

Estimating the Spectral Radius of a Graph by the Second Zagreb Index

HOSAM ABDO, DARKO DIMITROV

*Institute of Computer Science, Freie Universität Berlin, Takustraße 9, D-14195 Berlin,
Germany*

abdo@mi.fu-berlin.de, darko@mi.fu-berlin.de

TAMÁS RÉTI

Óbuda University, Bécsi út 96/B, H-1034 Budapest, Hungary

reti.tamas@bgk.uni-obuda.hu

DRAGAN STEVANOVIĆ¹

*Mathematical Institute, Serbian Academy of Science and Arts, Knez Mihajlova 36, Belgrade,
Serbia, and University of Primorska, Institute Andrej Marušič, Muzejjski trg 2, Koper, Slovenia*
dragance106@yahoo.com

(Received June 25, 2014)

Abstract

The *spectral radius* of the adjacency matrix of a molecular graph is a topological index that is related to the branching of the molecule. We show that the spectral radius can be very accurately estimated by another topological index, the *second Zagreb index*.

¹Corresponding author.

1 Introduction

Quantitative measures of branching of a molecule can be a valuable help in determining relations between the molecule's structure and its physico-chemical properties, which is the main task of QSAR (quantitative structure-activity relations) and QSPR (quantitative structure-property relations) studies. Mathematical quantities that describe the structure or shape of molecules are known as molecular descriptors [1]. Among them, so-called topological indices [2] play a significant role. Here, we consider two topological indices that are related to the molecular branching and are frequently used in QSAR and QSPR researches: the *spectral radius* and the *second Zagreb index* of a (molecular) graph.

The eigenvalues of a simple graph G are the eigenvalues of its adjacency matrix A , a $(0, 1)$ -matrix indexed by graph vertices that describes whether any two given vertices u, v are adjacent ($A_{uv} = 1$) or not ($A_{uv} = 0$). The largest eigenvalue, denoted by λ_1 , is called the spectral radius of G . Using the spectral radius of G as a measure of branching was proposed by Cvetković and Gutman in 1977 [3] and studied later in several works including [4–6].

For a simple graph $G = (V, E)$ with $n = |V|$ vertices and $m = |E|$ edges, the first Zagreb index $M_1(G)$ and the second Zagreb $M_2(G)$ index are defined as

$$M_1(G) = \sum_{v \in V} d(v)^2 \quad \text{and} \quad M_2(G) = \sum_{uv \in E} d(u)d(v),$$

where $d(v)$ is the degree of a vertex $v \in V$. The first and second Zagreb indices were introduced in 1972 by Gutman and Trinajstić [7] within the study of the dependence of total π -electron energy of molecular structure. In [8], it was shown that these terms are measures of branching of the molecular carbon-atom skeleton. For more details on the first and second Zagreb indices see the recent review [9].

Here, we propose an estimate of the spectral radius of a simple graph that depends on the second Zagreb index and the number of edges of the graph. The proposed estimate is

$$\lambda_1 \approx \sqrt{\frac{M_2(G)}{m}},$$

or shortly $\lambda_1 \approx \sqrt{M_2/m}$.

In Fig. 1 three molecular graphs: (a) coronene ($C_{24}H_{12}$), (b) benzene (C_6H_6), and (c) 2,3-dimethylpentane (C_7H_{16}), are depicted. It is fairly easy to obtain that M_2 , m and $\sqrt{M_2/m}$ in those cases are: (a) 204, 30, 2.6077, (b) 24, 6, 2, and (c) 26, 6, 2.0817,

respectively. On the other hand, the spectral radii in these cases are: (a) 2.6751, (b) 2, and (c) 2.0529. As it is evident from these examples, $\sqrt{M_2/m}$ is neither a lower, nor an upper bound on λ_1 .

2 Statistical tests

In order to test the quality of $\sqrt{M_2/m}$ as an estimate of λ_1 , we calculated the root-mean-square deviation between the values of λ_1 and $\sqrt{M_2/m}$ within the sets of small graphs:

- (a) connected graphs from 5 to 10 vertices (a total of 11,989,754 graphs),
- (b) connected graphs from 5 to 12 vertices and maximum vertex degree 4 (a total of 7,643,962 graphs),
- (c) trees from 10 to 20 vertices (a total of 1,345,929 trees),
- (d) trees from 10 to 20 vertices and maximum vertex degree 4 (a total of 617,975 trees),

and also within the sets of larger graphs, namely the Barabási-Albert preferential attachment graphs [10] on 50, 100, 150, 200, 250 and 300 vertices, respectively, in which:

- (e) each new node was preferentially connected to two existing nodes (1,000 graphs for each order), and
- (f) each new node was preferentially connected to four existing nodes (1,000 graphs for each order).

The root-mean-square deviation between $\sqrt{M_2/m}$ and λ_1 ranged from 0.01414 to 0.02952 for connected graphs, from 0.01414 to 0.05721 for bounded degree graphs, from 0.04201 to 0.10912 for trees and from 0.03472 to 0.09089 for bounded degree trees. For Barabási-Albert graphs, the root-mean-square deviation ranged from 0.08917 for graphs on 50 vertices to 0.47284 for graphs on 300 vertices.

For comparison purposes, we selected ten further estimates for λ_1 , mostly lower or upper bounds from literature, based on the simplicity of their expressions. Let $d_{2,u} = \sum_{v \in N_u} d_v$ be the sum of degrees of the neighbors of $u \in V$, and let $M_1 = \sum_{u \in V} d_u^2$ be the so-called first Zagreb index [9]. The following estimates of λ_1 were selected:

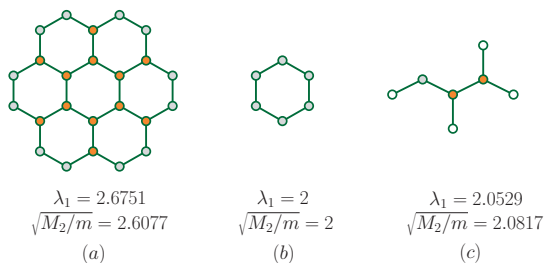


Figure 1: Examples of molecular graphs for which λ_1 is (a) larger than, (b) equal to or (c) smaller than $\sqrt{M_2/m}$.

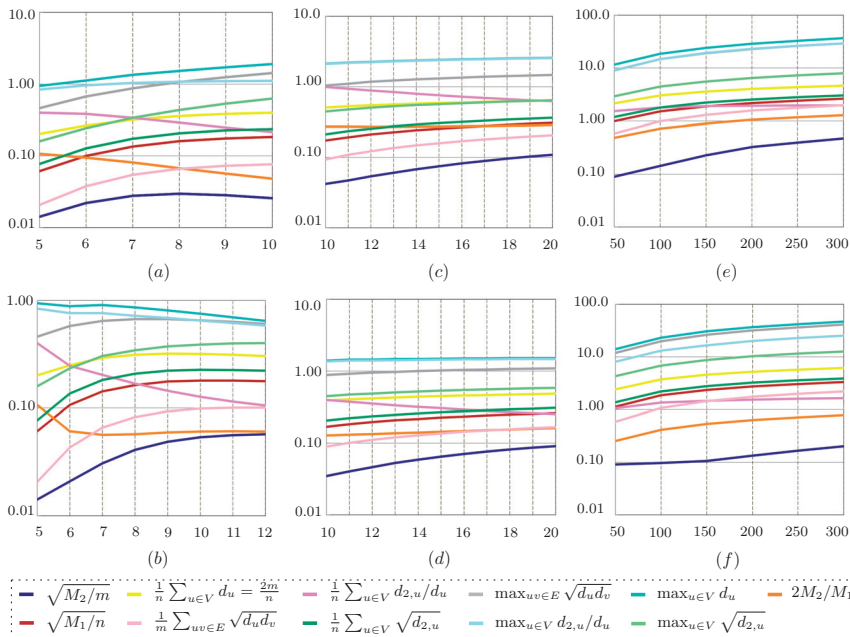


Figure 2: Diagrams of root-mean-square deviation for $\sqrt{M_2/m}$ and ten other estimates in sets of: (a) connected graphs, (b) connected graphs with maximum vertex degree four, (c) trees, (d) trees with maximum vertex degree four, (e) Barabási-Albert preferential attachment graphs in which each new node is connected to two existing nodes, and (f) Barabási-Albert preferential attachment graphs in which each new node is connected to four existing nodes. On the x-axes are the numbers of vertices of considered graphs; for details see text. Note that in all cases examined, the blue line lies significantly below the other lines.

- (a) $\frac{1}{n} \sum_{u \in V} d_u = \frac{2m}{n}$ [11],
- (b) $\frac{1}{n} \sum_{u \in V} d_{2,u}/d_u$,
- (c) $\frac{1}{n} \sum_{u \in V} \sqrt{d_{2,u}}$,
- (d) $\frac{1}{m} \sum_{uv \in E} \sqrt{d_u d_v}$ [12],
- (e) $\sqrt{M_1/n}$ [13],
- (f) $\max_{u \in V} d_u$ [11],
- (g) $\max_{u \in V} d_{2,u}/d_u$ [13],
- (h) $\max_{u \in V} \sqrt{d_{2,u}}$ [13],
- (i) $\max_{uv \in E} \sqrt{d_u d_v}$ [14],
- (j) $2M_2/M_1$ [15].

The root-mean-square deviation between all these estimates and the values of λ_1 for the above mentioned graph sets is depicted in Fig. 2. The diagrams in Fig. 2 are rather self-explanatory: $\sqrt{M_2/m}$ is the best among all 11 estimates in all graph sets. The next two estimates, that tend to take up the 2nd and the 3rd place throughout these graph sets, are $2M_2/M_1$ and $\frac{1}{m} \sum_{uv \in E} \sqrt{d_u d_v}$.

3 Common lower and upper bounds

A possible explanation for high similarity of λ_1 and $\sqrt{M_2/m}$ may stem from the fact that λ_1 and $\sqrt{M_2/m}$ have common and similar lower and upper bounds. As for the common lower bound, λ_1 is the supremum of the Rayleigh quotients

$$\lambda_1 = \sup_{x \neq 0} \frac{x^T A x}{x^T x} = \sup_{x \neq 0} \frac{2 \sum_{uv \in E} x_u x_v}{\sum_{u \in V} x_u^2}.$$

For a particular choice $x = (\sqrt{d_u})_{u \in V}$ we have that

$$\lambda_1 \geq \frac{2 \sum_{uv \in E} \sqrt{d_u d_v}}{\sum_{u \in V} d_u} = \frac{\sum_{uv \in E} \sqrt{d_u d_v}}{m}$$

due to $\sum_{u \in V} d_u = 2m$ (probably the earliest appearance of this bound in the literature was in [12, Corollary 4.5]). On the other hand, Cauchy-Schwarz inequality applied to sequences $(\sqrt{d(u)d(v)})_{uv \in E}$ and $(1)_{uv \in E}$ yields

$$\sqrt{M_2 m} = \sqrt{\left(\sum_{uv \in E} \sqrt{d_u d_v} \right)^2 \left(\sum_{uv \in E} 1^2 \right)} \geq \sum_{uv \in E} \sqrt{d_u d_v}.$$

Hence, both λ_1 and $\sqrt{M_2/m}$ are bounded from below by the average value of the sequence $(\sqrt{d(u)d(v)})_{uv \in E}$.

The common upper bound is provided by the maximum of the same sequence. As in [14, Lemma 2.1], let x be a positive eigenvector of A corresponding to λ_1 , and let $s, t \in V$ be such that $x_s = \max_{u \in V} x_u$ and $x_t = \max_{\{v: sv \in E\}} x_v$. Then from $\lambda_1 x = Ax$

follows

$$\begin{aligned} \lambda_1 x_s &= \sum_{\{v: sv \in E\}} x_v \leq d_s x_t, \\ \lambda_1 x_t &= \sum_{\{u: tu \in E\}} x_u \leq d_t x_s, \end{aligned}$$

wherefrom $\lambda_1^2 \leq d_s d_t$ and $\lambda_1 \leq \sqrt{d_s d_t} \leq \max_{uv \in E} \sqrt{d_u d_v}$. Further, trivially

$$\sqrt{\frac{M_2}{m}} = \sqrt{\frac{\sum_{uv \in E} d_u d_v}{m}} \leq \max_{uv \in E} \sqrt{d_u d_v}.$$

The previous two paragraphs can be summarized as

$$\frac{\sum_{uv \in E} \sqrt{d_u d_v}}{m} \leq \left\{ \lambda_1, \sqrt{\frac{M_2}{m}} \right\} \leq \max_{uv \in E} \sqrt{d_u d_v}.$$

4 Case of equality

Equality $\lambda_1 = \sqrt{M_2/m}$ is satisfied for several classes of graphs defined in terms of A and the all-one vector j :

- regular graphs, which satisfy $Aj = \lambda_1 j$,
- harmonic graphs [16], which satisfy $A^2 j = \lambda_1 A j$, and
- semiharmonic graphs [16], which satisfy $A^3 j = \lambda_1^2 A j$.

Note that a regular graph is also harmonic, and that a harmonic graph is also semiharmonic, so that it is enough to consider semiharmonic graphs only. The vector $d = (d_u)_{u \in V}$ of vertex degrees satisfies $d = A j$, so that

$$M_2 = \sum_{uv \in E} d_u d_v = \frac{1}{2} d^T A d = \frac{1}{2} j^T A^3 j.$$

Since further $m = \frac{1}{2} j^T A j$, we have that in a semiharmonic graph holds

$$\sqrt{\frac{M_2}{m}} = \sqrt{\frac{j^T A^3 j}{j^T A j}} = \sqrt{\frac{j^T (\lambda_1^2 A j)}{j^T A j}} = \lambda_1.$$

Let us recall that a graph is bipartite if its vertex set can be partitioned as $V = V_1 \cup V_2$, such that each of its edges joins vertices from different parts. Equality $\lambda_1 = \sqrt{M_2/m}$ is satisfied for two further classes of bipartite graphs as well:

- semiregular graphs, which satisfy $(Aj)_u = \begin{cases} p_1, & u \in V_1 \\ p_2, & u \in V_2 \end{cases}$ for some p_1, p_2 , and
- pseudosemiregular graphs [17], which satisfy $(A^2j)_u = \begin{cases} p_1(Aj)_u, & u \in V_1 \\ p_2(Aj)_u, & u \in V_2 \end{cases}$ for some p_1, p_2 .

Note that a semiregular graph is also pseudosemiregular, so that it is enough to consider pseudosemiregular graphs only. For a pseudosemiregular graph, if $u \in V_1$ we have

$$\begin{aligned} (A^3j)_u &= \sum_{v \in N_u} (A^2j)_v = p_2 \sum_{v \in N_u} (Aj)_v \\ &= p_2 \sum_{v \in V} A_{uv}(Aj)_v = p_2(A^2j)_u = p_2p_1(Aj)_u, \end{aligned}$$

while if $u \in V_2$ we have

$$\begin{aligned} (A^3j)_u &= \sum_{v \in N_u} (A^2j)_v = p_1 \sum_{v \in N_u} (Aj)_v \\ &= p_1 \sum_{v \in V} A_{uv}(Aj)_v = p_1(A^2j)_u = p_1p_2(Aj)_u. \end{aligned}$$

Hence, $A^3j = p_1p_2Aj$, so that pseudosemiregular graph is also semiharmonic. Here, Aj is a positive eigenvector corresponding to the eigenvalue p_1p_2 of the nonnegative matrix A^2 , so that $p_1p_2 = \lambda_1^2$ and, thus, $\sqrt{M_2/m} = \sqrt{p_1p_2} = \lambda_1$.

In addition, equality $\lambda_1 = \sqrt{M_2/m}$ may be satisfied for graphs that need not be semiharmonic. An example of such a graph is shown in Fig. 3, which has $\lambda_1 = 3$, $M_2 = 81$, and $m = 9$.

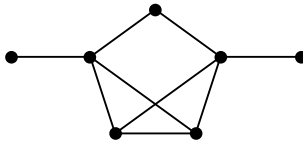


Figure 3: An example of a non-semiharmonic graph with $\lambda_1 = \sqrt{M_2/m}$.

A careful reader may have noticed in the above cases that M_2 is divisible by m . Indeed, this is true for any graph satisfying $\lambda_1 = \sqrt{M_2/m}$. Namely, the characteristic polynomial of G is a monic polynomial with integer coefficients, so that λ_1 is an algebraic integer. Consequently, λ_1^2 is an algebraic integer as well. On the other hand, $\lambda_1^2 = M_2/m$ is also a rational number, and it is well-known that the only algebraic integers in the set of rational numbers are integers. Thus, λ_1^2 is an integer, so that M_2 is divisible by m .

5 Outliers

Although statistical tests in Section 2 show that the root-mean-square deviation between λ_1 and $\sqrt{M_2/m}$ is consistently small across various graph sets, it is important to notice that the root-mean-square deviation, as the quadratic mean of the differences $\lambda_1 - \sqrt{M_2/m}$ for all graphs in a particular set, represents a sort of average value. As a consequence, this does not mean that the difference $|\lambda_1 - \sqrt{M_2/m}|$ will be small for each graph—as a matter of fact, it can be arbitrarily large for graphs with special structure. We have identified the double stars and the kites as examples of graphs with large discrepancy between λ_1 and $\sqrt{M_2/m}$.

The double star $DS_{a,a}$ is obtained from two copies of the star $K_{1,a}$, by connecting their centers with a new edge. It has $2a + 1$ edges and its second Zagreb index is $M_2(DS_{a,a}) = (a + 1)^2 + 2a(a + 1) = (a + 1)(3a + 1)$, so that

$$\sqrt{M_2/m} = \sqrt{\frac{(a + 1)(3a + 1)}{2a + 1}}.$$

The spectral radius λ_1 of $DS_{a,a}$ may also be determined theoretically. As all the leaves are similar to each other and two centers are similar to each other, we conclude that the principal eigenvector of $DS_{a,a}$ contains only two different components— l for the leaves and c for the centers. The eigenvalue equation at a leaf and at a center then yields

$$\lambda_1 l = c, \quad \text{and} \quad \lambda_1 c = c + al.$$

This system yields

$$\lambda_1^2 = \lambda_1 + a,$$

whose solutions are

$$\lambda_1 = \frac{1 \pm \sqrt{1 + 4a}}{2} = \pm \sqrt{a + \frac{1}{4}} + \frac{1}{2}.$$

The spectral radius of $DS_{a,a}$ is, of course, equal to the larger value (but note that the smaller value is also an eigenvalue of $DS_{a,a}$), so that if we know let $a \rightarrow \infty$, we obtain that

$$\lim_{a \rightarrow \infty} \frac{\sqrt{M_2/m}}{\lambda_1} = \sqrt{\frac{3}{2}}.$$

Therefore, for large value of a , the value of $\sqrt{M_2/m}$ will be approximately 22.47% larger than λ_1 for the double star $DS_{a,a}$. However, if in the above example the maximal degree does not exceed four, which is a case with molecular graphs, then the value of $\sqrt{M_2/m}$ will be at most 3.81% larger than λ_1 .

Another example is the kite graph $KP_{r,s}$, obtained from the complete graph K_r and the path P_s by identifying an endvertex of P_s with a vertex of K_r . A very small interval to which the spectral radius of $KP_{r,s}$, $r \geq 3$, belongs, has been found in [18]:

$$r - 1 + \frac{1}{r^2} + \frac{1}{r^3} < \lambda_1(KP_{r,s}) < r - 1 + \frac{1}{4r} + \frac{1}{r^2 - 2r}.$$

Hence, $\lim_{r \rightarrow \infty} \lambda_1(KP_{r,s})/(r - 1) = 1$, and, more importantly, it does not depend on s . The second Zagreb index of $KP_{r,s}$ is equal to

$$M_2 = \frac{1}{2}(r - 1)^3(r - 2) + r(r - 1)^2 + 2r + 4s - 6,$$

while it has

$$m = \frac{1}{2}r(r - 1) + s - 1$$

edges. If we now choose s to be much larger than r , then $\sqrt{M_2/m}$ will tend to a constant. For example, if we set $s = r^4$, then

$$\lim_{r \rightarrow \infty} \sqrt{\frac{M_2}{m}} = \sqrt{4\frac{1}{2}},$$

and

$$\lim_{r \rightarrow \infty} \frac{\sqrt{M_2/m}}{\lambda_1} = \lim_{r \rightarrow \infty} \frac{\sqrt{4\frac{1}{2}}}{r - 1} = 0.$$

Therefore, for large value of r and even larger value of $s(=r^4)$, the value of $\sqrt{M_2/m}$ will be negligible with respect to λ_1 for the kite graph $KP_{r,s}$. We would like to note that in this case, if we consider that the maximal degree is at most four, then the lower bound of $\sqrt{M_2/m}/\lambda_1$ is approximately 0.627451, obtained for $r=4$ and $s \rightarrow \infty$.

6 Conclusions

We have shown that the largest eigenvalue λ_1 of adjacency matrix of a simple graph can be very well estimated with a simple expression $\sqrt{M_2/m}$. Statistical tests have shown that this estimate consistently has smaller root-mean-square deviation than ten other estimates of λ_1 for sets of graphs with up to 10 vertices, trees with up to 20 vertices and sets of Barabási-Albert graphs with up to 300 vertices. Theoretical explanation for the quality of this estimate has been sought in the fact that both λ_1 and $\sqrt{M_2/m}$ are situated between the average and the maximum value of the sequence $(\sqrt{d_u d_v})_{uv \in E}$, and this argument was further supported by the fact that the equality between λ_1 and $\sqrt{M_2/m}$

holds for a number of graph classes. Certainly, as with any statistics, there exist outliers for which $\sqrt{M_2/m}$ can be very different from λ_1 , but it appears that they are quite rare. Moreover, it turns out that the deviations of the outliers are significantly smaller when one considers graphs with maximal degree four, which is a case with molecular graphs.

Acknowledgment

Dragan Stevanović is supported by Research Programme P1-0285 and Research Project J1-4021 of the Slovenian Research Agency, and Research Grant ON174033 of the Ministry of Education and Science of Serbia.

References

- [1] R. Todeschini, V. Consonni, *Molecular Descriptors for Chemoinformatics*, Wiley-VCH, Weinheim, 2009.
- [2] J. Devillers, A.T. Balaban (Eds.), *Topological Indices and Related Descriptors in QSAR and QSPR*, Wiley-VCH, Amsterdam, 1999.
- [3] D.M. Cvetković, I. Gutman, Note on branching, *Croat. Chem. Acta* **47** (1977) 115–121.
- [4] M. Randić, M. Vračko, M. Novič, Eigenvalues as molecular descriptors, in: M.V. Diudea (Ed.), *QSPR/QSAR Studies by Molecular Descriptors*, Nova, Huntington, 2001, pp. 147–211.
- [5] I. Gutman, D. Vidović, Two early branching indices and the relation between them, *Theor. Chem. Acc.* **108** (2002) 98–102.
- [6] M. Fischermann, I. Gutman, A. Hoffmann, D. Rautenbach, D. Vidović, L. Volkman, Extremal chemical trees, *Z. Naturforsch.* **57a** (2002) 49–52.
- [7] I. Gutman, N. Trinajstić, Graph theory and molecular orbitals. Total π -electron energy of alternant hydrocarbons, *Chem. Phys. Lett.* **17** (1972) 535–538.
- [8] I. Gutman, B. Ruščić, N. Trinajstić, C.F. Wilcox, Graph theory and molecular orbitals. XII. Acyclic polyenes, *J. Chem. Phys.* **62** (1975) 3399–3405.
- [9] I. Gutman, Degree-based topological indices, *Croat. Chem. Acta* **86** (2013) 351–361.
- [10] A.L. Barabási, R. Albert, Emergence of scaling in random networks, *Science* **286** (1999) 509–512.
- [11] D. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs—Theory and Applications*, Johann Ambrosius Barth Verlag, Heidelberg-Leipzig, 1995.
- [12] A.J. Schwenk, R.J. Wilson, Eigenvalues of Graphs, in: L.W. Beineke, R.J. Wilson (Eds.), *Selected topics in graph theory*, Academic Press, New York, 1978, pp. 307–336.
- [13] O. Favaron, M. Mahéo, J.-F. Saclé, Some eigenvalue properties in graphs (conjectures of Graffiti-II), *Discrete Math.* **111** (1993) 197–220.

- [14] A. Berman, X.-D. Zhang, On the spectral radius of graphs with cut vertices, *J. Combin. Theory, Ser. B* **83** (2001) 233–240.
- [15] B. Zhou, Zagreb indices, *MATCH Commun. Math. Comput. Chem.* **52** (2004) 113–118.
- [16] A. Dress, D. Stevanović, Hoffman-type identities, *Appl. Math. Lett.* **16** (2003) 297–302.
- [17] A. Yu, M. Lu, F. Tian, On the spectral radius of graphs, *Linear Algebra Appl.* **387** (2004) 41–49.
- [18] D. Stevanović, P. Hansen, The minimum spectral radius of graphs with a given clique number, *Electron. J. Linear Algebra* **17** (2008) 110–117.