Directed Zagreb Indices

Barbara M. Anthony and Alison M. Marr

Abstract Zagreb indices for undirected graphs were introduced nearly 50 years ago. Their original development was related to uses in chemistry, but over time mathematicians have also found them to be an interesting topic of study. We define and introduce Zagreb indices for directed graphs, give results that parallel many of the conjectures and theorems that exist for the original Zagreb indices, and produce results specific to the directed graph case.

Keywords Directed graphs · First Zagreb index · Second Zagreb index

1 Introduction

The Zagreb indices were first introduced [4] nearly 50 years ago. Since that time dozens of papers have been written comparing these two indices, finding bounds on their values, and generalizing these indices. The popularity of these indices stems from their applications to chemistry. For a general overview of the history of these indices and their applications to chemistry, see [8]. Additionally, the survey [7] by Liu and You summarizes some of the existing mathematical work in the field.

Let $G = (V, E)$ be a graph. Let $d(v)$ denote the degree of vertex v in the graph G. The classical definitions of the first and second Zagreb indices, developed by Gutman and Trinajstić [4], are as follows:

Definition 1 ([4]) The first Zagreb index on a graph G is defined as

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

B. M. Anthony (✉) · A. M. Marr
Southwestern University, Department of Mathematics and Computer Science, Georgetown, TX, USA
e-mail: anthonyb@southwestern.edu; marra@southwestern.edu

C. Gentile et al. (eds.), *Graphs and Combinatorial Optimization: from Theory to Applications*, AIRO Springer Series 5,
https://doi.org/10.1007/978-3-030-63072-0_15

The second Zagreb index on a graph G is defined as

$$M_2(G) = \sum_{e=(u,v)\in E(G)} d(u)d(v).$$

In this paper, we define and examine a new generalization of Zagreb indices by defining them on directed graphs. While work has been done on zeroth-order general Randić indices (also motivated by chemistry, giving the sum of bond contributions) on digraphs [9], this is the first time that directed Zagreb indices are being defined and studied. Throughout this work, we use *nodes* and *arcs* when speaking about directed graphs, and *vertices* and *edges* for undirected graphs. The number of nodes or vertices is denoted by n, while the number of arcs or edges in a graph or digraph is denoted by m. We use *digraph* and *directed graph* interchangeably. The *in-degree* of a node u is denoted by $d^-(u)$, and the *out-degree* by $d^+(u)$. A *source* is a node with in-degree zero, and a *sink* is a node with out-degree zero. Let $D = (N, A)$ be a directed graph. The in and out neighborhoods of a node u are defined, respectively, as $N^-(u) = \{v \in N(D) | (v, u) \in A(D)\}$ and $N^+(u) = \{v \in N(D) | (u, v) \in A(D)\}$.

Definition 2 The first Zagreb index on a directed graph D is defined as

$$\overrightarrow{M}_1(D) = \sum_{v\in N(D)} d^+(v)d^-(v).$$

The second Zagreb index on a directed graph D is defined as

$$\overrightarrow{M}_2(D) = \sum_{e=(u,v)\in A(D)} d^+(u)d^-(v).$$

We allow graphs to be connected or disconnected. We do not allow multiple arcs or loops. We do not allow isolated nodes: just as their inclusion does not alter the undirected Zagreb indices, nor do they change the directed Zagreb indices. When the digraph under consideration is obvious from the context, we may omit it, simply writing $\overrightarrow{M}_1$ instead of $\overrightarrow{M}_1(D)$ for either of the directed Zagreb indices, and similarly for the graph in undirected Zagreb indices.

An *oriented graph* is a digraph with no bidirected arcs, that is, if (u, v) is an arc in the digraph, (v, u) cannot be an arc in the digraph. A *cycle* is a graph (or digraph with no bidirected arcs) where $n = m \geq 3$. If n is odd, it is an *odd cycle*; otherwise, it is an *even cycle*. When the arcs in a cycle are all oriented in the same direction, it is a *directed cycle*. Given a graph G, we define G^* to be the digraph with bidirected arcs in G^* for every edge in G. Thus, for example, K_2^* is a pair of bidirected arcs. Given a digraph D, let $\overleftarrow{D}$ denote the digraph where the orientation of every arc in D is flipped. We define a *directed path* to be a path in which all arcs are oriented so that the destination of an arc in the path is the origin of the subsequent path arc.

2 Results

This section begins with some fundamental properties of these newly defined directed Zagreb indices. Next, we compare the two directed Zagreb indices and relate these results to previously know relationships about the (undirected) Zagreb indices. We examine the values of $\overrightarrow{M}_1(D)$ and $\overrightarrow{M}_2(D)$ for various cycles, stars, and paths, characterize many categories of graphs in terms of whether or not $\overrightarrow{M}_1$ and $\overrightarrow{M}_2$ are equal, and then explore the possible values for $\overrightarrow{M}_2(D) - \overrightarrow{M}_1(D)$.

2.1 Fundamental Properties of Directed Zagreb Indices

Though $\overrightarrow{M_1}$ is defined in terms of nodes, we show that we can in fact write $\overrightarrow{M}_1$ as a sum over arcs in the digraph. We then observe that flipping the orientation of all arcs in a digraph does not change the directed Zagreb indices. We give explicit formulas for the directed Zagreb indices on regular digraphs. Finally, we show (in Property 4) that the directed Zagreb indices of a disconnected graph are simply the sum of the directed Zagreb indices of the components.

Property 1 An alternative way to write $\overrightarrow{M}_1(D)$ is

$$\overrightarrow{M_1}(D) = \frac{1}{2} \sum_{\boldsymbol{e}=(u,v)\in A(D)} (d^-(u) + d^+(v)).$$

Proof Consider a node $x \in D$. In the proposed alternative, the contribution of the node x to $\overrightarrow{M}_1(D)$ comes from every arc that it is a part of. For those arcs in which x is the origin, we count the number of arcs that enter x, that is, $d^-(x)$, for every arc that starts at x, whose number is $d^+(x)$, giving a total of $d^+(x)d^-(x)$. For those arcs in which x is the destination, we count the number of arcs that leave x, that is, $d^+(x)$, and we count that for every entering arc, namely $d^-(x)$, again giving a total of $d^+(x)d^-(x)$. The division by two handles the double-counting. Ultimately, in the alternative representation, we have counted $d^+(x)d^-(x)$ for every node $x \in D$, precisely matching the definition of $\overrightarrow{M}_1(D)$. □

Property 2 $\overrightarrow{M}_1(D) = \overrightarrow{M}_1(\overleftarrow{D})$ and $\overrightarrow{M}_2(D) = \overrightarrow{M}_2(\overleftarrow{D})$

Property 3 Let $D = (N, A)$ be a regular digraph with $d^+(v) = d^-(v) = k \; \forall v \in N$. Then $\overrightarrow{M}_1(D) = nk^2$ and $\overrightarrow{M}_2(D) = mk^2$.

Corollary 1 *The complete digraph K_n^* has $\overrightarrow{M}_1(K_n^*) = n(n-1)^2$ and $\overrightarrow{M}_2(K_n^*) = n(n-1)^3$.*

Property 4 Let a directed graph D consist of two connected components, digraphs R and S. Then $\overrightarrow{M}_1(D) = \overrightarrow{M}_1(R) + \overrightarrow{M}_1(S)$, and $\overrightarrow{M}_2(D) = \overrightarrow{M}_2(R) + \overrightarrow{M}_2(S)$.

Proof Since nodes in D can be partitioned into nodes in R and nodes in S, $\overrightarrow{M}_1(D) = \sum_{v \in N(D)} d^+(v)d^-(v) = \sum_{v \in N(R)} d^+(v)d^-(v) + \sum_{v \in N(S)} d^+(v)d^-(v) = \overrightarrow{M}_1(R) + \overrightarrow{M}_1(S)$. Analogously, the arcs in D can be partitioned into the arcs in R and the arcs in S, yielding the desired result for $\overrightarrow{M}_2$. □

2.2 Comparing Directed Zagreb Indices

One of the most popular avenues of research when studying the first two Zagreb indices is to compare their values. For undirected Zagreb indices, M_1 and M_2 can be equal, $M_1 > M_2$ or $M_1 < M_2$. In [6], Horoldagva, Das, and Selenge show which classes of graphs fall into each of the three categories. We show that for directed Zagreb indices only two of these options are possible.

Theorem 1 *For any directed graph D, $\overrightarrow{M}_1(D) \leq \overrightarrow{M}_2(D)$.*

Proof Proof by induction on the number of arcs in D.

Base case: Trivially, if there are no arcs in D, then $\overrightarrow{M}_1(D) = 0 = \overrightarrow{M}_2(D)$. For illustration, if D contains a single arc, then $\overrightarrow{M}_1(D) = 0 \cdot 1 + 1 \cdot 0 = 0$, and $\overrightarrow{M}_2(D) = 1 \cdot 1 = 1$, so $\overrightarrow{M}_1(D) < \overrightarrow{M}_2(D)$.

Inductive hypothesis: We assume that for any digraph D with k arcs, $\overrightarrow{M}_1(D) \leq \overrightarrow{M}_2(D)$. Let $D^\wedge$ be a digraph with $k+1$ arcs. We want to show that $\overrightarrow{M}_1(D^\wedge) \leq \overrightarrow{M}_2(D^\wedge)$.

Pick an arbitrary arc $e = (u, v) \in D^\wedge$. Removing e from $D^\wedge$ yields a digraph D' with exactly k arcs, and thus $\overrightarrow{M}_1(D') \leq \overrightarrow{M}_2(D')$ by the inductive hypothesis. Thus, by construction, $e \notin D'$.

We now consider how $\overrightarrow{M}_1$ differs between D' and $D^\wedge$. The only terms in the sum which are altered are the terms contributed by the nodes u and v. In $D^\wedge$, the in-degree of node u is unchanged, and its out-degree increases by 1. Thus, the contribution of u to $\overrightarrow{M}_1$ was previously $d^-(u) \cdot d^+(u)$, and is now $d^-(u) \cdot (d^+(u)+1)$, showing that the change from the contribution of node u is exactly $d^-(u)$. Similarly, the change from the contribution of node v is exactly $d^+(v)$.

Thus $\overrightarrow{M}_1(D^\wedge) = \overrightarrow{M}_1(D') + d^-(u) + d^+(v)$.

Calculating $\overrightarrow{M}_2(D^\wedge)$, since $e \notin D'$, $\overrightarrow{M}_2(D^\wedge)$ is precisely $\overrightarrow{M}_2(D')$ plus the contribution from arc e, both to the new summand term from e, and potential increases to existing arcs in D'.

The new arc e generates a contribution of $(d^+(u) + 1)(d^-(v) + 1)$, along with additional nonnegative contributions to the terms for arcs leaving u and entering v, namely

$$\sum_{x \in N^+(u)} d^-(x) + \sum_{y \in N^-(v)} d^+(y).$$

Thus

$$\vec{M}_2(D^\wedge) = \vec{M}_2(D') + d^+(u) + d^-(v) + d^+(u)d^-(v) + 1 + \sum_{x \in N^+(u)} d^-(x) + \sum_{y \in N^-(v)} d^+(y)$$

$$\geq \vec{M}_1(D') + d^+(u) + d^-(v) + d^+(u)d^-(v) + 1 + \sum_{x \in N^+(u)} d^-(x) + \sum_{y \in N^-(v)} d^+(y)$$

$$= \vec{M}_1(D^\wedge) + d^+(u)d^-(v) + 1 + \sum_{x \in N^+(u)} d^-(x) + \sum_{y \in N^-(v)} d^+(y)$$

$$\geq \vec{M}_1(D^\wedge) \text{ since all terms are nonnegative.}$$

□

We now establish an explicit connection between the classical Zagreb indices on an undirected graph, and the directed Zagreb indices on the corresponding digraph with bidirected arcs for all edges in the undirected graph. That result, Proposition 1, combined with Theorem 1 lead us to an alternative proof of a known result, that for undirected graphs the first Zagreb index is at most twice the second Zagreb index. While the result was already presented in [2], we highlight it here in Corollary 2 because of how it further illustrates the connection between the directed and undirected Zagreb indices.

Proposition 1 *Let G be an arbitrary undirected graph. Then $M_1(G) = \vec{M}_1(G^*)$ and $2 \cdot M_2(G) = \vec{M}_2(G^*)$.*

Proof Recall that G^* is the directed graph with bidirected arcs in G^* for every edge in G. By construction of G^*, for any node $v \in G^*$ arising from a vertex $x \in G$, $d^+(v) = d^-(v) = d(x)$. The equality of $M_1(G) = \vec{M}_1(G^*)$ follows immediately, and $2 \cdot M_2(G) = \vec{M}_2(G^*)$ because there are two arcs in G^* for every edge in G. □

Corollary 2 *Let G be an arbitrary undirected graph. Then, $M_1(G) \leq 2M_2(G)$.*

As reported in Caporossi et al. [1], experiments with the AutoGraphiX system led to a conjecture that for undirected Zagreb indices $\frac{M_1(G)}{n} \leq \frac{M_2(G)}{m}$, which Pierre Hansen presented at the second meeting of the International Academy of Mathematical Chemistry in 2006. However, while it was shown the following year by Hansen and Vukicević [5] that the relationship always holds true for chemical graphs, they show that the conjecture is not true for general graphs. One such instance provided in [5] consists of a disconnected graph whose two components were a $K_{1,6}$ and C_3. We show that there is a natural transformation of that graph into the digraph $K^*_{1,6} \cup C^*_3$ that likewise disproves the analogous inequality for digraphs.

Lemma 1 *There exists a digraph such that* $\frac{\vec{M}_1(D)}{n} > \frac{\vec{M}_2(D)}{m}$.

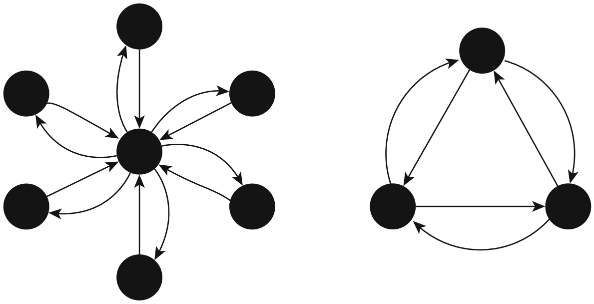

Fig. 1 A directed Zagreb instance where the first index divided by the number of nodes exceeds the second index divided by the number of arcs

Proof Let $D = K^*_{1,6} \cup C^*_3$, as shown in Fig. 1. Observe that $\overrightarrow{M}_1(D) = 36 + 6 \cdot 1 + 3 \cdot 4 = 54$, where the contributions come from the center of the star, the leaves of the star, and the nodes of the cycle, respectively. Observe also that $\overrightarrow{M}_2(D) = 6 \cdot 6 + 6 \cdot 6 + 6 \cdot 4 = 96$, with contributions from the six arcs directed out of the center of the star, the six arcs directed into the center of the star, and the six arcs in the C^*_3. Since the graph consists of 10 nodes and 18 arcs, $\dfrac{\overrightarrow{M}_1(D)}{n} = 5.4 > 5.333 = \dfrac{\overrightarrow{M}_2(D)}{m}$. □

2.3 Bounds on Directed Zagreb Indices

Considering all orientations on a particular graph, we can create bounds on the possible values for $\overrightarrow{M}_1(D)$ and $\overrightarrow{M}_2(D)$.

Proposition 2 *For any orientation of a $K_{1,n}$ (with no bidirectional arcs),* $0 \le \overrightarrow{M}_1(K_{1,n}) \le \lfloor \frac{n^2}{4} \rfloor$ *and* $\lceil \frac{n^2}{2} \rceil \le \overrightarrow{M}_2(K_{1,n}) \le n^2$.

Proof For the lower bound for $\overrightarrow{M}_1$, consider a star $K_{1,n}$ where all arcs are directed into the center. Then $\overrightarrow{M}_1(K_{1,n}) = 0$. Clearly a negative value is not possible.

For the upper bound for $\overrightarrow{M}_1$, consider a star $K_{1,n}$ where $\lfloor \frac{n}{2} \rfloor$ of the arcs are directed into the center, and the rest ($\lceil \frac{n}{2} \rceil$ arcs) are directed out of the center. This is the largest that $\overrightarrow{M}_1$ can be as the only contribution to $\overrightarrow{M}_1$ is at the center, and it is maximized when the in-degree and out-degree are as close as possible. Then $\overrightarrow{M}_1(K_{1,n}) = \lfloor \frac{n}{2} \rfloor * \lceil \frac{n}{2} \rceil$. If $n = 2s$ for some positive integer s, then $\lfloor \frac{n}{2} \rfloor * \lceil \frac{n}{2} \rceil = s * s = s^2 = n^2/4 = \lfloor \frac{n^2}{4} \rfloor$. If $n = 2s + 1$ for some positive integer s, then $\lfloor \frac{n}{2} \rfloor * \lceil \frac{n}{2} \rceil = \lfloor s + \frac{1}{2} \rfloor * \lceil s + \frac{1}{2} \rceil = s(s + 1) = s^2 + s = \lfloor \frac{4s^2+4s+1}{4} \rfloor = \lfloor \frac{n^2}{4} \rfloor$.

For $\overrightarrow{M}_2$, each arc contributes one times the in-degree (or out-degree) of the center. Arcs directed into the center contribute the in-degree of the center, and arcs directed out of the center contribute the out-degree of the center. Since the in-degree and out-degree of the center sums to n, $\overrightarrow{M}_2$ is maximized when either the in-degree or out-degree is maximized; that is, if all arcs are directed into the center of the star, or all are directed out of center of the star, $\overrightarrow{M}_2 = n^2$. Similarly, it is minimized when each is smallest, namely one is $\lfloor \frac{n}{2} \rfloor$ and the other is $\lceil \frac{n}{2} \rceil$. If $n = 2s$ for some positive integer s, then $\overrightarrow{M}_2 = \lfloor \frac{n}{2} \rfloor^2 + \lceil \frac{n}{2} \rceil^2 = s^2 + s^2 = 2s^2 = \lceil \frac{n^2}{2} \rceil$. If $n = 2s + 1$, then $\overrightarrow{M}_2 = \lfloor \frac{n}{2} \rfloor^2 + \lceil \frac{n}{2} \rceil^2 = \lfloor s + \frac{1}{2} \rfloor^2 + \lceil s + \frac{1}{2} \rceil^2 = s^2 + (s + 1)^2 = 2s^2 + 2s + 1 = \frac{4s^2+4s+1+1}{2} = \frac{n^2+1}{2} = \lceil \frac{n^2}{2} \rceil$. The result follows in either case. □

Proposition 3 *For any oriented P_n, $0 \leq \overrightarrow{M}_1(P_n) \leq n - 2$ (where there is some orientation which yields each possible integral value) and $n-1 \leq \overrightarrow{M}_2(P_n) \leq 4n-8$.*

Proof If arcs in P_n alternate directions, then $\overrightarrow{M}_1(P_n) = 0$. Clearly a negative value is not possible. Each endpoint of the path is either a source or a sink and thus does not contribute to $\overrightarrow{M}_1$. Each interior node on the path has either two arcs pointed in, contributing nothing, two arcs pointing out, contributing nothing, or one arc pointing in and one arc pointing out, contributing 1 to $\overrightarrow{M}_1$. Thus their sum, $\overrightarrow{M}_1(P_n)$ is maximized at $n - 2$ when all arcs are oriented in the same direction on the path, and integral values between the bounds can be obtained by the appropriate number of interior nodes with one arc pointing in and one arc pointing out.

For $\overrightarrow{M}_2(P_n)$, each of the $n-1$ arcs must contribute at least 1, and the lower bound of $n - 1$ is achieved when all arcs are oriented in the same direction on the path. Each arc can contribute at most 4 to $\overrightarrow{M}_2(P_n)$, which happens only if at each node the in-degree and out-degree on the path are both 2. The number of such occurrences is maximized when the arcs in P_n alternate directions. and all but the first and last arcs thus contribute 4, yielding $\overrightarrow{M}_2(P_n) = 4n - 8$. □

We next give results about when $\overrightarrow{M}_1(D) = 0$ and when $\overrightarrow{M}_1(D) \neq 0$.

Lemma 2 *$\overrightarrow{M}_1(D) = 0$ if and only if every node in D is either a source or a sink.*

Proof $\overrightarrow{M}_1(D) = 0$ means that each node contributes 0 to the sum which means either $d^+(v) = 0$ or $d^-(v) = 0$ for every node in D. Hence, each node is either a source or a sink. And if each node is a source or sink that implies that either $d^+(v) = 0$ or $d^-(v) = 0$ for every node v in D and hence $\overrightarrow{M}_1(D) = 0$. □

Proposition 4 *If a graph G has an odd cycle, then $\overrightarrow{M}_1(D) \neq 0$.*

Proof Consider an odd cycle C in G. There is no possible orientation of the arcs in C such that every node in C will be a sink or a source. That is, by a simple parity argument, some node must have an arc entering it and an arc leaving it. Thus, the directed graph D does not consist only of sources and sinks, and by Lemma 2, $\overrightarrow{M}_1(D) \neq 0$. □

Proposition 5 *If D contains no odd cycles, then there is an orientation of the arcs in D so that $\overrightarrow{M}_1(D) = 0$.*

Proof Since the digraph has no odd cycles, either it has no cycles, or its only cycles are even. We consider those cases separately.

Suppose the digraph has no cycles. Take a longest path in the tree and orient adjoining arcs in opposite directions. When all arcs on that path have been oriented, return to any node on that path that is incident to unoriented arcs, and orient any adjacent arcs in the same direction as all the others at that node (either all into or all out of the node) and then continue down each of those paths and orient adjoining arcs in opposite directions. Repeat until all arcs are oriented. Then, by construction every node in D is either a source or a sink, and by Lemma 2, $\overrightarrow{M}_1(D) = 0$.

Now suppose the digraph has at least one even cycle. Pick a largest even cycle, and orient adjoining arcs in that cycle in opposite directions. Then, continue this process on all remaining even cycles and/or paths until the graph clearly has only sources and sinks. If there are any arcs within this previously oriented cycle, they must only be connecting nodes that are an odd distance apart on the original cycle. Hence, the arcs inside the cycle can be oriented to keep all nodes being sources and sinks. Any paths that connect to a node of the original cycle and are not inside the original cycle can be oriented as described above starting with the same direction as the node where the path begins. Any additional cycles that might be adjoining the original cycle can also be oriented to keep all nodes sources and sinks as they are also even. □

2.4 Equality of Directed Zagreb Indices

We seek to fully characterize instances where $\overrightarrow{M}_1 = \overrightarrow{M}_2 \neq 0$. First we show that we need only focus on connected digraphs. Then we show that directed cycles and K_2^* have this property. However, we then show that digraphs where this equality holds are quite limited. We conjecture that directed cycles, K_2^*, and digraphs that are a disjoint union of these digraphs are in fact the only digraphs for which equality of $\overrightarrow{M}_1 = \overrightarrow{M}_2 \neq 0$ holds. Proving this conjecture remains an open question, but we make progress in that direction by showing that no digraph with a source and a sink will have $\overrightarrow{M}_1 = \overrightarrow{M}_2$, nor will oriented trees, nor a directed cycle plus an additional arc, nor cycles that are oriented but not directed.

Lemma 3 *If a disconnected graph has $\overrightarrow{M}_1 = \overrightarrow{M}_2 \neq 0$, then each of its connected components must also have $\overrightarrow{M}_1 = \overrightarrow{M}_2 \neq 0$.*

Proof By Property 4, the directed Zagreb indices of each component sum to the directed Zagreb index of the overall graph. Since Theorem 1 ensures that $\overrightarrow{M}_1 \leq \overrightarrow{M}_2$ for every digraph, the only way that the overall digraph can have $\overrightarrow{M}_1 = \overrightarrow{M}_2$ is thus if for each component equality holds. □

Lemma 4 *The directed cycle C_n has $\overrightarrow{M}_1(C_n) = \overrightarrow{M}_2(C_n) \neq 0$.*

Proof By Property 3, since the in-degree and out-degree of every node is $k = 1$, $\overrightarrow{M}_1(D) = nk^2 = n$ and $\overrightarrow{M}_2(D) = mk^2 = m$. Since $m = n$ in C_n with $n \geq 3$, the result is immediate. □

Lemma 5 *K_2^* has $\overrightarrow{M}_1(K_2^*) = \overrightarrow{M}_2(K_2^*) \neq 0$.*

Conjecture 1 The directed cycle and K_2^* (or graphs consisting solely of directed cycles and K_2^*) are the only graphs in which $\overrightarrow{M}_1 = \overrightarrow{M}_2 \neq 0$.

The following results lend support to the conjecture.

Property 5 It is NOT true that inserting an arc will always increase $\overrightarrow{M}_2$ by more than it increases $\overrightarrow{M}_1$.

Proof Consider a unidirectional path, that starts at node v_0 and ends at node v_{n-1}. Suppose we then insert a directed arc from v_{n-1} to v_0. The increase in $\overrightarrow{M}_1$ is 2, with one each contributed at v_0 and v_{n-1}. The increase in $\overrightarrow{M}_2$ is 1, the contribution from the new arc, as the sum from the other arcs does not change. □

Theorem 2 *Any digraph with a source and a sink cannot have* $\overrightarrow{M}_1 = \overrightarrow{M}_2$.

Proof Proof by contradiction. Let D be a digraph with a node u that is a source, and a node v that is a sink. Suppose that $\overrightarrow{M}_1 = \overrightarrow{M}_2$. Insert an arc from v to u. The increase in $\overrightarrow{M}_2$ is exactly 1, since the arc (v, u) contributes 1, but v as a sink had no other arcs out of it, and u as a source had no other arcs into it. But the increase in $\overrightarrow{M}_1$ is more than 1, since the increase is precisely the number of arcs into v (which is at least 1 as a sink) plus the number of arcs out of u (again, at least 1 as a source). Since the increase in $\overrightarrow{M}_1$ is more than the increase in $\overrightarrow{M}_2$, the values $\overrightarrow{M}_1$ and $\overrightarrow{M}_2$ could not have been equal, contradicting the original assumption. □

Since every oriented tree must contain both a source and a sink, we have the following corollary. We include the proof that every oriented tree must contain both a source and a sink for completeness.

Corollary 3 *Any oriented tree T with $n \geq 2$ has* $\overrightarrow{M}_1(T) < \overrightarrow{M}_2(T)$.

Proof Suppose our tree T has no sinks. Pick an arbitrary node, and follow an oriented edge (in the appropriate direction) out of that node. Repeat. Either we arrive at a node that has out-degree 0, which is thus a sink, or we return to a node we have already visited, which would mean there is a cycle, which is not possible in a tree.

Suppose instead our tree T has no source nodes. Reverse the orientation of all edges. Then our reversed graph would be a tree with no sinks. However, by the above argument, that is again impossible.

Thus, since every oriented tree has a source and a sink, we cannot have $\overrightarrow{M}_1 = \overrightarrow{M}_2$. □

Theorem 3 *Any digraph D which consists of solely a directed cycle and one additional arc has* $\overrightarrow{M}_1(D) < \overrightarrow{M}_2(D)$.

Proof First, recall that Lemma 4 ensures that $\overrightarrow{M}_1 = \overrightarrow{M}_2$ for any directed cycle. A graph that consists of a directed cycle and one additional arc can be constructed by the addition of an arc in one of the following ways:

1. as a disconnected arc,
2. as a chord in the cycle,

3. as an arc directed inward (or outward) into one node of the cycle and the other node would be a new node, or
4. an arc going in the opposite direction of one of the current arcs in the cycle.

Let e be the new arc in each case below.

Case 1 [Disconnected arc]: $\overrightarrow{M}_1(D+e) = \overrightarrow{M}_1(D)$ and $\overrightarrow{M}_1(D+e) = \overrightarrow{M}_2(D)+1$, by Property 4 hence $\overrightarrow{M}_1(D+e) < \overrightarrow{M}_2(D+e)$.

Case 2 [Chord in the cycle]: $\overrightarrow{M}_1(D+e) = \overrightarrow{M}_1(D)+1+1$ as two nodes will now be contributing 2 instead of 1. Similarly, $\overrightarrow{M}_2(D+e) = \overrightarrow{M}_2(D)+4+1+1$ where the 4 is from the new arc and the two 1s are from the additional in/out degree at the endpoints. Again, since $\overrightarrow{M}_1(D) = \overrightarrow{M}_2(D)$, $\overrightarrow{M}_1(D+e) < \overrightarrow{M}_2(D+e)$.

Case 3 [New node]: $\overrightarrow{M}_1(D+e) = \overrightarrow{M}_1(D)+1$ as only one node within the cycle will have a changed in- (or out-) degree. Similarly, $\overrightarrow{M}_2(D+e) = \overrightarrow{M}_2(D)+2+1$ where the 2 comes from the new arc and the 1 is how much the one arc in the cycle will change by.

Case 4 [Opposite direction]: $\overrightarrow{M}_1(D+e) = M_1(D)+2$ and $\overrightarrow{M}_2(D+e) = M_1(D)+4+1+1$ where the 4 is the new arc's contribution and each 1 is the amount two different arcs in the cycle will change.

In all cases, we see $\overrightarrow{M}_1(D+e) > \overrightarrow{M}_2(D+e)$. □

Theorem 4 *For any cycle C that is oriented but not directed, $\overrightarrow{M}_1(C) < \overrightarrow{M}_2(C)$.*

Proof Consider a cycle C of length n that is oriented but not directed. Let s be the number of maximal directed paths of length 1 in the cycle. Let t be the number of maximal directed paths of length greater than 1, but less than n in the cycle. Recall that since C is not directed, no directed path in the cycle can have length more than $n-1$.

To calculate $\overrightarrow{M}_1(C)$, first note that any node will either contribute 1 or 0. The node contributes 1 if the node is part of a unidirectional path (with in-degree and out-degree both one) and contributes 0 if it is a place where the direction of arcs in the cycle changes (that is, the in-degree and out-degree are not equal). The direction will change at $s+t$ places (where the paths change direction) and thus $\overrightarrow{M}_1(C) = n-(s+t) = n-s-t$. Note: $s+t$ must be even as it counts the number of direction changes and you cannot change direction an odd number of times and have a cycle. Furthermore, $s+t > 0$ as the cycle C is not unidirectional.

For $\overrightarrow{M_2}(C)$, any edge that is a path of length 1 will contribute 4 to the sum. Any path of length greater than 1 and less than n will have two arcs that each contribute 2 (the arcs at the start/end of the path). And, any arcs remaining will each contribute one to the sum. This gives: $\overrightarrow{M}_2(C) = 4s+4t+n-s-2t = n+3s+2t$.

Since s and t are nonzero, $n+3s+2t > n-s-t$ which ensures that $\overrightarrow{M}_1(C) < \overrightarrow{M}_2(C)$. □

2.5 Differences of Zagreb Indices

For undirected graphs, differences between the first two Zagreb indices were first studied in-depth by Furtula, Gutman, and Ediz in [3], who introduced the idea of the reduced second Zagreb index and studied this difference mainly on trees. In 2016, Das, Horoldagva, and Selenge [6] completely characterized which (undirected) graphs have $M_1 - M_2 = 0$, $M_1 - M_2 > 0$ and $M_1 - M_2 < 0$. In the directed case, we already know $\overrightarrow{M}_2 - \overrightarrow{M}_1 \geq 0$, and can in fact show that $\overrightarrow{M}_2 - \overrightarrow{M}_1$ can equal any nonnegative integer.

We know from Lemma 4 that $\overrightarrow{M}_2 - \overrightarrow{M}_1$ can equal zero. In addition we can also get $\overrightarrow{M}_2 - \overrightarrow{M}_1 = 1$ as if $D = P_n$ with all the arcs oriented in the same direction, $\overrightarrow{M}_2 - \overrightarrow{M}_1 = n - 1 - (n - 2) = 1$. We also know we can make $\overrightarrow{M}_2 - \overrightarrow{M}_1$ arbitrarily large by noting that $\overrightarrow{M}_2 - \overrightarrow{M}_1 = 4n$ for any $n \geq 3$ by using $D = C_n^*$, the cycle with all bidirectional edges present or we can get $\overrightarrow{M}_2 - \overrightarrow{M}_1 = n^2$ for $K_{1,n}$ with all the arcs directed out of the center. While these examples provide motivation that all nonnegative integer values are possible for this difference, the following theorem gives a construction technique for producing a digraph with any desired difference.

Theorem 5 *For all $s \in \mathbb{N}$, there exists a directed graph with $\overrightarrow{M}_2 - \overrightarrow{M}_1 = s$.*

Proof Let D be the digraph $K_{1,n}$ with x edges directed into the center and k edges directed out of the center where $x + k = n$ and $k \leq x$. Consider the collection of $k + 1$ digraphs $\{D = D_0, D_1, D_2, \ldots, D_k\}$ where D_i is the digraph formed from D by connecting i of the arcs directed out of the center to i different arcs directed into the center. An example of this construction can be seen in Fig. 2.

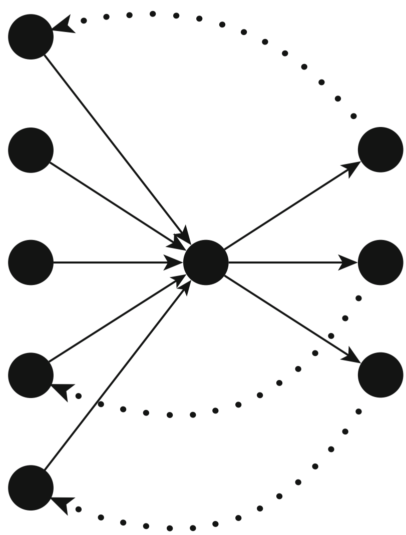

Fig. 2 A construction technique for digraphs with all possible values for the difference between the two Zagreb indices. The inclusion of any subset of the dotted edges leads to one of the digraphs in the collection

In general, $\overrightarrow{M}_2(D_i) - \overrightarrow{M}_1(D_i) = (x^2 + k^2 + i) - (xk + 2i) = x^2 + k^2 - xk - i$. Considering the case where $x = k$, $\overrightarrow{M}_2(D_i) - \overrightarrow{M}_1(D_i) = (2k^2 + i) - (k^2 + 2i) = k^2 - i$ and i ranges from 0 to k. Hence, in this case we get all the integers in the interval $[k^2 - k, k^2]$. Now, consider the case where $x = k + 1$. Similar calculations give $\overrightarrow{M}_2(D_i) - \overrightarrow{M}_1(D_i) = (k + 1)^2 + k^2 - k(k + 1) - i = k^2 + k + 1 - i$ and i ranges from 0 to k, so we get all the integers in the interval $[k^2 + 1, k^2 + k + 1]$. And if we move up to the next value for k, (so now $x = k + 1$ and k becomes $k + 1$), we get $\overrightarrow{M}_2(D_i) - \overrightarrow{M}_1(D_i) = (k + 1)^2 - i$ with i ranging from 0 to $k + 1$. So, we get the next interval to be $[k^2 + k, k^2 + 2k + 1]$. And, thus the overlap of intervals continues and we continue to increase the upper bound. If we plug in $k = 1$, we see that we start the interval at [0, 1] and hence can get any nonnegative integer values since these intervals line up and/or overlap and increase without bound. □

3 Conclusions and Open Questions

In this paper, we introduce the definition of first and second Zagreb indices on directed graphs. Initial propositions are given, relationships between the two indices are explored, and several classes of digraphs are studied in depth. We showed that the difference between $\overrightarrow{M}_2$ and $\overrightarrow{M}_1$ can take on any nonnegative integer value and state a conjecture on when this difference is zero. In particular, we believe that in all cases other than a directed cycle or K_2^* or disconnected combinations thereof, the difference between $\overrightarrow{M}_1$ and $\overrightarrow{M}_2$ is non-zero and inserting additional arcs or nodes will not result in equality of $\overrightarrow{M}_1$ and $\overrightarrow{M}_2$.

Another avenue of future research is motivated by Sect. 2.2. There we discuss $\overrightarrow{M}_1/n \le \overrightarrow{M}_2/m$. While this was shown to not be true for all digraphs, could it be true for all connected digraphs? Or even possibly for all digraphs where not all arcs are bidirectional?

Finally, since directed Zagreb indices do not have the same chemistry motivations of undirected Zagreb indices, they could be defined in many other ways or other indices described on undirected graphs could be generalized for digraphs. New definitions would prompt new results, propositions, and relationships, leading to additional areas for mathematical exploration of indices on digraphs.

References

1. Caporossi, G., Hansen, P., Vukicević, D.: Comparing Zagreb indices of cyclic graphs. Commun. Math. Comput. Chem. **63**, 441–451 (2010)
2. Fath-Tabar, G.H.: Old and new Zagreb indices of graphs. Commun. Math. Comput. Chem. **65**, 79–84 (2011)
3. Furtula, B., Gutman, I., Ediz, S.: On difference of Zagreb indices. Discrete Appl. Math. **178**, 83–88 (2014)

4. Gutman, I., Trinajstić, N.: Graph theory and molecular orbitals. Total φ-electron energy of alternant hydrocarbons. Chem. Phys. Lett. **17**, 535–538 (1972)
5. Hansen, P., Vukicević, D.: Comparing the Zagreb indices. Croat. Chem. Acta **80**, 165–168 (2007)
6. Horoldagva, B., Das, K.C., Selenge, T.A.: Complete characterization of graphs for direct comparing Zagreb indices. Discrete Appl. Math. **215**, 146–154 (2016)
7. Liu, B., You, Z.: A survey on comparing Zagreb indices. Commun. Math. Comput. Chem. **65**, 581–593 (2011)
8. Nikolić, S., Kovačević, G., Miličević, A., Trinajstić, N.: The Zagreb indices 30 years after. Croatica Chem. Acta **76**, 113–124 (2003)
9. Volkmann, L.: Sufficient conditions on the zeroth-order general Randić index for maximally edge-connected digraphs. Commun. Combin. Optim. **1**, 1–13 (2016)