.

A chemical tree is a tree in which no vertex has degree greater than four [34]. Among the n-vertex chemical trees, P_n has minimum Estrada index. For the

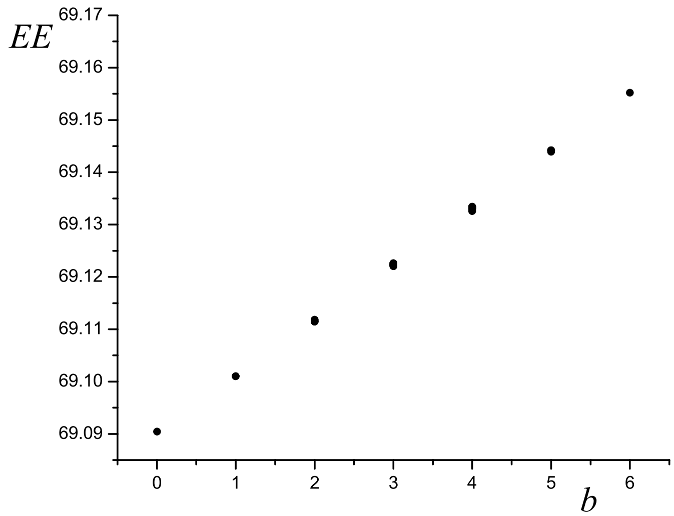

FIGURE 7. Correlation between the Estrada indices (EE) of the 36 catacondensed benzenoid systems with 6 hexagons and the number b of their bay regions.

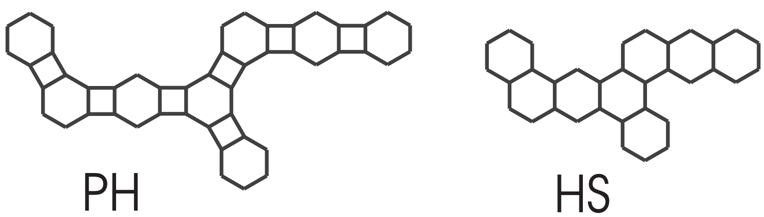

FIGURE 8. A phenylene (PH) and its hexagonal squeeze (HS).

Estrada index of chemical trees it was concluded [35] that the n-vertex chemical tree with the greatest Estrada index might be the Volkmann tree $VT_n(4)$. However, this assertion cannot be considered as proven in a rigorous mathematical manner. Such a proof awaits to be achieved in the future.

In the case of trees with a fixed number of vertices (including both chemical and non-chemical trees) it was found that EE increases with the increasing extent of branching [36]. This fact motivated investigations of the relation between EE and other branching indices. It was established that there is a linear correlation between EE and the quantity $D = \sum_{i=1}^{n}(d_i)^2$, earlier encountered in Theorem 3.4, see Figure 5.

The quantitative analysis of these correlations resulted in the following approximate expression:

$$EE \approx 1.735\, n - 0.13 + 0.11\, D.$$

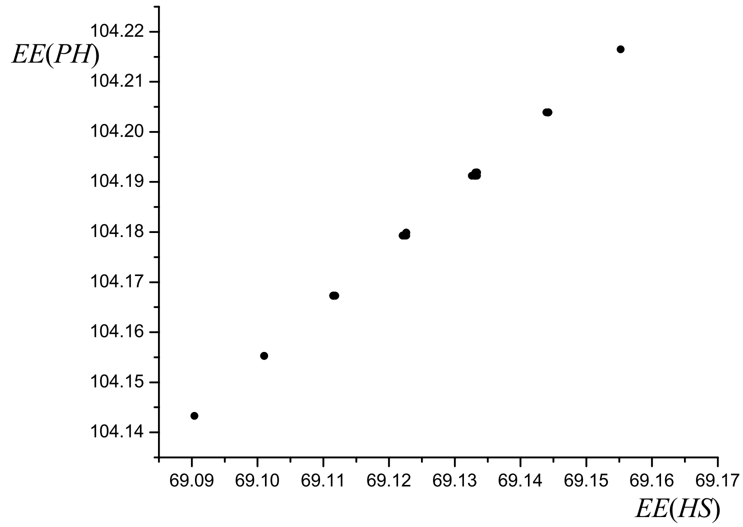

FIGURE 9. Correlation between the Estrada indices of phenylenes, $EE(PH)$, and the Estrada indices of the corresponding hexagonal squeezes, $EE(HS)$. The data points shown in this figure pertain to phenylenes with 6 hexagons; there are 37 species of this kind.

This formula is capable of reproducing EE with an error less than 0.1%.

The Estrada index of trees was also correlated with the greatest graph eigenvalue [35, 37]; a characteristic example of such correlations is shown in Figure 6. One can see that the EE/λ_1 relation is not simple. The fact that the (EE, λ_1) data points are grouped on several (almost) horizontal lines indicates that EE is much less sensitive to structural features than λ_1.

Empirical studies revealed that the number of vertices n and number of edges m are the main factors influencing EE-value of molecular graphs [29, 36, 38]. For benzenoid systems, (m, n)-type approximations are capable of reproducing over 99.8% of EE-value [29, 38]. In order to find some finer structural details on which EE depends, series of isomeric benzenoid systems (having equal n and m) were examined. The Estrada indices of benzenoid isomers vary only to a very limited extent. The main structural feature influencing these variations is the number of bay regions, b. (The quantity b is equal to the number of edges on the boundary of a benzenoid graph, connecting two vertices of degree 3; for details see [39].) Within sets of benzenoid isomers, EE is an increasing linear function of b, see Figure 7.

Phenylenes are molecular graphs consisting of hexagons and squares, joined in a manner that should be evident from the example depicted in Figure 8. To each phenylene a so-called "hexagonal squeeze" can be associated, containing only hexagons, cf. Figure 8.

The Estrada index of phenylenes was studied in [40]. Within sets of isomers (having equal number of hexagons) a good linear correlation exists between the Estrada index of phenylenes, $EE(PH)$ and of the corresponding hexagonal squeezes, $EE(HS)$, see Figure 9. Bearing in mind that the hexagonal squeezes are benzenoid systems, and that the structure-dependence of EE of benzenoids is almost completely understood, the good linear correlation between $EE(PH)$ and $EE(HS)$ resolves also the problem of structure–dependence of the Estrada index of phenylenes.

Concluding this section we wish to clearly emphasize that the relations established for molecular graphs, in particular those illustrated in Figures 5, 6, 7, and 9, are empirical findings that have not (yet) been proven in a rigorous mathematical manner. It should be a challenge for the reader of this article to accomplish the needed proofs.

References

[1] D. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs – Theory and Application, third ed.*, Johann Ambrosius Barth Verlag, Heidelberg, Leipzig, 1995.

[2] J. A. de la Peña, I. Gutman, J. Rada, *Estimating the Estrada index*, Lin. Algebra Appl. **427** (2007), 70–76.

[3] E. Estrada, *Characterization of 3D molecular structure*, Chem. Phys. Lett. **319** (2000), 713–718.

[4] E. Estrada, *Characterization of the folding degree of proteins*, Bioinformatics **18** (2002), 697–704.

[5] E. Estrada, *Characterization of the amino acid contribution to the folding degree of proteins*, Proteins **54** (2004), 727–737.

[6] E. Estrada, J. A. Rodríguez–Velázquez, *Subgraph centrality in complex networks*, Phys. Rev. **E 71** (2005), 056103–1–9.

[7] E. Estrada, J. A. Rodríguez–Velázquez, *Spectral measures of bipartivity in complex networks*, Phys. Rev. **E 72**(2005), 046105–1–6.

[8] E. Estrada, *Topological structural classes of complex networks*, Phys. Rev. **E 75** (2007), 016103–1–12.

[9] E. Estrada, J. A. Rodríguez–Velázquez, M. Randić, *Atomic branching in molecules*, Int. J. Quantum Chem. **106** (2006), 823–832.

[10] E. Estrada, N. Hatano, *Statistical–mechanical approach to subgraph centrality in complex networks*, Chem. Phys. Lett. **439** (2007), 247–251.

[11] E. Estrada, *Atom–bond connectivity and the energetic of branched alkanes*, Chem. Phys. Lett. **463** (2008), 422–425.

[12] R. Carbó-Dorca, *Smooth function topological structure descriptors based on graph–spectra*, J. Math. Chem. **44** (2008), 373–378.

[13] A. Gursoy, O. Keskin, R. Nussinov, *Topological properties of protein interaction networks from a structural perspective*, Biochem. Soc. Trans. **36** (2008), 1398–1403.

[14] C. Y. Lin, C. H. Chin, H. H. Wu, S. H. Chen, C. W. Ho, M. T. Ko, *HUBBA: Hubs objects analyzer – a framework of interactome hubs identification for network biology*, Nucleic Acids Res. **36** (2008), 438–443.

[15] A. Platzer, P. Perco, A. Lukas, B. Mayer, *Characterization of protein–interaction networks in tumors*, BMC Bioinformatics **8** (2007), 224–229.

[16] E. Zotenko, J. Mestre, D. P. O'Leary, T. M. Przytycka, *Why do hubs in the yeast protein interaction network tend to be essential: Re-examining the connection between the network topology and essentiality*, Computat. Biol. **4** (2008), 1–16.

[17] M. Jungsbluth, B. Burghardt, A. K. Hartmann, *Fingerprinting networks: Correlations of local and global network properties*, Physica **A381** (2007), 444–456.

[18] V. V. Kerrenbroeck, E. Marinari, *Ranking vertices or edges of a network by loops: A new approach*, Phys. Rev. Lett. **101** (2008), 098701
[19] T. Došlić, *Bipartiviy of fullerene graphs and fullerene stability*, Chem. Phys. Lett. **412** (2005), 336–340.
[20] L. Da Fontoura Costa, M. A. R. Tognetti, F. N. Silva, *Concentric characterization and classification of complex network nodes: Application to an institutional collaboration network*, Physica **A387** (2008), 6201–6214.
[21] I. Gutman, E. Estrada, J. A. Rodríguez–Velázquez, *On a graph–spectrum–based structure descriptor*, Croat. Chem. Acta **80** (2007), 151–154.
[22] G. H. Fath-Tabar, A. R. Ashrafi, I. Gutman, *Note on Estrada and L-Estrada indices of graphs*, Bull. Acad. Serbe Sci. Arts. (Cl. Math. Natur.), in press.
[23] B. Zhou, *On Estrada index*, MATCH Commun. Math. Comput. Chem. **60** (2008), 485–492.
[24] I. Gutman, *Lower bounds for Estrada index*, Publ. Inst. Math. (Beograd), **83** (2008), 1–7.
[25] I. Gutman, S. Radenković, *A lower bound for the Estrada index of bipartite molecular graphs*, Kragujevac J. Sci. **29** (2007), 67–72.
[26] I. Gutman, A. Graovac, *Estrada index of cycles and paths*, Chem. Phys. Lett. **436** (2007), 294–296.
[27] Y. Ginosar, I. Gutman, T. Mansour, M. Schork, *Estrada index and Chebyshev polynomials*, Chem. Phys. Lett. **454** (2008), 145–147.
[28] M. Robbiano, R. Jiménez, L. Medina, *The energy and an approximation to Estrada index of some trees*, MATCH Commun. Math. Comput. Chem. **61** (2009), 369–382.
[29] I. Gutman, S. Radenković, A. Graovac, D. Plavšić, *Monte Carlo approach to Estrada index*, Chem. Phys. Lett. **446** (2007), 233–236.
[30] T. Aleksić, I. Gutman, M. Petrović, *Estrada index of iterated line graphs*, Bull. Acad. Serbe Sci. Arts (Cl. Sci. Math. Natur.) **134** (2007), 33–41.
[31] H. Deng, *A proof of a conjecture on the Estrada index*, MATCH Commun. Math. Comput. Chem. **62** (2009), 599–606.
[32] H. Zhao, Y. Jia, *On the Estrada index of bipartite graphs*, MATCH Commun. Math. Comput. Chem. **61** (2009), 495–501.
[33] H. Deng, *A note on the Estrada index of trees*, MATCH Commun. Math. Comput. Chem. **62** (2009), 607–610.
[34] I. Gutman, O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer–Verlag, Berlin, 1986.
[35] I. Gutman, B. Furtula, V. Marković, B. Glišić, *Alkanes with greatest Estrada index*, Z. Naturforsch. **62a** (2007), 495–498.
[36] I. Gutman, B. Furtula, B. Glišić, V. Marković, A. Vesel, *Estrada index of acyclic molecules*, Indian J. Chem. **46** (2007), 723–728.
[37] I. Gutman, S. Radenković, B. Furtula, T. Mansour, M. Schork, *Relating Estrada index with spectral radius*, J. Serb. Chem. Soc. **72** (2007), 1321–1327.
[38] I. Gutman, S. Radenković, *Estrada index of benzenoid hydrocarbons*, Z. Naturforsch. **62a** (2007), 254–258.
[39] I. Gutman, S. J. Cyvin, *Introduction to the Theory of Benzenoid Hydrocarbons*, Springer–Verlag, Berlin, 1989.
[40] B. Furtula, I. Gutman, *Energy and Estrada index of phenylenes*, Indian J. Chem. **47A** (2008), 220–224.