On the difference between the (revised) Szeged index and the Wiener index of cacti *

Sandi Klavžar [a,b,c] Shuchao Li [d,†] Huihui Zhang [e]

[a] Faculty of Mathematics and Physics, University of Ljubljana, Slovenia
sandi.klavzar@fmf.uni-lj.si

[b] Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia

[c] Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

[d] Faculty of Mathematics and Statistics, Central China Normal University,
Wuhan 430079, P.R. China
lscmath@mail.ccnu.edu.cn

[e] Department of Mathematics, Luoyang Normal Univeristy, Luoyang 471002, P.R. China
zhanghhmath@163.com

Abstract

A connected graph is said to be a cactus if each of its blocks is either a cycle or an edge. Let $\mathscr{C}_n$ be the set of all n-vertex cacti with circumference at least 4, and let $\mathscr{C}_{n,k}$ be the set of all n-vertex cacti containing exactly $k \geqslant 1$ cycles where $n \geqslant 3k + 1$. In this paper, lower bounds on the difference between the (revised) Szeged index and Wiener index of graphs in $\mathscr{C}_n$ (resp. $\mathscr{C}_{n,k}$) are proved. The minimum and the second minimum values on the difference between the Szeged index and Wiener index of graphs among $\mathscr{C}_n$ are determined. The bound on the minimum value is strengthened in the bipartite case. A lower bound on the difference between the revised Szeged index and Wiener index of graphs among $\mathscr{C}_{n,k}$ is also established. Along the way the corresponding extremal graphs are identified.

Key words: Wiener index; Szeged index; Revised Szeged index; Extremal problem; Isometric cycle

AMS Subj. Class: 05C35, 05C12

*S.K. acknowledge the financial support from the Slovenian Research Agency (research core funding No. P1-0297). S.L. acknowledge the financial support from the National Natural Science Foundation of China (Grant Nos. 11271149, 11371062, 11671164), the Program for New Century Excellent Talents in University (Grant No. NCET-13-0817) and the Special Fund for Basic Scientific Research of Central Colleges (Grant No. CCNU15A02052)). H.Z. acknowledge the financial support from the excellent doctoral dissertation cultivation grant from Central China Normal University (Grant No. 2016YBZZ084).

†Corresponding author

1 Introduction

All graphs considered in this paper are finite, undirected and simple. If G is a graph, then its vertex set and edge set will be denoted V_G and E_G, respectively. The *distance*, $d_G(u,v)$, between vertices u and v of G is the length of a shortest u,v-path in G. The celebrated *Wiener index* (or *transmission*) $W(G)$ of G is the sum of distances between all pairs of vertices of G, that is,

$$W(G) = \sum_{\{u,v\} \subseteq V_G} d_G(u,v)\,. \tag{1}$$

This distance-based graph invariant was in chemistry introduced back in 1947 [36] and in mathematics about 30 years later [9]. Nowadays, the Wiener index is a well-established and much studied graph invariant; see the reviews [6, 7], a collection of papers dedicated to a half century of investigations of the Wiener index [12], and recent papers [25, 27, 35, 38].

Given an edge $e = uv$ of a graph G, set

$$N_u(e) = \{w \in V_G :\ d_G(u,w) < d_G(v,w)\}, \quad N_v(e) = \{w \in V_G :\ d_G(v,w) < d_G(u,w)\},$$
$$N_0(e) = \{w \in V_G :\ d_G(u,w) = d_G(v,w)\}.$$

Clearly, $N_u(e) \cup N_v(e) \cup N_0(e)$ is a partition of V_G with respect to e, where $N_0(e) = \emptyset$ if G is bipartite. For convenience, denote by $n_u(e)$, $n_v(e)$ and $n_0(e)$ the number of vertices of $N_u(e)$, $N_v(e)$ and $N_0(e)$, respectively. Thus, one has $n_u(e) + n_v(e) + n_0(e) = |V_G|$.

From [13, 36] we know that for a tree T,

$$W(T) = \sum_{e=uv \in E_T} n_u(e) n_v(e)\,.$$

Inspired by this fact, Gutman [10] introduced the *Szeged index* $Sz(G)$ of a graph G as

$$Sz(G) = \sum_{e=uv \in E_G} n_u(e) n_v(e)\,.$$

Furthermore, in order to involve also those vertices that are at equal distance from the endpoints of an edge, Randić [33] proposed the *revised Szeged index* $Sz^*(G)$ of a graph G as follows:

$$Sz^*(G) = \sum_{e=uv \in E_G} \left(n_u(e) + \frac{n_0(e)}{2}\right)\left(n_v(e) + \frac{n_0(e)}{2}\right).$$

Since $Sz(T) = W(T)$ holds for any tree T, a lot of research has been done on the difference between the Szeged index and the Wiener index on general graphs. If G is a graph, then $Sz(G) - W(G) \geqslant 0$ holds, a result conjectured in [10] and proved in [24]. Moreover, $Sz(G) = W(G)$ holds if and only if every block of G is a complete graph [8], see [17] for an alternative proof. Nadjafi-Arani et al. [30] investigated the structure of graphs G with $Sz(G) - W(G) = k$, where k is a positive integer. In particular, in [31] they characterized the graphs for which the difference is 4 and 5. The difference between $Sz(G)$ and $W(G)$ in networks was investigated in [20].

Based on the computer program AutoGraphiX, Hansen [14] presented nine conjectures on relations between the (revised) Szeged index and the Wiener index. Chen, Li and Liu [3, 4] proved three of these conjectures, while Li and Zhang [28] confirmed three additional above conjectures; these results deal with quotients between the (revised) Szeged index and the Wiener index of unicyclic graphs. Motivated by these conjectures, further relationship between the Wiener index and the (revised) Szeged index was established in [39]. For additional results on relations between the (revised) Szeged index and Wiener index see [2, 21, 22], and for more information about the (revised) Szeged index in general we refer to [1, 3, 15, 16, 26, 29, 32, 34, 37].

In this paper we continue the above direction of research by considering the difference between the Szeged index and the Wiener index on cacti. Since this difference is 0 if the circumference of a cactus is 3, we may and will restrict our attention to cacti with circumference at least 4. In the next section we give necessary definitions and state the main results of the paper. The first of them determines the minimum value on the difference between the Szeged index and the Wiener index of cacti, the second result strengthens this result in the bipartite case, while the third result determines the second minimum value on the difference between the Szeged index and the Wiener index of cacti. These three theorems are then proved in Section 4. The last main theorem that establishes a sharp lower bound on the difference between the revised Szeged index and the Wiener index is proved in Section 5. In Section 3 we recall some known results and prove new results that are needed for the proofs of the main results, while in the concluding section a couple of consequences are listed and a couple of problems are posed.

2 Main results

A *cactus* is a (connected) graph in which any two cycles have at most one common vertex, that is, a graph whose every block is either an edge or a cycle. A cycle in a cactus is called an *end-block cycle* if all but one vertex of this cycle have degree 2. The *circumference* of a graph is the length of its longest cycle. As already mentioned, since $Sz(G) - W(G) = 0$ if the circumference of G is 3, hence we set:

$$\mathscr{C}_n = \{G : \ G \text{ is a cactus of order } n \text{ and circumference at least } 4\}\,.$$

In addition, for integers $3 \leqslant r \leqslant n$, let $\mathscr{C}_n^r$ be the subset of n-vertex cacti defined as follows. If $r = n$, then set $\mathscr{C}_n^n = \{C_n\}$. Otherwise, $\mathscr{C}_n^r$ consists of the n-vertex cacti each of which is obtained from the cycle C_r, and either a cactus G' rooted at a vertex of the C_r or cacti G'' and G''' rooted at two adjacent vertices of the C_r, where G', G'', and G''' are cacti whose blocks are only K_2 or C_3. Using this notation our first main result reads as follows.

Theorem 2.1. *If $G \in \mathscr{C}_n$, then*

$$Sz(G) - W(G) \geqslant 2n - 5$$

with equality if and only if $G \in \mathscr{C}_n^5$.

In the bipartite case this result can be strengthened as follows. For a graph G let $\ell(G)$ denote the sum of the lengths of the cycles of G. Then:

Theorem 2.2. *If $G \in \mathscr{C}_n$ is bipartite, then*

$$Sz(G) - W(G) \geqslant \ell(G)(n - 2)$$

with equality if and only if each block of G is either a K_2 or an end-block C_4.

We note that a result closely related to Theorem 2.2 has been proved in [3, 4]. Namely, if G is a connected bipartite graph of order at least 4, then $Sz(G) - W(G) \geqslant 4n - 8$. Moreover, the bound is best possible when the graph is composed of C_4 and a tree on $n-3$ vertices sharing a single vertex. In the case of cacti this result coincides with Theorem 2.2 for $\ell(G) = 4$.

In our next result we establish a sharp lower bound on the difference between the Szeged index and the Wiener index of graphs from the family $\mathscr{C}_n \backslash \mathscr{C}_n^5$. In other words, the next theorem gives the second minimum value on the difference between Sz and W in the class of cacti. For the equality case we define $\mathcal{H}$ to be the set of graphs isomorphic to some graph from the two families of graphs that are schematically presented in Fig. 1.

Theorem 2.3. *If $G \in \mathscr{C}_n \backslash \mathscr{C}_n^5$, then*

$$Sz(G) - W(G) \geqslant 4n - 10$$

with equality if and only if $G \in \mathcal{H}$.

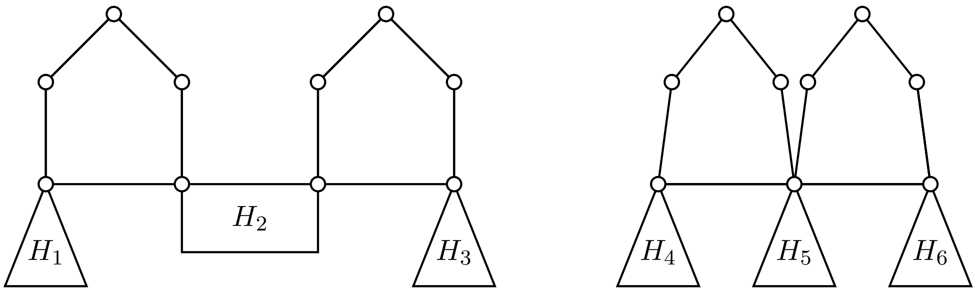

Figure 1: Two families of graphs that constitute $\mathcal{H}$; here each H_i, $1 \leqslant i \leqslant 6$, is either a trivial graph or each block of H_i is K_2 or C_3.

To formulate our last main result that deals with the difference between the revised Szeged index and the Wiener index in cacti, we define

$$\mathscr{C}_{n,k} = \{G : \ G \text{ is a cactus of order } n \text{ containing exactly } k \text{ cycles}\}.$$

Then our result reads as follows:

Theorem 2.4. *Let $G \in \mathscr{C}_{n,k}$, where $k \geqslant 1$ and $n \geqslant 3k+1$.*

(i) *If $4 \leqslant n \leqslant 9$, then*

$$Sz^*(G) - W(G) \geqslant \frac{k(n^2+4n-6)}{4}$$

with equality if and only if each block of G is either a K_2 or an end-block C_3.

(ii) *If $n \geqslant 10$, then*

$$Sz^*(G) - W(G) \geqslant k(4n-8)$$

with equality if and only if each block of G is either a K_2 or an end-block C_4.

It is interesting to observe that the extremal graphs of Theorem 2.2 and of Theorem 2.4 (ii) are the same.

3 Preliminary results

In this section, we give some preliminary results which will be used in the subsequent sections.

Simić et al. [34] rewrote the Szeged index as:

$$Sz(G) = \sum_{e=uv\in E_G} \sum_{\{x,y\}\subseteq V_G} \mu_{x,y}(e), \tag{2}$$

where $\mu_{x,y}(e)$ is the contribution of the vertex pair x and y to $n_u(e)n_v(e)$, that is,

$$\mu_{x,y}(e) = \begin{cases} 1, & \text{if } \begin{cases} x \in N_u(e) \text{ and } y \in N_v(e), \\ \text{or} \\ x \in N_v(e) \text{ and } y \in N_u(e), \end{cases} \\ 0, & \text{otherwise.} \end{cases}$$

Setting

$$\pi(x, y) = \sum_{e\in E_G} \mu_{x,y}(e) - d_G(x, y)\,,$$

it then follows from Equalities (1) and (2) that

$$Sz(G) - W(G) \;=\; \sum_{\{x,y\}\subseteq V_G} \pi(x, y)\,. \tag{3}$$

Recall that a subgraph H of a graph G is called *isometric* if the distance between any two vertices of H is independent of whether it is computed in the subgraph H or in G.

Lemma 3.1 ([39])**.** *Let C_r be an isometric cycle of a connected graph G, and let $x, y \in V_{C_r}$.*

(i) *If r is even, then $\pi(x, y) \geqslant \sum_{e\in E_{C_r}} \mu_{x,y}(e) - d_{C_r}(x, y) = d_{C_r}(x, y)$.*

(ii) *If r is odd, then $\pi(x, y) \geqslant \sum_{e\in E_{C_r}} \mu_{x,y}(e) - d_{C_r}(x, y) = d_{C_r}(x, y) - 1$.*

Lemma 3.2. *Let G_1, G_2 be vertex-disjoint, connected graphs of order at least 2. Let G be the graph obtained from G_1, G_2 by identifying one vertex of G_1 with one vertex of G_2, denote the new vertex by u_0. Then for all $x \in V_{G_1}\backslash\{u_0\}$ and $y \in V_{G_2}\backslash\{u_0\}$, one has $\pi(x, y) = \pi(x, u_0)+\pi(u_0, y)$.*

Proof. Consider a vertex pair $\{x, y\}$ with $x \in V_{G_1}\backslash\{u_0\}, y \in V_{G_2}\backslash\{u_0\}$. Note that u_0 is a cut vertex of G, we have $d_G(x, y) = d_G(x, u_0) + d_G(u_0, y)$. Next, we show that for any edge $e = uv \in E_{G_1}$,

$$\mu_{x,y}(e) = 1 \text{ if and only if } \mu_{x,u_0}(e) = 1. \tag{4}$$

First, suppose that $\mu_{x,y}(e) = 1$. Then we may assume without loss of generality that $x \in N_u(e)$ and $y \in N_v(e)$. Let P_k be a shortest path connecting y and u, and P_l be a shortest path joining y and v. As $y \in N_v(e)$, one has $l < k$. Since u_0 is a cut vertex of G, we have $u_0 \in V_{P_k} \cap V_{P_l}$. Thus, P_k and P_l can be written as $P_k = yP_au_0P_bu$ and $P_l = yP_au_0P_cv$, where P_a (resp. P_b, P_c) is a shortest path joining y and u_0 (resp. u_0 and u, u_0 and v). Therefore, $k = a + b, l = a + c$. As $l < k$, we have $c < b$, i.e., $d_G(u_0, v) < d_G(u_0, u)$. This implies that $u_0 \in N_v(e)$. Note that $x \in N_u(e)$, hence $\mu_{x,u_0}(e) = 1$. Similarly one shows that if $\mu_{x,u_0}(e) = 1$, then $\mu_{x,y}(e) = 1$.

In view of (4) and the definition of $\mu_{x,y}(e)$, it is clear that $\sum_{e\in E_{G_1}}\mu_{x,y}(e)=\sum_{e\in E_{G_1}}\mu_{x,u_0}(e)$. By a similar discussion as above, $\sum_{e\in E_{G_2}}\mu_{x,y}(e)=\sum_{e\in E_{G_2}}\mu_{u_0,y}(e)$ holds. Thus, one has

$$\begin{aligned}
\pi(x,y) &= \sum_{e\in E_G}\mu_{x,y}(e)-d_G(x,y)\\
&= \sum_{e\in E_{G_1}}\mu_{x,y}(e)+\sum_{e\in E_{G_2}}\mu_{x,y}(e)-d_G(x,u_0)-d_G(u_0,y)\\
&= \left(\sum_{e\in E_{G_1}}\mu_{x,u_0}(e)-d_G(x,u_0)\right)+\left(\sum_{e\in E_{G_2}}\mu_{u_0,y}(e)-d_G(u_0,y)\right)\\
&= \pi(x,u_0)+\pi(u_0,y),
\end{aligned}$$

as desired. □

Lemma 3.3. *Let G be an n-vertex cactus containing an even cycle C_r. Then*

$$Sz(G)-W(G)\geqslant\frac{2nr^2-r^3}{8}$$

with equality if and only if $G=C_r$ or G is composed from C_r and a graph G' on $n-r+1$ vertices sharing a single vertex, where each block of G' is a K_2 or a C_3.

Proof. For convenience, let $C_r=v_1v_2\dots v_rv_1$. Clearly, C_r is an isometric cycle. Let G_i be the component of $G-E_{C_r}$ containing the vertex v_i, $1\leqslant i\leqslant r$. Thus, $|V_{G_i}|\geqslant 1$ for all $1\leqslant i\leqslant r$. For each edge $e=uv\in E_{G_i}$, $1\leqslant i\leqslant r$, and every vertex pair $\{x,y\}\subseteq V_{C_r}$, it is straightforward to check that

$$x,y\in\begin{cases}N_u(e), & \text{if } v_i\in N_u(e);\\ N_v(e), & \text{if } v_i\in N_v(e);\\ N_0(e), & \text{if } v_i\in N_0(e).\end{cases}$$

This implies that $\mu_{x,y}(e)=0$. Therefore, for every vertex pair $\{x,y\}\subseteq V_{C_r}$, we have

$$\pi(x,y)=\sum_{e\in E_G}\mu_{x,y}(e)-d_G(x,y)=\sum_{e\in E_{C_r}}\mu_{x,y}(e)-d_{C_r}(x,y)=d_{C_r}(x,y). \tag{5}$$

This gives

$$\sum_{x,y\in V_{C_r}}\pi(x,y)=\sum_{x,y\in V_{C_r}}d_{C_r}(x,y)=\frac{r^3}{8}. \tag{6}$$

If $|V_{G_i}|\geqslant 2$, then, for every vertex pair $\{x,y\}$ with $x\in V_{C_r}, y\in V_{G_i}\backslash V_{C_r}$, one has $d_G(y,v_i)=\min_{z\in V_{C_r}}d_G(y,z)$. Note that v_i is a cut vertex of G. By Lemma 3.2, $\pi(x,y)=\pi(x,v_i)+\pi(v_i,y)$.

Thus,

$$\begin{aligned}
\sum_{x\in V_{C_r}}\sum_{y\in V_G\backslash V_{C_r}}\pi(x,y) &= \sum_{i=1}^{r}\sum_{x\in V_{C_r}}\sum_{y\in V_{G_i}\backslash V_{C_r}}(\pi(x,v_i)+\pi(v_i,y)) \\
&\geqslant \sum_{i=1}^{r}\sum_{x\in V_{C_r}}\sum_{y\in V_{G_i}\backslash V_{C_r}}\pi(x,v_i) \quad (\text{since } \pi(v_i,y)\geqslant 0) \qquad (7)\\
&= (n-r)\sum_{x\in V_{C_r}} d_{C_r}(x,v_i) \qquad (\text{by } (5))\\
&= \frac{(n-r)r^2}{4}, \qquad (8)
\end{aligned}$$

where the equality in (7) holds if and only if $\pi(v_i,y)=0$ for all $y\in V_{G_i}\backslash V_{C_r}, 1\leqslant i\leqslant r$.

Note that $\pi(x,y)\geqslant 0$ for every vertex pair $\{x,y\}\subseteq V_G\backslash V_{C_r}$. Together with (3), (6) and (8), we obtain that

$$\begin{aligned}
Sz(G)-W(G) &= \sum_{x,y\in V_{C_r}}\pi(x,y)+\sum_{x\in V_{C_r},y\in V_G\backslash V_{C_r}}\pi(x,y)+\sum_{x,y\in V_G\backslash V_{C_r}}\pi(x,y)\\
&\geqslant \frac{r^3}{8}+\frac{(n-r)r^2}{4}+0 \qquad (9)\\
&= \frac{2nr^2-r^3}{8},
\end{aligned}$$

where the equality in (9) holds if and only if $\pi(v_i,y)=0$ for all $y\in V_{G_i}\backslash V_{C_r}$, $1\leqslant i\leqslant r$, and $\pi(x,y)=0$ for every vertex pair $\{x,y\}\subseteq V_G\backslash V_{C_r}$.

Now, we show that if $Sz(G)-W(G)=\frac{2nr^2-r^3}{8}$, then G has exactly one cycle whose length is at least 4. Suppose on the contrary that G contains a cycle C_k, $k\geqslant 4$, different from C_r. Because G is a cactus, $|V_{C_r}\cap V_{C_k}|\leqslant 1$. This implies that $|V_{C_k}\backslash V_{C_r}|\geqslant 3$. Thus, there exist two vertices $u,v\in V_{C_k}\backslash V_{C_r}$ such that $d_{C_k}(u,v)=2$. By Lemma 3.1, we have $\pi(u,v)\geqslant d_{C_k}(u,v)-1=1$, which contracts the assumption that $\pi(x,y)=0$ for every vertex pair $\{x,y\}\subseteq V_G\backslash V_{C_r}$. Therefore, G contains only one cycle whose length is at least 4.

If there are two components, say G_a, G_b, of $G-E_{C_r}$ such that $|V_{G_a}|,|V_{G_b}|\geqslant 2$, then consider $x\in V_{G_a}\backslash\{v_a\}$ and $y\in V_{G_b}\backslash\{v_b\}$. Note that v_a and v_b are cut vertices of G. By Lemma 3.2, we have $\pi(x,y)=\pi(x,v_a)+\pi(v_a,v_b)+\pi(v_b,y)\geqslant\pi(v_a,v_b)\geqslant 1$, which is also a contradiction to the fact that $\pi(x,y)=0$ for every vertex pair $\{x,y\}\subseteq V_G\backslash V_{C_r}$. Therefore, there is at most one nontrivial component G_i. In this case, by direct calculation, we have $Sz(G)-W(G)=\frac{2nr^2-r^3}{8}$.

Hence, $\sum_{x,y\in V_G}\pi(x,y)=\frac{2nr^2-r^3}{8}$ if and only if G is composed of a cycle C_r on r vertices and a graph G' on $n-r+1$ vertices sharing a single vertex, where each block of G' is either a K_2 or a C_3. □

Lemma 3.4. *Let G be an n-vertex cactus containing an odd cycle $C_r=v_1v_2\dots v_rv_1$ of length at least 5. Then*

$$Sz(G)-W(G)\geqslant\frac{(r-1)(r-3)(2n-r)}{8}$$

with equality if and only if $G\in\mathscr{C}_n^r$.

Proof. Clearly, C_r is an isometric cycle. Let G_i be the component of $G - E_{C_r}$ containing the vertex $v_i, 1 \leqslant i \leqslant r$. Then $|V_{G_i}| \geqslant 1$ for any $1 \leqslant i \leqslant r$. For any edge $e = uv \in E_{G_i}$, $1 \leqslant i \leqslant r$, and every vertex pair $\{x, y\} \subseteq V_{C_r}$, by a discussion similar to the one from the beginning of the proof of Lemma 3.3, we obtain $\mu_{x,y}(e) = 0$. Thus, for every vertex pair $\{x, y\} \subseteq V_{C_r}$, we have

$$\pi(x,y) = \sum_{e\in E_G} \mu_{x,y}(e) - d_G(x,y) = \sum_{e\in E_{C_r}} \mu_{x,y}(e) - d_{C_r}(x,y) = d_{C_r}(x,y) - 1. \tag{10}$$

So,

$$\sum_{x,y\in V_{C_r}} \pi(x,y) = \sum_{x,y\in V_{C_r}} (d_{C_r}(x,y) - 1) = \frac{r(r-1)(r-3)}{8}. \tag{11}$$

For every vertex pair $\{x, y\}$ with $x \in V_{C_r}$, $y \in V_{G_i} \backslash V_{C_r}$, $1 \leqslant i \leqslant r$, we have $d_G(y, v_i) = \min_{z\in V_{C_r}} d_G(y, z)$. Since v_i is a cut vertex of G, Lemma 3.2 implies that $\pi(x, y) = \pi(x, v_i) + \pi(v_i, y)$. Then

$$\begin{aligned}
\sum_{x\in V_{C_r}} \sum_{y\in V_{G_i}\backslash V_{C_r}} \pi(x,y) &= \sum_{i=1}^{r} \sum_{x\in V_{C_r}} \sum_{y\in V_{G_i}\backslash V_{C_r}} (\pi(x,v_i) + \pi(v_i,y)) \\
&\geqslant \sum_{i=1}^{r} \sum_{x\in V_{C_r}} \sum_{y\in V_G\backslash V_{C_r}} \pi(x,v_i) \qquad (\text{since } \pi(v_i,y) \geqslant 0) \qquad (12) \\
&= (n-r) \sum_{x\in V_{C_r}} (d_{C_r}(x,v_i) - 1) \qquad (\text{by } (10)) \\
&= \frac{(n-r)(r-1)(r-3)}{4}, \qquad (13)
\end{aligned}$$

where the equality in (12) holds if and only if $\pi(v_i, y) = 0$ for all $y \in V_{G_i} \backslash V_{C_r}$, $1 \leqslant i \leqslant r$.

Note that $\pi(x, y) \geqslant 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash V_{C_r}$. Combining this fact with (3), (11) and (13), it follows that

$$\begin{aligned}
Sz(G) - W(G) &= \sum_{x,y\in V_{C_r}} \pi(x,y) + \sum_{x\in V_{C_r}, y\in V_G\backslash V_{C_r}} \pi(x,y) + \sum_{x,y\in V_G\backslash V_{C_r}} \pi(x,y) \\
&\geqslant \frac{r(r-1)(r-3)}{8} + \frac{(n-r)(r-1)(r-3)}{4} \qquad (14) \\
&= \frac{(r-1)(r-3)(2n-r)}{8},
\end{aligned}$$

where the equality in (14) holds if and only if $\pi(v_i, y) = 0$ for all $y \in V_{G_i}$, $1 \leqslant i \leqslant r$ and $\pi(x, y) = 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash V_{C_r}$.

We next show that if $Sz(G) - W(G) = \frac{(r-1)(r-3)(2n-r)}{8}$, then C_r is the only cycle of G with length at least 4. Otherwise, there exists another cycle C_k of length $k \geqslant 4$. Since G is a cactus, $|V_{C_r} \cap V_{C_k}| \leqslant 1$, which in turn implies that $|V_{C_k} \backslash V_{C_r}| \geqslant 3$. Thus, there are two vertices $u, v \in V_{C_k} \backslash V_{C_r}$ such that $d_{C_k}(u, v) = 2$. By Lemma 3.1, we have $\pi(u, v) \geqslant d_{C_k}(u, v) - 1 = 1$, this is a contradiction to $\pi(x, y) = 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash V_{C_r}$. Therefore, C_r is the only cycle of G with length at least 4.

Suppose that there exist two nontrivial components G_i, G_j of $G - E_{C_r}$ such that $|i - j| \neq 1, r - 1$. Then, take two vertices x, y with $x \in V_{G_i} \backslash \{v_i\}, y \in V_{G_j} \backslash \{v_j\}$. By (10), we have $\pi(v_i, v_j) \geqslant 1$. Note that v_i, v_j are cut vertices of G. By Lemma 3.2, we have

$$\pi(x, y) = \pi(x, v_i) + \pi(v_i, v_j) + \pi(v_j, y) \geqslant \pi(v_i, v_j) \geqslant 1,$$

a contradiction to the fact $\pi(x, y) = 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash V_{C_r}$. Therefore, there are just two nontrivial G_i, G_{i+1} or there is only one nontrivial G_i. For each of the above cases, by direct calculation, we have $Sz(G) - W(G) = \frac{(r-1)(r-3)(2n-r)}{8}$.

Hence, $Sz(G) - W(G) = \frac{(r-1)(r-3)(2n-r)}{8}$ holds if and only if $G \in \mathscr{C}_n^r$, as desired. □

4 Proofs of Theorems 2.1–2.3

4.1 Proof of Theorem 2.1

Recall the statement of the theorem to be proved in this subsection: If $G \in \mathscr{C}_n$, then $Sz(G) - W(G) \geqslant 2n - 5$ with equality if and only if $G \in \mathscr{C}_n^5$.

Let C_r be a longest cycle of G. Then $r \geqslant 4$ since the circumference of G is at least 4. If r is even, then by Lemma 3.3,

$$Sz(G) - W(G) \geqslant \frac{2nr^2 - r^3}{8} > 2n - 5$$

because $n \geqslant r \geqslant 4$. If r is odd, then by Lemma 3.4,

$$\begin{aligned} Sz(G) - W(G) &\geqslant \frac{(r-1)(r-3)(2n-r)}{8} && (15) \\ &\geqslant 2n - 5. \quad (\text{since } r \geqslant 5) && (16) \end{aligned}$$

Based on Lemma 3.4, the equality in (15) holds if and only if $G \in \mathscr{C}_n^r$; whereas the equality in (16) holds if and only if $r = 5$.

Hence, $Sz(G) - W(G) = 2n - 5$ holds if and only if $G \in \mathscr{C}_n^5$, as claimed.

4.2 Proof of Theorem 2.2

Recall the statement of Theorem 2.2: If $G \in \mathscr{C}_n$ is bipartite, then $Sz(G) - W(G) \geqslant \ell(G)(n - 2)$ with equality if and only if each block of G is either a K_2 or an end-block C_4, where $\ell(G)$ is the sum of the lengths of the cycles of G.

We first note that a bipartite cactus is a *partial cube*, that is, isometrically embeddable into a hypercube. One way to see it is by applying Djoković's characterization of partial cubes from [5] asserting that G is a partial cube if and only if G is bipartite and for any edge $e = uv$ the subgraphs of G induced by $N_u(e)$ and by $N_v(e)$ are convex.

Let $\mathcal{F}$ be the partition of E_G that consists of the singletons corresponding to the K_2-blocks of G, while each cycle C_{2k} contributes k pairs of opposite edges to $\mathcal{F}$. The partition $\mathcal{F}$ thus contains $\ell(G)/2$ sets of cardinality 2, the other sets are singletons. Then, applying the main theorem of [18],

$$W(G) = \sum_{F \in \mathcal{F}} n_u(e) n_v(e),$$

where, for a given $F \in \mathcal{F}$, the edge $e = uv$ is an arbitrary, fixed edge from F. Similarly, applying the main theorem from [11],

$$Sz(G) = \sum_{F \in \mathcal{F}} |F| n_u(e) n_v(e)\,,$$

where again the edge $e = uv$ is an arbitrary but fixed edge from F. (These two results are instances of the so called *standard cut method*, see the recent survey [23] for more information on the method.) Therefore,

$$\begin{aligned} Sz(G) - W(G) &= \sum_{F \in \mathcal{F}} (|F| - 1) n_u(e) n_v(e) = \sum_{F \in \mathcal{F}, |F|=2} (|F| - 1) n_u(e) n_v(e) \\ &= \sum_{F \in \mathcal{F}, |F|=2} n_u(e) n_v(e) \geqslant \sum_{F \in \mathcal{F}, |F|=2} 2(n-2) \\ &= \frac{\ell(G)}{2} 2(n-2) = \ell(G)(n-2)\,. \end{aligned}$$

The above inequality turns into equality if and only if every cycle C of G is a 4-cycle and $\{n_u(e), n_v(e)\} = \{2, n-2\}$ holds for any edge e of C. That is, the equality holds if and only if every cycle of G is an end-block C_4.

4.3 Proof of Theorem 2.3

We next prove Theorem 2.3 which asserts the following: If $G \in \mathscr{C}_n \backslash \mathscr{C}_n^5$, then $Sz(G) - W(G) \geqslant 4n - 10$ with equality if and only if $G \in \mathcal{H}$.

If G contains an even cycle C_r, then by Lemma 3.3, we have $Sz(G) - W(G) \geqslant \frac{2nr^2 - r^3}{8} > 4n - 10$, the last inequality follows since $r \geqslant 4$. Hence, we may assume in what follows that the lengths of all the cycles in G are odd.

Let C_r be one of the longest odd cycles of G. Then clearly $r \geqslant 5$. If $r \geqslant 7$, then by Lemma 3.4, we have $Sz(G) - W(G) \geqslant \frac{(r-1)(r-3)(2n-r)}{8} > 4n - 10$, as desired. Thus, it suffices to consider the remaining case $r = 5$. Note that $G \notin \mathscr{C}_n^5$. Hence, there exist at least two cycles of length 5, say C, and C' in G. For convenience, let $C = v_1 v_2 \ldots v_5 v_1$ and $C' = u_1 u_2 \ldots u_5 u_1$. Since G is a cactus, we have $|V_C \cap V_{C'}| \leqslant 1$. Thus, we proceed by considering the following two possible cases.

Case 1. $|V_C \cap V_{C'}| = 0$.
In view of (11), we have

$$\sum_{x,y \in V_C} \pi(x,y) = \sum_{x,y \in V_{C'}} \pi(x,y) = 5. \tag{17}$$

As G is a cactus, we may without loss of generality assume that u_1, v_1 are the vertices in G such that $d_G(v_i, u_1) = \min_{x \in V_{C'}} d_G(v_i, x)$ and $d_G(u_j, v_1) = \min_{x \in V_C} d_G(u_j, x)$ for all $1 \leqslant i, j \leqslant 5$. Based on (10), we have $\sum_{x \in V_C} \pi(x, v_1) = 2$ and $\sum_{y \in V_{C'}} \pi(y, u_1) = 2$. Since u_1 and v_1 are cut vertices of G, using Lemma 3.2 we infer that

$$\begin{aligned} \sum_{x \in V_C, y \in V_{C'}} \pi(x,y) &= \sum_{x \in V_C, y \in V_{C'}} (\pi(x, v_1) + \pi(v_1, u_1) + \pi(u_1, y)) \\ &\geqslant 5 \sum_{x \in V_C} \pi(x, v_1) + 5 \sum_{y \in V_{C'}} \pi(u_1, y) \quad (\text{since } \pi(v_1, u_1) \geqslant 0\ ) \qquad (18) \\ &= 20, \qquad (19) \end{aligned}$$

where the equality in (18) holds if and only if $\pi(v_1, u_1) = 0$.

Consider every vertex pair $\{x, y\}$ with $x \in V_C$, $y \in V_G \backslash (V_C \cup V_{C'})$. Assume that v_{i_0} is the vertex of C such that $d_G(y, v_{i_0}) = \min_{z \in V_C} d_G(y, z)$. In view of Lemma 3.2, we have $\sum_{x \in V_C} \pi(x, y) = \sum_{x \in V_C} (\pi(x, v_{i_0}) + \pi(v_{i_0}, y)) \geqslant \sum_{x \in V_C} \pi(x, v_{i_0}) = 2$. Thus,

$$\sum_{x \in V_C} \sum_{y \in V_G \backslash (V_C \cup V_{C'})} \pi(x, y) \geqslant \sum_{y \in V_G \backslash (V_C \cup V_{C'})} 2 = 2(n - 10). \tag{20}$$

The equality in (20) holds if and only if $\sum_{x \in V_C} \pi(x, y) = 2$ for all $y \in V_G \backslash (V_C \cup V_{C'})$. Similarly, we can also obtain that $\sum_{x \in V_{C'}} \sum_{y \in V_G \backslash (V_C \cup V_{C'})} \pi(x, y) \geqslant 2(n - 10)$ with equality if and only if $\sum_{x \in V_{C'}} \pi(x, y) = 2$ for all $y \in V_G \backslash (V_C \cup V_{C'})$.

Note that $\pi(x, y) \geqslant 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash (V_C \cup V_{C'})$. Combined with (17), (19) and (20), it follows that

$$\begin{aligned} Sz(G) - W(G) &= \sum_{x, y \in V_C} \pi(x, y) + \sum_{x, y \in V_{C'}} \pi(x, y) + \sum_{x \in V_C, y \in V_{C'}} \pi(x, y) + \sum_{x \in V_C} \sum_{y \in V_G \backslash (V_C \cup V_{C'})} \pi(x, y) \\ &\quad + \sum_{x \in V_{C'}} \sum_{y \in V_G \backslash (V_C \cup V_{C'})} \pi(x, y) + \sum_{x, y \in V_G \backslash (V_C \cup V_{C'})} \pi(x, y) \\ &\geqslant 5 + 5 + 20 + 2(n - 10) + 2(n - 10) \qquad (21) \\ &= 4n - 10, \end{aligned}$$

where the equality in (21) holds if and only if $\pi(v_1, u_1) = 0$, $\sum_{x \in V_C} \pi(x, y) = 2$, and $\sum_{x \in V_{C'}} \pi(x, y) = 2$ for all $y \in V_G \backslash (V_C \cup V_{C'})$, as well as $\pi(x, y) = 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash (V_C \cup V_{C'})$.

Next, we show that if $Sz(G) - W(G) = 4n - 10$, then G contains just two cycles of length 5. Otherwise, G contains a third cycle C'' of length 5. Note that $|V_{C''} \cap V_C| \leqslant 1$ and $|V_{C''} \cap V_{C'}| \leqslant 1$. Then there exist two vertices $x, y \in V_{C''} \backslash (V_C \cup V_{C'})$ such that $d_{C''}(x, y) = 2$. Since C'' is an isometric cycle, by Lemma 3.1(ii) we have $\pi(x, y) \geqslant d_{C''}(x, y) - 1 = 1$. This is a contradiction to the fact $\pi(x, y) = 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash (V_C \cup V_{C'})$.

Let G_i (resp. G'_i) be the component of $G - E_C$ (resp. $G - E_{C'}$) that contains the vertex v_i (resp. u_i), $1 \leqslant i \leqslant r$. Then $|V_{G_1}| \geqslant 2$ and $|V_{G_i}| \geqslant 1$ for $2 \leqslant i \leqslant 5$. Suppose that there exist components G_i and G_j with $|V_{G_i}| \geqslant 2$ and $|V_{G_j}| \geqslant 2$, where v_i and v_j are not adjacent. Select arbitrary vertices $x \in V_{G_i} \backslash \{v_i\}$ and $y \in V_{G_j} \backslash \{v_j\}$. Because v_i and v_j are cut vertices of G, applying Lemma 3.2 we get

$$\pi(x, y) = \pi(x, v_i) + \pi(v_i, v_j) + \pi(v_j, y) \geqslant \pi(v_i, v_j) \geqslant 1,$$

which is a contradiction to the fact $\pi(x, y) = 0$ for every vertex pair $\{x, y\} \subseteq V_G \backslash (V_C \cup V_{C'})$. Combined with $|V_{G_1}| \geqslant 2$, we obtain $|V_{G_2}| = |V_{G_3}| = |V_{G_4}| = 1$ or $|V_{G_3}| = |V_{G_4}| = |V_{G_5}| = 1$. Similarly, we can also show that $|V_{G'_2}| = |V_{G'_3}| = |V_{G'_4}| = 1$ or $|V_{G'_3}| = |V_{G'_4}| = |V_{G'_5}| = 1$. In each of the above subcases, by direct calculation we have $Sz(G) - W(G) = 4n - 10$.

Hence, $Sz(G) - W(G) = 4n - 10$ if and only if $G \in \mathcal{H}$, where $\mathcal{H}$ is depicted in Fig. 1.

Case 2. $|V_C \cap V_{C'}| = 1$.
By a similar discussion as in the proof of Case 1, we can show that $Sz(G) - W(G) \geqslant 4n - 10$ with equality if and only if $G \in \mathcal{H}$; see Fig. 1 again.

This completes the proof of Theorem 2.3.

5 Proof of Theorem 2.4

In this section we prove Theorem 2.4. Since $n_u(e) + n_v(e) + n_0(e) = n$ for $e = uv \in E_G$, it is a routine to check that

$$Sz^*(G) - W(G) = Sz(G) - W(G) + \sum_{e \in E_G} \left(\frac{n_0(e)}{2} n - \frac{n_0^2(e)}{4} \right). \tag{22}$$

In order to prove the theorem, we first demonstrate a couple of claims.

Claim 1. *If $G \in \mathscr{C}_{n,k}$ is such that $Sz^*(G) - W(G)$ is as small as possible, then each cycle of G is an end-block.*

Proof. Suppose on the contrary that G contains a cycle $C_r = v_1v_2 \dots v_rv_1$ which is not an end-block. Let G_i be the component of $G - E_{C_r}$ containing the vertex v_i, $1 \leqslant i \leqslant r$. As C_r is not an end-block, $G - E_{C_r}$ contains two nontrivial components, say G_a and G_b. We construct a new graph G' as follows:

$$G' = G - \bigcup_{i=2}^{r} \{v_ix : x \in N_{G_i}(v_i)\} + \bigcup_{i=2}^{r} \{v_1x : x \in N_{G_i}(v_i)\}.$$

Then G' is in $\mathscr{C}_{n,k}$. By direct calculation (based on (3) and Lemma 3.2), one has

$$\begin{aligned} Sz(G) - W(G) &= \sum_{i=1}^{r} \sum_{x,y \in V_{G_i}} \pi(x,y) + \sum_{j=1}^{r} \sum_{y \in V_{G_j} \backslash \{v_j\}} \sum_{i \neq j} \pi(v_i, y) + \sum_{x,y \in V_{C_r}} \pi(x,y) \\ &\quad + \sum_{1 \leqslant i < j \leqslant r} \sum_{\substack{x \in V_{G_i} \backslash \{v_i\}, \\ y \in V_{G_j} \backslash \{v_j\}}} \pi(x,y) \\ &= \sum_{i=1}^{r} \sum_{x,y \in V_{G_i}} \pi(x,y) + \sum_{j=1}^{r} \sum_{y \in V_{G_j} \backslash \{v_j\}} \sum_{i \neq j} \pi(v_i, y) + \sum_{x,y \in V_{C_r}} \pi(x,y) \\ &\quad + \sum_{1 \leqslant i < j \leqslant r} \sum_{\substack{x \in V_{G_i} \backslash \{v_i\}, \\ y \in V_{G_j} \backslash \{v_j\}}} (\pi(x, v_i) + \pi(v_i, v_j) + \pi(v_j, y)). \end{aligned} \tag{23}$$

Similarly, we have

$$\begin{aligned} Sz(G') - W(G') &= \sum_{i=1}^{r} \sum_{x,y \in V_{G_i}} \pi(x,y) + \sum_{j=1}^{r} \sum_{y \in V_{G_j} \backslash \{v_j\}} \sum_{i=2}^{r} \pi(v_i, y) + \sum_{x,y \in V_{C_r}} \pi(x,y) \\ &\quad + \sum_{1 \leqslant i < j \leqslant r} \sum_{\substack{x \in V_{G_i} \backslash \{v_i\}, \\ y \in V_{G_j} \backslash \{v_j\}}} (\pi(x, v_1) + \pi(v_1, y)). \end{aligned} \tag{24}$$

For convenience, set

$$\Delta_1 := (Sz(G) - W(G)) - (Sz(G') - W(G'))$$

and

$$\Delta_2 := \sum_{e \in E_G} \left(\frac{n_0(e)}{2} n - \frac{n_0^2(e)}{4} \right) - \sum_{e \in E_{G'}} \left(\frac{n_0(e)}{2} n - \frac{n_0^2(e)}{4} \right).$$

As $|V_{G_a}|, |V_{G_b}| \geqslant 2$, we have $|V_{G_a} \backslash \{v_a\}| \geqslant 1$, $|V_{G_b} \backslash \{v_b\}| \geqslant 1$. In view of (23) and (24), we have

$$\begin{aligned} \Delta_1 &= \sum_{1 \leqslant i < j \leqslant r} \sum_{\substack{x \in V_{G_i} \backslash \{v_i\}, \\ y \in V_{G_j} \backslash \{v_j\}}} \pi(v_i, v_j) \\ &\geqslant \sum_{\substack{x \in V_{G_a} \backslash \{v_a\}, \\ y \in V_{G_b} \backslash \{v_b\}}} \pi(v_a, v_b) \\ &\geqslant \pi(v_a, v_b). \end{aligned} \tag{25}$$

In what follows, we consider two possible cases according to the parity of r.

Case 1. r is even.
In this case, on the one hand, it is routine to check that $\Delta_2 = 0$. On the other hand, by Lemma 3.1(i), we have $\pi(v_a, v_b) \geqslant d_{C_r}(v_a, v_b) \geqslant 1$. In view of (25), we have $\Delta_1 \geqslant \pi(v_a, v_b) \geqslant 1$. Combined with (22), we have $(Sz^*(G) - W(G)) - (Sz^*(G') - W(G')) = \Delta_1 + \Delta_2 > 0$, a contradiction to the choice of G.

Case 2. r is odd.
Since C_r is an isometric cycle, by Lemma 3.1(ii) we have $\pi(v_a, v_b) \geqslant d_G(v_a, v_b) - 1 \geqslant 0$. In view of (25), we have $\Delta_1 \geqslant 0$. Next, we show that $\Delta_2 > 0$.

For the graph G, it is straightforward to check that

$$\begin{aligned}\sum_{e\in E_G}\left(\frac{n_0(e)}{2}n - \frac{n_0^2(e)}{4}\right) &= \sum_{i=1}^{r}\sum_{e\in E_{G_i}}\left(\frac{n_0(e)}{2}n - \frac{n_0^2(e)}{4}\right) + \sum_{e\in E_{C_r}}\left(\frac{n_0(e)}{2}n - \frac{n_0^2(e)}{4}\right)\\ &= \sum_{i=1}^{r}\sum_{e\in E_{G_i}}\left(\frac{n_0(e)}{2}n - \frac{n_0^2(e)}{4}\right) + \frac{n^2}{2} - \frac{\sum_{i=1}^{r}|V_{G_i}|^2}{4}. \qquad (26)\end{aligned}$$

Similarly, for the graph G', we have

$$\sum_{e\in E_{G'}}\left(\frac{n_0(e)}{2}n - \frac{n_0^2(e)}{4}\right) = \sum_{i=1}^{r}\sum_{e\in E_{G_i}}\left(\frac{n_0(e)}{2}n - \frac{n_0^2(e)}{4}\right) + \frac{n^2}{2} - \frac{r-1+(n-r+1)^2}{4}. \qquad (27)$$

Then, together with (26) and (27), it follows that

$$\Delta_2 = \frac{r-1+(n-r+1)^2}{4} - \frac{\sum_{i=1}^{r}|V_{G_i}|^2}{4}.$$

Suppose that there exist two components G_p, G_q satisfying $|V_{G_p}|, |V_{G_q}| \geqslant 2$ for $1 \leqslant p, q \leqslant r$. Let $\Delta_2' = \frac{r-1+(n-r+1)^2}{4} - \frac{\sum_{i\neq p,q}|V_{G_i}|^2}{4} - (|V_{G_p}| + |V_{G_q}| - 1)^2 - 1$. Then,

$$\Delta_2 - \Delta_2' = (|V_{G_p}| + |V_{G_q}| - 1)^2 + 1 - |V_{G_p}|^2 - |V_{G_q}|^2 = 2(|V_{G_p}| - 1)(|V_{G_q}| - 1) > 0.$$

The last inequality follows because $|V_{G_p}|, |V_{G_q}| \geqslant 2$. Thus, Δ_2 attains its minimum 0 if and only if $|V_{G_1}| = n - r + 1$, $|V_{G_i}| = 1$ for $2 \leqslant i \leqslant r$ up to isomorphism. Bearing in mind that $|V_{G_a}|, |V_{G_b}| \geqslant 2$, one can easily obtain that $\Delta_2 > 0$.

In view of (22), we have $(Sz^*(G) - W(G)) - (Sz^*(G') - W(G')) = \Delta_1 + \Delta_2 > 0$, i.e., $Sz^*(G') - W(G') < Sz^*(G) - W(G)$, which is a contradiction to the choice of G.

This completes the proof of Claim 1. □

Let each cycle of a graph $G_1 \in \mathscr{C}_{n,k}$ be an end-block cycle. Let C_p, C_q be two disjoint cycles of G_1 containing cut vertices u_0, v_0, respectively. Construct a new graph G_2 as follows:

$$G_2 = G_1 - \{v_0 x : x \in N_{C_q}(v_0)\} + \{u_0 x : x \in N_{C_q}(v_0)\}.$$

In other words, if v_0' and v_0'' are the neighbors of v_0 on C_q, then G_2 is obtained from G_1 by removing the edges $v_0 v_0'$ and $v_0 v_0''$ and adding the edges $u_0 v_0'$ and $u_0 v_0''$. Then G_2 is in $\mathscr{C}_{n,k}$ and each cycle of G_2 is also an end-block. Here, we show that this graph transformation keeps the value of $Sz^*(G) - W(G)$ unchanged.

Claim 2. *Let G_1 and G_2 be the graphs as defined above. Then $Sz^*(G_1) - W(G_1) = Sz^*(G_2) - W(G_2)$.*

Proof. First, we show that $\pi(x, y) = 0$ for every pair of cut vertices $x, y \in V_{G_1}$. For such cut vertices $x, y \in V_{G_1}$, there is a shortest path P_t connecting x and y. It is routine to check that for any edge $e \in E_{P_t}$,

$\mu_{x,y}(e) = 1$. Recall that each cycle of G_1 is an end-block. Assume that x_0 is a cut vertex of the cycle C_r in G_1. Then, for an edge $e = uv \in E_{C_r}$, we have

$$x, y \in \begin{cases} N_u(e), & \text{if } x_0 \in N_u(e); \\ N_v(e), & \text{if } x_0 \in N_v(e); \\ N_0(e), & \text{if } x_0 \in N_0(e). \end{cases}$$

This implies that $\mu_{x,y}(e) = 0$. For the remaining cut edge $e = uv \in E_{G_1} \backslash E_{P_t}$, it is routine to check that $x, y \in N_u(e)$ or $x, y \in N_v(e)$, which implies that $\mu_{x,y}(e) = 0$. Thus, $\pi(x, y) = \sum_{e \in E_{G_1}} \mu_{x,y}(e) - d_{G_1}(x, y) = \sum_{e \in E_{P_t}} \mu_{x,y}(e) - t + 1 = 0$.

Bearing in mind that u_0, v_0 are cut vertices, we have $\pi(u_0, v_0) = 0$. For convenience, denote by $V_1 = V_{G_1} \backslash (V_{C_p} \cup V_{C_q})$. In view of (3), one can obtain that

$$\begin{aligned} Sz(G_1) - W(G_1) &= \sum_{x,y \in V_{C_p}} \pi(x, y) + \sum_{x,y \in V_{C_q}} \pi(x, y) + \sum_{x,y \in V_1} \pi(x, y) + \sum_{x \in V_{C_p}, y \in V_1} \pi(x, y) \\ &\quad + \sum_{x \in V_{C_p}, y \in V_{C_q}} \pi(x, y) + \sum_{x \in V_{C_q}, y \in V_1} \pi(x, y) \\ &= \sum_{x,y \in V_{C_p}} \pi(x, y) + \sum_{x,y \in V_{C_q}} \pi(x, y) + \sum_{x,y \in V_1} \pi(x, y) + \sum_{x \in V_{C_p}, y \in V_1} \pi(x, y) \\ &\quad + \sum_{x \in V_{C_p}, y \in V_{C_q}} (\pi(x, u_0) + \pi(v_0, y)) + \sum_{x \in V_{C_q}, y \in V_1} (\pi(x, v_0) + \pi(v_0, y)) \end{aligned}$$

and

$$\begin{aligned} Sz(G_2) - W(G_2) &= \sum_{x,y \in V_{C_p}} \pi(x, y) + \sum_{x,y \in V_{C_q}} \pi(x, y) + \sum_{x,y \in V_1} \pi(x, y) + \sum_{x \in V_{C_p}, y \in V_1} \pi(x, y) \\ &\quad + \sum_{x \in V_{C_p}, y \in V_{C_q}} (\pi(x, u_0) + \pi(u_0, y)) + \sum_{x \in V_{C_q}, y \in V_1} (\pi(x, u_0) + \pi(u_0, y)). \end{aligned}$$

Thus, based on (5) or (10), we have

$$(Sz(G_1) - W(G_1)) - (Sz(G_2) - W(G_2)) \quad = \quad q \sum_{y \in V_1} (\pi(v_0, y) - \pi(u_0, y)). \tag{28}$$

Now, we prove that $\pi(v_0, y) = \pi(u_0, y)$ for all $y \in V_1$. If $y \in V_{C_l}$ for some cycle C_l in G_1 except C_p, C_q, and denote the unique cut vertex of C_l by w. By Lemma 3.2, we have $\pi(v_0, y) = \pi(v_0, w) + \pi(w, y) = \pi(w, y)$ and $\pi(u_0, y) = \pi(u_0, w) + \pi(w, y) = \pi(w, y)$. Thus, $\pi(v_0, y) = \pi(u_0, y)$. Otherwise, y isn't contained in any cycle. Then y is either a cut vertex of G_1 or $d_{G_1}(y) = 1$. In this case, we have $\pi(v_0, y) = \pi(u_0, y) = 0$.

Therefore, by (28), we know that $Sz(G_1) - W(G_1) = Sz(G_2) - W(G_2)$. Note that

$$\sum_{e \in E_{G_1}} \left(\frac{n_0(e)}{2} n - \frac{n_0^2(e)}{4} \right) = \sum_{e \in E_{G_2}} \left(\frac{n_0(e)}{2} n - \frac{n_0^2(e)}{4} \right).$$

In view of (22), we have $Sz^*(G_1) - W(G_1) = Sz^*(G_2) - W(G_2)$, as desired. □

Now all is ready for the proof of Theorem 2.4.

Choose a graph G in $\mathscr{C}_{n,k}$ such that $Sz^*(G) - W(G)$ is as small as possible. By Claim 1, each cycle of G is an end-block cycle. By a repeated application of the construction of Claim 2, we may assume that all the cycles of G have a common vertex. Denote the common vertex by u_0. For convenience, let $C_{r_1}, C_{r_2}, \ldots, C_{r_p}$ be the even cycles and $C_{t_1}, C_{t_2}, \ldots, C_{t_q}$ the odd cycles of G. Then $p + q = k$. In what follows, we first determine the lower bound on $Sz(G) - W(G)$.

For every vertex pair $\{x,y\} \subseteq V_{C_{r_i}}$, $1 \leqslant i \leqslant p$, by (6) we have

$$\sum_{x,y\in V_{C_{r_i}}} \pi(x,y) = \sum_{x,y\in V_{C_{r_i}}} d_{C_{r_i}}(x,y) = \frac{r_i^3}{8}. \tag{29}$$

Consider every vertex pair $\{x,\,y\}$ with $x \in V_{C_{r_i}}\backslash\{u_0\}$, $y \in V_{C_{r_j}}\backslash\{u_0\}$, $i \neq j$. Since u_0 is a cut vertex of G, by (5) and Lemma 3.2 we have $\pi(x,y) = \pi(x,u_0) + \pi(u_0,y) = d_{C_{r_i}}(x,u_0) + d_{C_{r_j}}(u_0,y)$. Then

$$\begin{aligned}\sum_{x\in V_{C_{r_i}}\backslash\{u_0\}}\ \sum_{y\in V_{C_{r_j}}\backslash\{u_0\}} \pi(x,y) &= \sum_{x\in V_{C_{r_i}}\backslash\{u_0\}}\ \sum_{y\in V_{C_{r_j}}\backslash\{u_0\}} (d_{C_{r_i}}(x,u_0) + d_{C_{r_j}}(u_0,y)) \\ &= \frac{r_i^2(r_j-1)}{4} + \frac{r_j^2(r_i-1)}{4}. \end{aligned} \tag{30}$$

For convenience, denote by $n_0 := r_1 + r_2 + \cdots + r_p$. Then, in view of (30), we have

$$\begin{aligned}\sum_{1\leqslant i<j\leqslant p}\ \sum_{x\in V_{C_{r_i}}\backslash\{u_0\}}\ \sum_{y\in V_{C_{r_j}}\backslash\{u_0\}} \pi(x,y) &= \sum_{1\leqslant i<j\leqslant p}\left(\frac{r_i^2(r_j-1)}{4} + \frac{r_j^2(r_i-1)}{4}\right) \\ &= \sum_{1\leqslant i\leqslant p} \frac{r_i^2(n_0 - r_i - p + 1)}{4}. \end{aligned} \tag{31}$$

Consider the remaining vertex pairs $\{x,\,y\}$ with $x \in \bigcup_{i=1}^p V_{C_{r_i}}$ and $y \in V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})$. Then from (5) and Lemma 3.2 we get $\pi(x,y) = \pi(x,u_0) + \pi(u_0,y) \geqslant \pi(x,u_0) = d_G(x,u_0)$. Consequently,

$$\begin{aligned}\sum_{x\in\bigcup_{i=1}^p V_{C_{r_i}}}\ \sum_{y\in V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})} \pi(x,y) &\geqslant \sum_{x\in\bigcup_{i=1}^p V_{C_{r_i}}}\ \sum_{y\in V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})} d_G(x,u_0) \qquad (\text{since } \pi(u_0,y)\geqslant 0) \\ &= (n - n_0 + p - 1) \sum_{x\in\bigcup_{i=1}^p V_{C_{r_i}}} d_G(x,u_0) \\ &= \frac{(n - n_0 + p - 1)\sum_{i=1}^p r_i^2}{4}. \end{aligned} \tag{32}$$

Note that $\pi(x,y) \geqslant 0$ for every vertex pair $\{x,y\} \subseteq V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})$. Together with (29), (31) and (32), it follows that

$$\begin{aligned} Sz(G) - W(G) &= \sum_{i=1}^p \sum_{x,y\in V_{C_{r_i}}} \pi(x,y) + \sum_{1\leqslant i<j\leqslant p}\ \sum_{x\in V_{C_{r_i}}\backslash\{u_0\}}\ \sum_{y\in V_{C_{r_j}}\backslash\{u_0\}} \pi(x,y) \\ &\quad + \sum_{x\in\bigcup_{i=1}^p V_{C_{r_i}}}\ \sum_{y\in V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})} \pi(x,y) + \sum_{x,y\in V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})} \pi(x,y) \\ &\geqslant \sum_{i=1}^p \frac{r_i^3}{8} + \sum_{i=1}^p \frac{r_i^2(n_0 - r_i - p + 1)}{4} + \sum_{i=1}^p \frac{(n - n_0 + p - 1)r_i^2}{4} \qquad (33) \\ &= \sum_{i=1}^p \frac{-r_i^3 + 2nr_i^2}{8} \\ &\geqslant p(4n-8), \qquad (\text{since } r_i \geqslant 4) \qquad (34) \end{aligned}$$

where the equality in (33) holds if and only if $\pi(u_0,y) = 0$ for any vertex $y \in V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})$ and $\pi(x,y) = 0$ for every vertex pair $\{x,y\} \subseteq V_G\backslash(\bigcup_{i=1}^p V_{C_{r_i}})$; whereas the equality in (34) holds if and only if $r_1 = r_2 = \cdots = r_p = 4$.

Next, we consider the value of $\sum_{e\in E_G}\left(\frac{n_0(e)}{2}n-\frac{n_0^2(e)}{4}\right)$. Bearing in mind that, for an edge e, if it is not contained in any odd cycle, then $n_0(e)=0$. So we consider that it is in some odd cycle, say C_{t_i}. One has $n_0(e)=n-t_i+1$ if $u_0\in N_0(e)$ and $n_0(e)=1$ otherwise. Thus, we have

$$\begin{aligned}
\sum_{e\in E_G}\left(\frac{n_0(e)}{2}n-\frac{n_0^2(e)}{4}\right) &= \sum_{i=1}^{q}\sum_{e\in E_{C_{t_i}}}\left(\frac{n_0(e)}{2}n-\frac{n_0^2(e)}{4}\right)\\
&= \sum_{i=1}^{q}\left(\frac{n^2}{2}-\frac{t_i-1+(n-t_i+1)^2}{4}\right)\\
&\geqslant \frac{q(n^2+4n-6)}{4}, \qquad \text{(since } t_i\geqslant 3\text{)} \qquad (35)
\end{aligned}$$

where the equality in (35) holds if and only if $t_1=t_2=\cdots=t_q=3$.

Together with (22), (34) and (35), it follows that

$$\begin{aligned}
Sz^*(G)-W(G) &= Sz(G)-W(G)+\sum_{e\in E_G}\left(\frac{n_0(e)}{2}n-\frac{n_0^2(e)}{4}\right)\\
&\geqslant p(4n-8)+\frac{q(n^2+4n-6)}{4}. \qquad (36)
\end{aligned}$$

Based on (34) and (35) we obtain that the equality in (36) holds if and only if $r_1=r_2=\cdots=r_p=4$ and $t_1=t_2=\cdots=t_q=3$.

Now, we give the proofs of (i) and (ii), respectively.

(i) If $4\leqslant n\leqslant 9$, then by direct calculation, we have $4n-8>\frac{n^2+4n-6}{4}$. In view of (36), one has $Sz^*(G)-W(G)\geqslant\frac{k(n^2+4n-6)}{4}$ with equality if and only if $t_1=t_2=\cdots=t_q=3$ and $q=k$, i.e., G is a graph satisfying each block of G being a K_2 or a C_3 and each cycle of G being an end-block. Thus, (i) holds.

(ii) If $n\geqslant 10$, then it is routine to check that $4n-8<\frac{n^2+4n-6}{4}$. In view of (36), one has $Sz^*(G)-W(G)\geqslant k(4n-8)$ with equality if and only if $r_1=r_2=\cdots=r_p=4$ and $p=k$, i.e., G is a graph satisfying each block of G being a K_2 or an end-block C_4. Hence, (ii) holds.

6 Concluding remarks

In this paper we have established lower bounds on the difference between the (revised) Szeged index and Wiener index of graphs in $\mathscr{C}_n$ (resp. $\mathscr{C}_{n,k}$). To conclude the paper we state two corollaries and two problems.

The following result follows from Theorem 2.1 and can also be deduced from [4, Theorems 3.1 and 3.2].

Corollary 6.1. *Let G be an n-vertex unicyclic graph with circumference at least 4. Then*

$$Sz(G)-W(G)\geqslant 2n-5$$

with equality if and only if G is a graph which is composed from C_5 and either a tree on $n-4$ vertices rooted at a vertex of C_5, or two trees rooted at two adjacent vertices of C_5.

Similarly, the following result follows from Theorem 2.4 and can alternatively be obtained from [4, Theorems 3.2 and 4.3].

Corollary 6.2. *Let G be a unicyclic graph on $n\geqslant 4$ vertices.*

(i) *If* $4 \leqslant n \leqslant 9$, *then*

$$Sz^*(G) - W(G) \geqslant \frac{n^2 + 4n - 6}{4}$$

with equality if and only if G *is composed from* C_3 *and a tree* T *on* $n-2$ *vertices sharing a single vertex.*

(ii) *If* $n \geqslant 10$, *then*

$$Sz^*(G) - W(G) \geqslant 4n - 8$$

with equality if and only if G *is composed from* C_4 *and a tree* T *on* $n-3$ *vertices sharing a single vertex.*

In Subsection 4.2 we have demonstrated that the cut method provides an efficient method to bound the difference between $Sz(G)$ and $W(G)$ when G is a bipartite cactus. Extensions of the cut method to general graphs are known, see [19, 23], hence the following problem appears natural in this context.

Problem 6.3. *Can the standard cut method for the Szeged index be extended to general graphs or to arbitrary cacti?*

We have identified the graphs G from $\mathscr{C}_{n,k}$ such that $Sz^*(G) - W(G)$ attains its minimum value. Hence we pose:

Problem 6.4. *Determine the second minimum value on the difference between the revised Szeged index and the Wiener index among the graphs from* $\mathscr{C}_{n,k}$.

Acknowledgments

The authors would like to express their sincere gratitude to all of the referees for their insightful comments and suggestions, which led to a number of improvements to this paper.

References

[1] M. Aouchiche, P. Hansen, On a conjecture about the Szeged index, European J. Combin. 31 (2010) 1662–1666.

[2] M. Bonamy, M. Knor, B. Lužar, A. Pinlou, R. Škrekovski, On the difference between the Szeged and Wiener index, Appl. Math. Comput., 312 (2017) 202-213.

[3] L.L. Chen, X.L. Li, M.M. Liu, I. Gutman, On a relation between the Szeged and the Wiener indices of bipartite graphs, Trans. Comb. 1 (2012) 43-49.

[4] L.L. Chen, X.L. Li, M.M. Liu, The (revised) Szeged index and the Wiener index of a nonbipartite graph, European J. Combin. 36 (2014) 237–246.

[5] D. Djoković, Distance preserving subgraphs of hypercubes, J. Combin. Theory Ser. B 14 (1973) 263–267.

[6] A.A. Dobrymin, R. Entriger, I. Gutman, Wiener index of trees: theory and applications, Acta Appl. Math. 66 (2001) 211–249.

[7] A.A. Dobrymin, I. Gutman, S. Klavžar, P. Žigert, Wiener index of hexagonal systems, Acta Appl. Math. 72 (2002) 247–294.

[8] A. Dobrynin, I. Gutman, Solving a problem connected with distances in graphs, Graph Theory Notes N. Y. 28 (1995) 21–23.

[9] R.C. Entringer, D.E. Jackson, D.A. Snyder, Distance in graphs, Czech. Math. J. 26 (1976) 283–296.

[10] I. Gutman, A formula for the Wiener number of trees and its extension to graphs containing cycles, Graph Theory Notes N. Y. 27 (1994) 9–15.

[11] I. Gutman, S. Klavžar, An algorithm for the calculation of the Szeged index of benzenoid hydrocarbons, J. Chem. Inf. Comput. Sci. 35 (1995) 1011–1014.

[12] I. Gutman, S. Klavžar, B. Mohar (Eds.), Fifty years of the Wiener index, MATCH Commun. Math. Comput. Chem. 35 (1997) 1–259.

[13] I. Gutman, O.E. Polansky, Mathematical Concepts in Organic Chemistry, Springer, Berlin, 1986.

[14] P. Hansen, Computers and conjectures in chemical graph theory, Plenary talk in the International Conference on Mathematical Chemistry, August 4-7, 2010, Xiamen, China.

[15] A. Ilić, Note on PI and Szeged indices, Math. Comput. Modelling 52 (2010) 1570–1576.

[16] S. Ji, M. Liu, J. Wu, A lower bound of revised Szeged index of bicyclic graphs, Appl. Math. Comput. 316 (2018) 480-487.

[17] H. Khodashenas, M.J. Nadjafi-Arani, A.R. Ashrafi, I. Gutman, A new proof of the Szeged-Wiener theorem, Kragujevac J. Math. 35 (2011) 165–172.

[18] S. Klavžar, I. Gutman, B. Mohar, Labeling of benzenoid systems which reflects the vertex–distance relations, J. Chem. Inf. Comput. Sci. 35 (1995) 590–593.

[19] S. Klavžar, On the canonical metric representation, average distance, and partial Hamming graphs, European J. Combin. 27 (2006) 68–73.

[20] S. Klavžar, M.J. Nadjafi-Arani, Wiener index versus Szeged index in networks, Discrete Appl. Math. 161 (2013) 1150–1153.

[21] S. Klavžar, M.J. Nadjafi-Arani, On the difference between the revised Szeged index and the Wiener index, Discrete Math. 333 (2014) 28–34.

[22] S. Klavžar, M.J. Nadjafi-Arani, Improved bounds on the difference between the Szeged index and the Wiener index of graphs, European J. Combin. 39 (2014) 148–156.

[23] S. Klavžar, M.J. Nadjafi-Arani, Cut method: update on recent developments and equivalence of independent approaches, Curr. Org. Chem. 19 (2015) 348–358.

[24] S. Klavžar, A. Rajapakse, I. Gutman, The Szeged and the Wiener index of graphs, Appl. Math. Lett. 9 (1996) 45–49.

[25] M. Knor, R. Škrekovski, A. Tepeh, Orientations of graphs with maximum Wiener index, Discrete Appl. Math. 211 (2016) 121–129.

[26] X.L. Li, M.M. Liu, Bicyclic graphs with maximal revised Szeged index, Discrete Appl. Math. 161 (2013) 2527–2531.

[27] S.C. Li, Y.B. Song, On the sum of all distances in bipartite graphs, Discrete Appl. Math. 169 (2014) 176–185.

[28] S.C. Li, H.H. Zhang, Proofs of three conjectures on the quotients of the (revised) Szeged index and the Wiener index and beyond, Discrete Math. 340 (2017) 311–324.

[29] H. Lei, T. Li, Y. Shi, H. Wang, Wiener polarity index and its generalization in trees, MATCH Commun. Math. Comput. Chem. 78(1) (2017) 199-212.

[30] M.J. Nadjafi-Arani, H. Khodashenas, A.R. Ashrafi, On the differences between Szeged and Wiener indices of graphs, Discrete Math. 311 (2011) 2233–2237.

[31] M.J. Nadjafi-Arani, H. Khodashenas, A.R. Ashrafi, Graphs whose Szeged and Wiener numbers differ by 4 and 5, Math. Comput. Modelling 55 (2012) 1644–1648.

[32] T. Pisanski, M. Randić, Use of the Szeged index and the revised Szeged index for measuring network bipartivity, Discrete Appl. Math. 158 (2010) 1936–1944.

[33] M. Randić, On generalization of Wiener index for cyclic structures, Acta Chim. Slov. 49 (2002) 483–496.

[34] S. Simić, I. Gutman, V. Baltić, Some graphs with extremal Szeged index, Math. Slovaca 50 (2000) 1–15.

[35] Z.K. Tang, H.Y. Deng, The (n, n)-graphs with the first three extremal Wiener indices, J. Math. Chem. 43 (2008) 60–74.

[36] H. Wiener, Structural determination of paraffin boiling points, J. Am. Chem. Soc. 69 (1947) 17–20.

[37] R.D. Xing, B. Zhou, On the revised Szeged index, Discrete Appl. Math. 159 (2011) 69–78.

[38] R.D. Xing, B. Zhou, On Wiener and hyper-Wiener indices of graphs with fixed number of cut vertices, Util. Math. 99 (2016) 121–130.

[39] H.H. Zhang, S.C. Li, L.F. Zhao, On the further relation between the (revised) Szeged index and the Wiener index of graphs, Discrete Appl. Math. 206 (2016) 152–164.