

# On the relationship between variable Wiener index and variable Szeged index<sup>☆</sup>



Stijn Cambie<sup>a,b</sup>, John Haslegrave<sup>c,d,\*</sup>

<sup>a</sup> Department of Mathematics, Radboud University Nijmegen, Postbus 9010, Nijmegen 6500 GL, the Netherlands

<sup>b</sup> Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea

<sup>c</sup> Mathematics Institute, University of Warwick, Coventry, UK

<sup>d</sup> Mathematical Institute, University of Oxford, UK

## ARTICLE INFO

### Article history:

Received 31 August 2021

Revised 6 June 2022

Accepted 8 June 2022

Available online 14 June 2022

MSC:

05C09

05C12

05C35

### Keywords:

Variable Wiener index

Variable Szeged index

Topological indices

## ABSTRACT

We resolve two conjectures of Hriňáková et al. (2019)[10] concerning the relationship between the variable Wiener index and variable Szeged index for a connected, non-complete graph, one of which would imply the other. The strong conjecture is that for any such graph there is a critical exponent in  $(0,1]$ , below which the variable Wiener index is larger and above which the variable Szeged index is larger. The weak conjecture is that the variable Szeged index is always larger for any exponent exceeding 1. They proved the weak conjecture for bipartite graphs, and the strong conjecture for trees.

In this note we disprove the strong conjecture, although we show that it is true for almost all graphs, and for bipartite and block graphs. We also show that the weak conjecture holds for all graphs by proving a majorization relationship.

© 2022 The Author(s). Published by Elsevier Inc.  
This is an open access article under the CC BY license  
(<http://creativecommons.org/licenses/by/4.0/>)

## 1. Introduction

Topological indices of graphs originate from chemical graph theory, and are descriptors of organic molecules that depend only on the graph of atoms and bonds and not on their physical arrangement. The oldest and most important of these is the Wiener index, or total distance, which was introduced by Wiener in 1947 and shown to accurately predict alkane boiling points [21]. Here, as is usual, the index is applied to the unweighted graph of carbon-carbon bonds, neglecting hydrogen atoms. Wiener subsequently established correlations between the Wiener index and other physical quantities (see e.g. Wiener [22]). The Wiener index was rediscovered by Rouvaray [19] in the 1970s, and applied to more general chemical graphs. It was independently introduced to the mathematics literature for trees by Zelinka [23], and subsequently extended to general graphs by Entringer et al. [4]. Since the late 1970s, it has seen an explosion of interest in both fields; see e.g. the survey article [2] from the mathematical side, and the collection [6] from the chemical side. Its importance has led to the study of many other topological indices, such as the Randić index [18], first and second Zagreb indices [7,8], and the Hosoya index [9].

<sup>☆</sup> This work has been supported by a Vidi Grant of the Netherlands Organization for Scientific Research (NWO), grant number 639.032.614, and UK Research and Innovation Future Leaders Fellowship MR/S016325/1.

\* Corresponding author at: Mathematical Institute, University of Oxford, UK.

E-mail address: [j.haslegrave@cantab.net](mailto:j.haslegrave@cantab.net) (J. Haslegrave).

Among topological indices, the Wiener index has a particularly simple and natural definition: it is the sum, over all unordered pairs of vertices, of the graph distance between those vertices (hence the alternative name of “total distance”; it is linearly related to the average distance when the order is fixed). We define the Wiener index only when the graph is connected, so all distances are finite; this is certainly the case for molecular graphs.

As a special case of the alternative method to compute  $W(G)$  by use of a canonical metric representation of a graph, as defined in Klavžar [12], it is well known that for a tree  $T$  one can compute the total distance as a sum of a certain quantity over the edges instead of the sum of distances over all pairs of vertices. For each edge  $e = uv$ , write  $n_u(v)$  for the number of vertices strictly closer to  $u$  than to  $v$  (including  $u$  itself), and  $n_v(u)$  for the number strictly closer to  $v$  than to  $u$ . Then

$$W(T) = \sum_{uv \in E(T)} n_x(y) \cdot n_y(x). \tag{1}$$

This fact, which follows from a simple double-counting, was observed by Wiener in his original paper [21], where only trees were considered. However, Eq. (1) fails to hold for most other graphs, owing to the fact that shortest paths are typically not unique.

For a general graph  $G$ , the right-hand side of (1) denotes a different topological index, introduced by Gutman [5] and subsequently called the Szeged index:

$$Sz(G) := \sum_{uv \in E(G)} n_u(v) \cdot n_v(u).$$

In general it is not necessarily true that  $Sz(G) = W(G)$ , but the inequality  $Sz(G) \geq W(G)$  holds for all connected graphs [13]. The graphs where equality occurs were classified by Dobrynin and Gutman [3]; these are precisely the block graphs (also known as clique trees), that is, graphs where every maximal biconnected subgraph is complete. The difference between the Szeged and Wiener indices can take any non-negative integer value other than 1 or 3 [16].

The study of topological indices encompasses variable forms, which depend not only on the graph  $G$  but on a parameter often denoted  $\alpha$ ; the terms which make up the index in question are raised to the power  $\alpha$  before summing. This allows the sensitivity of the topological index to extreme values to be altered. For example, the variable form of the Randić index was introduced in Bollobás and Erdős [1], and the variable Zagreb indices in Miličević and Nikolić [15].

Hriňáková et al. [10] generalized the Wiener and Szeged indices to analogous variable forms, and studied the relation between the variable Wiener index and variable Szeged index. For a parameter  $\alpha$ , these two graph parameters are defined as respectively

$$W^\alpha(G) := \sum_{u,v \in V} d(u,v)^\alpha \quad \text{and} \quad Sz^\alpha(G) := \sum_{uv \in E(G)} (n_u(v) \cdot n_v(u))^\alpha.$$

They gave the following two conjectures (in these conjectures, and the rest of this paper, graphs are tacitly assumed to be connected).

**Conjecture 1.1** ([10, Conjecture 5]). *For every non-complete graph  $G$  there is a constant  $\alpha_G \in (0, 1]$  such that*

$Sz^\alpha(G) > W^\alpha(G)$	if $\alpha > \alpha_G$ ;
$Sz^\alpha(G) = W^\alpha(G)$	if $\alpha = \alpha_G$ ;
$Sz^\alpha(G) < W^\alpha(G)$	if $\alpha < \alpha_G$ .

**Conjecture 1.2** ([10, Conjecture 6]). *For every non-complete graph  $G$  and every  $\alpha > 1$  we have  $Sz^\alpha(G) > W^\alpha(G)$ .*

Note that Conjecture 1.1, since it requires  $\alpha_G \leq 1$ , implies Conjecture 1.2. Complete graphs are excluded from the conjectures since if  $G$  is complete we have  $Sz^\alpha(G) = \binom{|G|}{2} = W^\alpha(G)$  for every  $\alpha$ .

If we consider  $(d(u,v))_{u,v \in V(G)}$  and  $(n_u(v) \cdot n_v(u))_{uv \in E(G)}$  as two sequences (of positive integers)  $(x_i)_i$  and  $(y_i)_i$  derived from a graph, then we are comparing the sums  $\sum f(x_i)$  and  $\sum f(y_i)$  where  $f(x) = x^\alpha$ . A famous inequality related to such sums is Karamata’s inequality [11], which is a generalization of the better known inequality of Jensen. Hriňáková et al. [10] used Karamata’s inequality to derive inequalities between  $Sz^\alpha(G)$  and  $W^\alpha(G)$  for trees and bipartite graphs (when  $\alpha > 1$ ), proving that the strong conjecture holds for trees and the weak conjecture for bipartite graphs.

In Section 2, we prove that the weak conjecture is indeed true by a small adaptation of Karamata’s inequality combined with a majorization result. In Section 4 we describe a family of counterexamples to the strong conjecture. However, we are able to prove the strong conjecture for some special families of graphs in Section 3. In particular we show that it is true for bipartite graphs, block graphs, and all graphs in a moderately sparse regime. We show that it is true for almost all graphs in a very strong sense: we prove that it is true with high probability for the random graph  $G(n, m)$  (that is, a uniformly random choice from the set of all connected graphs with  $n$  vertices and  $m = m(n)$  edges), even if  $m$  is chosen adversarially.

Since this work was submitted, a different and independent proof of the weak conjecture by Vukičević and Bulatović has appeared [20].

## 2. Proof of the weak conjecture

### 2.1. An adaptation of Karamata's inequality

Karamata's inequality generalises Jensen's inequality for convex functions to majorizing sequences. In this article, we will find it convenient to define the notion of majorizing sequences in a slightly more general way, in the sense that the sequences do not need to have the same length or sum. Throughout the whole article, a sequence will always be ordered (from largest to smallest).

**Definition 2.1** (majorizing sequences). Let  $(x_i)_{1 \leq i \leq n}$  and  $(y_i)_{1 \leq i \leq m}$  be two sequences. Then the sequence  $(x_i)_i$  majorizes  $(y_i)_i$  if for every positive integer  $k$  it is true that  $\sum_{1 \leq i \leq k} x_i \geq \sum_{1 \leq i \leq k} y_i$ . In this summation  $x_i$  and  $y_j$  are taken to be zero if  $i > n$  or  $j > m$ .

With this definition in place, Karamata's inequality is as follows.

**Theorem 2.2** (Karamata's inequality). Suppose  $(x_i)_{1 \leq i \leq n}$  and  $(y_i)_{1 \leq i \leq n}$  are decreasing sequences of equal length, taking values in some interval  $I$ , with  $(x_i)_i$  majorizing  $(y_i)_i$ . Let  $f : I \rightarrow \mathbb{R}$  be a convex function. If either of the following conditions are satisfied:

- (i)  $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$  or
- (ii)  $f$  is increasing on  $I$ ,

then  $\sum_{i=1}^n f(x_i) \geq \sum_{i=1}^n f(y_i)$ . Furthermore, provided  $f$  is strictly convex, we have equality only if the two sequences are identical.

The inequality under condition (i) is Karamata's original formulation [11]; the sufficiency of the alternative condition (ii) is well known and may be proved in the same way (see, for example, Pečarić et al. [17, Theorem 12.7]).

As a corollary, we obtain the following modified version for integer sequences.

**Corollary 2.3.** Let  $\vec{x} = (x_i)_{1 \leq i \leq n}$  and  $\vec{y} = (y_i)_{1 \leq i \leq m}$  be two sequences of nonnegative integers such that the former majorizes the latter and let  $t = \sum_i x_i - \sum_i y_i$ . Let  $f$  be a function satisfying  $f(0) = 0$ ,  $f(1) = 1$  and  $f'(x), f''(x) \geq 0$  for all  $x \in \mathbb{R}^+$ . Then

$$\sum_i f(x_i) \geq \sum_i f(y_i) + t.$$

If the two sequences are not equal (which is certainly the case when  $t > 0$ ) and  $f$  is strictly convex, the inequality is strict as well.

**Proof.** We can extend the sequence  $y_i$  with  $t$  values equal to 1 in the right spot (i.e. such that the sequence is still ordered) and add zeros to one of the two sequences to produce sequences  $\vec{x}', \vec{y}'$  of the same length. Suppose that  $y_k$  is the last positive term of  $\vec{y}$ , so that the  $t$  extra terms become  $y'_{k+1}, \dots, y'_{k+t}$ . Note that  $\sum_{i=1}^k y'_i \leq \sum_{i=1}^k x'_i$ , and for  $1 \leq s \leq t$ , we have  $\sum_{i=1}^{k+s} y'_i = s + \sum_{i=1}^k y_i$ . If  $x'_{k+s} > 0$  then we have  $\sum_{i=1}^{k+s} x'_i \geq s + \sum_{i=1}^k x_i \geq \sum_{i=1}^{k+s} y'_i$ , and if not we have  $\sum_{i=1}^{k+s} x'_i = \sum_{i \geq 1} x'_i \geq \sum_{i=1}^{k+s} y'_i$ . Consequently the updated sequence  $\vec{x}'$  does majorize  $\vec{y}'$  and so Karamata's inequality (i) says that

$$\sum_i f(x_i) \geq \sum_i f(y_i) + t. \quad \square$$

### 2.2. A majorization relationship

As the reader may expect, the crucial fact we will need is the following majorization between the two sequences derived from the graph. We will write  $m = |E(G)|$  and  $N = \binom{|G|}{2}$  for the natural lengths of the two sequences; we will frequently find it convenient to extend  $(n_e)_{e \in E(G)}$  to a sequence of length  $N$  by adding zero terms.

**Proposition 2.4.** For every graph  $G$ , the sequence  $(n_u(v) \cdot n_v(u))_{u,v \in E(G)}$  majorizes the sequence  $(d(u, v))_{u,v \in V}$ .

**Proof.** For every edge  $uv$ , let  $n_{uv} = n_u(v) \cdot n_v(u)$ . Let  $(n_i)_{1 \leq i \leq m}$  (where  $m = |E(G)|$ ) be the ordered sequence (from largest to smallest) of  $(n_e)_e$  and let  $(d_i)_i$  be the ordered sequence of the distances  $(d(u, v))_{u,v \in V}$ . The following claim essentially gives the intuition behind the majorization of the sequences, the property that the  $n_i$  are larger than  $d_i$  in some structured way.

**Claim 2.5.** For every pair of vertices  $x$  and  $y$  and every edge  $e = uv$  on a shortest path between  $x$  and  $y$ , we have  $d(x, y) \leq n_e$ .

**Proof.** Note that there are  $d(x, y) + 1$  vertices on a shortest path from  $x$  and  $y$ , all of them belonging to  $N_u(v)$  or  $N_v(u)$ . Hence  $d(x, y) + 1 \leq n_u(e) + n_v(e)$ , from which  $d(x, y) \leq n_u(v) \cdot n_v(u) = n_e$  follows.  $\square$

This implies that  $d(x, y) = \sum_{e \in P} \lambda_e n_e$ , where  $P$  is a shortest path (containing  $d(x, y)$  edges by definition) between  $x$  and  $y$  and  $\lambda_e = \frac{1}{n_e} \leq \frac{1}{d(x, y)}$  (by Claim 2.5), so  $\sum_{e \in P} \lambda_e \leq 1$ .

As a consequence, we have

$$\begin{aligned} \sum_{i=1}^k d_i &= \sum_{i=1}^k \sum_{e \in P_i} \lambda_e n_e \\ &= \sum_{e \in E} \Lambda_e n_e, \end{aligned}$$

where  $P_i$  is the chosen shortest path corresponding to the distance  $d_i$ . Here every  $\Lambda_e$  equals the number of shortest paths  $P_i$  through  $e$  (so at most  $n_e$ ) times  $\lambda_e$ , which is bounded by 1. Also we know  $\sum_{e \in E} \Lambda_e = \sum_{i=1}^k \sum_{e \in P_i} \lambda_e \leq k$ . This implies  $\sum_{e \in E} \Lambda_e n_e \leq \sum_{i=1}^k n_i$ .  $\square$

As a consequence, we note that [Conjecture 1.1](#) is true in a slightly stronger form.

**Theorem 2.6.** For every  $\alpha > 1$  and graph  $G$  which is not a complete graph, one has  $Sz^\alpha(G) - W^\alpha(G) > Sz(G) - W(G) \geq 0$ .

**Proof.** Since  $|E(G)| \neq \binom{n}{2}$ , the sequences  $(n_i)_i$  and  $(d_i)_i$  are not equal. Let  $t = \sum_i n_i - \sum_i d_i \geq 0$ . Since  $f(x) = x^\alpha$  satisfies  $f(0) = 0$  and  $f(1) = 1$  and is a positive, strictly convex increasing function for  $x \geq 0$  and  $\alpha > 1$ , by applying [Corollary 2.3](#) we have  $Sz^\alpha(G) > W^\alpha(G) + t$ .  $\square$

### 3. Graphs which satisfy the strong conjecture

In [Section 2](#), we noted that [Conjecture 1.2](#) was essentially an inequality for majorizing sequences instead of an inequality on graphs, once one observes [Proposition 2.4](#) is true. This cannot be the case for [Conjecture 1.1](#). One can take the sequences  $\bar{x} = \{625, 81, 81, 16\}$  and compare with  $\bar{y} = \{256, 256, 256, 1\}$ , the first one majorizing the second, but  $\sum x_i^\alpha = \sum y_i^\alpha$  being true for  $0, \frac{1}{4}$  and a number close to 0.88.

Fix a connected, non-complete graph  $G$  with  $n$  vertices and let  $h(\alpha) = Sz^\alpha(G) - W^\alpha(G)$ . Since  $h(\alpha)$  is a continuous function with  $h(0) < 0$  and  $h(1) \geq 0$ , by the intermediate value theorem there is at least one value of  $\alpha$  for which  $h(\alpha) = 0$  (and at least one such value lies in  $(0,1]$ ); the strong conjecture is therefore equivalent to  $\alpha$  being unique. We therefore give a sufficient condition, in terms of majorization, for this to be the case.

**Theorem 3.1.** Let  $\alpha$  be a value such that  $Sz^\alpha(G) = W^\alpha(G)$  for a non-complete graph  $G$ . If the sequence  $(n_i^\alpha)_i$  does majorize the sequence  $(d_i^\alpha)_i$ , then  $\alpha$  is unique.

**Proof.** Assume  $\beta$  also satisfies  $Sz^\beta(G) = W^\beta(G)$ . Let  $c = \frac{\beta}{\alpha}$ ; by [\[10, Proposition 8\]](#) we have  $\alpha, \beta > 0$  and so  $c > 0$ . Then

$$\sum_i (n_i^\alpha)^c = \sum_i (d_i^\alpha)^c, \tag{2}$$

where we extend the former sum with zero terms to give sequences of equal length. For  $c > 1$  the function  $f(x) = x^c$  is strictly convex, while for  $c < 1$  it is strictly concave. By [Theorem 2.2](#) (i), in both cases (2) would not be true, so this implies  $c = 1$  and hence  $\alpha = \beta$ .  $\square$

Together with [Proposition 2.4](#), this immediately gives the following.

**Corollary 3.2.** If  $G$  is a non-complete block graph then  $G$  satisfies [Conjecture 1.1](#) with  $\alpha_G = 1$ .

As further examples for the applicability, let us note that either of the following conditions on  $G$  is sufficient (but not necessary) to conclude that  $(n_i^\alpha)_i$  majorizes  $(d_i^\alpha)_i$ , where  $\alpha$  satisfies  $h(\alpha) = 0$ .

- I There exists some index  $j$  such that  $n_i \geq d_i$  if  $i \leq j$  and  $n_i \leq d_i$  if  $i > j$ .
- II For every  $1 \leq j \leq m$ , it holds that  $\prod_{i=1}^j n_i \geq \prod_{i=1}^j d_i$ .

To see that II is sufficient, note that it implies that  $(\log n_i)_{1 \leq i \leq m}$  majorizes  $(\log d_i)_{1 \leq i \leq m}$ , and hence (since  $f(x) = \exp(\alpha x)$  is increasing and convex) that  $(n_i^\alpha)_{1 \leq i \leq m}$  majorizes  $(d_i^\alpha)_{1 \leq i \leq m}$ . Since we also have  $\sum_{i=1}^k n_i^\alpha = Sz^\alpha(G) = W^\alpha(G) \geq \sum_{i=1}^k d_i^\alpha$  for each  $k > m$ , we have the required majorization property.

For graphs satisfying I, much more is true.

**Theorem 3.3.** If  $G$  satisfies condition I above, then for every value of  $\alpha > 0$  satisfying  $h(\alpha) \geq h(0)$  we have  $h'(\alpha) > 0$ .

This implies the strong conjecture holds for  $G$ , since there is some smallest positive value  $\alpha_G \leq 1$  for which  $h(\alpha_G) = 0 > h(0)$ , and thereafter the function is strictly increasing so strictly positive. Note that [Theorem 3.3](#) is best possible, in the sense that we may have  $h'(0) < 0$  (and hence  $h'(\epsilon) < 0$  with  $h(\epsilon)$  arbitrarily close to  $h(0)$ ). For example, this is the case for stars with at least 8 vertices.

**Proof.** Set  $\tilde{n}_i$  to be 1 if  $n_i = 0$  and  $n_i$  otherwise. Since  $d_i \geq 1$  for each index  $i$ , condition I still applies with  $n_i$  replaced by  $\tilde{n}_i$ , and ensures that  $\sum_{i \leq k} (\tilde{n}_i^\alpha - d_i^\alpha)$ , considered as a function of  $k$  for any fixed  $\alpha > 0$ , is increasing for  $k \leq j$  and decreasing

thereafter. Consequently,  $(\tilde{n}_i^\alpha)$  majorizes  $(d_i^\alpha)$  provided

$$\sum_{i=1}^N \tilde{n}_i^\alpha \geq \sum_{i=1}^N d_i^\alpha,$$

i.e. whenever  $h(\alpha) + N - e(G) \geq 0$ . Since  $h(0) = e(G) - N$ , this condition becomes  $h(\alpha) \geq h(0)$ . Now provided  $\alpha > 0$  we have

$$\begin{aligned} h'(\alpha) &= \sum_{i=1}^N (\log n_i \cdot n_i^\alpha - \log d_i \cdot d_i^\alpha) \\ &= \sum_{i=1}^N (\log \tilde{n}_i \cdot \tilde{n}_i^\alpha - \log d_i \cdot d_i^\alpha) \\ &= \alpha^{-1} \left( \sum_{i=1}^N f(\tilde{n}_i^\alpha) - \sum_{i=1}^N f(d_i^\alpha) \right), \end{aligned}$$

where  $f(x) = x \log x$  for  $x \geq 1$ . Since this is a strictly convex, increasing function, and the two sequences are distinct (the fact that  $\sum_i \tilde{n}_i > \sum_i d_i$  implies there is some  $i$  with  $\tilde{n}_i > d_i$ ), [Theorem 2.2](#) (ii) gives  $\sum_{i=1}^N f(\tilde{n}_i) > \sum_{i=1}^N f(d_i)$ , as required.  $\square$

In particular this implies the strong conjecture holds for several classes of graphs.

**Corollary 3.4.** *The graph  $G$  satisfies Conjecture 1.1 in any of the following cases:*

- $G$  is bipartite;
- $G$  is edge-transitive;
- $\text{diam}(G) = 2$ ; or
- $\text{diam}(G) = 3$  and  $m \leq N/2$ .

**Proof.** If  $G$  is bipartite then for each edge  $uv$ , each vertex is strictly closer to either  $u$  or  $v$ , since equality would create an odd cycle. Thus  $n_u(v) \cdot n_v(u) \geq |G| - 1 \geq d(x, y)$ , and I holds with  $j = m$ .

If  $G$  is edge-transitive then every edge is on a diameter, meaning  $n_i \geq \text{diam}(G) \geq d_i$  for  $1 \leq i \leq m$ .

If  $\text{diam}(G) = 2$  then  $d_i \in \{1, 2\}$  for each  $i$ . Take  $j$  maximal such that  $n_i \geq 2$ . Then  $n_i \geq d_i$  for  $i \leq j$ , and  $n_i \leq d_i$  for all  $i > j$ , as required. Finally, if  $\text{diam}(G) = 3$  and  $m \leq N/2$  then we have  $d_i \in \{2, 3\}$  for each  $i \leq m$ , and taking  $j$  maximal such that  $n_i \geq 3$  gives the required property.  $\square$

Another corollary of [Theorem 3.1](#) is the case when the  $n_i$  are “on average” at least equal to the diameter, which can be stated more precisely as follows with the use of the notion  $\mu_\alpha$  for the power mean with exponent  $\alpha$ , i.e. using  $N = \binom{n}{2}$  we have

$$\mu_\alpha(d_1, d_2, \dots, d_N) = \sqrt[\alpha]{\frac{\sum_{i=1}^N d_i^\alpha}{N}}.$$

**Corollary 3.5.** *Let  $\alpha$  be a value such that  $\text{Sz}^\alpha(G) = W^\alpha(G)$  for a non-complete graph  $G$ . If*

$$\mu_\alpha(d_1, d_2, \dots, d_N) \geq \left(\frac{m}{N}\right)^{\frac{1}{\alpha}} \text{diam}(G),$$

then this value  $\alpha$  is unique.

**Proof.** By [Theorem 3.1](#) it is sufficient to prove that  $(n_i^\alpha)_i$  majorizes  $(d_i^\alpha)_i$ , which is the case if for each  $k \leq m$  the first  $k$  terms of the first sequence are on average at least equal to the largest term of the second, namely,  $\text{diam}(G)^\alpha$ . As the sequence is ordered, it is sufficient to prove this for  $k = m$ , i.e.

$$\begin{aligned} \sum_{i=1}^m n_i^\alpha \geq m \text{diam}(G)^\alpha &\iff \sum_{i=1}^N d_i^\alpha \geq m \text{diam}(G)^\alpha \\ &\iff \sqrt[\alpha]{\frac{\sum_{i=1}^N d_i^\alpha}{N}} \geq \sqrt[\alpha]{\frac{m}{N}} \text{diam}(G). \end{aligned} \quad \square$$

[Corollary 3.5](#) is sufficient to prove [Conjecture 1.1](#) for all graphs lying in some sparse regime. Intuitively if  $N \gg m$ , the  $n_i$  are on average much larger than the  $d_i$  and so the majorization property will hold.

**Theorem 3.6.** *Conjecture 1.1 is true for graphs with  $m$  edges and  $n$  vertices whenever  $m \leq \frac{1}{4}(n^{4/3} - n^{1/3})$ .*

**Proof.** By [Corollary 3.5](#) we only need to prove that the inequality

$$\mu_\alpha := \mu_\alpha(d_1, d_2, \dots, d_N) \leq \left(\frac{m}{N}\right)^{\frac{1}{\alpha}} \text{diam}(G)$$

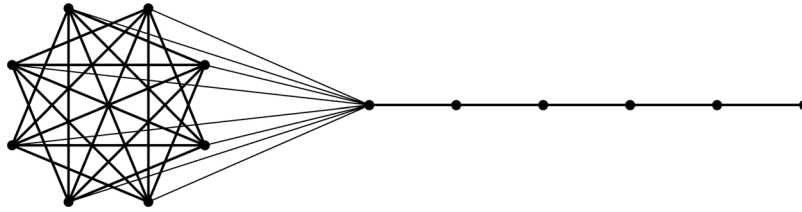


Fig. 1. The graph  $G_{k,\ell}$  for  $k = 8$  and  $\ell = 5$ .

is impossible. So assume this is the case, i.e. we have  $\text{diam}(G) \geq \left(\frac{N}{m}\right)^{\frac{1}{\alpha}} \mu_{\alpha} \geq (2n^{2/3})^{\frac{1}{\alpha}} \mu_{\alpha}$ . Note that for all edges in the middle of a diameter, the  $n_i$  are large.  $\square$

**Claim 3.7.** Let  $e_1 e_2 \dots e_{\text{diam}(G)}$  be the path between two vertices  $x$  and  $y$  for which  $d(x, y) = \text{diam}(G)$ . Then  $n_{e_j} \geq \frac{3}{16} \text{diam}(G)^2$  when  $\frac{\text{diam}(G)}{4} \leq j \leq \frac{3\text{diam}(G)}{4} + 1$ .

**Proof.** Let  $e_j = uv$  with  $u = e_j \cap e_{j-1}$  and  $v = e_j \cap e_{j+1}$ . Then  $N_u(e_j)$  contains at least the  $j$  vertices on the path from  $x$  to  $u$ , and similarly  $n_v(e_j) \geq \text{diam}(G) + 1 - j$ . Hence  $n_u(e_j)n_v(e_j) \geq j(\text{diam}(G) + 1 - j) \geq \frac{3}{16} \text{diam}(G)^2$ .

Taking into account that  $\alpha \leq 1$  and  $\mu_{\alpha} \geq 1$ , this implies

$$\begin{aligned} \sum_i n_i^{\alpha} &\geq \frac{\text{diam}(G)}{2} \left(\frac{3}{16} \text{diam}(G)^2\right)^{\alpha} \\ &\geq \frac{3}{32} \text{diam}(G)^{3\alpha} \\ &\geq \frac{3}{4} n^2 \mu_{\alpha}^{\alpha} \\ &> W^{\alpha}(G). \end{aligned}$$

This is the desired contradiction, so in the sparse regime Corollary 3.5 does apply.  $\square$

As a consequence, Conjecture 1.1 is a.a.s. true for any random graph of the form  $G_{n,p}$  or  $G_{n,m}$  (conditional on the graph being connected).

**Corollary 3.8.** For any value of  $m$  or  $p$ , Conjecture 1.1 is a.a.s. true for any random graph of the form  $G_{n,p}$  or  $G_{n,m}$  (conditional on the graph being connected).

**Proof.** By Theorem 3.6 the result is true in the sparse case, so we only have to consider the dense case where  $m > \frac{1}{4}(n^{4/3} - n^{1/3})$ . By a result of Klee and Larman [14], the diameter is a.a.s. at most 3, and when  $m \geq \frac{N}{3}$  it is a.a.s. equal to 2. Therefore Corollary 3.4 applies in this regime. Finally, the result for  $G(n, p)$  follows by conditioning on the number of edges.  $\square$

#### 4. Counterexamples to the strong conjecture

Let  $G_{k,\ell}$  be a graph constructed as follows. Start with a clique  $K_k$  and remove a  $k$ -cycle  $v_1 v_2 v_3 \dots v_k$  (a Hamiltonian path) from the clique. Next connect all its vertices with one endvertex  $u_1$  of a path  $u_1 u_2 \dots u_{\ell+1}$  of length  $\ell$  (i.e. of order  $\ell + 1$ ). See Fig. 1 for a small example.

Now we compute  $W^{\alpha}(G_{k,\ell})$  and  $\text{Sz}^{\alpha}(G_{k,\ell})$  (assuming  $k \geq 6$ ).

##### 4.1. Computation of $W^{\alpha}(G_{k,\ell})$ and $\text{Sz}^{\alpha}(G_{k,\ell})$

The number of edges of the graph equals  $\binom{k}{2} + \ell$ . There are  $k$  pairs of vertices which are at distance 2 in the clique. For each  $i \leq \ell + 1$ , there are  $k$  pairs with one vertex from the clique together with  $u_i$  which are at distance  $i$  from each other. There are also  $\ell + 1 - i$  pairs of vertices, both belonging to the path, which are at distance  $i$  from each other; these are  $(u_j, u_{j+i})$  for  $1 \leq j \leq \ell + 1 - i$ .

This implies that

$$W^{\alpha}(G_{k,\ell}) = \binom{k}{2} + \ell + (2k + \ell - 1)2^{\alpha} + \sum_{i=3}^{\ell+1} (k + \ell + 1 - i)i^{\alpha}.$$

Next, we compute  $\text{Sz}^{\alpha}(G_{k,\ell})$ .

For every pair of vertices in the clique of the form  $(v_i, v_{i+2})$  (indices modulo  $k$ ), we have  $N_{v_i}(v_{i+2}) = \{v_i, v_{i+3}\}$  and  $N_{v_{i+2}}(v_i) = \{v_{i-1}, v_{i+2}\}$ . So the corresponding value for  $n_{v_i v_{i+2}} = 4$ ; there are  $k$  such edges.

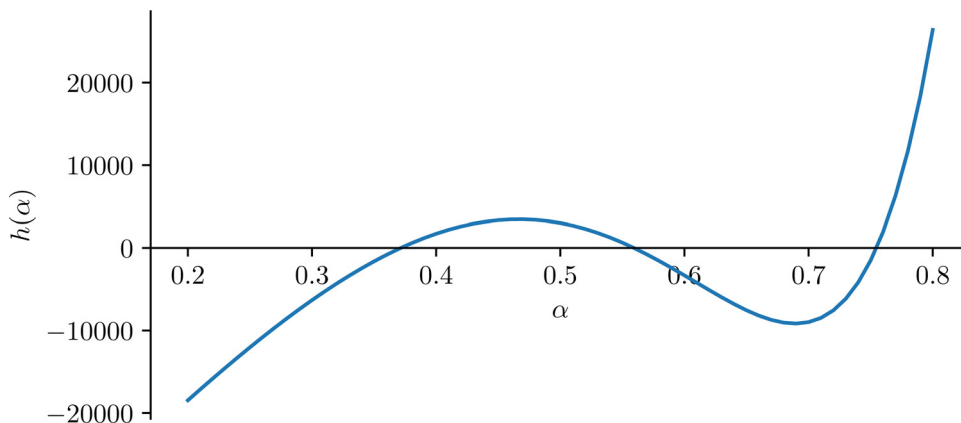


Fig. 2. The function  $h(\alpha)$  plotted for  $G_{520,82}$ .

For every edge  $v_i v_j$  with  $|i - j| \not\equiv 2 \pmod k$  we have  $N_{v_i}(v_j) = \{v_i, v_{j-1}, v_{j+1}\}$  and  $N_{v_j}(v_i) = \{v_j, v_{i-1}, v_{i+1}\}$ , so  $n_{v_i v_j} = 9$ . There are  $\binom{k}{2} - 2k$  edges of this form.

For every edge  $u_1 v_i$ , we have  $N_{v_i}(u_1) = \{v_i\}$  and  $N_{u_1}(v_i)$  contains all vertices from the path  $P_\ell$ , as well as  $v_{i-1}$  and  $v_{i+1}$ . So we have  $k$  edges here for which  $n_{v_i u_1} = \ell + 3$ .

For every edge of the form  $u_i u_{i+1}$  we have  $n_{u_i u_{i+1}} = (k + i)(\ell + 1 - i)$ .

Taking everything into account, we get

$$Sz^\alpha(G_{k,\ell}) = k4^\alpha + \left( \binom{k}{2} - 2k \right) 9^\alpha + k(\ell + 3)^\alpha + \sum_{i=1}^{\ell} ((k + i)(\ell + 1 - i))^\alpha.$$

#### 4.2. Some explicit counterexamples

There are multiple graphs of the form  $G_{k,\ell}$  which are counterexamples. All we need is that  $k$  is reasonably large, and then there are some corresponding  $\ell$  giving three values of  $\alpha$  for which  $Sz^\alpha(G_{k,\ell}) - W^\alpha(G_{k,\ell}) = 0$ . One example is  $k = 520$  and  $\ell = 82$ ; part of the function  $h(\alpha)$  showing that the conjecture does not hold is plotted in Fig. 2.

We can actually prove that there are infinitely many counterexamples. Since our construction is not a block graph, we know that  $h(0) < 0$  and  $h(1) > 0$ . For  $k$  tending to infinity and  $0 < \epsilon < 0.25$  fixed,  $0 \leq \alpha \leq 1$  and  $k^{0.5+\epsilon} < \ell < k^{1-\epsilon}$ , we note that

$$W^\alpha(G_{k,\ell}) = \binom{k}{2} \left( 1 + O\left(\frac{1}{k}\right) \right) + k \left( 1 + O\left(\frac{1}{k^\epsilon}\right) \right) \sum_{i=3}^{\ell+1} i^\alpha \sim \binom{k}{2} + k \frac{\ell^{\alpha+1}}{\alpha + 1}$$

and

$$Sz^\alpha(G_{k,\ell}) = \binom{k}{2} \left( 1 + O\left(\frac{1}{k^\epsilon}\right) \right) 9^\alpha + k^\alpha \left( 1 + O\left(\frac{1}{k^\epsilon}\right) \right) \sum_{i=1}^{\ell} i^\alpha \sim \binom{k}{2} 9^\alpha + k^\alpha \frac{\ell^{\alpha+1}}{\alpha + 1}.$$

This implies that  $h(\epsilon) \sim \binom{k}{2} (9^\epsilon - 1) > 0$  and  $h(1 - \epsilon) \sim -k \frac{\ell^{2-\epsilon}}{2-\epsilon} < 0$ . By applying the intermediate value theorem on  $h(0) < 0, h(\epsilon) > 0, h(1 - \epsilon) < 0$  and  $h(1) > 0$  we conclude that there are at least 3 values  $\alpha$  for which  $h(\alpha) = 0$ .

It is not hard to see that the same analysis gives the same conclusion when a matching instead of a cycle (for  $k$  even) was removed from the  $K_k$ , or any  $r$ -regular graph for fixed  $r$  (and  $k$  sufficiently large).

### 5. Conclusion

We have resolved both conjectures of Hriňáková et al. [10]. We proved the weaker Conjecture 1.2 by showing majorization of the sequence  $(n_i)_i$  with respect to  $(d_i)_i$ . The stronger Conjecture 1.1 is not true in general, as there are (infinitely many) sporadic counterexamples such as  $G_{520,82}$ , as explained in Section 4.

Nevertheless, it is true in many important cases as proven in this note. The conjecture is true for example for the following families:

- block graphs;
- graphs with diameter 2;
- any  $n$ -vertex graph with at most  $\frac{1}{2n^{2/3}} \binom{n}{2}$  edges;

- bipartite graphs;
- edge-transitive graphs; and
- the random graphs  $G(n, p)$  and  $G(n, m)$  a.a.s. (for the entire range of  $p$  or  $m$ ).

## Data Availability

No data was used for the research described in the article.

## References

- [1] B. Bollobás, P. Erdős, Graphs of extremal weights, *Ars Combin.* 50 (1998) 225–233.
- [2] A. Dobrynin, R. Entringer, I. Gutman, Wiener index of trees: theory and applications, *Acta Appl. Math.* 66 (3) (2001) 211–249.
- [3] A. Dobrynin, I. Gutman, On a graph invariant related to the sum of all distances in a graph, *Publ. Inst. Math. (Beograd) (N.S.)* 56 (70) (1994) 18–22.
- [4] R.C. Entringer, D.E. Jackson, D.A. Snyder, Distance in graphs, *Czechoslov. Math. J.* 26 (2) (1976) 283–296.
- [5] 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.
- [6] I. Gutman, S. Klavžar, B. Mohar (Eds.) Fifty years of the Wiener index, A. Kerber, Bayreuth, 1997. Papers in celebration of the 50th anniversary of the appearance of “Structural determination of paraffin boiling points” [J. Am. Chem. Soc. 69 (1947), no. 1, 17–20] by Harold Wiener, *Match No.* 35 (1997) (1997).
- [7] I. Gutman, B. Ruščić, N. Trinajstić, C.F. Wilcox, J. Chem, Graph theory and molecular orbitals. XII. Acyclic polyenes, *J. Chem. Phys.* 62 (1975) 3399–3405.
- [8] I. Gutman, N. Trinajstić, Graph theory and molecular orbitals. Total  $\pi$ -electron energy of alternant hydrocarbons, *Chem. Phys. Lett.* 17 (1972) 535–538.
- [9] H. Hosoya, Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons, *Bull. Chem. Soc. Jpn.* 44 (9) (1971) 2332–2339.
- [10] K. Hriňáková, M. Knor, R. Škrekovski, An inequality between variable Wiener index and variable Szeged index, *Appl. Math. Comput.* 362 (6) (2019) 124557.
- [11] J. Karamata, Sur une inégalité relative aux fonctions convexes, *Publ. Math. Univ. Belgrade* 1 (1932) 145–148.
- [12] S. Klavžar, On the canonical metric representation, average distance, and partial Hamming graphs, *Eur. J. Combin.* 27 (1) (2006) 68–73.
- [13] S. Klavžar, A. Rajapakse, I. Gutman, The Szeged and the Wiener index of graphs, *Appl. Math. Lett.* 9 (5) (1996) 45–49.
- [14] V. Klee, D. Larman, Diameters of random graphs, *Canadian J. Math.* 33 (3) (1981) 618–640.
- [15] A. Miličević, S. Nikolić, On variable Zagreb indices, *Croat. Chem. Acta* 77 (2004) 97–101.
- [16] M.J. Nadjafi-Arani, H. Khodashenas, A.R. Ashrafi, On the differences between Szeged and Wiener indices of graphs, *Discrete Math.* 311 (20) (2011) 2233–2237.
- [17] J.E. Pečarić, F. Proschan, Y.L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*, Mathematics in Science and Engineering, vol. 187, Academic Press, Inc., Boston, MA, 1992.
- [18] M. Randić, Characterization of molecular branching, *J. Am. Chem. Soc.* 97 (23) (1975) 6609–6615.
- [19] D.H. Rouvray, The value of topological indices in chemistry, *MATCH Commun. Math. Comput. Chem.* 1 (1975) 125–134.
- [20] Ž.K. Vukičević, L. Bulatović, On the variable Wiener–Szeged inequality, *Discrete Appl. Math.* 307 (2022) 15–18.
- [21] H. Wiener, Structural determination of paraffin boiling points, *J. Am. Chem. Soc.* 69 (1) (1947) 17–20.
- [22] H. Wiener, Relation of the physical properties of the isomeric alkanes to molecular structure. Surface tension, specific dispersion, and critical solution temperature in aniline, *J. Phys. Colloid Chem.* 52 (6) (1948) 1082–1089.
- [23] B. Zelinka, Medians and peripherians of trees, *Arch. Math. (Brno)* 4 (1968) 87–95.