MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **81** (2019) 133-148

ISSN 0340 - 6253

Maximum Wiener Indices of Unicyclic Graphs of Given Matching Number

Stijn Cambie *

Department of Mathematics, Radboud University Nijmegen, Postbus 9010, 6500 GL Nijmegen, The Netherlands
S.Cambie@math.ru.nl

(Received December 19, 2017)

Abstract

In this article, we determine the maximum Wiener indices of unicyclic graphs with given number of vertices and matching number. We also characterize the extremal graphs. This solves an open problem of Du and Zhou [2].

1 Introduction

Let G be a simple connected graph. We denote its vertex set by $V(G)$ and its edge set by $E(G)$. For two graphs G and H, $G+H$ denotes the graph obtained by adding all possible edges between the vertices of G and H. For an integer p, pG is the disjoint union of p copies of G. The matching number of a graph G is the size of a maximum independent edge subset of G, we will denote it by $m(G)$ or m. We will denote by $\mathbb{T}(n,m)$ and $\mathbb{U}(n,m)$ the set of all trees and unicyclic graphs respectively with n vertices and matching number m. Let $d(u,v)$ denote the distance between vertices u and v in a graph G. The Wiener index of a graph G equals the sum of distances between all unordered pairs of vertices, i.e.

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

*This work has been supported by a Vidi Grant of the Netherlands Organization for Scientific Research (NWO), grant number 639.032.614

The Wiener index, introduced in 1947 by Harry Wiener, is the oldest and one of the most important topological indices in chemical graph theory. This index can predict some chemical properties of molecules, e.g. the boiling point of alkanes, the density, the critical point and surface tension. It was used by chemists decades before it attracted the attention of mathematicians. An overview of the mathematical results, conjectures and problems with respect to the Wiener index can be found in [3]. In chemical graph theory, an important goal is to bound some important graph invariant like the Wiener index using some structural parameters. In this paper we consider the case of the matching number and we solve the remaining open case, Problem 11.4 of [3].

The minimum and maximum value for the Wiener index of a connected graph of order n with matching number m was determined by Dankelmann [1].

Theorem 1.1 (Dankelmann [1]) *Let G be a connected graph of order n and matching number m.*

- *If $m = \lfloor \frac{n}{2} \rfloor$, then $W(G) \geq \binom{n}{2}$. Equality holds iff G is the complete graph K_n.*
- *If $1 \leq m < \lfloor \frac{n}{2} \rfloor$, then $W(G) \geq 2\binom{n}{2} - mn + \binom{m}{2}$. Equality holds iff G is isomorphic to the graph obtained by $K_m + (n-m)K_1$.*
- *If n is even, then $W(G) \leq \binom{2m}{3} + \binom{2m}{2}(n-2m+1) + 2m\frac{n-2m+2}{2}\frac{n-2m}{2} + \frac{1}{2}(n-2m)^2$. Equality holds iff the graph is $A_{n,m}$, a path with $2m-1$ vertices, with one leaf of the path connected to $\frac{n-2m+2}{2}$ vertices and the other leaf with $\frac{n-2m}{2}$ vertices.*
- *If n is odd, then $W(G) \leq \binom{2m}{3} + \binom{2m}{2}(n-2m+1) + 2m(\frac{n-2m+1}{2})^2 + 4\binom{\frac{n-2m+1}{2}}{2}$. Equality holds iff the graph is $A_{n,m}$, a path with $2m-1$ vertices, with both leafs of the path connected to $\frac{n-2m+1}{2}$ different vertices, equivalently a path with $2m-3$ vertices with its ends concatenated to the centers of two stars of order $\frac{n-2m+3}{2}$.*

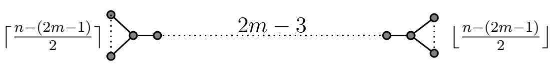

Figure 1. Extremal graph $A_{n,m}$

The graphs that attain the maximal Wiener indices are trees, while the graphs that attain the minimal Wiener indices contain many cycles. Hence two natural questions arise. What are the minimal Wiener indices for trees? What are the maximal Wiener indices for the graphs that are not trees? In the latter case, the graphs will be unicyclic. For graphs that contain multiple cycles, we can take a maximal matching, and then deleting an edge not contained in the matching such that the graph is still connected will increase the Wiener index.

Du and Zhou [2] solved the first question and also determined the minimal matching number for unicyclic graphs. Their results are stated in the following two theorems.

Theorem 1.2 (Du and Zhou [2]) *Let $G \in \mathbb{T}(n,m)$, where $2 \le m \le \lfloor \frac{n}{2} \rfloor$. Then $W(G) \ge n^2 + (m-3)n - 3m + 4$. Equality holds iff G is a star of order $n-m+1$ with an additional vertex connected to each of $m-1$ leaves.*

Theorem 1.3 (Du and Zhou [2]) *Let $G \in \mathbb{U}(n,m)$, where $2 \le m \le \lfloor \frac{n}{2} \rfloor$. If $(n,m) = (6,3)$, then $W(G) \ge 26$, with equality iff G is C_5 with a pendent vertex attached to it. In the other cases $W(G) \ge n^2 + (m-4)n - 3m + 6$. Equality holds for a star of order $n-m$ with a triangle attached to its center and an additional vertex connected to each of $m-2$ leaves, as well as for C_4, C_5, the graph C_5 with two pendent vertices attached to it, and C_5 with three pendent vertices attached to three consecutive vertices of the cycle.*

In this paper, we will prove the remaining question, which was left as an open problem in [2]. The result is summarized in the following theorem.

Theorem 1.4 *Let $G \in \mathbb{U}(n,m)$, where $2 \le m \le \lfloor \frac{n}{2} \rfloor$.*

- *If $n \le 2m+2$, then $W(G) \le 2 - \frac{8}{3}m^3 + 2m^2 + \frac{5}{3}m + 2m^2 n - 3mn - 2n + n^2$.*
- *If $n \ge 2m+3$ and n is odd, then $W(G) \le \frac{9}{2} - n - \frac{2}{3}m^3 + \frac{1}{2}n^2 - 2nm + 2m^2 + \frac{1}{2}mn^2 - \frac{11}{6}m$.*
- *If $n \ge 2m+4$ and n is even, then $W(G) \le 6 - n - \frac{2}{3}m^3 + \frac{1}{2}n^2 - 2nm + 2m^2 + \frac{1}{2}mn^2 - \frac{7}{3}m$.*

Moreover, these bounds are sharp.

The extremal graphs are members of two families of graphs, which are shown in Figure 2. We describe them explicitly in Theorem 7.1.

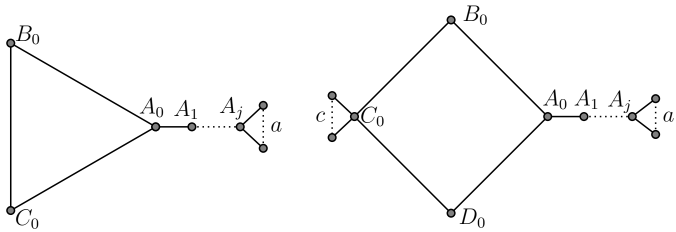

Figure 2. The graphs $G^3_{a,j}$ and $G^4_{a,c,j}$

So in this paper, we determine the maximum Wiener indices of graphs in $\mathbb{U}(n,m)$, where $m \geq 2$ and $n \geq 2m$ and characterize the extremal graphs. Note that the only nonempty remaining case is $(n,m) = (3,1)$, but then C_3 is the only unicyclic graph with these parameters. We will use the notation $W(\mathbb{U}(n,m)) = \max\{W(G) \mid G \in \mathbb{U}(n,m)\}$.

Our proof for the characterization of $W(\mathbb{U}(n,m))$ will actually first determine the characteristics of the extremal graphs and only after that, we calculate the sharp upper bound. It contains five main parts, which can be summarized as follows.

1. We prove that for fixed n, $W(\mathbb{U}(n,m))$ increases when m increases, as long as $m \leq \lfloor \frac{n}{2} \rfloor$. Intuitively, this is because a larger matching number implies that it is possible there are longer paths in the graph. A path is the graph with the largest Wiener index for a fixed order. The proof for this property is given in Section 2.

2. Second, we prove that the cycle in the extremal graph in $\mathbb{U}(n,m)$ for any n and m must be small. More precisely this cycle must be a triangle C_3 or quadrangle C_4. The possible constructions for this are given in Section 3. We need two possible constructions to ensure that we do not increase the matching number.

3. A unicyclic graph is a cycle and some attached trees. In Section 4 we prove that extremal graphs have attached trees which are of a certain form: they are the concatenation of a path and a star. Intuitively, this type of tree will have their vertices as far as possible from the remaining cycle and other attached trees. In this way the distances to the vertices outside the tree are as large as possible.

4. Using the two previous steps, we know that any extremal graph can be represented with only a few parameters. In Section 5 and Section 6 we determine some conditions on these parameters, which are the length and the number of leaves of the subtrees.

5. Finally, in Section 7 we calculate the Wiener index for the possible extremal graphs in $\mathbb{U}(n,m)$ and compare them to find the maximal value $W(\mathbb{U}(n,m))$.

In the first three main parts, we are proving some heuristic arguments. To prove some of these arguments, we use tree rearrangements. In particular, we are using some kind of subtree pruning and regrafting (SPR), which we define here.

Definition 1.5 (SPR) *Let G be a graph. Given a rooted subtree S of G, such that the root $d = S \cap H$. Pruning S from G is removing the whole structure S excluding the root d. Regrafting S at a vertex v, means that we are taking a copy S' of S which we insert at v, letting its root d' coincide with v. No additional edges are drawn in this process.*

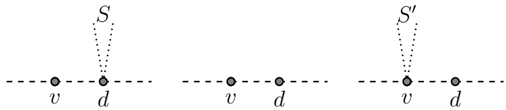

Figure 3. the graph G, S being pruned from G and S being regrafted at v

2 Monotonicity in the matching number

In this section, we will prove the following proposition.

Proposition 2.1 *For fixed n, when $m_1 < m_2$ and the sets $\mathbb{U}(n,m_1)$ and $\mathbb{U}(n,m_2)$ are both nonempty, then $W(\mathbb{U}(n,m_1)) < W(\mathbb{U}(n,m_2))$*

Assume this proposition is not true. In that case there exist some n and m such that $\mathbb{U}(n,m)$ and $\mathbb{U}(n,m+1)$ are both nonempty and $W(\mathbb{U}(n,m)) \geq W(\mathbb{U}(n,m+1))$. For some fixed n, we take the least integer m for which this holds and take an extremal graph $G \in \mathbb{U}(n,m)$ with $W(G) = W(\mathbb{U}(n,m))$. Note that G cannot be a cycle itself, since then $\mathbb{U}(n,m+1)$ would be empty. Hence G contains some cycle C_k where at each vertex r_i of the cycle there are attached some trees T_i (possibly consisting only of the vertex r_i).

Assume at least one of those trees attached is not a path. Then we can look to a vertex w on such a tree with degree at least 3 such that $d(w,C_k)$ is maximal. There are at least two leaves such that the shortest paths from these leaves to C_k contain w, by the choice of w. We call two of them l_1 and l_2.

Let S be the path from w to l_1. We prune S and regraft it at l_2.

The new graph will have matching number m or $m+1$ and a Wiener index which is strictly larger than that of G. This is a contradiction with $W(G) = W(\mathbb{U}(n,m)) \geq W(\mathbb{U}(n,m+1))$.

Hence all trees T_i are paths. Let $P_i = T_i \backslash \{r_i\}$. If there is only one P_i which is nonempty, then G has a maximum matching, which cannot be the case since $\mathbb{U}(n,m+1)$ is nonempty. Let I be the set of indices i such that P_i is nonempty. Take $i_1, i_2 \in I$ and wlog

$$\sum_{t \in I \backslash \{i_1,i_2\}} d(r_{i_1}, r_t)|P_t| \geq \sum_{t \in I \backslash \{i_1,i_2\}} d(r_{i_2}, r_t)|P_t| \tag{1}$$

Let l_{i_1} be the leaf of P_{i_1}. Note that $W(G)$ equals

$$\begin{aligned} W(G) = &\sum_{1\leq t\leq 2} \sum_{u,v\in P_{i_t}} d_G(u,v) + \sum_{u\in P_{i_1}} \sum_{v\in P_{i_2}} d_G(u,v) + \sum_{u,v\in G\backslash\{P_{i_1},P_{i_2}\}} d_G(u,v) \\ &+ \sum_{u\in\{P_{i_1},P_{i_2}\}} \sum_{v\in C_k} d_G(u,v) + \sum_{u\in\{P_{i_1},P_{i_2}\}} \sum_{t\in I\backslash\{i_1,i_2\}} \sum_{v\in P_{i_t}} d_G(u,v) \end{aligned} \tag{2}$$

Prune P_{i_2} and regraft it at l_{i_1}, to get a graph G'. This increases the matching number at most by 1. Using (2), we get that

$$\begin{aligned} W(G') - W(G) = &0 - (d(r_{i_1}, r_{i_2}) + 1)\,|P_{i_1}||P_{i_2}| + 0 + |G\backslash\{P_{i_1}, P_{i_2}\}||P_{i_1}||P_{i_2}| + \\ &\sum_{t\in I\backslash\{i_1,i_2\}} (d(r_{i_1}, r_t) - d(r_{i_2}, r_t))\,|P_t||P_{i_2}| \\ &\geq (k - d(r_{i_1}, r_{i_2}) - 1)\,|P_{i_1}||P_{i_2}| > 0 \end{aligned}$$

since $d(r_{i_1}, r_{i_2}) \leq k-2$, $|G\backslash\{P_{i_1}, P_{i_2}\}| \geq |C_k| = k$ and by (1). Hence G' satisfies $W(G') > W(G)$, while the matching number of G' is at most $m+1$. Since the assumption was that G had the largest Wiener index among all graphs on n vertices with matching number no larger than $m+1$, we get a contradiction. Hence the assumption was false and no such graph G did exist, which proves Proposition 2.1.

3 Reducing the cycle length

In this section, we prove that the extremal graphs cannot contain a cycle of length at least 5, as stated in the following proposition.

Proposition 3.1 *Given a graph in $\mathbb{U}(n,m)$ with Wiener index $W(\mathbb{U}(n,m))$, the cycle of G will be C_3 or C_4.*

Assume some extremal graph $G \in \mathbb{U}(n,m)$ contains a cycle C_k of length $k \geq 5$. Then there is some tree T_i attached to every vertex r_i of C_k. We will call these rooted trees T_1 up to T_k and their roots which lie on the k-gon C_k will be called r_1 up to r_k in cycle order. Note that it is possible that such a tree equals a single vertex, the root. With $|T_i|$ we will denote the number of vertices of the tree T_i and we assume

$$|T_3| = \max\{|T_1|, |T_2|, \ldots, |T_k|\} \tag{3}$$

In Figure 4 we see the graph G consisting of a k-cycle with the trees T_i attached to the vertices of the k-cycle. In Figure 5 there are drawn two possible modifications G_1, G_2 for this graph.

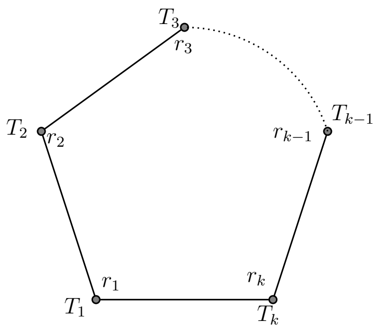

Figure 4. graph G containing C_k

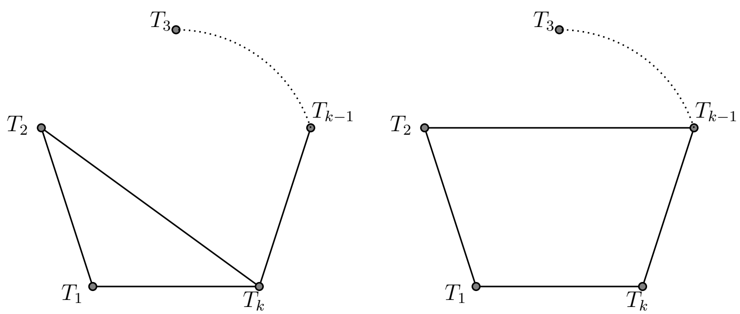

Figure 5. Modifications G_1 and G_2 for graph G

In words, G_1 (resp. G_2) is obtained from G by deleting the edge r_2r_3 and adding r_2r_k (resp. r_2r_{k-1}). Clearly, G_1 and G_2 gave the same number of edges as G does.

We will prove that at least one of G_1, G_2 has a higher Wiener index and a matching number which is at most the matching number of G, implying that G was not an extremal graph.

First we derive that both modifications have a higher Wiener index. Note that

$$W(G) = \sum_{1 \le i \le k} \sum_{u,v \in T_i} d_G(u,v) + \sum_{1 \le i < j \le k} \sum_{u \in T_i} \sum_{v \in T_j} d_G(u,v)$$

which implies that

$$W(G_2) - W(G) = \sum_{1 \le i < j \le k} |T_i||T_j| \left(d_{G_2}(r_i, r_j) - d_G(r_i, r_j) \right) \tag{4}$$

and

$$W(G_1) - W(G_2) = \sum_{1 \le i < j \le k} |T_i||T_j| \left(d_{G_1}(r_i, r_j) - d_{G_2}(r_i, r_j) \right). \tag{5}$$

We will prove that $W(G) < W(G_2) < W(G_1)$.

Note that $d_{G_2}(r_i, r_j) \le d_{G_1}(r_i, r_j)$ for all $1 \le i < j \le k$ except for $i = 2, j = k$. By (5), we get that $W(G_1) - W(G_2) \ge |T_2||T_{k-1}| + |T_2||T_3| - |T_2||T_k| \ge 1$ since $|T_i| \ge 1$ for all $1 \le i \le k$ and due to (3).

Note that $d_{G_2}(r_i, r_j) \ge d_G(r_i, r_j)$ when i and j are both different from 2, since there exists a shortest path in G_2 from r_i to r_j which is a subpath of the path $r_1 r_k r_{k-1} \dots r_3$ which also exists in G.

Note that

$$d_{G_2}(r_i, r_3) = d_G(r_i, r_3) + (k-4) \text{ for } i \in \{1, 2\}. \tag{6}$$

Next, we have to do a small case distinction between k even and k odd.

When k is odd, then $d_{G_2}(r_2, r_j) = d_G(r_2, r_j) - 2$ for $\frac{k+5}{2} \le j \le k-1$ and $d_{G_2}(r_2, r_j) = d_G(r_2, r_j) - 1$ for $j = \frac{k+3}{2}$. For the remaining values of j, one has $d_{G_2}(r_2, r_j) \ge d_G(r_2, r_j)$. Using (4) and (6), we have for k odd that

$$\begin{aligned} W(G_2) - W(G) &\ge (k-4)\left(|T_1| + |T_2|\right)|T_3| - 2 \sum_{\frac{k+5}{2} \le j \le k-1} |T_2||T_j| - |T_2||T_{\frac{k+3}{2}}| \\ &\ge (k-4)|T_1||T_3| > 0 \end{aligned}$$

due to (3), the fact that $|T_1| > 0$ and the fact that there are $\frac{k-5}{2}$ integral numbers in the interval $[\frac{k+5}{2}, k-1]$. When k is even, $d_{G_2}(r_2, r_j) = d_G(r_2, r_j) - 2$ for $2 + \frac{k}{2} \le j \le k-1$, and $d_{G_2}(r_2, r_j) \ge d_G(r_2, r_j)$ for the remaining values of j. Using (4) and (6), we have for k even that

$$W(G_2) - W(G) \geq (k-4)\left(|T_1| + |T_2|\right)|T_3| - 2 \sum_{\frac{k+4}{2} \leq j \leq k-1} |T_2||T_j| \quad \geq (k-4)|T_1||T_3| > 0$$

by (3), $|T_1| > 0$ and the fact that the sum is over $\frac{k-4}{2}$ integers.

Next, we show that the matching number of G is not smaller than the matching numbers of both G_1 and G_2. Assume to the contrary that the matching number of G is strictly smaller than the matching numbers of G_1 or G_2. It is clear that a maximal matching in G can be modified to a matching such that the submatchings are optimal for every tree T_i and do not use r_i if there exists a maximal matching for T_i without using r_i. So starting from the optimal matchings in every T_i such that r_i is not used when not necessary, the remaining task is to find an optimal matching between the r_i which are not used. Since the optimal matching in G_1 is strictly larger than the optimal matching in G, we will use r_2r_k in that matching and so r_1 has to be used in the optimal matching of T_1, since otherwise we could use r_2r_1 instead of r_2r_k in G and take the other pairs as in the matching of G_1. This implies also that r_2 and r_k are not used in the optimal matchings of T_2 and T_k respectively. Similarly we will use r_2r_{k-1} in the optimal covering of G_2. But since r_1 is used in the optimal covering of T_1 and r_k is not used in the optimal covering of T_k, we can replace r_2r_{k-1} by r_kr_{k-1} and so the same matching would work for G, contradiction.

We conclude that G_1 or G_2 have matching numbers not larger than that of G, but have a Wiener index which is greater than the one of G. Together with Proposition 2.1, this implies that it is impossible that $W(G) = W(\mathbb{U}(n,m))$. So the assumption at the beginning was wrong, implying that no extremal graph can contain a cycle of length at least 5, proving the proposition.

4 Optimal form of trees

In this section, we will prove that each tree T_i of an extremal graph is a concatenation of a path and a star, see figure 8.

Proposition 4.1 *For a graph G with $W(G) = W(\mathbb{U}(n,m))$, G is a cycle C_3 or C_4 with trees attached to it, each of which is a path attached to a star.*

We know already by Proposition 3.1 that an extremal graph G is a cycle C_3 or C_4 with some trees T_i attached.

For every i, we take a longest path $\mathcal{P}$ in T_i starting from r_i and call a leaf of that longest path l_i and the adjacent vertex to l_i on that path c_i.

Assume that some tree T_i is not a path attached to a star. Then there exists a nonempty rooted subtree S with root d_i (possibly equal to r_i) on $\mathcal{P}$ which is closest to c_i, as shown in Figure 6. Here $\mathcal{P} \cap S = d_i$ and $\mathcal{P} \cup S$ contain all edges adjacent to d_i.

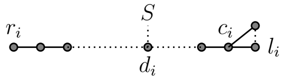

Figure 6. subtree T_i

The vertex d_i partitions the edge set of G into three parts: $E(H_1)$, $E(H_2)$ and $E(S)$ where H_1 is a tree containing l_i and d_i as leaves, H_2 is a unicyclic subgraph with d_i being a leaf.

Let $V_1 = V(H_1)\backslash\{d_i\}$ and $V_2 = V(H_2)\backslash\{d_i\}$ and $V_S = V(S)\backslash\{d_i\}$.

If $|V_1| \leq |V_2|$, then pruning S from G and regrafting S at c_i will strictly increase the Wiener index. Note for this that $\sum_{u\in V_S, v\in V_1} d(u,v)$ has decreased with at most $d(d_i,c_i)|V_1||V_S|$ and $\sum_{u\in V_S, v\in V_2} d(u,v)$ has increased with $d(d_i,c_i)|V_2||V_S|$, while $\sum_{u\in V_S} d(u,d_i)$ strictly increases by the value $|V_S|d(d_i,c_i)$.

Also, the matching number has not been increased. Doing this, we see the trees attached to the cycle, with possible one exception, are each a path attached to a star.

If $|V_1| > |V_2|$, then analogously pruning S and regrafting it at c_j (with $j \neq i$) will do the job (assuming the subtree T_j has length at least 1).

In the case there was only one tree attached to the cycle, we cannot do this and we have to use other replacements. If the cycle is C_4, we can prune S and regraft it at r_j (with r_j and r_i being opposite corners of C_4).

So from now on, we assume the cycle is C_3.

In the case there are multiple subtrees attached to $\mathcal{P}$ with the root not equal to c_i, if there is any subtree S' such that there is some maximum matching that does not use its root $d_i' \in \mathcal{P}$, we can prune S' and regraft is at some r_j $(j \neq i)$ and conclude.

we will assume S' is the rooted subtree with root d_i' which is second closest to c_i. We can prune S and regraft it at d_i', so the Wiener index increases, while the matching number does not and conclude again that the original graph was not extremal.

In the final case, we assume S is the only such subtree connected to $\mathcal{P}$. If $d_i = r_i$, we can prune and regraft S at r_j. If $d(d_i, r_i) > 2$, we prune and regraft it at a vertex v

which is an even distance closer to r_i, which also increases the Wiener index while the matching number is constant. If $d(d_i, r_i) = 1$, we easily can compare the configuration with an other one using a C_4 as cycle, as shown in Figure 7. Here v is the neighbour of d_i which is closer to c_i. The matching number of both configurations is the same, while the Wiener index increases with $(2|V_S| - 1)\,|V_1| - |V_S| > 0$, since $|V_S| \geq 1$ and $|V_1| > 3$.

Since removing S has increased the Wiener index, while the matching number has not been increased, we have a contradiction in our assumption that $W(G) = W(\mathbb{U}(n,m))$ due to Proposition 2.1.

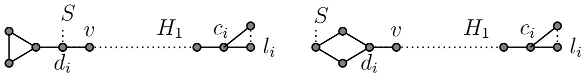

Figure 7. Constructing a better graph

5 Only one long tree

From the previous sections, we can conclude that the possible extremal graphs are as represented in Figure 8, here a, b, c, d are the numbers of leaves in every subtree. In this section, using some calculations we prove some extra conditions on the configuration of an extremal graph.

Proposition 5.1 *An extremal graph in $\mathbb{U}(n,m)$ has at most one tree with height larger than 1 connected to the cycle. When the cycle is C_3, then the extremal graph is isomorphic to a graph $G^3_{a,b,c,j}$ (see Figure 9) with $a \geq b \geq c$. Furthermore, when the cycle is C_4, the extremal graph is isomorphic to a graph $G^4_{a,c,j}$ (see Figure 9) where $a \in \{c, c+1\}$.*

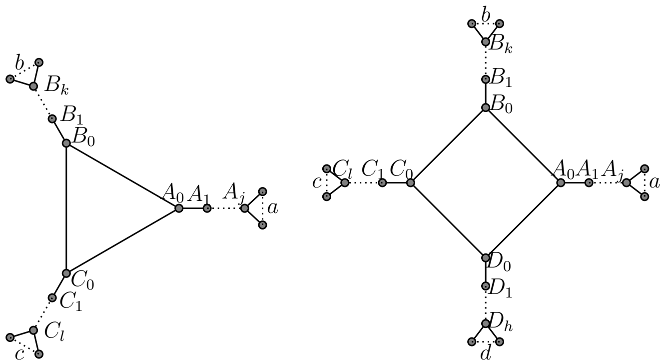

Figure 8. possible configurations

5.1 Case G contains C_3

Using the fact that $W(P_q) = \binom{q+1}{3}$, a calculation gives that the Wiener index of the first configuration in Figure 8 equals

$$\begin{aligned} W(G) =& \binom{k+l+3}{3} + \binom{j+k+3}{3} + \binom{l+j+3}{3} - \binom{k+2}{3} - \binom{l+2}{3} - \binom{j+2}{3} \\ &+ 2\left(\binom{a}{2} + \binom{b}{2} + \binom{c}{2}\right) + ac(l+j+3) + ab(k+j+3) + bc(k+l+3) \\ &+ (a+b+c)\left(\binom{j+2}{2} + \binom{k+2}{2} + \binom{l+2}{2}\right) \\ &+ a(k+l+2)(j+1) + b(k+1)(l+j+2) + c(l+1)(k+j+2) \end{aligned}$$

Wlog $a = \max\{a, b, c\}$. We will prove that $W(G') > W(G)$ when G' is the graph determined by $j' = j + k + l$ and $k' = l' = 0$ and $(a', b', c') = (a, b, c)$, if at most one of j, k, l equals zero. This is a consequence of the following calculation.

$$\begin{aligned} W(G') - W(G) =& jkl + jk + jl + kl + k(a-b)c + l(a-c)b \\ &+ akl + bjl + cjk + k(a-b) + l(a-c) > 0 \end{aligned}$$

since $a - b, a - c \geq 0$ and at least one of jk, jl, kl is strictly positive.

We see that the matching numbers of G' and G are the same. Note we can only use one leaf of the star at the end. Removing $\max\{a-1, 0\}$ edges from the right star and similarly for the two other stars, we have a triangle with some attached paths. Now a maximum matching uses all vertices, or all vertices minus one when the total number is odd. The same holds after the operation $j + k + l \to j'$. This implies that the extremal graphs containing C_3 have two trees of height at most 1. Due to the terms $k(a-b)$ and $l(a-c)$, the new configuration is strictly better if $a > b$ and $k > 0$ or $a > c$ and $l > 0$ originally. Hence $a = \max\{a, b, c\}$ is strictly necessary when we take $j > 0$ in the graph $G^3_{a,b,c,j}$.

5.2 Case G contains C_4

We calculate the Wiener index for the graph in function of the parameters a, b, c, d, h, j, k, l. We use formulas like $W(P_q) = \binom{q+1}{3}$,

$$\begin{aligned} \sum_{0 \leq x \leq j, 0 \leq y \leq k} d(A_x, B_y) &= \binom{k+j+3}{3} - \binom{j+2}{3} - \binom{k+2}{3} \\ &= (k+1)(j+1)\frac{k+j+2}{2} \end{aligned}$$

and

$$\sum_{0\le x\le h, 0\le y\le k} d(D_x, B_y) = (k+1)\binom{h+1}{2} + (h+1)\binom{k+1}{2} + 2(k+1)(h+1)$$
$$= (k+1)(h+1)\frac{k+h+4}{2}$$

multiple times.

$$\begin{aligned}W(G) =&\binom{h+2}{3} + \binom{j+2}{3} + \binom{k+2}{3} + \binom{l+2}{3}\\
&+ (k+1)(j+1)\frac{k+j+2}{2} + (k+1)(l+1)\frac{k+l+2}{2} + (k+1)(h+1)\frac{k+h+4}{2}\\
&+ (l+1)(j+1)\frac{l+j+4}{2} + (h+1)(j+1)\frac{h+j+2}{2} + (h+1)(l+1)\frac{l+h+2}{2}\\
&+ ab(k+j+3) + bc(k+l+3) + cd(l+h+3) + ad(h+j+3)\\
&+ ac(j+l+4) + bd(k+h+4)\\
&+ 2\left(\binom{a}{2} + \binom{b}{2} + \binom{c}{2} + \binom{d}{2}\right)\\
&+ (a+b+c+d)\left(\binom{h+2}{2} + \binom{j+2}{2} + \binom{k+2}{2} + \binom{l+2}{2}\right)\\
&+ a(j+1)(k+h+2) + b(k+1)(l+j+2) + c(l+1)(k+h+2)\\
&+ d(h+1)(l+j+2) + a(j+2)(l+1) + b(k+2)(h+1) + c(l+2)(j+1)\\
&+ d(h+2)(k+1).\end{aligned}$$

Given a graph G with parameters $\{a, b, c, d, h, j, k, l\}$, we construct the graph G' with parameters $\{\frac{a+b+c+d+\epsilon}{2}, 0, \frac{a+b+c+d-\epsilon}{2}, 0, 0, h+j+k+l, 0, 0\}$, where $\epsilon \in \{0, 1\}$ is chosen such that $\epsilon \equiv a+b+c+d \pmod 2$. We will prove that $m(G') \le m(G)$ and $W(G') \ge W(G)$.

Some elementary arithmetic operations show the following (which can be checked for instance by a standard symbolic manipulation program):

$$\begin{aligned}W(G') - W(G) =&hjk + hjl + hkl + jkl + 2hj + hk + 2hl + 2jk + jl + 2kl\\
&+ ahk + ahl + akl + bhj + bhl + bjl + chj + chk + cjk + djk\\
&+ djl + dkl + ah + ak + bj + bl + ch + ck + dj + dl\\
&+ \frac{1}{4}\left((d-a-b-c-1)^2 - (\epsilon-1)^2\right)h\\
&+ \frac{1}{4}\left((a-b-c-d-1)^2 - (\epsilon-1)^2\right)j\\
&+ \frac{1}{4}\left((b-a-c-d-1)^2 - (\epsilon-1)^2\right)k\\
&+ \frac{1}{4}\left((c-a-b-d-1)^2 - (\epsilon-1)^2\right)l\end{aligned}$$

$$+\frac{1}{2}\left((a-c)^2+(b-d)^2-\epsilon^2\right)\geq 0$$

The inequality holds since every term is positive, i.e. when $\epsilon = 1$, then $\max\{|a-c|, |b-d|\} \geq 1$ since $a+b+c+d$ is odd and when $\epsilon = 0$, then $|b-a-c-d-1| \geq 1$ since $a+b+c+d+1$ is odd and similarly for the three other differences of squares. Equality holds if and only if every term is equal to zero. This implies that at least three values in $\{h, j, k, l\}$ are zero. If $h = j = k = l = 0$, equality holds iff $|a-c| + |b-d| \leq 1$. When at least three values of a, b, c, d are nonzero, $m(G) \geq 3$ while $m(G') = 2$ and so G was not optimal by Proposition 2.1. In the other case, we note that $G \sim G'$. In the case one value in $\{h, j, k, l\}$ is nonzero, wlog $j > 0$, we get $b = d = 0$, $|a-c-1| = |\epsilon - 1|$ and $|a-c| = |\epsilon|$ implying $G' \sim G$ again.

We conclude that the extremal graphs containing C_3 or C_4 are isomorphic to a graph of the form $G^3_{a,b,c,j}$ or $G^4_{a,c,j}$ with $a \in \{c, c+1\}$. These graphs are shown in Figure 9.

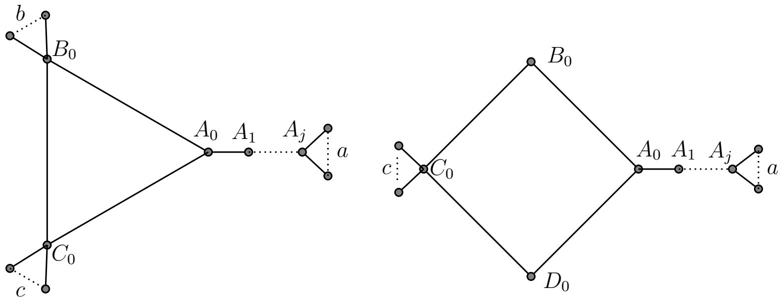

Figure 9. The graphs $G^3_{a,b,c,j}$ and $G^4_{a,c,j}$

6 Conditions on parameters of extremal $G^3_{a,b,c,j}$ and $G^4_{a,c,j}$

In this section, we prove the following proposition.

Proposition 6.1 *If the graph $G^3_{a,b,c,j}$ is an extremal graph, then $b = c = 0$. If the graph $G^4_{a,c,j}$ (with $a \geq c$) is an extremal graph with* $\max\{a, c, j\} \geq 1$, *then $c \geq 1$.*

Take some graph $G = G^3_{a,b,c,j}$.

If $a \geq b = 1 > c = 0$, we choose the graph $G' = G^3_{a,0,0,j+1}$. The graph G was not an extremal graph, since $m(G') = m(G)$ and $W(G') - W(G) = j + a > 0$. If $a \geq b > 1 > c = 0$, the graph $G' = G^4_{a,b-1,j}$ satisfies $m(G') = m(G)$ and $W(G') - W(G) =$

$(b-1)(a+j+1) > 0$. When $a \geq b \geq c \geq 1$, the graph $G' = G^4_{a,b+c-1,j}$ satisfies $m(G') \leq m(G)$ (equality when $2 \nmid j$) and $W(G') - W(G) = (a+j+1)(c+b-1) - bc \geq 1$. Hence G was not an extremal graph, due to Proposition 2.1. So a graph $G^3_{a,b,c,j}$ with $(b,c) \neq (0,0)$ can not be extremal graph. From now onwards we will write $G^3_{a,j}$ instead of $G^3_{a,0,0,j}$, this notation is shown in Figure 2.

Next, we prove that for an extremal graph $G^4_{a,c,j}$ with more than 4 vertices, we have that $c > 0$. Take an extremal graph $G^4_{a,c,j}$ Since an extremal graph satisfies $|a-c| \leq 1$ as shown in Subsection 5.2, we can take $a = c = 0$. (since $G^4_{1,0,j} = G^4_{0,0,j+1}$)

Observe that $m(G^4_{0,0,j}) = m(G^3_{0,j+1})$ and $W(G^4_{0,0,j}) < W(G^3_{0,j+1})$, from which the conclusion follows.

7 Calculating $W(\mathbb{U}(n,m))$

From the results of previous sections, the extremal graphs are of the form $G^3_{a,j}$, $G^4_{a,a-1,j}$ with $a \geq 2$ or $G^4_{a,a,j}$ with $a \geq 1$. We will determine the extremal values $W(\mathbb{U}(n,m))$ for $m \geq 2$.

When $G^3_{a,j}$ is a graph with n vertices and matching number m, then $n = a+j+3$ and $m = 2 + \lfloor \frac{j}{2} \rfloor$. So $j \in \{2m-4, 2m-3\}$ and $a = n-(j+3)$. Notice that $W\left(G^3_{n-2m+1,2m-4}\right) \leq W\left(G^3_{n-2m,2m-3}\right)$ with equality iff $n = 2m$, in which case $a = 0$ and so we actually look only to the same graph.

Next, note that $m(G^4_{a,a-1,j}) = m(G^4_{a,a,j}) = 2 + \lfloor \frac{j+1}{2} \rfloor$ and $n(G^4_{a,a-1,j}) + 1 = n(G^4_{a,a,j}) = j + 4 + 2a$. So $j \in \{2m-5, 2m-4\}$.

According to the parity of n, we can compare the Wiener index of the corresponding graphs $G^4_{a,a-1,j}$ and $G^4_{a,a,j}$. We check that

$$W\left(G^4_{n/2-m,n/2-m,2m-4}\right) - W\left(G^4_{n/2-m+1,n/2-m,2m-5}\right) = \frac{1}{4}(n-2m)(n+2m-4) \geq 0$$

where equality cannot occur, since $m \geq 2$ and $n \geq 2m+2$ as we need $n/2 - m \geq 1$.

When n is odd, we have that

$$\begin{aligned} &W\left(G^4_{n/2-m+1/2,n/2-m-1/2,2m-4}\right) - W\left(G^4_{n/2-m+1/2,n/2-m+1/2,2m-5}\right) \\ &\qquad = \frac{1}{4}(n-2m)(n+2m-4) + m - \frac{13}{4} \geq 0 \end{aligned}$$

since $n \geq 2m+1$ and $m \geq 2$. Equality occurs only when $n = 5$ and $m = 2$, but then $2m-5 < 0$, implying that the graph $G^4_{n/2-m+1/2,n/2-m+1/2,2m-5}$ does not exist.

To finish the search for the extremal graph, we have to compare the graphs of the form G^4 with the G^3 graph. When n is odd,

$$W\left(G^4_{n/2-m+1/2,n/2-m-1/2,2m-4}\right) - W\left(G^3_{n-2m,2m-3}\right)$$
$$= n + 2m^3 + nm + \frac{1}{2}mn^2 - \frac{1}{2}n^2 - \frac{7}{2}m - 2nm^2 + \frac{5}{2}.$$

Taking $n = 2m + k$, we get that this value equals $1 + \frac{1}{2}(k-3)(k+1)(m-1)$ which is strictly positive for $k \geq 3$ and strictly negative for $k = 1$. When n is even,

$$W\left(G^4_{n/2-m,n/2-m,2m-4}\right) - W\left(G^3_{n-2m,2m-3}\right) = 4 + n + 2m^3 + nm + \frac{1}{2}mn^2 - \frac{1}{2}n^2 - 4m - 2nm^2.$$

Taking $n = 2m + k$, this expression equals $2 + \frac{1}{2}(k^2 - 2k - 4)(m-1)$. This is smaller or equal than zero for $k \in \{0, 2\}$ with equality iff $m = 2$. When $k \geq 4$, it is strictly positive. We summarize all those results in the following theorem with Figure 2 showing the exact picture of the extremal graphs.

Theorem 7.1 *Let $G \in \mathbb{U}(n, m)$, where $2 \leq m \leq \lfloor \frac{n}{2} \rfloor$.*

- *If $n \leq 2m + 2$, then $W(G) \leq 2 - \frac{8}{3}m^3 + 2m^2 + \frac{5}{3}m + 2m^2 n - 3mn - 2n + n^2$ with equality iff $G = G^4_{0,0,0}, G^3_{0,1}$ for $(n, m) = (4, 2)$, $G = G^4_{1,1,0}, G^3_{2,1}$ for $(n, m) = (6, 2)$ and $G = G^3_{n-2m,2m-3}$ otherwise.*
- *If $n \geq 2m+3$ and n is odd, then $W(G) \leq \frac{9}{2} - n - \frac{2}{3}m^3 + \frac{1}{2}n^2 - 2nm + 2m^2 + \frac{1}{2}mn^2 - \frac{11}{6}m$ with equality iff $G = G^4_{n/2-m+1/2,n/2-m-1/2,2m-4}$.*
- *If $n \geq 2m+4$ and n is even, then $W(G) \leq 6 - n - \frac{2}{3}m^3 + \frac{1}{2}n^2 - 2nm + 2m^2 + \frac{1}{2}mn^2 - \frac{7}{3}m$ with equality iff $G = G^4_{n/2-m,n/2-m,2m-4}$.*

Acknowledgment: The author is very grateful to Jan Bok for presenting this problem in Ghent at GGTW 2017.

References

[1] P. Dankelmann, Average distance and independence number, *Discr. Appl. Math.* **51** (1994) 75-83.

[2] Z. Du, B. Zhou, Minimum Wiener indices of trees and unicyclic graphs of given matching number, *MATCH Commun. Math. Comput. Chem.* **63** (2010) 101–112.

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