MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **70** (2013) 885-900

ISSN 0340 - 6253

On Sphere-Regular Graphs and the Extremality of Information-Theoretic Network Measures

Veronika Kraus, Matthias Dehmer, Michael Schutte

Institute for Bioinformatics and Translational Research, UMIT, Eduard Wallnoefer Zentrum 1, 6060 Hall in Tyrol, Austria.

{veronika.kraus, matthias.dehmer,michael.schutte}@umit.at

(Received June 2, 2013)

Abstract

The entropy of a chemical graph can be interpreted as its structural information content. In this paper, we study extremality properties of graph entropies based on so-called information functionals. Based on different information functionals using metrical properties of graphs, we tackle the problem of determining classes of graphs which take maximal and minimal values. Also, we define a novel class of graphs which maximizes the structural information content based on the functional using i-spheres. Under certain assumptions, this class fully determines the class of maximal graphs based on their structural information content. For minimal graphs and other functionals, an analytic approach to the question failed. Hence we performed simulations and provide several conjectures on classes of extremal graphs by using our numerical results.

1 Introduction

The study of extremal properties of structural graph measures has been performed in terms of both theoretical and practical purposes, see [12, 22, 21]. Examples thereof are the well-known Randić index [16] and the Balaban J index where their extremality, by using several graph classes, has been investigated in [12] and [21], respectively. Similarly, [10] determined trees with minimal ABC-index which is based on vertex degrees. Other related work has been described in [6].

In this paper, we investigate extremality properties of graph entropy measures which are based on information functionals [4]. We emphasize that entropy-based graph measures have been defined in a broad variety, e.g., [14, 11, 7, 8, 5]. In many cases, the base for those measures is so-called Shannon entropy [18], a concept from information theory to measure the structural information content of arbitrary discrete random variables. In recent work, [6] tackled the problem of exploring extremal properties of graph entropies by inferring lower and upper bounds for certain graph classes. Also, *quasi-majorization* has been introduced and several extremality statements for graph entropies have been proven by using this method [6].

The main contribution of this paper is twofold: First, we introduce a novel graph class which we call *sphere-regular* and derive an inclusion criterion to classify these graphs. Second, we prove statements to find graphs possessing maximal and minimal entropy by using the class of entropy measures introduced in [4]. To prove the results, we put the emphasis on two information functionals for determining the entropy of the underlying graph topology; the first functional is based on i-spheres [4], the second is based on the vertex eccentricity. As a result, we find that the derived statements depend on the underlying weight sequences of the just mentioned information functionals (see Section 2) and, hence, proving quite general extremality results turned out to be very challenging. Also, we have performed simulations and provide several conjectures for classes of extremal graphs based on the obtained numerical results.

2 Notation

For a discrete probability distribution $\mathbf{p} = (p_1, \ldots, p_k)$, the Shannon entropy [18] is defined by

$$I(\mathbf{p}) = -\sum_{i=1}^{k} p_i \log p_i. \tag{1}$$

The Shannon entropy attains its maximum for the uniform distribution $\mathbf{p} = (\frac{1}{k}, \ldots, \frac{1}{k})$, where

$$I(\mathbf{p}) = -\sum_{i=1}^{k} \frac{1}{k} \log \frac{1}{k} = -\log \frac{1}{k} = \log k. \tag{2}$$

The minimum of $I(\mathbf{p})$ is taken at the distribution $\mathbf{p} = (1, 0, 0, \ldots, 0)$, where $I(\mathbf{p}) = 0$.

In the case of structural graph measures, we extract a vector

$$\mathbf{X} = (x_1, x_2, \ldots, x_k),$$

containing topological information, and define a probability vector by

$$\mathbf{p_X} = \left(\frac{x_1}{\sum_{j=1}^{k} x_j}, \ldots, \frac{x_k}{\sum_{j=0}^{k} x_j} \right), \tag{3}$$

to which we can apply Shannon's formula (1). This construction was first used by [17] and [19], who defined the vector $\mathbf{X}$ to be the sizes of the vertex orbits of a graph G. [14] was the first who used this measure to characterize the structural information content of a graph and proved properties thereof. Later, [2] defined graph entropies which are called magnitude-based information measures. Another approach to define the entropy of a graph in due to [11]. This quantity often called Körner entropy has been introduced to solve a coding problem and is therefore rooted in information theory, see [11].

Now, we state further definitions.

Definition 2.1. *[4] The structural information content of a graph G is given by*

$$I_f(G) = -\sum_{i=1}^{n} \frac{f(v_i)}{\sum_{j=1}^{n} f(v_j)} \log \frac{f(v_i)}{\sum_{j=1}^{n} f(v_j)}, \tag{4}$$

for some information functional $f : V \to \mathbb{R}$.

Throughout Sections 4 and 5 we will use the sphere-functional f_S. Let $\rho(G)$ denote the diameter of a graph G, and for i from 1 to $\rho(G)$, let $S_i(v)$ denote the i-th sphere of the vertex v, i.e. the set of all vertices at distance i of v.

Definition 2.2. *The sphere-functional f_S is defined by*

$$\begin{aligned} &f_S : V \to \mathbb{R} \\ &f_S(v) = c_1|S_1(v)| + c_2|S_2(v)| + \cdots c_{\rho(G)}|S_{\rho(G)}(v)|, \end{aligned} \tag{5}$$

where $(c_1, \ldots, c_{\rho(G)})$ is a real-valued weight sequence.

For our purposes, we will mainly use decreasing sequences $c_1, \ldots, c_{\rho(G)}$ of

constant decrease: $c_1 := S, c_2 := S - k, \ldots, c_{\rho(G)} := S - (\rho(G) - 1)k,$ (6a)

quadratic decrease: $c_1 := S^2, c_2 := (S - k)^2, \ldots, c_{\rho(G)} := (S - (\rho(G) - 1)k)^2$, or (6b)

exponential decrease: $c_1 := S, c_2 := Se^{-k}, \ldots, c_{\rho(G)} := Se^{-(\rho(G)-1)k}.$ (6c)

Intuitive choices for the parameters are $S = \rho(G)$ and $k = 1$.

The sphere-functional introduced above is of central interest in this paper, but the following eccentricity functional will be addressed in Section 6.

Definition 2.3. *The eccentricity $\sigma(v)$ of a vertex v is the greatest distance between v and any other vertex.*

Definition 2.4. *The eccentricity-functional f_σ is defined by*

$$\begin{aligned} &f_\sigma : V \to \mathbb{Z} \\ &f_\sigma(v) = \sigma(v). \end{aligned} \tag{7}$$

We now define several sub-classes of regular graphs, which will be of interest in Section 4. Recall that a (d)-regular graph is a graph, where every vertex has equal degree d. To our knowledge, the class of sphere-regular graphs (see Definition 2.5) has not been studied previously, while the class of distance-regular graphs has been introduced by [3] and studies thereof can be found in the literature.

Definition 2.5. *We call a graph* sphere-regular *if there exist positive integers $s_1, \ldots, s_\rho$, such that*

$$(|S_1(v)|, |S_2(v)|, \ldots, |S_{\rho(G)}(v)|) = (s_1, s_2, \ldots, s_\rho)$$

for all vertices $v \in V$, where ρ denotes the diameter $\rho(G)$ of G.

The following definition of a distance-regular graph is slightly different to the one given in [3], but it is easily seen that it is equivalent.

Definition 2.6. *A graph G is called* distance-regular *if for every $i = 0, \ldots, \rho(G)$ there exist integers b_i and c_i, such that for all pairs $v, w \in V$ with $d(v, w) = i$ holds*

$$|S_{i-1}(v) \cap S_1(w)| = b_i \text{ and} \tag{8a}$$

$$|S_{i+1}(v) \cap S_1(w)| = c_i. \tag{8b}$$

We will use the graph classes defined by Definitions 2.5 and 2.6 to find graphs of maximal entropy.

3 Software and Computation

Custom computer software has been developed and applied in order to determine the graphs among a group which maximize the entropy with regard to a given information functional. This has been achieved by calculating the entropy values for several graph sets (see Section 5) and displaying the graphs possessing extremal values (maximum and minimum).

In order to calculate the entropies, the existing routines from the `QuACN` package [15] based on the `R` programming language suggested a first implementation. This experimental early attempt turned out to be far too slow for an exhaustive analysis of larger data sets. As a meaningful optimization of the `R`-based solution would have been rather difficult, a different solution has been created in `C++`. The `LEMON` library has been chosen to represent the graphs at runtime, mainly because of its ability to read `graph6` files. These are readily available thanks to the `geng` tool from the `nauty` package [13], which has been used to generate the sets of all connected graphs (`geng -c $N`) and all trees (`geng -cbf $N $[N-1]`) of a given order `$N`. Our program loads one graph at a time, calculates its distance matrix and computes the entropy values. The algorithms to calculate the information functionals have been taken from `QuACN` and translated into `C++` code utilizing the STL.

The output of this program is a line-based text file with tab-separated fields for the considered information functionals, the diameter as well as an edge list representation for each graph. These intermediate output files, one per set, have been further processed and rendered with a `Python` script based on the `igraph` and `cairo` libraries. The final result is a collection of PDF files containing graphical representations of the extremal graphs, both for each set as a whole and for each subset of graphs with the same diameter.

4 Maximum entropy

In this section, we try to classify those graphs which return maximal entropy $I_{f_S}(G) = \log n$ for the sphere-functional and an arbitrary decreasing weight sequence. We will find that a full classification is only possible for special weight sequences such as the exponential sequence. In the following, we will speak about maximal graphs as those graphs having maximal entropy.

Proposition 4.1. *Every sphere-regular graph with n vertices has maximum entropy $I_f(G) = \log n$.*

Proof. Let $F = f(v) = \sum_{i=1}^{\rho(G)} c_i s_i$, then

$$I_f(G) = -\sum_{i=1}^{n} \frac{F}{nF} \log \frac{F}{nF} = -\log \frac{1}{n} = \log n.$$

□

Obviously, a trivial constraint for a sphere-regular graph is being regular, since the degree of a vertex corresponds to its first sphere. In contrast, regularity is not sufficient for sphere-regularity.

Proposition 4.2. *Not every regular graph is sphere-regular.*

Proof. The graph in Figure 1 proves the statement, since it is a regular graph with $d = 4$, but

$$S(v) = (4, 4, 4, 1),$$
$$S(w) = (4, 7, 2, 0).$$

□

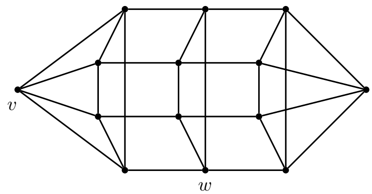

Figure 1: A regular graph which is not sphere-regular.

Still, a sufficient condition for sphere-regularity can be given using the notion of distance-regularity.

Theorem 4.3. *A distance-regular graph G is sphere-regular.*

Proof. We proof the statement inductively. For $S_1(v)$ let $i = 0$. Only a vertex v itself is at distance 0 from v, hence by the definition of distance regular graphs we have that

$$|S_1(v) \cap S_1(v)| = |S_1(v)| = c_0 = s_1$$

for all vertices v.

Now we assume that for all spheres $S_i(v)$ and all vertices v we have $S_i(v) = s_i$.

Note that, if $d(v, w) = i$, then

$$\begin{aligned} S_1(w) &= (S_1(w) \cap S_{i-1}(v)) \cup (S_1(w) \cap S_i(v)) \cup (S_1(w) \cap S_{i+1}(w)) \\ &= S_1(w) \cap (S_{i-1}(v) \cup S_i(v) \cup S_{i+1}(v)), \end{aligned} \tag{9}$$

because every neighbor of w is reachable in a maximum of $i+1$ steps via w, and if one of its neighbors was reachable in $i-j, j \geq 2$ steps, then w would be reachable in $i-j+1 \leq i-1$ steps via this vertex. Further, all vertices in $S_{i+1}(v)$ are neighbors of a vertex in $S_i(v)$, and they are those neighbor which are not at distance $i-1$ or i of v, that is

$$S_{i+1}(v) = \bigcup_{w \in S_i(v)} S_1(w) \backslash (S_i(v) \cup S_{i-1}(v)) = \bigcup_{w \in S_i(v)} (S_1(w) \cap S_{i+1}(v)).$$

By the inclusion-exclusion principle, this gives

$$|S_{i+1}| = \sum_{w \in S_i(v)} |S_1(w) \cap S_{i+1}(v)| + \sum_{\ell=2}^{|S_i(v)|} (-1)^{\ell+1} \bigcap_{\substack{(w_1,\ldots,w_\ell) \in S_i(v) \\ w_i \neq w_j}} S_1(w_i) \cap S_{i+1}(v).$$

Let $v \in V$ and $x \in V$ be at distance $i+1$ of v, i.e. $x \in S_{i+1}(v)$. Then, x must have at least one neighbor in $S_i(v)$, and by distance-regularity, $|S_i(v) \cap S_1(x)| = b_{i+1}$, that is, every vertex $x \in S_{i+1}(v)$ has a fixed number k_i of neighbors in $S_i(v)$, and hence it appears in $\binom{k_i}{\ell}$ intersections of spheres $S_{i+1}(v) \cap \bigcap_{j=1}^{\ell} S_1(w_j)$. Therefore,

$$\begin{aligned} |S_{i+1}(v)| &= \sum_{w \in S_i(v)} c_i + \sum_{\ell=2}^{k_i} (-1)^{\ell-1} | \bigcap_{w_1,\ldots,w_\ell \in S_i(v)} S_1(w_i) \cap S_{i+1}(v)| \\ &= c_i s_i + \sum_{\ell=2}^{k_i} (-1)^{\ell-1} |S_{i+1}(v)| \binom{k_i}{\ell}. \end{aligned}$$

Hence

$$|S_{i+1}(v)| = \frac{s_i c_i}{1 + \sum_{\ell=2}^{k_i} (-1)^{\ell} \binom{k_i}{\ell}}, \tag{10}$$

which is independent of v. Note therefore that the constant k_i is also determined by i. $\square$

Lemma 4.4. *Distance-regularity is not a necessary condition for sphere-regularity*

Proof. The statement is proven by the sphere-regular graph in Figure 2. For $i = 1$, we have

$$|S_2(3) \cap S_1(2)| = 2, \text{ but } |S_2(2) \cap S_1(1)| = 1.$$

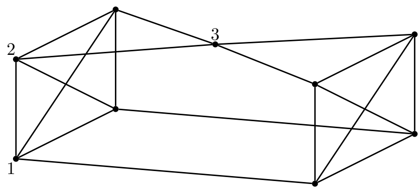

Figure 2: A sphere-regular graph which is not distance-regular.

$\square$

To summarize the results presented so far, we state the following Corollary.

Corollary 4.5. $\{$*regular graphs*$\} \supsetneqq \{$ *sphere-regular graphs* $\} \supsetneqq \{$ *distance-regular graphs*$\}$.

Apparently, the class of sphere-regular graphs is a novel class of graphs. Naturally, the question arises whether sphere-regular graphs are the only maximal graphs for I_{f_S}. The answer to this question depends heavily on the weight sequence used. The following Lemma answers the question for suitable sequences.

Lemma 4.6. *Sphere-regular graphs are the only maximal graphs for I_{f_S} when using a weight sequence such that there exist no numbers $a_i, i = 1, \ldots, \rho(G), a_i \in \mathbb{Z}$ with $\sum_{i=1}^{\rho(G)} a_i = 0$, where*

$$\sum_{i=1}^{\rho(G)} a_j c_j = 0. \tag{11}$$

Proof. Lets assume G is maximal but not sphere-regular, hence there is a vertex v with $S(v) = (s_1, s_2, \ldots, s_\rho)$ and a vertex w with $S(w) = (\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_\rho)$ with $s_i \neq \tilde{s}_i$ for at least one $i = 1, \ldots, \rho(G)$. Since the graph is maximal, we have

$$\sum_{i=1}^{\rho(G)} s_i c_i = \sum_{i=1}^{\rho(G)} \tilde{s}_i c_i,$$

and hence

$$\sum_{i=1}^{\rho(G)} (s_i - \tilde{s}_i) c_i = \sum_{i=1}^{\rho(G)} a_i c_i,$$

where the numbers a_i are integers which sum up to zero. □

An example of a sequence fulfilling the preliminaries of Lemma 4.6 is the exponential sequence, since obviously,

$$-a_i e^{-ik} \neq \sum_{j \neq i} a_j e^{-jk}$$

for any integers a_i, a_j.

For other (decreasing) sequences such as the linear or the quadratic sequence, linear combinations like equation (11) do exist (for the quadratic sequence there are fewer than for the linear one). It would be required to prove that no graph can actually take such a set of sphere-sequences, which can be disproved as follows.

Theorem 4.7. *There are maximal graphs with respect to I_{f_S} which are not sphere-regular.*

Proof. The graph depicted in Figure 3 is an example of a non sphere-regular maximal graph for I_{f_S}, using the linear sequence $c_1 = \rho(G), c_2 = \rho(G) - 1, \ldots, c_{\rho(G)} = 1$. It contains two types of nodes, namely nodes of type v and node of type w. We have

$$\begin{aligned} S(v) &= (4, 2, 2), \quad f_S(v) = 18, \\ S(w) &= (3, 4, 1), \quad f_S(w) = 18. \end{aligned} \tag{12}$$

□

Remark 4.1. *Note that even for the linear sequence, computations indicate that there is only a very small number of maximal graphs which are not sphere-regular. For graphs of size* 9*, for example, the graph depicted in Figure 3 is the only one among* 22 *maximal graphs.*

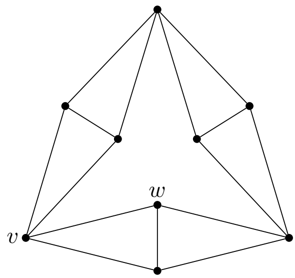

Figure 3: A maximal graph which is not regular.

In the following, we will present some restrictions on maximal graphs for I_{f_S}, which are valid for any decreasing weight sequence.

Lemma 4.8. *A graph of diameter* 2 *is maximal for* f_S *if and only if it is sphere-regular.*

Proof. We proved in Proposition 4.1 that a sphere-regular graph is maximal. Now assume G is a maximal graph with diameter $\rho(G) = 2$. Further assume it contains vertices v and w with different sphere-vectors, say $S(v) = (v_1, v_2)$ and $S(w) = (w_1, w_2)$ with $v_i \neq w_i$ for $i = 1, 2$. Since $v_1 + v_2 = w_1 + w_2 = (n-1)$ we have

$$\begin{aligned} c_1v_1 + c_2v_2 &= c_1w_1 + c_2w_2 \\ (n-1-v_2)c_1 + c_2v_2 &= (n-1-w_2)c_1 + w_2c_2 \\ (c_2-c_1)v_2 &= (c_2-c_1)w_2 \\ v_2 &= w_2, \end{aligned}$$

and also $v_1 = n-1-v_2 = n_1 - w_2 = w_1$. □

Lemma 4.9. *Maximal graphs cannot have unary nodes. Hence, in particular, trees cannot be maximal for* I_{f_S}.

Proof. Let v be a vertex of degree 1 and w be its unique neighbor. Further, let $S(w) = (w_1, w_2, \ldots, w_{d-1}, 0)$ be the sphere vector of w (The last entry has to be zero since any path of length r starting in w gives a path of length $r+1$ starting in v. Then,

$$s(v) = (v_1, \ldots, v_d) = (1, w_1 - 1, w_2, w_3, \ldots, w_{d-1}).$$

Given that the two vectors have the same weighted sum, we have

$$\begin{aligned} c_1 + (w_1-1)c_2 + c_3w_2 + \cdots c_dw_{d-1} &= c_1w_1 + c_2w_2 + \cdots + c_{d-1}w_{d-1} \\ (w_1-1)(c_1-c_2) + w_2(c_3-c_2) + \cdots + w_{d-1}(c_{d-1}-c_d) &= 0. \end{aligned} \tag{13}$$

This is a contradiction since all terms $c_i - c_{i+1}$ are positive due to the monotonicity of the sequence and $w_1 \geq 2$ if $n > 2$. □

Corollary 4.10. *The last nonzero entries of the sphere-sequence of a vertex of a maximal graph cannot be* 2 *or more consecutive ones.*

Proof. Let $\sigma(v)$ be the eccentricity of vertex v. If there is only one vertex at distance $\sigma(v) - 1$ and one vertex w at distance $\sigma(v)$ of v, then w must be a leaf. □

Lemma 4.11. *A maximal graph different from the complete graph K_n cannot contain a vertex of degree $n - 1$.*

Proof. A vertex of degree $n - 1$ has sphere-vector $S(v) = (n - 1, 0, \ldots, 0)$. For a vertex w with degree sequence $S(w) = (s_1, s_2, \ldots, s_\rho)$ we have

$$(n - 1 - s_1)c_1 = \sum_{i=2}^{\rho(G)} s_i c_i \leq (n - 1 - s_1)c_2,$$

since $c_1 > c_2$ the statement is proven. □

5 Minimum entropy

Very little is known about minimal entropy graphs. The absolute minimum of entropy is taken at the probability distribution $\mathbf{p} = (1, 0, 0, \ldots, 0)$, which is clearly not taken by any graph since the functional $f_S(v)$ never returns 0. Due to the limitations in size of computations on exhaustively generated graph classes [5], our numerical results provide no clear picture for universal conjectures.

Computations (see Section 3) give the minimal graphs of size 8 and 9 depicted in Figure 4 for I_{f_S}, where the left graph is minimal for the linear sequence and the right graph is minimal for the exponential sequence. Note therefore that the set N_8 of all non-isomorphic graphs on 8 vertices contains 11117 graphs, N_9 contains 261080 graphs and the exhaustively generated set on 10 vertices, N_{10}, has 11716571 graphs. Similarly sizes of exhaustive graph sets explode for larger numbers of vertices. Note that these graph classes have been recently used by [5] to determine the discrimination ability of topological graph measures.

Results for the linear sequence show that the minimal graph is not necessarily a tree, while for the exponential sequence the minimal graphs are closely related and it seems reasonable to conjecture that a minimal graph is a tree. Computations on exhaustively

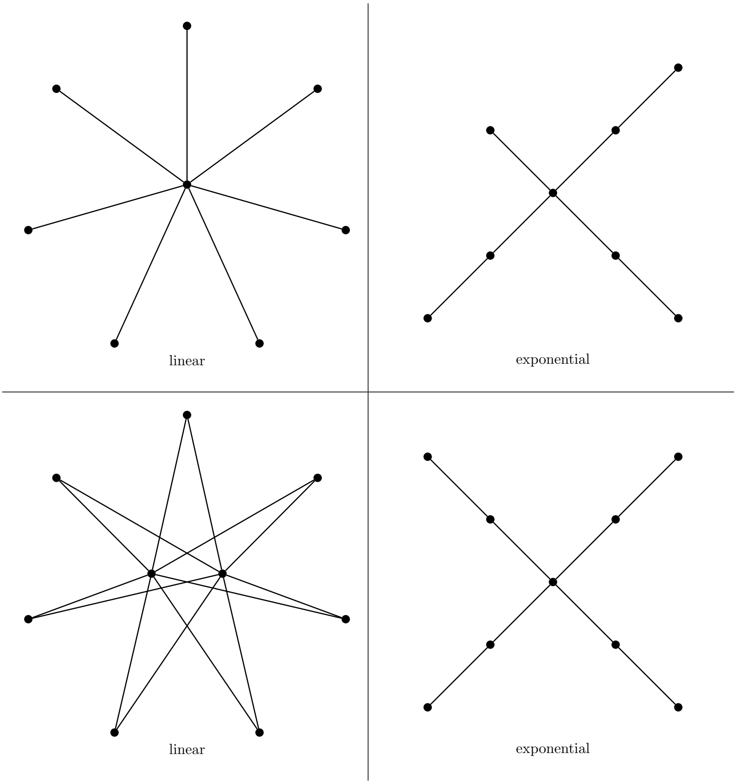

Figure 4: Minimal graphs of sizes 8 and 9.

generated classes of trees are less costly than those on all graphs, hence further computations on exhaustively generated trees up to size 20 provide a conjecture. We therefore call a tree a generalized star, if it has on central vertex, and a number of branches emerging from this vertex whose lengths differ by at most 1. Hence, the minimal graph for I_{f_S} with the exponential sequence is a generalized star for sizes 8 and 9, and so are the minimal trees up to size 20, as computations on trees show. In Figure 5 we depict the minimal tree on 20 vertices.

Conjecture 5.1. *The minimal graph for I_{f_S} with the exponential sequence is a tree. Further it is a generalized star of diameter approximately $\sqrt{2n}$ and, hence, with approximately $\sqrt{2n}$ branches.*

For the linear sequence, note that both minimal graphs are bipartite, in fact they are complete bipartite graphs $K_{1,7}$ and $K_{2,7}$, respectively. Still, further computations would

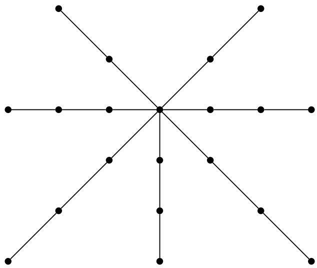

Figure 5: The minimal tree of size 20.

be required to determine the development of the minimal graph, which is, as mentioned previously, very costly on time and computational power due to the rapidly increasing sizes of sets of exhaustively generated graphs.

6 The eccentricity-based functional

In this section, we will focus on the second functional introduced in Section 2, namely the distance-functional f_σ, see Definition 7. The quite simple definition of the functional intuitively suggests that maxima and minima of I_{f_σ} might be easier to determine as those of I_{f_S}. In fact, this is not the case due to the high degeneracy [5] of the measure, that is, a lot of graphs which are not necessarily structurally related dispose of identical information content for I_{f_σ}.

6.1 Maximal graphs

We first present an elementary result on maximal graphs with respect to f_σ.

Lemma 6.1. *(i) A graph G is maximal with respect to $I_{f_\sigma}(G)$ if and only if every of its vertices is endpoint of a maximal path in G.*

(ii) A maximal graph different from the complete graph K_n cannot contain a vertex of degree $n-1$.

Proof. The first statement is obvious since to obtain a uniform distribution every vertex has to have the same eccentricity $\sigma(v) = \rho(G)$. There exists at least one graph with this property, namely the complete graph K_n. Hence the absolute maximum of the entropy, $\log n$, is indeed taken by a graph. The second statement follows immediately since a vertex of degree $n-1$ has eccentricity 1, but K_n is the only graph of diameter 1. □

Unfortunately, the above conditions are fulfilled by a large number of graphs, as mentioned in the following remark, which seemingly do not have further structural properties in common, a similarity is mentioned in the following.

Remark 6.1. *Computations indicate that there are* > 1500 *maximal graphs in* N_8 *and most probably more than* 50000 *in* N_9. *Also, note that many of these graphs are little connected, i.e. they contain a lot of vertices of degree* 2.

In Figure 6, we depict one of the large number of graphs on 9 vertices which are maximal for the functional f_σ. This graph can be considered a typical representant with a high number of vertices of degree 2. Interestingly, this graph minimizes the entropy with respect to the functional f_S together with a linear weight sequence, see Section 5.

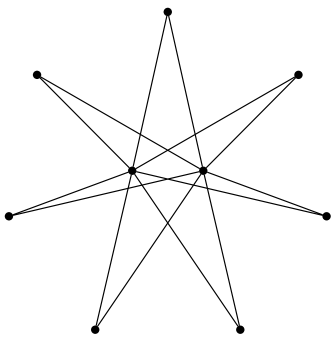

Figure 6: A maximal graph in N_9.

6.2 Minimal graphs

Based on the numerical results, we make the following conjecture for a minimal graph with respect to f_σ. We depict the minimal graphs on 8 vertices in Figure 7.

Conjecture 6.2. *A minimal graph for* I_{f_σ} *is a highly connected graph, i.e. it is a complete graph* K_n *where a small number of edges have been removed. In particular, we conjecture that a minimal graph on* n *vertices will have* $m \geq \frac{n}{2}$ *vertices of degree* $n-1$.

A graph described in Conjecture 6.2 has diameter 2, since a maximal path contains a vertex of degree $n-1$ at step 1. The vector of functionals is

$$\mathbf{f}(V) = (\underbrace{1, 1, \ldots, 1}_{m}, \underbrace{2, \ldots, 2}_{n-m}).$$

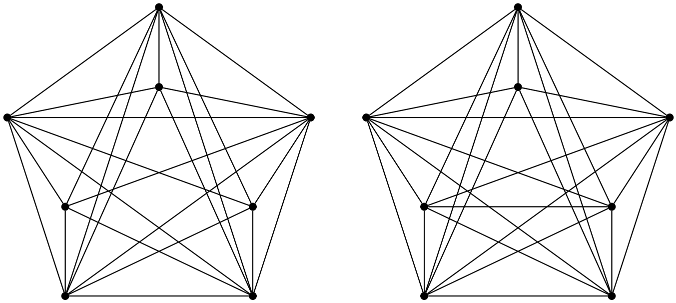

Figure 7: The minimal graphs of N_8.

Note that there is only a very small number of minimal graphs, especially in comparison to the number of maximal graphs. For N_8 and N_9, computations show that there are 2 minimal graphs. For $n = 8$, they are depicted in Figure 7, for $n = 9$ they contain 5 vertices of degree 8 each.

Further note that the range of values of I_{f_3} is quite small, i.e.

$$I_{\max}(N_9) - I_{\min}(N_9) \approx 0.085.$$

Even by performing thorough computations, we could not provide analytic results or strong conjectures for the case of f_σ. Therefore, note that the complete graph is a maximal graph for I_{f_σ}, while the complete graph K_8, where only two edges have been removed, has minimal entropy. The graph depicted in Figure 6 can be obtained from those in Figure 7 by further removal of edges, but returns maximal entropy. Hence, we can conclude that extremal behavior of information content based on the eccentricity functional will not be describable by edge-removal operations.

7 Summary and Conclusion

Now we summarize the results obtained in this paper. We have successfully classified the class of graphs which take the maximal value by using the entropy based on i-spheres and the exponential sequence (see Equations 4, 5). We further realized that the considerations are not completely true for a linear sequence, which indicates that the exponential sequence might also provide better uniqueness properties for non-extremal graphs. In parts, this has already been confirmed in a previous study to explore the discrimination ability (uniqueness) of entropic graph measures, see [5]. We found that for proving statements to find minimal graphs, the problem is much harder, and analytic proofs seem out of reach at the present state. A reason for this is surely the complexity of the used graph entropies. However, they have been proven useful in several disciplines such as structural chemistry and computational linguistics, see [9, 1].

By performing further computations on a large scale, we could get deeper insights to better understand the complexity of the problem. If we study the literature regarding contributions to prove extremal properties by using non-information-theoretic measures, e.g., see [10], we see that the problem of exploring extremal properties of parametric graph entropies remains quite intricate.

Acknowledgement: Matthias Dehmer and Veronika thank the Austrian Science Foundation (FWF) for supporting this work (project P22029-N13).

References

[1] O. Abramov, T. Lokot, Typology by means of language networks. Applying information theoretic measures to morphological derivation networks, in: M. Dehmer, F. Emmert–Streib, A. Mehler (Eds.), *Towards an Information Theory of Complex Networks: Statistical Methods and Applications*, Birkhäuser, Boston/Basel, 2011, pp. 321–346.

[2] D. Bonchev, N. Trinajstić, Information theory, distance matrix and molecular branching, *J. Chem. Phys.* **67** (1977) 4517–4533.

[3] A. Brouwer, A. Cohen, A. Neumaier, *Distance Regular Graphs*, Springer, Berlin, 1989.

[4] M. Dehmer, Information processing in complex networks: Graph entropy and information functionals, *Appl. Math. Comput.* **201** (2008) 82–94.

[5] M. Dehmer, M. Grabner, K. Varmuza, Information indices with high discriminative power for graphs, *PLoS ONE* **7** (2012) e31214.

[6] M. Dehmer, V. Kraus, On extremal properties of graph entropies, *MATCH Commun. Math. Comput. Chem.* **68** (2012) 889–912.

[7] M. Dehmer, A. Mowshowitz, A history of graph entropy measures, *Inf. Sci.* **1** (2011) 57–78.

[8] M. Dehmer, L. Sivakumar, K. Varmuza, Uniquely discriminating molecular structures using novel eigenvalue–based descriptors, *MATCH Commun. Math. Comput. Chem.* **67** (2012) 147–172.

[9] M. Dehmer, K. Varmuza, S. Borgert, F. Emmert–Streib, On entropy–based molecular descriptors: Statistical analysis of real and synthetic chemical structures, *J. Chem. Inf. Model.* **49** (2009) 1655–1663.

[10] B. Furtula, I. Gutman, M. Ivanović, D. Vukičević, Computer search for trees with minimal *ABC* index, *Appl. Math. Comput.* **219** (2012) 767–772.

[11] J. Körner, Coding of an information source having ambiguous alphabet and the entropy of graphs, *Transactions of the 6-th Prague Conference on Information Theory*, 1973, pp. 411–425.

[12] X. Li, I. Gutman, *Mathematical Aspects of Randić–Type Molecular Structure Descriptors*, Univ. Kragujevac, Kragujevac, 2006.

[13] B. D. McKay, Nauty, http://cs.anu.edu.au/∼bdm/nauty/ (2010).

[14] A. Mowshowitz, Entropy and the complexity of the graphs I: An index of the relative complexity of a graph, *Bull. Math. Biophys.* **30** (1968) 175–204.

[15] L. A. J. Müller, K. G. Kugler, A. Dander, A. Graber, M. Dehmer, QuACN - An R package for analyzing complex biological networks quantitatively, *Bioinformatics* (2010) 140–141.

[16] M. Randić, On characterization of molecular branching, *J. Amer. Chem. Soc.* **97** (1975) 6609–6615.

[17] N. Rashevsky, Life, information theory and topology, *Bull. Math. Biol.* **17** (1955) 229–235.

[18] C. Shannon, W. Weaver, *The Mathematical Theory of Communication*, Univ. Illinois Press, Champaign, 1949.

[19] E. Trucco, A note on the information content of graphs, *Bull. Math. Biol.* **18** (1956) 129–135.

[20] H. Wiener, Structural determination of paraffin boiling points, *J. Amer. Chem. Soc.* **69** (1947) 17–20.

[21] B. Zhou, Bounds on the Balaban index, *Croat. Chem. Acta* **81** (2008) 319–323.

[22] B. Zhou, The (n, n)-graphs with the first three extremal Wiener Indices, *J. Math. Chem.* **43** (2008) 60–74.