MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **72** (2014) 753-760

ISSN 0340 - 6253

The Second Zagreb Index of Molecular Graphs with Tree Structure

Ramin Kazemi

Department of Statistics
Imam Khomeini International University, Qazvin, Iran
r.kazemi@SCI.ikiu.ac.ir

(Received July 19, 2014)

Abstract

The second Zagreb index of (molecular) graph G is equal to the sum of the products of the degrees of pairs of adjacent vertices. In this paper, we study the second Zagreb index of molecular graphs with tree structure. Exactly, we obtain the average and variance of this index in a randomly chosen molecular graph with n vertices. Also, the asymptotic normality of the second Zagreb index is given.

1 Introduction

Let G be a molecular graph. Two vertices of G, connected by an edge, are said to be adjacent. The number of vertices of G, adjacent to a given vertex v, is the degree of this vertex, and will be denoted by $d(v)$. Analyzing the structure-dependency of total π-electron energy [6], an approximate formula was obtained in which terms of the form

$$M_1(G) = \sum_{v \in V(G)} (d(v))^2$$

and

$$M_2(G) = \sum_{uv \in E(G)} d(u)d(v)$$

occured where $V(G)$ and $E(G)$ are the vertex and edge sets of a graph G, respectively. It was immediately recognized that these terms increase with the increasing extent of branching of the carbon-atom skeleton, i.e., that these provide quantitative measures of molecular branching. Ten years later, in a review article, Balaban *et al.* included M_1 and M_2 among topological indices and named them *Zagreb group indices* [1]. The name Zagreb group index was soon abbreviated to *Zagreb index*, and nowadays M_1 is referred to as the first Zagreb index whereas M_2 as the second Zagreb index (For the chemical relevance of the second Zagreb index, see [5] and the references quoted therein). The research background of the Zagreb indices together with its generalization appears in chemistry or mathematical chemistry (see for examples: [1], [3], [4], [7], [12] and [13], and references therein). An illustrative example is provided in Figure 1.

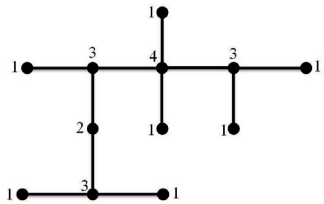

Figure 1: The molecular graph G with its node degrees indicated. In graph G, $M_1(G) = 54$ and $M_2(G) = 59$.

The structures of many molecules such as dendrimers and acyclic molecules are tree like. We present the following evolution process for random trees of size n, which turns out to be appropriate when studying the second Zagreb index of molecular graphs with tree structure.

Evolution process: Every size-n tree can be obtained uniquely by attaching nth node to one of the $n-1$ nodes in a tree of size $n-1$. It is of particular interest in applications to assume the random tree model and to speak about a random tree with n nodes, which means that all trees of size n are considered to appear equally likely. Equivalently one may describe random trees via the following tree evolution process, which generates random trees of arbitrary size n. At step 1 the process starts with the root. At step i the ith node is attached to any previous node v of the already grown tree T of size $i-1$ with probability $p_i(v) = \frac{1}{i-1}$. For applicability of our own results and specially connection with the chemical relevance, see [11].

Let Z_n be the first Zagreb index of a random tree of size n and $\mathcal{F}_n$ be the sigma-field generated by the first n stages of these trees. Let U_n be a randomly chosen node belong to a size-n tree. Kazemi [9] studied the first Zagreb index of random trees through the following recurrence equation:

$$Z_n = Z_{n-1} + 2(d(U_{n-1}) + 1) \tag{1}$$

where $d(U_{n-1})$ is the degree of the randomly chosen node U_{n-1}. He proved that [9, Corollary 2.6] for random trees (or molecular graphs with tree structure):

$$
\begin{aligned}
E(Z_n) &= 6(n-1) - 4H_{n-1} = 6n + \mathcal{O}(\log n) \\
Var(Z_n) &= 8n + \mathcal{O}(\log^2 n) \qquad n \geq 2
\end{aligned}
$$

where H_n is the nth Harmonic number.

The paper is organized as follows. We first show a recurrence for the second Zagreb index of molecular graphs with tree structure. Second, we give the mean and variance of the second Zagreb index. Then the asymptotic normality of the second Zagreb index is proved as the size grows to infinity.

2 The Main Results

Let M_n be the second Zagreb index of a random tree of size n. Let U_{n-1} be a randomly chosen node belong to a size-$n-1$ tree. Then by definition of the second Zagreb index and stochastic growth rule of the tree,

$$
M_n = M_{n-1} + \sum_{i=0}^{d(U_{n-1})} d(U_{n-1-i}) + 1 \, . \tag{2}
$$

Leaves have the degree 1 and another nodes have the degree>1. Thus we obtain the $E(M_n)$ and $Var(M_n)$ separably in two cases: 1) nth node attached to a leaf U_{n-1}, 2) nth node attached to a non-leaf node U_{n-1}. It is reasonable to consider the mean and variance of M_n as the average of the mean and variance of these cases, respectively. If nth node attached to a non-leaf node U_{n-1}, then the number of changes in M_{n-1} is equal to $d(U_{n-1}) + 1$. First, suppose nth node attached to a leaf U_{n-1}. Then

$$
M_n = M_{n-1} + d(U_{n-2}) + 2 \, , \tag{3}
$$

where $d(U_{n-2})$ is the degree of parent of leaf U_{n-1}. Hence

$$
E(M_n|\mathcal{F}_{n-1}) = M_{n-1} + E(d(U_{n-2})|\mathcal{F}_{n-1}) + 2 \, , \tag{4}
$$

since M_{n-1} is $\mathcal{F}_{n-1}$-measurable. The sum of the degrees of all nodes in a tree of size n is $2(n-1)$. Thus

$$
E(d(U_n)|\mathcal{F}_n) = \sum_{i=1}^{n} d(u_i)\frac{1}{n} = 2\Big(1 - \frac{1}{n}\Big) \, . \tag{5}
$$

Taking expectation of the relation (4) with respect to $\mathcal{F}_{n-2}$:

$$\begin{aligned} E(E(M_n|\mathcal{F}_{n-1})|\mathcal{F}_{n-2}) &= E(M_{n-1}|\mathcal{F}_{n-2}) \\ &+ E(E(d(U_{n-2})|\mathcal{F}_{n-1})|\mathcal{F}_{n-2}) + 2 \\ &= E(M_{n-1}|\mathcal{F}_{n-2}) + E(d(U_{n-2})|\mathcal{F}_{n-2}) + 2 \\ &= E(M_{n-1}|\mathcal{F}_{n-2}) + 2\Big(1 - \frac{1}{n-2}\Big) + 2\,, \end{aligned} \tag{6}$$

since $\mathcal{F}_{n-2} \subset \mathcal{F}_{n-1}$ [2]. Again, taking expectation of the relation (6):

$$E(M_n) = E(M_{n-1}) + 2\Big(1 - \frac{1}{n-2}\Big) + 2\,. \tag{7}$$

We have the initial conditions

$$M_k = \begin{cases} 0, & \text{if} \quad k = 1 \\ 1, & \text{if} \quad k = 2\,. \end{cases}$$

Thus, the recurrence equation (7) leads to

$$E(M_n) = 1 + 4(n-2) - 2H_{n-2} = 4n + \mathcal{O}(\log n)\,. \tag{8}$$

Now, taking expectation of the relation (2) with respect to $\mathcal{F}_{n-1}$, ..., $\mathcal{F}_{n-1-m}$, respectively

$$E(M_n|\mathcal{F}_{n-1}, d(U_{n-1}) = m) = M_{n-1} + \sum_{i=0}^{m} 2\Big(1 - \frac{1}{n-1-i}\Big) + 1\,, \tag{9}$$

since $\mathcal{F}_{n-1-m} \subset \cdots \subset \mathcal{F}_{n-1}$ and where $m \geq 2$. Then

$$\begin{aligned} E(M_n|d(U_{n-1}) = m) &= E(M_{n-1}) + 2(m+1) \\ &- 2(H_{n-(m+1)} - H_{n-2}) + 1\,. \end{aligned} \tag{10}$$

Again, the recurrence equation (10) leads to

$$\begin{aligned} E(M_n|d(U_{n-1}) = m) &= 1 + 2(m+1)(n-2) \\ &- \sum_{i=1}^{n-2} 2(H_{n-(m+i)} - H_{n-(i-1)}) + (n-2) \\ &= (2(m+1)+1)n + \mathcal{O}(\log n)\,. \end{aligned} \tag{11}$$

Hence, from (5), (8) and (11),

$$E(M_n) = \frac{11}{2} n + \mathcal{O}(\log n)\,, \tag{12}$$

since $E(d(U_{n-1})|\mathcal{F}_{n-1}) = 2(1 - (n-1)^{-1})$. Again, we first consider the simple case introduced in (3). Let $Y_1 = Y_2 = 0$ and for $n > 2$,

$$Y_n := M_n - M_{n-1} - 4 + \frac{2}{n-2}\,.$$

Then $E(Y_n|\mathcal{F}_{n-1}) = 0$. From (3),

$$
\begin{aligned}
E((M_n - M_{n-1} - 2)^2|\mathcal{F}_{n-1}) &= E(d^2(U_{n-2})) \\
&= \sum_{i=1}^{n-2} d^2(u_i)\frac{1}{n-2} \\
&= \frac{Z_{n-2}}{n-2} \, . \qquad (13)
\end{aligned}
$$

Also

$$
\begin{aligned}
\frac{Z_{n-2}}{n-2} &= E((M_n - M_{n-1} - 2)^2|\mathcal{F}_{n-1}) \\
&= E\left(\left(Y_n + 2 - \frac{2}{n-2}\right)^2 |\mathcal{F}_{n-1}\right) \\
&= E(Y_n^2|\mathcal{F}_{n-1}) + \left(2 - \frac{2}{n-2}\right)^2 . \qquad (14)
\end{aligned}
$$

Taking expectation of the relation (14),

$$
\begin{aligned}
E(Y_n^2) &= \frac{E(Z_{n-2})}{n-2} - \left(2 - \frac{2}{n-2}\right)^2 \\
&= \frac{6(n-3) - 4H_{n-3}}{n-2} - \left(2 - \frac{2}{n-2}\right)^2 \\
&= 2 + \mathcal{O}\left(\frac{\log n}{n}\right), \qquad n \geq 3 \, . \qquad (15)
\end{aligned}
$$

Now, from (15),

$$
Var(M_n) = \sum_{i=1}^{n} E(Y_i^2) = 2n + \mathcal{O}(\log^2 n)\, , \qquad (16)
$$

since for any $1 \leq i \neq j \leq n$, $E(Y_iY_j) = 0$ [2]. Now, we study the general case introduced in (2). Let $d(U_{n-1}) = m \geq 2$. Then

$$
\begin{aligned}
& E((M_n - M_{n-1} - 1)^2|\mathcal{F}_{n-1}) \\
=\; & E\Big(\sum_{i=0}^{m} d(U_{n-1-i})\Big)^2 \\
=\; & \sum_{i=0}^{m} \frac{Z_{n-1-i}}{n-1-i} + 2(m-1)\sum_{i=0}^{m} \frac{M_{n-1-i}}{n-1-i} \, . \qquad (17)
\end{aligned}
$$

Now, let $Y_1' = Y_2' = 0$ and for $n > 2$,

$$
Y_n' := M_n - M_{n-1} - 2((m+1)+1) + 2(H_{n-(m+1)} - H_{n-2}) \, .
$$

Thus

$$\begin{aligned} & E((M_n - M_{n-1} - 1)^2 | \mathcal{F}_{n-1}) \\ = \; & E\left(\left(Y_n' + 2(m+1) - 2(H_{n-(m+1)} - H_{n-2}) \right)^2 | \mathcal{F}_{n-1} \right) \\ = \; & E(Y_n'^2 | \mathcal{F}_{n-1}) + \left(2(m+1) - 2(H_{n-(m+1)} - H_{n-2}) \right)^2 . \end{aligned} \tag{18}$$

Hence

$$\begin{aligned} E(Y_n'^2) &= 6(m+1) + 11m(m-1) \\ &- \; 4(m+1)^2 + \mathcal{O}\Big(\frac{\log n}{n}\Big) . \end{aligned} \tag{19}$$

Now, from (5), (16) and (19),

$$Var(M_n) = 3n + \mathcal{O}(\log^2 n) ,$$

Thus we have the following theorem:

Theorem 1. *Let M_n be the second Zagreb index of a molecular graph with tree structure of size n. Then*

$$\begin{aligned} E(M_n) &= \frac{11}{2} n + \mathcal{O}(\log n) , \\ Var(M_n) &= 3n + \mathcal{O}(\log^2 n) . \end{aligned}$$

The process $M = (M_n - E(M_n), \; n \geq 1)$ is a martingale [2]. The following corollary is an immediate consequence of Chebyshev's inequality

Corollary 1. *We have [9, 10]:*

$$\begin{aligned} & \frac{M_n}{E(M_n)} \xrightarrow{P} 1 , \\ & \frac{M_n}{n} \xrightarrow{P} 4 , \\ & \frac{1}{n} \sum_{i=1}^{n} \frac{M_i - E(M_i)}{i} \xrightarrow{P} 0 . \end{aligned}$$

Let $d_{n,j}$ be the out-degree of jth node in a random tree of size n. Then $d_{n,j} \leq 1 + \sum_{i=j+1}^{n} I_{ji}$, where $I_{ji} = I(j\text{th node is the parent of } i\text{th node})$. It is obvious that I_{ji} are independent for fixed j and $I_{ji} \sim \mathrm{Ber}(\frac{1}{i-1})$. For any $\varepsilon > 0$,

$$\begin{aligned} P\Big(\max_j d_{n,j} > \varepsilon \sqrt{n} \Big) &\leq \frac{1}{\varepsilon^3 n^{3/2}} \sum_{j=1}^{n} E\Big(1 + \sum_{i=j+1}^{n} I_{ji} \Big)^3 \\ &\leq \frac{1}{\varepsilon^3 n^{3/2}} \sum_{j=1}^{n} \left(4 + 4E\Big(\sum_{i=j+1}^{n} I_{ji} - \frac{1}{i-1} + \frac{1}{i-1} \Big)^3 \right) \\ &= \frac{1}{\varepsilon^3 n^{3/2}} \sum_{j=1}^{n} \mathcal{O}\left(\log^3 n \right) \to 0 . \end{aligned} \tag{20}$$

Theorem 2. *As $n \to \infty$,*

$$\frac{M_n - \frac{11}{2}n}{\sqrt{3n}} \xrightarrow{D} N(0,1) \ .$$

Proof. Let

$$X_{n,i} = \frac{Y_i}{\sqrt{3n}}, \quad i = 1, 2, ..., n, n \geq 1 \ .$$

By [8, Corollary 3.1], it is sufficient to show that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P\Big(\Big|\sum_{i=1}^{n} E(X_{n,i}^2|\mathcal{F}_{i-1}) - 1 > \varepsilon\Big) = 0 \tag{21}$$

and

$$\lim_{n\to\infty} P\Big(\Big|\sum_{i=1}^{n} E(X_{n,i}^2 I(|X_{n,i}| > \varepsilon)|\mathcal{F}_{i-1}) > \varepsilon\Big) = 0 \ . \tag{22}$$

First,

$$\begin{aligned}
\sum_{i=1}^{n} E(X_{n,i}^2|\mathcal{F}_{i-1}) &= \frac{1}{3n}\sum_{i=1}^{n} E(Y_i^2|\mathcal{F}_{i-1}) \\
&= \frac{1}{3n}\sum_{j=1}^{n-2}\left(\frac{M_j}{j} - \left(2 - \frac{2}{j}\right)^2\right) \\
&= \frac{1}{3n}\sum_{j=1}^{n-2}\left(\frac{M_j - E(M_j)}{j}\right) \\
&+ \frac{1}{3n}\sum_{j=1}^{n-2}\left(\frac{E(M_j)}{j} - \left(2 - \frac{2}{j}\right)^2\right) \ .
\end{aligned}$$

Thus (21) follows by Corollary 1 and (20). Also

$$\begin{aligned}
& \lim_{n\to\infty} P\Big(\Big|\sum_{i=1}^{n} E(X_{n,i}^2 I(|X_{n,i}| > \varepsilon)|\mathcal{F}_{i-1}) > \varepsilon\Big) \\
\leq\ & P\Big(\max_j d_{n,j} > \varepsilon\sqrt{n}\Big) + I\Big(\max_j d_{n,j} > \varepsilon\sqrt{n}\Big) \\
\to\ & 0 \ .
\end{aligned}$$

We can obtain the same results if $X_{n,i} = \frac{Y_i'}{\sqrt{3n}}$ and thus proof is completed.

■

3 General Mathematical Formulation

Gutman [5] introduced the following general form for topological indices:

$$TI = TI(G) = \sum_{uv\in E(T)} F(d(u), d(v)) \ ,$$

where the summation goes over all pairs of adjacent nodes u, v of molecular graph G, and where $F = F(x, y)$ is an appropriately chosen function. In particular, $F(x, y) = (xy)^{-\frac{1}{2}}$ for Randić index, $F(x, y) = x + y$ for the first Zagreb index, $F(x, y) = xy$ for the second Zagreb index, $F(x, y) = |x - y|$ for the third Zagreb index, $F(x, y) = (xy)^{\lambda}$ $(\lambda \in R)$ for the second variable Zagreb index, $F(x, y) = ((x + y - 2)(xy)^{-1})^{\frac{1}{2}}$, for the ABC index, $F(x, y) = (xy(x + y - 2)^{-1})^3$, for the augmented Zagreb index, $F(x, y) = 2\sqrt{xy}(x + y)^{-1}$ for the geometric-arithmetic index, $F(x, y) = 2(x + y)^{-1}$ for the harmonic index and $F(x, y) = (x + y)^{-\frac{1}{2}}$ for the sum-connectivity index. With the same approach discussed here we can obtain similar results for these quantities in molecular graphs with tree structure.

References

[1] A. T. Balaban, I. Motoc, D. Bondchev, O. Mekenyan, Topological indices for structure–activity correlations, *Topics Curr. Chem.* **114** (1983) 21–55.

[2] P. Billingsley, *Probability and Measure*, Wiley, New York, 1885.

[3] K. C. Das, I. Gutman, B. Zhou, New upper bounds on Zagreb indices, *J. Math. Chem.* **46** (2009) 514–521.

[4] K. C. Das, I. Gutman, Some properties of the second Zagreb index, *MATCH Commun. Math. Comput. Chem.* **52** (2004) 103–112.

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

[6] I. Gutman, N. Trinajstić, Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons, *Chem. Phys. Lett.* **17** (1972) 535–538.

[7] I. Gutman, B. Ruščić, C. F. Wilcox, Graph theory and molecular orbitals. 12. Acyclic polyenes, *J. Chem. Phys.* **62** (1975) 3399–3405.

[8] P. Hall, C. C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, New York, 1980.

[9] R. Kazemi, Probabilistic analysis of the first Zagreb index, *Trans. Comb.* **2**(2) (2013) 35–40.

[10] R. Kazemi, Depth in bucket recursive trees with variable capacities of buckets, *Acta Math. Sin. Engl. Ser.* **30** (2014) 305–310.

[11] R. Kazemi, The eccentric connectivity index of bucket recursive trees, *Iranian J. Math. Chem.* **5** (2014) 77–83.

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

[13] Z. Yan, H. Liu, H. Liu, Sharp bounds for the second Zagreb index of unicyclic graphs, *J. Math. Chem.* **42** (2007) 565–574.