

City University of New York (CUNY)

## CUNY Academic Works

---

Publications and Research

City College of New York

---

2011

### Connections between Classical and Parametric Network Entropies

Matthias Dehmer

*Institute for Bioinformatics and Translational Research*

Abbe Mowshowitz

*CUNY City College*

Frank Emmert-Streib

*Queen's University Belfast*

[How does access to this work benefit you? Let us know!](#)

More information about this work at: [https://academicworks.cuny.edu/cc\\_pubs/60](https://academicworks.cuny.edu/cc_pubs/60)

Discover additional works at: <https://academicworks.cuny.edu>

---

This work is made publicly available by the City University of New York (CUNY).

Contact: [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu)

# Connections between Classical and Parametric Network Entropies

Matthias Dehmer<sup>1\*</sup>, Abbe Mowshowitz<sup>2</sup>, Frank Emmert-Streib<sup>3</sup>

**1** Institute for Bioinformatics and Translational Research, UMIT, Hall in Tirol, Austria, **2** Department of Computer Science, The City College of New York (CUNY), New York, New York, United States of America, **3** Computational Biology and Machine Learning, Center for Cancer Research and Cell Biology, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, United Kingdom

## Abstract

This paper explores relationships between classical and parametric measures of graph (or network) complexity. Classical measures are based on vertex decompositions induced by equivalence relations. Parametric measures, on the other hand, are constructed by using information functions to assign probabilities to the vertices. The inequalities established in this paper relating classical and parametric measures lay a foundation for systematic classification of entropy-based measures of graph complexity.

**Citation:** Dehmer M, Mowshowitz A, Emmert-Streib F (2011) Connections between Classical and Parametric Network Entropies. PLoS ONE 6(1): e15733. doi:10.1371/journal.pone.0015733

**Editor:** Fabio Rapallo, University of East Piedmont, Italy

**Received:** August 4, 2010; **Accepted:** November 22, 2010; **Published:** January 5, 2011

**Copyright:** © 2011 Dehmer et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defense and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defense or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute for Government purposes notwithstanding any copyright notation hereon. Matthias Dehmer thanks the Austrian Science Funds for supporting this work (project P22029-N13). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: matthias.dehmer@umit.at

## Introduction

Information theory has proven to be a useful tool in the analysis and measurement of network complexity [1]. In particular, many researchers have investigated the application of entropy measures to graphs, see [1–4]. A variety of entropy-based measures have been used to characterize networks associated with biological or chemical systems [5,6]; a recent application in computational biology uses an entropy measure to analyze metabolic networks [7,8]. In addition to the use of measures on graphs to analyze biological or chemical systems, information theory has been employed in network physics, see [1,9,10]. Arnaud et al. [1] provide a comprehensive review of Shannon entropy measures applied to network ensembles. The measures discussed in this review are based on probabilistic aspects of networks such as the expected number of links, the expected community structure, or the expected degree sequences that have been used to define probabilities. In addition, Arnaud et al. [1] compared Shannon entropies on network ensembles with Gibbs and von Neumann entropies by plotting numerical values of the respective measures. By contrast, we will establish connections between different definitions of entropy by means of inequalities. Sanchirico et al. [10] explored another problem in network theory, namely, characterizing complex networks based on degree distributions. In particular, entropy functions have been used to investigate scale-free networks, see [10]. Finally, Krawitz et al. [9] have applied the so-called Basin entropy to boolean networks, which have been shown to be of great value in analyzing biological [7] and related systems [9]. Krawitz et al. [9] applied the Basin entropy measure to specific components of boolean networks [9].

In these applications, entropy provides a measure of network connectivity. It is noteworthy that Krawitz et al. [9] were able to estimate the Basin entropy from time-series data, since the model thus becomes applicable to erroneous networks (i.e., graphs affected by measurement errors) which are of great importance in biology.

The information measures we want to consider in this paper represent the structural information content of a network [5,11–13]. In particular, they have been applied to special classes of graphs and have figured prominently in research on topological aspects of biological and chemical systems, e.g., see, [5,11–18]. Common to all such research is the use of Shannon's [19] classical measure to derive entropies of the underlying graph topology interpreted as the structural information content of a graph. [5,11–13]. Measures of this kind are functions of probability values that derive, in the classical case [5,12], from a graph invariant and an equivalence relation [11,12,20]. Thus far, a number of specialized measures have been developed that are used primarily to characterize the structural complexity of chemical graphs [11,21,22]. That is to say, these measures can be viewed as indexes of complexity based on certain structural features of a graph. In the classical cases, special graph invariants (e.g., number of vertices, edges, degrees, distances etc.) and equivalence relations have given rise to special measures of information contents [11,12,15].

Another class of graph entropies, not based on a graph invariant associated with an equivalence relation, has also been explored. These alternative measures are based on information functions [23] that assign a probability value to each vertex of a graph [23]. An interesting feature of these measures is that they are

parametric, see, e.g., [2,15], thus allowing the formulation of optimization problems involving the parameters for given data sets. This approach to measurement is applicable to research problems in graph complexity, data analysis, and machine learning. Furthermore, the measures are computable in polynomial time because they depend on determining metrical properties of graphs [24]. In view of the large number of existing quantitative measures of network structure [22,25], methods are needed for comparing the different indexes and investigating their interrelations. Such research on interrelations can be expected to yield new insights into complex systems that can be represented by graphs [22,26].

One promising direction is to infer inequalities between such indices describing network information contents. Relatively little work on this problem has appeared in the literature, see, e.g., [27,28]. Thus far we have studied in [2] so-called implicit information inequalities involving two parametric entropies using different information functions. General as well as special graphs have been considered [2]. The present paper deals mainly with inequalities between classical and parametric entropies. On the one hand, this gives rise to general information inequalities between measures; on the other hand, bounds for special classes of graphs can be obtained.

The paper is organized as follows: In Section ‘Methods and Results’, we describe the classes of information measures to be examined, and detail relevant properties. Also, we prove inequalities between classical and parametric entropies. The last section provides a summary and conclusion.

## Methods and Results

### Classical Measures and Parametric Graph Entropies

In this section, we sketch briefly known graph entropy measures for determining the information content of networks. As a preliminary remark,  $|A|$  denotes the cardinality of a given set  $A$ . Now, let  $G=(V,E)$  be a graph and let  $|V| := n$ . The existing graph entropy measures can be divided into two main classes: (i) Classical measures [14] and (ii) parametric measures [23]. Classical measures  $I(G,\tau)$  are defined relative to a partition of a set  $X$  of graph elements induced by an equivalence relation  $\tau$  on  $X$ . More precisely, let  $X$  be a set of graph elements (typically vertices), and let  $\{X_i\}$  for  $1 \leq i \leq k$ , be a partition of  $X$  induced by  $\tau$ . Suppose further that  $p_i := \frac{|X_i|}{|X|}$ . Then

$$I(G,\tau) = - \sum_{i=1}^k p_i \log(p_i). \tag{1}$$

Parametric measures are defined on graphs relative to information functions. Such functions are not identically zero and map graph elements (typically vertices) to the non-negative reals. For simplicity of description, we consider information functions defined on  $V$ . Let  $q_i := \frac{f(v_i)}{\sum_{j=1}^n f(v_j)}$  for  $1 \leq i \leq n$ . Clearly, the  $q_i$  form a probability distribution over the vertices. Then

$$I_f(G) := - \sum_{i=1}^n q_i \log(q_i). \tag{2}$$

In general, a detailed overview of graph entropy measures can be found in [2,11,14]. Note that various other graph entropies have also been developed, see, e.g., [3,4,29–32] but these won’t be discussed here.

The structural feature of a network captured by a classical information measure depends on the graph invariant and the equivalence criterion  $\tau$ . This is clear from Equation (1). The relationship between the quantitative measure and graph structure for classical measures is examined further by Nikolić [33].

For more general measures (Equation (2)), the structural feature depends on the information function used to define the measure. Examples are given by

$$f^1(v_i) := c_1 |S_1(v_i, G)| + c_2 |S_2(v_i, G)| + \dots + c_{\rho(G)} |S_{\rho(G)}(v_i, G)|, \tag{3}$$

$$c_k > 0, 1 \leq k \leq \rho(G),$$

$$f^2(v_i) := c_i \delta(v_i), c_k > 0, \tag{4}$$

$$f^3(v_i) := c_i \sigma(v_i), c_k > 0. \tag{5}$$

The  $c_k$  are positive coefficients used to weight structural differences in a graph [23] and  $|S_j(v_i, G)|$  are the  $j$ -sphere cardinalities.  $\delta(v_i)$  denotes the degree and  $\sigma(v_i)$  the eccentricity of the vertex  $v_i \in V$ .  $\rho(G)$  stands for the diameter of  $G$ . Such functions are used to obtain the vertex probabilities as explained in [23]

$$p^f(v_i) := \frac{f(v_i)}{\sum_{j=1}^n f(v_j)}. \tag{6}$$

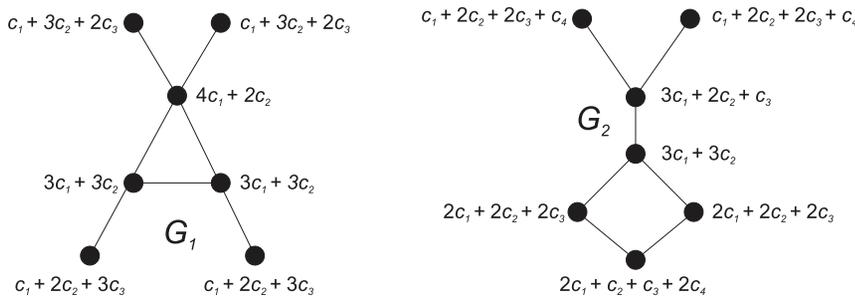
The family of graph entropies resulting from different probability distributions is represented by Equation (2). In the following, we provide examples of such an information function (choosing  $f^1$  as a special case) as well as of the resulting entropy measure. Furthermore, we compare this measure with a classical one using an identity graph as an example. Note that the information function  $f^1$  has already been used to characterize chemical structures [15]. But first consider the graphs in Figure 1 to explain the graph entropy measure  $I_{f^1}$  in more detail. For calculating this structural information content explicitly, we set

$$c_1 := \rho(G), c_2 := \rho(G) - 1, \dots, c_{\rho(G)} := 1. \tag{7}$$

The structural feature captured by this measure is linked to the following observation: The more the vertices differ with respect to their spherical neighborhoods, the smaller is the value and conversely. Hence,  $I_{f^1}$  detects a kind of inner symmetry of an underlying graph. By using  $f^1$  in Equation (2), regular graphs have a constant information content equal to the maximum entropy (for every information function). For example, the graph  $C_7$  gives rise to (see Figure 2)

$$p^{f^1}(v_i) = \frac{2c_1 + 2c_2 + 2c_3}{7(2c_1 + 2c_2 + 2c_3)} = \frac{1}{7}, \tag{8}$$

and finally  $I_{f^1}(C_7) = \log(7)$ . To compare the parametric with one of the classical measures, we consider a special case of Equation (1) in which the probabilities are determined by the respective sizes of the (vertex) orbits of the automorphism group, see [12]. The resulting graph entropy measure is denoted by  $I_a$ . Because  $C_7$  is vertex-transitive, there is only one orbit containing all vertices and



**Figure 1. We obtain  $2.78 = I_{f^1}(G_1) < I_{f^1}(G_2) = 2.79$ .**  
doi:10.1371/journal.pone.0015733.g001

therefore we obtain

$$I_a(C_7) = 0. \tag{9}$$

Now, we evaluate the two different graph entropy measures  $I_{f^1}$  and  $I_a$  for the identity graph depicted in Figure 3. This graph  $G_4$  has a trivial automorphism group (i.e., the identity graph) and, hence, all orbits are singleton sets. This implies

$$I_a(G_4) = \log(6). \tag{10}$$

But when calculating  $I_{f^1}$ , we get

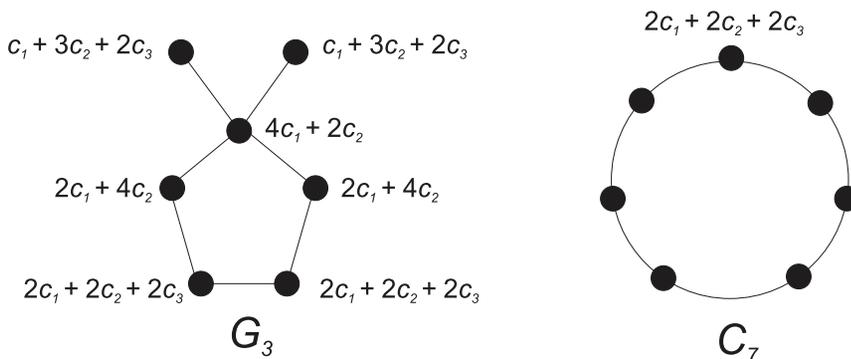
$$\begin{aligned}
 I_{f^1}(G_4) = & - \left[ \frac{c_1 + 3c_2 + c_3}{\sum_{j=1}^6 f^1(v_j)} \log \left( \frac{c_1 + 3c_2 + c_3}{\sum_{j=1}^6 f^1(v_j)} \right) \right. \\
 & + \frac{4c_1 + c_2}{\sum_{j=1}^6 f^1(v_j)} \log \left( \frac{4c_1 + c_2}{\sum_{j=1}^6 f^1(v_j)} \right) \\
 & + 2 \frac{3c_1 + 2c_2}{\sum_{j=1}^6 f^1(v_j)} \log \left( \frac{3c_1 + 2c_2}{\sum_{j=1}^6 f^1(v_j)} \right) \\
 & + \frac{c_1 + 2c_2 + 2c_3}{\sum_{j=1}^6 f^1(v_j)} \log \left( \frac{c_1 + 2c_2 + 2c_3}{\sum_{j=1}^6 f^1(v_j)} \right) \\
 & \left. + \frac{2c_1 + 2c_2 + c_3}{\sum_{j=1}^6 f^1(v_j)} \log \left( \frac{2c_1 + 2c_2 + c_3}{\sum_{j=1}^6 f^1(v_j)} \right) \right], \tag{11}
 \end{aligned}$$

where  $\sum_{j=1}^6 f^1(v_j) = 14c_1 + 12c_2 + 4c_3$ . Finally, we find that  $I_{f^1}(G_4) < \log(6)$ . In contrast, note that  $I_a$  represents a symmetry-

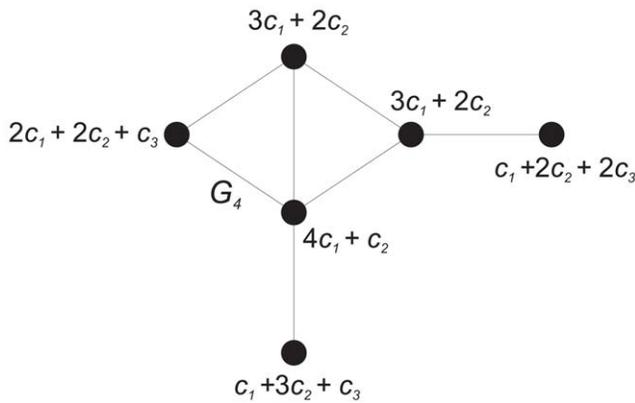
based complexity measure [12]. Other structural features could be chosen to provide other or more comprehensive measures of complexity. For example, Bertz [34] modified the total information content discussed by Bonchev [11] to obtain a different measure. Other approaches to tackle this challenging problem have been outlined by Nikolić [33]. To better understand the measure  $I_{f^1}$  and to get an intuitive sense of its complexity, we perform a parameter study. More precisely, we show the entropy represented by Equation (11) for different parameters. We plotted the entropy for constant values of  $c_1$  (0, 0.5 - first row, and 1, 3 - second row) independent of the other variables  $c_2$  and  $c_3$ , see Figure 4. Clearly, the positions of maximum entropy are shifted for different values of  $c_1$ ; and for higher values of  $c_1$  the location of the maximum approaches that of  $c_2 = c_3 = c_1$ .

### Inequalities for Parametric Graph Entropies and Classical Measures

Most of the graph entropy measures developed thus far have been applied in mathematical chemistry and biology [2,11,14]. These measures have been used to quantify the complexity of chemical and biological systems that can be represented as graphs. Given the profusion of such measures, it is useful, for instance, to prove bounds for special graph classes or to study interrelations among them. Such results might be useful to investigate network-based systems as well as to design new network measures more adequately. In terms of information-theoretic measures for graphs, relatively little attention has been paid to this effort. An early attempt in this direction was undertaken by Bonchev [27] when investigating inequalities between entropy-based network measures by considering special graph classes. In particular, Bonchev [27] used such inequalities to investigate the concept of branching



**Figure 2. We obtain  $2.79 = I_{f^1}(G_3) < I_{f^1}(C_7) = \log(7)$ .**  
doi:10.1371/journal.pone.0015733.g002



**Figure 3. A graph with identity group.**  
doi:10.1371/journal.pone.0015733.g003

[35] in molecules. A topic within this general framework which seems to be completely unexplored is an analysis (using inequalities) of formal relations between complexity measures. On the one hand, this could be done by starting from special graph classes which are interesting for practical applications. But, on the other hand, one can also infer more general interrelations between non-information-theoretic and information-theoretic measures (e.g., see Theorem (1)).

In [28], we have investigated so-called implicit information inequalities for graphs. Such information inequalities describe general interrelations between parametric measures based on arbitrary information functions. In this section, we demonstrate inequalities between classical graph entropies and the entropy families given by Equation (2). As mentioned earlier, numerous network information measures [11,14,22] have been developed, but their mathematical

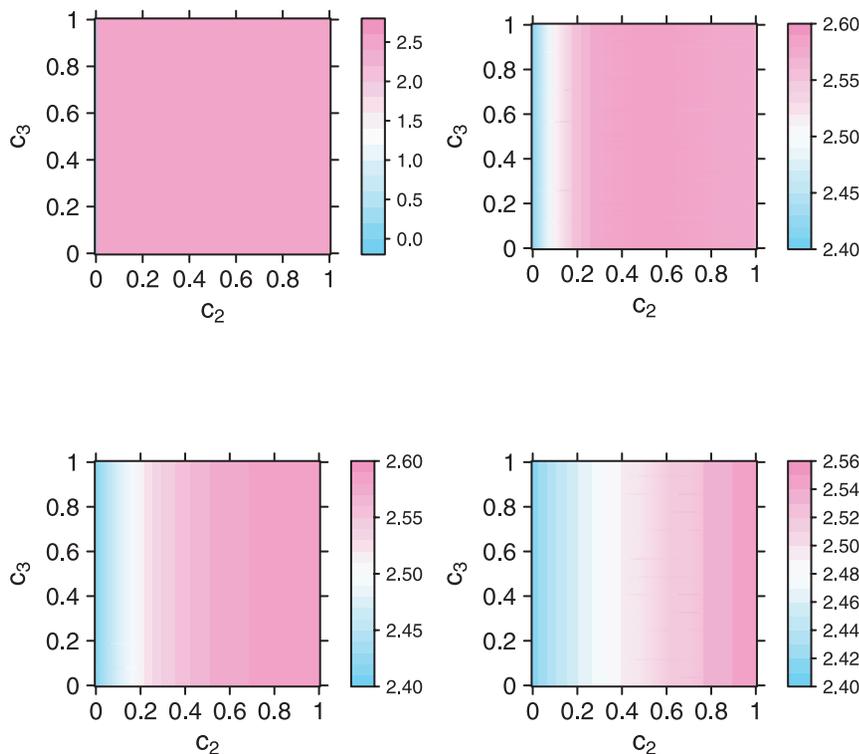
properties have yet to be studied in detail. Therefore, the results of this section can be considered as a first attempt to detail these properties. Some of the interrelations represent bounds which hold for special graph classes (with no assumptions about the parameters involved) when using a special information function.

We start with a more general statement expressing an interrelation between the parametric entropy and a classical entropy measure that is based on certain equivalence classes associated with an arbitrary equivalence relation. In particular, this interrelation can be stated as an upper bound of the parametric entropy depending on the classical entropy measure.

**Theorem 1.** Let  $G=(V,E)$  be an arbitrary graph, and let  $X_i, 1 \leq i \leq k$  be the equivalence classes associated with an arbitrary equivalence relation on  $X$ . Suppose further that  $f$  is an information function with  $f(v_i) > |X_i|$  for  $1 \leq i \leq k$ ,  $p(v_i) = \frac{f(v_i)}{\sum_{j=1}^n f(v_j)}$  and  $c := \frac{1}{\sum_{j=1}^n f(v_j)}$ . Then,

$$\begin{aligned} \frac{1}{|X|} I_f(G) &< c I(G, \tau) - \sum_{i=1}^k \frac{|X_i|}{|X|} c \log(c) - \frac{\log(|X|)}{|X|} \sum_{i=1}^k p(v_i) \\ &- \frac{1}{|X|} \sum_{i=k+1}^n p(v_i) \log(p(v_i)) \\ &+ \frac{1}{|X|} \sum_{i=1}^k p(v_i) \log\left(1 + \frac{|X|}{c f(v_i)}\right), \\ &+ \sum_{i=1}^k \log\left(\frac{p(v_i)}{|X|} + 1\right). \end{aligned} \tag{12}$$

**Proof.** Note that we use the simplified notation  $p(v_i)$  instead of  $p^f(v_i)$  because it is clear (by definition) that a vertex probability



**Figure 4. Entropy vs. Parameter Values.**  
doi:10.1371/journal.pone.0015733.g004

value depends on the information function  $f$ . Now, suppose  $f(v_i) > |X_i|$ ,  $i = 1, 2, \dots, k$ ,  $k \leq n$ . Then,

$$\frac{|X_i|}{|X|} c < \frac{1}{|X|} p(v_i) < \frac{1}{|X|} p(v_i) + 1 \tag{13}$$

and

$$\begin{aligned} \frac{|X_i|}{|X|} c \log\left(\frac{|X_i|}{|X|} c\right) &< \left(\frac{1}{|X|} p(v_i) + 1\right) \log\left(\frac{1}{|X|} p(v_i) + 1\right) \\ &= \frac{1}{|X|} p(v_i) \log\left(\frac{1}{|X|} p(v_i) + 1\right) + \log\left(\frac{1}{|X|} p(v_i) + 1\right). \end{aligned} \tag{14}$$

Assuming

$$\begin{aligned} \log\left(\frac{1}{|X|} p(v_i) + 1\right) &= \log\left(\frac{1}{|X|} p(v_i) \left(1 + \frac{|X|}{p(v_i)}\right)\right) \\ &= \log\left(\frac{1}{|X|}\right) + \log(p(v_i)) + \log\left(1 + \frac{|X|}{f(v_i)c}\right), \end{aligned} \tag{15}$$

and making use of Inequality (14) we derive

$$\begin{aligned} \frac{|X_i|}{|X|} c \log\left(\frac{|X_i|}{|X|}\right) + \frac{|X_i|}{|X|} c \log(c) &< \frac{1}{|X|} p(v_i) \log\left(\frac{1}{|X|}\right) + \frac{1}{|X|} p(v_i) \log(p(v_i)) \\ &+ \frac{1}{|X|} p(v_i) \log\left(1 + \frac{|X|}{f(v_i)c}\right) + \log\left(\frac{1}{|X|} p(v_i) + 1\right), \end{aligned} \tag{16}$$

or

$$\begin{aligned} -\frac{|X_i|}{|X|} c \log\left(\frac{|X_i|}{|X|}\right) - \frac{|X_i|}{|X|} c \log(c) &> -\frac{1}{|X|} p(v_i) \log\left(\frac{1}{|X|}\right) - \frac{1}{|X|} p(v_i) \log(p(v_i)) \\ &- \frac{1}{|X|} p(v_i) \log\left(1 + \frac{|X|}{f(v_i)c}\right) - \log\left(\frac{1}{|X|} p(v_i) + 1\right). \end{aligned} \tag{17}$$

Adding up these inequalities (i.e., by adding across the vertices), we obtain

$$\begin{aligned} cI(G, \tau) &> \sum_{i=1}^k \frac{|X_i|}{|X|} c \log(c) + \frac{\log(|X|)}{|X|} \sum_{i=1}^k p(v_i) \\ &+ \frac{1}{|X|} I_f(G) + \frac{1}{|X|} \sum_{i=k+1}^n p(v_i) \log(p(v_i)) \\ &- \frac{1}{|X|} \sum_{i=1}^k p(v_i) \log\left(1 + \frac{|X|}{f(v_i)c}\right) \\ &- \sum_{i=1}^k \log\left(\frac{1}{|X|} p(v_i) + 1\right). \end{aligned} \tag{18}$$

But this is Inequality (12).

In the following, we apply the assumption  $f(v_i) > |X_i|$  for  $1 \leq i \leq k$  to some special graph classes and using the proof technique of the previous theorem. The set  $X$  is taken to be  $V$ , and thus the equivalence relation induces a partition of  $V$  into equivalence classes of vertices. These assumptions allow for obtaining upper bounds on  $I_f(G)$  which can be stated as corollaries of Theorem (1).

**Corollary 2.** *Let  $S_n$  be a star graph having  $n$  vertices and suppose  $v_1$  is the vertex with degree  $n-1$ . The remaining  $n-1$  non-hub vertices are labeled arbitrarily.  $v_\mu$  stands for a non-hub vertex. Let  $f$  be an information function satisfying the conditions of Theorem (1). Let  $V_1 := \{v_1\}$  and  $V_2 := \{v_2, v_3, \dots, v_n\}$  denote the orbits of the automorphism group of  $S_n$  forming a partition of  $V$ . Then*

$$\begin{aligned} I_f(S_n) &< p(v_1) \log\left(1 + \frac{1}{p(v_1)}\right) + p(v_\mu) \log\left(1 + \frac{1}{p(v_\mu)}\right) \\ &+ \log(1 + p(v_1)) + \log(1 + p(v_\mu)) \\ &- \sum_{i=2}^n p(v_i) \log(p(v_i)) - (n-1) \cdot c \cdot \log[(n-1) \cdot c] - c \log(c). \end{aligned} \tag{19}$$

$i \neq \mu$

**Proof.** By hypothesis  $|V_1| = 1$  and  $|V_2| = n-1$  so that

$$n-1 < f(v_1) \quad \text{and} \quad 1 < f(v_\mu), \quad 2 \leq \mu \leq n. \tag{20}$$

The information functions given by Equation (3), (4), (5) can be seen to satisfy the above conditions by choosing appropriate  $c_k > 0$ . Again, with  $c := \frac{1}{\sum_{j=1}^n f(v_j)}$ , the Inequalities (20) yield

$$(n-1) \cdot c \cdot \log[(n-1) \cdot c] < (p(v_1) + 1) \log(p(v_1) + 1), \tag{21}$$

$$c \cdot \log(c) < (p(v_\mu) + 1) \log(p(v_\mu) + 1). \tag{22}$$

Now, applying the proof technique of Theorem (1) and performing some elementary transformations, we obtain Inequality (19).

**Corollary 3.**  *$G_n^I$  be an identity graph having  $n \geq 6$  vertices.  $G_n^I$  has only the identity automorphism and therefore each orbit is a singleton set, i.e.,  $|V_i| = 1, 1 \leq i \leq n$ . Let  $f$  be an information function satisfying the conditions of Theorem (1). Then,*

$$\begin{aligned} I_f(G_n^I) &< \sum_{j=1}^n p(v_j) \log\left(1 + \frac{1}{p(v_i)}\right) \\ &+ \sum_{j=1}^n \log(1 + p(v_i)) - n \cdot c \log(c). \end{aligned} \tag{23}$$

**Proof.** By hypothesis

$$1 < f(v_1), \tag{24}$$

$$1 < f(v_2), \tag{25}$$

⋮

$$1 < f(v_n). \tag{26}$$

$$2 < f(v_1), \tag{33}$$

Clearly,

$$c \cdot \log(c) < (p(v_1) + 1) \log(p(v_1) + 1), \tag{27}$$

$$2 < f(v_2), \tag{34}$$

$$c \cdot \log(c) < (p(v_2) + 1) \log(p(v_2) + 1), \tag{28}$$

⋮

$$2 < f(v_{\frac{n}{2}}). \tag{35}$$

⋮

$$c \cdot \log(c) < (p(v_n) + 1) \log(p(v_n) + 1). \tag{29}$$

Then, it is easy to see that

$$2c \cdot \log(2c) < (p(v_1) + 1) \log(p(v_1) + 1), \tag{36}$$

$$2c \cdot \log(2c) < (p(v_2) + 1) \log(p(v_2) + 1), \tag{37}$$

⋮

$$2c \cdot \log(2c) < (p(v_n) + 1) \log(p(v_n) + 1). \tag{38}$$

Once again applying the proof technique of Theorem (1) and performing some elementary transformations, we obtain Inequality (23).

Corollary (3) leads immediately to

**Corollary 4.** *Let  $G_n^I$  be an identity graph having  $n \geq 6$  satisfying the conditions of Corollary (3). Then,*

$$I_f(G_n^I) < n \log\left(1 + \frac{1}{p^*}\right) + n \log(2) + \frac{1}{\min_{1 \leq j \leq n} \{p(v_j)\}} \log\left(\sum_{j=1}^n f(v_j)\right), \tag{30}$$

When  $n$  is odd, we have

$$2 < f(v_1), \tag{39}$$

$$2 < f(v_2), \tag{40}$$

⋮

$$2 < f\left(v_{\lfloor \frac{n}{2} \rfloor - 1}\right), \tag{41}$$

$$1 < f\left(v_{\lfloor \frac{n}{2} \rfloor}\right), \tag{42}$$

where  $p^* := \min\{p(v_1), p(v_2), \dots, p(v_n)\}$ .

**Corollary 5.** *Let  $G_n^P$  be a path graph having  $n$  vertices and let  $f$  be an information function satisfying the conditions of Theorem (1). If  $n$  is even,  $G_n^P$  possesses  $\frac{n}{2}$  equivalence classes  $V_i$  and each  $V_i$  contains 2 vertices. Then,*

$$I_f(G_n^P) < \sum_{j=1}^{\frac{n}{2}} p(v_j) \log\left(1 + \frac{1}{p(v_i)}\right) + \sum_{j=1}^{\frac{n}{2}} \log(1 + p(v_i)) - \sum_{j=\frac{n}{2}+1}^n p(v_j) \log(p(v_i)) - n \cdot c \cdot \log(2c). \tag{31}$$

and

$$2c \cdot \log(2c) < (p(v_1) + 1) \log(p(v_1) + 1), \tag{43}$$

$$2c \cdot \log(2c) < (p(v_2) + 1) \log(p(v_2) + 1), \tag{44}$$

⋮

$$2c \cdot \log(2c) < \left(p\left(v_{n-\lfloor \frac{n}{2} \rfloor - 1}\right) + 1\right) \log\left(p\left(v_{n-\lfloor \frac{n}{2} \rfloor - 1}\right) + 1\right), \tag{45}$$

If  $n$  is odd, then there exist  $n - \lfloor \frac{n}{2} \rfloor$  equivalence classes  $n - \lfloor \frac{n}{2} \rfloor - 1$  that have 2 elements and only one class containing a single element. This implies,

$$I_f(G_n^P) < \sum_{j=1}^{n-\lfloor \frac{n}{2} \rfloor} p(v_j) \log\left(1 + \frac{1}{p(v_i)}\right) + \sum_{j=1}^{n-\lfloor \frac{n}{2} \rfloor} \log(1 + p(v_i)) - \sum_{j=n-\lfloor \frac{n}{2} \rfloor+1}^n p(v_j) \log(p(v_i)) - \left(n - \lfloor \frac{n}{2} \rfloor - 1\right) 2c \cdot \log(2c) - c \cdot \log(c). \tag{32}$$

**Proof.** By hypothesis

$$c \cdot \log(2c) < (p(v_{n-\lfloor \frac{n}{2} \rfloor}) + 1) \log(p(v_{n-\lfloor \frac{n}{2} \rfloor}) + 1). \quad (46)$$

Multiplying these inequality systems by -1 and performing the addition step (of the proof technique of Theorem (1) gives Inequality (31) and Inequality (32).

Assuming different initial conditions, we can derive additional inequalities between classical and parametric measures. We state the following theorems without proofs because the underlying technique is similar to the proofs of the previous assertions.

**Theorem 6.** Let  $G$  be an arbitrary graph and  $p(v_i) < |X_i|$ . Then,

$$\begin{aligned} \frac{1}{|X|} I_f(G) > I(G, \tau) - \frac{1}{|X|} \sum_{i=k+1}^n p(v_i) \log(p(v_i)) - \frac{\log(|X|)}{|X|} \sum_{i=1}^k p(v_i) \\ - \frac{1}{|X|} \sum_{i=1}^k |X_i| \log\left(1 + \frac{|X|}{|X_i|}\right) - \sum_{i=1}^k \log\left(1 + \frac{|X_i|}{|X|}\right). \end{aligned} \quad (47)$$

**Theorem 7.** Let  $G$  be an arbitrary graph with  $p_i$  being the probabilities satisfying Equation (1) such that  $p_i < f(v_i)$ . Then,

$$\begin{aligned} \frac{1}{c} I(G, \tau) > I_f(G) + \frac{\log(c)}{c} + \sum_{i=k+1}^n p(v_i) \log(p(v_i)) \\ - \sum_{i=1}^k \log(p(v_i)) - \sum_{i=1}^k \log\left(1 + \frac{1}{p(v_i)}\right) (1 + p(v_i)). \end{aligned} \quad (48)$$

For identity graphs, we can obtain a general upper bound for the parametric entropy.

**Corollary 8.** Let  $G_n^I$  be an identity graph having  $n$  vertices. Then,

$$\begin{aligned} I_f(G_n^I) < \log(n) - c \cdot \log(c) + \sum_{i=1}^n \log(p(v_i)) \\ + \sum_{i=1}^n \log\left(1 + \frac{1}{p(v_i)}\right) (1 + p(v_i)). \end{aligned} \quad (49)$$

## Discussion

Quantitative measures of network structure have been defined and applied in many different settings, see, e.g., [2,22,25]. For example, chemists represent molecules as graphs as an aid in distinguishing molecules and cataloguing their properties [36,37]; biologists model natural phenomena as complex networks in the study of brains and genetic information systems [38,39]; epidemiologists and computational biologists investigate the

## References

- Anand K, Bianconi G (2009) Entropy measures for networks: Toward an information theory of complex topologies. *Phys Rev E* 80: 045102(R).
- Dehmer M, Mowshowitz A (2010) A history of graph entropy measures. *Information Sciences*.
- Solé RV, Valverde S (2004) Information theory of complex networks: On evolution and architectural constraints. In: *Lecture Notes in Physics* volume 650: 189–207.
- Ulanowicz RE (2001) Information theory in ecology. *Computers and Chemistry* 25: 393–399.
- Rashevsky N (1955) Life, information theory, and topology. *Bull Math Biophys* 17: 229–235.
- Temkin O, Zeigarnik AV, Bonchev D (1996) *Chemical Reaction Networks. A Graph-Theoretical Approach*. CRC Press.
- Emmert-Streib F, Dehmer M (2010) Networks for systems biology: Conceptual connection of data and function. *IET Systems Biology*, accepted.
- Mazurie A, Bonchev D, Schwikowski B, Buck GA (2008) Phylogenetic distances are encoded in networks of interacting pathways. *Bioinformatics* 24: 2579–2585.
- Krawitz P, Shmulevich I (2007) Entropy of complex relevant components of boolean networks. *Phys Rev E* 76: 036115.
- Sanchirico A, Fiorentino M (2008) Scale-free networks as entropy competition. *Phys Rev E* 78: 046114.
- Bonchev D (1983) *Information Theoretic Indices for Characterization of Chemical Structures*. Chichester: Research Studies Press.
- Mowshowitz A (1968) Entropy and the complexity of the graphs I: An index of the relative complexity of a graph. *Bull Math Biophys* 30: 175–204.

spread of diseases in populations modeled as networks of individuals [40,41]; computer scientists design and build networks of artificial systems that support message exchange and distributed computation [42,43]. In each of these different settings, questions about the structure of networks arise, leading to the definition of mathematical functions designed to measure structural complexity. As a result of all these relatively independent scholarly efforts, many different measures [2,22,25,33] have been defined whose interrelations remain to be determined. This paper is intended as a contribution to the classification of these diverse measures of network structure. In particular, we have singled out two different classes of measures, namely classical and parametric entropy measures defined on graphs, and have examined their interrelations.

The approach taken in this paper is to establish inequalities between measures. As already mentioned, an early attempt in this direction has been undertaken by Bonchev [27] who proved inequalities in the course of investigating branching structures in molecules. Our aim here is somewhat broader, namely to establish general, formal relations between complexity measures defined on arbitrary graphs. Since complexity measures typically assign real numbers to networks, inequalities provide the foundation for constructing partial orders on sets of measures. Knowledge of such order relations enables us to use inequalities to infer bounds on specific measures of the structural complexity of graphs and networks. Knowing that measure  $a$  is bounded above by measure  $b$  whose maximum value has been established tells us that measure  $a$  has a maximum value less than that of measure  $b$ . Relatively little work on the problem of ordering entropy measures has appeared in the literature, see, e.g., [44,45].

The main focus of the paper has been on establishing inequalities between entropy-based measures, i.e., measures that make use of Shannon's classical entropy function. In particular, we examined inequalities involving classical and parametric information measures. Such measures have been used extensively to quantify the information content of systems represented by graphs [2,12,21,22]. For this reason, we believe that such inequalities are critical for a proper understanding of entropy-based measures.

The inequalities presented in this paper show interrelations between entropy-based measures applied to special classes of graphs. Establishing such inequalities for arbitrary graphs is a daunting task. The work reported here is thus a first step in that the methods employed can in principle be generalized to establish inequalities between information measures on arbitrary graphs. More research is clearly needed to extend the results to graphs in general, and ultimately to build a foundation for a unified interpretation of network complexity as measured by entropy-based functions.

## Author Contributions

Wrote the paper: MD AM FES.

13. Trucco E (1956) A note on the information content of graphs. *Bull Math Biol* 18: 129–135.
14. Bonchev D (2009) Information theoretic measures of complexity. In: Meyers R, ed. *Encyclopedia of Complexity and System Science*, Springer volume 5: 4820–4838.
15. Dehmer M, Varmuza K, Borgert S, Emmert-Streib F (2009) On entropy-based molecular descriptors: Statistical analysis of real and synthetic chemical structures. *J Chem Inf Model* 49: 1655–1663.
16. Emmert-Streib F, Dehmer M (2007) Information theoretic measures of UHG graphs with low computational complexity. *Appl Math Comput* 190: 1783–1794.
17. Kim J, Wilhelm T (2008) What is a complex graph? *Physica A* 387: 2637–2652.
18. Mehler A (2011) A quantitative graph model of social ontologies by example of Wikipedia. In: Dehmer M, Emmert-Streib F, Mehler A, eds. *Towards an Information Theory of Complex Networks: Statistical Methods and Applications*. Boston/Basel: Birkhäuser.
19. Shannon CE, Weaver W (1949) *The Mathematical Theory of Communication*. University of Illinois Press.
20. Mowshowitz A (1968) Entropy and the complexity of graphs IV: Entropy measures and graphical structure. *Bull Math Biophys* 30: 533–546.
21. Bonchev D (1979) Information indices for atoms and molecules. *Commun Math Comp Chem* 7: 65–113.
22. Bonchev D, Rouvray DH (2005) *Complexity in Chemistry, Biology, and Ecology*. Mathematical and Computational Chemistry. New York, NY, USA: Springer.
23. Dehmer M (2008) A novel method for measuring the structural information content of networks. *Cybernetics and Systems* 39: 825–843.
24. Skorobogatov VA, Dobrynin AA (1988) Metrical analysis of graphs. *Commun Math Comp Chem* 23: 105–155.
25. da F Costa L, Rodrigues F, Trivieso G (2007) Characterization of complex networks: A survey of measurements. *Advances in Physics* 56: 167–242.
26. Bonchev D (2003) *Complexity in Chemistry*. Introduction and Fundamentals. Boca Raton, FL, USA: Taylor and Francis.
27. Bonchev D, Trinajstić N (1977) Information theory, distance matrix and molecular branching. *J Chem Phys* 67: 4517–4533.
28. Dehmer M, Mowshowitz A (2010) Inequalities for entropy-based measures of network information content. *Applied Mathematics and Computation* 215: 4263–4271.
29. Körner J (1973) Coding of an information source having ambiguous alphabet and the entropy of graphs. *Transactions of the 6-th Prague Conference on Information Theory*. pp 411–425.
30. Simonyi G (1995) Graph entropy: A survey. In: Cook W, Lovász L, Seymour P, eds. *Combinatorial Optimization, DIMACS Series in Discrete Mathematics and Theoretical Computer Science* volume 20: 399–441.
31. Tutzauer F (2007) Entropy as a measure of centrality in networks characterized by path-transfer flow. *Social Networks* 29: 249–265.
32. Wilhelm T, Hollunder J (2007) Information theoretic description of networks. *Physica A* 388: 385–396.
33. Nikolić S, Trinajstić N (2000) Complexity of molecules. *J Chem Inf Comput Sci* 40: 920–926.
34. Bertz SH (1981) The first general index of molecular complexity. *Journal of the American Chemical Society* 103: 3241–3243.
35. Bonchev D (1995) Topological order in molecules I. molecular branching revisited. *Journal of Molecular Structure: THEOCHEM* 336: 137–156.
36. Balaban AT (1995) Chemical graphs: Looking back and glimpsing ahead. *Journal of Chemical Information and Computer Sciences* 35: 339–350.
37. Trinajstić N (1992) *Chemical Graph Theory*. Boca Raton, FL, USA: CRC Press.
38. Foulds LR (1992) *Graph Theory Applications*. Springer.
39. Mason O, Verwoerd M (2007) Graph theory and networks in biology. *IET Systems Biology* 1: 89–119.
40. Emmert-Streib F, Dehmer M (2008) *Analysis of Microarray Data: A Network-Based Approach*. Weinheim, Germany: Wiley-VCH.
41. Emmert-Streib F, Dehmer M (2009) Information processing in the transcriptional regulatory network of yeast: Functional robustness. *BMC Syst Biol* 3.
42. Patil A, Esfahanian AH, Xiao L, Liu Y (2007) Resource allocation using multiple edge-sharing multicast trees. *IEEE Transactions on Vehicular Technology* 58: 3178–3186.
43. Mehler A, Weiß P, Lücking A (2010) A network model of interpersonal alignment. *Entropy* 12: 1440–1483.
44. Lyons R (2010) Identities and inequalities for tree entropy. *Comb Probab Comput* 19: 303–313.
45. Zhang Z, Yeung RW (1998) On characterization of entropy functions via information inequalities. *IEEE Transactions on Information Theory* 44: 1440–1452.