Variable Neighborhood Search for Extremal Graphs. 6. Analyzing Bounds for the Connectivity Index

Pierre Hansen*

GERAD and École des Hautes Études Commerciales, 3000 Chemin de la Côte-Sainte-Catherine, Montréal H3T 2A7, Canada

Hadrien Mélot

Service d'Optimisation, Institut de Statistique et de Recherche Opérationnelle, Université Libre de Bruxelles, Belgium

Received December 20, 2001

Recently, Araujo and De la Peña gave bounds for the connectivity index of chemical trees as a function of this index for general trees and the ramification index of trees. They also gave bounds for the connectivity index of chemical graphs as a function of this index for maximal subgraphs which are trees and the cyclomatic number of the graphs. The ramification index of a tree is first shown to be equal to the number of pending vertices minus 2. Then, in view of extremal graphs obtained with the system *AutoGraphiX*, all bounds of Araujo and De la Peña are improved, yielding tight bounds, and in one case corrected. Moreover, chemical trees of a given order and a number of pending vertices with minimum and with maximum connectivity index are characterized.

1. INTRODUCTION

Computers are frequently used in chemical graph theory, to evaluate invariants as well as correlations between chemical activities or properties and those invariants or functions of them in *QSAR* and *QSPR* studies.[1−15] They can be used also to enhance chemical graph theory per se, e.g. in finding extremal graphs for invariants or new relations between invariants.

Several systems for computer-aided graph theory have been developed in the last two decades. They are based on different principles. The first approach, of which the *GRAPH* system of Cvetkovic and co-workers[16−18] is a prominent example, exploits interactive computation of invariants. The *ALGOR* component of *GRAPH* contains programs for computing order, size, spectrum of the adjacency matrix, number of triangles, components, bridges, maximum matching, and many other invariants as well as for testing properties such as planarity or hamiltonicity for graphs which can be modified on screen. Moreover, this system allows the generation of families of graphs. It contains also a bibliographic component (*BIBLIO*) and a computer-aided theorem proving one (*THEOR*). From 1982 to 1992 it led to 55 papers written by 16 mathematicians, and much further work has since been done with its help.

Other systems working along similar lines are the *Graph-Base* system of Knuth,[19] the *LINK* system of Berry and co-workers,[20−23] and the *VEGA* system developed by Pisanski.[24]

A second approach, illustrated by the *INGRID* system of Brigham and Dutton,[25−28] consists of building a database of graph theoretic relations and proceeding to algebraic manipulations to obtain new ones. This is motivated by specifying assumptions on the graphs under study in the form of values or intervals on invariants. Then these intervals are reduced by rules derived from the relations, thus revealing the influence of one invariant upon another. The *INGRID* system can find which relations may lead to a new one, refute conjectures by considering them temporarily as theorems and deriving a contradiction, and show that some relations are implied by one or several others.

A third approach initiated in the system *Graffiti* of Fajtlowicz[29−33] consists of generating a large series of simple a priori relationships between graph invariants, testing them on a database of examples, rejecting those which are falsified, shelving those which do not appear to be interesting or do not provide new information, and submitting the remaining ones to the mathematical (and more recently chemical) community, in the computer file "Written on the Wall".[34] This system has also attracted the attention of many mathematicians and led to several tens of papers.[35]

A fourth, more recent, approach, initiated by the system *AutoGraphiX* (*AGX*) of Caporossi and Hansen,[36] consists of generating a series of extremal or near-extremal graphs for some invariants or relations between invariants and then analyzing them, directly, interactively, or automatically to derive conjectures. Heuristic search, and more precisely the *Variable Neighborhood Search* metaheuristic,[37] is applied to obtain those graphs. The approach is presented in ref 36, where its main features are illustrated; it can (i) find graphs satisfying various constraints; (ii) find graphs with extremal values for some invariant, possibly subject to constraints; (iii) refute conjectures; (iv) suggest conjectures; and (v) give ideas of proofs or show that some such ideas are not likely to succeed.

Three ways to automate the generation of conjectures obtained by *AGX* are outlined in ref 38: (a) a numerical

* Corresponding author e-mail: pierreh@crt.umontreal.ca.

10.1021/ci010133j CCC: $25.00
Published on Web 11/28/2002

one, based on the mathematics of principal component analysis, where instead of explaining differences between extremal graphs (observations), resemblances between them are captured in the form of a basis of affine relations between invariants; (b) a geometric one in which the convex hull of points representing extremal graphs in invariants space is computed and provides linear inequalities corresponding to facets; and (c) an algebraic one which recognizes families of extremal graphs and manipulates known formulas for invariants on these families to obtain new relations.

Further papers describing mathematical applications of *AGX* are in refs 39−41. Several papers on the use of *AGX* in chemical graph theory have also appeared. In ref 42 a study is made of the energy of a graph. New relations between energy, size, and order of graphs were obtained and easily proved. They appear to be simpler in most cases than those of the literature and yet previously unnoticed. Conjectures on the structure of extremal graphs are also found and appear to be harder to prove. An example is as follows: "Among unicyclic graphs on n vertices, the cycle C_n has maximum energy if $n \leq 7$ and $n = 9, 11, 13, 15$. For all other values of n the unicyclic graph with maximal energy is composed of a 6-cycle and an appended path with $n - 6$ edges". Partial results toward a proof of this conjecture were very recently obtained by Gutman and Hou.[43]

In ref 44 chemical trees with minimum connectivity index are characterized; they belong (or are equivalent through a transformation) to three families, two of which are caterpillars. These families were easily obtained with *AGX*. The proof relies on a novel use of linear and mixed-integer programming, where variables are associated with numbers of edges with given end degrees instead of with individual edges, as in most graph-theoretic models. These results led to several further papers in which cyclic graphs are considered.[45−50]

In ref 51 chemical trees with maximum HOMO−LUMO gap are studied. For even order they are combs, a result already obtained by Shao and Hong.[52] For odd order their form is more complicated and does not appear to be unique.

The purpose of the present paper is to further illustrate the power of interactive use of *AGX* through a study of the connectivity index of trees, chemical trees, and chemical graphs as functions of the order and number of pending vertices or cyclomatic number of the tree or graph. Our point of departure is a recent paper of Araujo and De la Peña[1] on these problems, which presents several lower and upper bounds on the connectivity index. With the help of *AGX* all these results are improved (and in one case corrected). Let us now review the results of ref 1.

The *connectivity index* χ was conceived by Randić[53] and is also called the Randić index. Together with its generalizations it is certainly the molecular-graph-based structure-descriptor, that found the most numerous applications in organic chemistry, medicinal chemistry, and pharmacology. Given adjacent vertices v_1 and v_2 of a graph, the *weight* of the edge $\{v_1, v_2\}$ is defined by

$$\frac{1}{\sqrt{d_1 d_2}} \tag{1}$$

where d_i is the degree of the vertex v_i. The connectivity index of a graph is the sum of the weights of all edges of the graph.

In ref 1 the *ramification index* of a tree T is defined as

$$r(T) = \sum_{d_i \geq 3} (d_i - 2)$$

We will prove below that this index is equivalent, up to a constant, to the number of pending vertices of T. To this effect, we establish a more general result, extending to graphs the previous definition.

Theorem 1. *For all graphs G without isolated vertices, with n vertices, n_1 pending vertices, m edges, and ramification index r(G)*

$$r(G) = 2(m - n) + n_1$$

Proof. By definition of the ramification index

$$r(G) = \sum_{v_i \in V} (d_i - 2) - \sum_{v_i \in V, d_i < 3} (d_i - 2) \tag{2}$$

By easy observations when $d_i = 2$ or 1

$$\sum_{v_i \in V, d_i < 3} (d_i - 2) = -n_1 \tag{3}$$

and (2) becomes

$$r(G) = \sum_{v_i \in V} d_i - 2n + n_1 \tag{4}$$

As for every graph G

$$\sum_{v_i \in V} d_i = 2m \tag{5}$$

we get the result

$$r(G) = 2(m - n) + n_1 \tag{6}$$

Corollary 2. *If T is a tree*

$$r(T) = n_1 - 2$$

Proof. Replacing m by $n - 1$ in Theorem 1 gives the result.

Recall that the *cyclomatic number* $g = m - n + 1$ of a connected graph G is the maximum number of independent cycles of G.

Corollary 3. *If G is a connected graph*

$$r(G) = n_1 + 2g - 2$$

Proof. Replacing $m - n$ by $g - 1$ in Theorem 1 gives the result.

A *comet* is a tree composed of a star and an appended path. For any numbers n and $2 \leq n_1 \leq n - 1$, we denote by $T(n, n_1)$ the comet of order n with n_1 pending vertices, i.e., a tree formed by a path P_{n-n_1} of which one end vertex coincides with a pending vertex of a star S_{n_1+1} of order $n_1 + 1$ (see Figure 1).

By definition of the connectivity index, we have

$$\chi(P_n) = \frac{n}{2} - \left(\frac{3}{2} - \sqrt{2}\right) \tag{7}$$

and if $2 \leq n_1 < n - 1$

$$\chi(T(n, n_1)) = \sqrt{n_1} + \left(\frac{1}{\sqrt{2}} - 1\right)\frac{1}{\sqrt{n_1}} + \frac{n - n_1 - 2}{2} + \frac{1}{\sqrt{2}} \quad (8)$$

If $n_1 = n - 1$, T is a star S_n and $\chi(T) = \sqrt{n-1}$.

A first theorem of ref 1 bounds the connectivity index of a chemical tree T of order n in terms of its ramification index and of the connectivity indices of a comet and a path of the same order.

Theorem 4 ([1]). *Let T be a chemical tree with n vertices and r(T) = t − 2. Then*

$$\chi(T(n, t)) - c_0(r(T) - 1) \leq \chi(T) \leq \chi(P_n) - a_0 r(T)$$

where

$$a_0 = 1 - \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{6}} \ (\simeq 0.0144)$$

and

$$c_0 = 0 \textit{ if } r(T) = 0$$

and otherwise

$$c_0 = \frac{\sqrt{3}}{2} - \frac{3}{4} \ (\simeq 0.1160)$$

With the help of *AGX*, we improve these two bounds and characterize the corresponding extremal chemical trees. These results are presented in section 3.

It is known that $\chi(P_n)$ is maximum for all trees with n vertices (see refs 44, 45, and 54) and that $\chi(S_n)$ is minimal for all graphs of order n.[55] One may wonder if a similar property holds for $\chi(n, n_1)$. *AGX* suggests that the trees with minimal connectivity index and a fixed number of pending vertices are the comets. This is proved in section 2 (as this result is needed in section 3).

A second theorem of ref 1 gives bounds for the connectivity index of chemical graphs.

Theorem 5 ([1]). *Let G be a chemical graph and T be a maximal subgraph of G which is a tree. Then*

$$\chi(T) - d_0 g(G) \leq \chi(G) \leq \chi(T) + b_0 g(G)$$

where

$$b_0 = -\frac{1}{2\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{6}} - \frac{1}{2} \ (\simeq 0.1320)$$

and

$$d_0 = \sqrt{2} + \frac{1}{2\sqrt{3}} - \frac{3}{2} \ (\simeq 0.2029)$$

In particular if G has n vertices and r(T) = t − 2, then

$$\chi(T(n, t)) - c_0(r(T) - 1) - d_0 g(G) \leq \chi(G) \leq \chi(P_n) - a_0 r(T) + b_0 g(G)$$

Note that, in view of Corollary 2, G has t pending vertices. The coefficients a_0 and c_0 are the same as in Theorem 4.

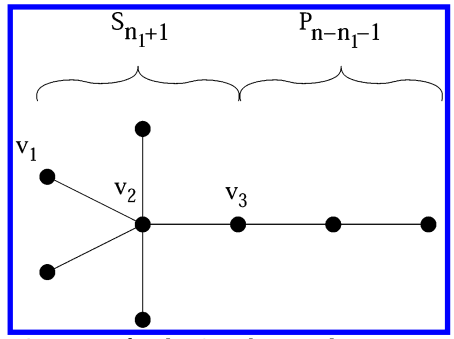

Figure 1. A comet of order 8 with 5 pending vertices.

We give counterexamples, correct and slightly strengthen Theorem 5 in section 4. Conclusions are drawn in section 5.

2. A NEW BOUND ON THE CONNECTIVITY INDEX OF TREES

As explained above, we used *AGX* to find chemical trees with minimal connectivity index. The trees obtained are comets for $n_1 \leq 4$. This fact led us to run *AGX* on trees that are not chemical, i.e., we minimized $\chi(T)$ where T is a tree with n and n_1 fixed. The system gave systematically extremal graphs which are comets, for all values of n and n_1 considered. We now prove that comets indeed minimize the connectivity index.

Theorem 3 of ref 55 says that every graph G of order n containing no isolated vertex has a connectivity index greater than or equal to $\sqrt{n-1}$ with equality if and only if G is a star.

Let $\alpha = \{v_1, v_2\}$ be a pending edge of a graph G; we note $G - v_1v_2$ the graph obtained by removing the vertex of degree 1 of α and the edge α.

The following lemma is a variant of the Lemma 1 of ref 55, which applies when G is a tree, and which uses the number of pending vertices.

Lemma 6. *Let $\{v_1, v_2\}$ be an edge of a tree T of order n with $n_1 \leq n - 2$ pending vertices. If $d_1 = 1$ and $d_2 \geq 2$ then*

$$\chi(T) - \chi(T - v_1v_2) \geq \sqrt{n_1} - \sqrt{n_1 - 1} + \left(\frac{1}{\sqrt{2}} - 1\right)\left(\frac{1}{\sqrt{n_1}} - \frac{1}{\sqrt{n_1 - 1}}\right)$$

with equality if and only if T is a comet and v_2 has maximum degree.

Proof. Denote by W_2 the sum of the weights of the edges, other than $\{v_1, v_2\}$, incident with vertex v_2. Note that, by definition of the connectivity index

$$\chi(T) - \chi(T - v_1v_2) = \frac{1}{\sqrt{d_2}} + W_2 - W_2\sqrt{\frac{d_2}{d_2 - 1}} \quad (9)$$

If all vertices adjacent to v_2, other than v_1, have a degree of 1, then W_2 is maximal. But in this case, we have a star and $n_1 = n - 1$, which is excluded by assumption. So, at least one vertex adjacent to v_2 has a degree $\geq$ 2. Thus

$$W_2 \leq \frac{d_2 - 2}{\sqrt{d_2}} + \frac{1}{\sqrt{2d_2}} \quad (10)$$

with equality if and only if all the vertices adjacent to v_2 are pending vertices, except one, noted v_3, with $d_3 = 2$ (see again Figure 1).

As

$$1 - \sqrt{\frac{d_2}{d_2 - 1}} < 0$$

majorizing W_2 with (10) in (9) yields

$$\chi(T) - \chi(T - v_1v_2) \geq \frac{1}{\sqrt{d_2}} + \left(\frac{d_2 - 2}{\sqrt{d_2}} + \frac{1}{\sqrt{2d_2}}\right)\left(1 - \frac{\sqrt{d_2}}{\sqrt{d_2 - 1}}\right) \quad (11)$$

or, after easy manipulations

$$\chi(T) - \chi(T - v_1v_2) \geq \sqrt{d_2} - \sqrt{d_2 - 1} + \left(\frac{1}{\sqrt{2}} - 1\right)\left(\frac{1}{\sqrt{d_2}} - \frac{1}{\sqrt{d_2 - 1}}\right) \quad (12)$$

Let $f(d_2)$ denote the right-hand side of (12). Then

$$f'(d_2) = \frac{1}{2}\left[\frac{1}{\sqrt{d_2}}\left(1 + \frac{\sqrt{2} - 1}{\sqrt{2}}\frac{1}{d_2}\right) - \frac{1}{\sqrt{d_2 - 1}}\left(1 + \frac{\sqrt{2} - 1}{\sqrt{2}}\frac{1}{d_2 - 1}\right)\right]$$

and as

$$\frac{1}{\sqrt{d_2}}\left(1 + \frac{\sqrt{2} - 1}{\sqrt{2}}\frac{1}{d_2}\right)$$

decreases when d_2 increases, $f'(d_2) < 0$. Hence, as $d_2 \leq \Delta \leq n_1$, where Δ denotes the maximum degree of T

$$\chi(T) - \chi(T - v_1v_2) \geq \sqrt{n_1} - \sqrt{n_1 - 1} + \left(\frac{1}{\sqrt{2}} - 1\right)\left(\frac{1}{\sqrt{n_1}} - \frac{1}{\sqrt{n_1 - 1}}\right) \quad (13)$$

with equality if and only if T is a comet, by construction of the graph obtained by taking equality in (10) and (11), where v_2 is the center of the star, and v_3 is the end vertex of the path which belongs to the star (see once more Figure 1).

We now give our result for trees with minimal connectivity index and fixed number of pending vertices.

Theorem 7. *Let T be a tree of order $n \geq 3$, with n_1 pending vertices. Then if $n_1 < n - 1$*

$$\chi(T) \geq \sqrt{n_1} + \left(\frac{1}{\sqrt{2}} - 1\right)\frac{1}{\sqrt{n_1}} + \frac{n - n_1 - 2}{2} + \frac{1}{\sqrt{2}} \quad (14)$$

with equality if and only if T is the comet $T(n, n_1)$.

Proof. If T is a comet, then we have equality, since we have seen in expression (8) that the connectivity index of a comet is equal to the right-hand side of (14).

We apply induction on n. It is easy to check that assertion (14) holds for the 4 trees of order 3, 4 and 5 which are not stars ($n_1 < n - 1$). These trees are P_3, P_4, P_5, and $T(5, 3)$,

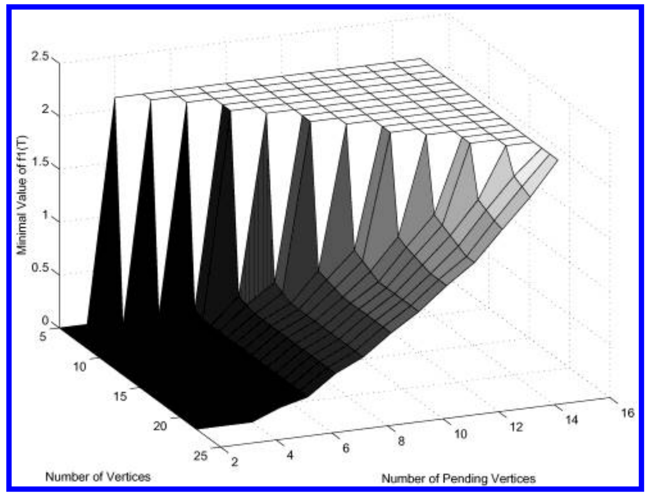

Figure 2. Minimal values for $f_1(T)$ found by *AGX*.

which are all comets. So let us assume that $n \geq 6$ and that the result holds for all smaller values of n.

It is well-known that every tree has at least two pending vertices and that if we remove a pending edge of a tree, the subgraph so obtained remains a tree. Let $\{v_1, v_2\}$ be a pending edge of T (with $d_1 = 1$). We next prove the induction, discussing the degree of v_2.

(i) If $d_2 = 1$, then T is the only tree with two vertices connected by one edge, which contradicts the assumption $n \geq 6$.

(ii) If $d_2 = 2$, we note v_3 ($\neq v_1$) the vertex adjacent to v_2. We have

$$\chi(T) - \chi(T - v_1v_2) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{d_3}}\left(\frac{1}{\sqrt{2}} - 1\right)$$

As $1/\sqrt{2} - 1$ is negative, this expression is minimal when $1/\sqrt{d_3}$ is maximal, i.e., when d_3 is minimal.

If $d_3 = 1$, T is the path P_3, which is impossible by the assumption $n \geq 6$. So, $d_3 = 2$ is minimal and

$$\chi(T) - \chi(T - v_1v_2) \geq \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}} - 1\right) = \frac{1}{2} \quad (15)$$

with equality if and only if $d_3 = 2$.

As $d_2 = 2$, the tree $T - v_1v_2$ has n_1 pending vertices and $n - 1$ vertices. In the case of $n_1 = n - 2$, the tree $T - v_1v_2$ is a star, so T is a comet and we can stop the induction here. If $n_1 < n - 2$, by induction the theorem holds and

$$\chi(T - v_1v_2) \geq \sqrt{n_1} + \left(\frac{1}{\sqrt{2}} - 1\right)\frac{1}{\sqrt{n_1}} + \frac{n - n_1 - 3}{2} + \frac{1}{\sqrt{2}} \quad (16)$$

with equality if and only if $T - v_1v_2$ is a comet. So, by expressions (15) and (16)

$$\chi(T) \geq \sqrt{n_1} + \left(\frac{1}{\sqrt{2}} - 1\right)\frac{1}{\sqrt{n_1}} + \frac{n - n_1 - 2}{2} + \frac{1}{\sqrt{2}}$$

with equality if and only if $T - v_1v_2$ is a comet and $d_3 = 2$, and that implies that we have equality if and only if T is also a comet.

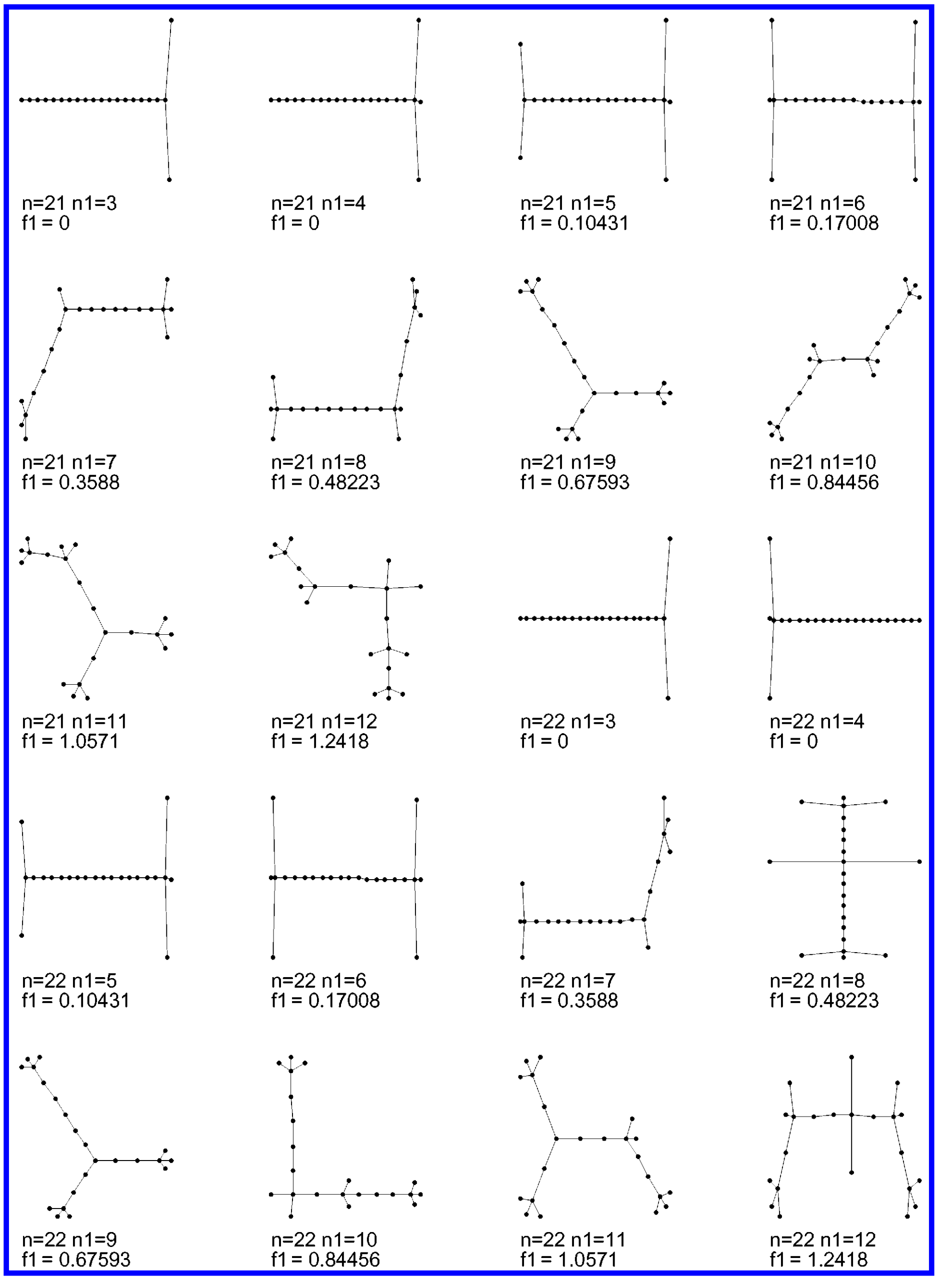

Figure 3. Chemical trees with presumably minimal values for $f_1(T)$ found by *AGX*.

(iii) If $d_2 \geq 3$, the tree $T - v_1v_2$ has $n_1 - 1$ pending vertices and $n - 1$ vertices. Then, by the induction hypothesis

$$\chi(T - v_1v_2) \geq \sqrt{n_1 - 1} + \left(\frac{1}{\sqrt{2}} - 1\right)\frac{1}{\sqrt{n_1 - 1}} + \frac{n - n_1 - 2}{2} + \frac{1}{\sqrt{2}} \quad (17)$$

with equality if and only if $T - v_1v_2$ is a comet. By Lemma 6, we have

$$\chi(T) \geq \chi(T - v_1v_2) + \sqrt{n_1} - \sqrt{n_1 - 1} + \left(\frac{1}{\sqrt{2}} - 1\right)\left(\frac{1}{\sqrt{n_1}} - \frac{1}{\sqrt{n_1 - 1}}\right) \quad (18)$$

with equality if and only if T is a comet. Thus, by (17) and (18)

$$\chi(T) \geq \sqrt{n_1} + \left(\frac{1}{\sqrt{2}} - 1\right)\frac{1}{\sqrt{n_1}} + \frac{n - n_1 - 2}{2} + \frac{1}{\sqrt{2}}$$

with equality if and only if T is a comet.

An immediate consequence of Theorem 7 is that the term $-c_0(r(T) - 1)$ can be disposed of in Theorem 4. However, we can do better, as shown in the next section.

3. IMPROVEMENTS OF THEOREM 4

We are now looking for the extremal chemical trees for the connectivity index.

3.1. Lower Bound. We have used the system *AGX* on chemical trees, iterating on the number n of vertices and the number n_1 of pending vertices, minimizing the function

$$f_1(T) = \chi(T) - \chi(T(n, n_1)) \quad (19)$$

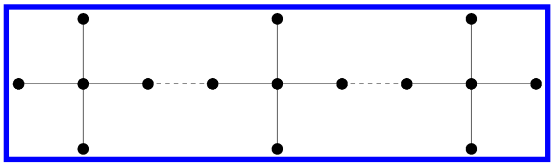

Figure 4. Structure of $L_e(n, n_1)$.

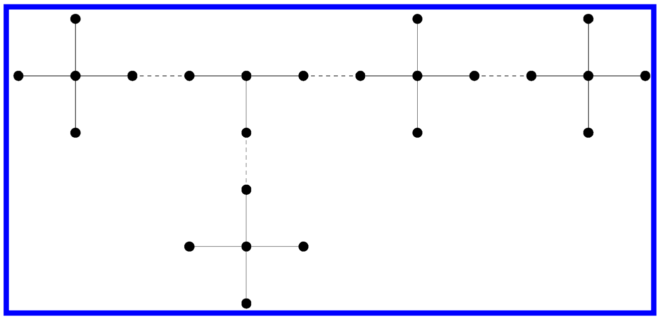

Figure 5. Structure of $L_o(n, n_1)$.

where T is a chemical tree (if T is not chemical, we know by Theorem 7 that the minimal trees are the comets).

Presumably minimal values for $f_1(T)$, depending of n and n_1 and found by *AGX*, are represented in Figure 2, for $5 \le n \le 21$ and $2 \le n_1 \le \min(n - 2, 12)$. Examples of the corresponding extremal trees are shown in Figure 3, for $21 \le n \le 22$ and $3 \le n_1 \le 12$.

An examination of these trees reveals some special structures. When $n_1 \le 4$, they are comets, which are minimal by Theorem 7. When $n_1 \ge 5$, the comets are no more chemical trees, and *AGX* gives trees with specific structures, depending of the parity of n_1. When n_1 is even, we denote them $L_e(n, n_1)$ and when n_1 is odd $L_o(n, n_1)$.

The structure of $L_e(n, n_1)$ is depicted in Figure 4. They are composed by subgraphs which are stars S_5, and these stars are connected by paths (the dotted lines in the figure), for which the lengths can be zero. The configuration is complete if $n \ge 9$ and $6 \le n_1 \le \lfloor (n+3)/2 \rfloor$ (and even).

We can compute $\chi(L_e(n, n_1))$. We see in Figure 4 that $L_e(n, n_1)$ is formed by $(n_1-2)/2$ stars S_5. This chemical tree has n_1 pending edges of weight 1/2, $n_1 - 4$ edges between the centers of the stars and the paths joining these stars of weight $1/\sqrt{8}$. The other edges are on the paths between the stars and have a weight of 1/2 also. As any tree has $n - 1$ edges, there are $n - 2n_1 + 3$ edges of this type (and so n_1 has to be less than or equal to $\lfloor (n+3)/2 \rfloor$).

So

$$\begin{aligned}\chi(L_e(n, n_1)) &= \frac{n_1}{2} + \frac{n_1 - 4}{2\sqrt{2}} + \frac{n - 2n_1 + 3}{2} \\ &= \frac{n}{2} + \frac{n_1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{3}{2} - \sqrt{2}\end{aligned}$$

The structure of $L_o(n, n_1)$ is depicted in Figure 5. This structure is similar to $L_e(n, n_1)$, but, in this case, we have not only 2 but 3 *pending stars* S_5, i.e., stars connected to the remainder of the tree through one vertex only. That implies that one vertex has degree 3. The dotted lines in Figure 5 are paths, for which the length can be zero. This configuration is not always complete: we should have $n \ge 16$ and $9 \le n_1 \le \lfloor (n+2)/2 \rfloor$ (and odd).

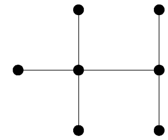

Figure 6. The only chemical tree with 7 vertices and 5 pending ones.

We can compute $\chi(L_o(n, n_1))$. We see in Figure 5 that this chemical tree has n_1 pending edges of weight 1/2, $n_1 - 6$ edges between the centers of the stars S_5 and the paths joining these stars of weight $1/\sqrt{8}$, and 3 edges adjacent to the vertex of degree 3 of weight $1/\sqrt{6}$. The other edges are on the paths represented by the dotted lines and have a weight of 1/2 also. The tree containing $n - 1$ edges, there are $n - 2n_1 + 2$ edges of this type (and so n_1 has to be less than or equal to $\lfloor (n+2)/2 \rfloor$).

So

$$\begin{aligned}\chi(L_o(n, n_1)) &= \frac{n_1}{2} + \frac{n_1 - 6}{2\sqrt{2}} + \frac{3}{\sqrt{6}} + \frac{n - 2n_1 + 2}{2} \\ &= \frac{n}{2} + \frac{n_1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + 1 + \frac{3}{\sqrt{6}} - \frac{3}{\sqrt{2}}\end{aligned}$$

Dropping the parity condition on n_1, we see that $\chi(L_e(n, n_1)) < \chi(L_o(n, n_1))$ because

$$\frac{3}{2} - \sqrt{2} \simeq 0.0858 < 1 + \frac{3}{\sqrt{6}} - \frac{3}{\sqrt{2}} \simeq 0.1034$$

All these observations on the results of *AGX* suggest the following theorem: $L_e(n, n_1)$ are the minimal chemical trees, when $n_1 \ge 15$.

Theorem 8. *Let T be a chemical tree of order n with $n_1 \ge 5$ pending vertices. Then*

$$\chi(T) \ge \frac{n}{2} + \frac{n_1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{3}{2} - \sqrt{2} \qquad (20)$$

with equality if and only if n_1 is even and T is isomorphic to $L_e(n, n_1)$.

Proof. First we apply induction, as in the proof of Theorem 7, to prove inequality (20).

By assumption $n_1 \ge 5$, and as $n_1 \le n - 1$, we have $n \ge 6$. If $n = 6$, we have the star S_6 which is not chemical. If $n = 7$, n_1 is equal to 5 or 6. The latter case is again a star (S_7) which is not chemical. It is easy to check that the only chemical tree T^* with 7 vertices and 5 pending vertices is that one of Figure 6. In this case, condition (20) is satisfied as

$$\begin{aligned}\chi(T^*) = \frac{3}{2} + \frac{2}{\sqrt{3}} + \frac{1}{\sqrt{12}} \simeq 2.9434 \ge \\ \frac{7}{2} + \frac{5}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{3}{2} - \sqrt{2} \simeq 2.8536\end{aligned}$$

So let us assume that $n \ge 8$ and (20) holds for all smaller values of n.

Let $\{v_1, v_2\}$ be a pending edge of a tree T with n vertices (with $d_1 = 1$). If $d_2 = 1$, T has 2 vertices which contradicts $n \ge 8$.

If $d_2 = 2$

$$\chi(T) - \chi(T - v_1v_2) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{d_3}}\left(\frac{1}{\sqrt{2}} - 1\right)$$

where d_3 is the degree of v_3, the vertex adjacent to v_2 other than v_1. Thus

$$\chi(T) - \chi(T - v_1v_2) \geq \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}} - 1\right) = \frac{1}{2}$$

as $d_3 \geq 2$.

As $d_2 = 2$, the tree $T - v_1v_2$ has n_1 pending vertices and $n - 1$ vertices.

By induction

$$\chi(T) \geq \frac{n-1}{2} + \frac{n_1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{3}{2} - \sqrt{2} + \frac{1}{2}$$

and (20) holds.

If $d_2 \geq 3$, the tree $T - v_1v_2$ has $n_1 - 1$ pending vertices and $n - 1$ vertices. Then, by induction

$$\chi(T - v_1v_2) \geq \frac{n}{2} + \frac{n_1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{3}{2} - \sqrt{2} + \frac{1}{2} - \frac{1}{\sqrt{2}} \quad (21)$$

We are in the same conditions as for Lemma 6, so we can use eq 13, replacing n_1 by 4 since $\Delta \leq 4$ in chemical trees. We obtain

$$\chi(T) - \chi(T - v_1v_2) \geq 2 - \sqrt{3} + \left(\frac{1}{\sqrt{2}} - 1\right)\left(\frac{1}{2} - \frac{1}{\sqrt{3}}\right) = \frac{3}{2} - \frac{2}{\sqrt{3}} - \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{8}} \quad (22)$$

Combining expressions (21) and (22), we have

$$\chi(T) \geq \frac{n}{2} + \frac{n_1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{3}{2} - \sqrt{2} + \frac{1}{2} - \frac{1}{\sqrt{2}} + \frac{3}{2} - \frac{2}{\sqrt{3}} - \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{8}}$$

and as $2 - 1/\sqrt{2} - 2/\sqrt{3} - 1/\sqrt{6} + 1/\sqrt{8}$ is positive, expression (20) holds again.

We have shown above that $\chi(L_e(n, n_1))$ is equal to the lower bound of (20). We now prove that trees $L_e(n, n_1)$ are the only ones to reach that bound.

For all chemical trees we have by definition of the connectivity index that

$$\chi(T) = \frac{x_{12}}{\sqrt{2}} + \frac{x_{13}}{\sqrt{3}} + \frac{x_{14}}{2} + \frac{x_{22}}{2} + \frac{x_{23}}{\sqrt{6}} + \frac{x_{24}}{2\sqrt{2}} + \frac{x_{33}}{3} + \frac{x_{34}}{2\sqrt{3}} + \frac{x_{44}}{4} \quad (23)$$

where x_{ij} is the number of edges connecting a vertex of degree i with a vertex of degree j.

Denoting by n_i the number of vertices of degree i, the following six (linearly independent) relations are obeyed (see ref 44):

$$x_{12} + x_{13} + x_{14} = n_1 \quad (24)$$

$$x_{12} + 2x_{22} + x_{23} + x_{24} = 2n_2 \quad (25)$$

$$x_{13} + x_{23} + 2x_{33} + x_{34} = 3n_3 \quad (26)$$

$$x_{14} + x_{24} + x_{34} + 2x_{44} = 4n_4 \quad (27)$$

$$n_1 + 2n_2 + 3n_3 + 4n_4 = 2(n-1) \quad (28)$$

$$n_1 + n_2 + n_3 + n_4 = n \quad (29)$$

If $\chi(T)$ is equal to the lower bound of (20), there are no terms of the form $1/\sqrt{3}$, $1/\sqrt{6}$, $1/2\sqrt{3}$, and so $x_{13} = x_{23} = x_{34} = 0$ for T. This implies that $x_{33} = 0$ because otherwise all vertices of degree 3 are connected together, and only to vertices of degree 3, which is impossible in a tree, so $n_3 = 0$.

We get the following relations:

$$x_{12} + x_{14} = n_1 \quad (30)$$

$$x_{12} + 2x_{22} + x_{24} = 2n_2 \quad (31)$$

$$x_{14} + x_{24} + 2x_{44} = 4n_4 \quad (32)$$

$$n_1 + 2n_2 + 4n_4 = 2(n-1) \quad (33)$$

$$n_2 = n - n_1 - n_4 \quad (34)$$

Replacing n_2 in (33) by (34) gives

$$n_4 = \frac{n_1 - 2}{2} \quad (35)$$

we can replace n_4 by this expression in (34) to obtain

$$n_2 = n - \frac{3n_1}{2} + 1 \quad (36)$$

Eliminating x_{14} by (30) and n_4 by (35) in (32) yields

$$x_{24} = n_1 - 4 - 2x_{44} + x_{12} \quad (37)$$

Then replacing x_{24} by (37) and n_2 by (36) in (31) gives

$$x_{22} = n - 2n_1 + 3 + x_{44} - x_{12} \quad (38)$$

We rewrite the connectivity index (23), replacing x_{13}, x_{23}, x_{33}, and x_{34} by 0

$$\chi(T) = \frac{x_{12}}{\sqrt{2}} + \frac{x_{14}}{2} + \frac{x_{22}}{2} + \frac{x_{24}}{2\sqrt{2}} + \frac{x_{44}}{4}$$

Using eqs (31), (37), and (38) to replace x_{14}, x_{24}, and x_{22}, respectively, we can write this index with the unknowns n, n_1, x_{12}, and x_{44}. We get

$$\chi(T) = \frac{x_{12}}{\sqrt{2}} + \frac{n_1}{2} - \frac{x_{12}}{2} + \frac{x_{44}}{2} - n_1 + \frac{n}{2} + \frac{3}{2} - \frac{x_{12}}{2} + \frac{n_1}{2\sqrt{2}} - \frac{2}{\sqrt{2}} - \frac{x_{44}}{\sqrt{2}} + \frac{x_{12}}{2\sqrt{2}} + \frac{x_{44}}{4}$$

So

$$\chi(T) = \frac{n}{2} + \frac{n_1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{3}{2} - \sqrt{2} + x_{12}\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{8}} - 1\right) + x_{44}\left(\frac{3}{4} - \frac{1}{\sqrt{2}}\right)$$

As $1/\sqrt{2} + 1/\sqrt{8} - 1 \simeq 0.0607$ and $3/4 - 1/\sqrt{2} \simeq 0.0429$ are strictly positive, we have to put $x_{12} = x_{44} = 0$, otherwise the lower bound will not be obtained.

In short, with all the information above we have minimal graphs with $x_{12} = x_{13} = x_{23} = x_{33} = x_{34} = x_{44} = 0$, $x_{14} = n_1$, $x_{22} = n - 2n_1 + 3$ and $x_{24} = n_1 - 4$. By eq 35 we know that the number of vertices of degree 4 is $(n_1-2)/2$, forming the stars S_5. Moreover, there is no vertex of degree 3, all pending vertices are connected to a vertex of degree 4, and the other vertices (of degree 2) form paths connecting the stars S_5. Because $n_4 = (n_1-2)/2$ has to be integer, n_1 has to be even. This is the configuration of $L_e(n, n_1)$.

This theorem is illustrated in Figure 7 (when $n = 22$), where the circles (joined by a dashed curve) represent the value of the minimal chemical trees obtained by *AGX*, the dotted line is the lower bound of Theorem 4 and the black line is the new lower bound proposed in Theorem 8. Observe that the bound is attained for four values of n_1, all even.

3.2. Upper Bound. We first introduce a new concept. The *ramification subgraph* of a tree T is the subgraph induced by the vertices of degree greater than or equal to 3. In other words, it is the subgraph obtained from T by deleting all vertice of degree ≤ 2 and their incumbent edges.

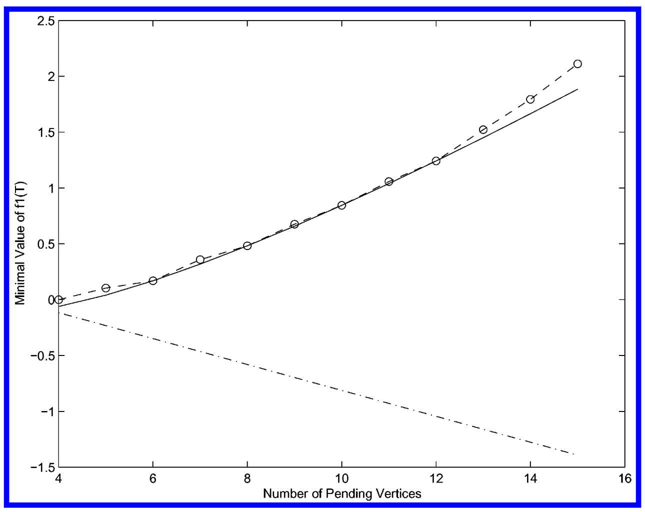

Figure 7. Minimal values for $f_1(T)$, $n = 22$.

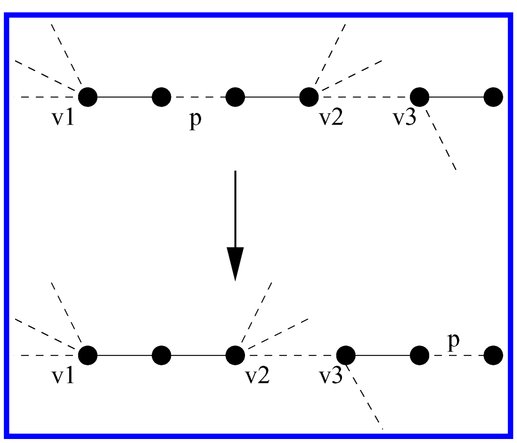

Figure 8. Moving the p inner edges of the path between v_1 and v_2.

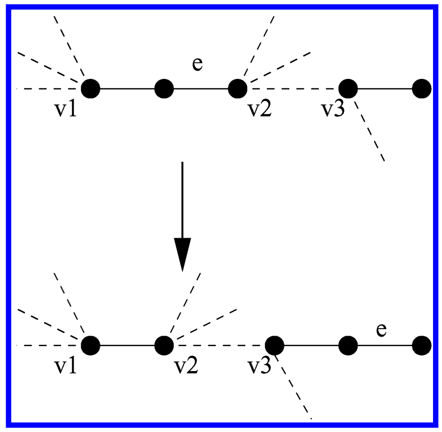

Figure 9. Moving edge e.

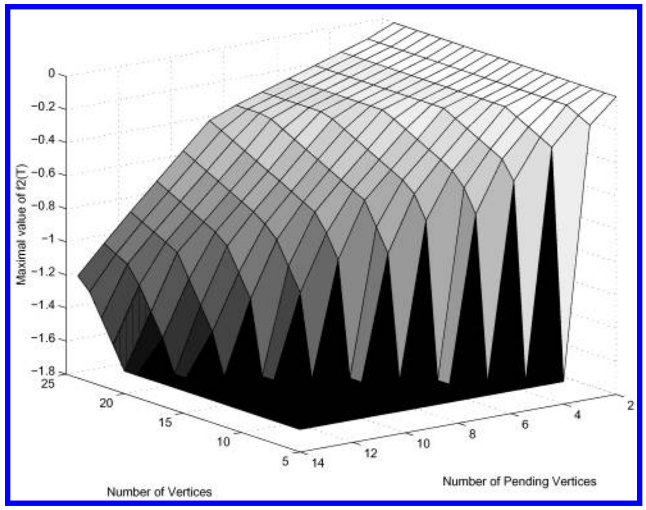

Figure 10. Maximal values for $f_2(T)$ found by *AGX*.

Theorem 9. *The ramification subgraph of a chemical tree T is a tree if $\chi(T)$ is maximal.*

Proof. If $r(T) = 1$, by definition of the ramification index, T has only one vertex of degree 3 forming the ramification subgraph, which is a tree.

If $r(T) \geq 1$ and the ramification subgraph of T is not a tree, we have at least two vertices of T of degree 3 or 4, say v_1 and v_2 that are connected by a path of length at least 2. Suppose that the number p of inner edges on this path is positive, we note T' the chemical tree obtained by moving these p edges to a pending vertex (adjacent to a vertex v_3) as shown in Figure 8. It is easy to see that

$$\chi(T') - \chi(T) = \frac{1}{\sqrt{2}} - \frac{1}{2} + \frac{1}{\sqrt{d_3}}\left(\frac{1}{\sqrt{2}} - 1\right)$$

as $1/\sqrt{2} - 1$ is negative, the previous expression is minimal when d_3 is minimal, i.e., $d_3 = 2$. Thus

$$\chi(T') - \chi(T) \geq \frac{1}{\sqrt{2}} - \frac{1}{2} + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}} - 1\right) = 0$$

and $\chi(T)$ is not maximal if $p \geq 1$ so we can suppose that $p = 0$. We can move one of the two edge between v_1 and v_2, say e, to a pending vertex, as shown in Figure 9. We note T'' the chemical tree obtained by this operation. By definition

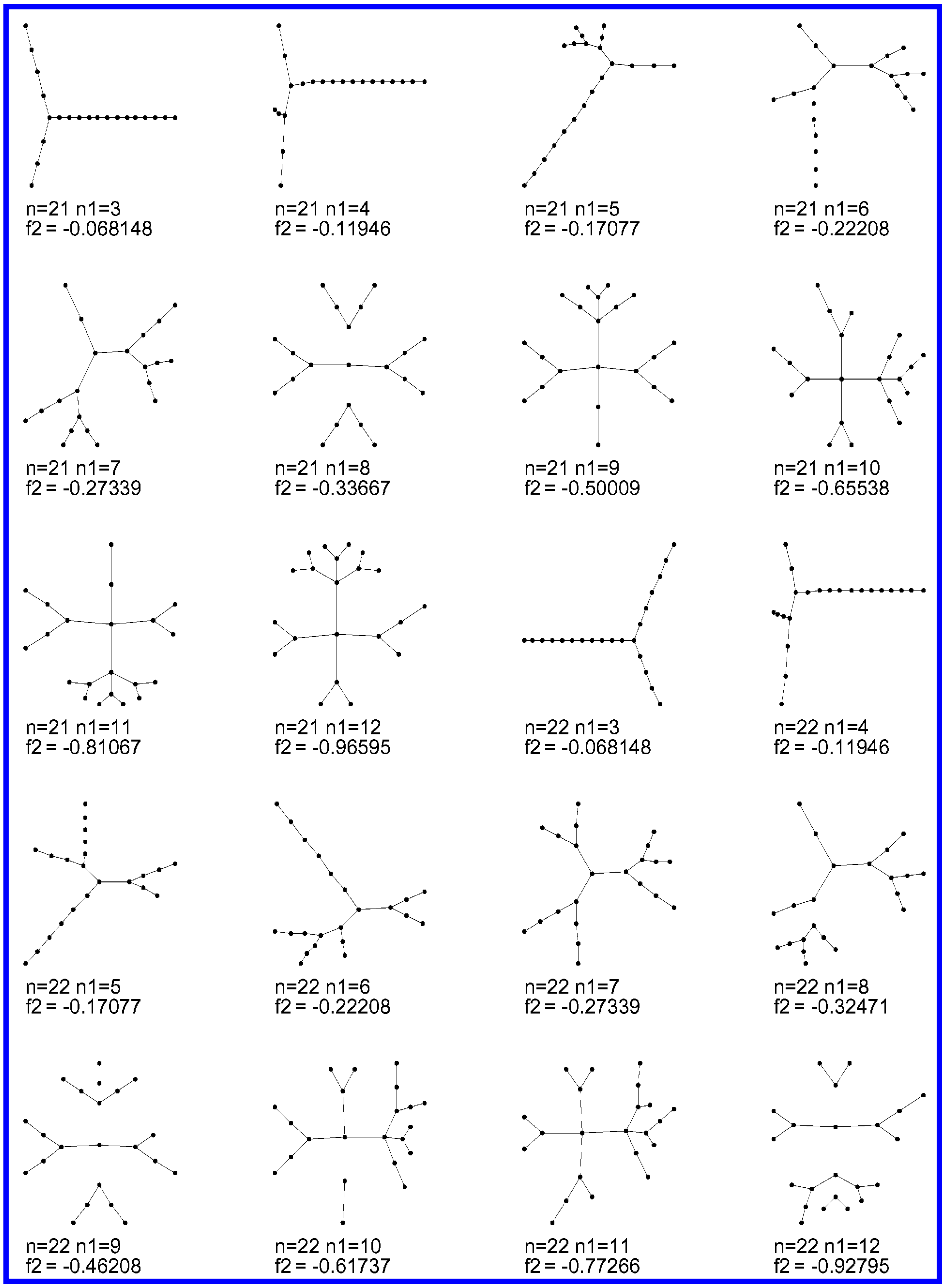

Figure 11. Graphs with presumably maximal values for $f_2(T)$ found by *AGX*.

of the connectivity index

$$\chi(T'') - \chi(T) = \frac{1}{\sqrt{d_3}}\left(\frac{1}{\sqrt{2}} - 1\right) + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{d_1 d_2}} - \frac{1}{\sqrt{2d_1}} - \frac{1}{\sqrt{2d_2}}$$

as $1/\sqrt{2} - 1$ is negative, this expression is minimal when d_3 is minimal, i.e., $d_3 = 2$. Moreover as d_1 and d_2 are equal to 3 or 4, there are three different cases: the two vertices have a degree 3; the two vertices have a degree 4; one vertex has a degree 3 and the other a degree 4. It is easy to check that the difference between $\chi(T'')$ and $\chi(T)$ is minimal when $d_1 = d_2 = 3$. Thus

$$\chi(T'') - \chi(T) \geq \frac{1}{2} + \frac{1}{3} - \frac{2}{\sqrt{6}} \simeq 0.0168$$

and T is not maximal for the connectivity index. The result follows.

Again, we have applied the system *AGX* on chemical trees, iterating on the number n of vertices and the number n_1 of pending vertices. We maximize the following objective function. Let T be a chemical tree

$$f_2(T) = \chi(T) - \chi(P_n) \tag{39}$$

The maximal values found by *AGX* for $f_2(T)$ are represented in Figure 10, for $5 \leq n \leq 24$ and $2 \leq n_1 \leq \min(n - 2, 14)$. We see that a plane can be taken for an upper bound, except if $n_1 = 2$, which is an easy case, i.e., the paths.

Examples of extremal graphs are shown in Figure 11, for $21 \leq n \leq 22$ and $3 \leq n_1 \leq 12$. An examination of these extremal trees reveals again a special structure. We denote $U(n, n_1)$ these specific chemicals trees. The trees $U(n, n_1)$ have a subgraph of $n_1 - 2$ vertices of degree 3 which is a tree; we note this vertex set V_3. In Figure 12, the vertices of V_3 are on a path, but it can be different. All these vertices are connected to another vertex of V_3 or to a path of length

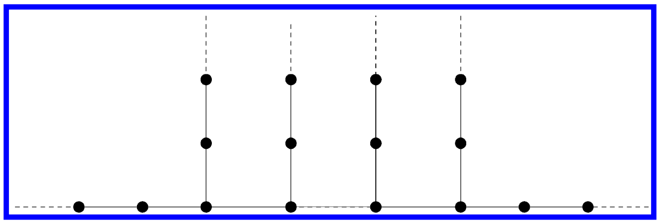

Figure 12. Structure of $U(n, n_1)$.

at least 2. The number of paths adjacent to the vertices of V_3 is $|V_3| + 2$, and the number of vertices of degree 2 is $n - 2n_1 + 2$.

This configuration is complete if $n1 \geq 7$ and $3 \leq n_1 \leq \lfloor (n+2)/3 \rfloor$, and that explains that we do not have an upper plane, going through many points in Figure 10 if the case $n_1 = 2$ is included.

We can compute $\chi(U(n, n_1))$. We see in Figure 12 that $U(n, n_1)$ has n_1 pending edges of weight $1/\sqrt{2}$, n_1 edges connecting the paths and the vertices of V_3 of weight $1/\sqrt{6}$, $n_1 - 3$ edges joining the vertices of V_3 of weight 1/3. The $n - 3n_1 + 2$ other edges are the inner edges of the paths (the dotted lines) of weight 12 (and thus $n_1 \leq \lfloor (n+2)/3 \rfloor$). So

$$\chi(U(n, n_1)) = \frac{n}{2} + n_1\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{6}} - \frac{7}{6}\right) \qquad (40)$$

Theorem 10. *Let T be a chemical tree of order n and with $n_1 \geq 3$ pending vertices. Then*

$$\chi(T) \leq \frac{n}{2} - a_0' n_1$$

where

$$a_0' = \frac{7}{6} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{6}} \simeq 0.0513$$

and with equality if and only if T is isomorphic to $U(n, n_1)$.

Proof. As in the proof of Theorem 8, we use expression (23) and relations (24)−(29).

The idea of this proof is similar to the proof of Theorem 2 of ref 44. We will consider eqs (24)−(29) as a system of six linear equations in the unknowns x_{12}, x_{22}, x_{33}, n_2, n_3, n_4 and solve it. The solutions thus obtained will then depend on the remaining parameters, namely on the following: x_{13}, x_{14}, x_{23}, x_{24}, x_{34}, x_{44}, n_1 and n; the graphs $U(n, n_1)$ have a majority of these latter parameters equal to zero.

Rewriting eqs (24)−(29) as

$$x_{12} = f_1$$
$$2n_2 - x_{12} - 2x_{22} = f_2$$
$$3n_3 - 2x_{33} = f_3$$
$$4n_4 = f_4$$
$$2n_2 + 3n_3 + 4n_4 = f_5$$
$$n_2 + n_3 + n_4 = f_6$$

where

$$f_1 = n_1 - x_{13} - x_{14}$$
$$f_2 = x_{23} + x_{24}$$
$$f_3 = x_{13} + x_{23} + x_{34}$$
$$f_4 = x_{14} + x_{24} + x_{34} + 2x_{44}$$
$$f_5 = 2n - n_1 - 2$$
$$f_6 = n - n_1$$

we readily obtain the required solutions:

$$x_{12} = f_1$$
$$x_{22} = 3f_6 - f_5 + \frac{f_4}{4} - \frac{f_2}{2} - \frac{f_1}{2}$$
$$x_{33} = -3f_6 + \frac{3}{2}f_5 - \frac{3}{4}f_4 - \frac{f_3}{2}$$
$$n_2 = 3f_6 - f_5 + \frac{f_4}{4}$$
$$n_3 = -2f_6 + f_5 - \frac{f_4}{2}$$
$$n_4 = \frac{f_4}{4}$$

In expression (23) only x_{12}, x_{22}, and x_{33} are needed in explicit form:

$$x_{12} = n_1 - x_{13} - x_{14} \qquad (41)$$

$$x_{22} = n - \frac{5}{2}n_1 + \frac{3}{4}x_{14} - \frac{1}{4}x_{24} + \frac{1}{4}x_{34} + \frac{1}{2}x_{44} + \frac{1}{2}x_{13} - \frac{1}{2}x_{23} + 2 \qquad (42)$$

and

$$x_{33} = \frac{3}{2}n_1 - \frac{3}{4}x_{14} - \frac{3}{4}x_{24} - \frac{5}{4}x_{34} - \frac{3}{2}x_{44} - \frac{1}{2}x_{13} - \frac{1}{2}x_{23} - 3 \qquad (43)$$

Substitution of the relations (41)−(43) back into eq 23 yields

$$\chi(T) = \frac{n}{2} + n_1\left(\frac{1}{\sqrt{2}} - \frac{3}{4}\right) + x_{13}\left(-\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{12}\right) + x_{14}\left(-\frac{1}{\sqrt{2}} + \frac{5}{8}\right) + x_{23}\left(\frac{1}{\sqrt{6}} - \frac{5}{12}\right) + x_{24}\left(\frac{1}{2\sqrt{2}} - \frac{3}{8}\right) + x_{34}\left(\frac{1}{2\sqrt{3}} - \frac{7}{24}\right) \qquad (44)$$

Computing and rounding these multipliers we get

$$\chi(T) = \frac{n}{2} + n_1\left(\frac{1}{\sqrt{2}} - \frac{3}{4}\right) - 0.0464x_{13} - 0.0821x_{14} \qquad (45)$$

$$-0.0084x_{23} - 0.0214x_{24} - 0.0030x_{34} \qquad (46)$$

From eq 46 it is clear that, for fixed n and n_1, the value of $\chi(T)$ will be maximal if the parameters x_{13}, x_{14}, x_{23}, x_{24}, x_{34}, and x_{44} are equal to zero, because all of their multipliers are negative, if it is possible.

We obtain the following solutions:

$$x_{12} = n_1 \tag{47}$$

$$x_{22} = n - \frac{5}{2}n_1 + 2 \tag{48}$$

$$x_{33} = \frac{3}{2}n_1 - 3 \tag{49}$$

$$n_2 = n - 2n_1 + 2 \tag{50}$$

$$n_3 = n_1 - 2 \tag{51}$$

$$n_4 = 0 \tag{52}$$

This implies that, if these solutions are feasible, the maximal graphs have no vertex of degree 4, and all vertices of degree 3 form an isolated component, which is impossible in a tree (where $n_3 \geq 1$ by (51) and by the assumption $n_1 \geq 3$). So it is clear that

$$x_{13} + x_{23} + x_{34} > 0 \tag{53}$$

As the solutions are unfeasible, we would like to modify them by increasing parameters with the smallest possible coefficient in (45). This is done with x_{44} for which the coefficient is equal to zero. But if $x_{44} > 0$, the vertices of degree 4 cannot be in an isolated component and $x_{14} + x_{24} + x_{34} > 0$. We prefer increasing x_{34} because of its coefficient in (45). So, for a fixed n_4, the maximal graph will contain a subtree of n_4 vertices of degree 4 (thus $x_{44} = n_4 - 1$) and $2n_4 + 2$ vertices of degree 3 adjacent to the vertices of degree 4 (and $x_{34} = 2n_4 + 2$). Replacing x_{44} and x_{34} by these values for $n_4 \geq 1$ fixed in (42) and (43) gives

$$x_{22} = n - \frac{1}{2}(5n_1 + x_{23} - x_{13}) + 2 + n_4 \tag{54}$$

$$x_{33} = \frac{3}{2}n_1 - \frac{x_{13}}{2} - \frac{x_{23}}{2} - 4 - 4n_4 \tag{55}$$

As x_{33} is non-negative, we get from (55)

$$x_{13} + x_{23} \leq 3n_1 - 8 - 8n_4 \tag{56}$$

As $x_{13} \geq 0$

$$x_{13} + x_{23} \geq x_{23} \geq x_{23} - x_{13}$$

By (56)

$$x_{23} - x_{13} \leq 3n_1 - 8 - 8n_4 \tag{57}$$

As x_{22} has to be a non-negative integer value, by (54), $5n_1 + x_{23} - x_{13}$ has to be even for all n_1, i.e., $x_{23} - x_{13} = (2k + 1)n_1$ (for $k = 0, 1,...$), by (57), the only possibility is that

$$x_{23} - x_{13} = n_1 \tag{58}$$

Remark that if $n_4 = 0$ (thus $x_{44} = x_{34} = x_{24} = x_{14} = 0$) we obtain the same result. In this case $x_{13} + x_{23} > 0$ by (53), and all the other parameters are equal to zero.

The solutions for x_{22} and x_{33} are

$$x_{22} = n + 2 - \frac{1}{2}(5n_1 + x_{23} - x_{13}) \tag{59}$$

$$x_{33} = \frac{3}{2}n_1 - \frac{x_{13}}{2} - \frac{x_{23}}{2} - 3 \tag{60}$$

As $x_{33} \geq 0$ and from (60)

$$x_{13} + x_{23} \leq 3n_1 - 6 \tag{61}$$

As $x_{13} \geq 0$

$$x_{23} - x_{13} \leq 3n_1 - 6 \tag{62}$$

We can apply the same scheme than above: as x_{22} has to be a non-negative integer value, by (59), $5n_1 + x_{23} - x_{13}$ has to be even for all n_1, i.e., $x_{23} - x_{13} = (2k + 1)n_1$ (for $k = 0, 1,...$), since $x_{23} - x_{13} \leq 3n_1 - 6$, the only one possibility is that $x_{23} - x_{13} = n_1$.

So, for all values for n_4, eq 58 is obeyed and we can replace x_{23} by $x_{13} + n_1$ in (44):

$$\chi(T) = \frac{n}{2} + n_1\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{6}} - \frac{7}{6}\right) + x_{13}\left(-\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{6}} - \frac{4}{12}\right) + x_{14}\left(-\frac{1}{\sqrt{2}} + \frac{5}{8}\right) + x_{24}\left(\frac{1}{2\sqrt{2}} - \frac{3}{8}\right) + x_{34}\left(\frac{1}{2\sqrt{3}} - \frac{7}{24}\right) \tag{63}$$

Computing and rounding these multipliers we get

$$\chi(T) = \frac{n}{2} + n_1\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{6}} - \frac{7}{6}\right) - 0.0548x_{13} - 0.0821x_{14} \tag{64}$$

$$-0.0214x_{24} - 0.0030x_{34} \tag{65}$$

As all multipliers are negative, the maximal value for $\chi(T)$ is obtained when $x_{13} = x_{14} = x_{24} = x_{34} = x_{44} = 0$ and

$$\chi(T) \leq \frac{n}{2} - n_1\left(\frac{7}{6} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{6}}\right)$$

which is the upper bound of the theorem. Moreover, we have the following solutions:

$$x_{12} = n_1 \tag{66}$$

$$x_{22} = n - 3n_1 + 2 \tag{67}$$

$$x_{23} = n_1 \tag{68}$$

$$x_{33} = n_1 - 3 \tag{69}$$

$$n_2 = n - 2n_1 + 2 \tag{70}$$

$$n_3 = n_1 - 2 \tag{71}$$

$$n_4 = 0 \tag{72}$$

It is easy to see that these values describes the graphs $U(n, n_1)$.

Figure 13 gives the results for $n = 24$ where the circles represent the value of the maximal chemical trees obtained by *AGX*, the dotted line is the upper bound of Theorem 4, and the black line is the new upper bound that we propose

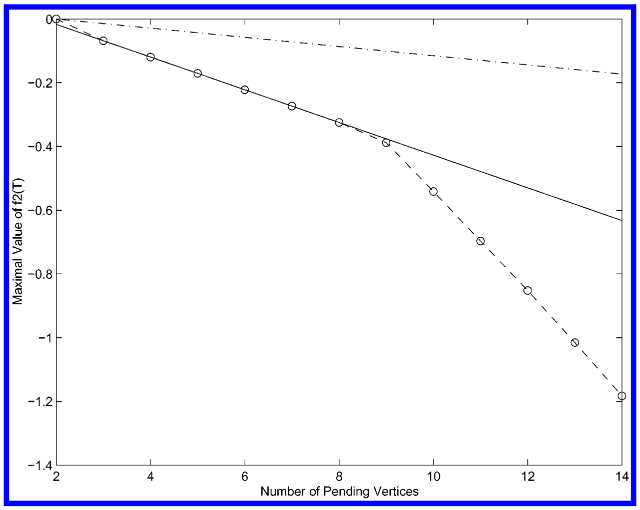

Figure 13. Maximal values for $f_2(T)$, $n = 24$.

in Theorem 10. For this value of n the bound is attained for six values of n_1. Observe that the bound cannot be extended to paths without losing much of its precision (as measured by the number of values of n_1 for which it is tight). Such border effects appear to be common in graph theoretic problems studied by *AGX*.

4. CORRECTION OF THEOREM 5

Let G be a simple graph and $\alpha = \{v_1, v_2\}$ an edge in G. Then $G(\alpha)$ is the graph obtained from G by deleting α. The following proposition is a correction of Proposition 3.3 of ref 1.

Proposition 11. *Let G be a chemical graph and* $\alpha = \{v_1, v_2\}$ *be an edge of G with* $d_1, d_2 \geq 2$. *Then*

$$\chi(G(\alpha)) - d_0' \leq \chi(G) \leq \chi(G(\alpha)) + b_0'$$

where

$$b_0' = (\sqrt{2} - 1)/2 \simeq 0.2071$$

and

$$d_0' = 2\sqrt{3} - \frac{13}{4} \simeq 0.2141$$

Proof. By definition of the connectivity index, we can check that

$$\chi(G) - \chi(G(\alpha)) = \sum_{v_1 - v_x \neq v_2} \frac{1}{\sqrt{d_x}}\left[\frac{1}{\sqrt{d_1}} - \frac{1}{\sqrt{d_1 - 1}}\right] + \sum_{v_2 - v_y \neq v_1} \frac{1}{\sqrt{d_y}}\left[\frac{1}{\sqrt{d_2}} - \frac{1}{\sqrt{d_2 - 1}}\right] + \frac{1}{\sqrt{d_1 d_2}}$$

The minimal value of this expression is reached when $d_x = 1$ for all $v_x \neq v_2$ adjacent to v_1 and $d_y = 1$ for all $v_y \neq v_1$ adjacent to v_2. The maximal value is reached when $d_x = 4$ for all $v_x \neq v_2$ adjacent to v_1 and $d_y = 4$ for all $v_y \neq v_1$ adjacent to v_2.

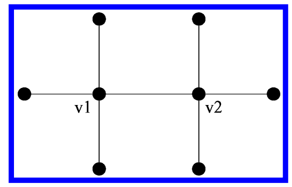

Figure 14. This graph reaches the Lower Bound of Proposition 11.

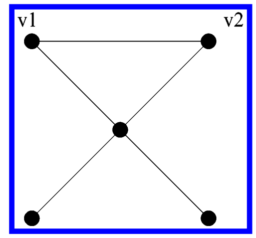

Figure 15. This graph reaches the Upper Bound of Proposition 11.

So

$$(d_1 - 1)\left[\frac{1}{\sqrt{d_1}} - \frac{1}{\sqrt{d_1 - 1}}\right] + (d_2 - 1)\left[\frac{1}{\sqrt{d_2}} - \frac{1}{\sqrt{d_2 - 1}}\right] + \frac{1}{\sqrt{d_1 d_2}} \leq \chi(G) - \chi(G(\alpha)) \leq \frac{1}{2}(d_1 - 1)\left[\frac{1}{\sqrt{d_1}} - \frac{1}{\sqrt{d_1 - 1}}\right] + \frac{1}{2}(d_2 - 1)\left[\frac{1}{\sqrt{d_2}} - \frac{1}{\sqrt{d_2 - 1}}\right] + \frac{1}{\sqrt{d_1 d_2}}$$

Next, we inspect all the possible cases for d_1 and d_2 ($\in\{2, 3, 4\}$ by assumption) for the lower bound and the upper bound. The minimum value for the difference ${}^1\chi(G) - {}^1\chi(G(\alpha))$ is reached when $d_1 = d_2 = 4$ and is equal to $-d_0'$. The maximum value for this difference is reached when $d_1 = d_2 = 2$ and is equal to b_0', and that shows the result.

The constants b_0' and d_0' are not the same than the constants b_0 and d_0 of Theorem 5. We have constructed examples for which the new bounds are reached. Thus these bounds are the best possible. The chemical graph of Figure 14 has a connectivity index which reaches the lower bound of Proposition 11, when deleting the edge $\{v_1, v_2\}$.

In the same way, the graph of Figure 15 has a connectivity index which reaches the upper bound.

Remark that these two graphs are counterexamples to Proposition 3.3 of ref 1.

We note that, by construction, the only possible graph which reaches the lower bound of Proposition 11 is that one described in Figure 14. This graph is a tree, and thus is not interesting for Theorem 5. This fact leads to a new proposition for chemicals graph with a cyclomatic number ≥ 1. The lower bound is better in this case.

Proposition 12. *Let G be a chemical graph for which the cyclomatic number* $g(G) \geq 1$. *Let* $\alpha = \{v_1, v_2\}$ *be an edge in a cycle of G. Then*

$${}^1\chi(G(\alpha)) - d_0^* \leq {}^1\chi(G) \leq {}^1\chi(G(\alpha)) + b_0'$$

where

$$d_0^* = \frac{2}{\sqrt{6}} + \frac{4}{\sqrt{3}} - \frac{1}{\sqrt{2}} - \frac{9}{4} \simeq 0.1688$$

Proof. Again, it is clear that

$$
{}^{1}\chi(G) - {}^{1}\chi(G(\alpha)) = \sum_{v_1 - v_x \neq v_2} \frac{1}{\sqrt{d_x}}\left[\frac{1}{\sqrt{d_1}} - \frac{1}{\sqrt{d_1 - 1}}\right] + \sum_{v_2 - v_y \neq v_1} \frac{1}{\sqrt{d_y}}\left[\frac{1}{\sqrt{d_2}} - \frac{1}{\sqrt{d_2 - 1}}\right] + \frac{1}{\sqrt{d_1 d_2}}
$$

Like in the proof of the previous Proposition, the maximal value is reached when $d_x = 4$ ($v_1 - v_x \neq v_2$) and $d_y = 4$ ($v_2 - v_y \neq v_1$). But the minimal value is no more reached when $d_x = d_y = 1$, because this leads to the special tree described in Figure 14. By hypothesis, G cannot be a tree because $g(G) \geq 1$. Moreover, the edge $\alpha = \{v_1, v_2\}$ must be on a cycle of G. So, there exists at least one vertex $v_x \neq v_2$, adjacent to v_1 and at least one vertex $v_y \neq v_1$, adjacent to v_2 for which d_x, $d_y \geq 2$, because v_x and v_y have to be connected by a path, different from $\{v_1, v_2\}$.

Hence

$$
\left(d_1 - 2 + \frac{1}{\sqrt{2}}\right)\left[\frac{1}{\sqrt{d_1}} - \frac{1}{\sqrt{d_1 - 1}}\right] + \left(d_2 - 2 + \frac{1}{\sqrt{2}}\right)\left[\frac{1}{\sqrt{d_2}} - \frac{1}{\sqrt{d_2 - 1}}\right] + \frac{1}{\sqrt{d_1 d_2}} \leq {}^{1}\chi(G) - {}^{1}\chi(G(\alpha))
$$

Again, by inspecting the possible values for d_1 and d_2 ($\in\{2, 3, 4\}$), we find d_0^* when $d_1 = d_2 = 4$.

The chemical graph of Figure 16 (with $g(G) \geq 1$) has a connectivity index which reaches the lower bound when deleting the edge $\{v_1, v_2\}$.

We can now give the following theorem, which is a correction of Theorem 5.

Theorem 13. *Let G be a chemical graph and T be a maximal subtree of G. Then*

$$
{}^{1}\chi(T) - d_0^* g(G) \leq {}^{1}\chi(G) \leq {}^{1}\chi(T) + b_0' g(G) \tag{73}
$$

In particular, if G has n vertices and T has n_1 pending vertices, then

$$
\frac{n}{2} + c_0' n_1 + d_0'' + d_0^* g(G) \leq \chi(G) \leq \frac{n}{2} - a_0' n_1 + b_0' g(G) \tag{74}
$$

where

$$
c_0' = \frac{1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) \simeq -0.1464
$$

and

$$
d_0'' = \frac{3}{2} - \sqrt{2} \simeq 0.0858
$$

Proof. If $g(G) = 0$, G is a tree, and assumption (73) is trivially verified because the only maximal subtree of a tree is the tree itself. If $g(G) = 1$, assumption (73) is correct by Proposition 12. By induction on $g(G)$ and by Proposition 12, assumption (73) still holds for $g(G) > 1$.

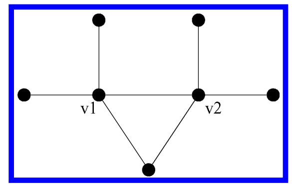

Figure 16. This graph reaches the Lower Bound of Proposition 12.

The second expression (74) is easily obtained when applying Theorem 8 to the lower bound of (73) and Theorem 10 to the upper bound of (73).

5. CONCLUSION

In this paper, the usefulness of the *AGX* system for computer-aided study of graphs is further illustrated by a discussion of its use in the analysis, correction and improvement of bounds on the connectivity index of chemical trees and graphs, recently proposed by Araujo and De la Peña.[1] The ramification index $r(T)$ of a tree T is shown to be equal to the number of pending vertices of T minus 2. General trees with minimum connectivity index, given order, and number of pending vertices are characterized, extending a result of Bollobás and Erdös:[55] they are comets.

In view of a series of presumably extremal graphs obtained with *AGX*, all bounds proposed by Araujo and De la Peña[1] are improved. Sharp lower and upper bounds on the connectivity index of trees as a function of this index for general trees and the ramification index (or, which is equivalent, the number of pending vertices) are obtained. Moreover, chemical trees of given order and number of pending vertices with minimum and with maximum connectivity index are characterized. Lower and upper bounds on the connectivity index of chemical graphs as a function of this index for induced maximal subgraphs which are trees are corrected and tight bounds obtained.

The theorems proved were first obtained as conjectures, which follow easily from the arrays of presumably extremal graphs given by *AGX*. These graphs can be recognized to belong to some known family, as, e.g., comets in the case leading to Theorem 7. Moreover, figures representing values of corresponding indices, or functions of indices, often lead to conjectures in a straightforward way. This is the case for Theorem 8 and especially for Theorem 10.

These results show that the computer can be used in many more ways than number crunching in the study of chemical graphs. Further automation of conjecture-finding functions of *AGX* is under way. We believe that *AGX*, as it stands, is already a very useful tool, which could be applied to the study of numerous graph invariants and relationships between them.

REFERENCES AND NOTES

(1) Araujo, O.; De la Peña, J. A. Some Bounds for the Connectivity Index of a Chemical Graph. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 827−831.
(2) *Structure-Based Drug Design: Experimental and Computational Approaches*; Codding, P. W., Ed.; Kluwer: Dordrecht, 1998.
(3) Devillers, J. *Neural Networks in QSAR and Drug Design*; Academic Press: London, 1996.
(4) *Topological Indices and Related Descriptors in QSAR and QSPR*; Devillers, J., Baladan, A., Eds; Gordon and Breach: Newark, 2000.
(5) Doucet, J.-P.; Weber, J. *Computer-Aided Molecular Design*; Academic Press: London, 1996.

(6) *Optimization in Computational Chemistry and Molecular Biology: Local and Global Approaches*; Floudas, C. A., Pardalos, P. M., Eds.; Kluwer: Dordrecht, 2000.
(7) Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR, Vols. I and II*; ACS Professional Reference Book, Oxford University Press: New York, 1995.
(8) Karelson, M. *Molecular Descriptors in QSAR/QSPR*; Wiley: New York, 2000.
(9) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
(10) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Wiley: New York, 1986.
(11) Kier, L. B.; Hall, L. H. *Molecular Structure Description*; Academic Press: London, 1999.
(12) Lutz, M.; Kenadin, T. *Quantitative Molecular Pharmacology and Informatics in Drug Discovery*; Wiley: New York, 1999.
(13) Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley: New York, 2001.
(14) *Structure−Property Correlations in Drug Research*; van de Waterbeemd, H., Ed.; Academic Press: London, 1996.
(15) *Computer-Assisted Lead Finding and Optimization*. van de Waterbeemd, H., Testa, B., Folkers, G., Eds.; Proceedings of the 11th European Symposium on QSAR, VCH: Weinheim, 1996.
(16) Cvetković, D.; Kraus, L. L.; Simic, S. K. Discussing Graph Theory with a Computer, I: Implementation of Graph Theoretic Algorithms. *Univ. Beograd Publ. Emektrotehn. Fak*. **1981**, 100−104.
(17) Cvetković, D.; Kraus, L. "Graph" an Expert System for the Classification and Extension of the Knowledge in the Field of Graph Theory, User's Manual. *Elektrotehn. Fak., Beograd*. **1983**.
(18) Cvetković, D.; Simic, S. Graph Theoretical Results Obtained by the Support of the Expert System "Graph". *Bull. l'Acad. Serbe Sci. Arts* (T. CVII): **1994**, 19−41.
(19) Knuth, D. *The Stanford Graphbase: A platform for Combinatorial Computing*; Addison − Wesley: Reading, MA, 1993.
(20) Berry, J. W. Improving Discrete Mathematics and Algorithms Curricula with LINK. *SIGCSE Bulletin (ACM Special Interest Group on Computer Science Education)* **1997**, *29*, 14−20.
(21) Berry, J. W. LINK: A Software System for Experimentation with Graphs and Hypergraphs. *SIAM Discrete Math. Newsletter* **1997**, *2*.
(22) Berry, J. W.; Dean, N.; Goldberg, M. K.; Shannon, G. E.; Skiena, S. Graph Drawing and Manipulation with LINK. *Lecture Notes Comput. Sci.* **1997**, *1353*, 425−437.
(23) Berry, J. W.; Dean, N.; Goldberg, M. K.; Shannon, G. E.; Skiena, S. LINK: A System for Graph Computation. *Software − Practice and Experience* **1999**.
(24) Vega. A System for Manipulating Discrete Mathematical Structures. Project Homepage: http://vega.ijp.si/.
(25) Dutton, R. D.; Brigham, R. C. Ingrid: A Software Tool for Extremal Graph Theory Research. *Congressum Numerantium* **1983**, *39*, 337−352.
(26) Brigham, R. C.; Dutton, R. D. A Compilation of Relations Between Graph Invariants. *Networks* **1985**, *15*, 73−107.
(27) Brigham, R. C.; Dutton, R. D.; Gomez, F. INGRID: A Graph Invariant Manipulator. *J. Symbolic Comput.* **1989**, *7*, 163−177.
(28) Brigham, R. C.; Dutton, R. D. A Compilation of Relations Between Graph Invariants. Supplement 1. *Networks* **1991**, *21*, 421−455.
(29) Fajtlowicz, S. On Conjectures of Graffiti. *Discrete Mathematics* **1988**, *72*, 113−118.
(30) Fajtlowicz, S. On Conjectures of Graffiti − II. *Congr. Numer.* **1987**, *60*, 187−197.
(31) Fajtlowicz, S. On Conjectures of Graffiti − III. *Congr. Numer.* **1988**, *66*, 23−32.
(32) Fajtlowicz, S. On Conjectures of Graffiti − IV. *Congr. Numer.* **1990**, *70*, 231−240.
(33) Fajtlowicz, S. On Conjectures of Graffiti − V. In *Seventh International Quadrennial Conference on Graph Theory*; 1995; Vol. 1, pp 367−376.
(34) Fajtlowicz, S. Written on the Wall. A regulary updated file accessible from http://www.math.uh.edu/∼clarson/.
(35) De la Vina, E. Bibliography on Conjectures of Graffitti.http://cms.dt.uh.edu/faculty/delavinae/research/wowref.htm, 2000.
(36) Caporossi, G.; Hansen, P. Variable Neighborhood for Extremal Graphs 1. The System AutoGraphiX. *Discr. Math.* **2000**, *212*, 29−44.
(37) Hansen, P.; Mladenovic, N. Variable Neighborhood Search: Principles and Applications. *Eur. J. Oper. Res.* **2001**, *130*, 449−467.
(38) Caporossi, G.; Hansen, P. Variable Neighborhood for Extremal Graphs 5. Three Ways to Automate Finding Conjectures. *Les Cahiers du GERAD*, G-2000-58. To appear in *Discrete Mathematics*.
(39) Caporossi, G.; Dobrynin, A. A.; Hansen, P.; Gutman, I. Trees with Palindromic Hosoya Polynomials. *Graph Theory Notes N. Y.* **1999**, *37*, 10−16.
(40) Caporossi, G.; Cvetkovic, D.; Hansen, P.; Simic, S. Variable Neighborhood Search for Extremal Graphs 3. On the Largest Eigenvalue of Color-Constrained Trees. *Lin. Multilin. Algebra* **2001**, *49*, 143−160.
(41) Aouchiche, M.; Caporossi, G.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 8. Variations on Graffiti 105. *Congressus Numerantium* **2001**, *148*, 129−144.
(42) Caporossi, G.; Cvetkovic, D.; Gutman, I.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 2. Finding Graphs with Extremal Energy. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 984−996.
(43) Gutman, I.; Hou, Y. Bipartite Unicyclic Graphs with Greatest Energy. *Comm. Math. Comput. Chem.* **2001**, *43*, 17−28.
(44) Caporossi, G.; Gutman, I.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 4. Chemical Trees with Extremal Connectivity Index. *Comput. Chem.* **1999**, *23*, 469−477.
(45) Caporossi, G.; Gutman, I.; Hansen, P.; Pavlovic, L. Graphs with Maximum Connectivity Index. *Les Cahiers du GERAD*, G-2001-50, 2001. *Computers and Chemistry* (in press).
(46) Gutman, I.; Miljković, O.; Caporossi, G.; Hansen, P. Alkanes with Small and Large Randic Connectivity Index. *Chem. Phys. Lett.* **1999**, *306*, 366−372.
(47) Gutman, I.; Araujo, O.; Morales, D. A. Bounds for the Randic Connectivity Index. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 572−598.
(48) Gutman, I.; Araujo, O.; Morales, D. A. Estimating the Connectivity Index of a Saturated Hydrocarbon. *Indian J. Chem.* **2000**, *39A*, 381−385.
(49) Gutman, I.; Araujo, O.; Rada, J. *Acta Chim. Hungarica Models Chem.* **2000**, *137*, 653.
(50) Gutman, I.; Miljkovic, O. Molecules with Smallest Connectivity Indices. *Comm. Math. Chem. (MATCH)* **2000**, *41*, 57−70.
(51) Fowler, P. W.; Hansen, P.; Caporossi, G.; Soncini, A. Polyenes with Maximum HOMO−LUMO Gap (Variable Neighborhood for Extremal Graphs 7.). *Chem. Physics Lett.* **2001**, *342*, 105−112.
(52) Shao, J. Y.; Hong, Y. Bounds on the Smallest Positive Eigenvalue of a Tree with Perfect Matchings. *Chinese Sci. Bull.* **1992**, *37*, 713−717.
(53) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.
(54) Yu, P. (in chinese). *J. Math. Studies* **1998**, *31*, 225.
(55) Bollobás, B.; Erdös, P. Graphs of Extremal Weights. *Ars Combinatoria* **1998**, *50*, 225−233.

CI010133J