

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/112625>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

The Quantum Theil Index: Characterizing Graph Centralization using von Neumann Entropy

DAVID E. SIMMONS*,

JUSTIN P. COON

Department of Engineering Science

University of Oxford

Oxford, UK, OX1 3PJ

* Corresponding author: david.simmons@eng.ox.ac.uk
justin.coon@eng.ox.ac.uk

AND

ANIMESH DATTA

Department of Physics

University of Warwick

Coventry, UK, CV4 7AL

animesh.datta@warwick.ac.uk

February 6, 2018

Abstract

We show that the von Neumann entropy (from herein referred to as the von Neumann index) of a graph's trace normalized combinatorial Laplacian provides structural information about the level of centralization across a graph. This is done by considering the Theil index, which is an established statistical measure used to determine levels of inequality across a system of 'agents', e.g., income levels across a population. Here, we establish a Theil index for graphs, which provides us with a macroscopic measure of graph centralization. Concretely, we show that the von Neumann index can be used to bound the graph's Theil index, and thus we provide a direct characterization of graph centralization via the von Neumann index. Because of the algebraic similarities between the bound and the Theil index, we call the bound the von Neumann Theil index. We elucidate our ideas by providing examples and a discussion of different $n = 7$ vertex graphs. We also discuss how the von Neumann Theil index provides a more comprehensive measure of centralization when compared to traditional centralization measures, and when compared to the graph's classical Theil index. This is because it more accurately accounts for macro-structural changes that occur from micro-structural changes in the graph (e.g., the removal of a vertex). Finally, we provide future direction, showing that the von Neumann Theil index can be generalized by considering the Rényi entropy. We then show that this generalization can be used to bound the negative logarithm of the graph's Jain fairness index. Graph; structure; von Neumann; entropy; complexity; centralization.

2000 Math Subject Classification: 34K30, 35K57, 35Q80, 92D25

1 Introduction

Recently, there have been many attempts to increase our understanding of the structure of graphical models by studying entropic quantities relating to the graph's underlying constructs. Many approaches to studying graph entropy have been proposed, and to the best of the authors' knowledge, it is not clear if/how many of these entropic

quantities relate to each other. For example, in [1] an information function was defined on each of the vertices of a graph in order to associate an entropic quantity with the network; in [2], the entropic properties of graph ensembles have been studied; while in [3, 4] Shannon entropy was used to study the topological uncertainty for networks embedded within a spatial domain. There have also been studies of the von Neumann entropy¹ [7] (from herein referred to as the von Neumann index) of graphs [2, 8, 9, 10, 11, 6, 12, 13, 14, 15, 16]. This work focuses on such a study.

1.1 Graphs and the von Neumann Index

The mathematical theory of quantum mechanics allows us to view quantum states of finite-dimensional systems as Hermitian, positive semi-definite matrices with unit trace [7]. It is also known that the combinatorial Laplacian of a graph is a symmetric positive semi-definite matrix. Thus, normalizing this matrix by its trace allows us to associate a quantum state with a graph, [8]. A natural next step is to study how the structure of the graph is related to the von Neumann index of the graph's trace normalized Laplacian, [8]. This idea was introduced in [8]. To the best of the authors' knowledge, a precise relation is yet to be established, [17], and the goal of this work is to shed further light on this topic.

Let us discuss related work in this area. In [9, 17] it was noted that the von Neumann index may be interpreted as a measure of network regularity. In [10], it was then shown that for scale free networks the von Neumann index of a graph is linearly related to the Shannon entropy of the graph's ensemble. Correlations between these entropies were observed when the graph's degree distribution displayed heterogeneity in [2]. In [6], it was shown that the von Neumann index of an n vertex graph is almost always minimized by the star graph. In that work, the star graph was also conjectured to minimize the entropy over connected graphs. Other studies of the von Neumann entropy of Laplacians have been performed in [11, 12, 13, 14, 15, 18]. Also of interest is [16], in which it is shown that the quantum relative entropy of a graph's Laplacian with another particular quantum state can be related to the number of spanning trees of the graph.

1.2 Graphs, Centrality, and Centralization

Centrality in the context of graphical models was initially investigated in the pioneering work of [19]. It is concerned with understanding and measuring the *importance* or *centrality of particular* vertices within a graph. It is therefore related to the microscopic properties of the graph. The method that is used to define the importance of a vertex is, of course, dependent upon the nature of the underlying system that the graph represents. Because of this, there are many different measures of vertex centrality that have been proposed in the literature [20], e.g., authority centrality, betweenness centrality, closeness centrality, degree centrality, eigenvector centrality, flow betweenness, hub centrality, and Katz centrality. In [21], Laplacian centrality was presented. Laplacian centrality is a vertex specific measure that is obtained by calculating the normalized Laplacian energy of the graph before and after a vertex is removed, and computing the difference between these two values.

Related to centrality is the concept of graph centralization. Graph centralization is concerned with understanding the overall structure of the graph, and determining whether it possesses a high or low level of centralization [22]. It is therefore related to the macroscopic properties of the graph. As with centrality, there are different ways that graph centralization can be measured. Freeman [22] showed that any measure of vertex centrality can be extended to a measure of graph centralization. He constructed three centralization measures, each taking values between 0 and 1. Using his measures, he demonstrated that the star graph obtains the maximum possible value, while the complete graph achieves the minimum possible value. A more in depth discussion of graph centralization can be found in [23].

In this work, we focus on a novel measure of graph centralization. This measure is not explicitly derived from a vertex centrality measure as per Freeman's approach, [22]. Instead, it is related to the Theil index [24] (introduced

¹If we define the von Neumann entropy $H(G)$ of a graph to be the von Neumann entropy of its trace normalized Laplacian, then - for unions of disjoint graphs we will not have the sub-additivity property that is typically required from entropic quantity be satisfied: i.e., the von Neumann entropy of the union may be larger than the the sum of the von Neumann entropies of its disjoint parts. This point was discussed in [5] and can also be seen by observing Proposition 1.5 of [6] and considering the disjoint union of two identical graphs, each with von Neumann entropy greater than one. Consequently, from herein we refer to the graph's von Neumann entropy as its von Neumann index.

below) and the graph’s von Neumann index. To the best of the authors’ knowledge, this is the first time such a measure has been proposed.

1.3 The Theil Index

The Theil index T has been a topic of interest since its inception in 1972 [24]. The Theil index [24, 25] is an established observable that has been used to measure, e.g., inequality within economies, as well as racial segregation within social environments. Formally, the Theil index is defined on a population of n agents, each with some characteristic $x_i \geq 0$, $i \in \{1, \dots, n\}$. The Theil index is then given by [25]

$$T := \frac{1}{n} \sum_{i=1}^n \frac{x_i}{\mu} \log \left(\frac{x_i}{\mu} \right), \quad (1)$$

where μ is defined to be

$$\mu := \frac{1}{n} \sum_{i=1}^n x_i. \quad (2)$$

The Theil index can be used to characterize the uniformity of the characteristics among the agents, and is bounded according to

$$0 \leq T \leq \log n.$$

If $T = 0$ then the characteristics are uniform among the agents; if $T = \log n$ then $x_i \neq 0$ for only one i . Some examples of characteristics from the literature are income [26] or total wetland [27].

Through inspection, it is easy to see that the Theil index is closely related to the Shannon entropy. Specifically, $x_i/(\mu n)$ defines a probability distribution on the agents (because $\sum_i x_i/(\mu n) = 1$ and $x_i/(\mu n) \geq 0$), so that we can define the Shannon entropy [28]

$$H(X) := - \sum_{i=1}^n \frac{x_i}{\mu n} \log \left(\frac{x_i}{\mu n} \right), \quad (3)$$

where X is the random variable with distribution defined by $x_i/(\mu n)$. The Theil index can then be written as

$$T = \log n - H. \quad (4)$$

Alternatively, we may write

$$T = S(X \| U_n), \quad (5)$$

where $S(A \| B)$ denotes the relative entropy [29] between two random variables A and B , and X is the random variable described immediately above while U_n is the discrete uniform random variable on n points. Equation (5) provides an information theoretic interpretation of the Theil index as the average number of extra bits of information required to encode a source distributed according to X given that we assume it to be uniformly distributed (the logarithm being base two here). This also adds to our intuition of the Theil index as a measure of uniformity/inequality.

1.4 Contributions

In this work, we provide an interpretation of the von Neumann index of the graph’s Laplacian as a measure of the graph’s centralization². This is done by defining a Theil index for the graph in terms of the degrees of the graph’s vertices (equation (23) below), which we can use to determine the level of centralization across the vertices of the graph. Our main result is to show that the graph’s von Neumann index can be used to bound the graph’s Theil index (Theorem 1 below). The bound is a particular instance of the quantum relative entropy. Because of the relationship between relative entropy and the Theil index (5), this motivates us to refer to the bound as the graph’s

²Our work differs from [21], in which centrality was established from properties of the graph’s Laplacian. First, [21] proposed a centrality measure, not a graph centralization measure (for a discussion of centrality and centralization, see section 1.2). Second, that work did not consider the von Neumann entropy when establishing their centrality measure.

von Neumann Theil index (Definition 2 below). A corollary to our main result (Corollary 3) also determines a relationship between graph von Neumann entropy and the graph entropy considered in [1, 30]. To elucidate our ideas, we provide examples and a discussion of different $n = 7$ vertex graphs in section 5. In these examples, we also consider Freeman’s centralization measures [22].

2 Preliminary Material

The role of this section is to present background material on spectral graph theory and density operators, and to present the concept of Theil indices for graphs in terms of the graph’s entropy [1].

2.1 Spectral graph theory and density operators

Consider an undirected graph $G = (V, E)$, which consists of a vertex set

$$V = \{v_1, \dots, v_n\}, \tag{6}$$

with $|V| = n$, and an edge set

$$E = \{e_1, \dots, e_m\}, \tag{7}$$

with $|E| = m$. Here, we assume that the element of the edge set (say the k th element) connecting vertex pair (i, j) can be written as

$$e_k = v_{i,j}. \tag{8}$$

The degree matrix of the graph is given by

$$\Delta(G) = \text{diag} \{d_1, \dots, d_n\}, \tag{9}$$

where d_i is the degree of v_i . The (i, j) th element of a graph’s adjacency matrix is given by

$$[A(G)]_{i,j} = 1 \tag{10}$$

if vertex pair (i, j) is connected by an edge. The Laplacian of the graph is then defined to be [31]

$$L(G) = \Delta(G) - A(G). \tag{11}$$

A graph’s Laplacian is useful when studying the connectedness properties of the graph. For example, the number of zero eigenvalues of the Laplacian’s spectrum is known to categorize the number of disconnected regions within a graph [31]. Also, Kirchoff’s theorem tells us that the (absolute value of the) determinant of any first order minor associated with a graph’s Laplacian is the number of spanning trees of the graph [32].

More, recently, graph Laplacians have been studied using tools from quantum information theory. Specifically, we can view quantum states as positive semi-definite matrices with unit trace. It is well known that the Laplacian of a graph is a Hermitian positive semi-definite matrix. Thus, after normalizing this matrix by its trace we obtain a new matrix that has all the properties of a density matrix (e.g., a quantum state):

$$\rho_G = \frac{L(G)}{\text{Tr}\{L(G)\}}. \tag{12}$$

The idea of measuring the von Neumann index of (12) was originally posed in [8]:

$$H(G) = H(\rho_G) := -\text{Tr} \{\rho_G \log \rho_G\}. \tag{13}$$

Note, throughout the paper, we use $H(G)$ and $H(\rho_G)$ interchangeably. With (13), we can identify the quantum relative entropy between ρ_G and I/n , which is given by

$$D(\rho_G \| I_n) := \text{Tr} \{\rho_G (\log \rho_G - \log I/n)\} = \log n - H(G). \tag{14}$$

The main result of this work will be to show that $D(\rho_G \| I_n)$ bounds a centralization measure (the graph’s Theil index, discussed below) on the network. Because of the relationship between relative entropy and the Theil index (5), our main result motivates us to refer to the bound as the graph’s von Neumann Theil index (Definition 2 below).

2.2 The Theil Index of a Graph

As was discussed in the introduction, the Theil index T assigns a value to a set of characteristics $\{x_1, \dots, x_n\}$ (see (1) and (2)). In light of this, here we assign a Theil index to a graph by extracting characteristics from the graph's vertices. Specifically, we define a functional

$$f : V \mapsto \mathbb{R}^+, \quad (15)$$

on the vertices of an n vertex graph, G . If we then let the n characteristics be

$$x_{G,i} := f(v_i), \quad i \in \{1, \dots, n\},$$

we obtain the Theil index of the graph G

$$T_f(G) := \frac{1}{n} \sum_{i=1}^n \frac{x_{G,i}}{\mu} \log \left(\frac{x_{G,i}}{\mu} \right), \quad (16)$$

where

$$\mu := \frac{1}{n} \sum_{i=1}^n x_{G,i}. \quad (17)$$

It is important to acknowledge that our work does not establish a relationship between this Theil index and the graph's von Neumann index when $x_{G,i}$ is arbitrary. Instead, we establish such a relationship when $x_{G,i}$ takes a specific form determined by the degrees of each of the graph's vertices. It is also important to acknowledge that this is *not* the first time such a functional f has been proposed in the literature. In particular, it was used in [1] to define a concept of graph entropy. Specifically, [1] gave the following definition.

Definition 1. *Let G be a graph on n vertices. Let f be the functional (15), which assigns an arbitrary positive number to each vertex v_i . Then the entropy of the graph with respect to f is defined to be*

$$H_f(G) := - \sum_{i=1}^n \frac{f(v_i)}{\sum_j f(v_j)} \log \frac{f(v_i)}{\sum_j f(v_j)}. \quad (18)$$

With this definition, we can now write the Theil index of the graph as

$$T_f(G) = \log n - H_f(G). \quad (19)$$

One natural example of a functional is the vertex degree power,

$$f(v_i) = d_i^k, \quad (20)$$

$k \in \mathbb{R}$, which was originally considered in [30]. Here, we call the corresponding entropy the *degree entropy*, which is given by

$$H_{d,k}(G) := - \sum_i \tilde{d}_i^k \log \tilde{d}_i^k, \quad (21)$$

where

$$\tilde{d}_i^k := \frac{d_i^k}{\sum_j d_j^k} \quad (22)$$

is the k th power relative degree of vertex i . We can then write the Theil index corresponding to the graph's degree entropy as

$$T_{d,k}(G) := \log n - H_{d,k}(G). \quad (23)$$

2.3 The Theil index and Graph Centralization

Naturally, we can use the Theil index of the graph corresponding to the degree entropy, (23), to measure the level of centralization present across a network. Specifically, because the Theil index is a measure of relative entropy, we can use it to measure how ‘close’ the relative degree distribution (22) is to the uniform distribution: regular graphs have a uniform degree distribution, while star graphs may be perceived as having the ‘least uniform’ degree distribution. It is easy to show that these examples, respectively, induce minimum and maximum values for the Theil index. Intuitively, the complete graph (which is a k -regular graph) may be thought of as being the least centralized graph, while the star graph may be thought of as being the most centralized graph.

3 The Graph Theil Index and von Neumann Entropy

Having developed the Theil index of the graph (19), we now present the following Theorem, which provides a relationship between the graph’s Theil index³ and the quantum relative entropy defined in (14). From section 2.3, this then identifies a relationship between the quantum relative entropy of the graph and its centralization. Because of these intricate relationships, and the relationship that the Theil index has with relative entropy (see (5)), before presenting the theorem we first define the von Neumann Theil index of the graph.

Definition 2. Let ρ_G be given by (12) and I be an equivalently dimensioned identity matrix. The von Neumann Theil index of the graph $T_Q(G)$ is defined to be the quantum relative entropy between ρ_G and I/n , i.e.,

$$T_Q(G) := D(\rho_G \| I/n) := \text{Tr} \{ \rho_G (\log \rho_G - \log I/n) \} = \log n - H(G). \quad (24)$$

Theorem 1. Let G be a connected graph on n vertices, $T_{d,k}(G)$ be the graph’s Theil index corresponding to its degree entropy (23), and $T_Q(G)$ be the graph’s von Neumann Theil index (24). Also, let $M := \{d_i : d_i = \max_j \{d_j\}\}$, so that $|M|$ is the multiplicity of $\max_i \{d_i\}$. Exactly one of the following statements holds:

A: $\log n - \log |M| \geq T_Q$ if and only if there exists k such that

$$T_{d,1}(G) \leq T_Q(G) \leq T_{d,k'}(G) \quad \forall k' \geq k. \quad (25)$$

Furthermore,

$$\text{if } \frac{(\sum_i d_i)^2}{\sum_i d_i + \sum_i d_i^2} \geq |M| \quad \text{then} \quad \log n - \log |M| \geq T_Q. \quad (26)$$

B: $\log n - \log |M| \leq T_Q$ if and only if

$$T_{d,k}(G) \leq T_Q(G) \quad \forall k. \quad (27)$$

Proof. The proof is given in section 4 below. □

3.1 Discussion of Theorem 1

Theorem 1 deserves discussion. First and foremost, this theorem highlights a fundamental relationship between graph von Neumann entropy (and the von Neumann Theil index) and graph centralization. Specifically, it tells us that $T_Q(G)$ (see (14)) *always* upper bounds the centralization measure when $k = 1$. Because $T_{d,k}(G)$ is a non decreasing function of k (Lemma 1), it also tells us that, given $\log n - \log |M| \geq T_Q(G)$, $T_Q(G)$ will (for a specific $k > 1$) equal the centralization measure exactly. Equation (26) provides information about when we can be sure $\log n - \log |M| \geq T_Q$. These ideas are formalized in the following corollary.

Corollary 1. We have

$$T_{d,1}(G) \leq T_Q(G). \quad (28)$$

Furthermore, if $\log n - \log |M| \geq T_Q$ then there exists $k \in \mathbb{R}^+$ such that

$$T_{d,k}(G) = T_Q(G). \quad (29)$$

³Specifically, the Theil index corresponding to its degree entropy (23).

Interestingly, k -regular graphs always satisfy the first statement of Theorem 1. This is because for any graph G , $H(G) \leq \log(n-1)$ [9], while $\log|M| = \log n$ when G is k -regular. We therefore reach a second corollary of the theorem.

Corollary 2. *If G is k -regular then*

$$0 = T_{d,k}(G) \leq T_Q(G) \forall k \geq 1. \quad (30)$$

It is important to note that Theorem 1 could have been stated as per the following corollary. However, doing so would not have elucidated the connection between von Neumann entropy and graph centralization.

Corollary 3. *Consider an identical hypothesis as that of Theorem 1. Then exactly one of the following statements holds:*

A $H(G) \leq \log|M|$ if and only if

$$H_{d,k}(G) \geq H(G) \forall k. \quad (31)$$

B $H(G) \geq \log|M|$ if and only if there exist k such that

$$H_{d,1}(G) \geq H(G) \geq H_{d,k'}(G) \forall k' \geq k. \quad (32)$$

Thus, Corollary 3 provides an alternative relationship between the relative degree entropy and the von Neumann entropy of a graph.

4 Proof of Theorem 1

To prove (25) of statement **A**, we begin by showing that $H_{d,1}(G) \geq H(G)$ (i.e., $T_{d,1}(G) \leq T_Q(G)$). If we select a set of positive operator valued measures $\{\Lambda_i\}$, $i \in \{1, \dots, n\}$, such that $[\Lambda_i]_{k,k} = 1$ and all other entries are zero, with ρ_G given by (12) we find that

$$H_{d,1}(G) = - \sum_i \text{Tr}\{\Lambda_i \rho_G \log(\Lambda_i \rho_G)\}. \quad (33)$$

The first inequality of (25) then follows from [29, Theorem 11.1.1], which shows that for an arbitrary quantum state ρ the von Neumann entropy can be written as

$$H(\rho) = \min_{\{\Lambda_y\}} \left[- \sum_y \text{Tr}\{\Lambda_y \rho \log(\Lambda_y \rho)\} \right], \quad (34)$$

where the minimum is taken over all rank 1 positive operator valued measures. The second inequality of (25) follows immediately from Lemma 1 below.

To prove (26) of statement **A**, we consider the Rényi 2 entropy of G , which is given by [33, Equation (2)]

$$H_2(G) = \log \left(\frac{(\sum_i d_i)^2}{\sum_i d_i + \sum_i d_i^2} \right). \quad (35)$$

We also note that $H_2(G) \leq H(G) \forall G$, [33]. Consequently, the following implications hold:

$$\frac{(\sum_i d_i)^2}{\sum_i d_i + \sum_i d_i^2} \geq |M| \Leftrightarrow \log \frac{(\sum_i d_i)^2}{\sum_i d_i + \sum_i d_i^2} \geq \log |M| \Leftrightarrow H_2(G) \geq \log |M| \Rightarrow H(G) \geq \log |M|. \quad (36)$$

where the second bidirectional implication follows from (35).

Statement **B** follows immediately from Lemma 1 below.

Lemma 1. For all graphs G on n vertices, the Theil degree index of the graph $T_{d,k}$ is a nondecreasing function of k , i.e.,

$$\frac{dT_{d,k}}{dk} \geq 0. \quad (37)$$

Also, with $M = \{d_i : d_i = \max_j \{d_j\}\}$, we have

$$\lim_{k \rightarrow \infty} T_{d,k}(G) = \log n - \log |M|. \quad (38)$$

Proof. We begin by proving (37). The basis of the proof will be in the construction of the bound in (41) and (42) below. To aid our construction of the bound, we note that in general $\frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right)$ can be either positive or negative, and for any particular $k = k'$, we can construct two sets $\mathcal{N}_{k'}$ and $\mathcal{P}_{k'}$ (corresponding to negative and positive elements, respectively), where

$$\mathcal{N}_{k'} = \left\{ i : \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \Big|_{k=k'} < 0 \right\} \quad \text{and} \quad \mathcal{P}_{k'} = \left\{ i : \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \Big|_{k=k'} \geq 0 \right\}. \quad (39)$$

With the above, we are in a position to construct the bound:

$$\begin{aligned} \frac{dH_{d,k}}{dk} \Big|_{k=k'} &= - \sum_i \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \log \frac{d_i^k}{\sum_j d_j^k} \right) \Big|_{k=k'} \\ &= - \sum_i \left[\frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \log \frac{d_i^k}{\sum_j d_j^k} + \frac{d_i^k}{\sum_j d_j^k} \frac{d}{dk} \left(\log \frac{d_i^k}{\sum_j d_j^k} \right) \right] \Big|_{k=k'} \\ &= - \sum_i \left[\frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \log \frac{d_i^k}{\sum_j d_j^k} + \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \right] \Big|_{k=k'} \\ &= - \sum_i \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \Big|_{k=k'} \log \frac{d_i^{k'}}{\sum_j d_j^{k'}} - \frac{d}{dk} \left(\frac{\sum_y d_y^k}{\sum_j d_j^k} \right) \Big|_{k=k'} \\ &= - \sum_i \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \Big|_{k=k'} \log \frac{d_i^{k'}}{\sum_j d_j^{k'}} \\ &= \sum_{i \in \mathcal{N}_{k'}} \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \Big|_{k=k'} \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} + \sum_{t \in \mathcal{P}_{k'}} \frac{d}{dk} \left(\frac{d_t^k}{\sum_j d_j^k} \right) \Big|_{k=k'} \log \frac{\sum_j d_j^{k'}}{d_t^{k'}} \\ &\leq \min_{i \in \mathcal{N}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} \right\} \sum_{i \in \mathcal{N}_{k'}} \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \Big|_{k=k'} + \max_{i \in \mathcal{P}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_t^{k'}} \right\} \sum_{t \in \mathcal{P}_{k'}} \frac{d}{dk} \left(\frac{d_t^k}{\sum_j d_j^k} \right) \Big|_{k=k'}, \quad (41) \end{aligned}$$

where the final inequality follows because the first and second summations are (by the definition of the sets in (39)) negative and positive, respectively. From Lemma 2, we have

$$\max_{i \in \mathcal{P}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} \right\} \leq \min_{i \in \mathcal{N}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} \right\},$$

so that we can extend the final inequality of (41) to

$$\begin{aligned}
\left. \frac{dH_{d,k}}{dk} \right|_{k=k'} &\leq \max_{i \in \mathcal{P}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} \right\} \sum_{i \in \mathcal{N}_{k'}} \left. \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \right|_{k=k'} + \max_{i \in \mathcal{P}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} \right\} \sum_{t \in \mathcal{P}_{k'}} \left. \frac{d}{dk} \left(\frac{d_t^k}{\sum_j d_j^k} \right) \right|_{k=k'} \\
&= \max_{i \in \mathcal{P}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} \right\} \left(\sum_{i \in \mathcal{N}_{k'}} \left. \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \right|_{k'=k} + \sum_{t \in \mathcal{P}_{k'}} \left. \frac{d}{dk} \left(\frac{d_t^k}{\sum_j d_j^k} \right) \right|_{k=k'} \right) \\
&= \max_{i \in \mathcal{P}_{k'}} \left\{ \log \frac{\sum_j d_j^{k'}}{d_i^{k'}} \right\} \left. \frac{d}{dk} \left(\frac{\sum_i d_i^k}{\sum_j d_j^k} \right) \right|_{k=k'} = 0.
\end{aligned} \tag{42}$$

Substituting this into (23) gives the result.

We now prove (38). For arbitrary k , we have

$$H_{d,k}(G) = - \sum_i \frac{d_i^k}{\sum_j d_j^k} \log \frac{d_i^k}{\sum_j d_j^k} \tag{43}$$

$$\rightarrow - \sum_M \frac{1}{|M|} \log \frac{1}{|M|} \tag{44}$$

$$= \log |M|, \tag{45}$$

where the second line is obtained by letting $k \rightarrow \infty$. □

The following lemma is used within the proof of Lemma 1.

Lemma 2. *Let*

$$\mathcal{N}_{k'} = \left\{ i : \left. \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \right|_{k=k'} < 0 \right\} \quad \text{and} \quad \mathcal{P}_{k'} = \left\{ i : \left. \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \right|_{k=k'} \geq 0 \right\}. \tag{46}$$

Then

$$\max_{i \in \mathcal{N}_{k'}} \{d_i\} \leq \min_{i \in \mathcal{P}_{k'}} \{d_i\}. \tag{47}$$

Proof. We have

$$\left. \frac{d}{dk} \left(\frac{d_i^k}{\sum_j d_j^k} \right) \right|_{k=k'} = d_i^{k'} \left(\frac{\left(\sum_j d_j^{k'} \right) \log d_i - \left. \frac{d}{dk} \left(\sum_j d_j^k \right) \right|_{k=k'}}{\left(\sum_j d_j^{k'} \right)^2} \right),$$

which is negative if and only if

$$\left(\sum_j d_j^{k'} \right) \log d_i - \left. \frac{d}{dk} \left(\sum_j d_j^k \right) \right|_{k=k'} < 0, \quad \text{if and only if} \quad d_i < \exp \left(\frac{\left. \frac{d}{dk} \left(\sum_j d_j^k \right) \right|_{k=k'}}{\sum_j d_j^{k'}} \right).$$

Thus, by the definition of the sets $\mathcal{N}_{k'}$ and $\mathcal{P}_{k'}$

$$\max_{i \in \mathcal{N}_{k'}} \{d_i\} < \exp \left(\frac{\left. \frac{d}{dk} \left(\sum_j d_j^k \right) \right|_{k=k'}}{\sum_j d_j^{k'}} \right) \leq \min_{i \in \mathcal{P}_{k'}} \{d_i\}.$$

This proves the lemma. □

5 Examples and Discussion

In [22], it was argued that the centralization of a graph should index the tendency of a single vertex within the graph to be more central than other vertices in the graph. This type of graph centralization measure is typically based on differences between the centrality of the most central point and that of all others [22]. Here, we claim that such an approach gets part way to understanding how centralized a graph is, but (in many scenarios) fails to establish the tendency of the graph's centralization to change *significantly* under *small* structural changes. An example of such a graph might be the circle graph, the vertices of which have equal centrality measures throughout. Thus, with the centralization measure proposed by [22], the circle graph's centralization will be 0. However, the removal of any single vertex will result in a path graph, which canonically possesses significantly more centralization. Interestingly, the Theil index also suffers from this type of problem; however, the von Neumann Theil index does not.

In what follows, we present examples of 10 common graphs with $n = 7$ vertices in FIGs 5.1, 5.2, 5.3, and 5.4. These figures serve to illustrate the above discussion. Each of the figures is a particular ordering of the 10 graphs, where the ordering is determined by:

1. two centralization measures proposed by Freeman [22], based, respectively, on degree centrality (FIG 5.1a to 5.1j)

$$C_D = \frac{\sum_{i=1}^n \max\{d_j\} - d_i}{n^2 - 3n + 2}, \quad (48)$$

and betweenness centrality⁴ (FIG 5.2a to 5.2j)

$$C_B = \frac{\sum_{i=1}^n \max\{g_j\} - g_i}{n^3 - 4n^2 + 5n - 2}; \quad (49)$$

2. the Theil index (23) with $k = 1$ (FIG 5.3a to 5.3j);
3. the von Neumann Theil index (24) (FIG 5.4a to 5.4j).

These orderings provide us with a visualization of a 7 vertex graph as it becomes more decentralized when measured using each of these metrics. As can be seen from these figures, the different measures impose different orderings on the graphs. However, all measures determine the star graph to be the most centralized and the complete graph to be the least centralized. Note, it was shown in [6] that almost all graphs achieve a larger von Neumann entropy (and hence smaller von Neumann Theil index) than the star graph. It was also conjectured that the star graph minimizes the von Neumann entropy among all connected graphs.

Let us discuss our thesis in more detail. As was discussed in sections 1.3 and 2.2, the Theil index depends only on the relative degrees of the vertices, and provides a measure of their skewness. However, it fails to capture richer structural details about the underlying connectedness of the vertices. Similar ideas hold for Freeman's degree centralization measure, and to a lesser extent his betweenness centralization measure. Crucially, they may all be considered as somewhat crude measures of centralization when viewed in the context of graphical models. Our rationale for this stems from our observation that the von Neumann Theil index is able to capture the tendency of a graph's macro-structure to change under small micro-structural changes (e.g., vertex removals). The most obvious example of this is the circle and complete graph (although our arguments can be extended to the other graph's too), which are both k -regular, but the vertices of the circle graph are much more dependent on some vertices than others. This is not so for the complete graph. Arguably then, the circle graph is considerably more centralized than the complete graph. Despite this, both of Freeman's measures, and the Theil index, rank these two graphs as being the most decentralized out of all graphs, whereas the von Neumann Theil index rightly identifies the more centralized nature of the circle graph while categorizing the complete graph as the most decentralized. Similar discussions can be had towards the other graphs.

The von Neumann Theil index is able to capture the centralization properties that we desire from the Theil index (i.e., it bounds the skew measured by the Theil index, which we interpret as a crude measure of centralization, Theorem 1), but it also depends on the underlying connectedness of the graph in a way that the other measures do not. In the quantum sense, the von Neumann Theil index is just the quantum relative entropy between the quantum

⁴In (49), g_i is defined to be the betweenness centrality [20] of vertex i .

Graph ordering according to the degree centralization measure C_D (48)

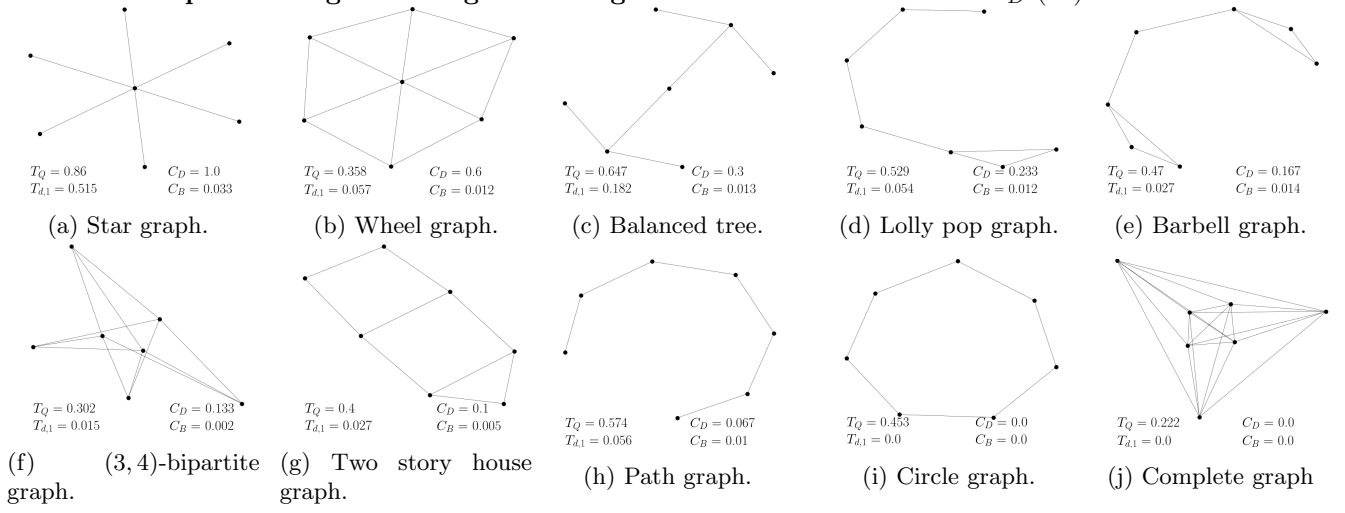


Figure 5.1: $n = 7$ vertex graphs arranged according to decreasing order of the degree centralization measure C_D proposed by Freeman (48).

state corresponding to the graph's Laplacian and a maximally mixed quantum state with density operator I/n . Consequently, it tells us how 'close' the underlying quantum state corresponding to the graph is to the maximally mixed state, i.e., the complete graph. This discussion goes further to explaining our ideas.

6 Future Direction

In the previous sections, we demonstrated that the von Neumann index of a graph could be used to bound the graph's Theil index (Theorem 1). The Theil index was shown to be a canonical (but crude) measure of centralization (sections 2.2 and 2.3). We therefore concluded that the graph's von Neumann index (specifically, its *von Neumann Theil index*) provided key structural information about the level of centralization present across the graph. In this section, we show that the von Neumann Theil index can be generalized by considering the Rényi entropy. We call this the generalized von Neumann Theil index. The generalized von Neumann Theil index is shown to bound the negative logarithm of the graph's Jain fairness index.

6.1 Generalizing the von Neumann Theil index

Let ρ_G be the graph's trace normalized Laplacian (12). Above (Definition 2), the von Neumann Theil index was defined to be

$$T_Q(G) = \log n - H(G), \quad (50)$$

where

$$H(G) = -\text{Tr} \{ \rho_G \log \rho_G \}.$$

Equation (50) can be trivially generalized by considering the Rényi entropy [34] of ρ_G , which is defined to be

$$H^{(p)} := \frac{1}{1-p} \log \text{Tr} \{ \rho_G^p \}.$$

Using L'Hopital's rule [35], we find that

$$H(G) = \lim_{p \rightarrow 1} H^{(p)}. \quad (51)$$

Graph ordering according to the betweenness centrality measure C_B (49)

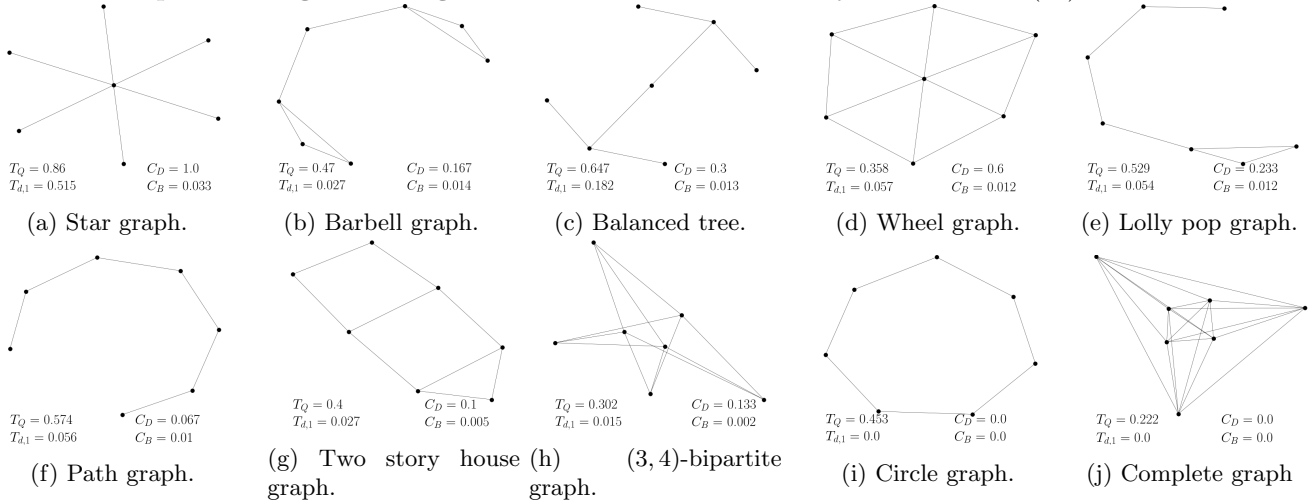


Figure 5.2: $n = 7$ vertex graphs arranged according to decreasing order of the betweenness centralization measure C_B proposed by Freeman (49).

Graph ordering according to the Theil index $T_{d,1}$ (23)

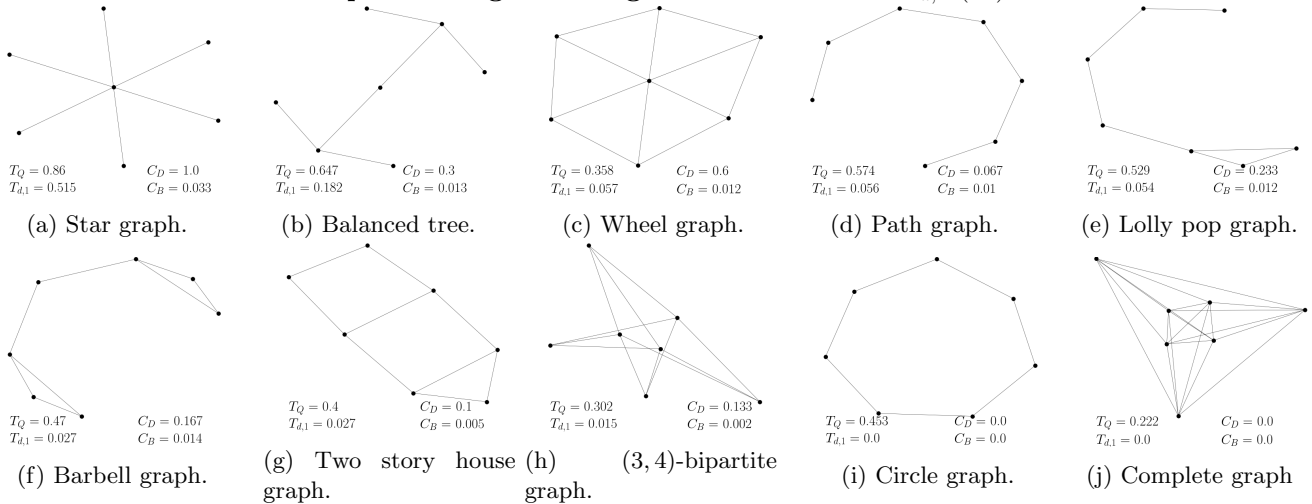


Figure 5.3: $n = 7$ vertex graphs arranged according to decreasing order of the Theil index $T_{d,1}$ (23).

Graph ordering according to the von Neumann Theil index T_Q (24)

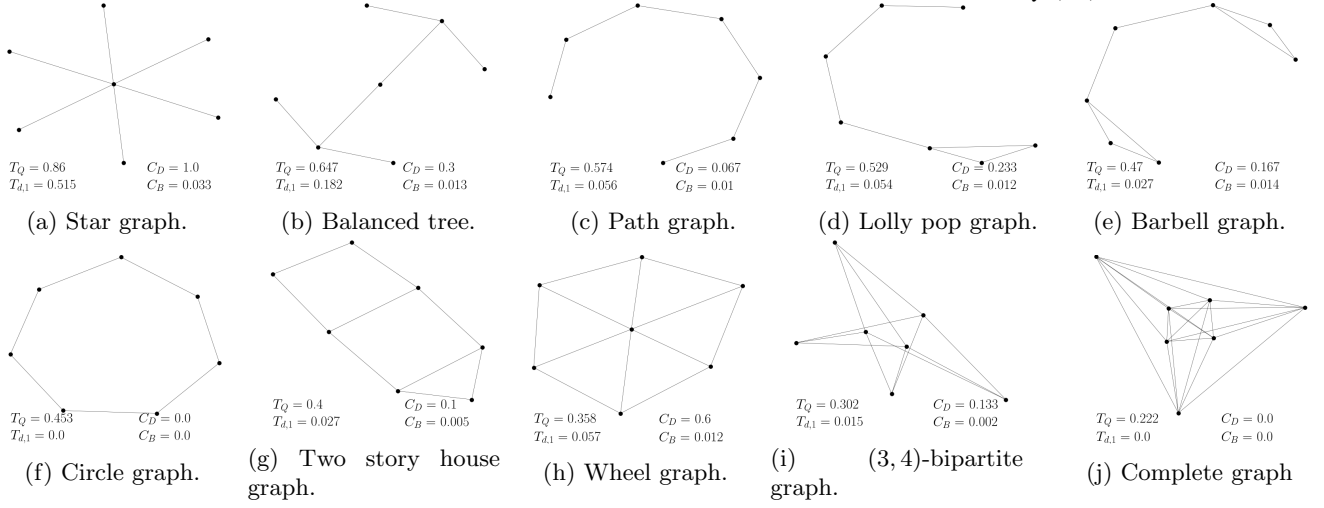


Figure 5.4: $n = 7$ vertex graphs arranged according to decreasing order of the von Neumann Theil index T_Q (24).

We can therefore define the graph's generalized von Neumann Theil index to be

$$T_Q^{(p)} := \log n - H^{(p)} \quad (52)$$

and from (51) we have

$$T_Q(G) = \lim_{p \rightarrow 1} T_Q^{(p)}(G). \quad (53)$$

Note, because $H^{(p)}$ is a monotonically decreasing function of p , the generalized von Neumann Theil index is a monotonically increasing function of p .

6.2 Bounding the Jain Fairness Index

In the following lemma, we show that (52) can be used to bound the negative logarithm of the Jain fairness index [36] (an index used to measure the *fairness* of a distribution of characteristics among a set of nodes). This adds further credit to the idea that von Neumann entropy (and its generalization) characterises graph centralization.

Lemma 3. *Let G be a connected graph on n vertices, where the degree of vertex i is given by d_i . Define the Jain fairness index to be [36]*

$$\mathcal{J}(d_1, \dots, d_n) = \frac{(\sum d_i)^2}{n \sum d_i^2}. \quad (54)$$

Then the following equality holds:

$$T_Q^{(2)} \geq -\log \mathcal{J}(d_1, \dots, d_n).$$

Proof. This follows from

$$\begin{aligned} T_Q^{(2)} &= \log n - \log \left(\frac{(\sum d_i)^2}{\sum d_i + \sum d_i^2} \right) \geq \log n - \log \left(\frac{(\sum d_i)^2}{\sum d_i^2} \right) \\ &= -\log \left(\frac{(\sum d_i)^2}{n \sum d_i^2} \right) \\ &= -\log \mathcal{J}(d_1, \dots, d_n). \end{aligned}$$

The first equality follows from [33, equation (2)] and the last line follows from (54). \square

7 Conclusions

In this work, we define an n point distribution on an n vertex graph (which we call the *relative* degree distribution), where each point of the distribution is proportional to a positive power of the corresponding vertex degree. We then use this distribution to construct a Theil index for the graph. Naturally, this Theil index can be interpreted as a measure of graph centralization. We then show that the von Neumann index of the graph's trace normalized combinatorial Laplacian can be used to bound the graph's Theil index. This is the first time that von Neumann index has been related to the Theil index. We refer to the bound as the von Neumann Theil index. We conclude that the graph's von Neumann index (specifically, the von Neumann Theil index) provides structural information about the level of centralization within the graph. We also demonstrate that the von Neumann Theil index provides a more precise measure of centralization when compared to traditional centralization measures and the graph's classical Theil index. As a final note, we demonstrate that the von Neumann Theil index can be generalized by considering the quantum Rényi- p entropy. We show that with $p = 2$, this generalization bounds the negative logarithm of the graph's Jain fairness index, further demonstrating the relationship between centralization and quantum graph entropy.

Funding

This work was supported by the EPSRC under grant number (EP/K04057X/2) and the UK National Quantum Technologies Programme under grant number (EP/M013243/1).

References

- [1] Matthias Dehmer. Information processing in complex networks: Graph entropy and information functionals. *Applied Mathematics and Computation*, 201(1):82–94, 2008.
- [2] Kartik Anand, Ginestra Bianconi, and Simone Severini. Shannon and von neumann entropy of random networks with heterogeneous expected degree. *Physical Review E*, 83(3):036109, 2011.
- [3] Justin P Coon. Topological uncertainty in wireless networks. In *Global Communications Conference (GLOBECOM), 2016 IEEE*, pages 1–6. IEEE, 2016.
- [4] Arta Cika, Justin P Coon, and Sunwoo Kim. Effects of directivity on wireless network complexity. In *Workshop on Spatial Stochastic Models for Wireless Networks (SpaSWiN),, 2017 IEEE*, pages 1–7. IEEE, 2017.
- [5] Manlio De Domenico, Vincenzo Nicosia, Alexandre Arenas, and Vito Latora. Structural reducibility of multi-layer networks. *Nature communications*, 6:6864, 2015.
- [6] Michael Dairyko, Leslie Hogben, Jephian C-H Lin, Joshua Lockhart, David Roberson, Simone Severini, and Michael Young. Note on von neumann and rényi entropies of a graph. *Linear Algebra and its Applications*, 521:240–253, 2017.
- [7] John Von Neumann. *Mathematical foundations of quantum mechanics*. Number 2. Princeton university press, 1955.
- [8] Samuel L Braunstein, Sibasish Ghosh, and Simone Severini. The laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states. *Annals of Combinatorics*, 10(3):291–317, 2006.
- [9] Filippo Passerini and Simone Severini. The von neumann entropy of networks. *Available at SSRN 1382662*, 2008.
- [10] Kartik Anand and Ginestra Bianconi. Entropy measures for networks: Toward an information theory of complex topologies. *Physical Review E*, 80(4):045102, 2009.

- [11] Niel de Beaudrap, Vittorio Giovannetti, Simone Severini, and Richard Wilson. Interpreting the von neumann entropy of graph laplacians, and coentropic graphs. *A Panorama of Mathematics: Pure and Applied*, 658:227, 2016.
- [12] Abdelilah Belhaj, Adil Belhaj, Larbi Machkouri, Moulay Brahim Sedra, and Soumia Ziti. Weighted graph theory representation of quantum information inspired by lie algebras. *arXiv preprint arXiv:1609.03534*, 2016.
- [13] Supriyo Dutta, Bibhas Adhikari, and Subhashish Banerjee. A graph theoretical approach to states and unitary operations. *Quantum Information Processing*, 15(5):2193–2212, 2016.
- [14] Supriyo Dutta, Bibhas Adhikari, and Subhashish Banerjee. Zero discord quantum states arising from weighted digraphs. *arXiv preprint arXiv:1705.00808*, 2017.
- [15] David E. Simmons, Justin P. Coon, and Animesh Datta. Symmetric laplacians, quantum density matrices and their von-neumann entropy. *Linear Algebra and its Applications*, 532(Supplement C):534 – 549, 2017.
- [16] Vittorio Giovannetti and Simone Severini. The kirchhoff’s matrix-tree theorem revisited: counting spanning trees with the quantum relative entropy. *arXiv preprint arXiv:1102.2398*, 2011.
- [17] Wenxue Du, Xueliang Li, Yiyang Li, and Simone Severini. A note on the von neumann entropy of random graphs. *Linear Algebra and its Applications*, 433(11-12):1722–1725, 2010.
- [18] Manlio De Domenico and Jacob Biamonte. Spectral entropies as information-theoretic tools for complex network comparison. *Physical Review X*, 6(4):041062, 2016.
- [19] Alex Bavelas. Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, 22(6):725–730, 1950.
- [20] Mark Newman. *Networks: an introduction*. Oxford university press, 2010.
- [21] Xingqin Qi, Eddie Fuller, Qin Wu, Yezhou Wu, and Cun-Quan Zhang. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*, 194:240–253, 2012.
- [22] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.
- [23] John Scott. *Social network analysis*. Sage, 2017.
- [24] H. Theil. *Statistical decomposition analysis: With applications in the social and administrative sciences*. Studies in mathematical and managerial economics. North-Holland Pub. Co., 1972.
- [25] Pedro Conceicao and Pedro Ferreira. The young person’s guide to the theil index: Suggesting intuitive interpretations and exploring analytical applications. 2000.
- [26] Josef Novotny. On the measurement of regional inequality: does spatial dimension of income inequality matter? *The Annals of Regional Science*, 41(3):563–580, 2007.
- [27] Rajan K Sampath. Equity measures for irrigation performance evaluation. *Water International*, 13(1):25–32, 1988.
- [28] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley and Sons, 2012.
- [29] Mark M Wilde. From classical to quantum shannon theory. *arXiv preprint arXiv:1106.1445*, 2011.
- [30] Shujuan Cao, Matthias Dehmer, and Yongtang Shi. Extremality of degree-based graph entropies. *Information Sciences*, 278:22–33, 2014.
- [31] Michael William Newman. *The Laplacian spectrum of graphs*. PhD thesis, University of Manitoba, 2000.
- [32] Chris Godsil and Gordon F Royle. *Algebraic graph theory*, volume 207. Springer Science and Business Media, 2013.

- [33] Michael Dairyko, Leslie Hogben, Jephian C-H Lin, Joshua Lockhart, David Roberson, Simone Severini, and Michael Young. Note on von neumann and renyi entropies of a graph. *arXiv preprint arXiv:1609.00420*, 2016.
- [34] Martin Muller-Lennert, Frederic Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum renyi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, 2013.
- [35] Stefan Waner. *Introduction to analysis*. 1995.
- [36] Raj Jain, Dah-Ming Chiu, and William R Hawe. *A quantitative measure of fairness and discrimination for resource allocation in shared computer system*, volume 38. Eastern Research Laboratory, Digital Equipment Corporation Hudson, MA, 1984.