Social Network Analysis and Mining (2024) 14:183
https://doi.org/10.1007/s13278-024-01285-y

ORIGINAL ARTICLE

Score and rank semi-monotonicity for closeness, betweenness, and distance–decay centralities

Paolo Boldi[1] · Davide D'Ascenzo[1] · Flavio Furia[1] · Sebastiano Vigna[1]

Received: 21 March 2024 / Revised: 14 June 2024 / Accepted: 17 June 2024

Abstract
Among the properties describing the behavior of centrality measures with respect to network modifications, *score monotonicity* means that adding an arc increases the centrality score of the target of the arc; *rank monotonicity* means that adding an arc improves the importance of the target with respect to the remaining nodes. It is known (Boldi and Vigna Intern Math 10:222–262, 2014, Boldi et al. Netw Sci 5(4):529–550, 2017) that score and rank monotonicity hold in directed graphs for almost all the classical centrality measures. In undirected graphs one expects that the corresponding properties hold when adding a new edge—in this case, both endpoints of the new edge should enjoy the increase in score/rank. However, recent results (Boldi et al. in Netw Sci 11(3):1–23, 2023) have shown that this is not true: for many centrality measures, it is possible to find situations in which adding an edge *reduces* the rank of one of its two endpoints. In this paper we introduce a weaker property for undirected networks, *semi-monotonicity*, in which just one of the two endpoints of a new edge is required to enjoy score or rank monotonicity. We show that this property is satisfied by closeness centrality, by a large class of distance-based centralities, and (somehow surprisingly) by betweenness centrality. In the last two cases, we prove in fact a stronger property, *basin dominance*, which is of independent interest.

Keywords Centrality · Semi-monotonicity · Closeness · Betweenness · Graphs

1 Introduction

Centrality measures are a well-established tool for the analysis of complex networks. They are used to identify the most important nodes in a graph and to understand the structure of the network itself. In the last decades, many centrality measures have been proposed, and they have been used in a wide range of applications, from social network analysis to biology, from computer networks to the study of the World Wide Web. To gain a formal understanding as to which extent a centrality mimics the intuitive concept of importance it is useful to provide a set of axioms that a centrality measure should (or should not) satisfy.

Previous work (Boldi and Vigna 2014, 2019; Boldi et al. 2017) has identified two important properties that a centrality measure should capture: score monotonicity and rank monotonicity. The former states that adding an arc to the network should increase the score of the target of the arc, while the latter requires that adding an arc increases the rank of the target of the arc with respect to the other nodes in the network. In the case of directed networks, it is known that almost all classical centrality measures satisfy both properties (Boldi et al. 2017). In the case of undirected networks, though, the situation is different: while again score monotonicity is satisfied by almost all classical centrality measures, rank monotonicity is not (Boldi et al. 2023), meaning that it is possible to find situations where adding an edge *reduces* the rank of one of its two endpoints. In particular:

- Degree centrality and Seeley's index (Seeley 1949) are score and (strictly) rank monotone;
- Closeness (Bavelas 1948, 1950), harmonic centrality (Beauchamp 1965), and Katz's index (Katz 1953) are score monotone but not rank monotone;
- Betweenness (Anthonisse 1971; Freeman 1977), eigenvector centrality (Landau 1985; Berge 1958), and PageR-

This paper is an extended version of Boldi et al. (2024).

✉ Paolo Boldi
paolo.boldi@unimi.it

[1] Dipartimento di Informatica, Università degli Studi di Milano, Milano, Italy

Published online: 06 September 2024

ank (Page et al. 1999) are neither score nor rank monotone.

In terms of real-world networks, this means that getting a new follower is always a good thing, whereas getting a new friend might even be detrimental to one's importance.

One of the questions left open in Boldi et al. (2023) is whether the addition of an edge could result in a decrease in the rank of both endpoints. Said otherwise: if a new edge is added to an *undirected* network, does the importance of *at least one* of its two endpoints increase?

In this paper we provide a positive answer to this question by introducing *semi-monotonicity*, a weaker condition than monotonicity in which we require that *at least one* endpoint of the new edge in an undirected network enjoys monotonicity. The main results of this paper are that closeness centrality, distance-decay centralities (e.g., harmonic centrality) and betweenness do satisfy rank semi-monotonicity.

The paper is organized as follows. In Sect. 2, we introduce the concept of score semi-monotonicity and (strict) rank semi-monotonicity. Then, in Sect. 3, we introduce the concept of δ-monotonicity and δ-semi-monotonicity, providing connections between these two notions and rank (semi-)monotonicity. Finally, in Sect. 4 we introduce an even stronger property, basin dominance, which is the property we will prove to hold for almost all the centralities we consider.

We then show in Sect. 5 that basin dominance holds also for *distance-decay centralities* satisfying a (discrete) convexity condition on their decay function. The condition is necessary and sufficient, thus characterizing basin dominance for those centralities. From the characterization, we immediately obtain results for closeness, harmonic centrality, and for centralities with exponential or power-law decay functions. Finally, in Sect. 6 we show that betweenness is basin dominant. Along the way, we show that the semi-monotonicity of closeness and betweenness is not strict, contrarily to what happens for harmonic centrality, echoing similar results for the directed case. In Sect. 7, we end the paper by drawing some conclusions and interpretations of our results.

With respect to the conference paper (Boldi et al. 2024), the definitions of δ-(semi-)monotonicity are new; all results about distance-decay centralities (in particular, the characterization of basin dominance in terms of convexity of the decay function) are also new, and so is the proof of basin dominance for betweenness. We have discussed several detailed properties of basin dominance that were not mentioned in the conference paper. We fixed a minor mistake—closeness centrality is rank semi-monotone, but not basin dominant, and we provide a counterexample. The definitions and results about exponential/power-law decay functions are new. We added several relevant references to previous literature.

2 Score and rank semi–monotonicity

Although this paper is focused on undirected graphs, we will sometimes present definitions for both directed and undirected graphs, as the directed case is sometimes used for comparison; the following definitions about graphs are standard (see, for instance, Berge (1958)). A *directed graph* is a pair $G = (V_G, A_G)$ where V_G is a set of nodes and $A_G \subseteq V_G \times V_G$ is a set of ordered pairs, called *arcs*; we will denote the existence of an arc from x to y with the notation $x \to y$. An *undirected graph* is a directed graph such that $x \to y$ iff $y \to x$; when this happens, we say that there is an *edge* between x and y, denoted by the notation $x — y$, or, equivalently, we say that x and y are *adjacent*.

We denote with $G + x \to y$ the graph obtained by adding the arc $x \to y$ to a directed graph G, and with $G + x — y$ the graph obtained by adding the edge $x — y$ to an undirected graph G.

A (simple) *path* of length k from x to y in an undirected graph G is a sequence of distinct vertices $x_0, x_1, \cdots, x_k$ such that $x_0 = x$, $x_k = y$, and $x_{i-1} - x_i$ for every $1 \leq i \leq k$. The *distance* $d_G(x, y)$ between x and y in G is the length of a shortest path from x to y in G; if there is no such path, we set $d_G(x, y) = \infty$. We will use the notation d_{xy} to denote the distance between x and y in the undirected graph G when the graph is clear from the context.

A *centrality measure* is a function $c : V_G \to \mathbf{R}$, assigning real values, which we also refer to as *scores*, to all the vertices of a graph G, where vertices with larger scores should be interpreted as having a greater structural importance in the network. We assume all centrality measures to be invariant under isomorphism.

We start by recalling the definition of score and rank monotonicity on undirected graphs, which were introduced in Boldi et al. (2023) as a natural extension of score and rank monotonicity on directed graphs (Boldi and Vigna 2019, Boldi et al. 2017). Score monotonicity requires that when adding an edge to a graph, both endpoints enjoy an increase in score. Formally:

Definition 2.1 (Score monotonicity) Given an undirected graph G, a centrality c is said to be *score monotone on G* if for every pair of distinct,[1] non-adjacent vertices x and y we have that

$$c(x) < c'(x) \quad and \quad c(y) < c'(y),$$

where c' is the value of the centrality on the graph $G + x — y$.

Note that, in general, we will give definitions of properties for a centrality c on a graph G, and we will say that the same property holds *on a set of graphs* if it is true on all the graphs from the set.

A score increase does not imply that the rank relations between the two vertices involved in the new edge and the other vertices in the network remain unchanged. This observation motivates the definition of rank monotonicity (Boldi et al. 2017, 2023): this property requires that every vertex that used to be dominated by two non-adjacent vertices x and y is still dominated after the addition of the edge $x — y$.[2] Formally, we can consider two versions of rank monotonicity:

Definition 2.2 (Rank monotonicity) Given an undirected graph G, a centrality c is said to be *rank monotone on G* if for every pair of distinct, non-adjacent vertices x and y the following two statements hold:

- for all vertices $z \neq x, y$:

$$c(z) < c(x) \quad \text{implies} \quad c'(z) < c'(x) \text{ and}$$
$$c(z) = c(x) \quad \text{implies} \quad c'(z) \leq c'(x),$$

- for all vertices $z \neq x, y$:

$$c(z) < c(y) \quad \text{implies} \quad c'(z) < c'(y) \text{ and}$$
$$c(z) = c(y) \quad \text{implies} \quad c'(z) \leq c'(y).$$

where c' is the value of the centrality on the graph $G + x — y$.

A strict (in fact, simpler) version of the previous definition can be introduced, requiring that adding the new edge breaks all ties in favor of the two endpoints of the new edge itself:

Definition 2.3 (Strict rank monotonicity) Given an undirected graph G, a centrality c is said to be *strictly rank monotone on G* if for every pair of distinct, non-adjacent vertices x and y the following two statements hold:

- for all vertices $z \neq x, y$:
  $$c(z) \leq c(x) \quad \text{implies} \quad c'(z) < c'(x),$$
- for all vertices $z \neq x, y$:
  $$c(z) \leq c(y) \quad \text{implies} \quad c'(z) < c'(y).$$

where c' is the value of the centrality on the graph $G + x — y$.

Semi-monotonicity is a weaker condition than monotonicity, originating from the observation that in undirected networks some centrality measures do not satisfy score or rank monotonicity. Indeed, adding an edge to a graph can reduce the score or the rank of one of the two endpoints of the edge (Boldi et al. 2023). Thus, we just require that adding a new edge increases the score of *at least one* of the two endpoints:

Definition 2.4 (Score semi-monotonicity) Given an undirected graph G, a centrality c is said to be *score semi-monotone on G* if for every pair of distinct, non-adjacent vertices x and y we have that

$$c(x) < c'(x) \quad or \quad c(y) < c'(y).$$

where c' is the value of the centrality on the graph $G + x — y$.

Similarly, rank semi-monotonicity means that adding a new edge increases the rank of *at least one* of the two endpoints, and as before comes in two versions:

Definition 2.5 (Rank semi-monotonicity) Given an undirected graph G, a centrality c is said to be *rank semi-monotone on G* if for every pair of distinct, non-adjacent vertices x and y at least one of the following two statements holds:

- for all vertices $z \neq x, y$:

$$c(z) < c(x) \quad \text{implies} \quad c'(z) < c'(x) \text{ and}$$
$$c(z) = c(x) \quad \text{implies} \quad c'(z) \leq c'(x),$$

- for all vertices $z \neq x, y$:

$$c(z) < c(y) \quad \text{implies} \quad c'(z) < c'(y) \text{ and}$$
$$c(z) = c(y) \quad \text{implies} \quad c'(z) \leq c'(y).$$

where c' is the value of the centrality on the graph $G + x — y$.

[1] It is unfortunate that the assumption $x \neq y$ was not stated explicitly in the definition of score monotonicity given in Boldi and Vigna (2014), although being part of Sabidussi's original definition (Sabidussi 1966); the same restriction is necessary for strict rank monotonicity (Boldi et al. 2017). Indeed, adding loops cannot change the value of any centrality depending on shortest paths, so no such centrality could be score monotone or strictly rank monotone without the assumption $x \neq y$. On the other hand, the results in Boldi and Vigna (2019), Boldi et al. (2017) show that spectral centralities satisfy a stronger definition of strict rank monotonicity that includes the possibility of adding loops.

[2] The original definition of rank monotonicity for directed graphs was given in Chien et al. (2003).

Again, the strict version is simpler:

Definition 2.6 (Strict rank semi-monotonicity) Given an undirected graph G, a centrality c is said to be *strictly rank semi-monotone on* G if for every pair of distinct, non-adjacent vertices x and y at least one of the following two statements holds:

- for all vertices $z \neq x, y$:
  $c(z) \leq c(x)$ implies $c'(z) < c'(x)$,
- for all vertices $z \neq x, y$:
  $c(z) \leq c(y)$ implies $c'(z) < c'(y)$.

where c' is the value of the centrality on the graph $G + x — y$.

3 δ–monotonicity and δ–semi–monotonicity

When proving rank monotonicity of a centrality measure, working purely on order relationship on the whole graph can be tricky: it is often easier to prove a stronger but more amenable property that provides a sufficient condition for rank monotonicity based on value differences, rather than on order: indeed, the authors of Boldi et al. (2017) used such a property to prove rank monotonicity for closeness centrality, harmonic centrality, PageRank, and Katz's index, without explicitly naming it.

This observation motivates us to introduce the concepts of δ-monotonicity and δ-semi-monotonicity, and to show that they indeed provide a sufficient condition for, respectively, rank monotonicity and rank semi-monotonicity. The idea underlying δ-monotonicity is that the score increase of the target of a new arc should be at least as large as the score increase of all other vertices. This approach formalizes the idea underlying the proofs of Boldi et al. (2017) and factors out the common structure of the proofs of rank semi-monotonicity in the rest of the paper.

We start with the definition of δ-monotonicity for directed graphs:

Definition 3.1 (δ-monotonicity (directed)) Given a directed graph G, a centrality c is said to be *δ-monotone on* G if for every pair of distinct vertices x and y such that $x \nrightarrow y$ the following holds:

$$c'(z) - c(z) \leq c'(y) - c(y) \qquad \text{for every } z \neq y.$$

where c' is the value of the centrality on the graph $G + x \rightarrow y$. It is said to be *strictly δ-monotone* if the inequality is strict.

The undirected version of δ-monotonicity follows the same pattern of the monotonicity properties of the previous section:

Definition 3.2 (δ-monotonicity (undirected)) Given an undirected graph G, a centrality c is said to be *δ-monotone on* G if for every pair of distinct, non-adjacent vertices x and y the following holds:

$$\begin{aligned} c'(z) - c(z) &\leq c'(x) - c(x) \qquad \text{for every } z \neq x, y \qquad \text{and} \\ c'(z) - c(z) &\leq c'(y) - c(y) \qquad \text{for every } z \neq x, y. \end{aligned}$$

where c' is the value of the centrality on the graph $G + x — y$. It is said to be *strictly δ-monotone* if the inequalities are strict.

Finally, only for undirected graphs we define δ-semi-monotonicity, whose definition parallels the weakening of monotonicity we discussed in the previous section:

Definition 3.3 (δ-semi-monotonicity) Given an undirected graph G, a centrality c is said to be *δ-semi-monotone on* G if for every pair of distinct, non-adjacent vertices x and y the following holds:

$$\begin{aligned} c'(z) - c(z) &\leq c'(x) - c(x) \qquad \text{for every } z \neq x, y \qquad \text{or} \\ c'(z) - c(z) &\leq c'(y) - c(y) \qquad \text{for every} z \neq x, y. \end{aligned}$$

where c' is the value of the centrality on the graph $G + x — y$. It is said to be *strictly δ-semi-monotone* if the inequalities are strict.

We now prove that δ-monotonicity implies rank monotonicity, and δ-semi-monotonicity implies rank semi-monotonicity. This is intuitive, as the score increase of the target of a new arc (or of the endpoints of a new edge, in the undirected case) is at least as large as the increase in score of all other vertices, so its order relationship with other vertices can only improve.

Theorem 3.1 *If a centrality measure is* (*strictly*) *δ-monotone on a* (*directed or undirected*) *graph then it is* (*strictly*) *rank monotone on the same graph.*

Proof We prove this result for a directed graph: a similar proof can be obtained for the undirected case. By the definition of δ-monotonicity, we know that when adding an arc $x \rightarrow y$ to G, the inequality

$$c'(z) - c(z) \leq c'(y) - c(y)$$

holds for every $z \neq y$, the inequality being strict in the strict case. Adding this inequality to the left-hand sides of the implication appearing in the definition of rank monotonicity, we obtain

$$\begin{aligned} c(z) < c(y) \quad &\text{implies} \quad c'(z) < c'(y) \quad \text{for every } z \neq y, \\ c(z) = c(y) \quad &\text{implies} \quad c'(z) \leq c'(y) \quad \text{for every } z \neq y. \end{aligned}$$

The proof for the strict case is analogous. □

Theorem 3.2 *If a centrality measure is (strictly) δ-semi-monotone on a graph then it is (strictly) rank semi-monotone on the same graph.*

The proof can be straightforwardly derived as for Theorem 3.1 using the appropriate hypotheses and definitions.

4 Basin dominance

Finally, we introduce an even stronger property than δ-semi-monotonicity, *basin dominance*: this property is related to the notion of distance in a graph, and it will be instrumental in the proofs of rank semi-monotonicity for centralities based on shortest paths in the next sections.

We start with the notion of *basin*, formalizing the idea that given two vertices x, y in an undirected graph we can classify the vertices of the graph depending on whether they are closer to x or to y, with possibly an overlap for the vertices that are equidistant from x and y. This concept appeared implicitly in some recent works (Brandes et al. 2022; Skibski 2023) where graph distances are used to determine winners of pairwise comparisons, in a *graph-as-an-election* fashion. In particular, between two candidate vertices x and y of a graph, voters give their preference to the one whom they are strictly closer to. In fact, we can trace the notion back at least to [Entringer et al. (1976), Property 2.2] (albeit in Entringer et al. (1976), Brandes et al. (2022), Skibski (2023) equidistant vertices are not considered).

Definition 4.1 (Basin) Given an undirected graph G with vertex set V_G and two vertices x and y, we define the *basin of x (with respect to y)* B_{xy} and the *basin of y (with respect to x)* B_{yx} as

$$B_{xy} = \{\, u \in V_G \mid d_{ux} \le d_{uy} \,\}$$
$$B_{yx} = \{\, u \in V_G \mid d_{uy} \le d_{ux} \,\}$$

Figure 1 shows an example of a graph with the basins of two vertices; the vertices that are equidistant from x and y are included in both basins. We note a few useful facts:

- if x is not adjacent to y, adding the edge x — y to the graph leaves the basins unchanged; moreover, the shortest paths between x and the vertices in its basin B_{xy} do not change (the same is true for y);
- if z is equidistant from x and y and they are non-adjacent, the shortest paths between z and other vertices do not change because of the addition of the edge x — y; as a consequence, the score of z will remain the same after the addition of x — y in any centrality depending only on shortest paths;
- let u be a vertex in the basin of x; then, a shortest path between u and x passes exclusively through vertices in the basin of x; the same holds symmetrically for a shortest path between $v \in B_{yx}$ and y; moreover, a shortest path between u and y passes first exclusively through the basin of x, and then exclusively through the basin of y; again, everything holds symmetrically for a shortest path between $v \in B_{yx}$ and x;
- shortest paths between arbitrary vertices can zig-zag between basins multiple times (see Fig. 2);
- if x — y, the difference of the sum of distances from x and from y is the opposite of the difference of the sizes of their basins (Entringer et al. 1976):

$$\sum_z d_{zx} - \sum_z d_{zy} = |B_{yx}| - |B_{xy}|.$$

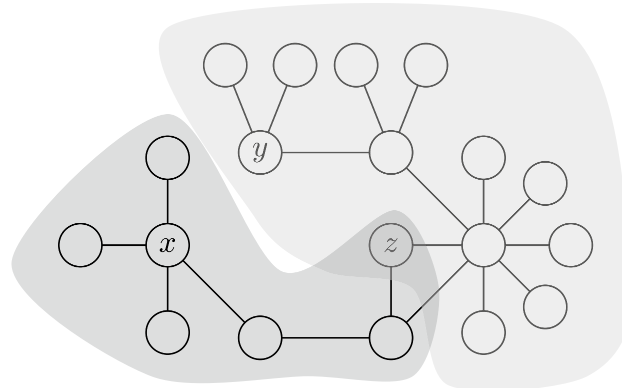

Fig. 1 An undirected graph G, with B_{xy} (the basin of x w.r.t. y) shown in dark grey and B_{yx} (the basin of y w.r.t. x) in light grey. Note that z is equidistant from x and y, and thus it is included in both basins

Given the notion of basin, we can define the *basin dominance* property, which requires that the increase in score of x and y is at least as large as (or larger than, in the strict case) the increase in score of all other vertices in their corresponding basin:

Definition 4.2 (Basin dominance) A centrality c is said to be *basin dominant* on an undirected graph G if for every pair of distinct, non-adjacent vertices x and y we have that

$$c'(u) - c(u) \le c'(x) - c(x) \qquad \text{for every } u \in B_{xy} \setminus \{x\} \text{and}$$
$$c'(v) - c(v) \le c'(y) - c(y) \qquad \text{for every } \ v \in B_{yx} \setminus \{y\}.$$

where c' is the value of the centrality on the graph $G + x$ — y. It is *strictly basin dominant* if both inequalities are strict.

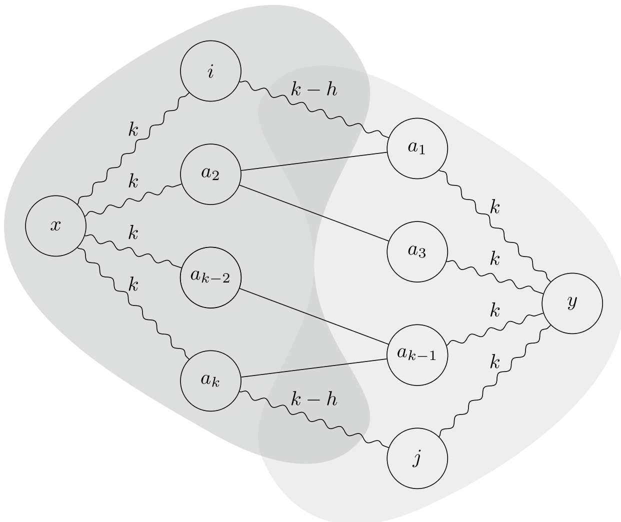

Fig. 2 A family of graphs where the shortest path between i and j can change basin $k+1$ times, for $k \geq 2$. B_{xy} is shown in dark grey and B_{yx} is shown in light grey. Each squiggly line with label ℓ represents a path of length ℓ between its two endpoints. Tuning the parameter h we can make the shortest path between i and j change basin $k+1$ times even after the addition of the edge x — y. For example, choosing $h = k-1$ we get $d_{ij} = d'_{ij} = k+1$ (where d'_{uv} is the distance between u and v *after* the addition of the edge x — y); moreover, $d_{ix} + d_{xj} = d_{iy} + d_{yj} = d'_{ix} + 1 + d'_{yj} = 2k+1$. On the other hand, choosing $h = 0$, the shortest path between i and j after the addition of the edge x — y will pass through x and y

At first glance, this property might seem weaker than δ-semi-monotonicity, as it requires a larger increase only for a subset of the vertices. However, it is immediate to notice that either x or y has the largest increase, and with respect to *that* vertex c will be δ-semi-monotone. Combining this observation with the results of the previous section we have that:

Theorem 4.1 *If a centrality measure is (strictly) basin dominant on a graph then it is (strictly) δ-semi-monotone. Hence, if a centrality measure is (strictly) basin dominant on a graph then it is (strictly) rank semi-monotone on the same graph.*

As it will become apparent in the following section, basins will enable us to reason about the score increase of a vertex in a more geometric way.

5 Distance–decay centralities

Conventions. In this section, and in the following ones, we will often be in a situation where x and y are distinct, non-adjacent vertices of an undirected graph G: we will then uniformly use c for the value of a centrality on G, and c' for the value of the same centrality on the graph $G + x$ — y. Analogously, we will denote distances in G with d_{uv} and in G' with d'_{uv}. More in general, we will use the *prime* symbol on quantities and functions related to G'.

Our first goal, now, is to prove that a large class of centrality measures are basin dominant, and thus δ-semi-monotone (hence, rank semi-monotone). In fact, we will be able to characterize basin dominance in terms of a discrete convexity condition. As a bonus, we will obtain indirectly that closeness centrality is rank semi-monotone.[3]

A *geometric centrality* (Boldi and Vigna 2014) is a centrality measure that depends only on distances between vertices. A special class of geometric centralities is the following:

Definition 5.1 (Distance-decay centrality) A *distance-decay centrality* is a centrality measure c for which there exists a nonincreasing *decay function* $\alpha : \mathbf{N} \setminus \{0\} \to \mathbf{R}$ such that[4]

$$c(v) = \sum_{\substack{u \neq v \\ d_{uv} < \infty}} \alpha(d_{uv}).$$

The idea that the influence of a vertex on the centrality of another is computed additively on some nonincreasing function of its distance is very natural and it appeared several times in the literature (Beauchamp 1965; Cohen et al. 2014; Dangalchev 2006; Jackson 2008; Harris 1954; Pan 2011, just to cite a few), but there is no standard definition. We took the name and definition used from Cohen et al. (2014), where it is declined as “distance-decay closeness”, but given that closeness does not satisfy the definition (see Sect. 5.1), we prefer to use “distance-decay centrality”.[5]

To state our results, we need to recall the definition of the *discrete derivative* operator Δ (Graham et al. 1994), which given a function $f : \mathbf{N} \to \mathbf{R}$ returns the function $\mathbf{N} \to \mathbf{R}$ defined by

$$(\Delta f)(i) = f(i+1) - f(i).$$

[3] After submitting this paper, we noticed that some of the results of this section were previously and independently obtained in Kishi (1981).

[4] In the directed case we have a *positive* and a *negative* (distance-decay) centrality, the latter being usually that of interest, depending on whether we use d_{vu} or d_{uv} in the definition.

[5] The reader should not confuse the term *distance-decay centrality* with *decay centrality* (Jackson 2008; Dangalchev 2006), a measure that became popular in the economic literature. Note that, in Sect. 5.3, we will identify *decay centrality* as a special case of distance-decay centrality exhibiting *exponential decay*.

The operator can be iterated, and we denote with Δ^k the k-th iteration of the operator. As in the case of the standard derivative, $\Delta f \le 0$ if and only if f is nonincreasing.[6]

While the operator is normally defined for functions with domain $\mathbf{N}$, we will use it for functions with domain $\mathbf{N} \setminus \{0\}$ to avoid shifting the input of decay functions. For instance, the condition on the decay function is just that $\Delta\alpha \le 0$. Observe that when adding an edge x — y all distances decrease, so no score can decrease. Thus, the condition of the derivative of α being nonnegative is fairly natural; otherwise, adding a new edge might decrease the score of some vertex—a quite counterintuitive behavior.

By the same considerations, if we further require that $(\Delta\alpha)(1) < 0$ we have the following result:

Theorem 5.1 *A distance-decay centrality measure c is score monotone on connected undirected graphs iff the first derivative of its decay function α is negative at* 1, *that is, iff α satisfies* $(\Delta\alpha)(1) < 0$.

Proof For the right-to-left implication, we have $\alpha(1) > \alpha(2) \ge \alpha(k)$ for all $k \ge 2$. Thus, when adding the edge x — y to the graph, for all $z \ne x, y$ we have $\alpha\left(d'_{xz}\right) \ge \alpha(d_{xz})$, and $\alpha\left(d'_{xy}\right) > \alpha(d_{xy})$. For the reverse, assume by contradiction $\alpha(1) = \alpha(2)$ and consider the graph x — u — y: when adding an edge $x - y$ we have $c'(x) = 2\alpha(1) = \alpha(1) + \alpha(2) = c(x)$. □

The reader could be puzzled by the fact that a condition at a single point is equivalent to score monotonicity. However, the gap between $\alpha(1)$ and $\alpha(2)$ has a special role because when adding an edge x — y there is exactly one distance that turns into 1, that is, d_{xy}. Since in the graph x — u — y that distance is 2, the gap between $\alpha(1)$ and $\alpha(2)$ must be larger than zero for score monotonicity to happen. On the other hand, since all other coefficients are smaller than or equal to $\alpha(2)$ by definition, the nonzero gap between $\alpha(1)$ and $\alpha(2)$ induces a nonzero gap between $\alpha(1)$ and all other coefficients, which makes the condition sufficient. We will observe a similar phenomenon with strictness in basin dominance.

Interestingly, the *second* derivative gives us further insights into the inner workings of the centrality. As in the continuous case, a *convex* function f is a function such that $\Delta^2 f \ge 0$. Convexity gives us information about what happens when a distance changes because of the addition of an edge.

Indeed, since $(\Delta^2 f)(i) = f(i+2) - 2f(i+1) + f(i)$, we have that $\Delta^2 f \ge 0$ if and only if

$$f(i) - f(i+1) \ge f(i+1) - f(i+2) \quad \text{for every } i, \qquad (1)$$

that is, shortening a distance from $i+2$ to $i+1$ provides an increase in score that is not larger than the one given by a distance shortening from $i+1$ to i. We can generalize this fact by making the shortened distances more far apart, and the gap in the shortenings different:

Lemma 5.1 *If $f : \mathbf{N} \setminus \{0\} \to \mathbf{R}$ is convex, then for every $i \le j, k \ge 0$, and $0 \le \ell \le j - i$ we have*

$$f(i+k) - f(j+k-\ell) \le f(i) - f(j).$$

Moreover, if $i < j$, $k > 0$, and $\left(\Delta^2 f\right)(i) > 0$ (i.e., f is strictly convex at i), then the inequality is strict.

Proof By telescoping,

$$\begin{aligned} f(i) - f(j) &= \\ &= f(i) - f(i+1) + f(i+1) - f(i+2) + f(i+2) - \\ &\quad \cdots - f(j-1) + f(j-1) - f(j) \\ &\ge f(i+k) - f(i+k+1) + f(i+k+1) - f(i+k+2) \\ &\quad - \cdots - f(j+k) \\ &= f(i+k) - f(j+k) \ge f(i+k) - f(j+k-\ell), \end{aligned}$$

where the second inequality is strict if f is strictly convex at i, as

$$f(i) - f(i+1) > f(i+1) - f(i+2) \ge \cdots \ge f(i+k) - f(i+k-1).$$

□

We remark that if α is strictly convex at 1, that is, $\left(\Delta^2\alpha\right)(1) > 0$, we have that $\alpha(1) - \alpha(2) > \alpha(2) - \alpha(3) \ge 0$ by definition, so $(\Delta\alpha)(1) = \alpha(2) - \alpha(1) < 0$; we just proved that

Theorem 5.2 *If the decay function α of a distance-decay centrality is strictly convex at* 1, *that is,* $\left(\Delta^2\alpha\right)(1) > 0$, *then c is score monotone on connected undirected graphs.*

This is not all: convexity will give us results about basin dominance, and thus about rank monotonicity. This happens because Lemma 5.1, when read in the context of a decay function, tells us that when adding an arc from x to y, if some distance shortens from j to i, then it provides an increase in score that is at least the one given by a distance shortening from $j+k$ to $j+k-\ell$, and in fact more in the strict case. Thus, x will benefit of a new edge x — y more than any other vertex in the basin of x with respect to y, and analogously for y.

Using this observation, we can prove the following result:

Theorem 5.3 *If the decay function α of a distance-decay centrality c is convex, then c is basin dominant on connected*

[6] For a function $f : \mathbf{N} \setminus \{0\} \to \mathbf{R}$, we write $f \le 0$ ($f \ge 0$, respectively) if $f(x) \le 0$ ($f(x) \ge 0$, respectively) holds for every $x \in \mathbf{N} \setminus \{0\}$.

undirected graphs. Thus, it is δ-semi-monotone and rank semi-monotone. If it is furthermore strictly convex at 1, *that is*, $\left(\Delta^2\alpha\right)(1) > 0$, *then c is strictly basin dominant. Thus, it is strictly δ-semi-monotone and strictly rank semi-monotone.*

Proof Let c be a distance-decay centrality, and let x and y be two distinct non-adjacent vertices. We have to prove that $c'(u) - c(u) \le c'(x) - c(x)$ for every vertex $u \in B_{xy}$ (or <, if $\left(\Delta^2\alpha\right)(1) > 0$); we have that

$$c'(u) - c(u) = \sum_{z \neq u} \alpha\left(d'_{uz}\right) - \sum_{z \neq u} \alpha\left(d_{uz}\right) = \sum_{z \neq u}\left(\alpha\left(d'_{uz}\right) - \alpha\left(d_{uz}\right)\right).$$

We now consider each term of the summation, assuming without loss of generality that $c'(u) - c(u) > 0$ (the case $c'(u) = c(u)$ yields obviously the result we want, using Theorem 5.2 in the strict case).

If $d'_{uz} = d_{uz}$ then $\alpha\left(d'_{uz}\right) - \alpha\left(d_{uz}\right) = 0 \le \alpha\left(d'_{xz}\right) - \alpha\left(d_{xz}\right)$ (because adding a new edge cannot make the distance increase, and α is nonincreasing).

If instead $d'_{uz} < d_{uz}$, then $d'_{uz} = d'_{ux} + d'_{xz} > d'_{xz}$. Moreover, $d_{uz} \le d_{ux} + d_{xz}$ and $d_{ux} = d'_{ux}$, so subtracting $d_{uz} \le d_{ux} + d_{xz}$ from $d'_{uz} = d'_{ux} + d'_{xz}$ we obtain

$$d'_{uz} - d_{uz} \ge d'_{xz} - d_{xz}.$$

We now note that by the convexity of α, we can apply Lemma 5.1 with $i = d'_{xz}$, $j = d_{xz}$, $k = d'_{uz} - i$ and $\ell = j + k - d_{uz}$, obtaining

$$\alpha\left(d'_{uz}\right) - \alpha\left(d_{uz}\right) \le \alpha\left(d'_{xz}\right) - \alpha\left(d_{xz}\right),$$

where the inequality is strict if $\left(\Delta^2\alpha\right)(1) > 0$ and $z = y$.

Hence, by combining the two cases we have

$$\begin{aligned} c'(u) - c(u) &= \sum_{z \neq u,x}\left(\alpha\left(d'_{uz}\right) - \alpha(d_{uz})\right) + \alpha\left(d'_{ux}\right) - \alpha(d_{ux}) \\ &\le \sum_{z \neq x,u}\left(\alpha\left(d'_{xz}\right) - \alpha(d_{xz})\right) + \alpha\left(d'_{xu}\right) - \alpha(d_{xu}) \\ &= c'(x) - c(x), \end{aligned}$$

where, once again, inequality is strict if $\left(\Delta^2\alpha\right)(1) > 0$, as it is strict when $z = y$. □

We can also prove the converse of Theorem 5.3, showing that basin dominance and convexity conditions are tightly coupled:

Theorem 5.4 *If a distance-decay centrality c is basin dominant, then its decay function is convex. If it is strictly basin dominant, its decay function is further strictly convex at* 1, *that is*, $\left(\Delta^2\alpha\right)(1) > 0$.

Proof By contradiction, assume that for some $i \ge 1$ we have $\alpha(i) - \alpha(i+1) < \alpha(i+1) - \alpha(i+2)$. For the case $i \ge 2$, consider the graph on the left in Fig. 3. We have that after the addition of the edge $x — y$, the variation of the centrality of x is of the form $t_x + k \cdot (\alpha(i) - \alpha(i+1))$, whereas the variation of the centrality of u is of the form $t_u + k \cdot (\alpha(i+1) - \alpha(i+2))$. For k sufficiently large, the latter is greater than the former.

For the case $i = 1$, assume again by contradiction $\alpha(1) - \alpha(2) < \alpha(2) - \alpha(3)$ and consider the graph on the right in Fig. 3. We have

$$\begin{aligned} c(x) &= 2\alpha(1) + \alpha(2) \qquad & c'(x) &= 3\alpha(1) \\ c(u) &= \alpha(1) + \alpha(2) + \alpha(3) \qquad & c'(u) &= \alpha(1) + 2\alpha(2) \end{aligned}$$

so $c'(x) - c(x) = \alpha(1) - \alpha(2) < \alpha(2) - \alpha(3) = c'(u) - c(u)$.

The last statement is easily proved using the same argument we just used for $i = 1$, but assuming by contradiction $\alpha(1) - \alpha(2) = \alpha(2) - \alpha(3)$ and concluding $c'(x) - c(x) = c'(u) - c(u)$ as a consequence. □

In the next few sections, we will apply our characterization of basin dominance to a few geometric centrality measures.

5.1 Closeness centrality

Closeness centrality (Bavelas 1948, 1950) is one of the oldest centrality measures in the literature. It was shown to be score monotone but not rank monotone on connected undirected networks (Boldi et al. 2023). Moreover, in the latter paper it was left as an open problem whether closeness was (in our terminology) rank semi-monotone or not. In the rest of this section, we will solve this open problem by showing that closeness is in fact rank semi-monotone, but not in a strict way.

Recall that the *peripherality* of a vertex x is the sum of the distances between v and all the other vertices of G:

$$p(v) = \sum_u d_{uv}.$$

As usual, in the directed case there is a positive and a negative peripherality, but we will be concerned with the connected, undirected case only, for which the two notions coincide. The *closeness centrality* of v is just the reciprocal of its peripherality:

$$C(v) = \frac{1}{p(v)}.$$

To prove rank semi-monotonicity for closeness, we will have to pass through the centrality defined by negated peripherality. This is because

$$-p(v) < -p(u) \iff C(v) < C(u),$$

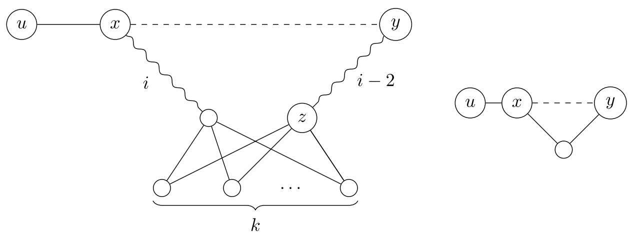

Fig. 3 Graphs used in the proof of Theorem 5.4: the one on the right is used for the case $i = 1$, whereas the left one covers $i \geq 2$. Note that, when $i = 2$, the vertices labeled with y and z on the left graph collapse to the same vertex. Each squiggly line with label ℓ represents a path of length ℓ between its two endpoints

so negated peripherality is (strictly) semi-rank monotone if and only if closeness is.

Lemma 5.2 *Negated peripherality is basin dominant on connected undirected graphs. Thus, it is δ-semi-monotone and rank semi-monotone on the same graphs.*

Proof By Theorem 5.3, since negated peripherality is a distance-decay centrality with decay function $\alpha(i) = -i$, which satisfies $\Delta^2\alpha = 0$. □

We thus obtain

Theorem 5.5 *Closeness centrality is rank semi-monotone on connected undirected graphs.*

Note that, maybe counterintuitively, closeness centrality is *not* basin dominant. An erroneous statement to this effect appears in [Boldi et al. (2024), Lemma 1], but the counterexample in Fig. 4 shows that this is not the case, as

$$C'(u) - C(u) = \frac{1}{k+5} - \frac{1}{k+6} > \frac{1}{2k+3} - \frac{1}{2k+4} = C'(x) - C(x)$$

for all $k \geq 3$. The main statement of [Boldi et al. (2024), Theorem 3], however, is correct: closeness is rank semi-monotone, as we have just stated.

The lack of basin dominance is part of the counterintuitive behavior of closeness: it is not strictly rank monotone, not even on directed graphs (Boldi et al. 2023); it is not score monotone, either, on directed graphs unless the graph is strongly connected (Boldi and Vigna 2014).

The problem lies in the reciprocation used to make closeness increase when peripherality decreases: because of reciprocation, when we add an edge $x — y$ the distances that are shortened by a certain amount d have the same influence on peripherality, but the effect on scores will depend on the original score of the vertex. In the example, u has a greater centrality than x, and thus benefits more than x by the reduction of 1 of peripherality.

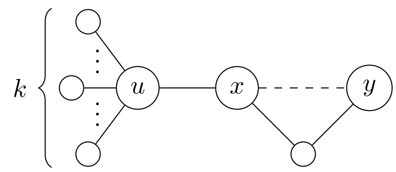

Fig. 4 A graph showing that closeness centrality is not basin dominant. For all $k \geq 3$, the closeness centrality of u increases more than that of x when we add the edge $x — y$

Harmonic centrality (see the next section) solves this problem by reciprocating distances instead of the whole sum.

We conclude this section by showing that:

Theorem 5.6 *Closeness centrality is not strictly rank semi-monotone on (an infinite family of) connected undirected graphs.*

Proof Consider the graphs in Fig. 5, where $u \in B_{xy}$. This is an infinite family of graphs with a parameter k which controls the sizes of the two stars around vertices w and y. Computing the peripheralities of u, x and y before and after the addition of $x — y$, we obtain

$$p(u) = 2 \cdot (k+4) + 4 \cdot k + 13 \qquad p'(u) = 2 \cdot (k+4) + 3 \cdot k + 12$$
$$p(x) = 3 \cdot (k+4) + 3 \cdot k + 9 \qquad p'(x) = 3 \cdot (k+4) + 2 \cdot k + 8$$
$$p(y) = 4 \cdot (k+4) + k + 15 \qquad p'(y) = 4 \cdot (k+4) + k + 12.$$

For all $k \geq 10$, we have that

$$p(x) = p(u), \quad p'(x) = p'(u), \quad p(y) \leq p(u), \quad p'(y) > p'(u),$$

showing that closeness is not semi-monotone at y (because y used to be at least as central as u, but it is less central after the addition of the edge) and not strictly rank semi-monotone at x (it is always as central as x, before and after adding the edge). □

Note how the counterexample shows a graph where y has a smaller basin than x but a greater score, losing rank after the addition of the edge (see the last fact we reported about basins in Sect. 4). Moreover, note that the degree of y can become arbitrarily large, while the degree of x and u are fixed. Nevertheless, x and u have a greater score than y in G' for $k \geq 10$, notwithstanding its degree.

We will see that harmonic centrality behaves very differently for the graph in Fig. 5, yet showing some counterintuitive results.

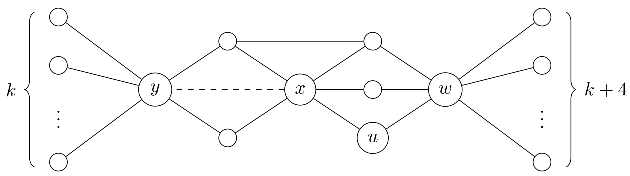

Fig. 5 A counterexample to strict rank semi-monotonicity for closeness centrality. For all $k \geq 10$, u and x have the same score before and after the addition of the edge x — y. Moreover, u has the same score of y (or smaller) before the addition, but a higher score after the addition, breaking strict rank semi-monotonicity

5.2 Harmonic centrality

Harmonic centrality (Beauchamp 1965) solves the issue of unreachable vertices in closeness centrality. In particular, if we assume $\infty^{-1} = 0$, we can define it as

$$h(v) = \sum_{u \neq v} \frac{1}{d_{uv}},$$

so that unreachable vertices have no impact on the summation and, thus, on the final centrality score of the node. On undirected graphs, it is score monotone but not rank monotone, as shown in Boldi et al. (2023), where the same counterexample disproving rank monotonicity for closeness centrality also shows that harmonic centrality fails at satisfying this axiom. We can however leverage our results on distance-decay centralities to prove strict basin dominance:

Theorem 5.7 *Harmonic centrality is strictly basin dominant on connected undirected graphs. Thus, it is strictly δ-semi-monotone and strictly rank semi-monotone on the same graphs.*

Proof By Theorem 5.3, as harmonic centrality is a distance-decay centrality with decay function $\alpha(i) = 1/i$, which is strictly convex, as $1/i - 1/(i+1) > 1/(i+1) - 1/(i+2)$. The rest follows by Theorem 4.1. □

The stronger result we can give for harmonic centrality should be compared to the fact that on strongly connected graphs harmonic centrality is strictly rank monotone, whereas closeness centrality is just rank monotone (Boldi et al. 2023). Following the example in Fig. 5, we can show different behaviors for harmonic centrality as k varies. Computing the harmonic centrality of u, x and y before and after the addition of x — y as a function of k, we have:

$$h(u) = \frac{1}{2}\cdot(k+4) + \frac{1}{4}\cdot k + \frac{13}{3} \qquad h'(u) = \frac{1}{2}\cdot(k+4) + \frac{1}{3}\cdot k + \frac{9}{2}$$

$$h(x) = \frac{1}{3}\cdot(k+4) + \frac{1}{3}\cdot k + 6 \qquad h'(x) = \frac{1}{3}\cdot(k+4) + \frac{1}{2}\cdot k + \frac{13}{2}$$

$$h(y) = \frac{1}{4}\cdot(k+4) + k + 4 \qquad h'(y) = \frac{1}{4}\cdot(k+4) + k + \frac{29}{6}.$$

Note that for such a graph, the increase in score of y is a constant, while the increase in score of x and u is a function of k. For $k \geq 2$, we have that $h'(y) - h(y) \leq h'(x) - h(x)$. In particular, for $k = 4$, $h(y) = h(x)$ hence $h'(y) < h'(x)$, even if the size of the neighborhood of y is larger than that of x. For $k \geq 5$, $h'(y) > h'(x)$ even if y increase its score less than x. Finally, for $k \geq 8$, $h'(y) > h'(u)$. This example shows how the behavior of harmonic centrality can be quite tricky to predict.

5.3 Further distance–decay centralities

As we already mentioned, other decay functions have been considered, for example *exponential decay*, where $\alpha(i) = \xi^i$ for some $0 < \xi < 1$ (Jackson 2008; Dangalchev 2006). One can also consider *power-law decay*, where $\alpha(i) = 1/i^k$ for some $k > 0$ (note that the case $k = 1$ is harmonic centrality). Armed with our results, we can easily prove the following theorem:

Theorem 5.8 *Distance-decay centralities with exponential decay or power-law decay are score monotone, strictly basin dominant, strictly δ-semi-monotone, and strictly rank semi-monotone.*

Proof By Theorem 5.2 and Theorem 5.3, as

$$\xi^i - \xi^{i+1} > \xi\left(\xi^i - \xi^{i+1}\right) = \xi^{i+1} - \xi^{i+2}$$

and

$$\frac{1}{i^k} - \frac{1}{(i+1)^k} = \frac{(i+1)^k - i^k}{i^k(i+1)^k} = \left(\left(\frac{i+1}{i}\right)^k - 1\right)\frac{1}{(i+1)^k}$$
$$> \left(\left(\frac{i+2}{i+1}\right)^k - 1\right)\frac{1}{(i+2)^k} = \frac{(i+2)^k - (i+1)^k}{(i+1)^k(i+2)^k}$$
$$= \frac{1}{(i+1)^k} - \frac{1}{(i+2)^k},$$

so in both cases the decay function is strictly convex everywhere. □

6 Betweenness centrality

Betweenness centrality (Anthonisse 1971; Freeman 1977) tries to capture the idea that a vertex is central if it lies on many shortest paths between other vertices. It does not focus on the length of shortest paths but on how many of them involve a given node, trying to estimate the amount of flow passing through nodes in a network. For this reason, betweenness is not a geometric measure.

Formally, if we call σ_{ij} the number of shortest paths between two vertices i and j and $\sigma_{ij}(v)$ the number of such paths passing through a vertex v, then we can define the betweenness centrality of v as

$$b(v) = \sum_{\substack{i,j \neq v \\ \sigma_{ij} > 0}} \frac{\sigma_{ij}(v)}{\sigma_{ij}}.$$

A few variants of this definition appear in the literature:

- Contrarily to Anthonisse's original definition (Anthonisse 1971), which was stated on directed graphs in the form above, Freeman's later definition (Freeman 1977) was stated on undirected graphs using *unordered* pairs, and indeed his summations have the condition $i < j$. The form above, however, is the one usually found in current literature. With respect to Freeman's definition, however, the value is doubled; this difference is irrelevant as far as score and rank monotonicity are concerned.
- Some authors exclude explicitly the case $i = j$, others do not. Since $\sigma_{ii} = 1$ and $\sigma_{ii}(v) = 0$ for all $v \neq i$, the resulting value does not change, and we will thus assume $i \neq j$ in our proofs.
- Some authors give the definition without the condition $\sigma_{ij} > 0$, in which case the definition applies only to (strongly) connected graphs. In the undirected case, however, the resulting definition is equivalent to defining the centrality on each connected component separately, so we will consider the connected case only, and drop the condition $\sigma_{ij} > 0$ in the proofs.

As in the previous sections, we denote with σ and σ' the number of shortest paths before and after the addition of an edge $x — y$, and with b and b' the betweenness centrality before and after the addition of the edge.

We know from Boldi et al. (2023) that betweenness is neither rank nor score monotone. Nonetheless, we can show that the betweenness of two vertices can never decrease after we link them with a new edge: this result was also proved in [Bergamini et al. (2018), Theorem 5.2], although with a slightly weaker statement; we thus provide a full proof here for the sake of completeness:

Lemma 6.1 *The following property holds when adding the edge $x — y$:*

$$\frac{\sigma'_{ij}(x)}{\sigma'_{ij}} - \frac{\sigma_{ij}(x)}{\sigma_{ij}} \geq 0 \qquad \text{for all } i,j \neq x.$$

As a consequence, $b'(x) \geq b(x)$. The same statements are true for y.

Proof Given vertices $i,j \neq x$, let us call p_x ($p_{\bar{x}}$, respectively) the number of shortest paths between i and j passing (not passing, resp.) through x in G. As usual, let us refer to the same quantities in G' with p'_x ($p'_{\bar{x}}$, resp.). We have to show that the following holds:

$$\frac{\sigma'_{ij}(x)}{\sigma'_{ij}} - \frac{\sigma_{ij}(x)}{\sigma_{ij}} = \frac{p'_x}{p'_x + p'_{\bar{x}}} - \frac{p_x}{p_x + p_{\bar{x}}} \geq 0.$$

Summing over all $i,j \neq x$ proves the second part of the statement.

We consider two cases:

- if $d'_{ij} < d_{ij}$, all shortest paths in G' between i and j pass through the edge $x — y$ (hence, through x). Thus, we obtain:

$$1 - \frac{p_x}{p_x + p_{\bar{x}}} \geq 0,$$

  which is clearly true.
- if $d'_{ij} = d_{ij}$, then all shortest paths between i and j in G are still shortest paths in G', but there may be some new ones passing through x, so $p'_{\bar{x}} = p_{\bar{x}}$ and $p'_x \geq p_x$. Letting $\alpha = p'_x - p_x \geq 0$, we obtain:

$$\frac{p_x + \alpha}{p_x + \alpha + p_{\bar{x}}} - \frac{p_x}{p_x + p_{\bar{x}}} = \frac{p_x^2 + p_x p_{\bar{x}} + \alpha p_x + \alpha p_{\bar{x}} - (p_x^2 + \alpha p_x + p_x p_{\bar{x}})}{(p_x + \alpha + p_{\bar{x}})(p_x + p_{\bar{x}})} = \frac{\alpha p_{\bar{x}}}{(p_x + \alpha + p_{\bar{x}})(p_x + p_{\bar{x}})} \geq 0,$$

  which is again true, concluding the proof.

□

Incidentally, observe that we can always tell if the betweenness centrality of a vertex is zero without actually computing it. In fact, denoting with $N_G(v)$ the *neighborhood* of v in G, that is, the set of vertices z such that $v - z$, we have:

Lemma 6.2 *Let G be a connected undirected graph and v a vertex of G. Then $b(v) = 0$ iff v is simplicial, that is, iff the subgraph induced by $N_G(v)$ is a clique.*

Proof If $b(v) = 0$ no shortest paths are passing through v: but then any two neighbors of v must be adjacent, or otherwise they would have distance 2, and there would be a path through v of length 2. Conversely, suppose that v is simplicial and consider vertices $i,j \neq v$. A shortest path between i

and j cannot involve v, otherwise it would touch two neighbors of v, say i', j', and we might shorten it by skipping v. □

An immediate consequence is the following result:

Theorem 6.1 *Betweenness centrality is neither score semi-monotone, nor strictly rank semi-monotone on (an infinite family of) connected undirected graphs.*

Proof Consider the family of graphs shown in Fig. 6, in which we have a clique connected to three vertices u, x and y (the unnamed nodes form a clique of arbitrary size k, in the picture we show the case $k = 4$). Since they all have the same neighborhood, and that neighborhood is a clique, their betweenness is zero. But when we add the edge $x — y$ this property is still true, so their betweenness remains zero. □

We now prove a sufficient condition for the betweenness of x and y to remain unchanged when adding the edge $x — y$:

Lemma 6.3 *Let x and y be two distinct non-adjacent vertices of G. If $N_G(x) \setminus \{x\} = N_G(y) \setminus \{y\}$ then $b'(x) = b(x)$ and $b'(y) = b(y)$.*

Proof Assume by contradiction, without loss of generality (see Lemma 6.1), that $b'(x) > b(x)$. Then, there must be at least one pair of distinct vertices i, j with a shortest path linking them passing through both x and y via the new edge $x — y$. Thus, $d_{ij} \geq d'_{ij} = d_{iu} + 2 + d_{yj}$ for some neighbor u of x. However, at the same time $d_{ij} \leq d_{iu} + 1 + d_{yj}$ because u is also a neighbor of y, leading to a contradiction. □

Note that we remove x and y from their neighborhoods to avoid that loops prevent the application of the lemma.

We are now going to show that betweenness centrality is rank semi-monotone on connected undirected graphs. In fact, we show that it even enjoys basin dominance, which is rather surprising given that basin dominance depends on the length of shortest paths, while betweenness depends on the fraction of shortest paths passing through a vertex. Moreover, it is rather hard to find interesting axioms satisfied by betweenness (Boldi and Vigna 2014).

We start by proving a technical lemma that relates certain products of numbers of shortest paths. There are several cases to consider, but the most interesting one is depicted in Fig. 7. Consider a situation where i is strictly closer to x than to y, and j is strictly closer to y than to x. Moreover, fix some (arbitrary) subset of vertices α in the basin of x that we need to go through, and some subset of vertices μ in the basin of x that we need to avoid. The dotted line represents

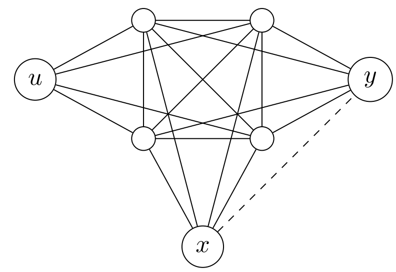

Fig. 6 Simple counterexample for score semi-monotonicity and strict rank semi-monotonicity for betweenness centrality. The dashed edge is the $x — y$ edge that we add to G, obtaining G'. The betweenness score of vertices x, y and u is 0 both in G and G'. This is true regardless of the size k of the clique (in the picture, $k = 4$)

a (generic) path from[7] i to j satisfying these traversal conditions and passing through x, whereas the dashed line represents a (generic) shortest path from i to j passing through x.

The key observation is that when we add the edge $x — y$, the only part of shortest paths going from i to j and passing through x that can change is the part from x to j (regardless of whether we are looking at the paths of dotted or dashed type). So the number of dashed/dotted paths in G or G', given that they are all nonzero, is the product of the number of the paths from i to x by the number of the paths from x to j, and only the second factor changes when we add the edge $x — y$. Moreover, this second factor is the same for both counts. Thus, if we multiply the counts of the dashed and dotted paths, as long as one count is taken in G and the other is taken in G' the product is the same—we're just pairing the paths from i to x and from x to j in different ways, or, with a slogan, we can "move the prime":

Lemma 6.4 *Let G be a connected undirected graph, $x, y \in V_G$ two non-adjacent vertices and $i, j \in V_G$ two vertices. Let also $\alpha, \mu \subseteq B_{xy} \setminus \{x\}$ be two subsets of the basin of x not containing x. Call p_x ($p_{\bar{x}}$, respectively) the number of shortest paths from i to j in G passing through x (not passing through x, resp.). Let also $p_{\alpha\bar{\mu}x}$ be the number of paths from i to j passing through all the vertices of α and then through x, but never passing through any of the vertices of μ before reaching x, and $p_{\alpha\bar{\mu}\bar{x}}$ the number of paths from i to j passing through all the vertices of α but never passing through x or any of the vertices of μ. Define similarly p'_x, $p'_{\bar{x}}$ $p'_{\alpha\bar{\mu}x}$, and $p'_{\alpha\bar{\mu}\bar{x}}$ for the paths satisfying the same conditions, but in G'. Then, the following properties hold:*

[7] Note that paths are sequences of vertices, so they have a direction also in an undirected graph. Of course, any path from i to j can be reversed in an undirected graph, obtaining a path from j to i.

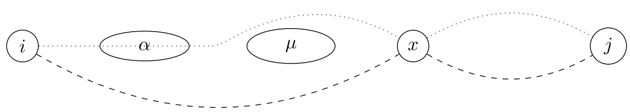

Fig. 7 A graph showing the main case of Lemma 6.4

(1.) $p'_{\alpha\bar{\mu}x}p_x = p_{\alpha\bar{\mu}x}p'_x$

(2.) $p'_{\alpha\bar{\mu}\bar{x}}p_{\bar{x}} = p_{\alpha\bar{\mu}\bar{x}}p'_{\bar{x}}$

(3.) $p'_{\alpha\bar{\mu}\bar{x}}p_x \leq p_{\alpha\bar{\mu}\bar{x}}p'_x$

(4.) $p'_{\alpha\bar{\mu}x}p_{\bar{x}} \geq p_{\alpha\bar{\mu}x}p'_{\bar{x}}$

Proof We can assume w.l.o.g. that $i \in V_G \backslash B_{yx}$ and $j \in V_G \backslash B_{xy}$, otherwise everything holds trivially because the new edge x — y does not create new shortest paths from i to j, hence all the counts of shortest paths from i to j in G and in G' are the same.

We now prove separately each of the four statements.

(1.) If $p_x = 0$ then $p_{\alpha\bar{\mu}x} = 0$, so the statement holds, and the same happens for $p'_x = 0$, so we can assume $p_x, p'_x \neq 0$, and in particular $d_{ij} = d_{ix} + d_{xj}$, $d'_{ij} = d'_{ix} + d'_{xj}$, and $d'_{ix} = d_{ix}$, as shortest paths between i and x are not affected by the new edge.

Thus, $p_{\alpha\bar{\mu}x} \neq 0$ iff there is a path in G from i to x of length d_{ix} that passes through α but not through μ, but since $d'_{ix} = d_{ix}$ this is true iff there is a path in G' from i to x of length d'_{ix} that passes through α but not through μ, that is, iff $p'_{\alpha\bar{\mu}x} \neq 0$. Thus, we can also assume $p_{\alpha\bar{\mu}x}, p'_{\alpha\bar{\mu}x} \neq 0$.

Now let s be the number of shortest paths in G from i to x, and $s_{\alpha\bar{\mu}}$ the number of such paths that pass through vertices of α and do not pass through any of the vertices of μ (see Fig. 7): these values do not change in G' because the new edge x — y can only shorten the paths from x to j. Let also t and t' be the number of shortest paths from x to j in G and G', respectively. Then,

$$p'_{\alpha\bar{\mu}x}p_x = s_{\alpha\bar{\mu}}t'st = s_{\alpha\bar{\mu}}tst' = p_{\alpha\bar{\mu}x}p'_x.$$

(2.) For the shortest paths not passing through x, there are two possibilities:

- If $d'_{ij} = d_{ij}$ then $p'_{\alpha\bar{\mu}\bar{x}} = p_{\alpha\bar{\mu}\bar{x}}$ and $p'_{\bar{x}} = p_{\bar{x}}$, since all the shortest paths in G are also shortest in G', and more are possibly added but they all pass through x.
- If $d'_{ij} < d_{ij}$ then $p'_{\alpha\bar{\mu}\bar{x}} = 0 = p'_{\bar{x}}$, because in G' all the shortest paths between i and j must pass through x.

In both cases, the equality holds.

(3.) When we add x — y, we have two possibilities:

- If $d'_{ij} = d_{ij}$ then $p'_{\alpha\bar{\mu}\bar{x}} = p_{\alpha\bar{\mu}\bar{x}}$ and $p_x \leq p'_x$;
- If $d'_{ij} < d_{ij}$ then $p'_{\alpha\bar{\mu}\bar{x}} = 0$.

In both cases, the inequality holds.

(4.) Multiplying (1.) and (2.) we have $p'_{\alpha\bar{\mu}x}p_x p'_{\alpha\bar{\mu}\bar{x}}p_{\bar{x}} = p_{\alpha\bar{\mu}x}p'_x p_{\alpha\bar{\mu}\bar{x}}p'_{\bar{x}}$, and dividing by (3.) we obtain the statement. □

We are now ready to prove basin dominance for betweenness centrality: the proof goes through a case-by-case analysis of the contribution of each summand to the difference in centrality, with the main case being covered by Lemma 6.4:

Theorem 6.2 *Betweenness centrality is basin dominant on connected undirected graphs. Thus, betweenness centrality is δ-semi-monotone and rank semi-monotone on connected undirected graphs.*

Proof Let us call $\Delta_u = b'(u) - b(u)$ the score difference for a vertex u and for every pair of distinct vertices $i \neq j$, let also

$$\Delta_u(i,j) = \frac{\sigma'_{ij}(u)}{\sigma'_{ij}} - \frac{\sigma_{ij}(u)}{\sigma_{ij}}.$$

Obviously

$$\Delta_u = \sum_{i,j \neq u} \Delta_u(i,j).$$

We want to show that $\Delta_u \leq \Delta_x$ for every $u \in B_{xy}$. For vertices u being equidistant from x and y, everything holds trivially because the new edge x — y does not create any shortest path between i and j passing through u. Then, we can restrict our attention to the case where u is strictly closer to x than to y, that is, $u \in V_G \backslash B_{yx}$.

The two summations giving Δ_u and Δ_x happen on a different set of pairs of indices, and we treat the common and non-common pairs separately.

The easiest case is that of pairs i, j that appear in the summation of Δ_x but not in the summation of Δ_u, because we know from Lemma 6.1 that those summands are non-negative.

Then we consider the pairs i, j that appear in the summation of Δ_u but not in the summation of Δ_x, that is, those where either i or j are equal to x; without loss of generality let us assume $j = x$. We want to show that in this case, instead, we have

$$\Delta_u(i,x) = \frac{\sigma'_{ix}(u)}{\sigma'_{ix}} - \frac{\sigma_{ix}(u)}{\sigma_{ix}} \leq 0.$$

When $i \in B_{xy}$, the new x — y edge does not create any new shortest path between i and x, so $\Delta_u(i,x) = 0$. Conversely, when $i \notin B_{xy}$ new shortest paths cannot pass through u (remember that $u \in B_{xy}$); thus,

- if $d_{ix} = d'_{ix}$ then $\sigma_{ix} \leq \sigma'_{ix}$ and $\sigma_{ix}(u) = \sigma'_{ix}(u)$;
- if $d_{ix} > d'_{ix}$ then $\sigma'_{ix}(u) = 0$.

We are now left with the pairs i, j that appear in both summations, that is, $i,j \neq u$ and $i,j \neq x$. In this case, we want to prove a term-by-term bound, that is:

$$\Delta_u(i,j) = \frac{\sigma'_{ij}(u)}{\sigma'_{ij}} - \frac{\sigma_{ij}(u)}{\sigma_{ij}} \leq \frac{\sigma'_{ij}(x)}{\sigma'_{ij}} - \frac{\sigma_{ij}(x)}{\sigma_{ij}} = \Delta_x(i,j). \qquad (2)$$

Note that if i and j belong to the same basin the edge x — y does not create any new shortest path between i and j, so (2) holds because both sides are zero; this includes the case in which i and j are in the intersection of the basins, so we can assume, without loss of generality, that $i \in V_G \backslash B_{yx}$ and $j \in V_G \backslash B_{xy}$.

The case $\Delta_u(i,j) \leq 0$ is trivial because of Lemma 6.1; we now analyze the case $\Delta_u(i,j) > 0$.

Since $i, u \in V_G \backslash B_{yx}$ and $j \in V_G \backslash B_{xy}$, we can state two facts:

- all shortest paths in G and G' from i to j passing through x and y must pass through x before y;
- all shortest path in G and G' from i to j passing through u and x must pass through u before x.

As mentioned in Sect. 4, a shortest path between i and y passes exclusively through vertices in the basin of x and then exclusively through vertices in the basin of y, until it reaches y. Moreover, a shortest path between j and y passes exclusively through vertices in the basin of y. Combining these two facts is enough to show that the first statement holds.

The second one, instead, can be proven by contradiction: if u is after x in a shortest path from i to j, then u is necessarily between x and y. This means that $d'_{ij} < d_{ij}$ since we might shorten the path from x to y passing through u by taking x — y. Hence $\sigma'_{ij}(u) = 0$ and thus $\Delta_u(i,j) < 0$—a contradiction.

We now define a few counters of shortest paths from i to j in G satisfying certain conditions:

- p_x: passing through x;
- $p_{\bar{x}}$: not passing through x;
- p_{ux}: passing through u and then through x;
- $p_{\bar{u}x}$: passing through x but not through u;
- $p_{u\bar{x}}$ passing through u but not through x;
- $p_{\bar{u}\bar{x}}$ passing through neither.

The same notations are used for G', but we use p' instead of p.

With these notations, we can write the single terms appearing in (2) as follows:

$$\sigma_{ij}(u) = p_{ux} + p_{u\bar{x}} \quad \sigma'_{ij}(u) = p'_{ux} + p'_{u\bar{x}} \quad \sigma_{ij} = p_x + p_{\bar{x}}$$
$$\sigma_{ij}(x) = p_{ux} + p_{\bar{u}x} \quad \sigma'_{ij}(x) = p'_{ux} + p'_{\bar{u}x} \quad \sigma'_{ij} = p'_x + p'_{\bar{x}}$$

which makes us able to rewrite (2) as

$$\frac{p'_{ux} + p'_{u\bar{x}}}{\sigma'_{ij}} - \frac{p_{ux} + p_{u\bar{x}}}{\sigma_{ij}} \leq \frac{p'_{ux} + p'_{\bar{u}x}}{\sigma'_{ij}} - \frac{p_{ux} + p_{\bar{u}x}}{\sigma_{ij}},$$

which is equivalent to

$$\frac{p'_{u\bar{x}} - p'_{\bar{u}x}}{\sigma'_{ij}} \leq \frac{p_{u\bar{x}} - p_{\bar{u}x}}{\sigma_{ij}}.$$

Multiplying both sides by $\sigma'_{ij}\sigma_{ij}$, we obtain

$$p'_{u\bar{x}} \cdot \sigma_{ij} - p'_{\bar{u}x} \cdot \sigma_{ij} \leq p_{u\bar{x}} \cdot \sigma'_{ij} - p_{\bar{u}x} \cdot \sigma'_{ij}.$$

Now, it is enough to show that the two following inequalities hold:

$$p'_{u\bar{x}} \cdot (p_x + p_{\bar{x}}) \leq p_{u\bar{x}} \cdot (p'_x + p'_{\bar{x}})$$
$$p'_{\bar{u}x} \cdot (p_x + p_{\bar{x}}) \geq p_{\bar{u}x} \cdot (p'_x + p'_{\bar{x}}),$$

for which, in turn, it is sufficient to show that the following four statements hold:

$$p'_{u\bar{x}}p_x \leq p_{u\bar{x}}p'_x \quad p'_{u\bar{x}}p_{\bar{x}} = p_{u\bar{x}}p'_{\bar{x}}$$
$$p'_{\bar{u}x}p_x = p_{\bar{u}x}p'_x \quad p'_{\bar{u}x}p_{\bar{x}} \geq p_{\bar{u}x}p'_{\bar{x}}.$$

The latter are immediate by Lemma 6.4 if we set appropriately α and μ to $\emptyset$ or $\{u\}$. □

7 Conclusions and future work

This paper answers positively some questions raised in Boldi and Vigna (2014), Boldi et al. (2017, 2023) about closeness, harmonic centrality and betweenness. Table 1 puts the positive results about rank semi-monotonicity of this paper in context with the positive results of the directed case. It is interesting to note that rank semi-monotonicity is in fact the only property of this kind that is true for betweenness, both in the directed and in the undirected case. All the results are consequences of a stronger result of a similar type about basin dominance.

However, we also proved results about basin dominance, score monotonicity, and rank semi-monotonicity for distance-decay centralities whose decay function is convex;

Table 1 Summary of the results about semi-monotonicity obtained in this paper (in boldface) corresponding to negative results about monotonicity in Boldi et al. (2023). All results are about (strongly) connected graphs, except for the monotonicity property of harmonic centrality on directed graphs, which is true on all graphs

	Undirected		Directed (Boldi and Vigna 2014; Boldi et al. 2017)	
	Score	Rank	Score	Rank
Closeness	Monotone (Boldi et al. 2023)	**Semi-monotone**	Monotone	Monotone
Harmonic centrality	Monotone (Boldi et al. 2023)	**Strictly semi-mon.**	Monotone	Strictly monotone
Betweenness	**Not semi-monotone**	**Semi-monotone**	Not monotone	Not monotone

this result enabled us to prove rank semi-monotonicity for other centralities defined in the literature.

Our negative results are about the strictness of rank semi-monotonicity, in particular of closeness and betweenness, and lack of score semi-monotonicity for betweenness. For all the negative results, we have an infinite family of counterexamples.

The notion of basin dominance turned out to be the key idea in all proofs of semi-monotonicity. It would be interesting to investigate whether basin dominance applies to other geometric measures, or even other centrality measures based on shortest paths, as in that case one gets immediately rank semi-monotonicity.

Proving or disproving score and (strict) rank semi-monotonicity for other measures (in particular, for the spectral ones that were shown not to be rank monotone in Boldi et al. (2023), such as PageRank) remains an open problem.

Acknowledgements This work was supported in part by project SERICS (PE00000014) under the NRRP MUR program funded by the EU–NGEU. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the Italian MUR. Neither the European Union nor the Italian MUR can be held responsible for them. Davide D'Ascenzo has been financially supported by the Italian National PhD Program in Artificial Intelligence (DM 351 intervento M4C1–Inv. 4.1–Ricerca PNRR), funded by EU–NGEU.

Author contributions All authors contributed equally to the writing.

Funding Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement.

Declarations

Competing interests The authors declare no competing interests.

References

Anthonisse JM (1971) The rush in a directed graph. J Comput Phys 1971:1–10

Bavelas A (1948) A mathematical model for group structures. Hum Organ 7:16–30

Bavelas A (1950) Communication patterns in task-oriented groups. Journal of the Acoustical Society of America 22:725–730

Beauchamp MA (1965) An improved index of centrality. Behavioral Science 10(2):161–163

Bergamini E, Crescenzi P, D'Angelo G, Meyerhenke H, Severini L, Velaj Y (2018) Improving the betweenness centrality of a node by adding links. Journal of Experimental Algorithmics (JEA) 23:1–32

Berge C (1958) Théorie des graphes et ses applications. Dunod, Paris, France

Boldi P, D'Ascenzo D, Furia F, Vigna S (2024) Score and rank semi-monotonicity for closeness, betweenness and harmonic centrality. In: Cherifi H, Rocha LM, Cherifi C, Donduran M (eds) Complex networks & their applications XII, number 1143 in studies in computational intelligence. Springer, Cham, pp 102–113

Boldi P, Furia F, Vigna S (2023) Monotonicity in undirected networks. Netw Sci 11(3):1–23

Boldi P, Luongo A, Vigna S (2017) Rank monotonicity in centrality measures. Netw Sci 5(4):529–550

Boldi P, Vigna S (2014) Axioms for centrality. Intern Math 10:222–262

Boldi P, Vigna S (2019) Rank monotonicity in centrality measures-Corrigendum. Netw Sci 7(2):265–268

Brandes U, Laußmann C, Rothe J (2022) Voting for centrality. In: Proceedings of the 21st international conference on autonomous agents and multiagent systems, AAMAS '22, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, pp 1554–1556

Chien S, Dwork C, Kumar R, Simon D, Sivakumar D (2003) Link evolution: analysis and algorithms. Intern Math 1:277–304

Cohen E, Delling D, Pajor T, Werneck RF (2014) Computing classic closeness centrality, at scale. In: Proceedings of the 2nd ACM conference on online social networks, Association for Computing Machinery, New York, NY, USA, pp 37–50

Dangalchev C (2006) Residual closeness in networks. Phys A Stat Mech Appl 365(2):556–564

Entringer RC, Jackson DE, Snyder DA (1976) Distance in graphs. Czechoslov Math J 26(2):283–296
Freeman L (1977) A set of measures of centrality based on betweenness. Sociometry 40:35–41
Graham LR, Knuth DE, Oren P (1994) Concrete mathematics, 2nd edn. Addison-Wesley, Melbourne
Harris CD (1954) The market as a factor in the localization of industry in the United States. Ann Assoc Am Geograph 44(4):315–348
Jackson MO (2008) Social and economic networks, vol 3. Princeton University Press, Princeton
Katz L (1953) A new status index derived from sociometric analysis. Psychometrika 18:39–43
Kishi G (1981) On centrality functions of a graph. In: Saito N, Nishizeki T (eds) Graph theory and algorithms. Springer, Berlin Heidelberg, pp 45–52
Landau E (1985) Zur relativen wertbemessung der turnierresultate. Deutsch Wochenschach 11:366–369
Page L, Brin S, Motwani R, Winograd T (1999) The pagerank citation ranking: bringing order to the web. Stanford InfoLab, Stanford
Pan RK, Saramäki J (2011) Path lengths, correlations, and centrality in temporal networks. Phys Rev E 84(1):016105
Sabidussi G (1966) The centrality index of a graph. Psychometrika 31:581–603
Seeley JR (1949) The net of reciprocal influence a problem in treating sociometric data. Can J Exp Psychol 3:234
Skibski O (2023) Closeness centrality via the condorcet principle. Soc Netw 74:13–18