MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **89** (2023) 327–342
ISSN: 0340–6253
doi: 10.46793/match.89-2.327G

On the Number of All Substructures Containing at Most Four Edges

Shicai Gong*, Liping Zhang, Changbao Su

School of Science, Zhejiang University of Science and Technology, Hangzhou, 310023, P. R. China

schangbao20@163.com, zlp92066@sina.com, schangbao20@163.com

(Received May 29, 2022)

Abstract

Let G be a simple graph with order n, $n \geq 5$, and adjacency matrix $\mathbf{A}(G)$. In this paper, we determine the number of all substructures having at most four edges in terms of its adjacency matrix $\mathbf{A}(G)$ together with some graph invariants determined by $\mathbf{A}(G)$. Then, as applications, we provide an algebraic expression for the second Zagreb index and $||\mathbf{A}^4||$ of a graph.

1 introduction

For the purposes of this paper, we assume that all graphs are finite and simple. Let G be a graph with order n and size m. As usual, we use $V(G)$ and $E(G)$ to denote the vertex set and edge set of G, respectively. For $v \in V(G)$, let $N_G(v)$ (or $N(v)$ for short) be the *neighbors* of v in G and let $d_v = |N(v)|$, the *degree* of v. The *cycle*, the *star* and the *path* of order n are denoted by C_n, S_n and P_n, respectively.

Let $G = (V(G), E(G))$ be a graph with vertex set $V(G) = \{1, 2, \ldots, n\}$. The *adjacency matrix* $\mathbf{A} = \mathbf{A}(G) = (a_{ij})_{n\times n}$ associated to G is defined as $a_{ij} = 1$ if and only if i is adjacent to j, and $a_{ij} = 0$ otherwise. The matrix

*Corresponding author.

$diag\{d_1, d_2, \ldots, d_n\}$, denoted by $\mathbf{D}(G)$, is referred as the degree diagonal matrix of G. Obviously, for each i, d_i is the i-th row (or column) sum of $\mathbf{A}$. $\mathbf{L}(G) := \mathbf{D}(G) - \mathbf{A}(G)$ and $\mathbf{Q}(G) := \mathbf{D}(G) + \mathbf{A}(G)$ are called the *Laplacian matrix* and the *signless Laplacian matrix* of G, respectively. For any given graph G, the degree diagonal matrix, $\mathbf{D}(G)$, is determined by its adjacency matrix, $\mathbf{A}(G)$, then both of $\mathbf{L}(G)$ and $\mathbf{Q}(G)$ are determined by $\mathbf{A}(G)$.

Given two disjoint graphs $G_1 = (V(G_1), E(G_1))$ and $G_2 = (V(G_2), E(G_2))$, the graph $G = (V(G), E(G))$ is called the *sum* of G_1 and G_2 if $V(G) = V(G_1) \cup V(G_2)$ and $E(G) = E(G_1) \cup E(G_2)$. We shall write simply $G_1 \cup G_2 \cup \cdots \cup G_k$ as kG_1 if $G_1 = G_2 = \cdots = G_k$. Usually, the subgraph kP_2 of the graph G is called a k-matching of G. A graph G is called H-free if G does not contain the graph H as its subgraph.

It is well known that the computation of the number of general substructures, such as kP_2 and kP_3, may be very difficult, and is known to be NP-Complete; see e.g., [12, 13]. Therefore it is interesting to use some graph invariants or topological indices to determine the number of substructures of a graph that have specific graph properties. From Lemma 8.1.2 in [5] (or Corollary 8.1.3 in [5]), the number of the triangles of any graph can be obtained from the trace of the 3-moment (or the third adjacency coefficient) of its adjacency matrix. Recently, Lei et al. [16] computed the numbers of various subgraphs of order 4 in several common lattice graphs. Lemma 2.1 in [1] tells us that the number of P_3 and of $2P_2$ substructures are determined by the degree sequence of such a graph. In 2009, Farrell and Guo (see [4]) found a formula that calculates the number of 3-matchings in graphs in terms of its degrees sequence and the number of triangles. Then Behmaram [1] established a formula for the number of 4-matchings in triangular-free graph in terms of its order, degrees sequence and the number of quadrangles. For more results on the number of other substructures such as 5-matchings and 6-matchings of some special classes of graphs, one can see [7, 14, 18, 19].

In this paper, we continue to investigate the number of substructures having few edges. Especially, we will deduce formulas on the number of all substructures having at most four edges in terms of the invariants and indices determined by its adjacency matrix $A(G)$. In this paper we assume the degree sequence, the degree diagonal matrix, the Laplacian matrix and the signless Laplacian matrix of a given graph are known, as the degree diagonal matrix is determined by its adjacency matrix. Therefore, our formulas include the degrees sequence, the k-moment of the adjacency matrix, adjacency coefficients, Laplacian coefficients and signless Laplacian coefficients of such a graph, all which will be defined in below.

The paper is organized as follows. In Section 2, we will introduce some preliminary results and some lemmas. Especially, we will introduce a matrix operation, from which we can obtain the number of substructures $S_4 \cup P_2$, the sum of the star S_4 and the path P_2. In addition, we will provide an algorithm for computing the number of $S_4 \cup P_2$ substructures contained in any given graph G in this section. In Section 3, we first study the number of substructures having at most three edges. Then we determine the number of all substructures containing exactly four edges. In Section 4 as applications, we provide an algebraic expression for the second Zagreb index of graphs and an expression for $||A^4||$ in terms of the degree sequence of such a graph.

2 Preliminary

For a square matrix A, we use A^k, $(A^k)_{ij}$, $tr(A)$ and $||A||$ to denote the k-moment of A, the (i,j)-element of A^k, the trace and the sum of all entries of A, respectively. Let G be a graph with vertex set $V = \{v_1, v_2, \cdots, v_n\}$ and $W = v_0 v_1 \cdots v_k$ (perhaps $v_i = v_j$ for $i \neq j$) be a *walk* joining vertices v_1 and v_k. The integer k is referred as the length of the walk W. We begin our investigation with the following well known result.

Lemma 1. ([6, Lemma 2.2.1], or [5, Lemma 8.1.2]) Let G be a graph with order n, vertex set $V = \{1, 2, \cdots, n\}$ and adjacency matrix $\mathbf{A}$. Then

$$(\mathbf{A}^k)_{ij}$$

denotes the number of walks with length k from the vertex i to the vertex j.

For convenience, denote by n_H the number of the substructures H

contained in G. Denote by P_4^* and C_3^* graphs obtained from P_4 and C_3 by adding a pendent edge to one of those 2-degree vertices, respectively. Applying lemma 1, some concise relationships between $tr(A^k)$ (or $||A^k||$) and the number of some substructures, when k is small, can be obtained as follows.

Corollary. *Let G be a graph with order n, size m and adjacency matrix $\mathbf{A}$. Then*

(i). $tr(\mathbf{A}^2) = 2m = 2n_{P_2}$;

(ii). $tr(\mathbf{A}^3) = 6n_{C_3}$;

(iii). $tr(\mathbf{A}^4) = 8n_{C_4} + 4n_{P_3} + 2n_{P_2}$;

(iv). $tr(\mathbf{A}^5) = 10n_{C_5} + 30n_{C_3} + 10n_{C_3^*}$;

(v). $||\mathbf{A}^3|| = 6n_{C_3} + 2n_{P_4} + 2n_{P_2} + 4n_{P_3}$;

(vi). $||\mathbf{A}^4|| = 8n_{C_4} + 2n_{P_2} + 8n_{P_3} + 2n_{P_5} + 4n_{P_4} + 4n_{C_3^*} + 18n_{C_3} + 6n_{S_4}$.

Proof. (i) is trivial; (ii) follows from Corollary 8.1.3 in [5] directly.

(iii). Note that $tr(\mathbf{A}^4) = \sum_{i=1}^{n}(\mathbf{A}^4)_{ii}$, then $tr(\mathbf{A}^4)$ denotes the sum of all number of walks with length 4 from i to i by lemma 1. In view of each walk with length 4, from i to i, may be formed by C_4, P_2 or P_3, thus we assume that

$$tr(\mathbf{A}^4) = xn_{C_4} + yn_{P_3} + zn_{P_2}.$$

Let $C_4 = ijkli$ be a given 4-cycle. Then there are exactly 8 walks with length 4 related to such a cycle, named as $ijkli$, $jklij$, $klijk$, $lijkl$, $ilkji$, $jilkj$, $kjilk$ and $lkjil$. Let $P_3 = ijk$ be a given 3-path. Then there are exactly 4 walks with length 4 related to such a 3-path, named as $ijkji$, $kjijk$, $jijkj$ and $jkjij$. Let $P_2 = ij$ be a given 2-path. Then there are exactly 2 walks, $ijiji$ and $jijij$, with length 4 related to such a 2-path. Thus $x = 8$, $y = 4$ and $z = 2$. Consequently, we have

$$tr(\mathbf{A}^4) = 8n_{C_4} + 4n_{P_3} + 2n_{P_2}.$$

(iv). By lemma 1, $tr(\mathbf{A}^5)$ denotes the sum of all number of walks with length 5 from i to i for $i = 1, 2, \ldots, n$. In view of each walk of length 5, from i to i, may be formed by C_5, C_3 or C_3^*, we assume that

$$tr(\mathbf{A}^5) = xn_{C_5} + yn_{C_3} + zn_{C_3^*}.$$

By the discussion similar to above, we have $x = 10$, $y = 30$ and $z = 10$. Thus

$$tr(\mathbf{A}^5) = 10n_{C_5} + 30n_{C_3} + 10n_{C_3^*}.$$

(v). Likewise, we have $\sum_{i \neq j}(\mathbf{A}^3)_{ij} = 2n_{P_4} + 2n_{P_2} + 4n_{P_3}$. Consequently, combining with (ii), we have

$$||\mathbf{A}^3|| = tr(\mathbf{A}^3) + \sum_{i\neq j}(\mathbf{A}^3)_{ij} = 6n_{C_3} + 2n_{P_4} + 2n_{P_2} + 4n_{P_3}.$$

(vi). Note that $||\mathbf{A}^4|| = tr(\mathbf{A}^4) + \sum_{i\neq j}(\mathbf{A}^4)_{ij}$ and by lemma 1 $(\mathbf{A}^4)_{ij}$ $(i \neq j)$ denotes the sum of the number of walks with length 4 from the vertex i to the vertex j with $i \neq j$, then we assume that

$$\sum_{i\neq j}(\mathbf{A}^4)_{ij} = xn_{P_5} + yn_{P_4} + zn_{C_3^*} + pn_{C_3} + qn_{S_4} + rn_{P_3},$$

in view of each walk of length 4, from i to j, $i \neq j$, is formed by one of the subgraphs P_5, P_4, C_3^*, C_3, S_4 and P_3. By a similar approach, we have $x = 2$, $y = 4$, $z = 4$, $p = 18$, $q = 6$ and $r = 4$. Consequently, we have

$$||\mathbf{A}^4|| - tr(\mathbf{A}^4) = 2n_{P_5} + 4n_{P_4} + 4n_{C_3^*} + 18n_{C_3} + 6n_{S_4} + 4n_{P_3}.$$

Therefore, the proof is complete. ■

Let G be a graph with order n, adjacency matrix $\mathbf{A}$ and degree diagonal matrix $\mathbf{D}$. The adjacency polynomial, the Laplacian polynomial and the signless Laplacian polynomial of G are defined as

$$\begin{aligned}
\phi(G;\lambda) &= det(\lambda\mathbf{I} - \mathbf{A}) = \textstyle\sum_{i=0}^{n} a_i\lambda^{n-i};\\
\varphi(G;\lambda) &= det(\lambda\mathbf{I} - \mathbf{L}) = det(\lambda\mathbf{I} - \mathbf{D} + \mathbf{A}) = \textstyle\sum_{i=0}^{n}(-1)^i l_i\lambda^{n-i};\\
\psi(G;\lambda) &= det(\lambda\mathbf{I} - \mathbf{Q}) = det(\lambda\mathbf{I} - \mathbf{D} - \mathbf{A}) = \textstyle\sum_{i=0}^{n}(-1)^i q_i\lambda^{n-i},
\end{aligned} \tag{2.1}$$

respectively. Hereafter, a_i, l_i and q_i are called the i-th adjacency coefficient, i-th Laplacian coefficient and i-th signless Laplacian coefficient of G, respectively.

A subgraph H of G is called an elementary subgraph if each component of H is either an edge or a cycle. For an elementary subgraph H, denote by $c(H)$ and $c_1(H)$ the number of components which are cycles and edges, respectively. The following result determines all adjacency coefficients of graphs in terms of elementary subgraphs.

Lemma 2. [6, Theorem 3.10] Let G be a graph with order n and adjacency polynomial $\phi(G;\lambda) = \sum_{i=0}^{n}(-1)^i a_i\lambda^{n-i}$, defined as Eq.(2.1). Then

$$a_i = \sum(-1)^{c_1(H)+c(H)}2^{c(H)},$$

where the summation is over all the elementary subgraphs H of G having i vertices, $i = 1, 2, \ldots, n$.

Note that each elementary subgraph of order 5 is either C_5 or $C_3 \cup P_2$, then applying Lemma 2, we have the following result immediately.

Corollary. *Let G be a graph with the fifth adjacency coefficient a_5. Then*

$$a_5 = 2n_{C_3 \cup P_2} - 2n_{C_5}. \tag{2.2}$$

Suppose that F is a spanning forest of the graph G with k components, $T_i (i = 1, 2, \ldots, k)$. For each i, denote by n_i the number of vertices contained in T_i. Let $\gamma(F) = \prod_{i=1}^{k} n_i$. Due to Kelmans and Chelnokov, the Laplacian coefficient l_{n-k} is given as follows.

Lemma 3. [15, p203] *Let G be a graph with order n and Laplacian polynomial $\varphi(G;\lambda) = \sum_{i=0}^{n} (-1)^i l_i \lambda^{n-i}$, defined as Eq.(2.1). Then*

$$l_{n-k} = \sum_{F \in \mathfrak{F}_k} \gamma(F),$$

where $\mathfrak{F}_k$ denotes the set of all spanning forests of G with exactly k components.

As a consequence of Lemma 3, we have

Corollary. *Let G be a graph with the fourth Laplacian coefficient l_4. Then*

$$l_4 = 5n_{S_5} + 8n_{S_4 \cup P_2} + 9n_{2P_3} + 12n_{P_3 \cup 2P_2} + 8n_{P_4 \cup P_2} + 5n_{P_5} + 5n_{P_4^*} + 16n_{4P_2}. \tag{2.3}$$

Proof. Since each spanning forest having $n - 4$ components has exactly four edges, by direct verification we can see the components make up one of the following graphs

$$\{S_5, S_4 \cup P_2, 2P_3, P_3 \cup 2P_2, P_4 \cup P_2, P_5, P_4^*, 4P_2\}$$

together with isolated vertices appropriately. Then applying Lemma 3, the result follows. ■

Let G be a graph. A spanning subgraph of G whose each connected component is either a tree or an odd unicyclic graph is called a *TU*-subgraph of G. Suppose that a *TU*-subgraph H of G contains c odd unicyclic graphs and s trees named as $T_1, T_2, \ldots, T_s$. Then the weight of H, denoted by $W(H)$, is defined by $W(H) = 4^c \prod_{i=1}^{s} n_i$, in which n_i is the number of the vertices of T_i. Due to Cvetković, Rowlinson and Simić, the signless Laplacian coefficient q_i is deduced in terms of the weight of *TU*-subgraphs of G as follows.

Lemma 4. [3, Theorem 4.4] *Let G be a connected graph with order n and signless Laplacian polynomial $\psi(G;\lambda) = \sum_{i=0}^{n}(-1)^i q_i \lambda^{n-i}$, defined as Eq.(2.1). Then for each $i (i = 0, 1, 2, \ldots, n)$*

$$q_i = \sum_{H_i} W(H_i),$$

where the summation runs over all TU-subgraphs H_i of G having i edges.

As a consequence of Lemma 4, we have

Corollary. *Let G be a simple graph with the fourth Laplacian coefficient l_4 and the fourth signless Laplacian coefficient q_4. Then*

$$q_4 - l_4 = 4n_{C_3^*} + 8n_{C_3 \cup P_2}. \tag{2.4}$$

Proof. According to Lemma 4, we know

$$q_4 = \sum_{H_4} W(H_4),$$

where the summation runs over all TU-subgraphs of G having 4 edges. By direct verification, each TU-subgraph of G with 4 edges is one of the graphs

$$\{S_5, S_4 \cup P_2, 2P_3, P_3 \cup 2P_2, P_4 \cup P_2, P_5, P_4^*, 4P_2, C_3^*, C_3 \cup P_2\},$$

in which there has exactly two graphs C_3^* and $C_3 \cup P_2$ are non-bipartite. Consequently, combining with Corollary 2, we have

$$q_4 = l_4 + 4n_{C_3^*} + 8n_{C_3 \cup P_2}.$$

Thus the proof is complete. ∎

In addition, we need to introduce a matrix operation, from which we can count the number of the subgraphs $S_4 \cup P_2$. Let G be a graph with vertex set $V(G) = \{1, 2, \ldots, n\}$ and adjacency matrix $\mathbf{A}$. Suppose that, for each i, the vertex i corresponds to the i-th row, as well as the i-th column, of $\mathbf{A}$. Let i be a vertex with degree at least 3 and let the vertices j, k and l be the neighbors of i. Denote by $G\backslash\{i; j, k, l\}$ the graph obtained from G by deleting vertices $\{i, j, k, l\}$ and all edges incident to them. We should point out that the subgraph $G\backslash\{i; j, k, l\}$ is different slightly from the subgraph $G\backslash\{i, j, k, l\}$. In $G\backslash\{i, j, k, l\}$, the degree of the vertex i may be less than 3, that is, $G\backslash\{i, j, k, l\} = G\backslash\{j, i, k, l\}$. However, from the definition above, $G\backslash\{i; j, k, l\}$ and $G\backslash\{j; i, k, l\}$ are two distinct graphic representations even if $d_i \geq 3$, $\{j, k, l\} \subseteq N_i$ and $d_j \geq 3$, $\{i, k, l\} \subseteq N_j$, although the resultant graphs $G\backslash\{i; j, k, l\}$ and $G\backslash\{j; i, k, l\}$ have no distinction. Correspondingly, the adjacency matrix associated to $G\backslash\{i; j, k, l\}$ is denoted by $\mathbf{A}\{i; j, k, l\}$.

For simplicity, let

$$\alpha(G) := \frac{1}{2} \sum_{d_i \geq 3, \{j,k,l\} \subseteq N_i} ||A\{i; j, k, l\}||,$$

where the summation runs over all vertices i whose degree is at least 3 and $\{j, k, l\} \subseteq N_i$. Then we have

Theorem 1. *Let G be a graph with vertex set $V(G) = \{1, 2, \ldots, n\}$ and adjacency matrix $\mathbf{A}$. Then*

$$n_{S_4 \cup P_2} = \alpha(G).$$

Proof. Obviously, the expression $\frac{1}{2}||A\{i; j, k, l\}||$ denotes the number of edges contained in the subgraph $G \backslash \{i; j, k, l\}$. By the definition above, $G \backslash \{i; j, k, l\}$ is the graph by removing the subgraph S_4, the star with central vertex i and pendent vertices j, k, l, and all edges incident to them. So $\frac{1}{2}||A\{i; j, k, l\}||$ is the number of $S_4 \cup P_2$, where the star S_4 is fixed whose central vertex is i and pendent vertices are vertices j, k, l. On the other hand, each star S_4 corresponds a vertex with degree at least 3 together with its three neighbors. Consequently, the proof is complete. ■

In the final of this section, we give an algorithm for computing the graph invariant $\alpha(G)$ as follows:

Algorithm.

Input: The adjacency matrix $\mathbf{A} = \mathbf{A}(G) = (a_{ij})_{n \times n}$ with respect to a given graph G of order n.

Output: $2\alpha(G)$.

Step1. Constructing vectors $R = (r_1, r_2, \ldots, r_k)$ and $C^i = (c_1^i, c_2^i, \ldots, c_{s_i}^i)$ $(i = 1, \ldots, k)$.

(1.1) Initially $k = 0$ and $s_i = 0, i = 1, \ldots, k$.

(1.2) Traverse each row of the matrix $\mathbf{A}$ and obtain the degree d_i of the vertex i for $i = 1, 2, \ldots, n$.

(1.3) $k \leftarrow k + 1$ and $r_k \leftarrow i$ if $d_i \geq 3$; otherwise, return nothing and continue loop. When $i = n$, stop.

(1.4) $s_i \leftarrow s_i + 1$ and $c_s^i \leftarrow j$ if $a_{r_i,j} = 1$; otherwise, return nothing and continue loop. When $i = k$ and $j = n$, stop.

Step2.

(2.1) Initially $a = 0$.

(2.2) For an arbitrary triple array (p,q,h) of the vector C^i, $a = a + \mathbf{A}\{r_i; p,q,h\}$, where p,q and h are three distinct entries of C^i;

(2.3) When all the triple arrays of C^i, $(i = 1,\ldots,k)$, are traversed, the loop ends and the whole algorithm is complete.

Remark. In this algorithm, the traversal of step 1 does not involve nesting, the algorithm complexity of this step is $O(k)$, and the traversal of step 3 is nested in step 2, so the algorithm complexity of these two steps is $O(k|C|^3)$. Therefore, the final algorithm complexity is $O(k|C|^3)$.

3 The number of all substructures containing at most four edges

Let G be a graph with degree sequence $(d_1,d_2,\ldots,d_n)$ and size m. We in the following sometimes set

$$\beta_k(G) := \sum_{i=1}^{n} \binom{d_i}{k} \qquad (k = 2,\ldots,n-1)$$

for simplicity.

3.1 The number of all substructures containing at most three edges

From the definition of the adjacency matrix, we have

$$n_{P_2} = m = \frac{1}{2} tr(A^2). \tag{3.1}$$

We can see there are exactly two types of graphs, P_3 and $2P_2$, having exactly two edges, then we have

$$n_{P_3} = \sum_{i=1}^{n} \binom{d_i}{2} = \beta_2(G). \tag{3.2}$$

and

$$n_{2P_2} = \binom{m}{2} - \beta_2(G). \tag{3.3}$$

Furthermore, one can see that there are exactly five graphs containing three edges, named as

$$\{S_4, C_3, P_4, P_3 \cup P_2, 3P_2\}.$$

The following proposition tells us the number of those substructures.

Proposition 2. *Let G be a graph with order n and size m. Then*

(i). $n_{C_3} = \frac{1}{6}tr(A^3)$;

(ii). $n_{S_4} = \beta_3(G)$;

(iii). $n_{P_4} = \frac{1}{2}||A^3|| - \frac{1}{2}tr(A^3) - 2\beta_2(G) - m$;

(iv). $n_{P_3 \cup P_2} = -||A^3|| + \frac{1}{2}tr(A^3) - 3\beta_3(G) + (m+2)\beta_2(G) + 2m$;

(v). $n_{3P_2} = \binom{m}{3} + \frac{1}{2}||A^3|| - \frac{1}{6}tr(A^3) + 2\beta_3(G) - m\beta_2(G) - m$.

Proof. (i). The result follows from (ii) of Corollary 2, or Corollary 8.1.3 in [5] directly.

(ii). The result follows from the fact that all edges of the star S_4 have exactly one central vertex and the number of the stars S_4 formed by the fixed vertex i is $\binom{d_i}{3}$.

(iii). From Corollary 2 (v), we have

$$n_{P_4} = \frac{1}{2}||A^3|| - 3n_{C_3} - n_{P_2} - 2n_{P_3}.$$

Then, combining Eq.s (3.1), (3.2) and (i), we have

$$n_{P_4} = \frac{1}{2}||A^3|| - \frac{1}{2}tr(A^3) - m - 2\beta_2(G).$$

(iv). The expression

$$(m-2)\sum_{i=1}^{n}\binom{d_i}{2}$$

has a combinatorial interpretation as follows: we first select two edges having a common vertex, a path P_3, and then select an arbitrary edge from the remaining $m-2$ edges. In view of the resultant graph may be one of the graphs $\{C_3, P_4, S_4, P_3 \cup P_2\}$, we assume that

$$(m-2)\sum_{i=1}^{n}\binom{d_i}{2} = xn_{P_4} + yn_{C_3} + zn_{S_4} + wn_{P_3 \cup P_2}.$$

Let $ie_1je_2ke_3l$ be a given path P_4. Then the path P_4 has two different structures: first select the path ie_1je_2k and then select the edge e_3, or first select the path je_2ke_3l and then select the edge e_1. Consequently, $x = 2$. Similarly, we have $y = 3, z = 3$ and $w = 1$. Therefore, we have

$$(m-2)\beta_2(G) = 2n_{P_4} + 3n_{C_3} + 3n_{S_4} + n_{P_3 \cup P_2},$$

which implies that

$$n_{P_3\cup P_2} = (m-2)\beta_2(G) - 2n_{P_4} - 3n_{C_3} - 3n_{S_4}.$$

Then, combining with (i), (ii) and (iii), the result follows. (v). Applying Theorem 2.3 in [1] and Corollary 2, we have

$$n_{3P_2} = \binom{m}{3} - n_{P_3\cup P_2} - n_{P_4} - n_{S_4} - n_{C_3}.$$

Then, combining with (i), (ii), (iii) and (iv), the result follows. ■

3.2 The number of all substructures having exactly four edges

By a direct verification, we find that there are exactly eleven graphs containing four edges, named as

$$\{C_4, C_3^*, C_3\cup P_2, S_5, S_4\cup P_2, P_5, P_4^*, 2P_3, P_3\cup 2P_2, P_4\cup P_2, 4P_2\}.$$

From Theorem 1, we have $n_{S_4\cup P_2} = \alpha(G)$, where the invariant $\alpha(G)$ is defined in section 2. Moreover, we have

Theorem 3. *Let G be a graph with order n and size m. Then*

(i). $n_{C_3^*} = \frac{1}{5}tr(A^5) - tr(A^3) + \frac{1}{4}(l_4 - q_4) + a_5$;

(ii). $n_{C_3\cup P_2} = -\frac{1}{10}tr(A^5) + \frac{1}{2}tr(A^3) + \frac{1}{4}(q_4 - l_4) - \frac{1}{2}a_5$;

(iii). $n_{S_5} = \beta_4(G)$;

(iv). $n_{C_4} = \frac{1}{8}tr(A^4) - \frac{1}{2}\beta_2(G) - \frac{1}{4}m$;

(v). $n_{P_5} = \frac{1}{2}||A^4|| - ||A^3|| - \frac{2}{5}tr(A^5) - \frac{1}{2}tr(A^4) + \frac{3}{2}tr(A^3) - \frac{1}{2}(l_4 - q_4)$
$-2a_5 - 3\beta_3(G) + 2\beta_2(G) + 2m.$

Proof. (i)-(ii). From Eq.s (2.2), (2.4) and Corollary 2 (iv), we have

$$\begin{cases} q_4 - l_4 = 4n_{C_3^*} + 8n_{C_3\cup P_2} \\ a_5 = 2n_{C_3\cup P_2} - 2n_{C_5} \\ tr(A^5) = 10n_{C_5} + 30n_{C_3} + 10n_{C_3^*}. \end{cases}$$

Then $n_{C_3^*}$ and $n_{C_3\cup P_2}$ can be solved directly, we omit the details.

(iii). The formula is established in a routine manner to Proposition 2 (ii), we also omit the detail.

(iv). From corollary 2 (iii), we have

$$n_{C_4} = \frac{1}{8}tr(A^4) - \frac{1}{4}n_{P_2} - \frac{1}{2}n_{P_3}.$$

Combining with (3.1) and (3.2), we have

$$n_{C_4} = \frac{1}{8}tr(A^4) - \frac{1}{4}m - \frac{1}{2}\beta_2(G).$$

(v). From corollary 2 (vi), we have

$$n_{P_5} = \frac{1}{2}||A^4|| - 4n_{C_4} - n_{P_2} - 4n_{P_3} - 2n_{P_4} - 2n_{C_3^*} - 9n_{C_3} - 3n_{S_4}$$

Combining with (i), (iv), proposition 2 (i) and (iii), we have

$$n_{P_5} = \frac{1}{2}||A^4|| - ||A^3|| - \frac{2}{5}tr(A^5) - \frac{1}{2}tr(A^4) + \frac{3}{2}tr(A^3) - \frac{1}{2}(l_4 - q_4)$$

$$-2a_5 - 3\beta_3(G) + 2\beta_2(G) + 2m.$$

Therefore, the proof is complete. ■

Up to now, the formulas on the number of five graphs $\{C_3^*, C_3 \cup P_2, S_5, C_4, P_5, S_4 \cup P_2\}$ are given. To determine the number of the remaining five subgraphs having 4 edges, we need the following equations.

Lemma 5. *Let G be a graph having n vertices and m edges. Then the following equations hold*

$$\sum_{i=1}^{n}(m - d_i)\binom{d_i}{3} = n_{P_4^*} + n_{S_4 \cup P_2} + n_{C_3^*};$$

$$(m-3)n_{P_4} = n_{P_4 \cup P_2} + 2n_{P_5} + 2n_{P_4^*} + 4n_{C_4} + 2n_{C_3^*};$$

$$\sum_{1 \le i < j \le n} \binom{d_i}{2}\binom{d_j}{2} = n_{2P_3} + n_{P_5} + 2n_{C_4} + n_{P_4} + 2n_{C_3^*} + n_{P_4^*} + 3n_{C_3};$$

$$\sum_{i=1}^{n}\binom{m-d_i}{2}\binom{d_i}{2} = n_{P_3 \cup 2P_2} + 3n_{P_5} + 2n_{P_4 \cup P_2} + 2n_{2P_3} + 4n_{C_4} + 2n_{C_3^*} + n_{P_4^*} + 3n_{C_3 \cup P_2};$$

$$(m-3)n_{3P_2} = 4n_{4P_2} + 2n_{P_3 \cup 2P_2} + n_{P_4 \cup P_2}.$$

Proof. the proof of all those equations are analogous, we only give the proof of the first one and omit all others. A combinatorial interpretation on the expression

$$\sum_{i=1}^{n}(m - d_i)\binom{d_i}{3}$$

is: we first select three edges incident to the vertex i and then select an arbitrary edge from $E(G)\backslash N_i$. Then using the method similar to Proposition 2 (iv) we have

$$\sum_{i=1}^{n}(m - d_i)\binom{d_i}{3} = n_{P_4^*} + n_{S_4 \cup P_2} + n_{C_3^*}.$$

Thus the first equation follows. ■

Consequently, we have

Theorem 4. *Let G be a graph with order n, size m and adjacency matrix A. Then we have*

$$\begin{array}{lcl}
(i).\ n_{P_4^*} & = & \sum_{i=1}^n (m-d_i)\binom{d_i}{3} - \frac{1}{5}tr(A^5) + tr(A^3) - \frac{l_4-q_4}{4} - a_5 - \alpha(G);\\
(ii).\ n_{P_4\cup P_2} & = & -||A^4|| + \frac{1}{2}(m+1)||A^3|| + \frac{4}{5}tr(A^5) + \frac{1}{2}tr(A^4) + l_4 - q_4\\
 & - & \frac{1}{2}(m+3)tr(A^3) + 4a_5 - m^2 - 2\sum_{i=1}^n (m-d_i)\binom{d_i}{3}\\
 & + & 2\alpha(G) + 6\beta_3(G) - (2m-4)\beta_2(G);\\
(iii).\ n_{2P_3} & = & -\frac{1}{2}||A^4|| + \frac{1}{2}||A^3|| + \frac{1}{5}tr(A^5) + \frac{1}{4}tr(A^4) - \frac{1}{2}tr(A^3) + a_5\\
 & + & \frac{1}{4}(l_4 - q_4) - \frac{1}{2}m + \sum_{1\le i<j\le n}\binom{d_i}{2}\binom{d_j}{2}\\
 & + & \alpha(G) + 3\beta_3(G) + \beta_2(G) - \sum_{i=1}^n (m-d_i)\binom{d_i}{3};\\
(iv).\ n_{P_3\cup 2P_2} & = & \frac{3}{2}||A^4|| - (m-1)||A^3|| - \frac{7}{10}tr(A^5) - \frac{1}{2}tr(A^4)\\
 & - & \frac{1}{2}(l_4-q_4) + \sum_{i=1}^n \binom{m-d_i}{2}\binom{d_i}{2} + 5\sum_{i=1}^n (m-d_i)\binom{d_i}{3}\\
 & - & 2\sum_{1\le i<j\le n}\binom{d_i}{2}\binom{d_j}{2} - 5\alpha(G) + (4m-14)\beta_2(G)\\
 & + & (m-1)tr(A^3) + 2m^2 - 4m - \frac{7}{2}a_5 - 9\beta_3(G);\\
(v).\ n_{4P_2} & = & -\frac{1}{2}||A^4|| + \frac{1}{2}(m-2)||A^3|| + \frac{3}{20}tr(A^5) + \frac{1}{8}tr(A^4) + \frac{3}{4}a_5\\
 & + & \frac{1}{12}(12-5m)tr(A^3) - \frac{1}{2}\sum_{i=1}^n \binom{m-d_i}{2}\binom{d_i}{2} + 2\alpha(G)\\
 & + & \sum_{1\le i<j\le n}\binom{d_i}{2}\binom{d_j}{2} - 2\sum_{i=1}^n (m-d_i)\binom{d_i}{3} + \frac{1}{2}(m+3)\beta_3(G)\\
 & - & \frac{1}{4}(m^2+3m-24)\beta_2(G) + \frac{1}{4}(m-3)\binom{m}{3} - m^2 + \frac{11m}{4}.
\end{array}$$

Proof. From Lemma 5, we have

$$n_{P_4^*} = \sum_{i=1}^n (m-d_i)\binom{d_i}{3} - n_{S_4\cup P_2} - n_{C_3^*};$$

$$n_{P_4\cup P_2} + 2n_{P_4^*} = (m-3)n_{P_4} - 2n_{P_5} - 4n_{C_4} - 2n_{C_3^*};$$

$$n_{2P_3} + n_{P_4^*} = \sum_{1\le i<j\le n}\binom{d_i}{2}\binom{d_j}{2} - n_{P_5} - 2n_{C_4} - n_{P_4} - 2n_{C_3^*} - 3n_{C_3};$$

$$n_{P_3\cup 2P_2} + 2n_{P_4\cup P_2} + 2n_{2P_3} + n_{P_4^*} = \sum_{i=1}^n \binom{m-d_i}{2}\binom{d_i}{2} - 3n_{P_5} - 4n_{C_4} - 2n_{C_3^*} - 3n_{C_3\cup P_2};$$

$$4n_{4P_2} + 2n_{P_3\cup 2P_2} + n_{P_4\cup P_2} = (m-3)n_{3P_2}.$$

Combining with Proposition 2 and Theorem 3, the number of each graph of $\{P_4^*, P_4\cup P_2, 2P_3, P_3\cup 2P_2, 4P_2\}$ can be obtained one by one. Thus the proof is complete. ■

4 Another expression on $||\mathbf{A}^3||$ and $||\mathbf{A}^4||$

Corollary 2 has deduced an expression on $||\mathbf{A}^3||$ and $||\mathbf{A}^4||$ in terms of the number of some substructures with small order, respectively.

Let G be a graph with n vertices and degree sequence $(d_1, d_2, \ldots, d_n)$.

The first Zagreb index and the second Zagreb index are defined as

$$M_1 = M_1(G) = \sum_{i=1}^{n} d_i^2 \quad and \quad M_2 = M_2(G) = \sum_{ij \in E(G)} d_i d_j.$$

The two Zagreb indices are the oldest vertex degree based molecular structure descriptors, invented in the 1970s [10,11]. Details of their theory can be found in the recent reviews [2,8].

One can easily verify that

$$||\mathbf{A}^2|| = 2n_{P_2} + n_{P_3} = \sum_{i=1}^{n} d_i^2 = M_1(G).$$

Interestingly, as by-product, we in the following show that $||\mathbf{A}^3|| = 2M_2(G)$, which provides an algebraic expression for the second Zagreb index of graphs. Moreover, we establish an expressions on $||\mathbf{A}^4||$ in terms of the degree sequence of such a graph.

Proposition 5. *Let G be a graph with degree sequence $(d_1, d_2, \ldots, d_n)$ and adjacency matrix A. Then*

1). $||\mathbf{A}^3|| = 2M_2(G)$;

2). $||\mathbf{A}^4|| = \sum_{u \in V(G)} (2 \sum_{i,j \in N(u)} d_i d_j + d_u^3)$.

Proof. 1). Let $e = uv$ be an arbitrary edge of G with $u_1 \in N(u) \backslash \{v\}$ and $v_1 \in N(v) \backslash \{u\}$. Then we find that if $u_1 = v_1$, say $u_1 = v_1 =: w$, then $G[u, v, w]$ forms a triangle of G; and if $u_1 \neq v_1$, then the edge set $\{u_1u, uv, vv_1\}$ forms a path P_4. Therefore, path P_4 has exactly one central edge and each edge of a triangle, C_3 ,can be considered as its central edge, we have

$$\sum_{uv \in E(G)} (d_u - 1)(d_v - 1) = 3n_{C_3} + n_{P_4}.$$

On the other hand, we find that

$$\sum_{uv \in E(G)} (d_u - 1)(d_v - 1) = \sum_{uv \in E(G)} (d_u d_v - d_u - d_v + 1) = M_2(G) - M_1(G) + n_{p_2}.$$

Consequently, the result follows applying Corollary 2(v).

2). Let $i, j \in N(u)$ be two neighbors, not necessarily distinct, of u with $i^* \in N(i)$ and $j^* \in N(j)$. In view of $||\mathbf{A}^4||$ equals the number of all walks having length 4 from i^* to j^* with $i^*, j^* \in V(G)$, then $||\mathbf{A}^4||$ equals the number of walks i^*iujj^*, where u runs over all vertices of G, i^* runs over all neighbors of i and j^* runs over all neighbors of i. Furthermore, we find that if $i \neq j$, then the number of those walks is

$$2 \sum_{u \in V(G)} \sum_{i,j \in N(u)} d_i d_j,$$

and if $i = j$, then the number of those walks is

$$\sum_{u \in V(G)} d_u^3.$$

Consequently, the result follows. ■

5 Concluding remarks

For any graph G with adjacency matrix $\mathbf{A}(G)$, we established formulas on the number of all substructures having at most four edges in terms of $\mathbf{A}(G)$ together with those graph invariants determined by $\mathbf{A}(G)$, including the degree sequence, k-moment of $\mathbf{A}(G)$, adjacency coefficients, Laplacian coefficients and signless Laplacian coefficients. Especially, we introduce a new matrix operation and the invariant $\alpha(G)$; see Section 2, avoiding all known equations are dependent.

***Acknowledgment*:** The paper is Supported by Zhejiang Provincial Natural Science Foundation of China (LY20A010005).

References

[1] A.Behmaram, On the number of 4-matchings in graphs, *MATCH Commun. Math. Comput. Chem.* **62** (2009) 381–388.

[2] B. Borovićanin, B. Furtula, On extremal Zagreb indices of trees with given domination number, *Appl. Math. Comput.* **279** (2016) 208–218.

[3] D. Cvetković, P. Rowlinson, S. Simić, Signless Laplacians of finite graphs, *Lin. Algebra Appl.* **423** (2007) 155–171.

[4] E. Farrell, J. Guo, On matching coefficients, *Discr. Math.* **89** (1991) 203–210.

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

[6] C. Godsil, *Algebraic Combinatorics*, Chapmam & Hall, New York, 1993.

[7] S. C. Gong, L. P. Zhang, S. W. Sun, On bipartite graphs having minimum fourth adjacency coefficient, *Graphs Comb.* **38** (2022) #60.

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

[9] I. Gutman, On the origin of two degree-based topological indices, *Bull. Acad. Serbe Sci. Arts (Cl. Sci. Math. Natur.)* **146** (2014) 39–52.

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

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

[12] P. Hell, D. Kirkpatrick, On the complexity of general graph factor problems, *SIAM J. Comput.* **12** (1983) 601–609.

[13] M. Jerrum, Two-dimensional monomer-dimer systems are computationally intractable, *J. Stat. Phys.* **48** (1987) 121–134.

[14] Y. Jiang, A. Tang, R. Hoffmann, Evaluation of moments and their application to Hückel molecular orbital theory, *Theor. Chim. Acta* **65** (1984) 255–265.

[15] A. Kel'mans, V. Chelnokov, A certain polynomial of a graph and graphs with an extremal number of trees. *J. Comb. Theory Ser. B* **16** (1974) 197–214.

[16] H. Lei, T. Li, Y. Ma, H. Wang, Analyzing lattice networks through substructures, *Appl. Math. Comput.* **329** (2018) 297–314.

[17] C. Oliveira, N. de Abreu, S. Jurkiewicz, The characteristic polynomial of the Laplacian of graphs in (a; b)-linear classes, *Lin. Algebra Appl.* **356** (2002) 113–121.

[18] R. Vesalian, F. Asgari, On the number of 5-matchings in graphs, *MATCH Commun. Math. Comput. Chem.* **69** (2013) 33–46.

[19] R. Vesalian, R. Namazi, F. Asgari, On the number of 6-matchings in graphs, *MATCH Commun. Math. Comput. Chem.* **73** (2015) 239–265.

[20] G. F. Wang, A proof of Boeschi's conjecture, *Networks* **24** (1994) 277–284.