Absolute Expressiveness of Subgraph Motif Centrality Measures

Andreas Pieris  
University of Edinburgh & University of Cyprus  
Cyprus  
apieris@inf.ed.ac.uk

Jorge Salas  
Pontificia Universidad Católica de Chile & University of Edinburgh  
Chile  
jusalas@uc.cl

ABSTRACT
In graph-based applications, a common task is to pinpoint the most important or “central” vertex in a (directed or undirected) graph, or rank the vertices of a graph according to their importance. To this end, a plethora of so-called centrality measures have been proposed in the literature that assess which vertices in a graph are the most important ones. Riveros and Salas, in an ICDT 2020 paper, proposed a family of centrality measures based on the following intuitive principle: the importance of a vertex in a graph is relative to the number of “relevant” connected subgraphs, known as subgraph motifs, surrounding it. We refer to the measures derived from the above principle as subgraph motif measures. It has been convincingly argued that subgraph motif measures are well-suited for graph database applications. Although the ICDT paper studied several favourable properties enjoyed by subgraph motif measures, their absolute expressiveness remains largely unexplored. The goal of this work is to precisely characterize the absolute expressiveness of the family of subgraph motif measures.

1 INTRODUCTION
Graphs are well-suited for representing complex networks such as biological networks, cognitive and semantics networks, computer networks, and social networks, to name a few. In many applications that involve (directed or undirected) graphs, a crucial task is to pinpoint the most important or “central” vertex in a graph, or rank the vertices of a graph according to their importance. Indeed, these graph-theoretic tasks naturally appear in many different contexts, for example, finding people who are more likely to spread a disease in the event of an epidemic [4], highlighting cancer genes in proteomic data [9], assessing the importance of subgraph motifs, surrounding it. We refer to the measures derived from the above principle as subgraph motif measures. It has been convincingly argued that subgraph motif measures are well-suited for graph database applications. Although the ICDT paper studied several favourable properties enjoyed by subgraph motif measures, their absolute expressiveness remains largely unexplored. The goal of this work is to precisely characterize the absolute expressiveness of the family of subgraph motif measures.

Centrality Measures and Graph Databases. As said above, measuring the importance of a vertex in a graph is an important task in many graph-based applications, and graph databases is no exception. For example, Neo4j, one of the main graph database management systems, has adopted and implemented several centrality measures and algorithms in its Graph Data Science (GDS) library such as Eigenvector [2], PageRank [15], Closeness [19], etc.1 A more principled approach towards centrality measures that are well-suited for graph database applications has been recently introduced by Riveros and Salas [17]. In particular, they proposed a family of measures based on the following intuitive principle: the importance of a vertex in a graph is relative to the number of “relevant” connected subgraphs surrounding it. More precisely, the importance of a vertex is defined as the logarithm of the number of relevant subgraphs surrounding it. As explicitly discussed in [17], the choice of applying the logarithmic function is purely for technical simplicity, and one could adopt any function, which we call filtering function, that leads to a richer family of centrality measures. Let us clarify that [17] considered only undirected graphs, but the above general principle can be naturally transferred to directed graphs. The notion of relevant subgraph is reminiscent of the concept of subgraph motif, that is, recurrent and statistically significant connected subgraphs of a larger graph. Subgraph motifs have recently attracted considerable attention as a useful tool to reveal structural design principles of complex networks [12]. Due to this conceptual similarity, we refer to the graph centrality measures (for directed or undirected graphs) derived from the principle by Riveros and Salas as subgraph motif measures. We further refer to monotonic subgraph motif measures if the filtering function is monotonic. Note that the idea of exploiting subgraph motifs for defining centrality measure has been already considered in the literature; see, e.g., [11].

It has been convincingly argued in [17] that (monotonic) subgraph motif measures are conceptually relevant for graph database applications. Consider, for example, a property graph G, which is essentially a finite directed graph, and a language L of basic graph patterns [1]. The evaluation of a query Q from L over G, denoted Q(G), is the set of vertices of G that comply with the graph pattern expressed by Q. In some scenarios, it is reasonable to assume that the more queries Q from L such that v ∈ Q(G), the more important v is in G (relative to L). This way of defining the importance of a vertex follows the general principle discussed above, where the relevant subgraphs (or subgraph motifs) are the basic graph patterns from the language L, and the filtering function is any monotonic function (e.g., the logarithm). Analogously, in a different scenario where the less queries Q from L such that v ∈ Q(G), the more important v is in G (relative to L), a subgraph motif measure where the underlying subgraph motifs are the graph patterns of L, and the filtering function is a decreasing one, could be adopted.

Our Main Objective. With the family of subgraph motif measures in place, Riveros and Salas went on to isolate favourable properties enjoyed by subgraph motif measures depending on the underlying subgraph motifs, that is, the set of relevant subgraphs [17]. Actually, their analysis focused on subgraph motif measures for undirected graphs with the filtering function being always the logarithmic function. In this work, instead, our main objective is to

---
1https://neo4j.com/docs/graph-data-science/current/algorithms/centrality/
delineate the limits of the family of (monotonic) subgraph motif measures as a whole (i.e., without fixing a priori the filtering function) for both directed and undirected graphs. More precisely, we would like to understand when an arbitrary centrality measure is a (monotonic) subgraph motif one, or when it induces the same ranking as a (monotonic) subgraph motif one.

Our Contributions. In Section 4, we provide a precise characterization of when an arbitrary centrality measure is subgraph motif. More precisely, we isolate a “bounded value” property $P$ over centrality measures, which essentially states that the total number of distinct values that can be assigned to vertices surrounded by a certain number of connected subgraphs is bounded, and then show that a centrality measure can be expressed as a subgraph motif one iff it enjoys $P$. We then proceed in Section 5 to characterize when an arbitrary centrality measure induces the same ranking as a subgraph motif measure. In this case, we isolate a “graph coloring” property $P$ over centrality measures, and then show that a centrality measure can be expressed as a subgraph motif one relative to the induced ranking iff it enjoys $P$. In Section 6, we focus on the family of monotonic subgraph motif measures, and provide analogous characterizations via refined properties in the spirit of the “bounded value” property discussed above. An interesting finding of our analysis is that in the case of connected graphs, every centrality measure can be expressed as a monotonic subgraph motif measure relative to the induced ranking. Having the above results, we proceed to determine if established measures (such as PageRank, Eigenvector, and many others) are (monotonic) subgraph motif (relative to the induced ranking). Such a classification, apart from being interesting in its own right, provides insights on the structural similarities and differences among the considered centrality measures.

Clarification Remark. In the rest of the paper, due to space constraints and for the sake of clarity, we focus on undirected graphs, but all the notions and results can be transferred to the case of directed graphs under the standard notion of weak connectedness. The case of directed graphs is deferred to the appendix.

2 PRELIMINARIES

We recall the basics on undirected graphs and graph centrality measures. In the rest of the paper, we assume the countable infinite set $V$ of vertices. For $n > 0$, let $[n] = \{1, \ldots, n\}$.

Undirected Graphs. An undirected graph (or simply graph) $G$ is a pair $(V, E)$, where $V$ is a finite non-empty subset of $V$ (the set of vertices of $G$), and $E \subseteq \{[u, v] \mid u, v \in V\}$ (the set of edges of $G$). For notational convenience, given a graph $G$, we write $V(G)$ and $E(G)$ for the set of its vertices and edges, respectively. We denote by $G$ the set of all graphs, and by $\text{VG}$ the set of vertex-graph pairs $\{(u, G) \mid u \in V(G)\}$. The neighbourhood of a vertex $v \in V(G)$ in $G$, denoted $N_G(v)$, is the set $\{u \in V(G) \mid [u, v] \in E(G)\}$. For $u \in N_G(v)$, we say that $v$ and $u$ are adjacent in $G$. For a vertex $v \in V$, we write $G_v$ for the graph $(v, \emptyset)$.

A subgraph of a graph $G$ is a graph $G'$ such that $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$; we write $G' \subseteq G$ to indicate that $G'$ is a subgraph of $G$. Note that the binary relation $\subseteq$ over graphs forms a partial order. We denote by Sub$(G)$ all the subgraphs of $G$, that is, the set of graphs $\{G' \mid G' \subseteq G\}$. Given a set of vertices $S \subseteq V(G)$, the subgraph of $G$ induced by $S$, denoted $G[S]$, is the subgraph $G'$ of $G$ such that $V(G') = S$, and $E(G') = \{[u, v] \in E(G) \mid u, v \in S\}$.

A path in $G$ is a sequence of vertices $\pi = v_0, v_1, \ldots, v_n$, for $n \geq 0$, such that $\{v_i, v_{i+1}\} \in E(G)$ for every $0 \leq i < n$. We further say that $\pi$ is a path from $v_0$ to $v_n$. The length of $\pi$, denoted $|\pi|$, is the number of edges in $\pi$. By convention, there exists a path of length 0 from a vertex to itself. The distance between two vertices $u, v \in V(G)$ in $G$, denoted $d_G(u, v)$, is defined as the length of a shortest path from $u$ to $v$ in $G$; if there is no path, then $d_G(u, v) = \infty$. We denote by $S_G(u, v)$ the set of all the shortest paths from $u$ to $v$ in $G$, that is, the set $\{\pi \mid \pi$ is a path from $u$ to $v$ in $G$ with $|\pi| = d_G(u, v)\}$.

A graph $G$ is connected if, for every two distinct vertices $u, v \in V(G)$, there exists a path from $u$ to $v$. We denote by $A(v, G)$ the set of all connected subgraphs of $G$ that contain $v$, that is, the set $\{G' \subseteq G \mid v \in V(G') \text{ and } G' \text{ is connected}\}$. By abuse of notation, we may treat $A(\cdot, \cdot)$ as a function of the form $VG \rightarrow \mathcal{P}(G)$; as usual, we write $\mathcal{P}(S)$ for the powerset of a set $S$. A connected component (or simply component) of $G$ is an induced subgraph $G[S]$ of $G$, where $S \subseteq V(G)$, such that $G[S]$ is connected, and, for every $v \in V(G) \setminus S$, there is no path in $G$ from $v$ to a vertex of $S$. It is clear that when $G$ is connected, the only component of $G$ is itself. We denote by $\text{Comp}(G)$ all the components of $G$, that is, the set of graphs $\{G' \mid G'$ is a component of $G\}$. Let $K_v(G)$ be the set of vertices of the component of $G$ containing the vertex $v$.

Centrality Measures. A centrality measure assigns a score to a vertex $v$ in a graph $G$, which reflects the importance of $v$ in $G$. In other words, we adopt the standard assumption that the higher the score of a vertex $v$ in $G$, the more important or “central” $v$ is in $G$. Formally, a centrality measure (or simply measure) is a function $C : VG \rightarrow \mathbb{R}$. We proceed to recall three measures that will be used throughout the paper; more measures are discussed in Section 7.

Stress. This is a well-known centrality measure introduced in the 1950s [20]. It measures the centrality of a vertex by essentially counting the number of shortest paths that go via that vertex. Formally, for a graph $G$ and a vertex $v \in V(G)$, the stress centrality of $v$ in $G$ is

$$
\text{Stress}(v, G) = \sum_{u, w \in V(G) \setminus \{v\}} |S^v_{G}(u, w)|,
$$

where $S^v_{G}(u, w) = \{\pi \in S_G(u, w) \mid \pi \text{ contains } v\}$.

All-Subgraphs. This measure was recently introduced in the context of graph databases [17]. It essentially states that a vertex is more central if it participates in more connected subgraphs. Formally, given a graph $G$ and a vertex $v \in V(G)$, the all-subgraphs centrality of $v$ in $G$ is defined as

$$
\text{All-Subgraphs}(v, G) = \log_{|G|}|A(v, G)|.
$$

Closeness. This is a well-known measure introduced back in 1960s [19]. It is usually called a geometrical measure since it relies on the distance inside a graph. It essentially states that the closer a vertex is to everyone in the graph the more central it is. Formally, given a graph $G$ and a vertex $v \in V(G)$, the closeness centrality of $v$ in $G$ is defined as the ratio

$$
\text{Closeness}(v, G) = \frac{1}{\sum_{u \in K_v(G)} d_G(v, u)}.
$$
Let us clarify that we define the sum of distances inside a component of \( G \) since the distance between two vertices in different components of \( G \) is by definition infinite.

### 3 SUBGRAPH MOTIF MEASURES

As already discussed in the Introduction, a natural way of measuring the importance of a vertex in a graph is to count the relevant connected subgraphs (i.e., the subgraph motifs) surrounding it, and then apply a certain filtering function from the non-negative integers to the reals on top of the count. Of course, the relevant subgraphs and the adopted filtering function are determined by the intention of the centrality measure. Interestingly, both the stress and the all-subgraphs centrality measures, discussed in the previous section, are actually based on this principle. Let us elaborate further on this. Consider a graph \( G \) and a vertex \( v \in V(G) \):

- For the stress centrality, the important subgraphs for \( v \) in \( G \) are the shortest paths that go via \( v \) in \( G \), and the filtering function is \( f_{\text{Stress}}(x) = 2x \) since each shortest path is counted twice. In other words, with \( G_v \) being the graph that corresponds to a path \( \pi \), it is clear that
  \[
  \text{Stress}(v, G) = f_{\text{Stress}} \left( \bigcup_{u,w \in V(G) \setminus \{v\}} \{ G_\pi \mid \pi \in S^\pi_G(u,w) \} \right).
  \]
- For the all-subgraphs centrality, the important subgraphs for \( v \) in \( G \) are the connected subgraphs of \( G \) that contain \( v \), that is, the set \( A(v, G) \), and the filtering function is \( \log_2 \).

Indeed, by definition, we have that
\[
\text{All-Subgraphs}(v, G) = \log_2 |A(v, G)|.
\]

We proceed to formalize the above simple principle, originally introduced in [17], which gives rise to a family of centrality measures, and then highlight our main research questions.

#### Centrality Measures based on Subgraph Motifs

We first need a way to specify what are the important subgraphs for a vertex \( v \) in a graph \( G \). This is done via the notion of subgraph family, which is defined as a function from vertex-graph pairs to sets of graphs \( F : V G \to 2^P(G) \) such that, for every \( (u, G) \in V G \), \( F(u, G) \subseteq A(u, G) \), that is, \( F \) assigns to each \( (u, G) \in V G \) a set of connected subgraphs of \( G \) surrounding \( u \). We also need the notion of filtering function, which, as said above, is simply a function of the form \( f : \mathbb{N} \to \mathbb{R} \).

We are now ready to define our centrality measures:

**Definition 3.1 (Subgraph Motif Measure).** Consider a subgraph family \( F \) and a filtering function \( f \). The \( (F, f) \)-measure is the function \( C(F, f) : V G \to \mathbb{R} \) such that, for \( (u, G) \in V G \),
\[
C(F, f)(u, G) = f(|F(u, G)|).
\]

A centrality measure \( C \) is a subgraph motif centrality measure if there are a subgraph family \( F \) and a filtering function \( f \) such that \( C \) is the \( (F, f) \)-measure, namely \( C = C(F, f) \).

Returning back to our discussion on stress and all-subgraph centralities, assuming that \( S \) is the subgraph family such that
\[
S(v, G) = \bigcup_{u,w \in V(G) \setminus \{v\}} \{ G_\pi \mid \pi \in S^\pi_G(u,w) \}.
\]

it is straightforward to see that
\[
\text{Stress} = C(S, f_{\text{Stress}}) \quad \text{and} \quad \text{All-Subgraphs} = C(A, \log_2).
\]

#### Main Research Questions

Having the family of subgraph motif centrality measures in place, the natural question that comes up concerns its absolute expressiveness. In other words, we are interested in the following research question:

**Question I: When is a centrality measure a subgraph motif centrality measure?**

One may wonder whether the above question is conceptually trivial in the sense that every centrality measure can be expressed as a subgraph motif centrality measure by choosing the subgraph family and the filtering function in the proper way as done for Stress and All-Subgraphs. It turns out that there are measures that are not subgraph motif, which renders the above question non-trivial.

**Proposition 3.2.** There exists a centrality measure that is not a subgraph motif centrality measure.

**Proof.** Let \( \tilde{G} \) be the graph \( \{(v_1, v_2, v_3), \{(v_1, v_2)\}) \). Consider the centrality measure \( C \) defined as follows:
\[
C(v, G) = \begin{cases}
1 & G \neq \tilde{G} \\
1 + 1 & G = \tilde{G} \text{ and } v = v_i \text{ for } i \in [3].
\end{cases}
\]

It suffices to show that \( C \) is not subgraph motif even if we focus on the set of graphs \( G^* = \{G_v \mid v \in V \} \cup \{\tilde{G}\} \subseteq G \). By contradiction, assume that \( C \) is a subgraph motif measure over \( G^* \). Thus, there is a subgraph family \( F \) and a filtering function \( f \) such that, for every \( G \in G^* \) and \( v \in V(G) \), \( C(v, G) = C(F, f)(v, G) \). We observe that:

1. For every \( (v, G) \in V \times G^* \) with \( v \in V(G) \), it holds that \( C(F, f)(v, G) \in \{1, \ldots, 4\} \), i.e., we have four distinct values. This follows by the definition of \( C = C(F, f) \).

2. For every \( (v, G) \in V \times G^* \), it holds that \( |F(v, G)| \in \{0, 1, 2\} \), i.e., we have three possible sizes for the sets of subgraph motifs. Indeed, \( |F(v, G)| \in \{0, 1, 2\} \) for \( v \in V \), \( |F(v_1, \tilde{G})| = 3 \), and \( |F(v_3, \tilde{G})| \in \{0, 1\} \).

Now, by the pigeonhole principle, we can conclude that there are at least two distinct pairs \( (v, G), (u, G') \in V \times G^* \) such that \( C(F, f)(v, G) \neq C(F, f)(u, G') \). But this contradicts the fact that \( f \) is a function, and the claim follows.

As we shall see, not only artificial measures as the one employed in the proof of Proposition 3.2, but also well-known centrality measures from the literature (such as Closeness) are not subgraph motif. We are going to prove such inexpressibility results by using the technical tools developed towards answering Question I.

In several applications that involve graphs, we are interested in the relative than the absolute importance of a vertex in a graph. More precisely, we are interested in the ranking of the vertices of a graph induced by a measure \( C \) and not in the absolute value assigned to a vertex by \( C \). This brings us to the next notion:

**Definition 3.3 (Induced Ranking).** Let \( C \) be a centrality measure. The ranking induced by \( C \), denoted Rank\((C)\), is the binary relation
\[
\{(u, G), (v, G) \mid u, v \in V(G) \text{ and } C(u, G) \leq C(v, G)\}
\]
over \( V G \). We say that \( C \) is a subgraph motif centrality measure relative to the induced ranking if there are a subgraph family \( F \) and a filtering function \( f \) with \( \text{Rank}(C) = \text{Rank}(C(F, f)) \).
Interestingly, although the measure employed in the proof of Proposition 3.2 is not subgraph motif, we can show that it is a subgraph motif measure relative to the induced ranking. In particular, with \( \hat{G} \) being the graph \( \{ \{ v_1, v_2, v_3 \}, \{ \{ v_1, v_2 \} \} \} \) used in the proof of Proposition 3.2, by defining the subgraph family \( F \) as

\[
F(v, G) = \begin{cases} 
\emptyset & G \neq \hat{G} \\
\emptyset & G = \hat{G} \text{ and } v = v_1 \\
\{ G_{\hat{G}}(C_2(\hat{G})) \} & G = \hat{G} \text{ and } v = v_2 \\
\{ G_v \} & G = \hat{G} \text{ and } v = v_3,
\end{cases}
\]

and the filtering function \( f \) as \( f(0) = 1, f(1) = 3 \) and \( f(2) = 2 \), it is not difficult to see that \( \text{Rank}(G) = \text{Rank}(C(F, f)) \). This observation brings us to our next research question:

**Question II:** When is a centrality measure a subgraph motif centrality measure relative to the induced ranking?

We conclude this section by showing that the above question is conceptually non-trivial, i.e., there are measures that are not subgraph motif measures relative to the induced ranking. In particular, with \( \hat{G} = (\{ v_1, v_2, v_3, v_4 \}, \{ \{ v_1, v_2 \} \}) \), for the measure \( C \) defined as

\[
C(v, G) = \begin{cases} 
1 & G \neq \hat{G} \\
| i + 1 | & G = \hat{G} \text{ and } v = v_i \text{ for } i \in [4]
\end{cases}
\]

we can show that, for every subgraph family \( F \) and filtering function \( f \), \( \text{Rank}(C) \neq \text{Rank}(C(F, f)) \). Therefore:

**Proposition 3.4.** There is a measure that is not a subgraph motif centrality measure relative to the induced ranking.

Let us stress that, not only artificial measures as the one employed in the proof of Proposition 3.4, but also established measures (such as Closeness) are not subgraph motif centrality measures relative to the induced ranking. Such inexpressibility results are shown by exploiting the tools developed towards answering Question II.

## 4 CHARACTERIZING SUBGRAPH MOTIF MEASURES

We proceed to provide an answer to Question I. More precisely, our goal is to isolate a structural property \( P \) over centrality measures that precisely characterizes subgraph motif measures, that is, for an arbitrary measure \( C \), \( C \) is a subgraph motif measure iff \( C \) enjoys the property \( P \). Interestingly, the desired property can be somehow extracted from the proof of Proposition 3.2. The crucial intuition provided by that proof is that the absolute expressiveness of subgraph motif measures is tightly related to the amount of connected subgraphs that are available for assigning different centrality values to vertices. In other words, a measure that assigns "too many" values among vertices that are surrounded by "too few" connected subgraphs cannot be expressed as a subgraph motif measure. We proceed to formalize this intuition.

We first collect all the different values assigned by a centrality measure \( C \) to the vertices of a graph \( \hat{G} \) that are surrounded by a bounded number of connected subgraphs of \( \hat{G} \). In particular, for \( n > 0 \), we define the set of real values

\[
\text{Val}_G^n(C) = \{ C(v, G) \mid v \in V(G) \text{ and } |A(v, G)| \leq n \}.
\]

We can then easily collect all the values assigned by \( C \) to the vertices of \( V \) that are surrounded by a bounded number of connected subgraphs in some graph. In particular, for \( n > 0 \),

\[
\text{Val}_G^n(C) = \bigcup_{G \in G} \text{Val}_G^n(C).
\]

We now define the following property over measures:

**Definition 4.1 (Bounded Value Property).** A measure \( C \) enjoys the bounded value property if, for every \( n > 0 \), \( |\text{Val}_G^n(C)| \leq n + 1 \). ■

The bounded value property captures the key intuition discussed above. It actually bounds the number of different values that can be assigned among vertices that are surrounded by a limited number of connected subgraphs; hence the name "bounded value property". Observe that the measure \( C \) devised in the proof of Proposition 3.2 does not enjoy the bounded value property; indeed, \( |\text{Val}_G^2(C)| = 4 > 3 \). Surprisingly, the bounded value property is all we need towards a precise characterization of subgraph motif measures.

**Theorem 4.2.** Consider a centrality measure \( C \). The following statements are equivalent:

1. \( C \) is a subgraph motif centrality measure.
2. \( C \) enjoys the bounded value property.

The above characterization, apart from giving a definitive answer to Question I, it provides a useful tool for establishing inexpressibility results. To show that a centrality measure \( C \) is not a subgraph motif measure it suffices to show that there exists an integer \( n > 0 \) such that \( |\text{Val}_G^n(C)| > n + 1 \). For example, in the case of Closeness, we can show that \( |\text{Val}_G^5(\text{Closeness})| \) > 6, and therefore:

**Proposition 4.3.** Closeness is not a subgraph motif measure.

Without Theorem 4.2 in place, it is completely unclear how one can prove that Closeness (or any other established measure) is not a subgraph motif measure. More inexpressibility results concerning well-established centrality measures are discussed in Section 7.

## 5 CHARACTERIZING SUBGRAPH MOTIF MEASURES RELATIVE TO THE INDUCED RANKING

We now focus on Question II. Our goal is to isolate a structural property \( P \) over centrality measures that precisely characterizes subgraph motif measures relative to the induced ranking, i.e., for an arbitrary measure \( C \), \( C \) is subgraph motif relative to the induced ranking iff \( C \) enjoys the property \( P \). Despite our efforts, we have not managed to define \( P \) in the spirit of the bounded value property (see Definition 4.1). It turns out, however, that \( P \) can be defined by exploiting a certain notion of graph coloring relative to a centrality measure. At the end of this section, we discuss how the coloring-based property helps to obtain a bounded-value-like property that is a necessary condition for a measure being subgraph motif relative to the induced ranking. As we shall see in Section 7, the latter provides a convenient tool for showing that several existing measures are not subgraph motif relative to the induced ranking.

**Graph Colorings.** The high-level idea is to consider the sizes of the available subgraph families that can be assigned to a vertex \( v \)
in a graph $G$, i.e., the set of integers $\{0, \ldots, |A(v, G)|\}$, as available colors. We can then refer to a precoloring of $VG$ (i.e., of all the possible graphs) as a function $pc : VG \to \mathbb{N}$ that assigns to each vertex $v$ in a graph $G$ only available colors from $\{0, \ldots, |A(v, G)|\}$. Then, the goal is to isolate certain properties of such a precoloring of $VG$ that leads to the desired characterization, i.e., a measure $C$ is a subgraph motif relative to the induced ranking iff there exists a precoloring of $VG$ that enjoys the properties in question. Such a characterization tells us that for a centrality measure being subgraph motif relative to the induced ranking tantamount to the fact that there are enough colors (i.e., sizes of subgraph families, but without considering their actual topological structure) that allow us to color $VG$ in a valid way, namely in a way that the crucial properties are satisfied. We proceed to formalize the above discussion about colorings.

Given a set $S \subseteq VG$, a precoloring of $S$ is a function $pc : S \to \mathbb{N}$ such that, for every $(v, G) \in S$, $pc(v, G) \in \{0, \ldots, |A(v, G)|\}$. It is very natural to expect from such a precoloring of $S$ to respect the values assigned by a measure $C$ to the vertices of a certain graph $G$, i.e., vertices of $G$ with different centrality values get different colors. This leads to the following property over precolorings:

**Definition 5.1 (Non-Uniform C-Injectivity).** Consider a set $S \subseteq VG$, and a precoloring $pc : S \to \mathbb{N}$ of $S$. Given a centrality measure $C$, we say that $pc$ is non-uniformly $C$-injective if, for every $(u, G), (v, G) \in S$, $C(u, G) = C(v, G)$, $pc(u, G) \neq pc(v, G)$. ■

The term non-uniform in the above definition refers to the fact that $C$-injectivity is only enforced inside a certain graph, and not across all the graphs mentioned in $S$, i.e., it might be the case that a non-uniformly $C$-injective precoloring of $S$ assigns to $(u, G), (u', G')$, where $G \neq G'$ and $C(u, G) = C(u', G')$, the same color.

It also natural to expect from a precoloring of $S$ to be somehow consistent with the induced ranking, not only inside a certain graph, but also among different graphs mentioned in $S$. In other words, if $(u, G)$ comes before $(v, G)$, and $(u', G')$ comes before $(v', G')$, then one of the following should hold: $(u, G)$ and $(v', G')$ get different colors, or $(u', G')$ and $(v', G')$ get different colors.

**Definition 5.2 (C-Consistency).** Consider a set $S \subseteq VG$, and a precoloring $pc : S \to \mathbb{N}$ of $S$. Given a centrality measure $C$, we say that $pc$ is $C$-consistent if, for every $(u, G), (v, G), (u', G'), (v', G') \in S$, the following holds: if $C(u, G) < C(v, G)$ and $C(u', G') < C(v', G')$, then $pc(u, G) \neq pc(v', G')$ or $pc(u', G') \neq pc(v', G')$. ■

Putting together the above two properties over precolorings, we get the notion of $C$-colorability of a set $S \subseteq VG$.

**Definition 5.3 (C-Colorability).** We say that a set $S \subseteq VG$ is $C$-colorable, for some measure $C$, if there exists a precoloring of $S$ that is non-uniformly $C$-injective and $C$-consistent. ■

**The Characterization.** Interestingly, $C$-colorability is all we need towards the desired characterization, namely a measure $C$ is a subgraph motif relative to the induced ranking iff $VG$ (i.e., all possible graphs) is $C$-colorable. We further show that the $C$-colorability of $VG$ is equivalent to the $C$-colorability of every finite set $S \subseteq VG$. The latter, apart from being interesting in its own right, it also provides a technical tool that is somehow more convenient than the $C$-colorability of the whole set $VG$ for classifying centrality measures as subgraph motif relative to the induced ranking.

**Theorem 5.4.** Consider a centrality measure $C$. The following statements are equivalent:

1. $C$ is a subgraph motif measure relative to the induced ranking.
2. Every finite set $S \subseteq VG$ is $C$-colorable.
3. $VG$ is $C$-colorable.

To show the above characterization, it suffices to establish the sequence of implications $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1)$. We proceed to discuss the key ingredients underlying the proofs of the above implications; the full proofs can be found in the appendix.

Implication $(1) \Rightarrow (2)$

This implication is a rather easy one to prove. Since, by hypothesis, $C$ is a subgraph motif measure relative to the induced ranking, there is a subgraph family $F$ and a filtering function $f$ such that $\text{Rank}(C) = \text{Rank}(C(F, f))$. Given a finite set $S \subseteq VG$, we define the function $pc_S : S \to \mathbb{N}$ as follows: for every $(v, G) \in S$, $pc_S(v, G) = |F(v, G)|$. It is clear that $pc_S$ is a precoloring of $S$ since, by definition, $F(v, G) \subseteq A(v, G)$, and thus, $pc_S(v, G) \in \{0, \ldots, |A(v, G)|\}$. It is also not difficult to show that $pc_S$ is non-uniformly $C$-injective and $C$-consistent, which in turn implies that $S$ is $C$-colorable.

Implication $(2) \Rightarrow (3)$

The proof of this implication heavily relies on an old result that goes back in 1949 by Rado [16] known as Rado’s Selection Principle. We write $\mathcal{P}_{\text{fin}}(A)$ for the finite powerset of a set $A$, i.e., the set that collects all the finite subsets of $A$. Furthermore, given a function $f : A \to B$, we write $f_C$ for the restriction of $f$ to $C \subseteq A$.

**Theorem 5.5 (Rado’s Selection Principle).** Let $A$ and $B$ be arbitrary sets. Assume that, for each $C \in \mathcal{P}_{\text{fin}}(A)$, $f_C$ is a function $C \to B$ (a so-called “local function”). Assume further that, for every $x \in A$, the set $\{f_C(x) \mid C \in \mathcal{P}_{\text{fin}}(A) \}$ and $x \in C$ is finite. Then, there is a function $f : A \to B$ (a so-called “global function”) such that, for every $C \in \mathcal{P}_{\text{fin}}(A)$, there is $D \in \mathcal{P}_{\text{fin}}(A)$ with $C \subseteq D$ and $f_D = f_C|_D$.

Several proofs and applications of Rado’s Theorem can be found in [13]. We proceed to discuss how it is used to prove $(2) \Rightarrow (3)$. By hypothesis, for each $S \in \mathcal{P}_{\text{fin}}(VG)$, there exists a precoloring of $S$, i.e., a function $pc_S : S \to \mathbb{N}$, that is non-uniformly $C$-injective and $C$-consistent. Since, for every $(v, G) \in VG$, $A(v, G)$ is finite, we can conclude that the following holds: for every $(v, G) \in VG$, the set $\{pc_S(v, G) \mid S \in \mathcal{P}_{\text{fin}}(VG) \}$ and $(v, G) \in S$ is finite. This allows us to apply Theorem 5.5 with $A = VG$ and $B = \mathbb{N}$. Therefore, there exists a function $f : VG \to \mathbb{N}$ such that, for every $S \in \mathcal{P}_{\text{fin}}(VG)$, there exists $S' \in \mathcal{P}_{\text{fin}}(VG)$ with $S \subseteq S'$ and $f_{S'} = pc_{S'}|_{S}$. Interestingly, by exploiting the latter property of the function $f$ guaranteed by Theorem 5.5, and the fact that, for each $S \in \mathcal{P}_{\text{fin}}(VG)$, $pc_S$ is a precoloring of $S$ that is non-uniformly $C$-injective and $C$-consistent, it is not difficult to show that $f$ is a precoloring of $VG$ that is non-uniformly $C$-injective and $C$-consistent, and item (3) follows.

Implication $(3) \Rightarrow (1)$

We finally discuss the proof of the last implication. The goal is to devise a subgraph family $F$ and a filtering function $f$ such that...
Rank(C) = \text{Rank}(C(F, f)), which in turn proves item (1). By hypothesis, there exists a precoloring pc of VG that is non-uniformly C-injective and C-consistent. We define F in such way that, for every \((u, G) \in \text{VG}, |F(u, G)| = \text{pc}(u, G)\); note that such a subgraph family exists since \(\text{pc}(u, G) \in \{0, \ldots, |A(u, G)|\}\). Now, defining the filtering function \(f\) is a non-trivial task. Let \(R_{pc}\) be the relation

\[
\{(i, j) \in \mathbb{N} \times \mathbb{N} \mid \text{there are } (u, G), (o, G) \in \text{VG such that } C(u, G) < C(o, G), \text{pc}(u, G) = i, \text{ and } \text{pc}(o, G) = j\}.
\]

The fact that \(\text{pc}\) is non-uniformly C-injective allows us to conclude that \(R_{pc}\) is irreflexive. Moreover, the C-consistency of \(\text{pc}\) implies that \(R_{pc}\) is asymmetric. Observe now that if we extend \(R_{pc}\) into a total order \(R^*_pc\) over \(\mathbb{N}\), and then show that \(R^*_pc\) can be embedded into a carefully chosen countable subset \(N \subseteq \mathbb{R}\), then we obtain the desired filtering function \(f\), which assigns real numbers to the sizes of the subgraph families assigned to the pairs of VG by \(F\) as dictated by the embedding of \(R^*_pc\) into \(N \subseteq \mathbb{R}\). Let us now briefly discuss how this is done. The binary relation \(R_{pc}\) is first extended into the strict partial order \(R^*_pc\) by simply taking its transitive closure. Now, the fact that \(R^*_pc\) can be extended into a total order \(R^*_pc\) over \(\mathbb{N}\) follows by the order-extension principle (a.k.a. Szpilrajn Extension Theorem), shown by Szpilrajn in 1930 [21], which essentially states that every partial order can be extended into a total order. Finally, the fact that \(R^*_pc\) can be embedded into \(N \subseteq \mathbb{R}\) is shown via the back-and-forth method, a technique for showing isomorphism between countably infinite structures satisfying certain conditions.

A Bounded-Value-Like Property. As discussed at the beginning of the section, despite our efforts, we have not managed to isolate a property in the spirit of the bounded value property that can characterize subgraph motif measures relative to the induced ranking. We managed, however, to isolate a bounded-value-like property that is a necessary condition for a measure being subgraph motif relative to the induced ranking. It should be clear that the bounded value property introduced in the previous section (see Definition 4.1) is not enough towards a necessary condition. This is because, as discussed in Section 3, there is a measure (see the one devised in the proof of Proposition 3.2) that is not subgraph motif, which means that it does not enjoy the bounded value property, but it is subgraph motif relative to the induced ranking. On the other hand, to our surprise, a non-uniform version of the bounded value property leads to the desired necessary condition. Let us make this more precise. The ranking induced by a measure \(C\) compares only the values of vertices of the same graph; a pair \(((u, G), (v, G'))\), where \(G \neq G'\), will never appear in \(\text{Rank}(C)\). This led us to conjecture that for characterizing subgraph motif measures relative to the induced ranking, it suffices to bound the number of different values that can be assigned among vertices inside the same graph that are surrounded by a limited number of connected subgraphs. This leads to the non-uniform version of the bounded value property:

Definition 5.6 (Non-Uniform Bounded Value Property). A measure \(C\) enjoys the non-uniform bounded value property if, for every integer \(n > 0\) and graph \(G \in \text{VG}, |\text{Val}_G^C(C)| \leq n + 1\).

We can then show the following implication:

**Proposition 5.7.** Consider a centrality measure \(C\). If there exists a precoloring of VG that is non-uniformly C-injective, then \(C\) enjoys the non-uniform bounded value property.

By combining Theorem 5.4 and Proposition 5.7, we get the following corollary, which establishes that indeed the non-uniform bounded value property leads to the desired necessary condition:

**Corollary 5.8.** If a centrality measure is a subgraph motif measure relative to the induced ranking, then it enjoys the non-uniform bounded value property.

The question whether the non-uniform bounded value property is also a sufficient condition is negatively settled by the next result:

**Proposition 5.9.** There exists a measure that is not a subgraph motif measure relative to the induced ranking, but it enjoys the non-uniform bounded value property.

Let us stress that Corollary 5.8 equips us with a convenient tool for showing that a centrality measure \(C\) is not a subgraph motif measure relative to the induced ranking: it suffices to show that there exist an integer \(n > 0\) and a graph \(G\) such that \(|\text{Val}_G^C(C)| > n + 1\). In the case of closeness, we can show that there exists a graph \(G\) such that \(|\text{Val}_G^C(Closeness)| > 6\), which in turn implies that:

**Proposition 5.10.** Closeness is not a subgraph motif measure relative to the induced ranking.

More inexpressibility results of the above form concerning established centrality measures are discussed in Section 7.

**Connected Graphs.** After a quick inspection of the proof of Proposition 4.3, one can conclude that Closeness is not a subgraph motif measure even if concentrated on connected graphs. On the other hand, the proof of Proposition 5.10 heavily relies on the fact that the employed graphs are not connected. This observation led us ask ourselves whether Closeness is a subgraph motif measure relative to the induced ranking if we consider only connected graphs. It turned out that, if we focus on connected graphs, not only Closeness, but every measure is a subgraph motif measure relative to the induced ranking. We proceed to formalize the above discussion.

Let \(\text{VCG} = \{(u, G) \in \text{VG} \mid G \text{ is connected}\}\). For an arbitrary centrality measure \(C\), its version that operates only on connected graphs is defined as the function \(\text{ConC} : \text{VCG} \rightarrow \mathbb{R}\) such that, for every \((u, G) \in \text{VCG}, C(u, G) = \text{ConC}(u, G)\), i.e., it is the restriction of \(C\) over \(\text{VCG}\). We then say that \(\text{ConC}\) is a subgraph motif measure (resp., subgraph motif measure relative to the induced ranking) if there exist a subgraph family \(F\) and a filtering function \(f\) such that \(\text{ConC} = \text{ConC}(F, f)\) (resp., \(\text{Rank}(C) \cap \text{VCG}^2 = \text{Rank}(C(F, f)) \cap \text{VCG}^2\)). We can then establish the following result:

**Theorem 5.11.** Consider a centrality measure \(C\). It holds that \(\text{ConC}\) is a subgraph motif measure relative to the induced ranking.

As discussed above, ConCloseness is not a subgraph motif measure (implicit in the proof of Proposition 4.3), whereas ConCloseness is a subgraph motif measure relative to the induced ranking (follows from Theorem 5.11). This reveals a striking difference between the two notions of expressiveness, that is, being subgraph motif or being subgraph motif relative to the induced ranking, when focusing on connected graphs.
We conclude this section by stressing that Theorem 5.11 is interesting in its own right as it provides a unifying framework for all centrality measures in a practically relevant setting: connected graphs and induced ranking. Indeed, graphs in real-life scenarios, although might be non-connected, they typically consists of one dominant connected component and several small components that are usually neglected as, by default, the most important vertex appears in the dominant component. Moreover, in real-life graph-based applications, we are usually interested in the induced ranking rather than the absolute centrality values assigned to vertices.

6 MONOTONIC FILTERING FUNCTIONS

Until now, we considered arbitrary filtering functions without any restrictions. On the other hand, the filtering functions based applications, we are usually interested in the induced rank-appears in the dominant component. Moreover, in real-life graphs and induced ranking. Indeed, graphs in real-life scenarios, that are usually neglected as, by default, the most important vertex beyond the scope of this work, and it remains the subject of future logarithmic functions, etc.). However, such a thorough analysis is natural to ask Questions I and II for Stress function, which are the following: with every vertex motif measures are monotonic, which means that the bounded value is a subgraph motif measure. Notice that, for monotonic subgraph motif measures; hence the name “monotonic bounded value property”.

Definition 6.2 (Monotonic Bounded Value Property). A centrality measure C enjoys the monotonic bounded value property if, for every integer \( n > 0 \), \( |\text{Val}_n^\text{B}(C)| \leq n + 1 - |\text{Val}_n^\text{B}(C)| \).

We can now define a refined version of the bounded value property, which provides a better upper bound for \( |\text{Val}_n^\text{B}(C)| \):

\[
\text{BVal}_n^\text{B}(C) = \bigcup_{G \in \mathbb{G}} \text{Val}_n^\text{B}(C).
\]

We can now define a refined version of the bounded value property, which provides a better upper bound for \( |\text{Val}_n^\text{B}(C)| \):
Theorem 6.5. Consider a centrality measure $C$. The following statements are equivalent:

1. $C$ is a monotonic subgraph motif centrality measure relative to the induced ranking.
2. $C$ enjoys the non-uniform monotonic bounded value property.

Connected Graphs. Recall that the family of subgraph motif measures relative to the induced ranking provides a unifying framework for all centrality measures whenever we concentrate on connected graphs (see Theorem 5.11). Interestingly, a careful inspection of the proof of Theorem 5.11 reveals that this holds even for the family of monotonic subgraph motifs relative to the induced ranking.

Theorem 6.6. For a measure $C$ it holds that $\text{ConC}$ is a monotonic subgraph motif measure relative to the induced ranking.

7 CLASSIFICATION

The results of the previous sections provide tools that allow us to establish whether a measure can be expressed as a (monotonic) subgraph motif measure (relative to the induced ranking). With such technical tools available, the next step is to determine whether existing measures belong to the family of (monotonic) subgraph motif measures (relative to the induced ranking). Such a classification, apart from being interesting in its own right, will provide insights on the structural similarities and differences among existing centrality measures. To this end, we focus on established measures from the literature (including Stress, All-Subgraphs, and Closeness), and provide a rather complete classification depicted in Tables 1 and 2; the formal definitions of the considered measures are omitted, and can be found in the appendix. The second (resp., third) column determines whether the measure $C$ stated in the first column is subgraph motif (resp., subgraph motif relative to the induced ranking): $\checkmark$ means that it is, $\times$ means that it is not, $\times[\text{trees}]$ means that it is not even for trees, $\checkmark[\text{con}]$ means that it is over connected graphs, $\checkmark[\text{trees}]$ means that it is over trees, and ? means that it is open. Concerning Table 2, $\ast$ refers to any measure considered in Table 1 apart from Betweenness, and $\times[\text{con}]$ means that the respective measure (i.e., Betweenness) is not monotonic subgraph motif even for connected graphs. Note that Table 2 is identical to Table 1, apart from Betweenness, which is provably not monotonic subgraph motif (relative to the induced ranking).

We would like to remark that the result $\checkmark[\text{con}]$ for Eigenvector in both tables holds for a broader class of graphs than connected graphs. Moreover, we can show that Betweenness is a (monotonic) subgraph motif measure (relative to the induced ranking) for a class of graphs that captures the class of trees and is incomparable to the class of connected graphs. Nevertheless, for the sake of readability, we state our expressibility results only for trees and connected graphs; further details can be found in the appendix.

Take-home Messages. We proceed to highlight the key take-home messages of the above classification, which we believe they provide further insights concerning the centrality measures in question:

1. If we focus on the induced ranking rather than the absolute values over connected graphs, then the family of monotonic subgraph motif measures should be understood as a unifying framework that incorporates every other measure.
2. Our classification excludes a priori the adoption of certain centrality measures (e.g., Closeness, Harmonic, etc.) in applications where the importance of a vertex should be measured based on the subgraph motifs surrounding it.
3. Betweenness, which computes the percentage of the shortest paths in a graph going through a vertex, is of different nature compared to all the other measures. Notably, although it looks similar to Stress, it behaves in a significantly different way. The relationship of Betweenness with (monotonic) subgraph motif measures deserves further investigation.
4. There is a notable difference between the two feedback measures considered in our classification, i.e., PageRank and Eigenvector, that deserves further exploration. As mentioned above, Eigenvector is a (monotonic) subgraph measure relative to the induced ranking over a broader class $C$ of graphs than connected graphs, whereas PageRank is provable not a subgraph motif measure over the class $C$.

A Note on Directed Graphs. As discussed in the clarification remark at the end of Section 1, although our analysis (including the classification of this section) focused on undirected graphs, all the notions and results can be transferred to the case of directed graphs under the notion of weak connectedness. The only exception is the negative result $\times[\text{trees}]$ for Eigenvector in both Tables 1 and 2. Although we can show that for directed graphs, Eigenvector is not a (monotonic) subgraph motif centrality measure, it remains open whether these negative results hold even for directed trees (i.e., directed graphs whose underlying undirected graph is a tree).

8 CONCLUSIONS

We have provided a rather complete picture concerning the absolute expressiveness of the family of (monotonic) subgraph motif
centrality measures (relative to the induced ranking) by establishing precise characterizations. We have also presented a detailed classification of standard centrality measures by using the tools provided by the aforementioned characterizations. Recall that, although our development focused on undirected graphs, all the notions and results can be transferred to the case of directed graphs under the standard notion of weak connectedness.

At this point, we would like to stress that the machinery on graph colorings, introduced in Section 5, can be used to provide characterizations for all the families considered in the paper, and not only for the family of subgraph motif measures relative to the induced ranking. For example, you can show that a measure $C$ is subgraph motif iff there exists a precoloring of $V_G$ that is uniformly $C$-injective; the latter is defined as non-uniform $C$-injectivity with the difference that $C$-injectivity is enforced across all the graphs (not only inside a certain graph).

The obvious question that remains open is whether we can isolate a bounded-value-like property that characterizes subgraph motif measures relative to the induced ranking. We strongly believe that the coloring-based characterization established in this work (Theorem 6.5) is a useful tool towards such a bounded-value-like characterization. Finally, towards a deeper understanding of the family of subgraph motif measures, one should perform a more refined analysis by focussing on restricted classes of subgraph families and filtering functions that enjoy desirable structural properties.

REFERENCES

[1] Renzo Angles, Marcelo Arenas, Pablo Barceló, Aidan Hogan, Juan L. Reutter, and Domagoj Vrgoc. 2017. Foundations of Modern Query Languages for Graph Databases. ACM Comput. Surv. 50, 5 (2017), 68:1–68:40.

[2] Phillip Bonacich. 1987. Power and centrality: A family of measures. American journal of sociology 92, 5 (1987), 1170–1182.

[3] Stephen P. Borgatti and Martin G. Everett. 2006. A Graph-theoretic perspective on centrality. Soc. Networks 28, 4 (2006), 466–484.

[4] Zoltán Dezső and Albert-László Barabási. 2002. Halting viruses in scale-free networks. Phys. Rev. E 65 (2002), 055103. Issue 5.

[5] Martin G Everett and Stephen P Borgatti. 1998. Analyzing clique overlap. Connections 21, 1 (1998), 49–61.

[6] Linton C Freeman. 1977. A set of measures of centrality based on betweenness. Sociometry (1977), 35–41.

[7] Aric Hagberg, Pieter Swart, and Daniel S Chult. 2008. Exploring network structure, dynamics, and function using NetworkX. Technical Report.

[8] Xinyu Huang, Dongming Chen, Dongqi Wang, and Tao Ren. 2020. Identifying Influencers in Social Networks. Entropy 22, 4 (2020), 450.

[9] Gábor Iván and Vince Grolmusz. 2010. When the Web meets the cell: using personalized PageRank for analyzing protein interaction networks. Bioinformatics 27, 3 (2010), 405–407.

[10] Mitri Kitti. 2016. Axioms for centrality scoring with principal eigenvectors. Social Choice and Welfare 46, 3 (2016), 639–653.

[11] Dirk Koschützki, Henning Schwoebbermeyer, and Falk Schreiber. 2007. Ranking of network elements based on functional substructures. Journal of Theoretical Biology 248, 3 (2007), 471–479.

[12] A. Masoudi-Nejad, F. Schreiber, and Z.R.M. Kashani. 2012. Building Blocks of Biological Networks: A Review on Major Network Motif Discovery Algorithms. IET Systems Biology 6, 5 (2012), 164–174.

[13] Leonid Mirsky. 1971. Transversal Theory: An Account of Some Aspects of Combinatorial Mathematics. Academic Press.

[14] Mark Newman. 2018. Networks. Oxford University Press.

[15] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. The PageRank Citation Ranking: Bringing Order to the Web. Technical Report 1999-66.

[16] Richard Rado. 1949. Axiomatic treatment of rank in infinite sets. Canad. J. Math. (1949), 337–343.

[17] Cristian Riveros and Jorge Salas. 2020. A Family of Centrality Measures for Graph Data Based on Subgraphs. In ICDT 23:1–23:18.

[18] Yannick Rochat. 2009. Closeness centrality extended to unconnected graphs: The harmonic centrality index. Technical Report.

[19] Gert Sabidussi. 1966. The Centrality Index of a Graph. Psychometrika 31, 4 (1966), 581–603.

[20] Alfonso Shimbol. 1953. Structural Parameters of Communication Networks. Bull. Math. Biophysics 15 (1953), 501–507.

[21] Edward Sapielrajn. 1950. Sur l’extension de l’ordre partiel. Fundamenta Mathematicae 16 (1930), 386–389.

[22] René van den Brink and Robert P. Gilles. 2000. Measuring domination in directed networks. Social Networks 22, 2 (2000), 141–157.

[23] Tomasz Was and Oskar Skibski. 2018. Axiomatization of the PageRank Centrality. In IJCAI 3898–3904.
A PRELIMINARY NOTIONS

In this preliminary section, we present some auxiliary notions that have not been introduced in the main body of the paper, but are needed for the proofs given in the appendix. In particular, we are going to introduce some special classes of graphs (lines and stars), and a couple of useful properties of centrality measures (closure under isomorphism and locality).

Special Graphs

The line graph with \( m > 0 \) vertices, denoted \( L_m \), is the undirected graph \((V, E)\), where
\[
V = \{1, \ldots, m\} \quad \text{and} \quad E = \bigcup_{i \in \{1, \ldots, m-1\}} \{(i, i + 1)\}.
\]
The star graph with \( m > 0 \) vertices, denoted \( S_m \), is the undirected graph \((V, E)\), where
\[
V = \{1, \ldots, m\} \quad \text{and} \quad E = \bigcup_{i \in \{2, \ldots, m\}} \{(1, i)\}.
\]
We refer to the vertex \( 1 \) as the center of \( S_m \).

Closure Under Isomorphism

We say that two graphs \( G_1 \) and \( G_2 \) are isomorphic, denoted \( G_1 \cong G_2 \), if there exists a bijective function \( h : V(G_1) \rightarrow V(G_2) \) such that \( \{u, v\} \in E(G_1) \) iff \( \{h(u), h(v)\} \in E(G_2) \). Furthermore, given the vertices \( v_1 \in V(G_1) \) and \( v_2 \in V(G_2) \), we say that the pairs \( (v_1, G_1) \) and \( (v_2, G_2) \) are isomorphic, denoted \( (v_1, G_1) \cong (v_2, G_2) \), if \( G_1 \cong G_2 \) witnessed by \( h \) and \( h(v_1) = v_2 \). We are now ready to define the property of being closed under isomorphism for centrality measures.

Definition A.1. Consider a centrality measure \( C \). We say that \( C \) is closed under isomorphism if, for every two pairs \( (v_1, G_1) \in VG \) and \( (v_2, G_2) \in VG \), \( (v_1, G_1) \cong (v_2, G_2) \) implies \( C(v_1, G_1) = C(v_2, G_2) \).

In simple words, the above property states that the values assigned by a measure to the vertices of a graph do not depend on the names of the vertices, but only on the structure of the graph; hence, the above property is sometimes called anonymity.

Locality

We finally recall the so-called locality property. Recall that, given a graph \( G \) and a set of vertices \( S \subseteq V(G) \), we write \( G[S] \) for the subgraph of \( G \) induced by \( S \). Moreover, for a vertex \( v \in V(G) \), we write \( K_v(G) \) for the set of vertices occurring in the component of \( G \) containing the vertex \( v \). The definition of locality follows:

Definition A.2. A centrality measure \( C \) is local if, for every pair \( (v, G) \in VG \), \( C(v, G) = C(v, G[K_v(G)]) \).

In simple words, the above property tells us that the value assigned by a measure \( C \) to a vertex \( v \) in a graph \( G \) depends only on the component \( G' \in \text{Comp}(G) \) that contains \( v \), i.e., during the calculation of \( C(v, G) \), \( C \) considers only the component \( G' \).

B PROOFS OF SECTION 3

Proof of Proposition 3.4

Let \( \hat{G} = (\{v_1, v_2, v_3, v_4\}, \{(v_1, v_2)\}) \). Consider the centrality measure \( C \) defined as follows:
\[
C(v, G) = \begin{cases} 
1 & G \neq \hat{G} \\
1 + i & G = \hat{G} \text{ and } v = v_i \text{ for } i \in [4].
\end{cases}
\]
It suffices to focus on the set of graphs \( \mathcal{G}^* = \{G_v \mid v \in V\} \cup \{\hat{G}\} \subseteq G \). By providing an argument similar to that in the proof of Proposition 3.2, we can show that, for every subgraph family \( F \) and filtering function \( f \), there exist \( 1 \leq i < j \leq 4 \) such that \( C(F, f)(v_i, \hat{G}) = C(F, f)(v_j, \hat{G}) \). Therefore, \( \{(v_i, \hat{G}), (v_j, \hat{G})\} \subseteq \text{Rank}(C(F, f)) \). On the other hand, only \((v_i, \hat{G}), (v_j, \hat{G})\) belongs to \( \text{Rank}(C) \). Consequently, for every subgraph family \( F \) and filtering function \( f \), \( \text{Rank}(C) \neq \text{Rank}(C(F, f)) \), and the claim follows.

C PROOFS OF SECTION 4

Proof of Theorem 4.2

(1 \( \Rightarrow \) 2). By contradiction, assume that \( C \) does not enjoy the bounded value property, namely there exists an integer \( n \geq 1 \) such that \( |\text{Val}(C)| > n+1 \). By hypothesis, \( C \) is a subgraph motif centrality measure, and thus, there exist a subgraph family \( F \) and a filtering function \( f \) such that the following holds: for every \( (v, G) \in VG \), \( C(v, G) = C(F, f)(v, G) \). We now define the set
\[
B_n = \{|F(v, G)| \mid (v, G) \in VG \text{ and } |A(v, G)| \leq n\}.
\]
Clearly, \( |B_n| \leq n + 1 \) since \( F(v, G) \subseteq A(v, G) \). We also define the function \( h : Val^n(C) \to B_n \) such that
\[
h(C(v, G)) = |F(v, G)|.
\]
By the pigeonhole principle, \( h \) is not injective, i.e., there exist \( C(v_1, G_1) \) and \( C(v_2, G_2) \) such that \( C(v_1, G_1) \neq C(v_2, G_2) \) but \( |F(v_1, G_1)| = |F(v_2, G_2)| \). This contradicts the fact that \( C(v_1, G_1) = f(F(v_1, G_1)) \) \( \neq f(F(v_2, G_2)) = C(v_2, G_2) \).

\((2 \implies 1)\). The goal is to show that there exist a subgraph family \( F \) and a filtering function \( f \) such that, for every \( (v, G) \in VG \), \( C(v, G) = C(F, f)(v, G) \). We start by defining a total order \( \preceq \) over the set of values \( Val(C) = \bigcup_{i=1}^{\infty} Val^n(C) \). By definition, for every \( n, m > 0 \) such that \( n \leq m \), it holds that \( Val^m(C) \subseteq Val^n(C) \). In other words, as we increase the integer \( n \) we are adding new values to the set \( Val^n(C) \). We can now define the binary relation \( \preceq \) over \( Val(C) \) as follows: for each \( a, b \in Val(C) \), \( a \preceq b \) if \( a \leq b \), or there exists \( n \) such that \( a \in Val^n(C) \) but \( b \notin Val^n(C) \). It is easy to see that \( \preceq \) is a total order over \( Val(C) \), and thus, it is a total order over \( Val^n(C) \) for each \( n > 0 \). For notational convenience, in the rest of the proof we assume that \( Val(C) = \{a_1, a_2, a_3, \ldots\} \) and \( a_1 \preceq a_2 \preceq a_3 \preceq \cdots \).

By exploiting the total order \( \preceq \) over \( Val(C) \), we now define a subgraph family \( F \). Consider an arbitrary \( (v, G) \in VG \), and let \( n = |A(v, G)| \). By hypothesis, \( C \) enjoys the bounded value property, which in turn implies that \( |Val^m(C)| \leq n + 1 \). Therefore, \( C(v, G) \in \{a_1, a_2, \ldots, a_{|Val^n(C)|}\} = Val^n(C) \). We further observe that \( A(v, G) \) is a finite set, and assume that \( A(v, G) = \{S_1, S_2, \ldots, S_n\} \). The subgraph family \( F \) is defined as follows:
\[
C(v, G) = a_i \implies F(v, G) = \{S_1, \ldots, S_{i-1}\}.
\]
This is indeed a subgraph family since \( F(v, G) \subseteq A(v, G) \). Notice that \( |F(v, G)| = i - 1 \) for \( i \in \{1, \ldots, |Val^n(C)| + 1\} \). Finally, we define the filtering function \( f : \mathbb{N} \to Val(C) \) as follows: for each \( i \in \mathbb{N} \),
\[
f(i) = a_{i+1}.
\]
We proceed to show the following technical lemma, which essentially states that \( F \) and \( f \) capture our intention:

**Lemma C.1.** For every \( (v, G) \in VG \), \( C(v, G) = C(F, f)(v, G) \).

**Proof.** Let \( n = |A(v, G)| \). If \( C(v, G) = a_i \in Val^n(C) \), then \( |F(v, G)| = i - 1 \). Therefore, \( |F(v, G)| = i - 1 \). Conversely, if \( C(F, f)(v, G) = a_i \), then \( |F(v, G)| = i - 1 \), and thus, by construction, \( C(v, G) = a_i \).

By Lemma C.1, we get that \( C \) is a subgraph motif centrality measure, and Theorem 4.2 follows.

**Proof of Proposition 4.3**

Our goal is to show that Closeness does not enjoy the bounded value property, and thus, by Theorem 4.2, it is not a subgraph motif measure, as needed. To this end, we need to show that there exists an integer \( n > 0 \) such that \( |Val^n(\text{Closeness})| > n + 1 \). For notational convenience, given a pair \( (v, G) \in VG \), let \( SD(v, G) = \sum_{u \in E(v, G)} d_G(v, u) \). We observe that, for each \( n > 0 \),
\[
\sum_{x \in n}|Val^n(\text{Closeness})| > n + 1 \quad \text{iff} \quad \left| \left\{ (SD(v, G) \mid v \in V(G) \text{ and } |A(v, G)| \leq n) \right\} \right| > n + 1.
\]

Therefore, it suffices to show that there is \( n > 0 \) such that \( |E_G| > n + 1 \). The rest of the proof proceeds in two main steps:

1. For every \( n > 0 \), it is easy to see that \( SD(1, L_1) = \frac{n(n+1)}{2} \) and \( |A(1, L_1)| = n \). Therefore, it is clear that, for every \( i, j \in [n] \) with \( i \neq j \), \( SD(1, L_i) \neq SD(1, L_j) \), \( |A(1, L_i)| = n \), and \( |A(1, L_j)| = n \). Therefore, \( |E_G| > n \).

2. It remains to show that there exists \( n > 0 \) for which there are two pairs \((v_1, G_1), (v_2, G_2)\) of \( VG \) such that \( |A(v_1, G_1)| \leq n \) and \( |A(v_2, G_2)| \leq n \), but \( SD(v_1, G_1) \neq SD(v_2, G_2) \) and \( SD(1, L_i) \), for every \( i \leq n \), which implies that \( |E_G| > n + 1 \), as needed. We proceed to show that the latter holds for \( n = 5 \). It is easy to verify that the center of the star graph \( S_7 \) participates in eight connected subgraphs. Thus, any leaf vertex participates in \( 1 + 2^2 = 5 \) connected subgraphs. Therefore, \( |E(1, S_7)| \leq 5 \) for \( i \in \{2, 3, 4\} \). On the other hand, if we look at the center vertex of \( L_3 \), that is, vertex 2 of \( L_3 \), it participates in four connected subgraphs, and thus, \( |E(2, L_3)| \leq 5 \). Finally, we have that \( SD(1, S_7) = 5 \) whereas \( SD(2, L_3) = 2 \). Observe that there is no \( i \in \{1, \ldots, 5\} \) such that \( SD(1, L_i) \in \{2, 5\} \), and thus, \( |E_G| > 6 \).

**D PROOFS OF SECTION 5**

**Proof of Theorem 5.4**

\((1 \implies 2)\). Since, by hypothesis, \( C \) is a subgraph motif measure relative to the induced ranking, there are a subgraph family \( F \) and a filtering function \( f \) such that \( \text{Rank}(C) = \text{Rank}(C(F, f)) \). Given a finite set \( S \subseteq VG \), we define the function \( p_{C_S} : S \to \mathbb{N} \) as follows: for every \( (v, G) \in S \), \( p_{C_S}(v, G) = |F(v, G)| \). It is clear that \( p_{C_S} \) is a precoloring of \( S \) since, by definition, \( F(v, G) \subseteq A(v, G) \), and thus, \( p_{C_S}(v, G) \in \{0, \ldots, |A(v, G)|\} \).

It remains to show that \( p_{C_S} \) is (i) non-uniformly \( C \)-injective, and (ii) \( C \)-consistent, which in turn implies that \( S \) is \( C \)-colorable:

1. Since \( \text{Rank}(C) = \text{Rank}(C(F, f)) \), for every \( (u_1, G), (u_2, G) \in S \), it holds that \( C(u_1, G) \neq C(u_2, G) \) if \( C(F, f)(u_1, G) \neq C(F, f)(u_2, G) \).

Therefore, \( p_{C_S}(u_1, G) = |F(u_1, G)| \neq |F(u_2, G)| = p_{C_S}(u_2, G) \), which implies that \( p_{C_S} \) is non-uniformly \( C \)-injective.
The fact that $\{v_1, G_1\}$, which essentially states that every partial order can be extended into a total order. It remains to show that $\{v_2, E\}$ into some countable set $\mathbb{N}$.

In the remaining three cases, where $\{v_1, G_1\}$ and $\{v_2, G_2\}$, we use the fact that $\text{Rank}(C) = \text{Rank}(C(F,f))$, $C(F,f)(v_1, G_1) < C(F,f)(v_2, G_2) = C(F,f)(v_2, G_2)$. Therefore, $|F(v_1, G_1)| = |F(v_2, G_2)| = |F(v_2, G_2)|$. Consequently, using the fact that $\text{Rank}(C) = \text{Rank}(C(F,f))$, $C(F,f)(v_1, G_1) < C(F,f)(v_2, G_2) = C(F,f)(v_2, G_2)$. Therefore, $F(v_2, G_2)$, which is clearly a contradiction. Hence, $PC_x$ is also C-consistent.

$(2 \Rightarrow 3)$. The proof of this implication heavily relies on Rado’s Selection Principle (see Theorem 5.5). By hypothesis, for each $S \in \mathcal{P}_\text{fin}(VG)$, there exists a preordering of $S$, i.e., a function $p_{PC_x} : S \rightarrow \mathbb{N}$ that is uniformly C-injective and C-consistent. Since for every $(v, G) \in VG$, $A(v, G)$ is finite, we can conclude that the following holds: for every $(v, G) \in VG$, the set $\{p_{PC_x}(v, G) \mid S \in \mathcal{P}_\text{fin}(VG) \land (v, G) \in S\}$ is finite. This allows us to apply Theorem 5.5 with $A = VG$ and $B = \mathbb{N}$. Therefore, there exists a function $f : VG \rightarrow \mathbb{N}$ such that, for every $S \in \mathcal{P}_\text{fin}(VG)$, there exists $S' \in \mathcal{P}_\text{fin}(VG)$ with $S \subseteq S'$ and $f_{|S} = p_{PC_x}|_{S'}$. By exploiting the latter property of the function $f$ guaranteed by Theorem 5.5, and the fact that, for each $S \in \mathcal{P}_\text{fin}(VG)$, $p_{PC_x}$ is a preordering of $S$ that is uniformly C-injective and C-consistent, we can show that (i) $f$ is a preordering of $VG$ that is (ii) non-uniformly C-injective, and (iii) C-consistent, which in turn implies item (3):

(i) For every pair $(v, G) \in VG$, we have that $f(v, G) = p_{PC_x}(v, G)$ for some set $S \in \mathcal{P}_\text{fin}(VG)$ such that $\{(v, G)\} \subseteq S$. Since $p_{PC_x}$ is a preordering of $S$, we get that $f(v, G) = p_{PC_x}(v, G) \in \{0, \ldots, |A(v, G)|\}$, and thus, $f$ is a preordering of $VG$.

(ii) For every two pairs $(v_1, G)$ and $(v_2, G)$ such that $C(v_1, G) \neq C(v_2, G)$, we know there exists a finite set $S$ such that $\{(v_1, G), (v_2, G)\} \subseteq S$ with $p_{PC_x}(v_i, G) = f(v_i, G)$ for $i \in \{1, 2\}$. Therefore, since $C(v_1, G) \neq C(v_2, G)$ and $p_{PC_x}$ is uniformly C-injective, we get that $f(v_1, G) = p_{PC_x}(v_1, G) \neq p_{PC_x}(v_2, G) = f(v_2, G)$. This implies that $f$ is uniformly C-injective.

(iii) Finally, by contradiction, assume there are four pairs $(v_1, G_1), (v_2, G_1), (u_1, G_2)$, and $(u_2, G_2)$ in $VG$ such that $C(v_1, G_1) < C(v_2, G_2)$ and $C(u_1, G_2) < C(u_2, G_2)$, but $f(v_1, G_1) = f(u_1, G_2)$ and $f(v_2, G_1) = f(u_2, G_2)$. We know that there exists a finite set $S'$ such that $S = \{(v_1, G_1), (v_2, G_2), (u_1, G_2), (u_2, G_2)\} \subseteq S'$ with $f(v_1, G_1) = p_{PC_x}(v_1, G_1) = p_{PC_x}(u_1, G_2)$ for every $(v, G) \in S$. This implies that $p_{PC_x}$ is not C-consistent, which is clearly a contradiction.

This completes the proof of the second implication.

$(3 \Rightarrow 1)$. The goal is to devise a subgraph family $F$ and a filtering function $f$ such that $\text{Rank}(C) = \text{Rank}(C(F,f))$, which in turn proves item (1). By hypothesis, there exists a preordering $p_{PC}$ of $VG$ that is uniformly C-injective and C-consistent. We define $F$ in such a way that, for every $(v, G) \in VG$, $|F(v, G)| = p_{PC}(v, G)$; note that such a subgraph family exists since $p_{PC}(v, G) \in \{0, \ldots, |A(v, G)|\}$. We now proceed to define the filtering function $f$. Let $R_{PC}$ be the binary relation

\[ \{(i, j) \in \mathbb{N} \times \mathbb{N} \mid \text{there are } (u, G), (v, G) \in VG \text{ such that } C(u, G) < C(v, G), p_{PC}(u, G) = i, \text{ and } p_{PC}(v, G) = j\}. \]

The fact that $p_{PC}$ is non-uniformly C-injective allows us to conclude that $R_{PC}$ is irreflexive. Moreover, the $C$-consistency of $p_{PC}$ implies that $R_{PC}$ is asymmetric. Observe now that if we extend $R_{PC}$ into a total order $R_{PC}^*$ over $\mathbb{N}$, then show that $R_{PC}^*$ can be embedded into a preorder.

The binary relation $R_{PC}$ is first extended into the strict partial order $R_{PC}$ by simply taking its transitive closure. Now, the fact that $R_{PC}$ can be extended into a total order $R_{PC}^*$ over $\mathbb{N}$ follows by the order-extension principle (a.k.a. Szpilrajn Extension Theorem), shown by Szpilrajn in 1930 [21], which essentially states that every partial order can be extended into a total order. It remains to show that $R_{PC}^*$ can be embedded into a total order $\mathbb{N}$.

Case 1. Assume first that there exists a minimum element in the ordered set $(\mathbb{N}, R_{PC}^*)$, denoted $n_0$ and $n_1$, respectively. Consider a close range of the rational numbers $[a, b] \subseteq \mathbb{Q}$ with $a < b$. Fix any order of the set $\mathbb{N} \setminus \{n_0, n_1\} = \{a_1, a_2, \ldots\}$, and fix a different order of $[a, b] = \{b_1, b_2, \ldots\}$. This is possible because both sets are countable. We proceed to define the order preserving embedding of $(\mathbb{N}, R_{PC}^*)$ to $([a, b], <)$. First, define $f(n_0) = a$ and $f(n_1) = b$. Now, the back and forth method goes as follow:

(1) Take $a_i$ such that $i$ is the first index where $f(a_i)$ is assigned. Define $f(a_i) = b_j$ such that $b_j$ is the first element in the list without pre image from $f$, and respects the order embedding up to this point. In other words, for every $a_k_i$ and $b_k$, such that $f(a_k_i)$ and $f(b_k)$ are defined and $a_k_i R_{PC}^* a_k b_k$, it holds that $f(a_k_i) < b_j < f(b_k)$.

(2) Repeat the previous step until we cover the whole set $\mathbb{N}$.

Given $a_p$ and $a_q$ such that $a_p R_{PC}^* a_q$, where $f(a_p)$ and $f(a_q)$ are both assigned, since $([a, b], <)$ is a dense order, for every two elements there is always another element in between. Therefore, we can always find $b_j$ as stated in the previous steps. In conclusion, we can map $(\mathbb{N}, R_{PC}^*)$ to $([a, b], <)$ respecting both orders, and $f$ is precisely the filtering function that we are looking for.

Cases 2, 3 and 4. In the remaining three cases, where $(\mathbb{N}, R_{PC}^*)$ has only a minimum element, only a maximum element, and no extreme elements, respectively, the same argument as above applies using the sets $\mathbb{Q}^+$, $\mathbb{Q}^-$ and $\mathbb{Q}$, respectively, instead of $[a, b] \subseteq \mathbb{Q}$.

It remains to show that $\text{Rank}(C(F,f)) = \text{Rank}(C)$ for the family $F$ and filtering function $f$ defined above. Take any two pairs $(v_1, G)$ and $(v_2, G)$ such that $C(v_1, G) < C(v_2, G)$. By definition, we have that $(p_{PC}(v_1, G), p_{PC}(v_2, G)) \in R_{PC}^*$. This is equivalent to $f(p_{PC}(v_1, G)) = f(|F(v_1, G)|) < f(p_{PC}(v_2, G)) = f(|F(v_2, G)|)$. Therefore, $C(F,f)$ and $C$ induce the same ranking inside each graph, and the claim follows.
Proof of Proposition 5.7
Let \( pc \) be the non-uniform C-injective precoloring of \( C \), which exists by hypothesis. Consider a graph \( G \), and an integer \( n > 0 \). Define the set

\[
S_n = \{ pc(v, G) \mid |\Delta v (G, G)| \leq n \}.
\]

In other words, \( S_n \) collects all the colors assigned by \( pc \) to vertices with at most \( n \) connected subgraphs around them. We then have that \( |\text{Val}_G^n(C)| \leq |S_n| \) since \( pc \) is non-uniformly C-injective. Since \( pc \) is a precoloring, \( S_n \subseteq \{0, \ldots , n\} \), and thus, \( |S_n| \leq n + 1 \), as needed.

Proof of Proposition 5.9
We are going to devise a centrality measure that is not subgraph motif relative to the induced ranking, but it enjoys the non-uniform bounded value property. Consider the graphs

\[
G_1 = (\{v_1, v_2, v_3\}, \{\{v_2, v_3\}\}) \quad G_2 = (\{u_1, u_2, u_3\}, \{\{u_2, u_3\}\}) \quad G_3 = (\{w_1, w_2, w_3\}, \{\{w_2, w_3\}\}).
\]

Observe that these three graphs have the same structure, that is, one isolated node and two nodes connected by an edge. We now define the following centrality measure:

\[
C_{(\{v_i\}, G)} = \begin{cases} 1 & (v, G) \in (\{v_1, G_1\}, (u_2, G_2), (w_3, G_3)) \\ 2 & (v, G) \in (\{v_2, G_1\}, (u_3, G_2), (w_1, G_3)) \\ 3 & (v, G) \in (\{v_3, G_1\}, (u_1, G_2), (w_2, G_3)) \\ 1 & \text{otherwise.} \end{cases}
\]

Roughly speaking, \( C \) is defined in such a way that the position of the isolated vertex in the induced ranking is shifted. Indeed, in \( \text{Rank}(C) \), \((v_1, G_1)\) comes first, \((w_1, G_3)\) comes second, and \((u_1, G_2)\) comes third. It can be shown that \( C \) is not subgraph motif relative to the induced ranking since any subgraph motif measure will need too many values in \( \text{Val}_G^n(C) \), for some \( n \geq 1 \), in order to mimic \( \text{Rank}(C) \). On the other hand, for every \( i \in \{1, 2, 3\} \), \( |\text{Val}_G^n(C)| \leq n + 1 \), and thus, \( C \) enjoys the non-uniform bounded value property, as needed.

Proof of Proposition 5.10
We are going to show that there exists a graph \( G \) such that \( |\text{Val}_G^n(\text{Closeness})| > 6 \). This implies that Closeness does not enjoy the non-uniform bounded value property, and thus, by item (2) of Theorem 5.4, it is not a subgraph motif measure relative to the induced ranking, as needed. Consider the graphs \( L_1, L_2, \ldots , L_5, S_4 \). We define the graph \( G \) as the minimal graph with the following structural property: for every graph \( G' \in \{L_1, L_2, \ldots , L_5, S_4\} \), there exists a component \( G'' \in \text{Comp}(G) \) that is isomorphic to \( G' \). By employing an argument similar to that in the proof of Proposition 4.3, and exploiting the fact that Closeness is closed under isomorphism (see Definition A.1) and local (see Definition A.2), we can show that \( |\text{Val}_G^n(\text{Closeness})| > 6 \).

Proof of Theorem 5.11
We are going to define a subgraph family \( F \) and a filtering function \( f \) such that \( \text{Rank}(C) \cap \text{VCG}^2 = \text{Rank}(C(F, f)) \cap \text{VCG}^2 \), which in turn implies that ConC is a subgraph motif measure relative to the induced ranking, as needed. Consider an arbitrary connected graph \( G \). We first observe that, for every \( v \in V(G) \), it holds that \( |\Delta v (G, G)| \geq |V(G)| \) since every path from \( v \) to any other vertex in \( G \) is a connected subgraph containing \( v \). We then define the equivalence relation \( \equiv_{G} \) over \( V(G) \) as follows: \( v \equiv_{G} u \) if \( C(v, G) = C(u, G) \). Let \( V(G)/\equiv_{G} = \{C_1, \ldots , C_m\} \) be the equivalence classes of \( \equiv_{G} \). We can assume, without loss of generality, that, for every \( i, j \in [m] \), with \( C_i = [v]_{\equiv_{G}} \) and \( C_j = [u]_{\equiv_{G}} \), \( i < j \) implies \( C(u, G) \subset C(v, G) \). We then define the subgraph family \( F \) in such a way that, for every vertex \( v \in V(G) \), if \( |\delta v (G)| = i + 1 \), we define the filtering function \( f \) in such a way that, for every \( i \in \{0, \ldots , m - 1\} \), \( f(i) = i + 1 \). It is now not difficult to verify that indeed \( \text{Rank}(C) \cap \text{VCG}^2 = \text{Rank}(C(F, f)) \cap \text{VCG}^2 \), and the claim follows.

E PROOFS OF SECTION 6

Proof of Proposition 6.1
Let \( \hat{G} \) be the graph \( (\{v_1, v_2, v_3\}, \{\{v_2, v_3\}\}) \). Consider the centrality measure \( C \) defined as follows:

\[
C(v, G) = \begin{cases} 1 & G \neq \hat{G} \text{ or } (G = \hat{G} \text{ and } v = v_3) \\ 2 & G = \hat{G} \text{ and } v = v_1 \\ 3 & G = \hat{G} \text{ and } v = v_2. \end{cases}
\]

We first show that \( C \) is a subgraph motif measure. Notice that, for every vertex \( v \in V \), \( C(v, G_v) = C(v_3, \hat{G}) = 1 \). Hence, we have only two options concerning the set of connected subgraphs assigned to the vertices of \( \hat{G} \) by a subgraph family, and the filtering function, which are

\footnote{Note that for pairs \((u, H)\), where \( H \) is a non-connected graph, we can simply define \( F(u, H) \) as the empty set since it is irrelevant what \( F \) does over non-connected graphs.}
the following: with $G_{uv}$ being the single-edge graph $\{\{u, v\}\}$, either

$$F_1(u, \hat{G}) = \begin{cases} \emptyset & v = v_1 \\ \{G_{o}, G_{v_{o_{3}}n}\} & v = v_2 \\ \{G_{n}\} & v = v_3 \end{cases}$$

with $f_1(0) = 3$, $f_1(1) = 1$ and $f_1(2) = 2$, or

$$F_2(u, \hat{G}) = \begin{cases} \emptyset & v = v_1 \\ \{G_{o}, G_{v_{o_{2}}n}\} & v = v_2 \\ \emptyset & v = v_3 \end{cases}$$

with $f_2(0) = 1$, $f_2(1) = 3$ and $f_2(2) = 2$. Now, by extending $F_1$ and $F_2$ in such a way that, for $(v, G) \in VG$ with $G \neq \hat{G}$, $F_1(v, G) = \{G_{o}\}$ and $F_2(v, G) = \emptyset$, it is clear that $C = C(F_1, f_1) = C(F_2, f_2)$. Observe, however, that $C(F_1, f_1)$ and $C(F_2, f_2)$ are not monotonic, and the claim follows.

**Proof of Theorem 6.3**

$$(1 \Rightarrow 2).$$ By contradiction, assume that $C$ does not enjoy the monotonic bounded value property, namely there exists an integer $n \geq 1$ such that $|Val^p(C)| > n + 1 - |BV_\text{Val}^p(C)|$. By hypothesis, $C$ is a monotonic subgraph motif centrality measure, and thus, there exist a subgraph family $F$ and a monotonic filtering function $f$ such that the following holds: for every $(v, G) \in VG$, $C(v, G) = C(F, f)(v, G)$. By the pigeonhole principle, and the fact that $Val^p(C) \cap BV_\text{Val}^p(C) = \emptyset$, we can conclude that there is a value $b \in (Val^p(C) \cup BV_\text{Val}^p(C))$ such that the following holds: there exists a pair $(v, G) \in VG$ with $C(v, G) = b$ and $|f((v, G)) > n$. Observe that $b \notin Val^p(C)$ since this would imply that $|f((v, G)) \leq n$, which contradicts the fact that $|f((v, G)) > n$; therefore, $b \in BV_\text{Val}^p(C)$; This means that $b < \max Val^p(C)$. Therefore, for some pair $(v', G') \in VG$, we have that $f((v', G')) = \max Val^p(C) > f((v, G)) = b$, but $|f((v', G')) \leq |A(v', G')| \leq n < |f((v, G))|$. This in turn implies that $f$ is not a monotonic filtering function, which leads to a contradiction.

$$(2 \Rightarrow 1).$$ We show that there exist a subgraph family $F$ and a monotonic filtering function $f$ such that, for every $(v, G) \in VG$, $C(v, G) = C(F, f)(v, G)$. We start by defining a total order $\leq_C$ over the set of values $Val(C) = \bigcup_{i=1}^{n} Val^i(C)$. For some $n \geq 1$, let $D_n(C) = Val^p(C) \cup BV_\text{Val}^p(C)$. By hypothesis, $C$ enjoys the monotonic bounded value property, and thus,

$$|D_n(C)| = |Val^p(C)| + |BV_\text{Val}^p(C)| = |Val^p(C)| + |BV_\text{Val}^p(C)| \leq n + 1.$$ 

We proceed to show that, for every $a \in D_n(C)$ and $b \in D_{n+1}(C) \setminus D_n(C)$, it holds that $a < b$. By contradiction, assume that there are $a \in D_n(C)$ and $b \in D_{n+1}(C) \setminus D_n(C)$ such that $b < a \leq \max Val^p(C)$. We proceed by considering two cases:

- The first case is when $b \in Val^{p+1}(C)$. In this case, by definition, $b \in BV_\text{Val}^p(C)$. Therefore, $b \in D_n(C)$, which contradicts the fact that $b \notin D_{n+1}(C) \setminus D_n(C)$.
- The second case is when $b \in BV_\text{Val}^{p+1}(C)$. In this case, $b \in Val(C) \setminus Val^{p+1}(C) \leq Val(C)/Val(C)$ since $Val^p(C) \subseteq Val^{p+1}(C)$. Since $b < \max Val^p(C)$, by definition, we have that $b \in BV_\text{Val}^p(C)$. Therefore, $b \in D_n(C)$, which contradicts the fact that $b \notin D_{n+1}(C) \setminus D_n(C)$.

We have just seen that for every $n \geq 1$, $D_n(C)$ is finite. Moreover, every value of $D_{n+1}(C) \setminus D_n(C)$ is strictly greater than every value of $D_n(C)$. With the above two facts in place, we can easily define the desired total order $\leq_C$ over $Val(C)$: for each $a, b \in Val(C)$, $a \leq_C b$ if $a < b$. For notational convenience, in the rest of the proof we assume that $Val(C) = \{a_1, a_2, a_3, \ldots\}$ and $a_1 \leq_C a_2 \leq_C a_3 \leq_C \cdots$.

Having the total order $\leq_C$ over $Val(C)$ in place, we can now define the desired subgraph family $F$ and monotonic filtering function $f$. For an arbitrary pair $(v, G) \in VG$, we define $F(v, G)$ in such a way that

$$C(v, G) = a_i \quad \text{implies} \quad |f((v, G))| = i - 1.$$ 

We proceed to show that $F$ is well-defined. Assume that there is $(v, G) \in VG$ such that $|A(v, G)| \leq i - 1$ but $C(v, G) = a_i$; for brevity, let $n = |A(v, G)|$. We define the set $S_i = \{a \in D_n(C) \mid a \leq a_i\}$. Observe that

$$n + 1 < i = |S_i| \leq |D_n(C)| = |Val^p(C)| + |BV_\text{Val}^p(C)|,$$

which contradicts the fact that $C$ enjoys the monotonic bounded value property. We finally define the monotonic filtering function $f : \mathbb{N} \rightarrow Val(C)$ as follows: for each $i \in \mathbb{N},$

$$f(i) = a_{i+1}.$$ 

It is straightforward to see that, for every $(v, G) \in VG$, $C(v, G) = C(F, f)(v, G)$. Thus, $C$ is a monotonic subgraph motif centrality measure, and Theorem 6.3 follows.
Proof of Theorem 6.5

(1 \Rightarrow 2). Consider a centrality measure C that is monotonic subgraph motif relative to the induced ranking. By definition, there are a subgraph family F and a monotonic filtering function f such that \( \text{Rank}(C) = \text{Rank}(C(F, f)) \). Assume, by contradiction, that C does not enjoy the non-uniform monotonic bounded value property, i.e., there exist a graph G and an integer n ≥ 1 such that \(|\text{Val}_G^p(C)| + |BVal_G^q(C)| > n + 1\). Since \(|\text{Val}_G^p(C(F, f))| = |\text{Val}_G^q(C)|\) and \(|BVal_G^q(C(F, f))| = |BVal_G^p(C)|\),

\[
|\text{Val}_G^p(C(F, f))| + |BVal_G^q(C(F, f))| > n + 1.
\]

This means that C(F, f) does not enjoy the non-uniform bounded value property. Hence, by Theorem 6.3, we get that C(F, f) is not a subgraph motif measure relative to the induced ranking, which leads to contradiction. Consequently, C enjoys the non-uniform bounded value property, as needed.

(2 \Rightarrow 1). We need to show that there exist a subgraph family F and a monotonic filtering function f such that \( \text{Rank}(C) = \text{Rank}(C(F, f)) \). For a graph G, we define the equivalence relation \( \equiv_G \) over \( V(G) \), and we order the equivalence classes of \( V(G)/\equiv_G \) as done in the proof of Theorem 5.11, namely with \( V(G)/\equiv_G = \{C_1, \ldots, C_m\} \), we assume that, for every \( i, j \in [m] \), with \( C_i = [v]\equiv_G \) and \( C_j = [u]\equiv_G \), \( i < j \) implies \( C(v, G) < C(u, G) \). We then define the subgraph family F and the monotonic filtering function f as in the proof of Theorem 5.11: for every \( v \in V(G) \), \( |F(v, G)| = i - 1 \) if \( [v]\equiv_G = C_i \), and \( f(i) = i + 1 \). It is not difficult to verify that indeed \( \text{Rank}(C) = \text{Rank}(C(F, f)) \). It remains to show that the subgraph family F always exists.\(^1\) Assume, by contradiction, that there is a vertex \( v \in V(G) \) such that \( |A(v, G)| = n < i - 1 \). We can show, by case analysis, that

\[
[C(w, G) | C(w, G) < C(v, G)] \subseteq \text{Val}_G^p(C) \cup BVal_G^q(C).
\]

Indeed, if \( |A(w, G)| \leq |A(v, G)| = n \), then \( C(w, G) \in \text{Val}_G^p(C) \); otherwise, if \( |A(w, G)| > |A(v, G)| \), then, by definition, \( C(w, G) \in BVal_G^q(C) \). Therefore, we can conclude that

\[
n + 1 < i = |\{C_1, \ldots, C_i\}| = |\{C(w, G) | C(w, G) < C(v, G)\}| \leq |\text{Val}_G^p(C) \cup BVal_G^q(C)| = |\text{Val}_G^p(C)| + |BVal_G^q(C)|,
\]

which contradicts the fact the C enjoys the non-uniform monotonic bounded value property, and the claim follows.

Proof of Theorem 6.6

It is enough to observe that the filtering function f employed in the proof of Theorem 5.11 is monotonic.

F PROOFS OF SECTION 7

In this final section of the appendix, we provide formal proofs for the expressibility/inexpressibility results presented in Section 7 (see Tables 1 and 2). The fact that Stress and All-Subgraphs are (monotonic) subgraph motif (and thus, (monotonic) subgraph motif relative to the induced ranking) has been already observed in the main body of the paper. Indeed, we have seen that

\[
\text{Stress} = C(S, f_{\text{Stress}}) \quad \text{and} \quad \text{All-Subgraphs} = C(A, \log_2),
\]

where both \( f_{\text{Stress}} \) and \( \log_2 \) are monotonic filtering functions.

Concerning Closeness, we have already shown that it is not a subgraph motif measure (see Proposition 4.3), and thus, it is not a monotonic subgraph motif measure. A simple inspection of the proof of Proposition 4.3, reveals that actually Closeness is not a subgraph motif measure (and thus, not a monotonic subgraph motif measure) even if we concentrate on trees, as claimed in Tables 1 and 2. The fact that Closeness is not a subgraph motif measure (and thus, not a monotonic subgraph motif measure) relative to the induced ranking has been established by Proposition 5.10.

Finally, the fact that every centrality measure considered in the classification of Section 7 is (monotonic) subgraph motif relative to the induced ranking if we concentrate on connected graphs, is a consequence of Theorems 5.11 and 6.6. We proceed to establish the claims stated in Tables 1 and 2, but are not covered by the above discussion.

Degree

The degree centrality measure is one of the simplest measures that can be found in the literature. It states that the bigger the neighborhood of a vertex, the most central it is. Formally, given a graph \( G \) and \( v \in V(G) \), the degree centrality of \( v \) in \( G \) is

\[
\text{Degree}(v, G) = |N_G(v)|.
\]

We proceed to show the following result:

**Proposition F.1.** Degree is a monotonic subgraph motif measure.

\(^1\)Note that here the argument for this statement is slightly more complex than the one given in the proof of Theorem 5.11 for a similar statement since we do not deal only with connected graphs.
Proof. For the degree centrality, the important subgraphs for a vertex $v$ in a graph $G$ are simply the edges of the form $\{u, v\}$. Formally, we define the subgraph family $F_{\text{Degree}}$ in such a way that, for every $(v, G) \in VG$,

$$F_{\text{Degree}}(v, G) = \{ \{u, v\}, \{\{u, v\}\} \mid \{u, v\} \in E(G)\}.$$  

We also consider the filtering function $id$ that is the identity over $N$, i.e., $id(x) = x$, which is monotonic. Clearly,

$$\text{Degree}(v, G) = id(|F_{\text{Degree}}(v, G)|)$$

which in turn implies that

$$\text{Degree} = C(F_{\text{Degree}}, id)$$

and the claim follows.

Cross-Clique

The cross-clique centrality measure [5], was introduced in the context of analyzing virus and information propagation. It is defined as the amount of cliques containing a certain vertex. Given a graph $G$ and a vertex $v \in V(G)$, we define the set $\mathcal{C}(v, G) = \{S \in A(v, G) \mid S$ is a clique $\}$. The cross-clique centrality of $v$ in $G$ is

$$\text{Cross-Clique}(v, G) = \left|\mathcal{C}(v, G)\right|.$$  

It is straightforward to see that the following holds:

**Proposition F.2.** Cross-Clique is a monotonic subgraph motif measure.

Harmonic

This centrality measure is commonly considered as an extension of Closeness to disconnected graphs. Given a graph $G$ and a vertex $v \in V(G)$, the harmonic centrality [18] of $v$ in $G$ is defined as

$$\text{Harmonic}(v, G) = \begin{cases} 0 & \text{if } v \text{ is isolated in } G \\ \sum_{u \in V(G) \setminus \{v\}} \frac{1}{d_G(v, u)} & \text{otherwise.} \end{cases}$$

Recall that, for two nodes $v, u \in V(G)$, if there is no path between them, then $d_G(v, u) = \infty$. Therefore, a vertex $u \in V(G)$ that is not connected with $v$ in $G$ via a path, has no impact on the harmonic centrality of $v$ in $G$. Just as Closeness, Harmonic is usually called a geometrical measure since it relies on the distance inside a graph. We proceed to show the following:

**Proposition F.3.** The following hold:

1. Harmonic is not a subgraph motif measure, even if we focus on trees.
2. Harmonic is not a subgraph motif measure relative to the induced ranking.

**Proof.** Item (1). We follow the same strategy as in the proof of Proposition 4.3. We first observe that $\text{Harmonic}(1, L_t) = \sum_{j=1}^{\lfloor \frac{n}{j} \rfloor} \frac{1}{j}$. Therefore, for $n > 1$, we have that $\text{Harmonic}(1, L_n) = \text{Harmonic}(1, L_n) + \frac{1}{n}$. This means $\text{Harmonic}(1, L_i) \neq \text{Harmonic}(1, L_j)$ for every $i, j$ with $i \neq j$. Hence, $|\text{Val}^R(\text{Harmonic})| \geq n - 1$. Since $\text{Harmonic}(1, L_1) = 0$, we get that $|\text{Val}^R(\text{Harmonic})| \geq n$. Now, let $n = 5$. There are two pairs left: $(2, L_3)$ and $(1, S_3)$. We first observe that, for every $n > 0$,

$$\text{Harmonic}(2, L_3) = 2 > \sum_{i=1}^{n-1} \frac{1}{i} = \text{Harmonic}(1, L_n).$$

Moreover, we have that

$$\text{Harmonic}(1, S_3) = 1 + \frac{1}{2} + \frac{1}{2} = 2.$$  

Therefore, $|\text{Val}^R(\text{Harmonic})| = 6$, which means Harmonic does not violate the bounded value property for $n = 5$. However, for $n = 6$ we get that $|\text{Val}^R(\text{Harmonic})| \geq 7$ since, by definition, $\text{Val}^R(\text{Harmonic}) \subseteq \text{Val}^R(\text{Harmonic})$. Now, for the pair $(2, L_4)$ we have $|\text{Val}^R(\text{Harmonic})| = 6$, whereas $\text{Harmonic}(2, L_4) = 2.5$. Therefore, $|\text{Val}^R(\text{Harmonic})| > 7$. Hence, Harmonic violates the bounded value property. Thus, by Theorem 4.2, we conclude that Harmonic is not a subgraph motif measure. This hold even if we concentrate on trees since the line and star graphs used above are trees.

Item (2). By Theorem 5.4, it suffices to show that Harmonic does not enjoy the non-uniform bounded value property. To this end, we follow the same idea as in the proof of Proposition 5.10 to, roughly speaking, gather all the graphs employed above to prove that $|\text{Val}^R(\text{Harmonic})| > 7$ in one disconnected graph $G$. Then, by exploiting the fact that Harmonic is closed under isomorphism (see Definition A.1) and local (see Definition A.2), we can show that $|\text{Val}^R(\text{Harmonic})| > 7.$  

\[ \blacksquare \]
PageRank

This measure, introduced in [15], is used in Google’s web searching engine. Given a graph \(G\) with \(V(G) = \{v_1, \ldots, v_{|V(G)|}\}\), we define the adjacency matrix of \(G\) as the matrix \(A_G\) such that \(A_{Gij} = 1\) if \(\{u_i, u_j\} \in E(G)\). We then define \(P\) as the column-stochastic matrix such that:

\[
P_{ij} = \frac{A_{ij}}{|N_G(i)|}
\]

In other words, \(P\) contains the probabilities of jumping from node \(i\) to node \(j\) during a random walk over \(G\). Let \(v\) be a stochastic vector \((1^T v = 1)\), and let \(0 < \alpha < 1\) be a teleportation parameter. Then, the \((v, \alpha)\)-PageRank vector is defined as the vector \(x_{v,\alpha}\) that solves the equation \((I - \alpha P)x = (1 - \alpha)v\). Now, given a graph \(G\) with \(V(G) = \{v_1, \ldots, v_{|V(G)|}\}\), the PageRank centrality of \(v_i\) in \(G\) is defined as

\[
\text{PageRank}(v_i, G) = x_{0.8,v}[1].
\]

where \(u = 1 \cdot \frac{1}{|V(G)|}\). Several variations of PageRank have been proposed in the literature. In particular, if we do not force \(x_{0.8,u}\) to be a probability distribution, we can ask that the sum of centralities inside a connected component is equal to the number of nodes in that component. Therefore, this give rise to the local version of PageRank, denoted LPR. We call it local since the values in a connected component are always the same regardless of the graph.

**Proposition F.4.** The following hold:

1. PageRank is not a subgraph motif measure, even if we focus on trees.
2. PageRank is not a subgraph motif measure relative to the induced ranking.

**Proof.** For both items, it suffices to consider the local version of PageRank since \(\text{Rank(PageRank)} = \text{Rank(LPR)}\).

**Item (1).** It can verified that \(\text{Val}^\delta_{L^\delta}(\text{LPR}) = \{1\}\), \(\text{Val}^\delta_{L^\delta_1}(\text{LPR}) = \{0.5, 0.33\}\), \(\text{Val}^\delta_{L^\delta_2}(\text{LPR}) = \{0.17, 0.25\}\), \(\text{Val}^\delta_{L^\delta_3}(\text{LPR}) = \{0.1\}\), \(\text{Val}^\delta_{L^\delta_4}(\text{LPR}) = \{0.07\}\), and \(\text{Val}^\delta_{L^\delta_5}(\text{LPR}) = \{0.2\}\). Therefore,

\[
\{0.5, 0.33, 0.17, 0.25, 0.1, 0.07, 0.2\} \subseteq \text{Val}^\delta(\text{LPR})
\]

Thus, \(|\text{Val}^\delta(\text{LPR})| > 7\), which implies that LPR does not enjoy the bounded value property. Hence, by Theorem 4.2, we conclude that LPR is not a subgraph motif measure. This holds even if we focus on trees since the line graphs considered above are trees.

**Item (2).** By Theorem 5.4, it suffices to show that LPR does not enjoy the non-uniform bounded value property. To this end, we follow the same idea as in the proof of Proposition 5.10 to, roughly speaking, gather all the graphs employed above to prove that \(|\text{Val}^\delta(\text{LPR})| > 7\) in one disconnected graph \(G\). Then, by exploiting the fact that LPR is closed under isomorphism (see Definition A.1) and local (see Definition A.2), we can show that \(|\text{Val}^\delta_G(\text{LPR})| > 7\), and the claim follows.

**Eigenvector**

Given a graph \(G\) with \(V(G) = \{u_1, \ldots, u_{|V(G)|}\}\), we define the adjacency matrix of \(G\) as the matrix \(A_G\) such that \(A_{Gij} = 1\) if \(\{u_i, u_j\} \in E(G)\). Let \(\lambda_{\max}(G)\) be the greatest eigenvalue of \(A_G\). We define the unique \(L^2\) normalized eigenvector associated with \(\lambda_{\max}(G)\) as \(v_{\max}\). The eigenvector centrality [2] of \(u_i\) in \(G\) is the \(i\)-th entrance of \(v_{\max}\):

\[
\text{Eigenvector}(u_i, G) = v_{\max}^i.
\]

We proceed to show the following:

**Proposition F.5.** Eigenvector is not a subgraph motif measure, even if we focus on trees.

**Proof.** The power method is the usual way to compute Eigenvector [7]. With an initial vector \(v = 1\), a tolerance factor \(\epsilon = 1.0e - 6\), and precision of three decimals, we get the following values:

\[
\begin{align*}
\text{Val}_{L^\delta}^\delta_1 (\text{Eigenvector}) &= \{0.5, 0.707\} & \text{Val}_{L^\delta_1}^\delta (\text{Eigenvector}) &= \{0.371, 0.601\} & \text{Val}_{L^\delta_1}^\delta (\text{Eigenvector}) &= \{0.288\} \\
\text{Val}_{L^\delta_2}^\delta (\text{Eigenvector}) &= \{0.231\} & \text{Val}_{L^\delta_2}^\delta (\text{Eigenvector}) &= \{0.191\} & \text{Val}_{L^\delta_2}^\delta (\text{Eigenvector}) &= \{0.408\}
\end{align*}
\]

and, for \(T = \{\{1, 2, 3, 4, 5\}, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{2, 5\}\}\),

\[
\text{Val}_{L^\delta}^\delta (\text{Eigenvector}) = \{0.353\}
\]

Therefore, Eigenvector does not enjoy the bounded value property, and thus, by Theorem 4.2, it is not a subgraph motif measure. This holds even for trees since all the graphs used above are indeed trees.

By Theorem 6.6, Eigenvector is a subgraph motif measure relative to the induced ranking over connected graphs. We proceed to strengthen this result by showing that it holds over a broader class of graphs than the class of connected graphs. Let \(K_1 = \{G \in G \mid \text{for every } G_1, G_2 \in \text{Comp}(G), \lambda_{\max}(G[V(G_1)]) = \lambda_{\max}(G[V(G_2)]) = \lambda_{\max}(G) \text{ implies } G_1 \cong G_2\}\).

It is easy to verify that every connected graph belongs to \(K_1\), and that \(K_1\) may contain disconnected graphs. We show that:
Proposition F.6. Eigenvector is a subgraph motif measure relative to the induced ranking when restricted to $K_1$.

Proof. From the proof of Proposition F.5, we know that for every component of $M(G) = \{S \in \text{Comp}(G) \mid \lambda_{\text{max}}(G[S]) = \lambda_{\text{max}}(G)\}$, the eigenvector centrality assigns the same amount of values to the connected component as it seen as a connected graph. On the other hand, if $S \notin M(G)$, then, for every vertex $v \in V(S)$, we have that Eigenvector$(v, G) = 0$. Note that, whenever two graphs are isomorphic they have the same values in every eigenvector. Values are just permuted depending on the isomorphism you map each graph to the other. Therefore, we have that, for every graph $G \in K_1$, there exists $S \in M(G)$ such that $\text{Val}_{G}^{|V(S)|}$(Eigenvector) = $\text{Val}_{G}^{|V(S)|}$(Eigenvector). In other words, when Eigenvector is restricted to $K_1$, it behaves like ConEigenvector. Hence, we can use the same subgraph family and filter function employed in the proof of Proposition 5.11 with the addition of value 0 for vertices outside $S$ in $M(G)$.

Betweenness

This centrality measure is similar to Stress with the difference that we take the proportion of shortest paths going through a certain vertex. Formally, given a graph $G$ and a vertex $v \in V(G)$, the betweenness centrality [6] of $v$ in $G$ is

$$\text{Betweenness}(v, G) = \sum_{(u, w) \in (K_n(G) \setminus \{v\})^2} \frac{|S_G(v, u, w)|}{|S_G(u, w)|}$$

We proceed to show the following result:

Proposition F.7. Betweenness is a subgraph motif measure when restricted to trees.

Proof. We need to show that there exist a subgraph family $F$ and a filtering function $f$ such that Betweenness = $C(F, f)$. Since we focus on trees, there is natural subgraph family that does the job. For every tree $T$ and $v \in V(T)$, we define $F$ as follows:

$$F(v, T) = \{p_{uw} \in A(v, T) \mid u, w \in V(T) \setminus \{v\} \text{ and } p_{uw} \in S_T(u, w)\}.$$  

For the filtering function we simply use the function $f_{c, 2}$, i.e., $f_{c, 2}(x) = 2x$. It remains to show that, for every $v \in V(T)$, Betweenness$(v, T) = f_{c, 2}(|F(v, T)|)$. In trees we know that, for every pair of vertices $u, v$, it holds that $|S_G(u, v)| = 1$. Therefore, for any $v \in V(T)$, we have that $\text{Betweenness}(v, T) = \text{Stress}(v, T) = f_{c, 2}(|F(v, T)|)$, and the claim follows.

One can define a broader class of graphs than the class of trees over which Betweenness still a subgraph motif measure. Let

$$C_1 = \{G \in G \mid \text{ for every } v, u \in V(G), |S_G(v, u)| \leq 1\},$$

i.e., it collects all the graphs such that, for every two nodes $v, u$ in $G$, there exists exactly one shortest path from $v$ to $u$, a property that trees trivially satisfy. It is then not difficult to show the following result by exploiting the subgraph family and the filtering function used in the proof of Proposition F.7:

Proposition F.8. Betweenness is a subgraph motif measure when restricted to $C_1$.

It turned out that the question whether Betweenness is a subgraph motif in general is a non-trivial one that remains open. We conclude this section by dealing with rankings. We proceed to show the following inexpressibility result:

Proposition F.9. The following hold:

1. Betweenness is not a monotonic subgraph motif measure, even if we focus on connected graphs.
2. Betweenness is not a monotonic subgraph motif measure relative to the induced ranking.

Proof. (1) We are going to show that Betweenness does not enjoy the monotonic bounded value property, and thus, by Theorem 6.3, it is not a monotonic subgraph motif measure. For $n \geq 1$, consider the connected graph $H_n = (V_n, E_n)$, where

$$V_n = \{1, 2, \ldots, n + 2\} \quad \text{and} \quad E_n = \{(1, i) \mid i \in \{3, 4, \ldots, n + 2\} \cup \{(2, i) \mid i \in \{3, 4, \ldots, n + 2\}\}.$$  

It can be verified that $\text{Betweenness}(3, H_n) = \frac{2}{n}$. However, for $n > 1$, $|\mathcal{A}(3, H_n)| > 4$. Therefore, we have that

$$|\text{Val}^1(\text{Betweenness})| + |\text{BVal}^1(\text{Betweenness})| > 5,$$

which implies that Betweenness does not enjoy the monotonic bounded value property, as needed.

Item(2). By Theorem 5.4, it suffices to show that Betweenness does not enjoy the non-uniform bounded value property. Since Betweenness is closed under isomorphism and local, we will use the same strategy as we did in analogous proofs. In particular, we define the graph $G$ as the minimal graph with the following property: for every graph $G' \in \{H_1, H_2, \ldots, H_5, H_6\}$, there exists a component $G'' \in \text{Comp}(G)$ such that $G''$ is isomorphic to $G'$. We can then show that

$$|\text{Val}^1_G(\text{Betweenness})| + |\text{BVal}^1_G(\text{Betweenness})| > 5,$$

which implies that Betweenness does not enjoy the non-uniform monotonic bounded value property, and the claim follows.
G  DIRECTED GRAPHS

Preliminaries

A directed graph (or digraph for short) \(G\) is a pair \((V, E)\), where \(V\) is a finite non-empty subset of \(V\) (the set of vertices of \(G\)), and \(E \subseteq \{(u, v) \mid u, v \in V\}\) (the set of edges of \(G\)). As for undirected graphs, we write \(V(G)\) and \(E(G)\) for the set of vertices and edges of \(G\), respectively. A multi-graph \(M\) is defined in the same way as undirected graphs, with the key difference that the edges from a multiset (or a bag), i.e., we can have more than one undirected edges between two vertices. The notations \(V(M)\) and \(E(M)\) can be naturally defined for multi-graphs. Given a directed graph \(G\), the undirected version of it, denoted \(\overline{G}(G)\), is defined as the multi-graph \((V(G), \{(u, v) \mid (u, v) \in E(G)\text{ or } (v, u) \in E(G)\})\).

Note that we use the notation \([\cdot]\) to denote bags. For example, \([a, a, b]\) is the bag consisting of two occurrences of \(a\) and one occurrence of \(b\). We say that a directed graph \(G\) is connected if the multi-graph \(\overline{G}(G)\) is connected. This notion of connectedness is known in the literature as weak connectedness. In simple words, \(G\) is connected if the multi-graph obtained after removing the directions of the edges is connected.

For a vertex \(v\) in a directed graph \(G\), we define the children of \(v\) in \(G\) as the set \(\text{OUT}(v, G) = \{u \in V(G) \mid (u, v) \in E(G)\}\) and the parents of \(v\) in \(G\) as the set \(\text{IN}(v, G) = \{u \in V(G) \mid (v, u) \in E(G)\}\). A connected subgraph of \(G\) is any graph \(G'\) such that \(V(G') \subseteq V(G)\) and \(E(G') \subseteq E(G) \cap (V(G') \times V(G'))\), and we write \(G' \subseteq G\) to indicate that \(G'\) is a subgraph of \(G\). Given a vertex \(v \in V(G)\), we define the set of all connected subgraphs containing \(v\) in \(G\) as \(\mathcal{A}(v, G) = \{G' \subseteq G \mid v \in V(G')\text{ and }G'\text{ is connected}\}\). Given two vertices \(u, v\) in \(V(G)\), a path connecting \(u\) and \(v\) is a sequence of vertices \(\pi_{uv} = w_1 w_2 \ldots w_n\) such that \(w_1 = u\), \(w_n = v\) and, for every \(i \in \{1, ..., n - 1\}\), \((w_i, w_{i+1}) \in E(G)\). The size of \(\pi_{uv}\) is defined as \(|\pi_{uv}| = n - 1\). By convention, \(\pi_{uu} = u\) and \(|\pi_{uu}| = 0\). A connected component (or simply component) of \(G\) is an induced subgraph \(G[S]\) of \(G\), where \(S \subseteq V(G)\), such that \(G[S]\) is connected, and, for every \(v \in V(G)\) \(\setminus S\), there is no path in \(G\) from \(v\) to a vertex in \(S\) neither a path from any vertex in \(S\) to \(v\). We denote by \(\text{Comp}(G)\) all the components of \(G\), that is, the set of graphs \(\{G' \mid G'\text{ is a component of }G\}\). We define the reachable component of a node \(v\) in \(G\) as the set \(\mathcal{K}_v(G) \subseteq V(G)\) such that \(\mathcal{K}_v(G) = \{u \in V(G) \mid \text{there exists a path from } u \text{ to } v\}\). We denote by \(\mathcal{K}\) the set of all digraphs, and by \(\mathcal{V}\) the set of vertex-digraph pairs \((u, G) \in \mathcal{V} \times \mathcal{G}\) \(\{u \in V(G)\}\). We note that we overload the notation and we use \(\mathcal{G}\) and \(\mathcal{V}\), as for undirected graphs, to refer to the set of all digraphs and the set of vertex-digraph pairs, respectively. This will not cause any confusion as the rest of the appendix is about digraphs.

A centrality measure for digraphs is a function \(C : \mathcal{V} \rightarrow \mathbb{R}\). In the rest of the appendix, we may say graph meaning digraph, and we only refer to centrality measures for digraphs.

Special Graphs

The line graph with \(m > 0\) vertices, denoted \(L_m\), is the digraph \((V, E)\), where

\[ V = \{1, \ldots, m\} \quad \text{and} \quad E = \bigcup_{i \in \{1, \ldots, m-1\}} \{(i, i + 1)\}. \]

The star graph with \(m > 0\) vertices, denoted \(S_m\), is the digraph \((V, E)\), where

\[ V = \{1, \ldots, m\} \quad \text{and} \quad E = \bigcup_{i \in \{2, \ldots, m\}} \{(i, 1)\}. \]

We refer to the vertex 1 as the center of \(S_m\).

Closure Under Isomorphism

We say that two digraphs \(G_1\) and \(G_2\) are isomorphic, denoted \(G_1 \cong G_2\), if there exists a bijective function \(h : V(G_1) \rightarrow V(G_2)\) such that \((v, w) \in E(G_1)\) iff \((h(v), h(w)) \in E(G_2)\). Furthermore, given the vertices \(v_1 \in V(G_1)\) and \(v_2 \in V(G_2)\), we say that the pairs \((v_1, G_1)\) and \((v_2, G_2)\) are isomorphic, denoted \((v_1, G_1) \cong (v_2, G_2)\), if \(G_1 \cong G_2\) witnessed by \(h\), and \(h(v_1) = v_2\). We are now ready to define the property of being closed under isomorphism for centrality measures.

**Definition G.1.** Consider a centrality measure \(C\). We say that \(C\) is closed under isomorphism if, for every two pairs \((v_1, G_1) \in \mathcal{V}\) and \((v_2, G_2) \in \mathcal{V}\), \((v_1, G_1) \cong (v_2, G_2)\) implies \(C(v_1, G_1) = C(v_2, G_2)\).

In simple words, the above property states that the values assigned by a measure to the vertices of a graph do not depend on the names of the vertices, but only on the structure of the graph; hence, the above property is sometimes called anonymity.

Locality

We finally recall the so-called locality property. Recall that, given a graph \(G\) and a set of vertices \(S \subseteq V(G)\), we write \(G[S]\) for the subgraph of \(G\) induced by \(S\). The definition of locality follows:

**Definition G.2.** A centrality measure \(C\) is local if, for every pair \((v, G) \in \mathcal{V}\), \(C(v, G) = C(v, G[K_v(G)])\).

In simple words, the above property tells us that the value assigned by a measure \(C\) to a vertex \(v\) in a graph \(G\) depends only on the component \(G' \in \text{Comp}(G)\) that contains \(v\), i.e., during the calculation of \(C(v, G)\), \(C\) considers only the component \(G'\).
We now state our main characterizations concerning monotonic subgraph motif measures for the case of directed graphs, which are similar to the ones for undirected graphs. Nevertheless, for the sake of completeness and clarity, we give here all the definitions. A subgraph family as a function $F : VG \rightarrow \mathcal{P}(G)$ such that for every pair $(v, G) \in VG$, $F(v, G) \subseteq A(v, G)$. A filtering function is a function $f : \mathbb{N} \rightarrow \mathbb{R}$.

Definition G.3 (Subgraph Motif Measure (for Directed Graphs)). Consider a subgraph family $F$ and a filtering function $f$. The $(F, f)$-measure is the function $C(F, f) : VG \rightarrow \mathbb{R}$ such that, for $(v, G) \in VG$, $C(F, f)(v, G) = f(|F(v, G)|)$. A centrality measure $C$ is a subgraph motif centrality measure if there is a subgraph family $F$ and a filtering function $f$ such that $C$ is the $(F, f)$-measure, namely $C = C(F, f)$. ■

We can define the sets $Val^p_G(C)$ and $Val^p(C)$ in exactly the same way as for undirected graphs. The key difference is that now we might have more values for a fixed integer $n$. In other words, in general, there are more pairs $(v, G)$ such that $|A(v, G)| \leq n$ when $G$ is a directed graph rather than being undirected. Analogously, we can define the notion of $C$-colorability for a measure $C$ over directed graphs. We are now ready to state our main characterizations for the case of directed graphs, which are similar to the ones for undirected graphs.

**Theorem G.4.** Consider a centrality measure $C$ for digraphs. The following statements are equivalent:

1. $C$ is a subgraphs motif centrality measure for digraphs.
2. For every $n > 0$, $|Val^p(C)| \leq n + 1$.

**Theorem G.5.** Consider a centrality measure $C$ for digraphs. The following statements are equivalent:

1. $C$ is a subgraph motif centrality measure for digraphs relative to the induced ranking.
2. Every finite set $S \subseteq VG$ is $C$-colorable.
3. $VG$ is $C$-colorable.

The above statements look the same as Theorems 4.2 and 5.4. Note, however, that they are semantically different since all the underlying notion (such as $Val^p(C)$ and $VG$ are for directed graphs). Interestingly, the proofs are identical with those for Theorems 4.2 and 5.4. We can finally state a result, analogous to Proposition 5.7, for directed graphs, which essentially states that the non-uniform bounded value property for directed graphs is a necessary condition for a measure for digraphs being subgraph motif relative to induced ranking.

**Proposition G.6.** Consider a centrality measure $C$ for digraphs. If there exists a precoloring of $VG$ that is non-uniformly $C$-injective, then $C$ enjoys the non-uniform bounded value property.

**Monotonic Subgraph Motif Measures for Digraphs**

We now state our main characterizations concerning monotonic subgraph motif measures for the case of directed graphs, which are similar to the ones for undirected graphs. Note that the sets $BVal^p_G(C)$ and $BVal^p(C)$ are defined in exactly the same way as for undirected graphs.

**Theorem G.7.** Consider a centrality measure $C$ for digraphs. The following statements are equivalent:

1. $C$ is a monotonic subgraph motif centrality measure for digraphs.
2. For every $n > 0$, $|Val^p(C)| \leq n + 1 - |BVal^p(C)|$.

**Theorem G.8.** Consider a centrality measure $C$ for digraphs. The following statements are equivalent:

1. $C$ is a monotonic subgraph motif centrality measure for digraphs relative to the induced ranking.
2. For every digraph $G$ and $n > 0$, $|Val^p_G(C)| \leq n + 1 - |BVal^p_G(C)|$.

As for the non-monotonic case, the proofs of the above results are identical with those for Theorems 6.3 and 6.5, respectively. This is because the proofs only exploit the sizes of the sets $Val^p_G(C)$, $Val^p(C)$, $BVal^p_G(C)$ and $BVal^p(C)$ rather than the structure of their members.

We can finally show a result, analogous to Theorem 5.11, for directed graphs, which essentially provides a unifying framework for all centrality measures over digraphs as long as we focus on connected graphs, and we are only interested in the induced ranking. The notation $ConC$ for a centrality measures $C$ for digraphs is defined in the same way as for undirected graphs.

**Theorem G.9.** For a centrality measure $C$ for digraphs it holds that $ConC$ is a monotonic subgraph motif measure for digraphs relative to the induced ranking.

**Proof.** The proof is along the lines of the proof for Theorem 5.11. Given a connected digraph $G$, we build the equivalence relation $\equiv_G$ over $V(G)$ in the same way as in the proof of Theorem 5.11. Due to the fact that $G$ is a connected digraph, for every vertex $v \in V(G)$, $|A(v, G)| \geq |V(G)|$ since between $v$ and any other node $u \in V(G)$ there exists a connected subgraph containing $v$ and $u$. Such a subgraph can be seen as an undirected path in the multi-graph $UN(G)$. Now, we can define a subgraph family $F$ such that, for every $(v, G)$ with $[v]_{\equiv_G} = C_i$, $|F(v, G)| = i - 1$. Finally, with the filtering function $f$ such that $f(i) = i + 1$, it is easy to verify that $\text{Rank}(C) = \text{Rank}(C(F, f))$.

---

**Subgraph Motif Measures for Directed Graphs**

We proceed to introduce the family of subgraph motif measures for directed graphs. This is actually done in the exactly the same way as for undirected graphs. Nevertheless, for the sake of completeness and clarity, we give here all the definitions. A subgraph family as a function $F : VG \rightarrow \mathcal{P}(G)$ such that for every pair $(v, G) \in VG$, $F(v, G) \subseteq A(v, G)$. A filtering function is a function $f : \mathbb{N} \rightarrow \mathbb{R}$.
We can show the following result: All-Subgraphs. The all-subgraphs centrality measure can be defined for digraphs in the obvious way. It is easy to see that:

**Degree.** Given a digraph $G$ and a vertex $v \in V(G)$, the degree centrality of $v$ in $G$ is

$$\text{Degree}(v, G) = |\text{OUT}(v, G)| + |\text{IN}(v, G)|.$$  

We can show the following result:

**Proposition G.12.** Degree is a monotonic subgraph motif measure for directed graphs.

**Proof.** Consider the subgraph family $D$ such that, for every $(v, G) \in V G$,  

$$D(v, G) = \{((v, u), ((v, u)) | u \in \text{OUT}(v, G)) \cup \{((v, u), (u, v)) | u \in \text{IN}(v, G)\}$$

and the filtering function $id$ that is the identity over $N$, i.e., $id(x) = x$, which is monotonic. It is clear that $\text{Degree}(v, G) = id(|D(v, G)|)$, and the claim follows.

**Cross-Clique.** The cross-clique centrality measure [5] was introduced in the context of analyzing virus and information propagation. It is defined as the amount of cliques containing a certain vertex. Given a digraph $G$ and a vertex $v \in V(G)$, we define the set $\text{Cl}(v, G) = \{ S \in A(v, G) | S \text{ is a clique} \}$. The cross-clique centrality of $v$ in $G$ is

$$\text{Cross-Clique}(v, G) = |\text{Cl}(v, G)|.$$  

It is straightforward to see that the following holds:

**Proposition G.13.** Cross-Clique is a subgraph motif measure for directed graphs.

**Proof.** We consider the subgraph family $K(v, G) = \{ S \in A(v, G) | S \text{ is a clique} \}$ and the identity filtering function $id$. Clearly, $\text{Cross-Clique}(v, G) = id(|K(v, G)|)$, for every pair $(v, G)$, and the claim follows.
Closeness. For a digraph $G$ and vertices $u,v \in V(G)$, the distance from $u$ to $v$ in $G$ is $d_G(u,v) = \min\{ |\pi_{uv}| \mid \pi_{uv} \text{ is a path from } u \text{ to } v \}$. Then, we define the sum of distances for $v$ in $G$ as $SD(v,G) = \sum_{u \in K_G(G) \setminus \{v\}} d_G(u,v)$. Finally, we define closeness centrality as

$$Closeness(u,G) = \frac{1}{SD(v,G)}.$$ 

By definition, the Closeness centrality for a pair $(v,G)$ such that $K_G(G) = 1$ is 0.

**Proposition G.14.** Closeness is not a subgraph motif measure (relative to the induced ranking) for directed graphs.

**Proof.** As we did for Proposition 4.3, we will show that Closeness does not satisfy the bounded value property nor the non-uniform bounded value property. For the former, we consider a set $S \subseteq VG$ such that, with $A_n = \{\text{Closeness}(v,G) \mid (v,G) \in S \text{ and } |A(v,G)| \leq n\}$, it holds that $|A_n| > n + 1$. This is enough to prove that $|Val^p(\text{Closeness})| > n + 1$ since $A_n \subseteq Val^p$. Consider the set $\Sigma_n = (SD(v,G) \mid (v,G) \in S \text{ and } |A(v,G)| \leq n)$. Observe that $|Val^p(\text{Closeness})| \leq n + 1$ if $\Sigma_n \leq n + 1$ because Closeness$(v,G) = \frac{1}{SD(v,G)}$ for every $(v,G) \in VG$. Therefore, it suffices to show that $|\Sigma_n| > n + 1$. The rest of the proof proceed in two main steps:

1. For every $j > 0$, it is easy to see that for the extreme of the line graph $SD(1,L_j) = \frac{j \cdot j - 1}{2}$ while $|A(1,L_j)| = j$. Therefore, it is clear that for $i, j \in |n|$ such that $i \neq j$, $SD(1,L_i) \neq SD(1,L_j)$. Therefore, $|\Sigma_n| > n$ if we add the isolated vertex and the pairs $\{(1,L_j) \mid j \in \{1, \ldots, n\}\}$ to our set $S$.

2. It remains to show that there exists $n > 0$ for which $|\Sigma_n| > n + 1$. Therefore, we just need to find one pair $(o',G')$ such that $|A(o',G')| \leq n$ and $SD(o',G') \notin (SD(1,L_i) \mid i \in \{1, \ldots, n\})$. This is the case for $n = 5$. Observe that for the center of the star graph with 3 nodes $(1, 5, 3)$ we have that $A(1, 5, 3) = 4$, whereas $SD(2, 3) = 2$. Consider now the digraph $H = \{(1, 2, 3, 4), (1, 2, 3, 4)\}$. For the leaf node 1 in $H$ we have $|A(1, H)| = 5$, whereas $SD(1, H) = 5$. At the same time, there is no $i \in \{1, \ldots, 5\}$ such that $SD(1, L_i) \in \{2, 5\}$. This means for $S = \{(1, L_1), (1, L_2), (1, L_3), (1, L_4), (2, 3, 3, 4), (2, 3, 3, 4)\}$, $|Val^p(\text{Closeness})| = |\Sigma_n| \geq |SD(v,G) \mid (v,G) \in S| > 5$.

To show that Closeness does not satisfy the non-uniform bounded value property we exploit the fact that it is local and closed under isomorphism. Gather all the pairs of $S$ inside one disconnected graph, where each connected component is isomorphic to one of the graphs used for the previous proof. In such a graph, $|Val^p_2| = 6$, which leads to a contradiction due to Proposition G.6. \qed

Harmonic. The harmonic centrality for directed graphs is defined as follows:

$$\text{Harmonic}(v,G) = \begin{cases} 0 & \text{if } |K_G(G)| = 1 \\ \sum_{u \in V(G) \setminus \{v\}} \frac{1}{d_G(v,u)} & \text{otherwise.} \end{cases}$$

Notice that whenever $u \notin K_G(G)$, $d_G(v,u) = \infty$, and thus, we assume that $\frac{1}{\sum_{u \notin K_G(G)}} = 0$.

**Proposition G.15.** The following hold:

1. Harmonic is not a subgraph motif measure for directed graphs.
2. Harmonic is not a subgraph motif measure relative to the induced ranking for directed graphs.

**Proof.** Item (1). We follow the same strategy as in the proof of Proposition F.3. We first observe that $\text{Harmonic}(1, L_i) = \sum_{j=1}^{i-1} \frac{1}{j}$. Therefore, for $i \geq 1$, we have that $\text{Harmonic}(1, L_{i+1}) = \text{Harmonic}(1, L_i) + \frac{1}{i}$. This means that $\text{Harmonic}(1, L_i) \neq \text{Harmonic}(1, L_j)$ for every $i, j$ with $i \neq j$. Hence, $|Val^p(\text{Harmonic})| \geq n - 1$. Since $\text{Harmonic}(1, L_i) = 0$, we get that $|Val^p(\text{Harmonic})| \geq n$. Now, let $n = 6$. Consider the following pairs: $(2, H_1)$ and $(2, H_2)$ where $H_1 = \{(1, 2, 3), (2, 1, 2, 3)\}$ and $H_2 = \{(1, 2, 3, 4), (2, 1, 2, 3, 3, 4)\}$. We first observe that, for $n > 0$,

$$\text{Harmonic}(2, H_1) = 2 > \sum_{i=1}^{n-1} \frac{1}{i} = \text{Harmonic}(1, L_n).$$

Moreover, we have that

$$\text{Harmonic}(2, H_2) = 1 + 1 + \frac{1}{2} = 2.5.$$ 

Therefore, $|Val^p(\text{Harmonic})| \geq 7$, which means Harmonic violates the bounded value property for $n = 6$. Thus, by Theorem G.4, we conclude that Harmonic is not a subgraph motif measure for directed graphs. This hold even for trees since every graph used above is a tree.

Item (2). By Proposition G.6, it suffices to show that Harmonic does not enjoy the non-uniform bounded value property. To this end, we follow the same idea as in the proof of Proposition 5.10. Roughly speaking, we gather all the graphs employed above to prove that
It can be verified that \( \text{Val}^6 \) (Harmonic) \( \geq 7 \) in one disconnected graph \( G \). Then, by exploiting the fact that Harmonic is closed under isomorphism (see Definition A.1) and local (see Definition A.2), we can show that \( \text{Val}^6_G \) (Harmonic) \( \geq 7 \).

**PageRank.** Given a graph \( G \) with \( V(G) = \{v_1, \ldots, v_{|V(G)|}\} \), we define the adjacency matrix of \( G \) as the matrix \( A_G \) such that \( A_{Gij} = 1 \) if \( (v_i, v_j) \in E(G) \). We then define \( P \) as the column-stochastic matrix such that:

\[
P_{ij} = \frac{A_{Gij}}{|NG(i)|}
\]

In other words, \( P \) contains the probabilities of jumping from node \( i \) to node \( j \) during a random walk over \( G \). Let \( v \) be a stochastic vector (\( 1^T v = 1 \)), and let \( 0 < \alpha < 1 \) be a teleportation parameter. Then, the \((\nu, \alpha)\)-PageRank vector is defined as the vector \( x_{\nu, \alpha} \) that solves the equation \( (I - \alpha P)x = (1 - \alpha)v \). Now, given a graph \( G \) with \( V(G) = \{v_1, \ldots, v_{|V(G)|}\} \), the PageRank centrality of \( v_i \) in \( G \) is defined as

\[
\text{PageRank}(v_i, G) = \nu_{0.8, 1} \cdot [1],
\]

where \( \nu = 1 \cdot \frac{1}{|V(G)|} \). Several variations of PageRank have been proposed in the literature. In particular, if we do not force \( x_{0.8, u} \) to be a probability distribution, we can show that the sum of centrality values inside a connected component is equal to the number of nodes in that component. Therefore, this give rise to the local version of PageRank, denoted LPR. We call it local since the values in a connected component are always the same regardless of the graph.

**Proposition G.16.** The following hold:

1. PageRank is not a subgraph motif measure for directed graphs.
2. PageRank is not a subgraph motif measure relative to induced ranking for directed graphs.

**Proof.** For both items, it suffices to consider the local version of PageRank since \( \text{Rank(PageRank)} = \text{Rank(LPR)} \).

**Item (1).** It can be verified that \( \text{Val}^2 \) (LPR) = \{1\}, \( \text{Val}^4 \) (LPR) = \{1, 1.80\}, \( \text{Val}^6 \) (LPR) = \{1, 1.80, 2.44\}, \( \text{Val}^8 \) (LPR) = \{1, 1.80, 2.44, 2.95\}, \( \text{Val}^{10} \) (LPR) = \{1, 1.80, 2.44, 2.95, 3.36\} and \( \text{Val}^{12} \) (LPR) = \{1, 2.60\}. Moreover, for the digraph \( H = (\{1, 2, 3, 4\}, (\{2, 1\}, \{2, 3\}, \{2, 4\})) \), we have that \( \text{Val}^6_H \) (LPR) = \{1.0, 1.27\}. Therefore,

\[
\{1, 1.80, 2.44, 2.95, 3.36, 2.60, 1.27\} \subseteq \text{Val}^6(LPR)
\]

Thus, \( |\text{Val}^6(LPR)| \geq 7 \), which implies that LPR does not enjoy the bounded value property. Hence, by Theorem G.4, we conclude that LPR is not a subgraph motif measure. This holds even if we focus on trees since every graph considered is a tree.

**Item (2).** By Proposition G.6, it suffices to show that LPR does not enjoy the non-uniform bounded value property. To this end, we follow the same idea as in the proof of Proposition 5.10. Roughly speaking, we gather all the graphs employed above to prove that \( |\text{Val}^6(LPR)| > 7 \) in one disconnected graph \( G \). Then, by exploiting the fact that LPR is closed under isomorphism (see Definition G.1) and local (see Definition G.2), we can show that \( |\text{Val}^6_G(LPR)| \geq 7 \), and the claim follows.

**Eigenvector.** Given a graph \( G \) with \( V(G) = \{u_1, \ldots, u_{|V(G)|}\} \), we define the adjacency matrix of \( G \) as the matrix \( A_G \) such that \( A_{Gij} = 1 \) if \( (u_i, u_j) \in E(G) \). Let \( \lambda_{\max} \) be the greatest eigenvalue of \( A_G \). We define the unique \( L^2 \)-normalized eigenvector associated with \( \lambda_{\max}(G) \) as \( \nu_{\max} \). The eigenvector centrality \([2]\) of \( u_i \) in \( G \) is the \( i \)-th entry of \( \nu_{\max} \):

\[
\text{Eigenvector}(u_i, G) = \nu_i^{\max}.
\]

**Proposition G.17.** Eigenvector is not a subgraph motif measure for directed graphs.

**Proof.** We first prove the following: given a disconnected graph \( G \) with \( n = |\text{Comp}(G)| \), for each component \( S \) of the set \( M = \{S \in \text{Comp}(G) \mid \lambda_{\max}(G[V(S)]) = \lambda_{\max}(G)\} \), we have that \( |\{\text{Eigenvector}(u, G) \mid v \in S\}| = |\text{Val}^6_G(V(S))| \) (Eigenvector). In other words, each component with the same maximum eigenvalue as \( G \) will have the same amount of values as the component seen as a connected graph. On the other hand, we know that Eigenvector must be a unitary vector. Therefore, if \( G \) has multiple connected components with the maximum eigenvalue \( \lambda_{\max} \), in order to give values to every vertex in those components, the vector is forced to give different values than those given to the component seen as a connected graph.

Given a graph \( G \), we can order its vertices \( V(G) = \{v_1, \ldots, v_n\} \) in such a way that, if \( \text{Comp}(G) = \{S_1, \ldots, S_k\} \), then, for every \( m \in \{1, \ldots, k\} \), there exists \( |m| = \{m_1, m_2 + 1, m_2 + 2, \ldots, m_n + 1\} \subseteq \{1, \ldots, n\} \) such that \( \{v_1, \ldots, v_{|m|}\} = S_m \). Therefore, the rows and columns of the adjacency matrix of \( G, A_G \), are ordered such that there are square sections in the diagonal in \( A_G \) where the \( i \)-th square section is the adjacency matrix of \( G[V(S_i)] \), \( A_G[V(S_i)] \). Out of these square sections the matrix \( A_G \) just have values 0 since vertices in different components have no connections between them. Now, for each \( S_m \in M \), we know that \( b^\max_m \) is the eigenvector associated with \( \lambda_{\max} \) in \( G[V(S_m)] \), then the vector \( v^m_{\max} = (0, 0, \ldots, b^\max_m(1), \ldots, b^\max_m(|S_m|), 0, 0, \ldots) \) is an eigenvector with eigenvalue \( \lambda_{\max} \) for \( A_G \) such that \( v^m_{\max}(i + j) = b^\max_m(j + 1) \) for \( i \in \{0, \ldots, |S_m| - 1\} \). This implies that any linear combination of the vectors \( v^m_{\max} \) will be an eigenvector for \( A_G \) with eigenvalue \( \lambda_{\max} \). Finally, we have that the unique vector \( v^\ast = c_1 v^1_{\max} + c_2 v^2_{\max} + \ldots + c_m v^m_{\max} \) such that \( \|v^\ast\| = 1 \) is the eigenvector for Eigenvector centrality in \( G \). In other words, for each node \( v_i \in V(G) \), then Eigenvector\((v_i, G) = v^\ast(i). \) As we showed, the vector \( v^\ast \) assign values for
every connected component \( S_m \in M \) which are proportional to \( b^\text{max}_m \). Therefore, we can conclude that for every \( n \), \( \|\text{Val}^4_{G}(\text{Eigenvector})\| \geq 1 \cup \bigcup_{S \in G} \text{Val}^4_{G[V(S)]}(\text{Eigenvector}) \). Now, we will construct a sequence of graphs \( G_i \) in order to prove that \( \|\text{Val}^4_{G}(\text{Eigenvector})\| \geq 5 \). For \( i > 1 \), define the digraph \( G_i = \{(0, 1, \ldots, i), \{(i, i + 1), (i + 1, i) \mid i \mod 2 = 0\} \). For \( i \geq 2 \), we have that \( G_i \) contains \([i/2] \) connected components. Each one is isomorphic to cycle with two nodes \( S_2 \) whenever \( i \mod 2 = 0 \). Since the cycle with two nodes has maximum eigenvalue \( \lambda_{\text{max}} = 1 \) the eigenvector \( \sigma^* = (1/\sqrt{2}, 1/\sqrt{4}, \ldots, 1/\sqrt{4}) \) is the eigenvector centrality vector for \( G_i \). This fact implies that \( \|\text{Val}^4_{G}(\text{Eigenvector})\| \geq 1 \cup \bigcup_{j=1}^{i} \text{Val}^4_{G_{j}}(\text{Eigenvector}) = 1 \) for every \( i > 0 \). In other words, the set \( \text{Val}^4_{G}(\text{Eigenvector}) \) has infinitely many values which contradicts the bounded value property. Thus, by Theorem G.4, we conclude that Eigenvector is not a subgraph motif measure for directed graphs. \( \square \)

By Theorem G.9, Eigenvector is a subgraph motif measure relative to the induced ranking over connected graphs. We proceed to strengthen this result by showing that it holds over a broader class of graphs than the class of connected graphs. Let

\[
K_1 = \{ G \in G \mid \text{ for every } G_1, G_2 \in \text{Comp}(G), \lambda_{\text{max}}(G[V(G_1)]) = \lambda_{\text{max}}(G[V(G_2)]) = \lambda_{\text{max}}(G) \text{ implies } G_1 \cong G_2 \}.
\]

It is easy to verify that every connected graph belongs to \( K_1 \), and that \( K_1 \) may contain disconnected components. We show that:

**Proposition G.18.** Eigenvector is a subgraph motif measure relative to the induced ranking when restricted to \( K_1 \).

**Proof.** From the proof of Proposition G.17, we know that for every component of \( M(G) = \{ S \in \text{Comp}(G) \mid \lambda_{\text{max}}(G[V(S)]) = \lambda_{\text{max}}(G) \} \), the eigenvector centrality assigns the same amount of values to the connected component as it seen as a connected graph. On the other hand, if \( S \notin M(G) \), then, for every vertex \( v \in V(S) \), we have that Eigenvector\( (v, G) = 0 \). Note that, whenever two graphs are isomorphic they have the same values in every eigenvector. Values are just permuted depending on the isomorphism you map each graph to the other. Therefore, we have that, for every graph \( G \in K_1 \), there exists \( S \in M(G) \) such that \( \text{Val}^4_{G}(\text{Eigenvector}) = \text{Val}^4_{G[V(S)]}(\text{Eigenvector}) \). In other words, when Eigenvector is restricted to \( K_1 \) it behaves like ConEigenvector. Hence, we can use the same subgraph family and filtering function employed in the proof of Theorem G.9 with the addition of value 0 for vertices outside \( S \) in \( M(G) \).

**Betweenness.** Betweenness centrality for \( v \) in \( G \) is defined as

\[
\text{Betweenness}(v, G) = \sum_{(u, w) \in (V(G)/\{v\})^2} \frac{|S^0_{G}(u, w)|}{|S_{G}(u, w)|}.
\]

If for some pair \((u, w)\) the set \(|S^0_{G}(u, w)| = 0\), then \( \frac{|S^0_{G}(u, w)|}{|S_{G}(u, w)|} = 0 \) by definition.

**Proposition G.19.** Betweenness is a subgraph motif measure for directed graphs when restricted to trees.

**Proof.** Proof is analogous to the one of Proposition F.7. For any tree \( T \), and any pair of nodes \( u \neq v \) in \( V(T) \), \( |\text{EF}(u, v)| \leq 1 \). Therefore, we can use the family of subgraphs defined in Proposition G.10 and the identity filtering function over \( \mathbb{N} \) to prove that Betweenness is a subgraph motif measure when restricted to trees. \( \square \)

**Proposition G.20.** The following hold:

1. Betweenness is not a monotonic subgraph motif measure for directed graphs, even if we focus on connected graphs.
2. Betweenness is not a monotonic subgraph motif measure relative to the induced ranking for directed graphs.

**Proof.** Item (1). We are going to show that Betweenness does not enjoy the monotonic bounded value property, and thus, by Theorem G.7, it is not a monotonic subgraph motif measure. For \( n \geq 1 \), consider the connected graph \( H_n = (V_n, E_n) \), where

\[
V_n = \{1, 2, \ldots, n + 2\} \quad \text{and} \quad E_n = \{(1, i) \mid i \in \{3, 4, \ldots, n + 2\}\} \cup \{(2, i) \mid i \in \{3, 4, \ldots, n + 2\}\}.
\]

It can be verified that \( \text{Betweenness}(3, H_n) = \frac{2}{n} \). However, for \( n > 1 \), \( |A(3, H_n)| > 4 \). Therefore, we have that

\[
|\text{Val}^4_{4}(\text{Betweenness})| + |B\text{Val}^4_{4}(\text{Betweenness})| > 5,
\]

which implies that Betweenness does not enjoy the monotonic bounded value property, as needed.

Item(2). By Theorem G.8, it suffices to show that Betweenness does not enjoy the non-uniform bounded value property. Since Betweenness is closed under isomorphism and local, we will use the same strategy as we did in analogous proofs. In particular, we define the graph \( G \) as the minimal graph with the following property: for every graph \( G' \in \text{Comp}(G) \) such that \( G' \) is isomorphic to \( G \). We can then show that

\[
|\text{Val}^4_{G}(\text{Betweenness})| + |B\text{Val}^4_{G}(\text{Betweenness})| > 5,
\]

which implies that Betweenness does not enjoy the non-uniform monotonic bounded value property, and the claim follows. \( \square \)