UMONS

Université de Mons
Faculté des Sciences
Département d'Informatique

Faculté des Sciences

Extremal properties of average-based and chemical graph invariants

Sébastien Bonte

Thesis submitted in fulfilment of the requirements for the degree of
Docteur en Sciences

Jury

Pr. Dr. Hadrien Mélot, Director	Université de Mons
Pr. Dr. Alain Hertz, Co-Director	Polytechnique Montréal
Pr. Dr. Jef Wijsen, President	Université de Mons
Pr. Dr. Véronique Bruyère, Secretary	Université de Mons
Pr. Dr. Boris Furtula	University of Kragujevac, Serbia
Pr. Dr. Jan Goedgebeur	KU Leuven
Pr. Dr. Martine Labbé	Université Libre de Bruxelles

Thanks

I would first like to thank the referees, Dr. Véronique Bruyère, Dr. Jef Wijsen, Dr. Jan Goedgebeur, Dr. Boris Furtula, and Dr. Martine Labbé for their time in reading this thesis and their valuable feedback.

Je voudrais ensuite remercier mes collègues de recherche, et ça débute bien évidemment par Hadrien. Le monde de la recherche m'a rapidement intrigué lors de mon mémoire en Master, mais la réalité est que je ne m'imaginais pas vraiment rejoindre ce monde. Si tu ne m'y avais pas invité et convaincu de poursuivre une thèse, je ne l'aurais probablement jamais découvert. Et pourtant, quelle découverte que le monde des graphes. Ma vision du monde en a été partiellement transformée. Merci pour ça, et pour ton accueil, ta bienveillance et ton accompagnement constants.

Je souhaite également remercier chacun de mes autres collègues de recherche. Merci Alain, j'apprécie ton intérêt et ton talent dans des domaines si diversifiés. Je trouve superbe la façon dont tu as parfois pu les unir. Je pense par exemple à l'écriture et la théorie des graphes. Merci Gauvain, tu as toujours pris le temps de m'expliquer les choses et de m'aider, que ce soit pour toutes les tâches administratives ou pour certains détails techniques. Je ne l'oublie pas et je l'apprécie beaucoup. Merci Pierre, j'admire beaucoup ton talent tant sur le plan technique qu'en mathématiques. Merci Valentin, je t'adresse les mêmes compliments qu'à Pierre, et je suis admiratif devant la vitesse à laquelle tu as su t'intégrer à l'équipe et t'adapter. Travailler avec vous tous a toujours été un réel plaisir, que ce soit lors des réunions où les idées fusent de toute part, ou les conférences où l'on a partagé de bons moments. Je suis très fier des résultats que l'on a pu produire, et de la rigueur scientifique dont on a fait preuve.

Je voudrais également remercier mes autres collègues à la faculté des sciences et plus particulièrement au département d'informatique. J'ai grandement apprécié l'ambiance et la culture de ce département. Je sais que cette ambiance n'existe pas partout, et je tenais donc à souligner cette atmosphère collaborative, où chacun a quelque chose à apporter. Je remercie en particulier les personnes avec qui j'ai collaboré dans mes tâches pédagogiques. Je remercie les enseignants

et professeurs, mais aussi et surtout les autres assistants et doctorants. Merci de nouveau à Gauvain, Valentin et Pierre pour leur aide précieuse également sur ce plan, ainsi qu'à Jérémy, Guillaume, Horacio et Aziz, avec qui j'ai eu le plaisir de gérer ces TPs. Merci enfin à tous les élèves-assistants qui y ont participé.

Bien sûr, je veux également remercier ceux qui m'ont soutenu en dehors de l'université, tout au long de mon parcours. Je pense à ma famille, à mes amis, dont le soutien a été précieux. Je pense tout particulièrement à Elina Thibeau-Sutre, qui a été à mes côtés et m'a soutenu avec beaucoup de bienveillance durant ces dernières années de thèse.

Contents

Chapter 1

Introduction

This chapter serves as an introduction to this thesis. In Section 1.1, we introduce the motivations for our work, although for a more rigorous introduction to the different topics, the reader is advised to check the corresponding chapters. Section 1.2 mentions our contributions towards our objectives, and the publications concerning those contributions. Finally, Section 1.3 details the structure of this thesis.

1.1 Motivation and objectives

Graph theory is the study of graphs, which are mathematical structures used to model the pairwise relations between objects. A graph contains vertices and edges. The *vertices* are the objects whose relationships we want to model, and the *edges* are the aforementioned relationships. This simple, yet somewhat vague, definition can model a large number of situations. These models allow us to solve many problems with real-world applications.

In our studies, we take a special interest in graph *invariants*, which are properties of graphs that only depend on their abstract structure. This means that two graphs that represent the same structure will have the same values for a given invariant, which can take many forms. More specifically, our interest lies in *extremal graph theory*. It is a branch of graph theory that studies the bounds on numerical graph invariants, typically under specified constraints, such as bounds on other invariants.

In the same way that graphs can model a vast number of situations, our research has taken us in many different directions, which we organized in three categories for clarity.

Our first set of objectives concerns what we call average-based invariants. We define *counting-based invariants* as invariants whose values are obtained by

counting the number of occurrences of a specific configuration or substructure in the graph. Then, we define *average-based invariants* as invariants whose values are obtained by computing the average size of such configurations and substructures. This set of objectives is classical in extremal graph theory, as we focus on establishing lower and upper bounds for these invariants.

The second set of objectives takes place in the context of *chemical graph theory*, which is the study of *chemical graphs*. These are graphs used to model molecules, using vertices to represent atoms and edges to represent bonds. Within chemical graph theory, we focus on *topological indices*. These are invariants that characterize a graph's topology, thereby allowing the study of the properties of the molecule represented by the graph.

This set of objectives also started as a classical example of extremal graph theory, though it quickly evolved into something more ambitious. There are many topological indices. Indeed, there are so many that several prominent researchers in the field have argued that the proliferation of indices has become counterproductive. One of our objectives is also the study of how one might identify and discourage redundant or uninformative topological indices.

Finally, while these two first parts focus on theoretical results, our third set of objectives is more practical. Indeed, much of our time is also spent on developing tools to assist us in our research. This is, of course, something many researchers do, but we go further by making these tools publicly available. The objective is to enable other researchers to use them in their own work and, in some cases, to reproduce parts of our results or the process we used to obtain them in a practical and accessible way.

1.2 Contributions and publications

Our work has led to many contributions, many of which have already resulted in publications. Here, we mention the eight papers we worked on and the results they present.

Average-based invariants

Our first contribution regarding average-based invariants is our work on the average number of non-equivalent colorings of a graph. Firstly, we conjecture several lower bounds on this invariant. In an effort to prove them, we determine the value of this graph invariant for some classes of graphs. We then prove the validity of the conjectures for specific families of graphs such as chordal graphs and graphs with maximum degree at most 2. This is contained in a paper [118] that we worked on and that was published in *Discrete Applied Mathematics*.

Note that, as the proofs contained in this paper were already published in another thesis, we simply provide our final results in this document. The reader is referred to the paper for full proofs and intermediate results.

Secondly, we give a general upper bound that is valid for all graphs G. Since that upper bound is rather trivial, we then give a more precise one for graphs of order n and maximum degree $\Delta(G) \in \{1, 2, n-2\}$. This work is published in a second paper [119] which we participated in and which was published in *Graphs and Combinatorics*. Again, for the same reason, we provide only our final results in this document. Similarly, the reader is referred to the paper for full proofs and intermediate results.

Our second contribution in terms of average-based invariants concerns the average size of maximal matchings. While the lower and upper bounds of respectively $\frac{1}{2}$ and 1 of this invariant are rather trivial, we give a technique to characterize its asymptotic behavior for specific classes of graphs. We apply this technique for many different families of graphs, showcasing a varied spectrum of values between the two bounds. This work has also led to a publication [115] in which we participated and which was published in the *Journal of Combinatorial Optimization*.

Chemical graph theory and topological indices

Our first contribution to chemical graph theory is the study of the arithmetic-geometric index, a topological index. We give a sharp upper bound on its value for connected chemical (maximum degree at most 4) graphs of a given order and size. We characterize the connected chemical graphs that reach this given bound. Finally, we prove that removing the constraint of connectivity does not improve the bound. We published these results in a paper [114] in *MATCH Communications in Mathematical and in Computer Chemistry*.

Trying to apply the same techniques for other topological indices led us to our second contribution. We characterize extremal chemical (maximum degree at most 3) graphs that minimize or maximize 33 degree-based topological indices. We give five families of graphs that are sufficient to characterize those chemical graphs for 29 of those 33 degree-based topological indices. This shows that, in terms of extremal properties, topological indices can be very redundant. This is a powerful result, as it provides a criterion to examine the utility of new topological indices before they are submitted. This is explained in our paper [21] in *MATCH Communications in Mathematical and in Computer Chemistry*.

Our final contribution to this field goes even further. We give a complete polyhedral description of chemical graphs of maximum degree at most 3. This allows us, for any given order, size and degree-based topological index, to give

the corresponding sharp lower and upper bounds. We can also characterize the set of graphs that reach those bounds through the parameters m_{ij}, which count edges between vertices of degrees i and j. This result is found in a submitted paper [75]. We do not provide the proofs of our results in this thesis, as they will be published in the thesis of another co-author. You can, of course, find them in our paper. We also submitted a second paper [22] on this topic, written with chemists in mind, so that they can apply our results.

Tools for extremal graph theory

PHOEG, an online tool for Extremal Graph Theory, existed prior to our involvement. At that time, interaction with the system required direct SQL queries to the database. We therefore developed an accessible web interface, which is now publicly available, and submitted an accompanying article [24].

Our paper [22] dedicated to chemists on the complete polyhedral description of chemical graphs with maximum degree at most 3 introduces ChemicHull, a tool we developed to present and visualize these results. We hope that chemists will use this tool in their research on extremal graphs for degree-based topological indices.

1.3 Structure

After this introductory chapter, Chapter 2 presents the necessary definitions and notation. It also offers a more detailed introduction to graph theory, including a brief history of the field and formal definitions of the topics studied in this thesis.

After that, this document is structured into three parts matching our three categories of objectives.

Part I is about our work on average-based invariants. Chapter 3 defines average-based invariants and gives many examples of such invariants while mentioning studies about these invariants. It is not a survey, but it serves as a solid introduction to this part of the thesis. Chapter 4, is about the average number of colors in the non-equivalent colorings of a graph. Chapter 5 is about the average size of maximal matchings in graphs.

Part II is about chemical graph theory, although it is also mainly about topological indices. Chapter 6 introduces the topic of chemical graph theory and topological indices. Chapter 7 is based on our work on the extremal chemical graphs for the arithmetic-geometric index. Chapter 8 describes our work on extremal graphs of maximum degree at most 3 for 33 degree-based topological indices. Finally, Chapter 9 further generalizes our work by giving

a complete polyhedral description for chemical graphs of maximum degree at most 3.

Part III is dedicated to our tools developed for extremal and chemical graph theory. Chapter 10 details our work on the new web interface of PHOEG, an online tool for Extremal Graph Theory. Chapter 11 presents ChemicHull, an online tool enabling the visualization of the 96 polytopes previously described in Chapter 9.

Finally, Chapter 12 concludes our work. There, we summarize our work and the opportunities it presents for further research.

Chapter 2

Definitions and notations

This chapter provides basic notions, notations and definitions of graph theory useful for understanding the rest of this thesis. Most of these come from Diestel's Graph Theory [68]. Notation and definitions that are specific to individual chapters will be introduced separately within those chapters.

Section 2.1 introduces graph theory, presenting a brief historical overview along with examples and applications. Section 2.2 defines graphs as we will use them in this document as a whole. It also introduces some types of graphs commonly found in graph theory. In Section 2.3, we define what invariants are and give a few examples. In particular, we introduce the two types of invariants central to this work: average-based invariants and chemical invariants. Finally, Section 2.4 defines the field of extremal graph theory in which our research took place.

2.1 Graph theory and its usefulness

Graph theory is the study of *graphs*, which are mathematical structures used to model the pairwise relationships between objects. Graphs consist of vertices, which represent the objects, and edges, which represent the relationships between them.

Although this description may seem vague, its generality is intentional: many real-world systems can be represented as graphs. For example, Figure 2.1 illustrates a small social network with three women: Alice, Brooke, and Caroline. In this graph, vertices represent people and edges indicate friendships. Brooke is friends with both Alice and Caroline, while Alice and Caroline are not friends. Graphs like this can be used to model relationships in larger networks.

Another common example is modelling physical locations and the connections between them. To illustrate, we take cities as vertices, and if two

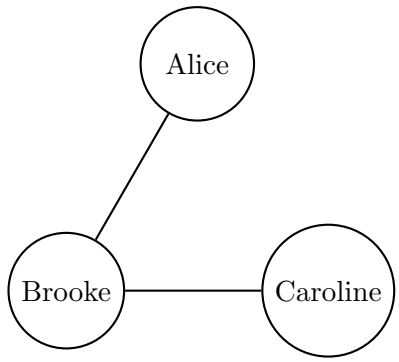


Figure 2.1: An example of a graph modelling friendship

cities are linked by a train route, an edge is drawn between them. Figure 2.2 demonstrates this concept by showing an approximation of a small section of the Belgian railway network. For instance, it indicates that one can travel between Tournai and Mons by train. This approach can, of course, be scaled to represent the entire Belgian railway network.

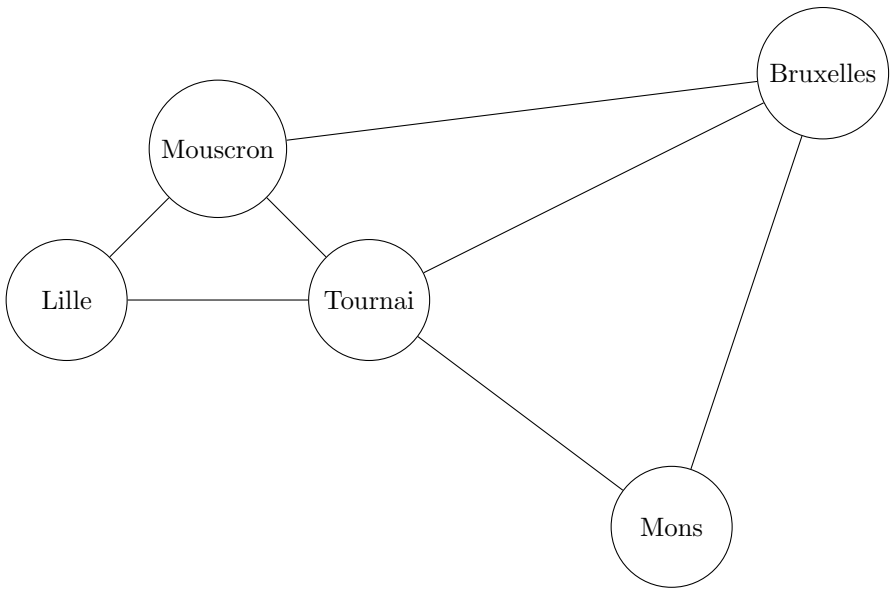


Figure 2.2: An example of a graph modelling part of the Belgian railway

The roots of graph theory are often traced to an example that also studies physical locations and the links between them. This famous example is called

the Königsberg bridge problem. This problem, solved in 1735, concerns the city of Königsberg, which was then part of the Kingdom of Prussia, but is now called Kaliningrad and is situated in Russia. The city was built around a river called the Pregel, and exactly seven bridges were built in order to cross the river and connect different part of the cities.

This situation is visually represented in Figure 2.3, where the river and the bridges are highlighted in blue and green, respectively. It is also modelled as a graph in Figure 2.4, where the vertices represent the different parts of the city and the edges correspond to the bridges linking them.

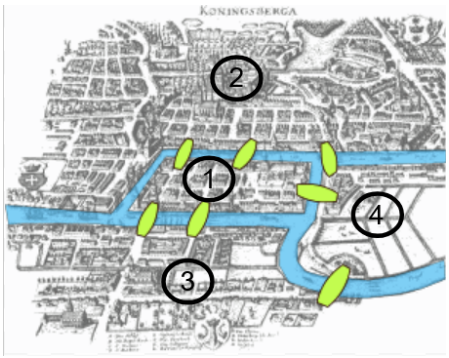


Figure 2.3: Drawing of Königsberg[1]and its seven bridges

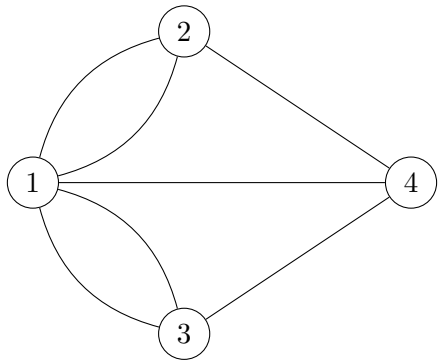


Figure 2.4: Graph modelling the Königsberg bridge problem

The puzzle asked whether one could cross all seven bridges of the city exactly once in a single trip, without revisiting a bridge. The trip must also finish right where it started. This is equivalent to finding an Eulerian cycle in the graph. Indeed, an *Eulerian cycle* is a cycle that visits every edge exactly once. As its names imply, it was Leonhard Euler [79] that solved this problem in 1736 by proving that no such trip exists. At the time, he conjectured that such a cycle exists if and only if each vertex has an even number of edges adjacent to it.

The intuitive idea behind this rule is that whenever you enter a part of the city, you must also leave it via another bridge. The only exceptions to this are the start and end of the trip, where you only need to respectively leave and enter that part of the city. Since the trip starts and ends at the same location, this applies to every part of the city. Consequently, each vertex in the corresponding graph must have an even number of edges, which ensures that you can enter and leave each vertex without getting stuck.

In terms of graphs, this equates to the conjecture stated above, which was

[1]Original drawing by Bogdan Giusca

only proved approximately 130 years later by Carl Hierholzer. This result later got published posthumously in 1873 [122]. Of course, Euler didn't use the representations we employ here; these were only introduced later by Sylvester in 1878 [200] in the context of chemistry, and that is where we will leave it at in terms of history.

The beauty and utility of graph theory lie in its ability to model a wide variety of problems. Even more interestingly, links can be made between similar problems that may not initially appear related. As a result, efficient algorithms or interesting results might already exist or be closely related to the problem we are looking at through the lens of graph theory. This means that it finds applications in many diverse fields. We list a few below, though countless others exist.

Graphs are widely used across the exact sciences. Of course, this includes mathematics and computer science, since graph theory first and foremost belongs to these fields, but they also have applications in subfields such as game theory. Graphs are important in chemistry as well, particularly in chemical graph theory, where molecules can be efficiently modeled as graphs. This thesis presents several such results, in the area of chemical graph theory.

It is also useful in biology, where it is, among other things, used to model migration patterns, protein relationships, evolution trees or ecological networks. In physics, it finds applications in quantum field theory and statistical physics, modeling connections between different parts of a system.

Graphs are also widely used for optimization tasks, such as scheduling (assigning tasks to timeslots) or logistics (planning supplies, transportation, delivery routes). This makes them useful in various engineering applications, particularly when designing networks of interconnected objects.

Finally, to illustrate the diversity of applications, we mention a few other wildly different ones: in epidemiology, graphs can model the spread of diseases; in sociology, they can represent the diffusion of rumors or patterns of social interaction; in linguistics, graphs are used to analyse the syntax or semantics of languages.

2.2 Graphs

As mentioned in Section 2.1, a graph is a structure used to model the relationships between vertices and edges. Formally, a graph G is defined as $G = (V, E)$ where V is the set of its vertices and $E \in V \times V$ is the set of its edges. The *order* of G, denoted $n = |V|$ or $|G|$, is its number of vertices. The *size* m of G, denoted $m = |E|$ or $||G||$, is its number of edges.

For example, the graph in Figure 2.5 has $V = \{1, 2, 3, 4, 5, 6\}$ and $E =$

$\{(1,2),(1,5),(1,6),(2,3),(2,4),(3,4),(4,5),(5,6)\}$. It has an order of 6 and a size of 8. We will refer to this graph as X, and it will serve as a running example throughout this chapter.

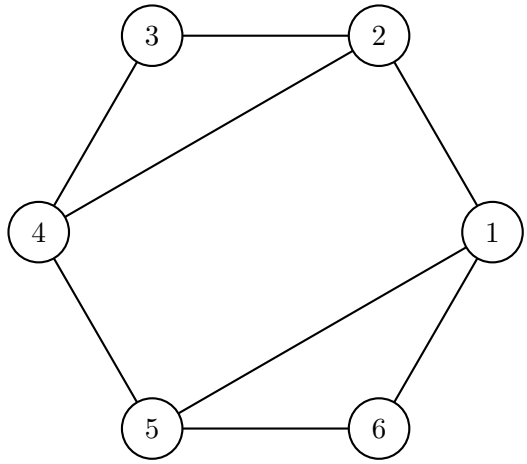


Figure 2.5: Example graph X

As mentioned, an edge links two vertices together. These two vertices are that edge's *extremities* or *endpoints*. They are *adjacent* to each other, which makes them each other's *neighbor*. We use the notation $N(v)$ for the set of neighbors of a vertex v. We also say that an edge is *incident* to both of its endpoints. The *degree* of a vertex u is the number of vertices adjacent to it and is denoted d_u. A vertex of degree 0 is an *isolated* vertex. A *dominant* vertex is a vertex that is adjacent to every other vertex in the graph.

In graph X, the vertices 1 and 2 are the endpoints of the edge $(1,2)$, and conversely the edge $(1,2)$ is incident to vertices 1 and 2 which are neighbors and adjacent to each other. The neighborhood of 1 is $\{2,5,6\}$; therefore the degree of 1 is 3.

A *simple* graph is a graph in which each pair of vertices is connected by at most one edge and no vertex is connected to itself. An *undirected* graph is a graph in which E consists of unordered pairs of vertices. Thus, (u,v) and (v,u) represent the same edge. By default, graphs are assumed to be simple undirected graphs. For example, Graph X is a simple undirect graph.

A graph can be drawn in different ways without changing which vertices are connected to each other. Two graphs are said to be *isomorphic* if they differ only by a relabeling of their vertices. Mathematically, two graphs $G = (V,E)$ and $H = (V',E')$ are isomorphic, denoted by $G \simeq H$, if and only if there exists a bijection $f : V \to V'$ such that $\forall u,v \in V, (u,v) \in E \iff (f(u),f(v)) \in E'$.

This bijection is called a *graph isomorphism*. For example, all three graphs shown in Figure 2.6 are isomorphic.

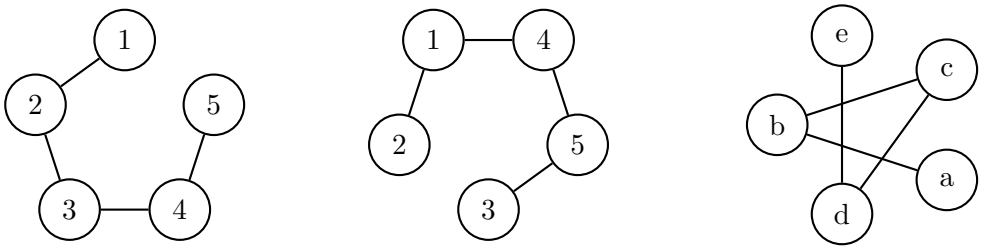


Figure 2.6: Three isomorphic graphs

Certain graphs arise frequently in theorems and proofs in graph theory, and this thesis is no exception. We therefore define them here.

The *complete graph*, denoted by K_n, is the graph of order n in which every pair of distinct vertices is joined by an edge. Figure 2.7 shows the complete graph on 5 vertices. Similarly, any subset of vertices where there is an edge between every pair of vertices is called a *clique*. The graph with no edges, denoted by $\overline{\mathsf{K}}_n$, is the *empty graph*. The empty graph on 5 vertices is drawn in Figure 2.8.

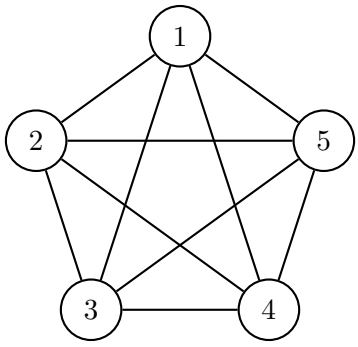


Figure 2.7: K_5

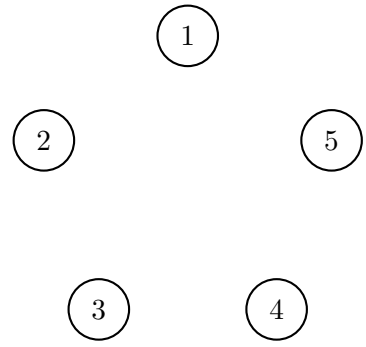


Figure 2.8: $\overline{\mathsf{K}}_5$

We denote by P_n the *path* of order n, a graph with $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{(v_i, v_{i+1}) | 1 \leq i \leq n-1\}$. Figure 2.9 shows the path on 5 vertices. By adding an edge (v_1, v_n) to E, we construct the *cycle* of order n, denoted by C_n. Figure 2.10 shows the cycle on 5 vertices.

A graph $G = (V, E)$ is *bipartite* if V can be partitioned into two disjoint sets U and W such that $U \cup W = V$, and every edge in E connects a vertex in U to a vertex in W. It is a *complete bipartite graph* if all possible edges between U and W are present. We denote by $\mathsf{K}_{a,b}$ the complete bipartite graph

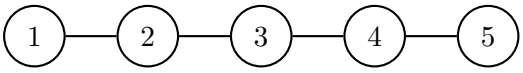


Figure 2.9: P_5

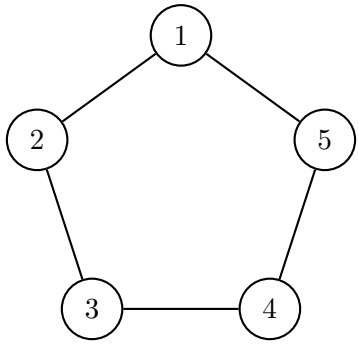


Figure 2.10: C_5

in which $|U| = a$ and $|W| = b$. The complete bipartite graph $\mathsf{K}_{2,4}$ is shown in Figure 2.11.

The *star* graph on n vertices, denoted by S_n, consists of a central vertex connected by an edge to every other vertex. It is equivalent to $\mathsf{K}_{1,n-1}$. Figure 2.12 shows the star on 6 vertices.

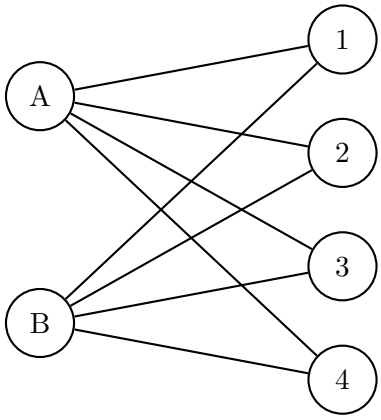


Figure 2.11: $\mathsf{K}_{2,4}$

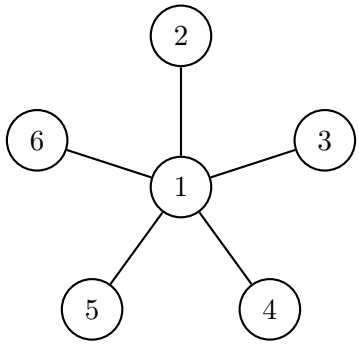


Figure 2.12: S_6 / $\mathsf{K}_{1,5}$

A *subgraph* $G' = (V', E')$ of a graph $G = (V, E)$ is formed by selecting a subset of the vertices and edges of G, with $V' \subseteq V$, $E' \subseteq E$ and every edge in E' having both endpoints in V'. The graph G' is an *induced subgraph* of G if, in addition, E' contains all edges of E whose endpoints are in V'. It is denoted $G[V']$.

For example, in Figure 2.13, $X' = (V', E')$ is a subgraph of $X = (V, E)$, since $V' \subseteq V$ and $E' \subseteq E$. However, X' is not an induced subgraph; for example, the edge $(1, 5) \notin E'$ even though $1, 5 \in V'$. An example of an induced

subgraph is shown in Figure 2.14, where X'' is an induced subgraph of X.

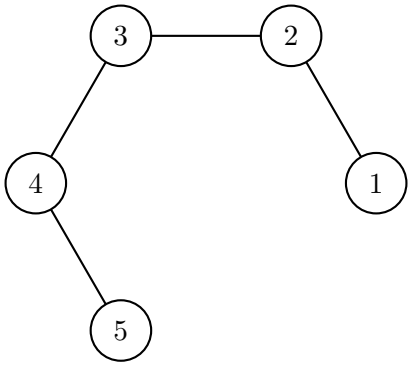


Figure 2.13: X', subgraph of X

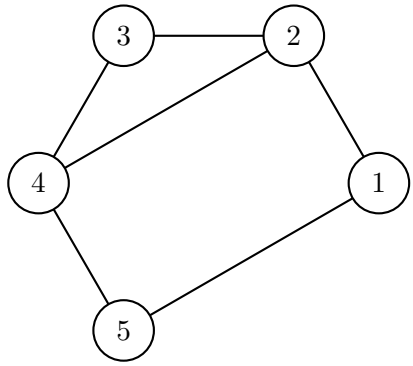


Figure 2.14: X'', induced subgraph of X

A graph $G = (V, E)$ is *connected* if every pair of vertices in V is *connected*; that is, if there exists a path between every pair of vertices. It is *acyclic* if there does not exist any induced cycle in G. A graph G is a *tree* if it is connected and acyclic. An example of a tree can be seen in Figure 2.15.

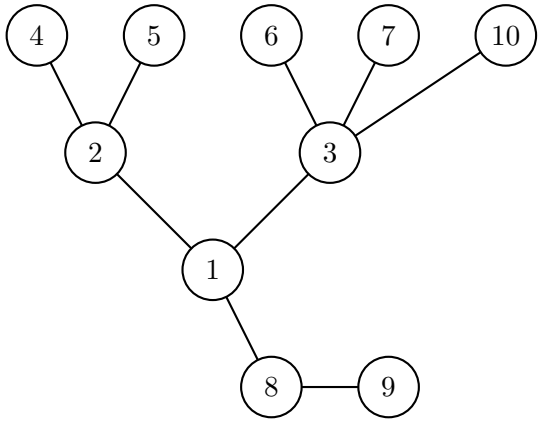


Figure 2.15: A tree

2.3 Invariants

An *invariant* of a graph is a property that depends only on the connections between its vertices, not on how the graph is drawn. Thus, two isomorphic graphs have the same value for any given invariant. More formally, a function

f on graphs is a graph invariant if, for any pair of graphs G_1 and G_2 such that $G_1 \simeq G_2$, it holds that $f(G_1) = f(G_2)$.

These values can belong to various types. For example, they can be numerical (such as $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{R}$), booleans, sets, or sequences. The number of graphs invariants one can imagine is infinite, although only a subset are of interest to researchers and thus studied.

A simple example of a graph invariant is the order n of a graph. Instead of just counting vertices, we can consider the properties of the vertices themselves. One such property is the degree of a vertex, but this is not a graph invariant, since the value depends on which vertex is chosen. However, the sorted sequence of the degrees of all vertices in a graph is indeed a graph invariant.

In the same way, the *minimum* and *maximum* degree of a graph, which are respectively the smallest and largest degree in a graph, are also invariants. For example, the minimum degree of X is 2 and the maximum degree is 3. Note that the maximum degree of a graph G is often denoted by $\Delta(G)$, while the minimum degree of G is denoted by $\delta(G)$.

There are invariants that measure the size of certain substructures of a graph. For example, the *girth* is the length of the shortest cycle in a graph, and the *circumference* is the length of the longest cycle. In graph X, the girth is 3 while the circumference is 6. The corresponding shortest cycle is shown in Figure 2.16 and the longest cycle is shown in Figure 2.17.

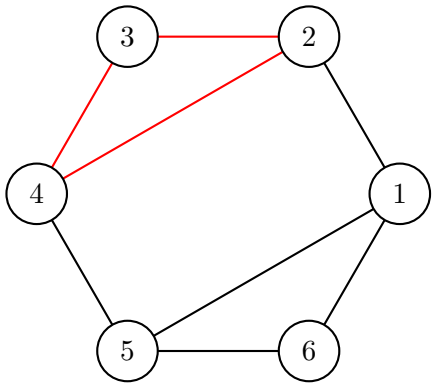


Figure 2.16: One of the shortest cycles of X

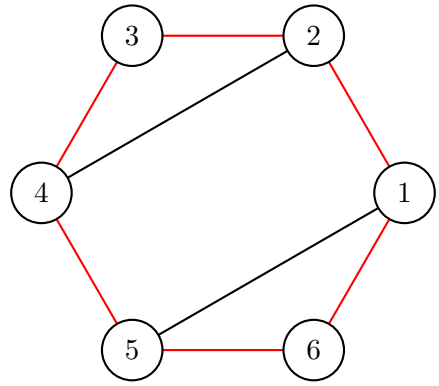


Figure 2.17: Longest cycle of X

The *distance* between two vertices is the number of edges in a shortest path connecting them. The *diameter* of a graph is the largest distance between any pair of its vertices. The diameter of X is 3, which is the distance between vertices 3 and 6.

Certain substructures in graphs are important in optimization problems.

One such substructure is a *matching*, which is a set of edges for which no two edges are incident. The *matching number* of a graph is the size of a largest matching in a graph. In graph X, the matching number is 3, and $\{(1,2),(3,4),(5,6)\}$ is one possible largest matching. This matching is illustrated in Figure 2.18.

An *independent set* is a set of vertices in which no two vertices are adjacent. The *independence number* of a graph is the size of a largest independent set. The independence number of X is 2, and $\{1,4\}$ is one possible largest independent set. This independent set is illustrated in Figure 2.19.

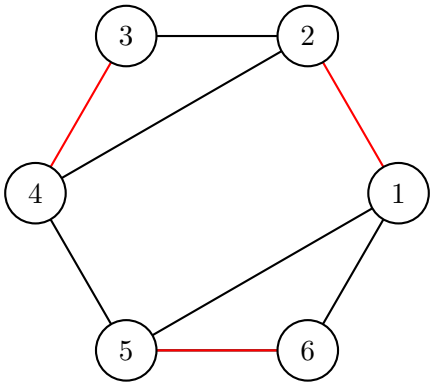


Figure 2.18: One of the largest matchings of X

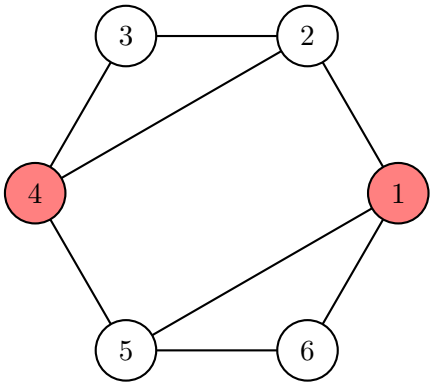


Figure 2.19: One of the largest independent sets of X

Graph coloring, a well-known problem in graph theory, also gives rise to important invariants. A *vertex coloring* is an assignment of colors to the vertices of a graph. Such a coloring is *proper* if no two adjacent vertices have the same color. The *chromatic number* of a graph is the smallest number of colors needed for a proper vertex coloring, often denoted by $\chi(G)$. A similar concept applies to edges, resulting in edge colorings and the associated invariants.

Graph X has a chromatic number of 3. For example, we can color vertices 1 and 4 with a first color, vertices 2 and 5 with a second color, and vertices 3 and 6 with a third color. It is not possible to properly color the graph with only 2 colors, since vertices 1, 5, and 6 are adjacent to one another. This coloring is illustrated in Figure 2.20.

As observed, there are many kinds of invariants, but this thesis focuses on two specific classes. In Part I, we study *average-based invariants*, which is not a term traditionally used in the literature. This class is defined based on another class, called *counting-based invariants*. Counting-based invariants are functions that assign to a graph G a value obtained by counting the number of distinct configurations, substructures, or combinatorial patterns of a specific

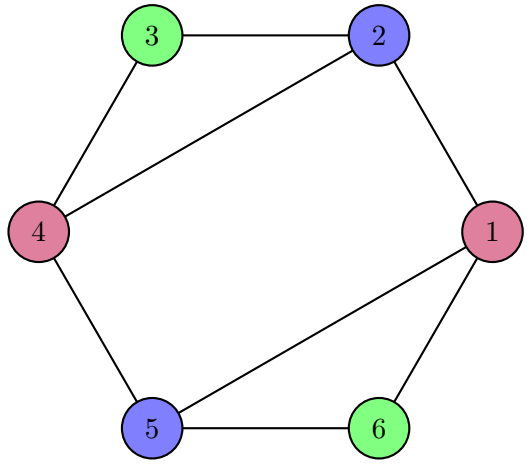


Figure 2.20: Example of a proper, minimal vertex coloring of X

type within G. Average-based invariants are defined as the average size of such configurations. This means that we calculate it by dividing the total size of all of them by their number. This links average-based invariants directly to counting-based invariants, as the first is often calculated thanks to the second. However, this does not mean that they necessarily have the same extremal properties.

The two examples of this first class of invariants that we explore are the average number of colors in the non-equivalent colorings of a graph in Chapter 4 and the average size of maximal matchings in graphs in Chapter 5.

In Part II, we consider chemical graph invariants. *Chemical graph theory* is the study of *chemical graphs*, which are connected graphs of maximum degree 3 or 4, depending on the definition. This distinction is briefly discussed in Section 6.1. Chapter 7 uses the definition based on a maximum degree of 4 while Chapter 8 uses the definition based on a maximum degree of 3.

Our interest in this field is mainly in *topological indices*, also called *molecular descriptors*, which are numerical graph invariants typically used to characterize the topology of a graph. They are often used to study the physicochemical properties of the molecules represented by these graphs. A degree-based topological index of a graph $G = (V, E)$ is a topological index that is computed from a sum of weights assigned to the edges of E, where each edge's weight depends on the degrees of its extremities. We further introduce this notion in Section 6.2.

We first investigate a specific chemical invariant, the arithmetic-geometric index, which is a topological index. It is defined in more details in Chapter 7.

We then generalize our work to other degree-based topological indices. More specifically, in Chapter 8 we show that five graph families are sufficient to characterize certain properties of 29 out of 33 studied degree-based topological indices.

Finally, to generalize our results further, we give a complete polyhedral description of chemical graphs whose maximum degree is at most 3. This description allows us to determine the minimum and maximum values chemical graphs can attain for a given order, size and degree-based topological index, when such values exist. We can also characterize the set of graphs that achieve these extremal values through the parameters m_{ij}, which count edges between vertices of degrees i and j. This result is explained and proven in our paper [75] and is presented in Chapter 9.

2.4 Extremal graph theory

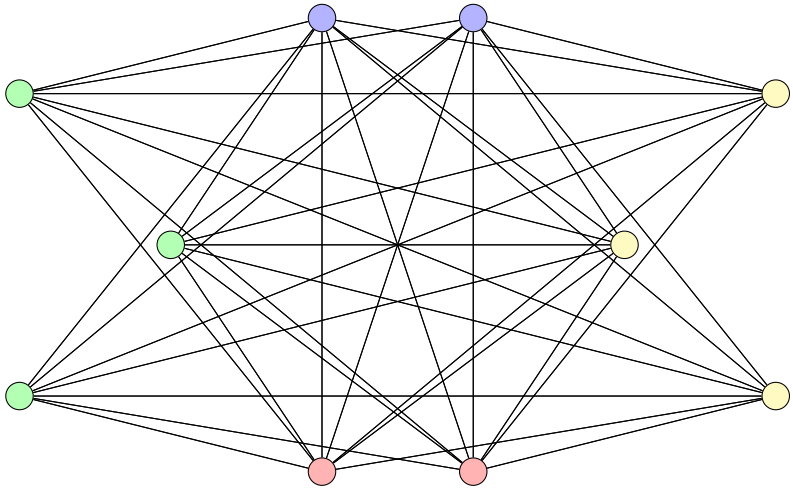
Extremal graph theory is a branch of combinatorics and graph theory that studies the maximum or minimum possible values of graph invariants under given constraints. In other words, it asks how large or small an invariant can be, given certain conditions. These constraints can be anything from a bound on the value of another invariant, or the presence or absence of specific subgraphs.

This can be viewed as an optimization problem, where the function being maximized or minimized is an invariant, and the set of feasible solutions consists of all graphs that satisfy the problem's constraints. Researchers are often also interested in characterizing the graphs that attain these bounds; such graphs are called *extremal graphs*.

A first simple example is the maximum value for the chromatic number, defined in Section 2.3. Without giving a detailed proof, it is clear that the chromatic number of a graph cannot exceed the maximum degree Δ plus one. The intuition is that if we had to use a $\Delta + 2^{th}$ color, some vertex would have degree $\Delta + 1$, which is impossible by the definition of Δ. This is a simple result, but additional constraints can yield tighter bounds. For example, Brooks' theorem [29] states that if we exclude complete graphs and cycles of odd order, the upper bound becomes Δ.

While such results are interesting in their own right, they can also be of practical use. For example, computing the chromatic number of a graph is an NP-complete problem, whereas computing its maximum degree is not. This allows us to obtain an upper bound much more quickly than computing the number. Similar situations arise for many other graph invariants. These bounds are valuable on their own, but they can also guide the choice of algorithm or be used to accelerate certain computational techniques.

Figure 2.21: $T(10, 4)$



Another example, which is a classical, well-known result in extremal graph theory, is Turán's theorem [206]. It states that for a given r with $r \leq n$, any graph of order n that does not contain a $(r+1)$-clique as a subgraph has at most as many edges as the Turán graph $T(n, r)$. In this context, the Turán graph is therefore an extremal graph.

Let q and s be integers such that $n = qr+s$. The Turán graph is constructed by forming $r - s$ independent sets of order q and s independent sets of order $q + 1$, which adds up to n vertices. Each vertex in a given independent set is adjacent to every vertex not in the same independent set. The Turán graph for $n = 10$ and $r = 4$ is shown in Figure 2.21 where $q = 2$ and $s = 2$ ($10 = 2 \times 4 + 2$).

The intuition here is that the Turán graph is as dense as possible while avoiding an $(r + 1)$-clique. Each independent set contains vertices that are not adjacent to each other, but are adjacent to every vertex outside this set, so adding any edge within one of these sets would create a clique of size $r + 1$.

Part I

About average-based invariants

Chapter 3

Average-based invariants

This chapter introduces average-based invariants, which we study in part I of this thesis. We first review counting-based invariants in Section 3.1, followed by average-based invariants in Section 3.2. In both sections, we provide definitions, followed by illustrative examples and related studies.

3.1 Counting-based invariants

This section defines counting-based invariants, presents representative examples from the literature, and defines different notions necessary to understand them.

An invariant is *counting-based* if its value for a graph G is determined by counting certain distinct substructures or combinatorial patterns within G. Choosing which structures are considered distinct can be nontrivial and may influence the invariant's properties. This effect is illustrated in Chapter 5, which concerns the average size of maximal matchings.

A *spanning tree* of a graph G is a subgraph of G that is a tree and includes all the vertices of G. The *complexity* of a graph $T(G)$ is the number of spanning trees in G. It is often simply called the number of spanning trees. This problem has a long history: the number of spanning trees can be computed using Kirchhoff's theorem [137], published in 1847. Since then, numerous studies [54, 55, 132, 133, 145, 154, 171, 180, 223] have investigated this problem, including recent work. This continued interest is motivated, for example, by practical applications, such as counting distinct communication or transportation routes, and modeling epidemic spread.

A *Hamiltonian cycle* of a graph G is a cycle that is a subgraph of G and contains all vertices of G. Some papers [87, 94, 112, 123, 190, 203] study bounds on the number of Hamiltonian cycles or how to generate graphs with a specified of them.

Ribeiro et al. [186] provide a comprehensive survey on counting induced or non-induced subgraphs in a graph. Although its primary focus is general subgraph counting, it also discusses results concerning the enumeration of specific subgraphs, including triads [191], stars [8, 96], and subtrees [143].

The *Hosoya-index* (or *Z-index*) $Z(G)$ of a graph G is an invariant equal to the number of matchings in G. It was introduced by Hosoya [124, 125] in 1971. Since then, many results [17, 38, 69, 121, 126, 127, 155, 156, 212, 213] have been published. We list only ten examples here; many more have appeared, including in recent years.

The *Fibonacci index* (or *Merrifield-Simmons index*) $F(G)$ or $\sigma(G)$ of a graph G is an invariant equal to the number of independent sets in G. Several articles consider both the Hosoya index and the Merrifield-Simmons index, which is unsurprising given their close relationship: each matching forms an independent set of edges. This index was introduced in 1982 by Prodinger and Tichy [182] and independently by Merrifield and Simmons [164] in 1989. The latter paper demonstrated correlations between the invariant and a chemical property. After that publication, many other results were published. For example, Véronique Bruyère and Hadrien Mélot [31] published some of these in 2009.

This last example was selected deliberately, as it partly motivated our work on similar invariants. In the aforementioned paper, Véronique Bruyère and Hadrien Mélot cite numerous other publications concerning the Fibonacci index in chemical graph theory [148, 150, 204, 212, 214, 215] and in extremal graph theory [120, 138, 178, 179, 208]. More recent results have also appeared concerning bounds on this invariant [90, 92, 169, 188, 189].

While there is a well-known duality between cliques and independent sets, the number of cliques in a graph has also been extensively studied in the literature. As usual, these papers[85, 86, 142, 158, 172, 217] investigate bounds on this invariant.

The *graphical Bell number* $\mathcal{B}(G)$ of a graph G is the number of non-equivalent colorings of G. This invariant is closely related to the *σ-polynomial*, first introduced in 1978 by Korfhage [139]. The σ-polynomial is defined as the polynomial in x whose coefficient of x^k equals the number of proper colorings of a graph using exactly k colors. Brenti [25, 26] extended this work and made further contributions. One might be tempted to think that $\mathcal{B}(G)$ is the value of the σ-polynomial where $x = 1$, however the σ-polynomial does not limit itself to non-equivalent colorings.

The graphical Bell number, as its name implies, is also closely related to the standard Bell number. Recall that the n^{th} Bell number represents the number of possible partitions of a set with n elements. Similarly, for an empty graph

on n vertices, the graphical Bell number counts the non-equivalent colorings of that graph. More generally, the graphical Bell number can be viewed as a constrained Bell number, where two vertices cannot be assigned the same color (i.e., belong to the same block of the partition) if they are adjacent in the graph.

Two recent results on this invariant are of particular interest to our research. In 2016, Alain Hertz and Hadrien Mélot [117] studied bounds on the graphical Bell number for graphs with a maximum degree constraint. In 2018, Romain Absil et al. [1] established a sharp lower bound on the number of non-equivalent colorings of graphs of order n and maximum degree $n-3$. We highlight these articles as they were authored by members of our research group. As noted in these two aforementioned papers, there have been other recent publications [73, 74, 91, 134] on this invariant.

3.2 Average-based invariants

In this second section, similarly to the first one, we define average-based invariants and present examples from the literature.

An invariant is *average-based* if, for a graph G, it is obtained by calculating the average size of distinct configurations, substructures, or combinatorial patterns of a specific type within G. Since they are closely related to counting-based invariants, we occasionally refer to invariants defined in Section 3.1.

Some of these invariants are obtained by averaging numerical characteristics of vertices that are defined relative to the rest of the graph. For example, the *eccentricity* of a vertex is the maximum distance from that vertex to any other vertex in the graph. The *average eccentricity* of a graph is defined as the average of the eccentricities of all its vertices. The notion of average eccentricity was introduced by Buckley and Harary [32] in 1990. Since its introduction, numerous works [9, 50, 52, 129, 177, 219] have investigated various properties and extremal aspects of this invariant.

Another invariant based on relations between vertices is the *mean distance*, which is the average distance between all pairs of vertices of the graph. This invariant was introduced by Doyle and Graver [72]. Since then, it has been studied extensively in the literature [42, 49, 51, 140, 168, 170].

Several publications [11, 58] investigate the average size of independent sets, which is an average-based analogue of the Merrifield-Simmons index. Similarly, some publications [27, 141] study the expected size of an independent set produced by a greedy algorithm. Although the notions of average size and expected size are distinct, the latter is a parallel problem often encountered alongside average-based invariants. This connection is natural, as the structures

underlying such invariants are frequently difficult to compute exactly and are therefore approached via greedy algorithms.

The graphical Bell number also has an associated average-based invariant, namely the average number of colors in the non-equivalent colorings of a graph. This invariant was first defined in a paper by our co-authors, Alain Hertz and Hadrien Mélot [116], as a way to derive inequalities for the Bell number. We, together with these co-authors, subsequently studied its lower [118] and upper [119] bounds, which are presented in greater detail in Chapter 4.

Another example of special interest is the average size of maximal matchings. This invariant has been studied in multiple works [12, 128]. In a similar way to the Merrifield-Simmons index, the expected size of a matching returned by a greedy algorithm was analyzed by Dyer and Frieze [76]. This is discussed in more detail about in Chapter 5, which presents our article [115] on this topic.

Chapter 4

On the average number of colors in the non-equivalent colorings of a graph

This chapter describes our study of the average number of colors in the non-equivalent colorings of a graph. It is based on our two articles [118, 119], although the proofs are not given here. While we contributed to this work, the proofs were already published in a co-author's thesis. The reader is invited to consult the respective papers for the proofs. This chapter introduces the problem and presents the main results.

Section 4.1 mentions previous works and provides an introduction to this invariant. Section 4.2 presents a detailed definition of the problem and establishes the necessary notation. Section 4.3 is devoted to the results of our first paper [118] on this topic, which concerns the lower bounds. Section 4.4 presents the upper bounds obtained in our second paper [119]. Finally, Section 4.5 summarizes the lower and upper bounds and concludes with an open problem.

4.1 Introduction

The total *number* $\mathcal{B}(G)$ *of non-equivalent colorings* (i.e., with different partitions into color classes) of a graph G is the number of partitions of the vertex set of G whose blocks are independent sets (i.e., sets of pairwise non-adjacent vertices). This invariant has been studied by several authors in the last few years [1, 73, 74, 91, 117, 134] under the name of (graphical) Bell number.

Recently, Hertz et *a*l. [116] have defined a new graph invariant $\mathcal{A}(G)$ which is equal to the average number of colors in the non-equivalent colorings of a

graph G.

It can be seen as a generalization of a concept linked to Bell numbers. More precisely, the Bell numbers $(\mathsf{B}_n)_{n\geq 0}$ count the number of different ways to partition a set that has exactly n elements. The 2-Bell numbers $(\mathsf{T}_n)_{n\geq 0}$ count the total number of blocks in all partitions of a set of n elements. Odlyzko and Richmond [175] have studied the average number A_n of blocks in a partition of a set of n elements, which can be defined as $\mathsf{A}_n = \frac{\mathsf{T}_n}{\mathsf{B}_n}$.

The graph invariant $\mathcal{A}(G)$ that we study in this chapter generalizes A_n. Indeed, when constraints (represented by edges in G) impose that certain pairs of elements (represented by vertices) cannot belong to the same block of a partition, $\mathcal{A}(G)$ is the average number of blocks in the partitions that respect all constraints. Hence, for a graph of order n, $\mathcal{A}(G) = \mathsf{A}_n$ if G is the empty graph of order n.

The close link between Bell numbers and graph colorings indicates that it is possible to use graph theory to discover nontrivial inequalities for the Bell numbers. For example, as shown in [116], $\mathcal{A}(\mathsf{P}_n) = \frac{\mathsf{B}_n}{\mathsf{B}_{n-1}}$ and $\mathcal{A}(\mathsf{P}_n) < \mathcal{A}(\mathsf{P}_{n+1})$ for $n \geq 1$, where P_n is the path on n vertices. This immediately implies $\mathsf{B}_n^2 < \mathsf{B}_{n-1}\mathsf{B}_{n+1}$, which means that the sequence $(\mathsf{B}_n)_{n\geq 0}$ is strictly log-convex. This result has also been proved recently by Alzer [10] using numerical arguments.

Very little is known about $\mathcal{A}(G)$ and there is to date no study concerning extremal properties of this graph invariant. The best possible upper bound for $\mathcal{A}(G)$ is clearly the order n of G since all colorings of G use at most n colors and $\mathcal{A}(\mathsf{K}_n) = n$ for the clique K_n of order n. It seems however much more complex to define a lower bound for $\mathcal{A}(G)$, as a function of n, which is reached by at least one graph of order n. We think that the best possible lower bound is reached by the empty graph $\overline{\mathsf{K}}_n$ of order n (i.e., the graph of order n with no edges).

4.2 Definitions

This section presents a few notations useful for our problem, and then presents a more formal definition.

Given two graphs G_1 and G_2 (with disjoint sets of vertices), we write $G_1 \cup G_2$ for the *disjoint union* of G_1 and G_2, and pG is the disjoint union of p copies of G. Also, $G \cup p\mathsf{K}_1$ is the graph obtained from G by adding p isolated vertices.

A vertex v is *isolated* if $|N(v)| = 0$. We write $\Delta(G)$ for the maximum degree of G. A vertex v of a graph G is *simplicial* if the induced subgraph $G[N(v)]$ of G is a clique. A graph is *chordal* if each of its induced subgraphs contains a simplicial vertex.

Let u and v be two vertices in a graph G of order n. Then, $G - v$ is the graph obtained from G by removing v and all its incident edges.

A *coloring* of a graph G is an assignment of colors to the vertices of G such that adjacent vertices have different colors. The *chromatic number* $\chi(G)$ of G is the minimum number of colors in a coloring of G. Two colorings are *equivalent* if they induce the same partition of the vertex set into color classes. Let $S(G, k)$ be the number of non-equivalent colorings of a graph G that use *exactly* k colors. Then, the total *number* $\mathcal{B}(G)$ *of non-equivalent colorings* of a graph G is defined by

$$\mathcal{B}(G) = \sum_{k=1}^{n} S(G, k) = \sum_{k=\chi(G)}^{n} S(G, k),$$

and the total *number* $\mathcal{T}(G)$ *of color classes in the non-equivalent colorings* of a graph G is defined by

$$\mathcal{T}(G) = \sum_{k=1}^{n} kS(G, k) = \sum_{k=\chi(G)}^{n} kS(G, k).$$

The *average number* $\mathcal{A}(G)$ *of colors in the non-equivalent colorings* of a graph G can therefore be defined as

$$\mathcal{A}(G) = \frac{\mathcal{T}(G)}{\mathcal{B}(G)}.$$

Note that $\mathcal{B}(\overline{\mathsf{K}}_n) = \mathsf{B}_n, \mathcal{T}(\overline{\mathsf{K}}_n) = \mathsf{T}_n$, and $\mathcal{A}(\overline{\mathsf{K}}_n) = \mathsf{A}_n$. As another example, consider the complement $\overline{\mathsf{P}}_5$ of a path on 5 vertices. As shown in Figure 4.1, there are three non-equivalent colorings of $\overline{\mathsf{P}}_5$ with 3 colors, four with 4 colors, and one with 5 colors, which gives $\mathcal{B}(\overline{\mathsf{P}}_5) = 8$, $\mathcal{T}(\overline{\mathsf{P}}_5) = 30$ and $\mathcal{A}(\overline{\mathsf{P}}_5) = \frac{30}{8} = 3.75$.

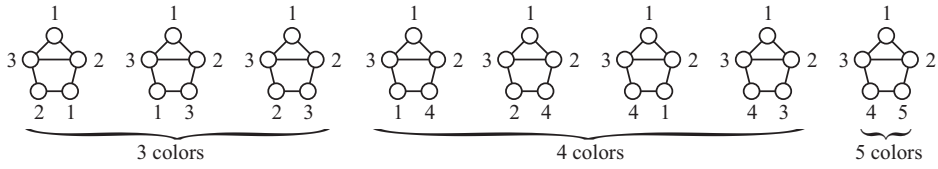


Figure 4.1: The non-equivalent colorings of $\overline{\mathsf{P}}_5$.

4.3 Lower bounds

This section concerns our results on the lower bounds on the average number of colors [118]. First, Subsection 4.3.1 introduces this topic and shows that it

is not as trivial as it may seem. Then, in Subsection 4.3.2, we dive into our results.

4.3.1 Conjectures

Conjecture 1. *Let G be a graph of order n. Then,*

$$\mathcal{A}(G) \geq \mathcal{A}(\overline{\mathsf{K}}_n)$$

with equality if and only if G is isomorphic to $\overline{\mathsf{K}}_n$.

Note that despite the apparent simplicity of Conjecture 1, its validity cannot be proven by simple intuitive means such as sequential edge removal. Indeed, there are graphs G for which the removal of any edge strictly increases $\mathcal{A}(G)$. This is the case, for example, for the complete bipartite graph with two vertices in one set of the bipartition and four vertices in the other set. Also, we cannot proceed by induction on the number of connected components of G. Indeed, there are pairs of graphs G_1, G_2 such that $\mathcal{A}(G_1) < \mathcal{A}(G_2)$ while $\mathcal{A}(G_1 \cup \mathsf{K}_1) > \mathcal{A}(G_2 \cup \mathsf{K}_1)$. For example, for $G_1 = \mathsf{K}_{2,3}$ and $G_2 = \mathsf{K}_3 \cup 2\mathsf{K}_1$, we have

$$\mathcal{A}(G_1) = 3.5 < 3.529 = \mathcal{A}(G_2),$$
$$\mathcal{A}(G_1 \cup \mathsf{K}_1) = 3.867 > 3.831 = \mathcal{A}(G_2 \cup \mathsf{K}_1).$$

The next two conjectures are stronger in the sense that it suffices to show that one of them is true to prove the validity of Conjecture 1. Let $G \cup p\mathsf{K}_1$ be the graph obtained from G by adding p isolated vertices, let K_n be the clique of order n, and let $\mathsf{K}_{1,n-1}$ be the star of order n (i.e. the graph with one vertex of degree $n-1$ and $n-1$ vertices of degree 1).

Conjecture 2. *Let G be a graph of order n. Then,*

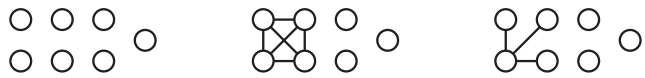
$$\mathcal{A}(G) \geq \mathcal{A}(\mathsf{K}_{\chi(G)} \cup (n - \chi(G))\mathsf{K}_1)$$

with equality if and only if G isomorphic to $\mathsf{K}_{\chi(G)} \cup (n - \chi(G))\mathsf{K}_1$.

Conjecture 3. *Let G be a graph of order n. Then*

$$\mathcal{A}(G) \geq \mathcal{A}(\mathsf{K}_{1,\Delta(G)} \cup (n - \Delta(G) - 1)\mathsf{K}_1)$$

with equality if and only if G is isomorphic to $\mathsf{K}_{1,\Delta(G)} \cup (n - \Delta(G) - 1))\mathsf{K}_1$, where $\Delta(G)$ is the maximum degree of G.



(a) (b) (c)

Figure 4.2: Three graphs that reach the lower bounds of Conjectures 1, 2 and 3.

The three conjectures come from the discovery systems *GraPHedron* [163] and *PHOEG* [66]. For illustration, by exhaustive enumeration, we have checked that:

- the graph of Figure 4.2 (a) minimizes $\mathcal{A}(G)$ among all graphs G of order 7;
- the graph of Figure 4.2 (b) minimizes $\mathcal{A}(G)$ among all graphs G of order 7 and chromatic number $\chi(G) = 4$;
- the graph of Figure 4.2 (c) minimizes $\mathcal{A}(G)$ among all graphs G of order 7 and maximum degree $\Delta(G) = 3$.

4.3.2 Results

The following Corollary is useful to better understand our results.

Corollary 4. *If v is a dominant vertex of a graph G, then,*

$$\mathcal{A}(G) = \mathcal{A}(G - v) + 1.$$

The lower bounds we are interested in depend on two parameters n and r with $1 \leq r \leq n$. They are equal to $\mathcal{A}(G)$ for some specific graphs G. More precisely, we define

- $L_1(n) = \mathcal{A}(\overline{\mathsf{K}}_n) = \dfrac{\mathsf{B}_{n+1} - \mathsf{B}_n}{\mathsf{B}_n}$,
- $L_2(n,r) = \mathcal{A}(\mathsf{K}_r \cup (n-r)\mathsf{K}_1) = \dfrac{\displaystyle\sum_{k=r}^{n} k \sum_{i=0}^{r} \binom{k-i}{r-i}\binom{r}{i}(r-i)!\left\{{n-r \atop k-i}\right\}}{\displaystyle\sum_{k=r}^{n}\sum_{i=0}^{r} \binom{k-i}{r-i}\binom{r}{i}(r-i)!\left\{{n-r \atop k-i}\right\}}$,
- $L_3(n,r) = \mathcal{A}(\mathsf{K}_{1,r-1} \cup (n-r)\mathsf{K}_1) = \dfrac{\displaystyle\sum_{i=0}^{n-r} \binom{n-r}{i}\mathsf{B}_{r+i}}{\displaystyle\sum_{i=0}^{n-r} \binom{n-r}{i}\mathsf{B}_{r+i-1}}$.

Given a graph G of order n, we are interested in the following inequalities:

$$\begin{aligned} L_1(n) &\leq \min\Big\{L_2(n,\chi(G)), L_3(n,\Delta(G)+1)\Big\} \\ &\leq \max\Big\{L_2(n,\chi(G)), L_3(n,\Delta(G)+1)\Big\} \\ &\leq \mathcal{A}(G). \end{aligned}$$

The first inequality is proved in the original paper. Hence, since $L_1(n) \leq \min\{L_2(n,\chi(G)), L_3(n,\Delta(G)+1)\}$, it suffices to show that one of these two conjectures is true to prove that Conjecture 1 is also true. Note that proving that Conjecture 3 is true for all graphs G of order n and maximum degree $\Delta(G) = n-1$ is as difficult as proving Conjecture 1. Indeed, let v be a vertex of degree $n-1$ in a graph G of order n. Since v is a dominant vertex of G, we know from Corollary 4 that $\mathcal{A}(G) = \mathcal{A}(G-v)+1$. Hence, minimizing $\mathcal{A}(G)$ is equivalent to minimizing $\mathcal{A}(G-v)$, with no maximum degree constraint on $G-v$.

Theorem 5. *Let G be a graph of order n with $\Delta(G) = 2$. Then,*

$$\mathcal{A}(G) \geq L_2(n,\chi(G)),$$

with equality if and only if $G \simeq \mathsf{K}_{\chi(G)} \cup (n-\chi(G))\mathsf{K}_1$. [118]

Theorem 6. *Let G be a graph of order n with $\Delta(G) = 2$. Then,*

$$\mathcal{A}(G) \geq L_3(n,\Delta(G)+1),$$

with equality if and only if $G \simeq \mathsf{K}_{1,2} \cup (n-3)\mathsf{K}_1$. [118]

As it has been explained multiple times, proving it for $\Delta(G) > 2$ is no simple task, but we do know it to be true for chordal graphs.

Theorem 7. *Conjectures 2 and 3 (and therefore 1) are true for chordal graphs. [118]*

4.4 Upper bounds

The previous section mentions that there is no known lower bound on $\mathcal{A}(G)$ which is a function of n and such that there exists at least one graph of order n which reaches it. The situation is not the same for the upper bound. Indeed, we state that there is an upper bound on $\mathcal{A}(G)$ which is a function of n and such that there exists exactly one graph of order n which reaches it.

Theorem 8. *Let G be a graph of order n, then,*

$$\mathcal{A}(G) \leq n,$$

with equality if and only if $G \simeq \mathsf{K}_n$. [119]

Since $\Delta(\mathsf{K}_n) = n - 1$ we immediately get the following corollary to Theorem 8.

Corollary 9. *Let G be a graph of order n and maximum degree $\Delta(G) = n-1$. Then, $\mathcal{A}(G) \leq n$, with equality if and only if $G \simeq \mathsf{K}_n$.*

We also give a sharper upper bound for graphs with maximum degree $\Delta(G) \in \{1, 2, n-2\}$, which are results from our paper [119] on this topic, where one could find their proofs.

Theorem 10. *Let G be a graph of order $n \geq 2$ and maximum degree $\Delta(G) = n - 2$. Then,*

$$\mathcal{A}(G) \leq \frac{n^2 - n + 1}{n},$$

with equality if and only if $G \simeq \mathsf{K}_{n-1} \cup \mathsf{K}_1$. [119]

Theorem 11. *Let G be a graph of order n and maximum degree $\Delta(G) = 1$. Then,*

$$\mathcal{A}(G) \leq \mathcal{A}(\left\lfloor \frac{n}{2} \right\rfloor \mathsf{K}_2 \cup (n \bmod 2)\mathsf{K}_1)$$

with equality if and only if $G \simeq \lfloor \frac{n}{2} \rfloor \mathsf{K}_2 \cup (n \bmod 2)\mathsf{K}_1$. [119]

Theorem 12. *If G is a graph of order $n \geq 3$ and maximum degree $\Delta(G) = 2$, then $\mathcal{A}(G) \leq \mathcal{A}(\mathsf{U}_n)$, with equality if and only if $G \simeq \mathsf{U}_n$, where*

$$\mathsf{U}_n = \begin{cases} \frac{n}{3}\mathsf{K}_3 & \text{if } n \text{ mod } 3 = 0, \text{ and } n \geq 3, \\ \frac{n-1}{3}\mathsf{K}_3 \cup \mathsf{K}_1 & \text{if } n = 4 \text{ or } n = 7, \\ \frac{n-4}{3}\mathsf{K}_3 \cup \mathsf{C}_4 & \text{if } n \text{ mod } 3 = 1, \text{ and } n \geq 10, \\ \frac{n-5}{3}\mathsf{K}_3 \cup \mathsf{C}_5 & \text{if } n \text{ mod } 3 = 2, \text{ and } n \geq 5. \text{ [119]} \end{cases}$$

4.5 Conclusion

We gave bounds for $\mathcal{A}(G)$, the number of colors in the non-equivalent colorings of a graph G. It is easy to prove that $\mathcal{A}(G) \leq \mathcal{A}(\mathsf{K}_n) = n$ for all graphs of order n, with equality if and only if $G \simeq \mathsf{K}_n$. Hence, n is the best possible upper bound on $\mathcal{A}(G)$ for a graph G of order n.

We think that the best possible lower bound on $\mathcal{A}(G)$ for a graph G of order n is $\mathcal{A}(\overline{\mathsf{K}}_n) = \frac{\mathsf{B}_{n+1} - \mathsf{B}_n}{\mathsf{B}_n}$. We have shown that despite its apparent simplicity,

this conjecture cannot be proven using simple techniques like sequential edge removal. We have then refined this conjecture by proposing lower bounds related to the chromatic number $\chi(G)$ and to the maximum degree $\Delta(G)$ of G. We have thus stated three open problems. We have shown that these three conjectures are true for chordal graphs and for graphs with maximum degree at most 2.

We have given a general upper bound on $\mathcal{A}(G)$ that is valid for all graphs G, and a more precise one for graphs of order n and maximum degree $\Delta(G) \in \{1, 2, n-2\}$. Note that there is no known lower bound on $\mathcal{A}(G)$ which is a function of n and such that there exists at least one graph of order n which reaches it.

The problem of finding a tight upper bound for graphs with maximum degree in $\{3, \ldots, n-3\}$ remains open. Since all graphs of order n and maximum degree $\Delta(G) \in \{1, n-2, n-1\}$ that maximize $\mathcal{A}(G)$ are isomorphic to $\left\lfloor \frac{n}{\Delta(G)+1} \right\rfloor \mathsf{K}_{\Delta(G)+1} \cup \mathsf{K}_{n \bmod (\Delta(G)+1)}$ (but this is not always true for $\Delta(G) = 2$), one could be tempted to think that this is also true when $3 \leq \Delta(G) \leq n-3$. We have checked this statement by enumerating all graphs having up to 12 vertices, using *PHOEG* [66]. We have thus determined that there is only one graph of order $n \leq 12$ and $\Delta(G) \neq 2$ (among more than 165 billion), namely $\overline{\mathsf{C}}_6 \cup \mathsf{K}_4$, for which such a statement is wrong. Indeed, $\mathcal{A}(\overline{\mathsf{C}}_6 \cup \mathsf{K}_4) = 5.979 > 5.967 = \mathcal{A}(2\mathsf{K}_4 \cup \mathsf{K}_2)$, which shows that $2\mathsf{K}_4 \cup \mathsf{K}_2$ does not maximize $\mathcal{A}(G)$ among all graphs of order 10 and maximum degree 3.

Chapter 5

The average size of maximal matchings in graphs

This chapter presents our work on the average size of maximal matchings in graphs. It is based on our article [115]. For the most part, the content of this chapter follows that article, apart from minor changes for coherence. Section 5.1 provides an introduction to the topic and outlines the remainder of the chapter.

5.1 Introduction

A *matching* in a graph $G = (V, E)$ is a set $M \subseteq E$ of edges without common vertices. If M is not a subset of any other matching in G, then it is *maximal* and a matching of maximum size is called a *maximum matching*. Maximal matchings are also called *independent edge dominating sets* [218]. Obviously, all maximum matchings are maximal but the converse does not always hold. Matching theory is a core subject in graph theory and has found applications in several domains such as networks, social science or chemistry (see the book of Lovász and Plummer [157] for a complete overview of matching theory). The main objective of this chapter is to answer the following natural question.

> *Given a graph, are there many maximal matchings that are significantly different in size than a maximum matching?*

Let $\nu(G)$ be the size of a maximum matching in G, let $\boldsymbol{M}(G)$ be the set of maximal matchings in G, let $\mathcal{T}_0(G) = |\boldsymbol{M}(G)|$ be the number of maximal matchings in G and let $\mathcal{T}_1(G)$ be the sum of the sizes of all maximal matchings in G. Then, the above question can be rephrased by asking if $\mathcal{I}(G) = \frac{\mathcal{T}_1(G)}{\nu(G)\mathcal{T}_0(G)}$

is close to 1. Similarly, for a parametrized family of graphs $\{G_n\}_{n\geq 0}$, one can wonder how close $\lim_{n\to\infty} \mathcal{I}(G_n)$ is to 1. A graph is *equimatchable* if every maximal matching is maximum. Hence, $\mathcal{I}(G) = 1$ for all equimatchable graphs G.

The graph invariants $\mathcal{T}_0(G)$ and $\mathcal{I}(G)$ are the subject of several recent studies. For example, Došlić and Zubac [71] show how to enumerate maximal matchings in several classes of graphs, and $\mathcal{I}(G)$ could be interpreted as the expected efficiency of packings of dimers in G. Also, using generating functions, they compute asymptotical values of $\mathcal{I}(G_n)$ for parametrized families $\{G_n\}_{n\geq 0}$ of graphs related to linear polymers. Ash and Short [196] determine the number of maximal matchings (i.e., $\mathcal{T}_0(G)$) in three types of chemical compounds that are polyphenylene chains. Huntemann and Neil [128] give the value of $\mathcal{I}(G)$ for some types of grid graphs G.

A well known notion in mathematical chemistry is the Hosoya index of graph [124] which is defined as the total number of (not necessarily maximal) matchings in it. Adriantiana *et al.* [12] give some properties of the average size of such matchings in G which we denote by $\mathcal{I}^{ARW}(G)$. We show in Section 5.3 that considering only *maximal* matchings gives a more general approach in the sense that properties related to $\mathcal{I}^{ARW}(G)$ can be inferred from properties related to $\mathcal{I}(G)$.

If all maximal matchings in a graph have an equal chance to be chosen, $\mathcal{I}(G)$ is the expected ratio of the size of a uniformly chosen maximal matching to the size $\nu(G)$ of a maximum matching in G. However, maximal matchings are often obtained by the application of greedy algorithms and the chances of obtaining a given maximal matching by applying such a procedure are not equal. For example, the simplest heuristic works as follows.

Algorithm 1 Randomized Greedy (RG)

set $M \leftarrow \emptyset$;
while G contains at least one edge **do**
  Choose an edge uv in G, with a uniform distribution and add it to M;
  Remove vertices u and v and all their incident edges from G;
end while

Dyer and Frieze [76] have analyzed the expected performance of the above algorithm. To this aim, they introduce the ratio $\mathcal{I}^{DF}(G)$ of the expected size $\mu(G)$ of a randomized application of the above greedy algorithm on a graph G to the maximum size $\nu(G)$ of a matching in G. Note that each maximal matching M produced by the RG algorithm can be obtained in $|M|!$ different ways. For example, for the path P_4 on 4 vertices v_1, v_2, v_3, v_4 and edges

v_1v_2, v_2v_3, v_3v_4, the maximal matching $\{v_1v_2, v_3v_4\}$ can be obtained by choosing either v_1v_2 or v_3v_4 as first edge. We therefore consider every output of the RG algorithm as an *ordered* matching and we denote by $\boldsymbol{M}^o(G)$ the set of ordered maximal matchings in G. For illustration, $\boldsymbol{M}(\mathsf{P}_4) = \{\{v_1v_2, v_3v_4\}, \{v_2v_3\}\}$ while $\boldsymbol{M}^o(\mathsf{P}_4) = \{(v_1v_2, v_3v_4), (v_3v_4, v_1v_2), (v_2v_3)\}$. For an ordered maximal matching $M = (u_1v_1, \ldots, u_{|M|}v_{|M|}) \in \boldsymbol{M}^o(G)$, let m_i $(1 \leq i \leq |M|)$ be the number of edges in the graph obtained from G by removing vertices u_j, v_j $(j = 1, \ldots, i-1)$ and all their incident edges, and let $p(M) = \prod_{i=1}^{|M|} m_i$. Dyer and Frieze define

$$\mu(G) = \sum_{M \in \boldsymbol{M}^o(G)} \frac{|M|}{p(M)} \quad \text{and} \quad \mathcal{I}^{DF}(G) = \frac{\mu(G)}{\nu(G)}.$$

They show that there are graphs G for which $\mathcal{I}^{DF}(G)$ is close to $\frac{1}{2}$. They also prove that $\mathcal{I}^{DF}(G) \geq \frac{6}{11}$ for all planar graphs G, and $\mathcal{I}^{DF}(F) \geq \frac{16}{21}$ for all forests F. In subsequent papers [15, 16, 77] the study of the expected performance of the RG algorithm was continued and compared to a slightly modified version where a vertex v is first chosen at random, and then a random edge incident to v is added to the matching. This is known as the *modified randomized greedy* (MRG) algorithm. These studies have shown that MRG seems to have a better worst-case performance than RG. Other versions of the greedy algorithm are proposed in the literature, such as Tinhofer's MinGreedy algorithm [205]. It is identical to MRG except that vertex v is chosen randomly among vertices of minimum degree. An analysis of the expected performance of these greedy algorithms and of other variants can be found in [19, 95, 181].

Some authors have studied the difference $\nu(G) - \mu(G)$ instead of the ratio of these two invariants. For example, Magun [159] has analyzed large random graphs having up to 10 000 vertices, to see how many edges are lost, on average, when using a randomized greedy algorithm instead of an exact algorithm for the maximum matching problem. There exist also studies on randomized greedy matchings on particular classes of graphs such as cubic or bipartite random graphs and on weighted graphs (see e.g., [14, 88, 166] and references therein).

The aim of this chapter is to compare the average size $\frac{1}{|\boldsymbol{M}(G)|}\sum_{M \in \boldsymbol{M}(G)} |M| = \frac{\mathcal{T}_1(G)}{\mathcal{T}_0(G)}$ with $\nu(G)$, independently of the application of any greedy algorithm that produces these maximal matchings. We also provide tools that help to study the asymptotic value $\lim_{n\to\infty} \mathcal{I}(G_n)$ for families $\{G_n\}_{n\geq 0}$ of graphs.

In the next section, we fix some notations. Then, in Section 5.3, we investigate similarities and differences between several ratios involving matchings. For example, we highlight that $\mathcal{I}(G)$ and $\mathcal{I}^{DF}(G)$ do not share the same extremal properties. The main result of this chapter is Theorem 13 presented

in Section 5.4. It gives a simple procedure to compute the asymptotic value of $\mathcal{I}(G_n)$ for many families $\{G_n\}_{n\geq 0}$ of graphs. The use of this procedure is illustrated in Section 5.5 for families of graphs, including paths, cycles, wheels, chains of cycles, chains of cliques, ladders, and trees.

5.2 Notation

We remind the reader of a few standard definitions and notations, and introduce some that will be used throughout this chapter. Let $G = (V, E)$ be a simple undirected graph. We write $G \simeq H$ if G and H are two isomorphic graphs. We denote by K_n (resp. P_n, C_n and W_n) the *complete graph* (resp. the *path*, *cycle* and the *wheel*) of order n. We write $\mathsf{K}_{a,b}$ for the complete bipartite graph where a and b are the cardinalities of the two sets of vertices of the bipartition. For a graph G of order n, we define $\widetilde{G}$ as the graph of order $2n$ obtained by adding a new vertex v' for every v of G and linking v to v'. It is the *thorn* graph of G with all parameters equal to 1 [98]. It can also be defined as the *corona product* $G \circ K_1$ of G and K_1. For illustration, $\widetilde{\mathsf{P}}_4$ and $\widetilde{\mathsf{K}}_3$ are depicted in Figure 5.1. Note that for a graph G of order n, there is only one maximum matching of size n in $\widetilde{G}$ which consists in taking the edges vv' for all v in G.

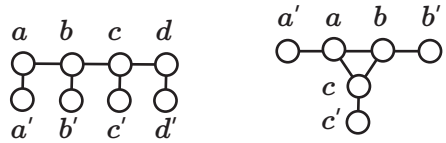


Figure 5.1: $\widetilde{\mathsf{P}}_4$ and $\widetilde{\mathsf{K}}_3$.

As mentioned in the previous section, $\boldsymbol{M}(G)$ is the set of maximal matchings in G, the maximum size of a matching in G is $\nu(G) = \max_{M\in\boldsymbol{M}(G)} |M|$, and the total number of maximal matchings in G is $\mathcal{T}_0(G) = |\boldsymbol{M}(G)|$. We can compute this last value in the following alternative way: if $S(G, k)$ is the number of maximal matchings of size k in G, then

$$\mathcal{T}_0(G) = \sum_{k=0}^{\infty} S(G, k) = \sum_{k=0}^{m} S(G, k),$$

where m is the number of edges in G. Also, the sum of the sizes of all maximal matchings in G is

$$\mathcal{T}_1(G) = \sum_{M\in\boldsymbol{M}(G)} |M| = \sum_{k=0}^{\infty} kS(G, k) = \sum_{k=0}^{m} kS(G, k).$$

We are interested in the following graph invariant $\mathcal{I}(G)$ which corresponds to the ratio of the average size of a maximal matching in a graph G to the size of a maximum matching in G:

$$\mathcal{I}(G) = \frac{\mathcal{T}_1(G)}{\nu(G)\mathcal{T}_0(G)}.$$

Observe that $\mathcal{I}(G) \in]\frac{1}{2}, 1]$ for all graphs G. Indeed, it is well known that $|M| \geq \frac{1}{2}\nu(G)$ for all maximal matchings $M \in \boldsymbol{M}(G)$, and since at least one of them is maximum we get $\mathcal{I}(G) > \frac{1}{2}$. It may happen that all maximal matchings are maximum, in which case $\mathcal{I}(G) = 1$. This is the case, for example, for cliques and complete bipartite graphs.

As already stressed in the introduction, other similar ratios appear in the literature. For example, if $\frac{1}{p(M)}$ is the probability that the RG algorithm outputs the ordered matching M from the set $\boldsymbol{M}^o(G)$ of ordered maximal matchings, then

$$\mathcal{I}^{DF}(G) = \frac{1}{\nu(G)} \sum_{M \in \boldsymbol{M}^o(G)} \frac{|M|}{p(M)}.$$

Also, Adriantiana *et al.* [12] have studied the average size of (not necessarily maximal) matchings. By denoting $S^{ARW}(G,k)$ the number of matchings of size k in G, we define $\mathcal{T}_0^{ARW}(G) = \sum_{k=1}^{m} S^{ARW}(G,k)$ and $\mathcal{T}_1^{ARW}(G) = \sum_{k=0}^{m} kS^{ARW}(G,k)$ which leads to

$$\mathcal{I}^{ARW}(G) = \frac{\mathcal{T}_1^{ARW}(G)}{\nu(G)\mathcal{T}_0^{ARW}(G)}.$$

Another possibility is to consider the average size of the ordered maximal matchings which lead to $\mathcal{T}_0^o(G) = |\boldsymbol{M}^o(G)|$, $\mathcal{T}_1^o(G) = \sum_{M \in \boldsymbol{M}^o(G)} |M|$ and

$$\mathcal{I}^o(G) = \frac{\mathcal{T}_1^o(G)}{\nu(G)\ \mathcal{T}_0^o(G)}.$$

By convention, $\mathcal{I}(G) = \mathcal{I}^{DF}(G) = \mathcal{I}^{ARW}(G) = \mathcal{I}^o(G) = 1$ for the empty graph G (i.e., when $\nu(G) = 0$). In the next section, we compare these four graph invariants and their extremal properties.

5.3 A comparison of similar graph invariants

There is a one to one correspondence between the matchings of size k in $G = (V, E)$ and the maximal matchings of size $|V| - k$ in $\widetilde{G}$. Indeed, consider a matching M in G and let W be the vertices of G that are not incident to

an edge in M. A maximal matching M' of $\widetilde{G}$ can be obtained by adding to M the edges vv' for the vertices $v \in W$. Since W contains $|V| - 2|M|$ vertices, we have $|M'| = |M| + |V| - 2|M| = |V| - |M|$. Conversely, if M' is a maximal matching in $\widetilde{G}$, let W be the set of vertices of G such that $vv' \in M'$. We can obtain a matching M of G by removing the edges vv' for all $v' \in W$. We thus get a matching of size $|M| = \frac{|V|-|W|}{2}$ while M' has $\frac{|V|-|W|}{2} + |W| = |V| - (\frac{|V|-|W|}{2}) = |V| - |M|$ edges.

The above relation between the matchings in G and the maximal matchings in $\widetilde{G}$ implies $S(\widetilde{G}, |V| - k) = S^{ARW}(G, k)$, $\mathcal{T}_0^{ARW}(G) = \mathcal{T}_0(\widetilde{G})$ and

$$\begin{aligned}\mathcal{T}_1^{ARW}(G) =& \sum_{k=0}^{\infty} kS^{ARW}(G,k) = \sum_{k=0}^{\infty} kS(\widetilde{G}, |V| - k) \\ =& \sum_{k=0}^{\infty} |V|S(\widetilde{G}, |V| - k) - \sum_{k=0}^{\infty} (|V| - k)S(\widetilde{G}, |V| - k) \\ =& |V|\mathcal{T}_0(\widetilde{G}) - \mathcal{T}_1(\widetilde{G}).\end{aligned}$$

Since $\nu(\widetilde{G}) = |V|$, this implies

$$\begin{aligned}&\mathcal{I}^{ARW}(G) = \frac{\mathcal{T}_1^{ARW}(G)}{\nu(G)\mathcal{T}_0^{ARW}(G)} = \frac{|V|\mathcal{T}_0(\widetilde{G}) - \mathcal{T}_1(\widetilde{G})}{\nu(G)\mathcal{T}_0(\widetilde{G})} = \frac{|V|(1 - \mathcal{I}(\widetilde{G}))}{\nu(G)} \\ \Leftrightarrow\quad &\mathcal{I}(\widetilde{G}) = 1 - \frac{\nu(G)\mathcal{I}^{ARW}(G)}{|V|}.\end{aligned}$$

For illustration, consider the clique K_2 on two vertices a, b. There are 2 matchings in K_2, namely the empty set of size 0, and $\{ab\}$ of size 1. Since $\nu(\mathsf{K}_2) = 1$ we get $\mathcal{I}^{ARW}(\mathsf{K}_2) = \frac{1}{2}$. Note that $\widetilde{\mathsf{K}}_2 \simeq \mathsf{P}_4$, which implies $\mathcal{I}(\mathsf{P}_4) = 1 - \frac{1}{2}\frac{1}{2} = \frac{3}{4}$. We could have computed $\mathcal{I}^{ARW}(\mathsf{K}_2)$ from $\mathcal{I}(\mathsf{P}_4)$ by observing that the two maximal matchings in P_4 are $\{aa', bb'\}$ and $\{ab\}$, which gives $\mathcal{I}(\mathsf{P}_4) = \frac{3}{4}$, and which implies $\mathcal{I}^{ARW}(\mathsf{K}_2) = 2(1 - \frac{3}{4}) = \frac{1}{2}$. Also, it is proven in [12] that $\lim_{n\to\infty} \nu(\mathsf{K}_n)\mathcal{I}^{ARW}(\mathsf{K}_n) = \frac{n}{2}$. The above link between $\mathcal{I}^{ARW}(G)$ and $\mathcal{I}(\widetilde{G})$ implies

$$\lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{K}}_n) \sim 1 - \frac{\frac{n}{2}}{n} = \frac{1}{2}.$$

Note that while $\mathcal{I}^{ARW}(G)$ can be derived from $\mathcal{I}(\widetilde{G})$, the converse is not always true since $\mathcal{I}(G)$ can be derived from $\mathcal{I}^{ARW}(H)$ only if there is such a graph H with $G \simeq \widetilde{H}$.

We are now comparing $\mathcal{I}(G)$, $\mathcal{I}^o(G)$ and $\mathcal{I}^{DF}(G)$ which are all a ratio between the average size of maximal matchings and the size of a maximum matching. For illustration, consider the graph $\widetilde{\mathsf{K}}_3$ of Figure 5.1:

- there are four maximal matchings, namely $\{aa', bb', cc'\}$ of size 3 and $\{aa', bc\}$, $\{bb', ac\}$, $\{cc', ab\}$ of size 2, and since $\nu(\widetilde{\mathsf{K}}_3) = 3$, we get $\mathcal{I}(\widetilde{\mathsf{K}}_3) = \frac{9}{12} = \frac{3}{4}$;

- the 12 ordered maximal matchings are

$$(aa', bb', cc'),\ (aa', cc', bb'),\ (bb', aa', cc'),\ (bb', cc', aa'),$$
$$(cc', aa', bb'),\ (cc', bb', aa'),\ (aa', bc),\ (bc, aa'),$$
$$(bb', ac),\ (ac, bb'),\ (cc', ab),\ (ab, cc').$$

  Hence, six of them are of size 3 and the six others are of size 2, which gives $\mathcal{I}^o(\widetilde{\mathsf{K}}_3) = \frac{30}{36} = \frac{5}{6}$;

- the probability that the output of the RG algorithm is (aa', bb', cc') is $\frac{1}{6}\frac{1}{3} = \frac{1}{18}$. It is the same for the five other ordered maximal matchings of size 3. The probability that RG produces (aa', bc) is also $\frac{1}{6}\frac{1}{3} = \frac{1}{18}$, and it is the same for (bb', ac) and (cc', ab). The probability that RG produces (bc, aa') is $\frac{1}{6}\frac{1}{1} = \frac{1}{6}$ and it is the same for (ac, bb') and (ab, cc'). Hence, $\mathcal{I}^{DF}(\widetilde{\mathsf{K}}_3) = \frac{1}{3}(6\frac{3}{18} + 3\frac{2}{18} + 3\frac{2}{6}) = \frac{7}{9}$.

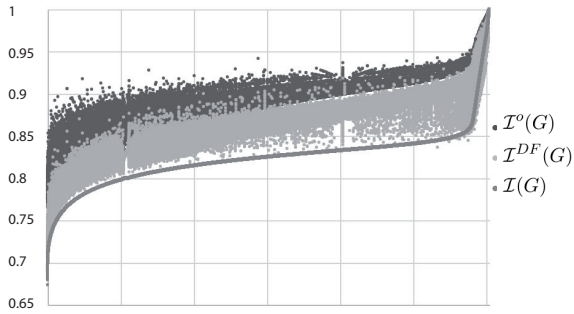


Figure 5.2: A comparison of $\mathcal{I}(G)$, $\mathcal{I}^o(G)$ and $\mathcal{I}^{DF}(G)$ on all graphs of order 10.

The above example illustrates well the fact that these three invariants differ from each other. To better observe these differences we have generated all non isomorphic graphs of order 10 and computed the values of the graph invariants using *PHOEG* [66]. There are 12 005 168 such graphs, and we have sorted them by non decreasing $\mathcal{I}(G)$ value. The values of $\mathcal{I}(G)$, $\mathcal{I}^o(G)$ and $\mathcal{I}^{DF}(G)$ are shown in Figure 5.2. We observe that $\mathcal{I}^o(G) \geq \mathcal{I}(G)$ for all G. This is not a surprise since there are exactly $k!S(G, k)$ ordered maximal matchings of size

k in G, which implies

$$\begin{aligned}\nu(G)(\mathcal{I}^o(G) - \mathcal{I}(G)) &= \frac{\sum_{k=0}^{\infty} kk!S(G,k)}{\sum_{k=0}^{\infty} k!S(G,k)} - \frac{\sum_{k=0}^{\infty} kS(G,k)}{\sum_{k=0}^{\infty} S(G,k)} \\ &= \frac{\sum_{k=0}^{\infty}\sum_{k'\geq k} S(G,k)S(G,k')\Big((kk!+k'k'!) - (kk'!-k'k!)\Big)}{\sum_{k=0}^{\infty} k!S(G,k)\sum_{k=0}^{\infty} S(G,k)} \\ &= \frac{\sum_{k=0}^{\infty}\sum_{k'\geq k} S(G,k)S(G,k')(k'-k)(k'!-k!)}{\sum_{k=0}^{\infty} k!S(G,k)\sum_{k=0}^{\infty} S(G,k)} \geq 0.\end{aligned}$$

When comparing $\mathcal{I}^{DF}(G)$ with the two other graph invariants, we observe that $\mathcal{I}^{DF}(G)$ is typically larger than $\mathcal{I}(G)$ and typically smaller than $\mathcal{I}^o(G)$. But there are some exceptions. We have found that $\mathcal{I}^{DF}(G) > \mathcal{I}^o(G)$ for 18 graphs of order 10 and $\mathcal{I}^{DF}(G) < \mathcal{I}(G)$ for 4 359 of them. In Figure 5.3, we illustrate examples of such exceptions for graphs of order 6. It is not difficult to check that $\mathcal{I}(G_1) = \frac{19}{20} > \frac{15}{16} = \mathcal{I}^{DF}(G_1)$ and $\mathcal{I}^o(G_2) = \frac{32}{39} < \frac{33}{40} = \mathcal{I}^{DF}(G_2)$. Note that the observed difference between $\mathcal{I}$ and $\mathcal{I}^{DF}$ dates back to the 1930s, as illustrated by the works of Flory [83] and Jackson and Montroll [130].

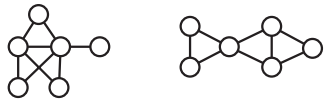


Figure 5.3: Two graphs G_1 and G_2 with $\mathcal{I}(G_1) > \mathcal{I}^{DF}(G_1)$ and $\mathcal{I}^o(G_2) < \mathcal{I}^{DF}(G_2)$.

Andriantiana *et al.*[12] have proven that the trees T of order n that maximize the average size $\mathcal{I}^{ARW}(T)\nu(T)$ of a (not necessarily maximal) matching are the paths P_n. It follows from the above equations that $\mathcal{I}(\widetilde{\mathsf{P}}_n) \leq \mathcal{I}(\widetilde{T})$ for all trees T of order n. Note that $\widetilde{\mathsf{P}}_n$ is a tree on $2n$ vertices. Dyer and Frieze [76] conjecture that $\mathcal{I}^{DF}(\widetilde{\mathsf{P}}_n) \leq \mathcal{I}^{DF}(T)$ for all trees T of order $2n$. We show in Section 5.5.5 that there are trees T of order $2n$ such that $\mathcal{I}(\widetilde{\mathsf{P}}_n) > \mathcal{I}(T)$. Hence, $\mathcal{I}$ and $\mathcal{I}^{DF}$ do not seem to share the same extremal properties.

5.4 Main result

The purpose of this section is to provide a general technique for determining $\lim_{n\to\infty} \mathcal{I}(G_n)$ for classes $\{G_n\}_{n\geq 0}$ of graphs where the number of maximal matchings in G_n linearly depends on the number of maximal matchings in graphs $G_{n'}$ with $n' < n$. For this purpose, Let us first prove a result which comes from the theory which relates sequences satisfying linear equations to bivariate generating functions (see for example Theorems IV.9 and IV.10 in [82]).

Theorem 13. *Let $(f_{n,k})$ be a sequence of numbers depending on two positive integer-valued indices, n and k, and such that*

$$f_{n,k} = \sum_{i=1}^{I} \sum_{j=1}^{n_i} a_{ij} f_{n-i,k-j}$$

for a strictly positive integer I, non-negative integers n_i $(i = 1, \ldots, I)$ and real numbers a_{ij} $(i = 1, \ldots, I; j = 1, \ldots, n_i)$. Let

$$F(x,y) = \frac{P(x,y)}{Q(x,y)} = \sum_{n\geq 0} \sum_{k\geq 0} f_{n,k} x^n y^k$$

be the associated bivariate generating function with $Q(x,y) = 1 - \sum_{i=1}^{I} \sum_{j=1}^{n_i} a_{ij} x^i y^j$.

If $Q(x,1)$ has a unique root ρ of smallest modulus with multiplicity 1, and if $P(\rho,1) \neq 0$, then

$$\lim_{n\to\infty} \frac{\sum_{k\geq 0} k f_{n,k}}{n \sum_{k\geq 0} f_{n,k}} = \frac{\sum_{i=1}^{I} \sum_{j=1}^{n_i} j a_{ij} \rho^i}{\sum_{i=1}^{I} \sum_{j=1}^{n_i} i a_{ij} \rho^i}.$$

Proof. Note first that the denominator of $F(x,1)$ can be factored as $Q(x,1) = (1 - x/\rho)q(x)$, where q only has roots whose modulus is greater than that of ρ. Dividing by $x - \rho$ and taking the limit $x \to \rho$, we obtain $\frac{\partial Q}{\partial x}(\rho,1) = -q(\rho)/\rho$ or equivalently $q(\rho) = \rho \frac{\partial Q}{\partial x}(\rho,1)$.

Since we assume $P(\rho,1) \neq 0$, $F(x,1) = \frac{P(x,1)}{Q(x,1)}$ has a simple pole at ρ and no other pole whose modulus is less than or equal to that of ρ. At the pole, we have

$$F(x,1) = \frac{P(x,1)}{(1-x/\rho)q(x)} \sim \frac{P(\rho,1)}{q(\rho)} (1 - x/\rho)^{-1}.$$

By partial faction expansion, one can express $F(x,1)$ as

$$F(x,1) = \frac{P(\rho,1)}{q(\rho)}(1-x/\rho)^{-1} + \tilde{F}(x),$$

where $\tilde{F}$ is holomorphic in a disc around 0 whose radius is greater than $|\rho|$. Applying singularity analysis in the rational/meromorphic case, it follows that

$$[x^n]F(x,1) = \sum_{k\geq 0} f_{n,k} \sim \frac{P(\rho,1)}{q(\rho)}\rho^{-n} = -\frac{P(\rho,1)}{\rho\frac{\partial Q}{\partial x}(\rho,1)}\rho^{-n}.$$

Likewise,

$$\frac{\partial F}{\partial y}(x,1) = \frac{Q(x,1)\frac{\partial P}{\partial y}(x,1) - P(x,1)\frac{\partial Q}{\partial y}(x,1)}{Q(x,1)^2}$$

has a pole (of multiplicity 2) at ρ, but no other poles in a disc around 0 whose radius is greater than ρ. At this pole, we have

$$Q(x,1)\frac{\partial P}{\partial y}(x,1) - P(x,1)\frac{\partial Q}{\partial y}(x,1) = -P(\rho,1)\frac{\partial Q}{\partial y}(\rho,1) + O(x-\rho)$$

and

$$Q(x,1)^2 = -(q(\rho)^2 + O(x-\rho))(1-x/\rho)^2,$$

thus

$$\frac{\partial F}{\partial y}(x,1) = -\frac{P(\rho,1)\frac{\partial Q}{\partial y}(\rho,1)}{q(\rho)^2}(1-x/\rho)^2 + O((x-\rho)^{-1}).$$

Applying singularity analysis now yields

$$\begin{aligned}
[x^n]\frac{\partial F}{\partial y}(x,1) &= \sum_{k\geq 0} k f_{n,k} \\
&= \left(\frac{-P(\rho,1)\,\frac{\partial Q}{\partial y}(\rho,1)}{q(\rho)^2}\,n + O(1)\right)\rho^{-n} \\
&= \left(\frac{-P(\rho,1)\,\frac{\partial Q}{\partial y}(\rho,1)}{\rho^2\left(\frac{\partial Q}{\partial x}(\rho,1)\right)^2}\,n + O(1)\right)\rho^{-n}.
\end{aligned}$$

Taking the quotient of the two asymptotic formulas, we obtain

$$\lim_{n\to\infty} \frac{\sum\limits_{k\geq 0} k f_{n,k}}{n\sum\limits_{k\geq 0} f_{n,k}} = \frac{\frac{\partial Q}{\partial y}(\rho,1)}{\rho\frac{\partial Q}{\partial x}(\rho,1)} = \frac{\sum\limits_{i=1}^{I}\sum\limits_{j=1}^{n_i} j a_{ij}\rho^i}{\sum\limits_{i=1}^{I}\sum\limits_{j=1}^{n_i} i a_{ij}\rho^i}.$$

■

Corollary 14. *Let $\{G_n\}_{n\geq 0}$ be a family of graphs such that*

$$S(G_n,k)=\sum_{i=1}^{I}\sum_{j=1}^{n_i}a_{ij}S(G_{n-i},k-j)$$

for a strictly positive integer I, non-negative integers n_i $(i=1,\dots,I)$ and real numbers a_{ij} $(i=1,\dots,I; j=1,\dots,n_i)$. Let $\alpha_i=\sum_{j=1}^{n_i}a_{ij}$ and $\beta_i=\sum_{j=1}^{n_i}ja_{ij}$, and let $c=\lim_{n\to\infty}\frac{\nu(G_n)}{n}$. If the equation $x^I-\sum_{i=1}^{I}\alpha_i x^{I-i}=0$ has a unique root r of maximum modulus with multiplicity 1 and if

$$\sum_{i=0}^{I-1}\frac{1}{r^i}\Big(T_0(G_i)-\sum_{j=1}^{i}\alpha_j T_0(G_{i-j})\Big)\neq 0$$

then

$$\lim_{n\to\infty}\mathcal{I}(G_n)=\frac{\sum_{i=1}^{I}\beta_i r^{I-i}}{c\sum_{i=1}^{I}i\alpha_i r^{I-i}}.$$

Proof. Let $f_{n,k}=S(G_n,k)$ and let

$$F(x,y)=\frac{P(x,y)}{Q(x,y)}=\sum_{n\geq 0}\sum_{k\geq 0}f_{n,k}x^n y^k$$

be the associated bivariate generating function with

$$Q(x,y)=1-\sum_{i=1}^{I}\sum_{j=1}^{n_i}a_{ij}x^i y^j.$$

Note that

$$Q(x,1)=1-\sum_{i=1}^{I}\sum_{j=1}^{n_i}a_{ij}x^i=1-\sum_{i=1}^{I}\alpha_i x^i.$$

Hence, assuming that $x^I-\sum_{i=1}^{I}\alpha_i x^{I-i}=0$ has a unique root r of maximum modulus and of multiplicity 1 is equivalent to assuming that $Q(x,1)$ has a unique root $\frac{1}{r}$ of smallest modulus with multiplicity 1. It is not difficult to check that

$$P(x,1)=\sum_{i=0}^{I-1}x^i\Big(T_0(G_i)-\sum_{j=1}^{i}\alpha_j T_0(G_{i-j})\Big).$$

Since we assume $P(\frac{1}{r},1)\neq 0$, we know from Theorem 13 that

$$\lim_{n\to\infty}\mathcal{I}(G_n)=\lim_{n\to\infty}\frac{\sum\limits_{k\geq 0}kf_{n,k}}{cn\sum\limits_{k\geq 0}f_{n,k}}=\frac{\sum\limits_{i=1}^{I}\sum\limits_{j=1}^{n_i}ja_{ij}r^{-i}}{c\sum\limits_{i=1}^{I}\sum\limits_{j=1}^{n_i}ia_{ij}r^{-i}}=\frac{\sum\limits_{i=1}^{I}\beta_i r^{I-i}}{c\sum\limits_{i=1}^{I}i\alpha_i r^{I-i}}.$$

■

5.5 Applications on some families of graphs

In this section we illustrate how Corollary 14 allows to determine the asymptotic behavior of the invariant $\mathcal{I}$ for various families of graphs. These families were chosen so as to cover a wide range of values in the interval $]\frac{1}{2}, 1]$.

5.5.1 Paths, cycles and wheels

The plastic number r is the root of maximum modulus and the unique real solution of the cubic equation $x^3 = x + 1$ [207]. Its value is $r = \sqrt[3]{\frac{9+\sqrt{69}}{18}} + \sqrt[3]{\frac{9-\sqrt{69}}{18}}$.

Theorem 15. $\lim_{n\to\infty} \mathcal{I}(\mathsf{P}_n) = \frac{2r+2}{2r+3}$ *where* r *is the plastic number.*

Proof. Let P_n be a path on n vertices $v_1, \ldots, v_n$ with edges v_iv_{i+1} $(i = 1, \ldots, n-1)$, and let M be a maximal matching of size k in P_n. If $v_1v_2 \in M$, then $M \setminus \{v_1v_2\}$ is a maximal matching in P_{n-2} obtained from P_n by removing v_1 and v_2. If $v_1v_2 \notin M$, then $v_2v_3 \in M$, which means that $M \setminus \{v_2v_3\}$ is a maximal matching in P_{n-3} obtained from P_n by removing v_1, v_2 and v_3. We therefore have

$$S(\mathsf{P}_n, k) = S(\mathsf{P}_{n-2}, k-1) + S(\mathsf{P}_{n-3}, k-1).$$

With the notations of Corollary 14, we have $I = 3, \alpha_1 = \beta_1 = 0$ and $\alpha_2 = \alpha_3 = \beta_2 = \beta_3 = 1$. The root of maximum modulus of the equation $x^3 - x - 1 = 0$ is the plastic number r. Moreover, since $T_0(\mathsf{P}_0) = T_0(\mathsf{P}_1) = T_0(\mathsf{P}_2) = 1$, we have

$$\sum_{i=0}^{I-1} \frac{1}{r^i}\Big(T_0(\mathsf{P}_i) - \sum_{j=1}^{i} \alpha_j T_0(\mathsf{P}_{i-j})\Big) = 1 + \frac{1}{r} \neq 0.$$

Hence the hypotheses of Corollary 14 are satisfied, and since $\nu(\mathsf{P}_n) = \lfloor \frac{n}{2} \rfloor$, we have $\lim_{n\to\infty} \frac{\nu(\mathsf{P}_n)}{n} = \frac{1}{2}$ which implies

$$\lim_{n\to\infty} \mathcal{I}(\mathsf{P}_n) = \frac{r+1}{\frac{1}{2}(2r+3)} = \frac{2r+2}{2r+3} \approx 0.8299.$$ ■

Corollary 16. $\lim_{n\to\infty} \mathcal{I}(\mathsf{C}_n) = \lim_{n\to\infty} \mathcal{I}(\mathsf{P}_n)$.

Proof. Let C_n be a cycle on n vertices $v_1, \ldots, v_n$ with edges v_iv_{i+1} $(i = 1, \ldots, n-1)$ and v_1v_n, and let M be a maximal matching of size k in C_n. If $v_1v_2 \in M$, then $M \setminus \{v_1v_2\}$ is a maximal matching of P_{n-2} obtained from C_n by removing v_1 and v_2. If $v_1v_2 \notin M$, then M contain at least one of the edges v_2v_3 and v_1v_n, and if M contains both of them, then $M \setminus \{v_2v_3, v_1v_n\}$ is a maximal matching in P_{n-4} obtained from C_n by removing v_1, v_2, v_3 and v_4. Hence,

$$\begin{aligned} S(\mathsf{C}_n, k) &= S(\mathsf{P}_{n-2}, k-1) + \Big(2S(\mathsf{P}_{n-2}, k-1) - S(\mathsf{P}_{n-4}, k-2)\Big) \\ &= 3S(\mathsf{P}_{n-2}, k-1) - S(\mathsf{P}_{n-4}, k-2). \end{aligned}$$

We know from Theorem 15 that $S(\mathsf{P}_n, k) = S(\mathsf{P}_{n-2}, k-1) + S(\mathsf{P}_{n-3}, k-1)$, which implies

$$\begin{aligned} S(\mathsf{C}_n, k) &= 3\Big(S(\mathsf{P}_{n-4}, k-2) + S(\mathsf{P}_{n-5}, k-2)\Big) \\ &\quad - \Big(S(\mathsf{P}_{n-6}, k-3) + S(\mathsf{P}_{n-7}, k-3)\Big) \\ &= S(\mathsf{C}_{n-2}, k-1) + S(\mathsf{C}_{n-3}, k-1). \end{aligned}$$

This is the same recurrence relation as in Theorem 15. Hence, $T_0(\mathsf{C}_n) = T_0(\mathsf{C}_{n-2}) + T_0(\mathsf{C}_{n-3})$. We can fix $T_0(\mathsf{C}_0) = 3$, $T_0(\mathsf{C}_1) = 0$ and $T_0(\mathsf{C}_2) = 2$ (even if $\mathsf{C}_0, \mathsf{C}_1$ and C_2 do not exist), so that the values of $T_0(\mathsf{C}_n)$) are the correct ones for $n \geq 3$. Hence,

$$\sum_{i=0}^{I-1} \frac{1}{r^i}\Big(T_0(\mathsf{C}_i) - \sum_{j=1}^{i} \alpha_j T_0(\mathsf{C}_{i-j})\Big) = 3 - \frac{1}{r^2} \neq 0,$$

where r is the plastic number. Since $\nu(\mathsf{C}_n)=\nu(\mathsf{P}_n)$, we conclude that

$$\lim_{n\to\infty} \mathcal{I}(\mathsf{C}_n) = \lim_{n\to\infty} \mathcal{I}(\mathsf{P}_n).$$

■

Corollary 17. $\lim_{n\to\infty} \mathcal{I}(\mathsf{W}_n) = \lim_{n\to\infty} \mathcal{I}(\mathsf{P}_n)$.

Proof. Let W_n be a wheel on n vertices $v_1, \ldots, v_n$ where v_1 is the center of the wheel. If n is even, then all maximal matchings in W_n contain an edge incident to v_1, while if n is odd, then there are exactly two maximal matchings with no edge incident to v_1, and both of them contain exactly $\frac{n-1}{2}$ edges. Since the graph obtained from W_n by removing the endpoints of an edge incident to v_1 is a path on $n-2$ vertices, we have

$$S(\mathsf{W}_n, k) = (n-1)S(\mathsf{P}_{n-2}, k-1) + \begin{cases} 2 & \text{if } n \text{ is odd and } k = \frac{n-1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$\mathcal{T}_0(\mathsf{W}_n) = (n-1)\mathcal{T}_0(\mathsf{P}_{n-2}) + \begin{cases} 2 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

and

$$\mathcal{T}_1(\mathsf{W}_n) = (n-1)(\mathcal{T}_1(\mathsf{P}_{n-2}) + \mathcal{T}_0(\mathsf{P}_{n-2})) + \begin{cases} n-1 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

Hence,

$$\frac{\mathcal{T}_1(\mathsf{W}_n)}{\mathcal{T}_0(\mathsf{W}_n)} = \frac{\mathcal{T}_1(\mathsf{P}_{n-2}) + \mathcal{T}_0(\mathsf{P}_{n-2}) + 1(n \bmod 2)}{\mathcal{T}_0(\mathsf{P}_{n-2}) + \frac{2}{n-1}(n \bmod 2)}.$$

Since $\nu(\mathsf{P}_{n-2}) = \nu(\mathsf{W}_n) - 1$, we have

$$\begin{aligned} \lim_{n\to\infty} \mathcal{I}(\mathsf{W}_n) &= \lim_{n\to\infty} \frac{\mathcal{T}_1(\mathsf{P}_{n-2}) + \mathcal{T}_0(\mathsf{P}_{n-2})}{\nu(\mathsf{W}_n)\mathcal{T}_0(\mathsf{P}_{n-2})} \\ &= \lim_{n\to\infty} \Big(\frac{(\nu(\mathsf{W}_n)-1)\mathcal{I}(\mathsf{P}_{n-2})}{\nu(\mathsf{W}_n)} + \frac{1}{\nu(\mathsf{W}_n)} \Big) \\ &= \lim_{n\to\infty} \Big(\mathcal{I}(\mathsf{P}_{n-2}) + \frac{1-\mathcal{I}(\mathsf{P}_{n-2})}{\nu(\mathsf{W}_n)} \Big) \\ &= \lim_{n\to\infty} \mathcal{I}(\mathsf{P}_n). \end{aligned}$$

■

5.5.2 Chains of cycles

The types of graphs analyzed in the following subsections are numerous, and to avoid giving a specific name to each of them, we will use the notation G_n^i for the ith type of graph, possibly with additional information next to the exponent i. Also, while we will always give the root r of maximum modulus mentioned in Corollary 14, we leave to the reader the task to check that $\sum_{i=0}^{I-1} \frac{1}{r^i} \Big(T_0(G_i) - \sum_{j=1}^{i} \alpha_j T_0(G_{i-j}) \Big) \neq 0$.

We first consider chains of hexagons studied in [196], where the cut-vertices of the hexagons are at distance 3, 2 or 1. More precisely, let $G_n^{1,s}$ be the graph obtained by considering the disjoint union of n hexagons $H_1, \ldots, H_n$, where $\{v_i^1, \ldots, v_i^6\}$ and $\{v_i^1v_i^2, \ldots, v_i^5v_i^6, v_i^1v_i^6\}$ are the vertex set and the edge set of H_i, and by adding an edge between v_i^{s+1} and v_{i+1}^1 for $i = 1, \ldots, n-1$. For illustration $G_3^{1,1}$, $G_3^{1,2}$ and $G_3^{1,3}$ are depicted in Figure 5.4. They are known as the ortho-, the meta-, and the para- phenylene chains, respectively [196]. Maximal matchings in analogous types of spiro-chains were considered in [70].

Approximate values for $\lim_{n\to\infty} \mathcal{I}(G_n^{1,s})$ are given in [196]. We show how to compute these values using the technique of Theorem 13. Since the authors

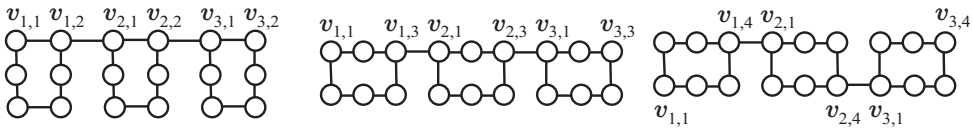


Figure 5.4: $G_3^{1,1}$, $G_3^{1,2}$ and $G_3^{1,3}$.

made calculation errors for $s = 2$, we give a detailed proof for this case, while fewer details are given for $s = 1$ and $s = 3$. So let's start with $s = 2$ and let $G_n^{1,2,1}$, $G_n^{1,2,2}$ and $G_n^{1,2,3}$ be defined as follows:

- $G_n^{1,2,1}$ is the graph obtained from $G_n^{1,2}$ by adding a vertex w_1 linked to v_1^1,
- $G_n^{1,2,2}$ is the graph obtained from $G_n^{1,2,1}$ by adding a vertex w_2 linked to w_1,
- $G_n^{1,2,3}$ is the graph obtained from $G_n^{1,2,2}$ by adding a vertex w_3 linked to w_1,

For illustration, $G_2^{1,2,1}$, $G_2^{1,2,2}$ and $G_2^{1,2,3}$ are depicted in Figure 5.5.

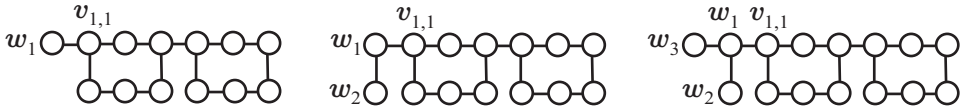


Figure 5.5: $G_n^{1,2,1}$, $G_n^{1,2,2}$ and $G_n^{1,2,3}$.

Theorem 18. $\lim_{n\to\infty} \mathcal{I}(G_n^{1,2}) = \frac{16r^2-3r+12}{3(7r^2-4r+6)}$ *where* $r = \frac{7+\sqrt[3]{307+9\sqrt{182}}+\sqrt[3]{307-9\sqrt{182}}}{3}$.

Proof. To simplify the notation, let's write $S(n,k)$ and $S_j(n,k)$ instead of $S(G_n^{1,2},k)$ and $S(G_n^{1,2,j},k)$. As mentioned in [196],

$$S(n,k) = 2S(n-1,k-2) + S_1(n-1,k-2) + 2S_2(n-1,k-2) \tag{5.1}$$

$$\begin{aligned} S_1(n,k) = {} & S(n-1,k-2) + S_1(n-1,k-2) \\ & + 3S_2(n-1,k-2) + S_3(n-1,k-2) \end{aligned} \tag{5.2}$$

$$S_2(n,k) = S(n,k-1) + S_2(n-1,k-2) + S_3(n-1,k-2) \tag{5.3}$$

$$S_3(n,k) = 2S(n,k-1) + S_2(n-1,k-2) + S_3(n-1,k-2). \tag{5.4}$$

Equations (5.3) and (5.4) show that $S_3(n,k) = S_2(n,k) + S(n,k-1)$ which implies that (5.4) can be rewritten as

$$\begin{aligned} S_2(n,k) + S(n,k-1) &= 2S(n,k-1) + S_2(n-1,k-2) \\ &\quad + (S_2(n-1,k-2) + S(n-1,k-3)) \\ \Leftrightarrow \qquad S_2(n,k) &= S(n,k-1) + S(n-1,k-3) + 2S_2(n-1,k-2). \end{aligned} \tag{5.4'}$$

Hence, Equation (5.2) can be rewritten as

$$\begin{aligned} S_1(n,k) &= S(n-1,k-2) + \Big(S(n,k) - 2S(n-1,k-2) - 2S_2(n-1,k-2)\Big) \\ &\quad + 3S_2(n-1,k-2) + \Big(S_2(n-1,k-2) + S(n-1,k-3)\Big) \\ &= S(n,k) - S(n-1,k-2) + S(n-1,k-3) + 2S_2(n-1,k-2) \\ &= S(n,k) - S(n-1,k-2) + S(n-1,k-3) \\ &\quad + (S_2(n,k) - S(n,k-1) - S(n-1,k-3)) \\ &= S(n,k) - S(n,k-1) - S(n-1,k-2) + S_2(n,k) \end{aligned} \tag{5.2'}$$

which implies that Equation (5.1) can be rewritten as

$$\begin{aligned} S(n,k) &= 2S(n-1,k-2) + 2S_2(n-1,k-2) \\ &\quad + \Big(S(n-1,k-2) - S(n-1,k-3) - S(n-2,k-4) + S_2(n-1,k-2)\Big) \\ &= 3S(n-1,k-2) - S(n-1,k-3) - S(n-2,k-4) + 3S_2(n-1,k-2) \end{aligned}$$

$$\Leftrightarrow 3S_2(n-1,k-2) = S(n,k) - 3S(n-1,k-2) + S(n-1,k-3) + S(n-2,k-4). \tag{5.1'}$$

Using Equation (5.1'), we can rewrite Equation (5.4') as follows:

$$\begin{aligned} &S(n+1,k+2) - 3S(n,k) + S(n,k-1) + S(n-1,k-2) \\ &\quad = 3S(n,k-1) + 3S(n-1,k-3) \\ &\qquad + \Big(2S(n,k) - 6S(n-1,k-2) + 2S(n-1,k-3) + 2S(n-2,k-4)\Big). \end{aligned}$$

We can then rewrite it further as follows:

$$\begin{aligned} S(n+1,k+2) &= 5S(n,k) + 2S(n,k-1) - 7S(n-1,k-2) \\ &\quad + 5S(n-1,k-3) + 2S(n-2,k-4) \end{aligned}$$

$$\Leftrightarrow \quad S(n,k) = 5S(n-1,k-2) + 2S(n-1,k-3) - 7S(n-2,k-4) + 5S(n-2,k-5) + 2S(n-3,k-6).$$

We have $I = 3$, $\alpha_1 = 7$, $\alpha_2 = -2$, $\alpha_3 = 2$, $\beta_1 = 16$, $\beta_2 = -3$ and $\beta_3 = 12$, and the root of maximum modulus of the equation $x^3 - 7x^2 + 2x - 2 = 0$ is

$$r = \frac{7 + \sqrt[3]{307 + 9\sqrt{182}} + \sqrt[3]{307 - 9\sqrt{182}}}{3}.$$

Since $\frac{\nu(G_n^{1,2})}{n} = 3$, it follows that

$$\lim_{n\to\infty} \mathcal{I}(G_n^{1,2}) = \frac{16r^2 - 3r + 12}{3(7r^2 - 4r + 6))} \approx 0.8064.$$ ■

Theorem 19.
$\lim_{n\to\infty} \mathcal{I}(G_n^{1,1}) = \frac{17r^2-13r+8}{3(7r^2-6r+3)}$ *where* $r = \frac{7+\sqrt[3]{262+6\sqrt{129}}+\sqrt[3]{262-6\sqrt{129}}}{3}$.
$\lim_{n\to\infty} \mathcal{I}(G_n^{1,3}) = \frac{13r^2+36r+28}{3(5r^2+16r+12)}$ *where* $r = \frac{5+\sqrt[3]{359+12\sqrt{78}}+\sqrt[3]{359-12\sqrt{78}}}{3}$.

Proof. The bivariate generating functions in [196] show that

$$\begin{aligned} S(G_n^{1,1},k) = {} & 4S(G_{n-1}^{1,1},k-2) + 3S(G_{n-1}^{1,1},k-3) - 4S(G_{n-2}^{1,1},k-4) \\ & + 3S(G_{n-2}^{1,1},k-5) - 2S(G_{n-2}^{1,1},k-6) + S(G_{n-3}^{1,1},k-6) \\ & - 2S(G_{n-3}^{1,1},k-7) + 2S(G_{n-3}^{1,1},k-8). \end{aligned}$$

We have $I = 3$, $\alpha_1 = 7$, $\alpha_2 = -3$, $\alpha_3 = 1$, $\beta_1 = 17$, $\beta_2 = -13$ and $\beta_3 = 8$, and the root of maximum modulus of the equation $x^3 - 7x^2 + 3x - 1 = 0$ is

$$r = \frac{7 + \sqrt[3]{262 + 6\sqrt{129}} + \sqrt[3]{262 - 6\sqrt{129}}}{3}.$$

Since $\nu(G_n^{1,1}) = 3n$, it follows from Theorem 13 that

$$\lim_{n\to\infty} \mathcal{I}(G_n^{1,1}) = \frac{17r^2 - 13r + 8}{3(7r^2 - 6r + 3))} \approx 0.8234.$$

Similarly, the bivariate generating functions in [196] show that

$$\begin{aligned} S(G_n^{1,3},k) = {} & 2S(G_{n-1}^{1,3},k-2) + 3S(G_{n-1}^{1,3},k-3) + 2S(G_{n-2}^{1,3},k-4) \\ & + 8S(G_{n-2}^{1,3},k-5) - 2S(G_{n-2}^{1,3},k-6) + 4S(G_{n-3}^{1,3},k-7). \end{aligned}$$

In this case, $I = 3$, $\alpha_1 = 5$, $\alpha_2 = 8$, $\alpha_3 = 4$, $\beta_1 = 13$, $\beta_2 = 36$ and $\beta_3 = 28$, and the root of maximum modulus of the equation $x^3 - 5x^2 - 8x - 4 = 0$ is

$$r = \frac{5 + \sqrt[3]{359 + 12\sqrt{78}} + \sqrt[3]{359 - 12\sqrt{78}}}{3}.$$

Since $\nu(G_n^{1,3}) = 3n$, we have

$$\lim_{n\to\infty} \mathcal{I}(G_n^{1,3}) = \frac{13r^2 + 36r + 28}{3(5r^2 + 16r + 12))} \approx 0.8257. \qquad \blacksquare$$

Instead of chains of hexagons, we could consider chains of cycles of order $r \neq 6$. For example, for $s \in \{1, 2\}$, let $G_n^{2,s}$ be the graph obtained by considering the disjoint union of n cycles on four vertices, where $\{v_{i,1}, \ldots, v_{i,4}\}$ and $\{v_{i,1}v_{i,2}, v_{i,2}v_{i,3}, v_{i,3}v_{i,4}, v_{i,1}v_{i,4}\}$ are the vertex set and the edge set of the ith cycle, and by adding an edge between $v_{i,s+1}$ and $v_{i+1,1}$ for $i = 1, \ldots, n-1$. Also, let $G_n^{2,s,1}$ be the graph obtained from $G_n^{2,s}$ by adding a vertex w_1 linked to $v_{1,1}$, and let $G_n^{2,s,2}$ be the graph obtained from $G_n^{2,s,1}$ by adding a vertex w_2 linked to w_1.

Theorem 20.
$\lim_{n\to\infty} \mathcal{I}(G_n^{2,1}) = \frac{6r^2+r+10}{2(3r^2+2r+6)}$ *where* $r = 1 + \sqrt[3]{\frac{45+\sqrt{1257}}{18}} + \sqrt[3]{\frac{45-\sqrt{1257}}{18}}$.
$\lim_{n\to\infty} \mathcal{I}(G_n^{2,2}) = \frac{5r+6}{2(3r+4)}$ *where* $r = \frac{3+\sqrt{17}}{2}$.

Proof. Let's start with $G_n^{2,1}$. Since exactly one of the edges $v_{1,1}v_{1,4}$ and $v_{1,3}v_{1,4}$ belongs to a maximal matching in $G_n^{2,1}$, and since the graph obtained from $G_n^{2,1}$ by removing $v_{1,4}$ and one of $v_{1,1}, v_{1,3}$ is $G_{n-1}^{2,1,2}$, we have

$$S(G_n^{2,1}, k) = 2S(G_{n-1}^{2,1,2}, k-1). \tag{5.5}$$

Let now M be a maximal matching of size k in $G_n^{2,1,2}$. If $w_1w_2 \in M$, then the other edges of M form a maximal matching in $G_n^{2,1}$. If $w_1w_2 \notin M$, then $w_1v_{1,1} \in M$ and there are two possible cases: if $v_{1,2}v_{1,3} \in M$, the other edges of M form a maximal matching in $G_{n-1}^{2,1}$; if $v_{1,2}v_{1,3} \notin M$, then $v_{1,3}v_{1,4} \in M$ and the other edges of M form a maximal matching in $G_{n-1}^{2,1,1}$. Hence,

$$S(G_n^{2,1,2}, k) = S(G_n^{2,1}, k-1) + S(G_{n-1}^{2,1}, k-2) + S(G_{n-1}^{2,1,1}, k-2). \tag{5.6}$$

Finally, let M be a maximal matching of size k in $G_n^{2,1,1}$. If $w_1v_{1,1} \in M$, then as in the previous case, M contains $k-2$ edges in $G_{n-1}^{2,1,1}$ or in $G_{n-1}^{2,1}$. If $w_1v_{1,1} \notin M$, there are two possible cases: if $v_{1,1}v_{1,2} \in M$, then $v_{1,3}v_{1,4} \in M$ and

the $k-2$ other edges of M form a maximal matching in $G_{n-1}^{2,1}$; if $v_{1,1}v_{1,2} \notin M$, then $v_{1,1}v_{1,4} \in M$ and the $k-1$ other edges of M form a maximal matching in $G_{n-1}^{2,1,2}$. Hence,

$$S(G_n^{2,1,1},k) = 2S(G_{n-1}^{2,1},k-2) + S(G_{n-1}^{2,1,1},k-2) + S(G_{n-1}^{2,1,2},k-1). \quad (5.7)$$

Playing with Equations (5.5)-(5.7) as we did in Theorem 18 leads to

$$S(G_n^{2,1},k) = 3S(G_{n-1}^{2,1},k-2)+3S(G_{n-2}^{2,1},k-3)-2S(G_{n-2}^{2,1},k-4)+2S(G_{n-3}^{2,1},k-5).$$

We thus have $I=3$, $\alpha_1=3$, $\alpha_2=1$, $\alpha_3=2$, $\beta_1=6$, $\beta_2=1$ and $\beta_3=10$, and the root of maximum modulus of the equation $x^3-3x^2-x-2=0$ is

$$r = 1 + \sqrt[3]{\frac{45+\sqrt{1257}}{18}} + \sqrt[3]{\frac{45-\sqrt{1257}}{18}}.$$

Since $\frac{\nu(G_n^{2,1})}{n} = 2$, it follows that

$$\lim_{n\to\infty} \mathcal{I}(G_n^{2,1}) = \frac{6r^2+r+10}{2(3r^2+2r+6))} \approx 0.8732.$$

We now consider $G_n^{2,2}$. A similar analysis as above leads to the following recurrence relation:

$$S(G_n^{2,2},k) = S(G_{n-1}^{2,2},k-1) + 2S(G_{n-1}^{2,2},k-2) + 2S(G_{n-2}^{2,2},k-3).$$

This time, we have $I=2$, $\alpha_1=3$, $\alpha_2=2$, $\beta_1=5$ and $\beta_2=6$, and the root of maximum modulus of the equation $x^2-3x-2=0$ is $r=\frac{3+\sqrt{17}}{2}$. Since $\frac{\nu(G_n^{2,2})}{n} = 2$, it follows that

$$\lim_{n\to\infty} \mathcal{I}(G_n^{2,2}) = \frac{5r+6}{2(3r+4))} = \frac{51+\sqrt{17}}{68} \approx 0.8106. \quad \blacksquare$$

Another chain of cycles is studied in [71], where two consecutive cycles share a common vertex. More precisely, let G_n^3 be the graph obtained by considering n disjoint copies of the cycle on 3 vertices, by choosing two vertices $v_{i,1}$ and $v_{i,2}$ in each cycle, and by merging (identifying) $v_{i,2}$ and $v_{i+1,1}$ $(i=1,\ldots,n-1)$. Also, let G_n^{3-} be the graph obtained from G_n^3 by removing $v_{1,1}$. For illustration, G_4^3 and G_4^{3-} are shown in Figure 5.6. While the following result is proven in [71], we give here a simple proof based on our technique.

Theorem 21.
$\lim_{n\to\infty} \mathcal{I}(G_n^3) = \frac{2r^2+r+2}{2r^2+3}$ *where* $r = \frac{1}{3}\Big(2 + \sqrt[3]{\frac{43+3\sqrt{177}}{2}} + \sqrt[3]{\frac{43-3\sqrt{177}}{2}}\Big)$.

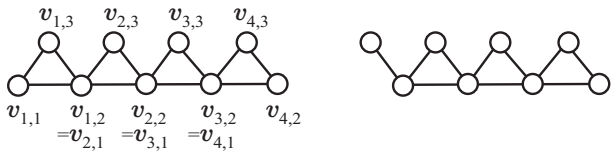


Figure 5.6: G_4^3 and G_4^{3-}.

Proof. Let M be a maximal matching of size k in G_n^3. Exactly one of $v_{1,1}v_{1,2}$, $v_{1,1}v_{1,3}$, $v_{1,2}v_{1,3}$ is in M. If $v_{1,1}v_{1,3} \in M$, the other edges in M form a maximal matching in G_{n-1}^3. If $v_{1,1}v_{1,2}$ or $v_{1,2}v_{1,3}$ belongs to M, the other edges in M form a maximal matching in G_{n-1}^{3-}. Hence,

$$S(G_n^3, k) = S(G_{n-1}^3, k-1) + 2S(G_{n-1}^{3-}, k-1). \tag{5.8}$$

Let now M be a maximal matching of size k in G_n^{3-}. If $v_{1,3}v_{2,1} \in M$, the other edges of M form a maximal matching in G_{n-1}^{3-}. If $v_{2,1}v_{2,3} \in M$, the other edges of M form a maximal matching in G_{n-2}^3. If $v_{2,1}v_{2,2} \in M$, the other edges of M form a maximal matching in G_{n-2}^{3-}. Hence,

$$S(G_n^{3-}, k) = S(G_{n-2}^3, k-1) + S(G_{n-1}^{3-}, k-1) + S(G_{n-2}^{3-}, k-1). \tag{5.9}$$

Combining (5.8) and (5.9) we get

$$\begin{aligned}
S(G_{n+1}^3, k+1) - S(G_n^3, k) &= 2S(G_{n-2}^3, k-1) + S(G_n^3, k) - S(G_{n-1}^3, k-1) \\
&\quad + S(G_{n-1}^3, k) - S(G_{n-2}^3, k-1) \\
\Leftrightarrow S(G_{n+1}^3, k+1) &= 2S(G_n^3, k) + S(G_{-1n}^3, k) - S(G_{n-1}^3, k-1) + S(G_{n-2}^3, k-1) \\
\Leftrightarrow S(G_n^3, k) &= 2S(G_{n-1}^3, k-1) + S(G_{n-2}^3, k-1) - S(G_{n-2}^3, k-2) + S(G_n^3, k-2).
\end{aligned}$$

Hence, $I = 3$, $\alpha_1 = 2$, $\alpha_2 = 0$, $\alpha_3 = 1$, $\beta_1 = 2$, $\beta_2 = -1$ and $\beta_3 = 2$, and the root r of maximum modulus of the equation $x^3 - 2x^2 - 1 = 0$ is

$$r = \frac{1}{3}\Big(2 + \sqrt[3]{\frac{43 + 3\sqrt{177}}{2}} + \sqrt[3]{\frac{43 - 3\sqrt{177}}{2}}\Big).$$

Since $\frac{\nu(G_n^3)}{n} = 1$, we have

$$\lim_{n\to\infty} \mathcal{I}(G_n^3) = \frac{2r^2 - r + 2}{2r^2 + 3} \approx 0.74817. \qquad \blacksquare$$

Note that G_n^3 can be considered as a chain of cliques since $\mathsf{C}_3 \equiv \mathsf{K}_3$. The next subsection will consider other chains of cliques.

5.5.3 Chains of cliques

Let $G_n^{4,s}$ be the graph obtained by taking n disjoint copies of a clique on $s \geq 1$ vertices, by choosing one vertex v_i $(i = 1, \ldots, n)$ in each clique, by adding n vertices $w_1, \ldots, w_n$ and linking w_i to v_i $(i = 1, \ldots, n)$, and by linking v_i to v_{i+1} $(i = 1, \ldots, n-1)$. For illustration, $G_5^{4,1}$ and $G_3^{4,3}$ are depicted in Figure 5.7. Observe that $G_n^{4,1} = \widetilde{\mathsf{P}}_n$. It is proven in Andriantiana *et al.* [12] that $\lim_{n\to\infty} \mathcal{I}^{ARW}(\mathsf{P}_n) = \frac{5-\sqrt{5}}{5}$. Therefore, the link between $\mathcal{I}^{ARW}(G)$ and $\mathcal{I}(\widetilde{G})$ established in Section 5.3 implies $\mathcal{I}(\widetilde{\mathsf{P}}_n) = 1 - \frac{\mathcal{I}^{ARW}(\mathsf{P}_n)\nu(\mathsf{P}_n)}{n}$. Since $\nu(\mathsf{P}_n) = \lfloor \frac{n}{2} \rfloor$, we have

$$\begin{aligned}\lim_{n\to\infty} \mathcal{I}(G_n^{4,1}) &= 1 - \lim_{n\to\infty} \frac{\mathcal{I}^{ARW}(\mathsf{P}_n)\nu(\mathsf{P}_n)}{n} = 1 - \lim_{n\to\infty} \frac{\frac{5-\sqrt{5}}{5}\lfloor \frac{n}{2} \rfloor}{n} \\ &= 1 - \frac{5-\sqrt{5}}{10} = \frac{5+\sqrt{5}}{10}.\end{aligned}$$

Let $f(n) = \mathcal{T}_0(\mathsf{K}_n)$ be the number of maximal matchings in a clique of order n. Clearly,

$$f(n) = \begin{cases} n!! & \text{if } n \text{ is odd,} \\ (n-1)!! & \text{if } n \text{ is even,} \end{cases}$$

where $n!!$ is the double factorial that equals $\prod_{k=0}^{\frac{n-1}{2}}(n-2k)$ for an odd number n. The following theorem gives the value of $\lim\limits_{n\to\infty} \mathcal{I}(G_n^{4,s})$ for all $s \geq 1$.

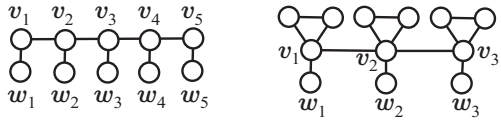


Figure 5.7: $G_5^{4,1}$ and $G_3^{4,3}$.

Theorem 22.

$$\lim_{n\to\infty} \mathcal{I}(G_n^{4,s}) = \frac{r\Big(\lfloor \frac{s+1}{2} \rfloor f(s-1) + \lceil \frac{s-1}{2} \rceil (s-1) f(s-2)\Big) + (2\lfloor \frac{s-1}{2} \rfloor + 1) f(s-1)^2}{\lceil \frac{s}{2} \rceil \Big(r\big(f(s-1) + (s-1)f(s-2)\big) + 2f(s-1)^2\Big)}$$

$$\begin{aligned}\textit{where } r = \frac{1}{2}\Big(& f(s-1) + (s-1)f(s-2) \\ & + \sqrt{5f(s-1)^2 + 2(s-1)f(s-1)f(s-2) + (s-1)^2 f(s-2)^2}\Big).\end{aligned}$$

Proof. Let M be a maximal matching of size k in $G_n^{4,s}$, let K be the clique of order s that contains v_1, and let K' be the clique of order s that contains v_2. There are three possible cases.

- If $v_1w_1 \in M$, then M contains also a maximal matching of $\lfloor \frac{s-1}{2} \rfloor$ edges in $K \setminus \{v_1\}$ and a maximal matching of $k - \lfloor \frac{s+1}{2} \rfloor$ edges in $G_{n-1}^{4,s}$. Since there are $f(s-1)$ matchings of size $\lfloor \frac{s-1}{2} \rfloor$ in $K \setminus \{v_1\}$, the number of matchings of size k in $G_n^{4,s}$ that contain v_1w_1 is equal to $f(s-1)S(G_{n-1}^{4,s}, k - \lfloor \frac{s+1}{2} \rfloor)$.

- If $s \geq 2$ and M contains one of the $s-1$ edges incident to v_1 in K, call it uv_1, then M also contains a maximal matching of $\lfloor \frac{s-2}{2} \rfloor$ edges in K and a maximal matching of $k - \lceil \frac{s-1}{2} \rceil$ edges in $G_{n-1}^{4,s}$. Since there are $f(s-2)$ matchings of size $\lfloor \frac{s-2}{2} \rfloor$ in $K \setminus \{v_1, u\}$, the number of matchings of size k in $G_n^{4,s}$ that contain an edge incident to v_1 in K is equal to $(s-1)f(s-2)S(G_{n-1}^{4,s}, k - \lceil \frac{s-1}{2} \rceil)$.

- If $v_1v_2 \in M$, then M also contains a maximal matching of $\lfloor \frac{s-1}{2} \rfloor$ edges in $K \setminus \{v_1\}$, a maximal matching of $\lfloor \frac{s-1}{2} \rfloor$ edges in $K' \setminus \{v_2\}$, and a maximal matching of $k - (2\lfloor \frac{s-1}{2} \rfloor + 1)$ edges in $G_{n-2}^{4,s}$. Since there are $f(s-1)$ matchings of size $\lfloor \frac{s-1}{2} \rfloor$ in $K \setminus \{v_1\}$ and in $K' \setminus \{v_2\}$, the number of matchings of size k in $G_n^{4,s}$ that contain v_1v_2 is equal to $f(s-1)^2 S(G_{n-2}^{4,s}, k - 2\lfloor \frac{s-1}{2} \rfloor - 1)$.

We have thus shown that

$$\begin{aligned} S(G_n^{4,s}) = {} & f(s-1)S(G_{n-1}^{4,s}, k - \left\lfloor \frac{s+1}{2} \right\rfloor) \\ & + (s-1)f(s-2)S(G_{n-1}^{4,s}, k - \left\lceil \frac{s-1}{2} \right\rceil) \\ & + f(s-1)^2 S(G_{n-2}^{4,s}, k - 2\left\lfloor \frac{s-1}{2} \right\rfloor - 1), \end{aligned}$$

which means that $I = 2$, $\alpha_1 = f(s-1) + (s-1)f(s-2)$, $\alpha_2 = f(s-1)^2$, $\beta_1 = \lfloor \frac{s+1}{2} \rfloor f(s-1) + \lceil \frac{s-1}{2} \rceil (s-1)f(s-2)$ and $\beta_2 = (2\lfloor \frac{s-1}{2} \rfloor + 1)f(s-1)^2$. The root of maximum modulus of the equation $x^2 - \alpha_1 x - \alpha_2 = 0$ is $r = \frac{1}{2}(\alpha_1 + \sqrt{\alpha_1^2 + 4\alpha_2})$ which gives

$$\begin{aligned} r &= \frac{1}{2}(f(s-1) + (s-1)f(s-2) \\ &\quad + \sqrt{(f(s-1) + (s-1)f(s-2))^2 + 4f(s-1)^2} \\ &= \frac{1}{2}(f(s-1) + (s-1)f(s-2) \end{aligned}$$

$$+\sqrt{5f(s-1)^2+2(s-1)f(s-1)f(s-2)+(s-1)^2f(s-2)^2}.$$

To conclude the proof, it is sufficient to observe that $\nu(G_n^{4,s}) = \lceil \frac{s}{2} \rceil n$, which implies $\lim_{n\to\infty} \frac{\nu(G_n^{4,s})}{n} = \lceil \frac{s}{2} \rceil$. ■

For illustration, for $s = 1$, Theorem 22 gives $\lim_{n\to\infty} \mathcal{I}(G_n^{4,1}) = \frac{r+1}{r+2}$ with $r = \frac{1+\sqrt{5}}{2}$, which, as expected, implies

$$\lim_{n\to\infty} \mathcal{I}(G_n^{4,1}) = \frac{3+\sqrt{5}}{5+\sqrt{5}} = \frac{5+\sqrt{5}}{10} \approx 0.7236.$$

The smallest value of $\lim_{n\to\infty} \mathcal{I}(G_n^{4,s})$ is obtained with $s = 3$, where $\lim_{n\to\infty} \mathcal{I}(G_n^{4,3}) = \frac{4r+3}{6r+4}$ with $r = \frac{3+\sqrt{13}}{2}$, which implies

$$\lim_{n\to\infty} \mathcal{I}(G_n^{4,3}) = \frac{18+4\sqrt{13}}{26+6\sqrt{13}} = \frac{39-\sqrt{13}}{52} \approx 0.6807.$$

The values of $\lim_{n\to\infty} \mathcal{I}(G_n^{4,s})$ for $s \leq 50$ are shown in Figure 5.8.

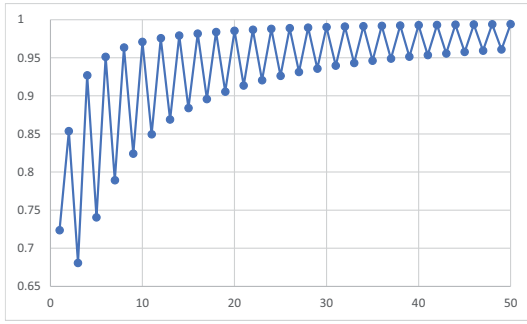


Figure 5.8: $\lim_{n\to\infty} \mathcal{I}(G_n^{4,s})$ for $s \leq 50$.

Corollary 23. $\lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{C}}_n) = \lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{P}}_n)$.

Proof. Let C_n be a cycle on n vertices $v_1, \ldots, v_n$ with edges v_iv_{i+1} $(i = 1, \ldots, n-1)$ and v_1v_n, and let w_i be the vertex of degree 1 linked to v_i $(i = 1, \ldots, n)$ in $\widetilde{\mathsf{C}}_n$. Also, let M be a maximal matching of size k in $\widetilde{\mathsf{C}}_n$. If $w_1v_1 \in M$, then $M \setminus \{w_1v_1\}$ is a maximal matching in $\widetilde{\mathsf{P}}_{n-1}$ obtained from $\widetilde{\mathsf{C}}_n$ by removing v_1 and w_1. If $w_1v_1 \notin M$, then M contains v_1v_2 or v_1v_n as well

as a maximal matching in $\widetilde{\mathsf{P}}_{n-2}$ obtained from $\widetilde{\mathsf{C}}_n$ by removing w_1, v_1 and one of v_2, v_n. Hence

$$S(\widetilde{\mathsf{C}}_n, k) = S(\widetilde{\mathsf{P}}_{n-1}, k-1) + 2S(\widetilde{\mathsf{P}}_{n-2}, k-1).$$

Theorem 22 implies

$$S(\widetilde{\mathsf{P}}_n, k) = S(\widetilde{\mathsf{P}}_{n-1}, k-1) + S(\widetilde{\mathsf{P}}_{n-2}, k-1)$$

which gives,

$$\begin{aligned} S(\widetilde{\mathsf{C}}_n, k) &= S(\widetilde{\mathsf{P}}_{n-2}, k-2) + S(\widetilde{\mathsf{P}}_{n-3}, k-2) \\ &\quad + 2S(\widetilde{\mathsf{P}}_{n-3}, k-2) + 2S(\widetilde{\mathsf{P}}_{n-4}, k-2) \\ &= S(\widetilde{\mathsf{C}}_{n-1}, k-1) + S(\widetilde{\mathsf{C}}_{n-2}, k-1). \end{aligned}$$

This is the same recurrence relation as in Theorem 22 for $s = 1$. Since $\nu(\widetilde{\mathsf{C}}_n) = \nu(\widetilde{\mathsf{P}}_n) = n$, we conclude that $\lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{C}}_n) = \lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{P}}_n)$. ■

We now consider another kind of chain of cliques. For $s \geq 2$ let $G_n^{5,s}$ be the graph obtained by taking n disjoint copies of a clique on s vertices, by choosing two vertices $v_{i,1}$ and $v_{i,2}$ in each clique, and by linking $v_{i,2}$ to $v_{i+1,1}$ $(i = 1, \ldots, n-1)$. Also, let $G_n^{5,s+}$ be the graph obtained from $G_n^{5,s}$ by adding a vertex w linked to $v_{1,1}$ and, if $s \geq 3$, let $G_n^{5,s-}$ be the graph obtained from $G_n^{5,s}$ by removing a vertex $w \neq v_{1,1}, v_{1,2}$ in the clique of s vertices that contains $v_{1,1}$ and $v_{1,2}$, For illustration, $G_3^{5,4}$, $G_3^{5,4+}$ and $G_3^{5,4-}$ are depicted in Figure 5.9.

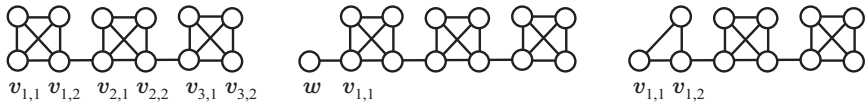


Figure 5.9: $G_3^{5,4}$, $G_3^{5,4+}$ and $G_3^{5,4-}$.

Theorem 24. *If $s \geq 3$ is odd, then*

$$\lim_{n\to\infty} \mathcal{I}(G_n^{5,s}) = \frac{r^2 \frac{s^2-1}{2} f(s-2) - r(s^2-s-1)f(s-2)^2 + \frac{2(s-1)}{2} f(s-2)^3}{\frac{s}{2}(r^2(s+1)f(s-2) - 2rsf(s-2)^2 + 3f(s-2)^3)}$$

where r is the root of maximum modulus of the equation

$$x^3 - (s+1)f(s-2)x^2 + sf(s-2)^2 x - f(s-2)^3 = 0,$$

while if $s \geq 2$ is even, then

$$\lim_{n\to\infty} \mathcal{I}(G_n^{5,s}) = \frac{r^2\frac{s^2}{2}f(s-2) + rs(s-1)(s-3)f(s-2)^2 + \frac{3s^2-5s+2}{2}f(s-2)^3}{\frac{s}{2}(r^2sf(s-2) + 2r(s^2-3s+1)f(s-2)^2 + 3(s-1)f(s-2)^3)}$$

where r is the root of maximum modulus of the equation

$$x^3 - sf(s-2)x^2 - (s^2-3s+1)f(s-2)^2x - (s-1)f(s-2)^3 = 0.$$

Proof. Let's first assume that s is odd, let M be a maximal matching of size k in $G_n^{5,s}$, and let K be the clique of order s that contains $v_{1,1}$ and $v_{1,2}$. If M contains one of the $s-1$ edges incident to $v_{1,2}$ in K, call it $uv_{1,2}$, then M also contains a maximal matching of $\frac{s-3}{2}$ edges in $K \setminus \{u, v_{1,2}\}$ and a maximal matching of $k - \frac{s-1}{2}$ edges in $G_{n-1}^{5,s}$. On the other hand, if M has no edge of K incident to $v_{1,2}$, then it contains a maximal matching of $\frac{s-1}{2}$ edges of $K \setminus \{v_{1,2}\}$ as well as a maximal matching of $k - \frac{s-1}{2}$ edges in $G_{n-1}^{5,s+}$. Since there are $f(s-2)$ maximal matchings in cliques of order $s-1$ and $s-2$, we get:

$$S(G_n^{5,s},k) = (s-1)f(s-2)S(G_{n-1}^{5,s}, k - \frac{s-1}{2}) + f(s-2)S(G_{n-1}^{5,s+}, k - \frac{s-1}{2}). \tag{5.10}$$

Consider now a maximal matching M of size k in $G_n^{5,s-}$. If M contains one of the $s-2$ edges incident to $v_{1,2}$ in K, call it $uv_{1,2}$, then M also contains a maximal matching of $\frac{s-3}{2}$ edges in $K \setminus \{w, u, v_{1,2}\}$ and a maximal matching of $k - \frac{s-1}{2}$ edges in $G_{n-1}^{5,s}$. Otherwise, $v_{1,2}v_{2,1} \in M$ and M contains a maximal matching of $\frac{s-3}{2}$ edges in $K \setminus \{w, v_{1,2}\}$ and a maximal matching of $k - \frac{s-1}{2}$ edges in $G_{n-1}^{5,s-}$. Since there are $f(s-3)$ maximal matchings in cliques of order $s-3$ and $f(s-2)$ maximal matchings in cliques of order $s-2$, we get:

$$\begin{aligned} S(G_n^{5,s-},k) = {} & (s-2)f(s-3)S(G_{n-1}^{5,s}, k - \frac{s-1}{2}) \\ & + f(s-2)S(G_{n-1}^{5,s-}, k - \frac{s-1}{2}). \end{aligned} \tag{5.11}$$

Consider finally a maximal matching M of size k in $G_n^{5,s+}$. If $wv_{1,1} \in M$, then M contains a maximal matching of size $k-1$ in $G_n^{5,s-}$. If $v_{1,1}v_{1,2} \in M$, then M contains a maximal matching of $\frac{s-3}{2}$ edges in $K \setminus \{v_{1,1}, v_{1,2}\}$ and a maximal matching of $k - \frac{s-1}{2}$ edges in $G_{n-1}^{5,s}$. The remaining case is when M contains one of the $s-2$ edges incident to $v_{1,1}$ in $K \setminus \{v_{1,1}, v_{1,2}\}$, call it $u, v_{1,1}$. There are two possibilities: if M contains an edge $u'v_{1,2}$ of $K \setminus \{u, v_{1,1}\}$, then $s \geq 5$ and M also contains a maximal matching of $\frac{s-5}{2}$ edges in $K \setminus \{u, u', v_{1,1}, v_{1,2}\}$ and a

maximal matching of $k-\frac{s-1}{2}$ edges in $G^{5,s}_{n-1}$; if no edge in M is incident to $v_{1,2}$ in K, then M contains a maximal matching of $\frac{s-3}{2}$ edges in $K\setminus\{u,v_{1,1},v_{1,2}\}$ and a maximal matching of $k-\frac{s-1}{2}$ edges in $G^{5,s+}_{n-1}$. Altogether this gives

$$\begin{aligned}S(G^{5,s+}_n,k) = S(G^{5,s-}_n,k-1) &+ f(s-2)S(G^{5,s}_{n-1},k-\frac{s-1}{2})\\ &+ (s-2)\Big((s-3)f(s-3)S(G^{5,s}_{n-1},k-\frac{s-1}{2})\\ &\quad + f(s-3)S(G^{5,s+}_{n-1},k-\frac{s-1}{2})\Big). \end{aligned} \tag{5.12}$$

Using the same proof technique as in Theorem 18, it is not difficult to show that playing with Equations (5.10)-(5.12) leads to

$$\begin{aligned}S(G^{5,s}_n,k) = (s+1)f(s-2)&S(G^{5,s}_{n-1},k-\frac{s-1}{2})\\ &- (s+1)f(s-2)^2S(G^{5,s}_{n-2},k-s+1)\\ &+ f(s-2)^2S(G^{5,s}_{n-2},k-s) + f(s-2)^3S(G^{5,s}_{n-3},k-\frac{3(s-1)}{2}).\end{aligned} \tag{5.13}$$

Hence, we have I=3, α_1=$(s+1)f(s-2)$, α_2= $-sf(s-2)^2$, α_3=$f(s-2)^3$, β_1=$\frac{s^2-1}{2}f(s-2)$, β_2=$(-s^2+s+1)f(s-2)^2$, and $\beta_3=\frac{3(s-1)}{2}f(s-2)^3$. Since $\nu(G^{5,s}_n)=\frac{ns-1}{2}$, we have $\lim_{n\to\infty}\frac{\nu(G^{5,s}_n)}{n}=\frac{s}{2}$, which means that by setting r equal to the root of maximum modulus of the equation $x^3-\alpha_1x^2-\alpha_2x-\alpha_3=0$, one gets the result in the statement of the theorem.
For s even, we have the following relations obtained with a similar analysis as for s odd:

$$\begin{aligned}S(G^{5,s}_n,k) = {}& f(s-1)S(G^{5,s}_{n-1},k-\frac{s}{2})\\ &+ (s-2)f(s-1)f(s-2)S(G^{5,s}_{n-1},k-s+1)\\ &+ f(s-1)f(s-2)S(G^{5,s+}_{n-2},k-s+1).\end{aligned} \tag{5.14}$$

$$\begin{aligned}S(G^{5,s+}_n,k) = {}& (f(s-1)+(s-2)f(s-2))S(G^{5,s}_{n-1},k-\frac{s}{2})\\ &+ (s-2)^2f(s-2)^2S(G^{5,s}_{n-2},k-s+1)\\ &+ f(s-2)S(G^{5,s+}_{n-1},k-\frac{s}{2}) + (s-2)f(s-2)^2S(G^{5,s+}_{n-2},k-s+1).\end{aligned} \tag{5.15}$$

By playing with these two equations as we did for s odd, we get

$$S(G^{5,s}_n,k) = sf(s-2)S(G^{5,s}_{n-1},k-\frac{s}{2}) + s(s-2)f(s-2)^2S(G^{5,s}_{n-2},k-s+1)$$

$$- (s-1)f(s-2)^2 S(G_{n-2}^{5,s}, k-s)$$
$$+ (s-1)f(s-2)^3 S(G_{n-3}^{5,s}, k-\frac{3s-2}{2}). \tag{5.16}$$

Hence, we have $I = 3$, $\alpha_1 = sf(s-2)$, $\alpha_2 = (s^2-3s+1)f(s-2)^2$, $\alpha_3 = (s-1)f(s-2)^3$, $\beta_1 = \frac{s^2}{2}f(s-2)$, $\beta_2 = s(s-1)(s-3)f(s-2)^2$, $\beta_3 = \frac{3s^2-5s+2}{2}f(s-2)^3$. Since $\frac{\nu(G_n^{5,s})}{n} = \frac{s}{2}$, by setting r equal to the root of maximum modulus of the equation $x^3 - \alpha_1 x^2 - \alpha_2 x - \alpha_3 = 0$, one gets the result in the statement of the theorem. ■

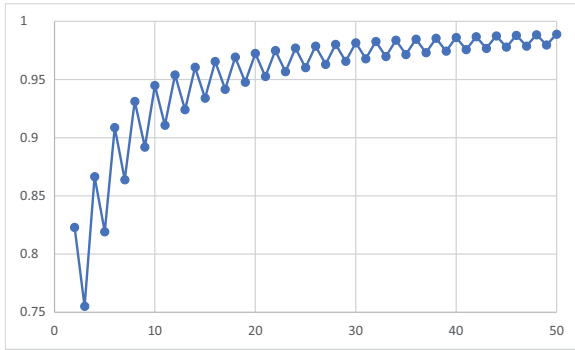


Figure 5.10: $\lim_{n\to\infty} \mathcal{I}(G_n^{5,s})$ for $s \leq 50$.

The values of $\lim_{n\to\infty} \mathcal{I}(G_n^{5,s})$ for $s \leq 50$ are shown in Figure 5.10. The lowest value is obtained with $s = 3$ for which we get the following recurrence equation

$$S(G_n^{5,3}, k) = 4S(G_{n-1}^{5,3}, k-1) - 4S(G_{n-2}^{5,3}, k-2) + S(G_{n-2}^{5,3}, k-3) + S(G_{n-3}^{5,3}, k-3).$$

The root with maximum modulus of the equation $x^3 - 4x^2 + 3x - 1 = 0$ is

$$r = \frac{1}{3}\left(4 + \sqrt[3]{\frac{47+3\sqrt{93}}{2}} + \sqrt[3]{\frac{47-3\sqrt{93}}{2}}\right)$$

which gives

$$\lim_{n\to\infty} \mathcal{I}(G_n^{5,3}) = \frac{2(4r^2-5r+3)}{3(4r^2-6r+3)} \approx 0.75503.$$

Note that $G_n^{5,2} \equiv \mathsf{P}_{2n}$ and, for $s = 2$, Equation (5.16) gives

$$S(G_n^{5,2}, k) = 2S(G_{n-1}^{5,2}, k-1) - S(G_{n-2}^{5,2}, k-2) + S(G_{n-3}^{5,2}, k-2)$$

which we could have derived from Theorem 15. Indeed:

$$S(G_n^{5,2}, k) = S(P_{2n}, k) = S(P_{2n-2}, k-1) + S(P_{2n-3}, k-1)$$

$$\begin{aligned}
&= S(P_{2n-2}, k-1) + S(P_{2n-5}, k-2) + S(P_{2n-6}, k-2) \\
&= S(P_{2n-2}, k-1) + (S(P_{2n-2}, k-1) - S(P_{2n-4}, k-2)) \\
&\quad + S(P_{2n-6}, k-2) \\
&= 2S(G_{n-1}^{5,2}, k-1) - S(G_{n-2}^{5,2}, k-2) + S(G_{n-3}^{5,2}, k-2).
\end{aligned}$$

5.5.4 Ladders

Let G_n^6 be the Cartesian product of P_2 with P_n which is a *ladder* with vertices v_i, w_i (i=1,...,n) and edges v_iv_{i+1}, w_iw_{i+1}(i=1,...,n−1) and v_iw_i (i=1,...,n). Also, let $G_n^{6,1}$ be the graph obtained from G_n^6 by adding a vertex u_1 linked to v_1 and let $G_n^{6,2}$ be the graph obtained from $G_n^{6,1}$ by adding a vertex u_2 linked to u_1. For illustration, G_3^6, $G_3^{6,1}$ and $G_3^{6,2}$ are depicted in Figure 5.11.

Theorem 25. $\lim_{n\to\infty} \mathcal{I}(G_n^6) = \frac{2r^4-r^2+3r+4}{2r^4+4r+5)}$ *where* $r = \frac{1}{3}\Big(1 + \sqrt[3]{\frac{47+3\sqrt{93}}{2}} + \sqrt[3]{\frac{47-3\sqrt{93}}{2}}\Big)$.

Proof. Let M be a maximal matching of size k in G_n^6. If $v_1w_1 \in M$, then $M \setminus \{v_1w_1\}$ is a maximal matching in G_{n-1}^6. If $v_1v_2 \in M$, then $M \setminus \{v_1v_2\}$ is a maximal matching in $G_{n-1}^{6,2}$. If $w_1w_2 \in M$, then $M \setminus \{w_1w_2\}$ is a maximal matching in $G_{n-2}^{6,2}$. If M contains both v_1v_2 and w_1w_2, then $M \setminus \{v_1v_2, w_1w_2\}$ is a maximal matching in G_{n-2}^6. In summary:

$$S(G_n^6, k) = S(G_{n-1}^6, k-1) - S(G_{n-2}^6, k-2) + 2S(G_{n-2}^{6,2}, k-1). \tag{5.17}$$

Let now M be a maximal matching of size k in $G_n^{6,1}$. If $u_1v_1 \in M$, then $M \setminus \{u_1v_1\}$ is a maximal matching in $G_{n-1}^{6,1}$. If $v_1v_2 \in M$, then $M \setminus \{v_1v_2\}$ is a maximal matching in $G_{n-2}^{6,2}$. If $v_1w_1 \in M$, then $M \setminus \{v_1w_1\}$ is a maximal matching in G_{n-1}^6. Hence

$$S(G_n^{6,1}, k) = S(G_{n-1}^6, k-1) + S(G_{n-1}^{6,1}, k-1) + S(G_{n-2}^{6,2}, k-1). \tag{5.18}$$

Finally, let M be a maximal matching of size k in $G_n^{6,2}$. If $u_1u_2 \in M$, then $M \setminus \{u_1u_2\}$ is a maximal matching in G_n^6. If $u_1v_1 \in M$, then $M \setminus \{u_1v_1\}$ is a maximal matching in $G_{n-1}^{6,1}$. Hence

$$S(G_n^{6,2}, k) = S(G_n^6, k-1) + S(G_{n-1}^{6,1}, k-1). \tag{5.19}$$

Equations (5.17) and (5.19) give

$$S(G_n^6, k) = S(G_{n-1}^6, k-1) + S(G_{n-2}^6, k-2) + 2S(G_{n-3}^{6,1}, k-2) \tag{5.20}$$

$$\begin{aligned}2S(G_n^{6,1},k) &= 2S(G_{n-1}^6,k-1)+2S(G_{n-1}^{6,1},k-1)\\ &\quad+\Big(S(G_n^6,k)-S(G_{n-1}^6,k-1)+S(G_{n-2}^6,k-2)\Big)\\ &= S(G_n^6,k)+S(G_{n-1}^6,k-1)+S(G_{n-2}^6,k-2)+2S(G_{n-1}^{6,1},k-1).\end{aligned} \tag{5.21}$$

By combining Equations (5.20) and (5.21) we get

$$\begin{aligned}S(G_n^6,k) &= \Big(2S(G_{n-1}^6,k-1)-\big(S(G_{n-2}^6,k-2)+S(G_{n-3}^6,k-3)+2S(G_{n-4}^{6,1},k-3)\big)\Big)\\ &\quad+S(G_{n-2}^6,k-2)+\Big(S(G_{n-3}^6,k-2)+S(G_{n-4}^6,k-3)+S(G_{n-5}^6,k-4)+2S(G_{n-4}^{6,1},k-3)\Big)\\ &= 2S(G_{n-1}^6,k-1)+S(G_{n-3}^6,k-2)-S(G_{n-3}^6,k-3)+S(G_{n-4}^6,k-3)+S(G_{n-5}^6,k-4).\end{aligned}$$

We have I=5, α_1=2, α_2=α_3=0, α_4=α_5=1, β_1=2, β_2=0, β_3= − 1, β_4=3 and $\beta_5 = 4$. The root of maximum modulus of the equation $x^5-2x^4-x-1=0$ is

$$r=\frac{1}{3}\Big(1+\sqrt[3]{\frac{47+3\sqrt{93}}{2}}+\sqrt[3]{\frac{47-3\sqrt{93}}{2}}\Big).$$

Since $\frac{\nu(G_n^6)}{n}=1$, we have

$$\lim_{n\to\infty}\mathcal{I}(G_n^6)=\frac{2r^4-r^2+3r+4}{2r^4+4r+5}\approx 0.8618.$$ ■

We now consider $\widetilde{G}_n^6$ as well as $\widetilde{G}_n^{6,+}$ which is obtained from $\widetilde{G}_n^6$ by adding a vertex u_1 linked to v_1 and a vertex u_2 linked to u_1. For illustration, $\widetilde{G}_3^6$ and $\widetilde{G}_3^{6+}$ are depicted in Figure 5.11.

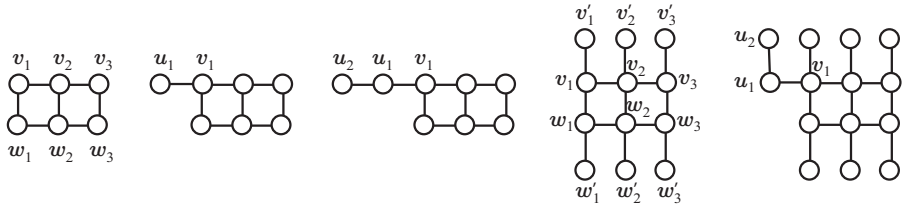


Figure 5.11: G_3^6, $G_3^{6,1}$, $G_3^{6,2}$, $\widetilde{G}_3^6$ and $\widetilde{G}_3^{6+}$.

Theorem 26. $\lim_{n\to\infty}\mathcal{I}(\widetilde{G}_n^6)=\frac{4r^2+3r-3}{2(3r^2+2r-3)}$ *where* $r=1+\frac{12+\sqrt[3]{54(-5+3i\sqrt{111})}}{3\sqrt[3]{27+3i\sqrt{111}}}$.

Proof. Let M be a maximal matching of size k in $\widetilde{G}_n^6$. If $v_1v_1'\in M$, then $M\setminus\{v_1v_1'\}$ is a maximal matching in $\widetilde{G}_{n-1}^{6+}$. If $v_1w_1\in M$, then $M\setminus\{v_1w_1\}$

is a maximal matching in $\widetilde{G}^6_{n-1}$. If $v_1v_2 \in M$, then there are two possibilies: if $w_1, w_1' \in M$, then $M \setminus \{v_1v_2, w_1w_1'\}$ is a maximal matching in $\widetilde{G}^{6+}_{n-2}$; if $w_1, w_2 \in M$, then $M \setminus \{v_1v_2, w_1w_2\}$ is a maximal matching in $\widetilde{G}^6_{n-2}$. In summary

$$\begin{aligned} S(\widetilde{G}^6_n, k) = {} & S(\widetilde{G}^6_{n-1}, k-1) + S(\widetilde{G}^6_{n-2}, k-2) \\ & + S(\widetilde{G}^{6+}_{n-1}, k-1) + S(\widetilde{G}^{6+}_{n-2}, k-2). \end{aligned} \tag{5.22}$$

Let M be a maximal matching of size k in $\widetilde{G}^{6+}_n$. If $u_1u_2 \in M$, then $M \setminus \{u_1u_2\}$ is a maximal matching in $\widetilde{G}^6_n$. If $u_1v_1 \in M$, then $M \setminus \{u_1v_1\}$ is a maximal matching in $\widetilde{G}^{6+}_{n-1}$. Hence

$$S(\widetilde{G}^{6+}_n, k) = S(\widetilde{G}^6_n, k-1) + S(\widetilde{G}^{6+}_{n-1}, k-1). \tag{5.23}$$

Equations (5.22) and (5.23) imply

$$\begin{aligned} S(\widetilde{G}^6_n, k) &= S(\widetilde{G}^6_{n-1}, k-1) + S(\widetilde{G}^6_{n-1}, k-2) + S(\widetilde{G}^6_{n-2}, k-2) + 2S(\widetilde{G}^{6+}_{n-2}, k-2) \\ &= 2S(\widetilde{G}^6_{n-1}, k-1) - \Big(S(\widetilde{G}^6_{n-2}, k-2) + S(\widetilde{G}^6_{n-2}, k-3) + S(\widetilde{G}^6_{n-3}, k-3) + 2S(\widetilde{G}^{6+}_{n-3}, k-3)\Big) \\ &\quad + S(\widetilde{G}^6_{n-1}, k-2) + S(\widetilde{G}^6_{n-2}, k-2) + 2\Big(S(\widetilde{G}^6_{n-2}, k-3) + S(\widetilde{G}^{6+}_{n-3}, k-3)\Big) \\ &= 2S(\widetilde{G}^6_{n-1}, k-1) + S(\widetilde{G}^6_{n-1}, k-2) + S(\widetilde{G}^6_{n-2}, k-3) - S(\widetilde{G}^6_{n-3}, k-3). \end{aligned}$$

We have $I = 3$, $\alpha_1 = 3$, $\alpha_2 = 1$, $\alpha_3 = -1$, $\beta_1 = 4$, $\beta_2 = 3$ and $\beta_3 = -3$, and the root of maximum modulus of the equation $x^3 - 3x^2 - x + 1 = 0$ is

$$r = 1 + \frac{12 + \sqrt[3]{54(-5 + 3i\sqrt{111})}}{3\sqrt[3]{27 + 3i\sqrt{111}}}.$$

Since $\frac{\nu(\widetilde{G}^6_n)}{n} = 2$, we have

$$\lim_{n\to\infty} \mathcal{I}(\widetilde{G}^6_n) = \frac{4r^2 + 3r - 3}{2(3r^2 + 2r - 3)} \approx 0.6968. \qquad \blacksquare$$

5.5.5 Trees

Let's now talk about trees. We have mentioned that Andriantiana *et al.*[12] have proven that the trees of order n that maximize the average size of a (not necessarily maximal) matching are the paths P_n. So let $\boldsymbol{T}_n$ be the set of trees of order n and let $\widetilde{\boldsymbol{T}}_n$ be the set of trees $\widetilde{T}$ with $T \in \boldsymbol{T}_n$. The result in [12] can ve rewritten as

$$\max_{T\in\boldsymbol{T}_n} \frac{\mathcal{T}_1^{ARW}(T)}{\mathcal{T}_0^{ARW}(T)} = \frac{\mathcal{T}_1^{ARW}(\mathsf{P}_n)}{\mathcal{T}_0^{ARW}(\mathsf{P}_n)} \quad \Leftrightarrow \quad \max_{T\in\boldsymbol{T}_n} \nu(T)\mathcal{I}^{ARW}(T) = \nu(\mathsf{P}_n)\mathcal{I}^{ARW}(\mathsf{P}_n).$$

Since $\nu(\widetilde{T}) = n$ for all $\widetilde{T} \in \widetilde{\boldsymbol{T}}_n$, we have

$$\min_{\widetilde{T}\in\widetilde{\boldsymbol{T}}_n} \mathcal{I}(\widetilde{T}) = 1 - \max_{T\in\boldsymbol{T}_n} \frac{\nu(T)\mathcal{I}^{ARW}(T)}{n} = 1 - \frac{\nu(\mathsf{P}_n)\mathcal{I}^{ARW}(\mathsf{P}_n)}{n} = \mathcal{I}(\widetilde{\mathsf{P}}_n).$$

Clearly, $\widetilde{\boldsymbol{T}}_n \subset \boldsymbol{T}_{2n}$ and Dyer and Frieze [76] conjecture that

$$\lim_{n\to\infty} \min_{T\in\boldsymbol{T}_{2n}} \mathcal{I}^{DF}(T) = \lim_{n\to\infty} \mathcal{I}^{DF}(\widetilde{\mathsf{P}}_n).$$

We show here that

$$\lim_{n\to\infty} \min_{T\in\boldsymbol{T}_{2n}} \mathcal{I}(T) < \lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{P}}_n).$$

For this purpose, let T_n be the tree with $4n-2$ vertices obtained from P_{3n-2} by adding vertices $w_1, \ldots, w_n$ so that w_i is linked to v_{3i-2}. Also, let T_n^1 be the graph obtained from T_n by adding a vertex u_1 linked to v_1, and let T_n^2 be the graph obtained from T_n^1 by adding a vertex u_2 linked to u_1. For illustration, T_3, T_3^1 and T_3^2 are drawn in Figure 5.12.

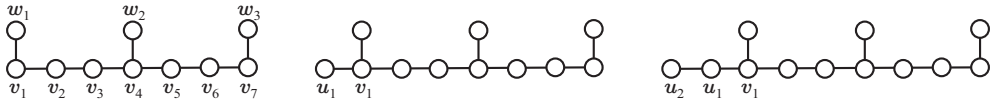


Figure 5.12: T_3, T_3^1 and T_3^2.

Theorem 27. $\lim_{n\to\infty} \mathcal{I}(T_n) = \frac{13}{18}$.

Proof. Let M be a maximal matching of size k in T_n. If $w_1v_1 \in M$, then $M \setminus \{w_1v_1\}$ is a maximal matching in T_{n-1}^2. If $v_1v_2 \in M$, then $M \setminus \{v_1v_2\}$ is a maximal matching in T_{n-1}^1. We therefore have

$$S(T_n, k) = S(T_{n-1}^1, k-1) + S(T_{n-1}^2, k-1). \tag{5.24}$$

Let now M be a maximal matching of size k in T_n^1. If M contains u_1v_1 or w_1v_1, then the $k-1$ other edges of M form a maximal matching in T_{n-1}^2. If $v_1v_2 \in M$, then $M \setminus \{v_1v_2\}$ is a maximal matching in T_{n-1}^1. Hence,

$$S(T_n^1, k) = S(T_{n-1}^1, k-1) + 2S(T_{n-1}^2, k-1) = S(T_n, k) + S(T_{n-1}^2, k-1). \tag{5.25}$$

Finally, let M be a maximal matching of size k in T_n^2. If $u_1u_2 \in M$, then $M \setminus \{u_1u_2\}$ is a maximal matching in T_n. If $u_1v_1 \in M$, then $M \setminus \{u_1v_1\}$ is a maximal matching in T_{n-1}^2. The linear equation is therefore

$$S(T_n^2, k) = S(T_n, k-1) + S(T_{n-1}^2, k-1). \tag{5.26}$$

Combining Equations (5.24) and (5.25) gives

$$\begin{aligned} S(T_n,k) &= \Big(S(T_{n-1},k-1)+S(T^2_{n-2},k-2)\Big) \\ &\quad + \Big(S(T_{n-1},k-2)+S(T^2_{n-2},k-2)\Big) \\ &= S(T_{n-1},k-1)+S(T_{n-1},k-2)+2S(T^2_{n-2},k-2). \end{aligned} \tag{5.27}$$

We can therefore rewrite Equation (5.26) as

$$\begin{aligned} &S(T_{n+2},k+2)-S(T_{n+1},k+1)-S(T_{n+1},k) \\ &\quad = 2S(T_n,k-1)+\Big(S(T_{n+1},k+1)-S(T_n,k)-S(T_n,k-1)\Big) \\ &\Leftrightarrow S(T_n,k) = 2S(T_{n-1},k-1)+S(T_{n-1},k-2)-S(T_{n-2},k-2)+S(T_{n-2},k-3). \end{aligned}$$

We have $I = 2$, $\alpha_1 = 3$, $\alpha_2 = 0$, $\beta_1 = 4$ and $\beta_2 = 1$, and $r = 3$ is the root of maximum modulus of the equation $x^2 - 3x = 0$. Since $\nu(T_n) = 2n+1$, we have $\lim_{n\to\infty} \frac{\nu(T_n)}{n} = 2$ and it follows that

$$\lim_{n\to\infty} \mathcal{I}(T_n) = \frac{4r+1}{2(3r)} = \frac{13}{18} \approx 0.7222. \qquad \blacksquare$$

Note that T_n has an even number of vertices, which implies

$$\lim_{n\to\infty} \min_{T\in \boldsymbol{T}_{2n}} \mathcal{I}(T) \le \frac{13}{18} < \frac{5+\sqrt{5}}{10} = \lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{P}}_n).$$

5.5.6 Thorns of complete bipartite graphs

We conclude our study with the thorn $\widetilde{\mathsf{K}}_{c,n}$ of complete bipartite graphs $\mathsf{K}_{c,n}$, where c is a constant. Assume without loss of generality that V_1, V_2 is the partition of $\mathsf{K}_{c,n}$ with $|V_1| = c$ and $|V_2| = n$.

Theorem 28. $\lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{K}}_{c,n}) = 1$ *for every integer constant* $c \ge 0$.

Proof. Clearly, a maximal matching in $\widetilde{\mathsf{K}}_{c,n}$ contains i edges linking V_1 to V_2 $(0 \le i \le c)$ and $(c-i)+(n-i) = n+c-2i$ edges linking a vertex v to its neighbor v' of degree 1. Hence, the number of maximal matchings of size $k = n+c-i$ is

$$S(\widetilde{\mathsf{K}}_{c,n},k) = \frac{c!n!}{(n+c-k)!(k-n)!(k-c)!}$$

which implies

$$\mathcal{T}_0(\widetilde{\mathsf{K}}_{c,n}) = \sum_{k=n}^{n+c} \frac{c!n!}{(n+c-k)!(k-n)!(k-c)!}$$

$$\text{and} \quad \mathcal{T}_1(\widetilde{\mathsf{K}}_{c,n}) = \sum_{k=n}^{n+c} \frac{kc!n!}{(n+c-k)!(k-n)!(k-c)!}.$$

Hence, $\lim_{n\to\infty} \mathcal{T}_0(\widetilde{\mathsf{K}}_{c,n}) \sim n^c$ and $\lim_{n\to\infty} \mathcal{T}_1(\widetilde{\mathsf{K}}_{c,n}) \sim n^{c+1}$ which implies

$$\lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{K}}_{c,n}) = \lim_{n\to\infty} \frac{\mathcal{T}_1(\widetilde{\mathsf{K}}_{c,n})}{\nu(\widetilde{\mathsf{K}}_{c,n})\mathcal{T}_0(\widetilde{\mathsf{K}}_{c,n})} = \lim_{n\to\infty} \frac{n^{c+1}}{(n+c)n^c} = 1. \qquad \blacksquare$$

5.6 Concluding remarks

It is well known that the size of a maximal matching in a graph G is at least half of the size $\nu(G)$ of a maximum matching in G. There are graphs in which almost all maximal matchings are maximum while for others, almost all maximal matchings have half of the size of a maximum matching. This prompted us to investigate the ratio $\mathcal{I}(G)$ of the average size of a maximal matching to the size of a maximum matching in G. Clearly, $\mathcal{I}(G) \approx 1$ if many maximal matchings have a size close to $\nu(G)$, while $\mathcal{I}(G) \approx \frac{1}{2}$ if many maximal matchings have a small size.

We have determined $\lim_{n\to\infty} \mathcal{I}(G_n)$ for many families $\{G_n\}_{n\geq 0}$ of graphs, showing that intermediate values between $\frac{1}{2}$ and 1 are reached. We show eighteen of them in Figure 5.13. These values were obtained using a general technique that can be applied when the number of maximal matchings in G_n linearly depends on the number of maximal matchings in graphs $G_{n'}$ of the same family, with $n' < n$. This technique allowed us to recalculate known values of $\lim_{n\to\infty} \mathcal{I}(G_n)$ which were obtained by other authors using different concepts such as (bivariate) generating functions.

A comparison of the curve in Figure 5.13 with the bottom curve of Figure 5.2 demonstrates that the asymptotic values $\lim_{n\to\infty} \mathcal{I}(G_n)$ for the classes of graphs studied in Section 5.5 are representative of the values of $\mathcal{I}(G)$ for any graphs G.

We conclude the chapter with an open problem on the class $\boldsymbol{T}_{2n}$ of trees of even order. Dyer and Frieze [76] conjecture that $\lim_{n\to\infty} \min_{T\in\boldsymbol{T}_{2n}} \mathcal{I}^{DF}(T) = \lim_{n\to\infty} \mathcal{I}^{DF}(\widetilde{\mathsf{P}}_n)$. In simpler words, they think that the worst case for $\mathcal{I}^{DF}$ on $\boldsymbol{T}_{2n}$ are the graphs obtained from P_n by adding a new vertex v' for every v of P_n and linking v to v'. We have shown in Section 5.5.5 that

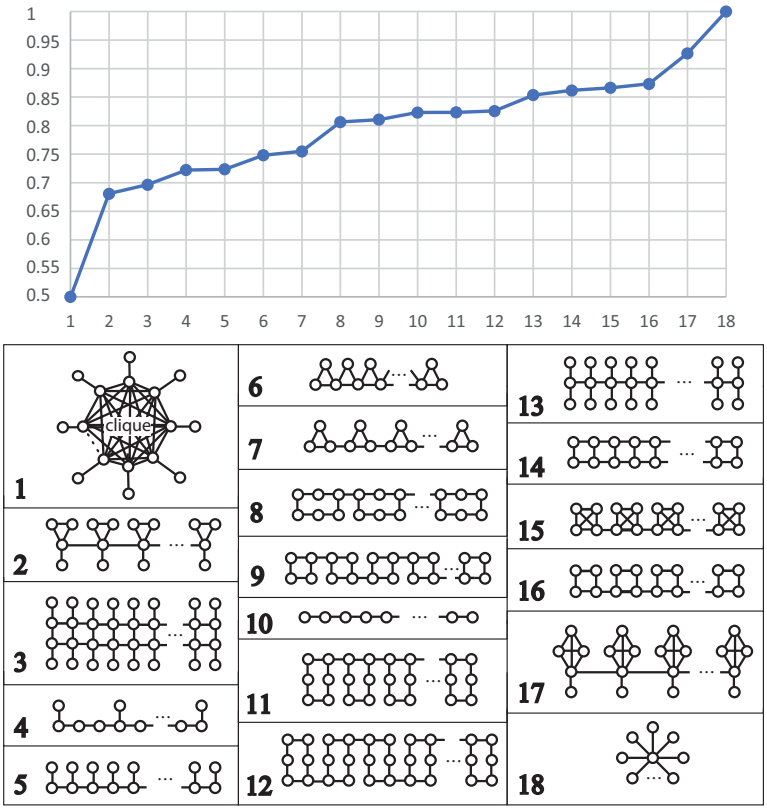


Figure 5.13: Various families $\{G_n\}_{n\geq 0}$ of graphs and their asymptotic value $\lim_{n\to\infty} \mathcal{I}(G_n)$.

$\lim_{n\to\infty} \min_{T\in \boldsymbol{T}_{2n}} \mathcal{I}(T) < \lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{P}}_n)$. Hence, the worst case for $\mathcal{I}$ on $\boldsymbol{T}_{2n}$ seems to be different from that of $\mathcal{I}^{DF}$ on $\boldsymbol{T}_{2n}$ and it would be interesting to characterize it.

Part II

Chemical graph theory

Chapter 6

Chemical graph theory

This chapter serves as an introduction to our work in chemical graph theory. Although this area constitutes an important part of our research, our focus is primarily on the graph-theoretical aspects of the results rather than on their chemical interpretation. Consequently, we do not discuss the chemical implications of the definitions or results in detail.

Section 6.1 provides a definition of the field of chemical graph theory. Section 6.2 introduces topological indices and presents our contributions in this area.

6.1 Definition

Chemical graph theory is the study of *chemical graphs*, which are connected graphs with maximum degree at most 3 or 4 depending on the definition adopted. As noted by Patrick Fowler [84]: "The definition of chemical graphs that is useful depends on context. Two definitions appropriate to different kinds of carbon framework can be found in the literature. The graphs that can be regarded as skeletons of saturated hydrocarbons (such as alkanes), are connected and have maximum degree at most 4. If instead the interest is in (unsaturated) conjugated π systems, such as alkenes, polyenes, benzenoids, and fullerenes, the maximum degree should be at most 3, since a conjugated carbon atom participates in at most three single bonds."

These graphs are used to model molecules, where vertices represent atoms and edges represent bonds. On the left side of Figure 6.1, a molecule of ethane is shown, and on the right side, the corresponding chemical graph appears. The two central vertices represent the carbon atoms, while the remaining vertices represent hydrogen atoms. The edges represent the bonds between the carbon and hydrogen atoms.

Figure 6.1: A molecule of ethane and the corresponding chemical graph

6.2 Topological indices

A *topological index*, also called a *molecular descriptor*, is a numerical graph invariant typically used to characterize a graph's topology. Such indices are frequently employed to study the physicochemical properties of the molecules represented by these graphs [43, 62, 63, 106, 131, 135, 136]. In this work, we focus on a particular subclass known as *degree-based topological indices*, which are defined as sums taken over all edges of the graph, where the contribution of each edge depends on the degrees of its endpoints.

One of the most well known topological index introduced in this form is the Randić index, proposed in 1975 by Randić [183]. For this index, each edge (u, v) contributes $\frac{1}{\sqrt{d_u d_v}}$. The interest generated by this index led to the introduction of many others. Among these, the arithmetic-geometric index, proposed by Shegehalli and Kanabur [193] in 2015, is of particular interest to us. For this invariant, each edge (u, v) contributes $\frac{d_u+d_v}{2\sqrt{d_u d_v}}$. In Chapter 7, which is based on our article [115], we investigate extremal chemical graphs with respect to this invariant. These results concern chemical graphs with maximum degree at most 4.

Without going into detail at this stage, we note that our results in this paper rely primarily on analyzing the contributions of edges whose endpoints have given degrees. After the publication of this aforementioned paper, we conjectured that the same techniques could be applied to other degree-based topological indices as well. This subsequent work is presented in Chapter 8, which is based on our article [21]. These results concern chemical graphs with maximum degree at most 3. There, we show that among the 33 degree-based topological indices that we identified in the literature, 5 families of chemical graphs suffice to characterize the extremal graphs for 29 of them.

This result is interesting as there have been questions on the usefulness of the growing number of degree-based topological indices being introduced. Indeed, I. Gutman [99] stated that "To use a mild expression, today we have far too many such descriptors, and there seems to lack a firm criterion to stop or slow down their proliferation.".

We extend this idea by generalizing the previous result. We provide a complete polyhedral description of chemical graphs of maximum degree at most 3. This means that for a given order, size and degree-based topological index, we can determine the exact minimum and maximum values that such chemical graphs can attain for that order and size, if they exist. We can also characterize the set of graphs that achieve these extremal values through the parameters m_{ij}, which count edges between vertices of degrees i and j. This work is further detailed in Chapter 9, based on our article [22], which is currently under review. Note that this article does not present the full mathematical details of how these results were obtained; instead, it focuses on presenting the results and their applications in chemical graph theory. Readers interested in those details are referred to our initial article on the topic [75], to which we also contributed. These details are omitted from this thesis in order to preserve them for inclusion in the thesis of another member of our research team.

A companion website for this last result was also developed by our team. This website, named ChemicHull, presents these results in a way that improves visual clarity and accessibility. It lists different polytopes that users can filter by specifying values of either n, m, or both. For each polytope, the website provides the list of vertices and facets, as well as a 3D visualization. Finally, when a topological index is provided, the tool displays the extremal vertices for that invariant. This tool is presented in Chapter 11 of Part III.

Chapter 7

Extremal chemical graphs for the arithmetic-geometric index

This chapter presents our work on extremal chemical graphs of maximum degree at most 4 for the arithmetic-geometric index. This work is based on our article [114]. In fact, most of this chapter follows that article, apart from minor changes for coherence. Section 7.1 introduces our work and the further sections of this chapter.

7.1 Introduction

The *arithmetic-geometric index* was proposed in 2015 by Shegehalli and Kanabur [193]. Let G be a graph with edge set E and let d_u and d_v be the degrees of the endpoints of an edge $uv \in E$. The arithmetic-geometric index $\mathtt{AG}(G)$ of G, is defined as

$$\mathtt{AG}(G) = \sum_{uv \in E} \frac{d_u + d_v}{2\sqrt{d_u d_v}}.$$

Shegehalli and Kanabur [193,194] give the value of this index for some families of graphs. The summand in the above formula is the ratio between the arithmetic and geometric means of d_u and d_v. If we replace each summand by its inverse, we obtain another graph invariant called the reciprocal of $\mathtt{AG}$ and also known as the geometric-arithmetic index. It was introduced in 2009 by Vukičević and Furtula [210] and studied for example in [13, 20, 53, 187].

In the past few year, an interest in the arithmetic-geometric index emerged in mathematical chemistry. The $\mathtt{AG}$ index of graphene, identified as the most

conductive material for electromagnetic interference shielding [197], was determined in [195]. Using the AG index, Zheng *et al.* [224] and Guo and Gao [97] study extremal properties of the spectral radius and energy of arithmetic-geometric matrix. In addition, Vujošević et al. [209] have characterized the chemical trees with maximum arithmetic-geometric index value.

Also, the relationship between the arithmetic-geometric index and other topological indices of interest in chemistry is studied, for example in [36, 47, 101, 144, 167]. In particular, it is of interest to determine relations between a topological index and its reciprocal [104]. Gutman [101] showed that AG and its reciprocal (the geometric-arithmetic index GA) will have similar predictive values in QSPR and QSAR applications. Hence, results on AG or on its reciprocal can both be useful for such applications. Note however that for fixed order and size, a graph maximizing AG is not necessarily a graph minimizing its reciprocal. For example, the results presented in this chapter show that the unique graph of order 6 and size 8 maximizing AG is the graph G_1 depicted in Figure 7.1, but this graph does not minimize GA among graphs of order 6 and size 8. Indeed, the graph G_2 in Figure 7.1 has $\mathtt{GA}(G_2) = \frac{8}{3}\sqrt{8} < \frac{18}{5} + \frac{16}{7}\sqrt{3} = \mathtt{GA}(G_1)$.

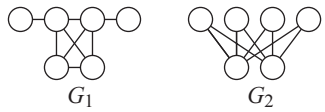


Figure 7.1: Graphs maximising AG do not necessarily minimize GA.

Several papers have focused on extremal properties of the arithmetic-geometric index. For example, some lower and upper bounds on AG are given in [47]. Also, lower bounds for graphs of fixed size (i.e., number of edges) are provided in [209], while an upper bound for graphs with fixed size and order (i.e., number of vertices) is established in [209]. Upper bounds for graphs of fixed size and fixed minimum and maximum degrees are given in [167]. The maximum value of the arithmetic-geometric index of graphs of fixed order is known for unicyclic graphs [211], bicyclic graphs [174], bipartite graphs and trees [209].

In this chapter, we prove the following upper bound on the value of the arithmetic index of a connected chemical graph G of order n and size m:

$$\mathtt{AG}(G) \leq \frac{2n+5m}{6} + \begin{cases} 0 & \text{if } 2m-n \equiv 0 \bmod 3, \\ \frac{3}{\sqrt{2}} - \frac{13}{6} & \text{if } 2m-n \equiv 1 \bmod 3, \\ \frac{21}{4\sqrt{3}} - \frac{37}{12} & \text{if } 2m-n \equiv 2 \bmod 3. \end{cases}$$

We show that with the exception of 22 (n, m) pairs, the bound is sharp, and we characterize the connected chemical graphs of order n and size $m \geq n-1$

that reach the bound. We also prove that no better value can be obtained by removing the constraint that the graph must be connected. Note that for $m = n - 1$, this gives a characterization of extremal chemical trees of fixed order n. While such a characterization is given in [209], we show that their result is not valid for 7 values of n.

In the next section we fix some notations, while Section 7.3 is devoted to observations that will motivate our characterization of connected chemical graphs with maximum arithmetic-geometric index value. Lemmas are proved in Section 7.4 and then used in Section 7.5 to prove the main theorem.

7.2 Notations

For basic notions of graph theory that are not defined here, we refer to Diestel [68]. Let $G = (V, E)$ be a simple undirected graph. The *order* $n = |V|$ of G is its number of vertices and the *size* $m = |E|$ of G is its number of edges. We write $G \simeq H$ if G and H are isomorphic. The *degree* of v, denoted d_v is the number of edges incident to v, and we say that v is *isolated* if $d_v = 0$.

A *chemical graph* is a graph whose vertices have degree at most 4. The arithmetic-geometric index $\mathtt{AG}(G)$ of a graph G can be seen as a sum of costs on the edges of G. In particular, if we deal with chemical graphs, there is a limited number of possible values for the costs since they are computed from the degrees of the endpoints of the edges. Let $c_{i,j} = \frac{i+j}{2\sqrt{ij}}$ be the cost of an edge with endpoints of degree i and j. The 4×4 cost matrix $C_{\mathtt{AG}}$ associated with the arithmetic-geometric index of chemical graphs is

$$C_{\mathtt{AG}} = \begin{pmatrix} 1 & \frac{3}{2\sqrt{2}} & \frac{2}{\sqrt{3}} & \frac{5}{4} \\ \frac{3}{2\sqrt{2}} & 1 & \frac{5}{2\sqrt{6}} & \frac{3}{2\sqrt{2}} \\ \frac{2}{\sqrt{3}} & \frac{5}{2\sqrt{6}} & 1 & \frac{7}{4\sqrt{3}} \\ \frac{5}{4} & \frac{3}{2\sqrt{2}} & \frac{7}{4\sqrt{3}} & 1 \end{pmatrix} \approx \begin{pmatrix} 1.0000 & 1.0607 & 1.1547 & 1.2500 \\ 1.0607 & 1.0000 & 1.0206 & 1.0607 \\ 1.1547 & 1.0206 & 1.0000 & 1.0104 \\ 1.2500 & 1.0607 & 1.0104 & 1.0000 \end{pmatrix}.$$

For a chemical graph G, let $n_i(G)$ $(i = 0, \ldots, 4)$ be the number of vertices of degree i and let $x_{i,j}(G)$ $(1 \leq i \leq j \leq 4)$ be the number of edges with extremities of degrees i and j in G. Then, since $C_{\mathtt{AG}}$ is symmetric, we have

$$\mathtt{AG}(G) = \sum_{1 \leq i \leq j \leq 4} c_{i,j}\, x_{i,j}(G).$$

In what follows, we say that a chemical graph G is *extremal* if $\mathtt{AG}(G) \geq \mathtt{AG}(G')$ for all chemical graphs G' with the same order and the same size as G.

7.3 Preliminaries

We begin this section with the definition of a class of chemical graphs which, as we will see, contains most of the extremal chemical graphs of order n and size m.

Definition 1. $\mathcal{G}_{n,m}$ is the set of chemical graphs of order n and size m, and such that $n_0(G) = 0$, $n_2(G) + n_3(G) \leq 1$ and all edges have at least one endpoint of degree 4.

For example, using *Nauty geng* [162] or *PHOEG* [66] to enumerate all chemical graphs having order n and size m, it can be observed that there is only one graph $G_{n,m}$ in $\mathcal{G}_{n,m}$ for $(n,m) = (5,4)$, $(6,10)$, $(7,7)$, $(7,9)$, $(8,9)$, $(9,8)$, $(9,9)$ and $(11,11)$, and there are two graphs in $\mathcal{G}_{12,11}$, one connected and one non-connected (see also Table 7.3 at the end of the chapter). These graphs are shown in Figure 7.2 and they were chosen because they will appear in the proofs of the next sections.

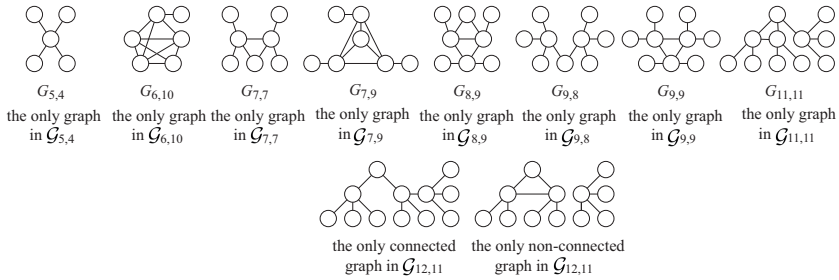


Figure 7.2: Examples of graphs in $\mathcal{G}_{n,m}$ for some pairs (n,m).

If all edges of a chemical graph G have at least one endpoint of degree 4, then $\mathtt{AG}(G) = c_{1,4}n_1(G) + 2c_{2,4}n_2(G) + 3c_{3,4}n_3(G) + (m - n_1(G) - 2n_2(G) - 3n_3(G))$ and since $2m = n_1(G) + 2n_2(G) + 3n_3(G) + 4n_4(G)$, we have

$$\begin{aligned}
\mathtt{AG}(G) &= c_{1,4}n_1(G) + 2c_{2,4}n_2(G) + 3c_{3,4}n_3(G) + 4n_4(G) - m \\
&= c_{1,4}n_1(G) + 2c_{2,4}n_2(G) + 3c_{3,4}n_3(G) + 4n_4(G) \\
&\quad -\tfrac{1}{2}(n_1(G) + 2n_2(G) + 3n_3(G) + 4n_4(G)) \\
&= \tfrac{3}{4}n_1(G) + (\tfrac{3}{\sqrt{2}} - 1)n_2(G) + (\tfrac{21}{4\sqrt{3}} - \tfrac{3}{2})n_3(G) + 2n_4(G).
\end{aligned}$$

For a pair (n,m) of integers, let $T_{n,m}$ be the set of quadruplets (t_1,t_2,t_3,t_4) of positive integers such that $\sum_{i=1}^{4} t_i = n$ and $\sum_{i=1}^{4} it_i = 2m$. Hence, we have $2m - n = t_2 + 2t_3 + 3t_4$. Note that $T_{1,0} = \emptyset$ since $\sum_{i=1}^{4} t_i \leq \sum_{i=1}^{4} it_i$. For $(t_1,t_2,t_3,t_4) \in T_{n,m}$, let $f(t_1,t_2,t_3,t_4)$ be defined as

$$f(t_1,t_2,t_3,t_4) = \tfrac{3}{4}t_1 + (\tfrac{3}{\sqrt{2}} - 1)t_2 + (\tfrac{21}{4\sqrt{3}} - \tfrac{3}{2})t_3 + 2t_4.$$

Clearly, if G is a chemical graph of order n and size m, with no isolated vertex and in which all edges have at least one endpoint of degree 4, then $(n_1(G), n_2(G), n_3(G), n_4(G))$ belongs to $T_{n,m}$ and we have observed above that

$$\mathtt{AG}(G) = f(n_1(G), n_2(G), n_3(G), n_4(G)).$$

Let (t_1, t_2, t_3, t_4) be a quadruplet in $T_{n,m}$ with $t_2 + t_3 \leq 1$:

- if $t_2 = 1$ then $t_3 = 0$, which means that $2m - n = 3t_4 + 1$ and $2m = t_1 + 2 + 4(n - t_1 - 1)$, or equivalently, $t_1 = \frac{4n-2m-2}{3}$;
- if $t_3 = 1$ then $t_2 = 0$, which means that $2m - n = 3t_4 + 2$ and $2m = t_1 + 3 + 4(n - t_1 - 1)$, or equivalently, $t_1 = \frac{4n-2m-1}{3}$;
- if $t_2 = t_3 = 0$, then $2m - n = 3t_4$ and $2m = t_1 + 4(n - t_1)$, or equivalently, $t_1 = \frac{4n-2m}{3}$.

Hence in all cases, we deduce the following property.

Property 29. *If (t_1, t_2, t_3, t_4) is a quadruplet in $T_{n,m}$ with $t_2 + t_3 \leq 1$, then*

- $t_1 = \lfloor \frac{4n-2m}{3} \rfloor$
- $t_2 = \begin{cases} 1 & \text{if } 2m - n \equiv 1 \bmod 3 \\ 0 & \text{otherwise} \end{cases}$
- $t_3 = \begin{cases} 1 & \text{if } 2m - n \equiv 2 \bmod 3 \\ 0 & \text{otherwise} \end{cases}$
- $t_4 = \lfloor \frac{2m-n}{3} \rfloor$.

Corollary 30. *There is at most one quadruplet (t_1, t_2, t_3, t_4) in $T_{n,m}$ with $t_2 + t_3 \leq 1$.*

Corollary 31. *If G is a graph in $\mathcal{G}_{n,m}$, then the quadruplet $(t_1, t_2, t_3, t_4) = (n_1(G), n_2(G), n_3(G), n_4(G))$ is the unique one in $T_{n,m}$ with $t_2 + t_3 \leq 1$.*

Proof. Let G be a graph in $\mathcal{G}_{n,m}$. Then $\sum_{i=1}^{4} n_i(G) = n$ and $\sum_{i=1}^{4} i n_i(G) = 2m$, which means that $(t_1, t_2, t_3, t_4) = (n_1(G), n_2(G), n_3(G), n_4(G))$ is a quadruplet in $T_{n,m}$ with $t_2 + t_3 \leq 1$. By Corollary 30, it is unique. ■

Some connected extremal chemical graphs have all edges with at least one endpoint of degree 4, but have $n_0(G) > 0$ or $n_2(G) + n_3(G) > 1$. Seven examples are shown in Figure 7.3 and we will prove that there are no other ones.

Also, some connected extremal chemical graphs have at least one edge with no endpoint of degree 4. Fifteen examples are shown in Figure 7.4 and we will prove that there are no other ones.

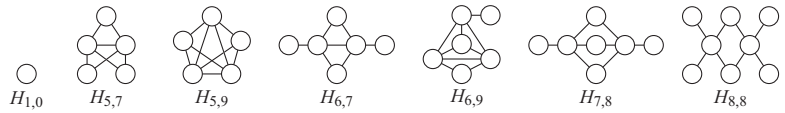


Figure 7.3: Seven connected extremal chemical graphs with all edges having at least one endpoint of degree 4 and with $n_0(G) > 0$ or $n_2(G) + n_3(G) > 1$.

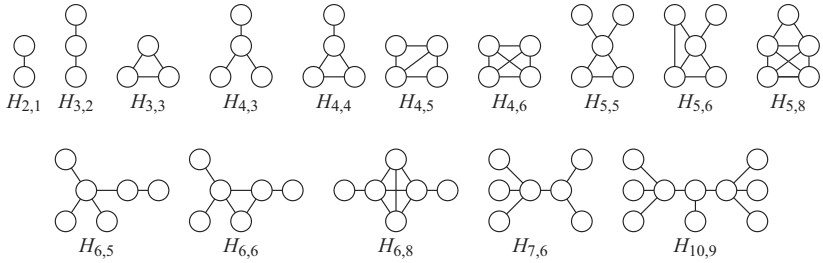


Figure 7.4: Fifteen connected extremal chemical graphs with at least one edge having no endpoint of degree 4.

For each pair (n, m) such that $H_{n,m}$ appears in Figure 7.3 or 7.4, we can enumerate all chemical graphs having order n and size m, using again *Nauty geng* [162] or *PHOEG* [66]. Table 7.1 gives the number of such graphs and it is therefore easy to verify that the following property holds.

Property 32. *The 22 graphs in Figures 7.3 and 7.4 are the only extremal graphs of their order and size.*

Table 7.1: Number N of chemical graphs for some orders n and sizes m.

n	1	2	3	3	4	4	4	4	5	5	5	5	5	6	6	6	6	6	7	7	8	10
m	0	1	2	3	3	4	5	6	5	6	7	8	9	5	6	7	8	9	6	8	8	9
N	1	1	1	1	3	2	1	1	6	6	4	2	1	14	20	22	20	15	38	82	188	883

The next property relates quadruplets in $T_{n,m}$ with connected graphs in $\mathcal{G}_{n,m}$. Note that a connected chemical graph of order n has m edges, with $n - 1 \leq m \leq \min\{2n, \frac{n(n-1)}{2}\}$.

Property 33. *Let n and m be two positive integers such that $n - 1 \leq m \leq \min\{2n, \frac{n(n-1)}{2}\}$ and (n, m) is not one of the 22 pairs for which there is a graph $H_{n,m}$ in Figure 7.3 or 7.4. If $(t_1, t_2, t_3, t_4) \in T_{n,m}$ and $t_2 + t_3 \leq 1$, then $\mathcal{G}_{n,m}$ contains at least one connected graph G with $n_i(G) = t_i$ $(i = 1, 2, 3, 4)$.*

Proof. Consider a pair (n, m) of positive integers such that $n - 1 \leq m \leq \min\{2n, \frac{n(n-1)}{2}\}$ and let (t_1, t_2, t_3, t_4) be any quadruplet in $T_{n,m}$. Note that

$n \geq 2$ since $T_{1,0} = \emptyset$. According to Property 29 and Corollary 31, a graph in $\mathcal{G}_{n,m}$ with $n_i(G) = t_i$ $(i = 1, 2, 3, 4)$ must have

- $n_1(G) = \lfloor \frac{4n-2m}{3} \rfloor$
- $n_2(G) = \begin{cases} 1 & \text{if } 2m - n \equiv 1 \bmod 3 \\ 0 & \text{otherwise} \end{cases}$
- $n_3(G) = \begin{cases} 1 & \text{if } 2m - n \equiv 2 \bmod 3 \\ 0 & \text{otherwise} \end{cases}$
- $n_4(G) = \lfloor \frac{2m-n}{3} \rfloor$.

Moreover, in order to impose that all edges in G have at least one endpoint of degree 4, we must have

- $x_{1,1}(G) = x_{1,2}(G) = x_{1,3}(G) = x_{2,2}(G) = x_{2,3}(G) = x_{3,3}(G) = 0$,
- $x_{1,4}(G){=}n_1(G)$, $x_{2,4}(G){=}2n_2(G)$, $x_{3,4}(G){=}3n_3(G)$, $x_{4,4}(G){=}m{-}n_1(G){-}2n_2(G){-}3n_3(G)$.

The following algorithm builds such a connected graph G, where V_i $(i = 1, \ldots, 4)$ is the set of vertices of degree i in G. It is illustrated in Figure 7.5.

1. Start from a graph of order n and size 0. Put $n_i(G)$ vertices in V_i, $i = 1, \ldots, 4$;
2. if $2m - n \equiv 1 \bmod 3$ then connect the vertex in V_2 to 2 vertices in V_4;
3. if $2m - n \equiv 2 \bmod 3$ then connect the vertex in V_3 to 3 vertices in V_4;
4. add $n_4(G){-}n_2(G){-}2n_3(G){-}1$ edges that link pairs of vertices in V_4 so that the graph induced by $V_2 \cup V_3 \cup V_4$ is a tree;
5. add $m - n_1(G) - n_2(G) - n_3(G) - n_4(G) + 1$ edges that link pairs of vertices in V_4 so that no vertex in V_4 is incident to more than 4 edges;
6. add edges linking each vertex of V_1 to a vertex of V_4 so that every vertex in V_4 has degree 4.

Steps 2 and 3 add the required number of edges with one endpoint of degree 4 and the other of degree 2 or 3, and Step 6 adds the required number of edges with one endpoint of degree 4 and the other of degree 1. Step 4 adds $n_4(G){-}n_2(G){-}2n_3(G){-}1$ edges linking pairs of vertices in V_4, while Step 5 adds $m - n_1(G) - n_2(G) - n_3(G) - n_4(G) + 1$ such edges. In total we will therefore have $m - n_1(G) - 2n_2(G) - 3n_3(G)$ edges linking pairs of vertices of V_4, which is the required number of edges with both endpoints of degree 4. It remains to prove that such a construction is always possible. For this purpose, the following constraints must be satisfied :

- $2n_2(G) \leq n_4(G)$ and $3n_3(G) \leq n_4(G)$ to ensure that Steps 2 and 3 can be performed;
- $m - n_1(G) - 2n_2(G) - 3n_3(G) \leq \frac{n_4(G)(n_4(G)-1)}{2}$ to avoid creating parallel edges in Steps 4 and 5;
- $n_1(G) \leq 4n_4(G)$, to ensure that Step 6 can be performed.

It is easy to check that these conditions are satisfied for all pairs (n, m) with $n \leq 13$ and $n - 1 \leq m \leq \min\{2n, \frac{n(n-1)}{2}\}$, except for the 22 pairs for which we have a graph $H_{n,m}$ in Figures 7.3 or 7.4. So assume $n \geq 14$. We then have

- $n_4(G) \geq \frac{2m-n-2}{3} \geq \frac{n-4}{3} > 3 \geq \max\{2n_2(G), 3n_3(G)\}$ (since $n_1(G) + n_2(G) \leq 1$).
- $$\begin{aligned} \frac{n_4(G)(n_4(G)-1)}{2} - x_{4,4}(G) &\geq \frac{1}{2}\left(\frac{2m-n-2}{3}\left(\frac{2m-n-2}{3}-1\right)\right) \\ &\quad - \left(m - \frac{4n-2m-2}{3}\right) \\ &= \frac{4m(m-n-11) + n^2 + 31n - 2}{18} \\ &\geq \frac{4(n-1)(n-1-n-11) + n^2 + 31n - 2}{18} \\ &= \frac{n^2 - 17n + 46}{18} > 0. \end{aligned}$$
- $4n_4(G) - n_1(G) \geq \frac{4(2m-n)}{3} - \frac{4n-2m+2}{3} = \frac{10m-8n-2}{3} \geq \frac{10(n-1)-8n-2}{3} = \frac{2n}{3} - 4 > 0$.

In summary, G has the right number of edges of each type and thanks to Step 4, it is connected. ■

The algorithm in the above proof is illustrated in Figure 7.5 for $n = m = 17$. In such a case, we have $n_1(G) = 11$, $n_2(G) = 0$, $n_3(G) = 1$, $n_4(G) = 5$ and $x_{4,4}(G) = 3$. In Step 1, we have represented the vertices of V_1 with the white color, while the vertex in V_3 is grey and the vertices in V_4 are black. Step 3 links the grey vertex to 3 black vertices. Step 4 adds 2 edges between black vertices. Step 5 adds the last edge between two black vertices and Step 6 adds the edges between the white and the black vertices. Notice that Step 4 was crucial to obtain a connected graph. Indeed another set of 3 edges linking black vertices could have produced a non-connected graph in $\mathcal{G}_{n,m}$ as illustrated at the bottom right of Figure 7.5.
The main objective of this chapter is to prove the following theorem.

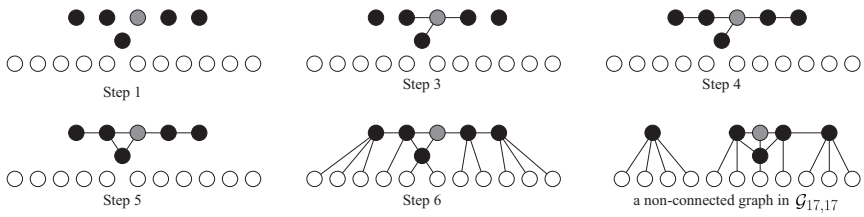


Figure 7.5: Illustration of the algorithm in the proof of Theorem 34

Theorem 34. *Let G be a connected chemical graph of order n and size m. If G is extremal, then either G is one of the 22 graphs $H_{n,m}$ of Figures 7.3 and 7.4, or G belongs to $\mathcal{G}_{n,m}$.*

To prove this theorem, we need some tools which are given in the next section.

7.4 Tools

Lemma 35. *Let G be a connected extremal chemical graph. Assume that G has a vertex u of degree 2 where v and w are its two neighbors.*

(a) If v and w are nonadjacent, then none of them has degree 3.

(b) If v and w are adjacent, $d_v \geq 3$ and $d_w \leq 3$, then no vertex nonadjacent to w has degree 2 or 3.

Proof.

(a) Assume that v and w are nonadjacent, and that one of them, say v, has degree 3. Let G' be the graph obtained from G by replacing uw with vw. Then, with $i = d_w$ and $j = d_x$, where x is any neighbor of v other than u, we have

$$\begin{aligned}\mathtt{AG}(G')-\mathtt{AG}(G) &\geq c_{1,4} - c_{2,3} + \min_{i=1,\ldots 4}(c_{4,i} - c_{2,i}) + 2\min_{j=1,\ldots 4}(c_{4,j}-c_{3,j}) \\ &\approx 0.1479 > 0.\end{aligned}$$

(b) Assume that v and w are adjacent with $d_v \geq 3$ and $d_w \leq 3$, and let x be a vertex nonadjacent to w such that $d_x = 2$ or 3. Let G' be the graph obtained from G by replacing uw with xw. Then, with $i = d_v$, $j = d_w$, $k = d_x$ and $\ell = d_y$, where y is any neighbor of x, we have

$$\begin{aligned}\mathtt{AG}(G')-\mathtt{AG}(G) \geq{}& \min_{i=3,4}(c_{1,i} - c_{2,i}) \\ &+ \min_{k=2,3}\left(\min_{j=2,3}(c_{j,k+1}-c_{j,2}) + k\min_{\ell=1,\ldots,4}(c_{\ell,k+1}-c_{\ell,k})\right) \\ \approx{}& 0.0128 > 0.\end{aligned}$$

In both cases, G' is connected and $\mathtt{AG}(G') > \mathtt{AG}(G)$, which means that G is not an extremal, connected chemical graph, a contradiction. ■

Lemma 36. *A connected extremal chemical graph does not contain a chain $v_1, v_2, \ldots, v_r$ as partial subgraph with v_1 nonadjacent to v_{r-1} and v_2 nonadjacent to v_r in G and with $d_{v_1} < d_{v_r}$, $d_{v_2} \leq 3$, and $d_{v_{r-1}} = 4$.*

Proof. Let G' be the graph obtained from G by replacing the edges v_1, v_2 and v_{r-1}, v_r by v_1, v_{r-1} and v_2, v_r. Then, with $d_{v_1} = i$, $d_{v_2} = j$ and $d_{v_r} = k$, we have

$$\mathtt{AG}(G') - \mathtt{AG}(G) \geq \min_{i=1,2,3} \min_{j=2,3} \min_{k=i+1,\ldots,4} (c_{i,4} + c_{j,k} - c_{i,j} - c_{4,k}) \approx 0.0207 > 0.$$

Since G' is connected and $\mathtt{AG}(G') > \mathtt{AG}(G)$, this means that G is not an extremal, connected chemical graph, a contradiction. ■

The next lemmas have a label (i, j) with $i < j$ to indicate that they state that a connected extremal chemical graph G has $x_{i,j}(G) = 0$, with a few exceptions.

Lemma (1,1). The only connected extremal chemical graph G with $x_{1,1}(G) > 0$ is $H_{2,1}$.

Proof. Let G be a connected extremal chemical graph with two adjacent vertices of degree 1. Since G is connected, it does not contain any other vertex, which means that $G \simeq H_{2,1}$. ■

Lemma (2,2). The only connected extremal chemical graphs G with $x_{2,2}(G)>0$ are $H_{3,3}$, $H_{4,4}$ and $H_{5,5}$.

Proof. Let u and v be two adjacent vertices of degree 2 in a connected extremal chemical graph G.

- Assume that u and v have a common neighbor w. If w has degree 2, then $G \simeq H_{3,3}$. So suppose w has degree at least 3. We know from Lemma 35(b) that all other vertices in the graph have degree 1 or 4. If they have all degree 1, then $G \simeq H_{4,4}$ (if w has degree 3) or $G \simeq H_{5,5}$ (if w has degree 4). So assume w is adjacent to a vertex x of degree 4. If x is adjacent to a vertex $y \neq w$ of degree 4, then Lemma 36 with the partial chain u, v, w, x, y contradicts the fact that G is a connected extremal chemical graph. Hence, all neighbors $y \neq w$ of x have degree 1. Similarly, if w has a second neighbor z of degree 4, then all neighbors of z, except w, have degree 1. There are therefore only three possible cases:

– if w has degree 3 then $\texttt{AG}(G) \approx 7.80 < 8.12 \approx \texttt{AG}(G_{7,7})$ (see Figure 7.2);

– if w has degree 4 and a neighbor of degree 1, then $\texttt{AG}(G) \approx 9.12 < 9.24 \approx \texttt{AG}(H_{8,8})$ (see Figure 7.3);

– if w has degree 4 and a second neighbor of degree 4, then $\texttt{AG}(G) \approx 12.62 < 12.78 \approx \texttt{AG}(G_{11,11})$ (see Figure 7.2);

In all cases G is not a connected extremal chemical graph, a contradiction.

- Assume that u and v have no common neighbor. Let $x \neq v$ (resp. $y \neq u$) be the second neighbor of u (resp. v). Let G' be the graph obtained from G by replacing xu with xv. Then, with $i = d_x$ and $j = d_y$, we have

$$\begin{aligned}\texttt{AG}(G') - \texttt{AG}(G) &\geq c_{1,3} - c_{2,2} + \min_{i=1,\ldots 4}(c_{3,i} - c_{2,i}) + \min_{j=1,\ldots 4}(c_{3,j} - c_{2,j}) \\ &\approx 0.0541 > 0.\end{aligned}$$

Hence, G is not extremal, a contradiction. ■

Lemma (1,2). The only connected extremal chemical graphs G with $x_{1,2}(G) > 0$ are $H_{3,2}$ and $H_{6,5}$.

Proof. Let G be a connected extremal chemical graph with two adjacent vertices u and v such that $d_u = 1$ and $d_v = 2$, and let w be the other neighbor of v. If w has degree 1 then $G \simeq H_{3,2}$. We know from Lemma (2,2) that w does not have degree 2, and from Lemma 35(a) that w does not have degree 3. Hence, $d_w = 4$. If w has three neighbors of degree 1, then $G \simeq H_{6,5}$; otherwise, w has a neighbor $x \neq v$ of degree at least 2 and Lemma 36 with the partial chain u, v, w, x contradicts the fact that G is a connected extremal chemical graph. ■

Lemma (3,3). The only connected extremal chemical graphs G with $x_{3,3}(G) > 0$ are $H_{4,5}$, $H_{4,6}$, $H_{5,8}$ and $H_{6,8}$.

Proof. Let G be a connected extremal chemical graph with two adjacent vertices u and v of degree 3. Let us first show that all vertices $x \neq u, v$ are either adjacent to both u and v, or to neither. If this is not the case then, we consider two cases.

- If u and v have a common neighbor w, then let $x \neq v$ be a vertex adjacent to u but not to v, and let $y \neq u$ be a vertex adjacent to v but not to u. Then G is not extremal. Indeed, let G' be the graph obtained from G by replacing xu with xv. Then, with $i = d_x$, $j = d_y$ and $k = d_w$, we have

$$\texttt{AG}(G') - \texttt{AG}(G) \quad \geq \quad c_{2,4} - c_{3,3} + \min_{i=1,2,3,4}(c_{4,i} - c_{3,i})$$

$$\begin{aligned} & + \min_{j=1,2,3,4}(c_{4,j}-c_{3,j}) + \min_{k=2,3,4}(c_{4,k}+c_{2,k}-2c_{3,k}) \\ \approx\ & 0.0593 > 0. \end{aligned}$$

- If u and v have no common neighbor, then there are two possible cases.
  - if all neighbors $x \neq u, v$ of u and v have degree 1, then G is not extremal since $\mathtt{AG}(G) \approx 5.62 < 5.87 \approx \mathtt{AG}(H_{6,5})$ (see Figure 7.4);
  - if at least one of u, v, say u, has a neighbor $x \neq v$ of degree at least 2, then G is not extremal. Indeed, let y be the other neighbor of u and let G' be the graph obtained from G by replacing yu with yv. Then, with $i = d_x$, $j = d_y$ and $k = d_z$, where z is any neighbor of v other than u, we have

$$\begin{aligned} \mathtt{AG}(G')-\mathtt{AG}(G) \ \geq\ & c_{2,4}-c_{3,3} + \min_{i=2,3,4}(c_{2,i}-c_{3,i}) \\ & + \min_{j=1,2,3,4}(c_{4,j}-c_{3,j}) + 2\min_{k=1,2,3,4}(c_{4,k}-c_{3,k}) \\ \approx\ & 0.0089 > 0. \end{aligned}$$

Hence, let x, y the the two common neighbors of u and v. We next show that at least one of x, y has degree 4, or G is $H_{4,5}$ or $H_{4,6}$.

- If x has degree 3, then x is adjacent to y since xu is an edge linking two vertices of degree 3 and we have seen that this implies that y cannot be adjacent to exactly one of u, x. Hence, either $G \simeq H_{4,6}$, or y has degree 4.
- If x has degree 2, then we know from the previous case that y is of degree 2 or 4. Hence, either $G \simeq H_{4,5}$, or y has degree 4.

So, without loss of generality, assume $d_x = 4$. Let $W = V \setminus \{u, v, x, y\}$ where V is the vertex set of G. Assume that W contains at least one vertex w in W of degree at least 3.

- If $d_w = 4$, then let z be a vertex adjacent to w but not to x and let G' be the graph obtained from G by replacing the edges ux and wz by uw and xz. Clearly, $\mathtt{AG}(G') = \mathtt{AG}(G)$, which means that G' is also a connected extremal chemical graph. But u and v are two adjacent vertices in G', and they are both of degree 3, while x is adjacent to u but not to v. We have shown above that this is impossible.
- If $d_w = 3$, then w is adjacent to x. Indeed, if this is not the case, then let z be any vertex in W adjacent to w and let G' be the graph obtained from G by replacing the edges ux and wz by uz and xw. Clearly, $\mathtt{AG}(G') = \mathtt{AG}(G)$ and u, v are two adjacent vertices of degree 3 in G' with

z adjacent to u but not to v, a contradiction. Moreover, all neighbors of w in W are adjacent to x. Indeed, assume that a vertex $z \in W$ is adjacent to w but not to x. Then $d_z \leq 2$ (since vertices of degree 3 in W are adjacent to x and no vertex in W has degree 4), and Lemma 36 with the partial chain z, w, x, u shows that G is not extremal, a contradiction. In summary, we have shown that w can have only one neighbor in W (else x would be of degree at least 5), which means that w is adjacent to y. We know from Lemma 35(b) that the neighbor z of w in W cannot be of degree 2 (since u has degree 3 and is not adjacent to w). Hence, $d_z = 3$, which means that w, x and y are its three neighbors (as z, x, y are the 3 neighbors of w). Hence, $\texttt{AG}(G) \approx 10.08 < 10.28 \approx \texttt{AG}(G_{6,10})$ (see Figure 7.2), a contradiction.

Hence, all vertices in W have degree 1 or 2. At least one vertex in W has degree 2, else

- if $d_y = 3$, then x is adjacent to y since uy is an edge linking two vertices of degree 3 and we have seen that this implies that x cannot be adjacent to exactly one of u, y. Hence, $\texttt{AG}(G) \approx 7.28 < 7.36 \approx \texttt{AG}(H_{5,7})$ (see Figure 7.3);
- if $d_y = 4$, then either $G \simeq H_{6,8}$, or $\texttt{AG}(G) \approx 10.04 < 10.12 \approx \texttt{AG}(G_{8,9})$ (see Figure 7.2).

So let z be a vertex of degree 2 in W. We know from Lemmas (2,2) and (1,2) that z is adjacent to x and y, which implies that y has degree 4, else u and y are two adjacent vertices of degree 3 and z is adjacent to y but not to u, which is impossible. There are therefore three possible cases:

- if W has two vertices of degree 2, then $\texttt{AG}(G) \approx 9.28 < 9.40 \approx \texttt{AG}(H_{6,9})$ (see Figure 7.3);
- if W has one vertex of degree 2 and x is not adjacent to y, then $\texttt{AG}(G) \approx 9.66 < 9.78 \approx \texttt{AG}(G_{7,9})$ (see Figure 7.2);
- if W has one vertex of degree 2 and x is adjacent to y, then $G \simeq H_{5,8}$. ■

Lemma (1,3). The only connected extremal chemical graphs G with $x_{1,3}(G) > 0$ are $H_{4,3}$, $H_{4,4}$, $H_{6,6}$, $H_{7,6}$ and $H_{10,9}$.

Proof. Let G be a connected extremal chemical graph with two adjacent vertices u and v such that $d_u = 1$ and $d_v = 3$. If v has a neighbor w of degree 2, we know from Lemma 35(a) that the second neighbor x of w is adjacent to v. Then, either $G \simeq H_{4,4}$ or it follows from Lemmas (2,2) and (3,3) that $d_x = 4$. If x has two neighbors of degree 1, then $G \simeq H_{6,6}$, else x has a neighbor $y \neq v, w$

such that $d_y \geq 2$, and Lemma 36 with the partial chain u, v, x, y shows that G is not extremal, a contradiction.

So assuming that G is not $H_{4,4}$ or $H_{6,6}$, we know that no neighbor of v has degree 2. It then follows from Lemma (3,3) that they all have degree 1 or 4. If v has a neighbor x of degree 4, then all neighbors $y \neq v$ of x that are also not adjacent to v have degree 1, else Lemma 36 with the partial chain u, v, x, y shows that G is not extremal. Hence there are only 4 cases:

- if all neighbors of v have degree 1, then $G \simeq H_{4,3}$;
- if v has only one neighbor of degree 4, then $G \simeq H_{7,6}$;
- if v has two non-adjacent neighbors of degree 4, then $G \simeq H_{10,9}$;
- if v has two adjacent neighbors of degree 4, then $\texttt{AG}(G) \approx 9.18 < 9.24 \approx \texttt{AG}(H_{8,8})$ (see Figure 7.3). ■

Lemma (2,3). The only connected extremal chemical graphs G with $x_{2,3}(G) > 0$ are $H_{4,4}$, $H_{4,5}$, $H_{5,6}$ and $H_{6,6}$.

Proof. Let G be a connected extremal chemical graph with two adjacent vertices u and v such that $d_u = 2$ and $d_v = 3$. We know from Lemma 35(a) that the second neighbor w of u is adjacent to v. If G is not equal to $H_{4,4}$ or $H_{4,5}$, it follows from Lemmas (2,2) and (3,3) that w has degree 4. Also, it follows from Lemmas (1,3) and (3,3) that if G is not $H_{6,6}$, then the third neighbor $x \neq u, w$ of v has degree 2 or 4.

- If $d_x = 2$, then x is adjacent to w since, by Lemma 35(a), the second neighbor of x must be adjacent to v. It follows from Lemma 35(b) that all vertices other than u, v, w, x have degree 1 or 4. Hence, the fourth neighbor $y \neq u, v, x$ of w has degree 1 or 4. If $d_y = 1$ then $G \simeq H_{5,6}$. If $d_y = 4$, then there are two cases:
  - if y has 3 neighbors of degree 1, then $\texttt{AG}(G) \approx 9.92 < 10.12 \approx \texttt{AG}(G_{8,9})$ (see Figure 7.2);
  - if y has a neighbor $z \neq w$ of degree 4, then Lemma 36 with the partial chain u, v, w, y, z shows that G is not extremal, a contradiction.
- If $d_x = 4$, then let $y \neq v, w$ be a neighbor of x. It follows from Lemmas 35(b) and 36 with the partial chain u, v, x, y that $d_y = 1$. Let G' be the graph obtained from G by replacing the edges uw and xy by ux and wy. Clearly, $\texttt{AG}(G') = \texttt{AG}(G)$, which means that G' is also a connected extremal chemical graph and since w is now the neighbor of v of degree 4 that is not adjacent to u, this means that all neighbors of w different from v, x have degree 1. Hence,

– if w is adjacent to x then $\texttt{AG}(G) \approx 8.85 < 8.86 \approx \texttt{AG}(H_{7,8})$ (see Figure 7.3);

– if w is not adjacent to x then $\texttt{AG}(G) \approx 10.35 < 10.5 \approx \texttt{AG}(G_{9,9})$ (see Figure 7.2). ■

7.5 Characterization of extremal chemical graphs

In this section, we characterize extremal chemical graphs of order n and size $m \geq n-1$. We first consider the connected extremal chemical graphs, and then the non-connected ones. We conclude the section with a property of extremal chemical graphs or order n and size $m \leq n-2$.

We start with the proof of Theorem 34 that states that a connected extremal chemical graph of order n and size m necessarily belongs to $\mathcal{G}_{n,m}$, except for 22 pairs (n, m).

Proof of Theorem 34. Observe first that if (t_1, t_2, t_3, t_4) is a quadruplet in $T_{n,m}$ with $t_2 + t_3 > 1$. Then there is $(s_1, s_2, s_3, s_4) \in T_{n,m}$ such that $s_2 + s_3 < t_1 + t_2$ and $f(s_1, s_2, s_3, s_4) > f(t_1, t_2, t_3, t_4)$. Indeed:

- If $t_2 \geq 2$ then set $s_1 = t_1 + 1$, $s_2 = t_2 - 2$, $s_3 = t_3 + 1$ and $s_4 = t_4$. We have, $\sum_{i=1}^{4} s_i = \sum_{i=1}^{4} t_i = n$ and $\sum_{i=1}^{4} i s_i = \sum_{i=1}^{4} i t_i = 2m$, which means that $(s_1, s_2, s_3, s_4) \in T_{n,m}$. Moreover,

$$f(s_1, s_2, s_3, s_4) - f(t_1, t_2, t_3, t_4) = c_{1,4} - 4c_{2,4} + 3c_{3,4} \approx 0.0384 > 0.$$

- if $t_3 \geq 2$ then set $s_1 = t_1$, $s_2 = t_2 + 1$, $s_3 = t_3 - 2$ and $s_4 = t_4 + 1$. We have $\sum_{i=1}^{4} s_i = \sum_{i=1}^{4} t_i = n$ and $\sum_{i=1}^{4} i s_i = \sum_{i=1}^{4} i t_i = 2m$, which means that $(s_1, s_2, s_3, s_4) \in T_{n,m}$. Moreover,

$$f(s_1, s_2, s_3, s_4) - f(t_1, t_2, t_3, t_4) = 2c_{2,4} - 6c_{3,4} + 4 \approx 0.0591 > 0.$$

- if $t_2 \geq 1$ and $t_3 \geq 1$, then set $s_1 = t_1 + 1$, $s_2 = t_2 - 1$, $s_3 = t_3 - 1$ and $s_4 = t_4 + 1$. Hence, $\sum_{i=1}^{4} s_i = \sum_{i=1}^{4} t_i = n$ and $\sum_{i=1}^{4} i s_i = \sum_{i=1}^{4} i t_i = 2m$, which means that $(s_1, s_2, s_3, s_4) \in T_{n,m}$. Moreover,

$$f(s_1, s_2, s_3, s_4) - f(t_1, t_2, t_3, t_4) = c_{1,4} - 2c_{2,4} - 3c_{3,4} + 4 \approx 0.0975 > 0.$$

Note that if $s_2 + s_3 > 1$, then we can repeat the same reasoning. We can therefore conclude that if (t_1, t_2, t_3, t_4) is a quadruplet in $T_{n,m}$ with $t_2 + t_3 > 1$, then there is $(s_1, s_2, s_3, s_4) \in T_{n,m}$ such that $s_2 + s_3 \leq 1$ and $f(s_1, s_2, s_3, s_4) > f(t_1, t_2, t_3, t_4)$.

So let G be a connected extremal chemical graph of order n and size m, and suppose that G is not one of the 22 graphs of Figures 7.3 and 7.4. It follows from the lemmas of the previous section that all edges in G have at least one endpoint of degree 4. Since $n_0(G) = 0$ (else $G \simeq H_{1,0}$), we have $\mathtt{AG}(G) = f(n_1(G), n_2(G), n_3(G), n_4(G))$. We have shown above that if $n_2(G)+n_3(G) > 1$, then there is a quadruplet (s_1, s_2, s_3, s_4) in $T_{n,m}$ such that $s_2 + s_3 \leq 1$ and $f(s_1, s_2, s_3, s_4) > f(n_1(G), n_2(G), n_3(G), n_4(G))$. Hence, if $n_2(G) + n_3(G) > 1$, then it follows from Property 33 that there is a connected chemical graph G' in $\mathcal{G}_{n,m}$ with $\mathtt{AG}(G') = f(s_1, s_2, s_3, s_4) > f(n_1(G), n_2(G), n_3(G), n_4(G)) = \mathtt{AG}(G)$, a contradiction. We can therefore conclude that $n_2(G) + n_3(G) \leq 1$, which implies that G belongs to $\mathcal{G}_{n,m}$. ■

It follows from Theorem 34 and Corollary 31 that if $1 \leq n - 1 \leq m$ and (n, m) is not a pair for which there is a graph $H_{n,m}$ in Figure 7.3 or 7.4 and if there exists a connected extremal graph of order n and size m, then all graphs in $\mathcal{G}_{n,m}$ are extremal and their arithmetic-geometric index is easy to compute since we know the number of edges with endpoints of degree i and j for all $1 \leq i \leq j \leq 4$. We can therefore state the following corollary.

Corollary 37. *Let G be a connected extremal chemical graph. If G is not one of the 22 graphs $H_{n,m}$ in Figure 7.3, then $\mathtt{AG}(G) = UB_{n,m}$, where*

$$UB_{n,m} = \frac{2n+5m}{6} + \begin{cases} 0 & \text{if } 2m-n \equiv 0 \bmod 3, \\ \frac{3}{\sqrt{2}} - \frac{13}{6} & \text{if } 2m-n \equiv 1 \bmod 3, \\ \frac{21}{4\sqrt{3}} - \frac{37}{12} & \text{if } 2m-n \equiv 2 \bmod 3. \end{cases}$$

Proof. Theorem 34 shows that G belongs to $\mathcal{G}_{n,m}$. Let us compute $\mathtt{AG}(G)$.

- If $2m - n \equiv 0 \bmod 3$, then $4n - 2m \equiv 0 \bmod 3$, which means that $n_1(G)=\frac{4n-2m}{3}$, $n_2(G)=n_3(G)=0$ and $n_4(G)=\frac{2m-n}{3}$. Hence,
$$\mathtt{AG}(G) = \tfrac{3}{4}\tfrac{4n-2m}{3} + 2\tfrac{2m-n}{3} = \tfrac{2n+5m}{6}.$$
- If $2m - n \equiv 1 \bmod 3$, then $4n - 2m \equiv 2 \bmod 3$, which means that $n_1(G)=\frac{4n-2m-2}{3}$, $n_2(G)=1, n_3(G)=0$ and $n_4(G)=\frac{2m-n-1}{3}$. Hence,
$$\mathtt{AG}(G) = \tfrac{3}{4}\tfrac{4n-2m-2}{3} + (\tfrac{3}{\sqrt{2}} - 1) + 2\tfrac{2m-n-1}{3} = \tfrac{2n+5m-13}{6} + \tfrac{3}{\sqrt{2}}.$$
- If $2m-n \equiv 2 \bmod 3$, then $4n-2m \equiv 1 \bmod 3$, which means that $n_1(G)=\frac{4n-2m-1}{3}$, $n_2(G)=0, n_3(G)=1$ and $n_4(G)=\frac{2m-n-2}{3}$. Hence,
$$\mathtt{AG}(G) = \tfrac{3}{4}\tfrac{4n-2m-1}{3} + (\tfrac{21}{4\sqrt{3}} - \tfrac{3}{2}) + 2\tfrac{2m-n-2}{3} = \tfrac{2n+5m}{6} + \tfrac{21}{4\sqrt{3}} - \tfrac{37}{12}.$$

■

As shown in Table 7.2, if (n, m) is one of the pairs for which there is a graph $H_{n,m}$ in Figure 7.3 or 7.4, then $\mathtt{AG}(H_{n,m}) < UB_{n,m}$. Hence, the connected graphs in $\mathcal{G}_{n,m}$ are the only connected chemical graphs G of order n and size m with $\mathtt{AG}(G) = UB_{n,m}$. The sharp upper bound $\mathtt{AG}(H_{n,m})$ for the 22 pairs (n, m) that are exceptions is slightly smaller than $UB_{n,m}$. We give in Table 7.2 the values of this upper bound as well as the differences between $UB_{n,m}$ and $\mathtt{AG}(H_{n,m})$. We observe that the largest difference is $\frac{1}{2}$ while the smallest is approximately equal to 0.0384.

Table 7.2: Sharp upper bound $\mathtt{AG}(H_{n,m})$ on the arithmetic-geometric index of graphs of order n and size m, for the 22 pairs (n, m) not included in Corollary 37, and difference with $UB_{n,m}$.

n	m	$\mathtt{AG}(H_{n,m})$	$UB_{n,m} - \mathtt{AG}(H_{n,m})$	
1	0	0	$-\frac{11}{4} + \frac{21}{4\sqrt{3}}$	≈ 0.2811
2	1	1	$\frac{1}{2}$	≈ 0.5000
3	2	$\frac{3}{\sqrt{2}}$	$\frac{1}{2}$	≈ 0.5000
3	3	3	$\frac{1}{2}$	≈ 0.5000
4	3	$\frac{6}{\sqrt{3}}$	$\frac{3}{4} - \frac{3}{4\sqrt{3}}$	≈ 0.3170
4	4	$1 + \frac{2}{\sqrt{3}} + \frac{5}{\sqrt{6}}$	$\frac{3}{2} + \frac{3}{\sqrt{2}} - \frac{2}{\sqrt{3}} - \frac{5}{\sqrt{6}}$	≈ 0.4254
4	5	$1 + \frac{10}{\sqrt{6}}$	$\frac{9}{2} - \frac{10}{\sqrt{6}}$	≈ 0.4175
4	6	6	$-\frac{11}{4} + \frac{21}{4\sqrt{3}}$	≈ 0.2811
5	5	$\frac{7}{2} + \frac{3}{\sqrt{2}}$	$-\frac{3}{4} - \frac{3}{\sqrt{2}} + \frac{21}{4\sqrt{3}}$	≈ 0.1598
5	6	$\frac{5}{4} + \frac{3}{\sqrt{2}} + \frac{7}{4\sqrt{3}} + \frac{5}{\sqrt{6}}$	$\frac{13}{4} - \frac{7}{4\sqrt{3}} - \frac{5}{\sqrt{6}}$	≈ 0.1984
5	7	$1 + \frac{9}{\sqrt{2}}$	$\frac{13}{2} - \frac{9}{\sqrt{2}}$	≈ 0.1360
5	8	$2 + \frac{3}{\sqrt{2}} + \frac{7}{\sqrt{3}}$	$\frac{13}{4} - \frac{3}{\sqrt{2}} - \frac{7}{4\sqrt{3}}$	≈ 0.1183
5	9	$3 + \frac{21}{2\sqrt{3}}$	$4 + \frac{3}{\sqrt{2}} - \frac{21}{2\sqrt{3}}$	≈ 0.0591
6	5	$\frac{15}{4} + \frac{3}{\sqrt{2}}$	$\frac{1}{4}$	≈ 0.2500
6	6	$\frac{5}{2} + \frac{3}{2\sqrt{2}} + \frac{15}{4\sqrt{3}} + \frac{5}{2\sqrt{6}}$	$\frac{9}{2} - \frac{3}{2\sqrt{2}} - \frac{15}{4\sqrt{3}} - \frac{5}{2\sqrt{6}}$	≈ 0.2537
6	7	$\frac{7}{2} + \frac{6}{\sqrt{2}}$	$\frac{5}{4} - \frac{6}{\sqrt{2}} + \frac{21}{4\sqrt{3}}$	≈ 0.0384
6	8	$\frac{9}{2} + \frac{7}{\sqrt{3}}$	$2 + \frac{3}{\sqrt{2}} - \frac{7}{\sqrt{3}}$	≈ 0.0799
6	9	$\frac{17}{4} + \frac{3}{\sqrt{2}} + \frac{21}{4\sqrt{3}}$	$\frac{21}{4} - \frac{3}{\sqrt{2}} - \frac{21}{4\sqrt{3}}$	≈ 0.0976
7	6	$\frac{15}{4} + \frac{4}{\sqrt{3}} + \frac{7}{4\sqrt{3}}$	$\frac{1}{2} - \frac{1}{2\sqrt{3}}$	≈ 0.2113
7	8	$\frac{5}{2} + \frac{9}{\sqrt{2}}$	$\frac{13}{2} - \frac{9}{\sqrt{2}}$	≈ 0.1360
8	8	$5 + \frac{6}{\sqrt{2}}$	$\frac{5}{4} - \frac{6}{\sqrt{2}} + \frac{21}{4\sqrt{3}}$	≈ 0.0384
10	9	$\frac{15}{2} + \frac{11}{2\sqrt{3}}$	$\frac{1}{4} - \frac{1}{4\sqrt{3}}$	≈ 0.1057

When $m = n - 1$, Corollary 37 gives an upper bound for chemical trees. More precisely, if T is a chemical tree of order n, then

$$\mathtt{AG}(T) \leq UB_{n,n-1} = \frac{7n-5}{6} + \begin{cases} 0 & \text{if } n \equiv 2 \bmod 3, \\ \frac{3}{\sqrt{2}} - \frac{13}{6} & \text{if } n \equiv 0 \bmod 3, \\ \frac{21}{4\sqrt{3}} - \frac{37}{12} & \text{if } n \equiv 1 \bmod 3. \end{cases}$$

and this bound is reached for all n, except for $n = 1, 2, 3, 4, 6, 7, 10$ since $H_{1,0}, H_{2,1}, H_{3,2}, H_{4,3}, H_{6,5}, H_{7,6}$ and $H_{10,9}$ appear in Figure 7.4. The same upper bound is given in [209], but the authors did not mention the 7 exceptions. For example, they state that when $n \equiv 1 \bmod 3$, there are $\frac{n-1}{3} - 1$ vertices of degree 4, and one vertex of degree 3 that must be adjacent to vertices of degree 4. This is clearly impossible for $n = 1, 4, 7$ and 10.

We now show that if we remove the constraint that extremal chemical graphs must be connected, then no better value of $\mathtt{AG}$ can be obtained.

Theorem 38. *Let G be a non-connected chemical graph of order n and size $m \geq n - 1$. If G is extremal, then G belongs to $\mathcal{G}_{n,m}$.*

Proof. Assume that the theorem is not valid and let G be a non-connected extremal chemical graph of order n and size $m \geq n-1$ that is a counterexample with the smallest number of connected components. It follows from Property 32 that (n, m) is not one of the 22 pairs for which there is a graph $H_{n,m}$ in Figure 7.3 or 7.4. Let $G_1, \ldots, G_k$ $(k \geq 2)$ be the connected components of G, and let N_i and M_i be the order and the size of G_i, respectively. Clearly, $\mathtt{AG}(G) = \sum_{i=1}^{k} \mathtt{AG}(G_i)$. Hence, since G is extremal, every G_i is a connected extremal graph of order N_i and size M_i. At least one G_i, say G_1, contains a cycle C. If C contains an edge xy with $d_x = 4$ and $d_y \geq 3$ then:

- if G_2 contains only one vertex z then let G' be the graph obtained from G by replacing the edge xy by the edge xz. Since y belongs to a cycle, at least one of its neighbors $z \neq x$ has degree at least 2. Hence, with $i = d_y$, $j = d_z$ and $k = d_u$, where u is any neighbor of y other than x and w, we have

$$\begin{aligned} \mathtt{AG}(G') - \mathtt{AG}(G) &\geq c_{1,4} + \min_{i=3,4}\Big(\min_{j=2,3,4}(c_{i-1,j} - c_{i,j}) - c_{4,i}\Big) \\ &\quad + \min_{i=3,4}\Big((i-2)\min_{k=1,2,3,4}(c_{i-1,k} - c_{i,k})\Big) \\ &\approx 0.0193 > 0. \end{aligned}$$

  Hence G is not extremal, a contradiction.

- if G_2 contains at least two vertices, then consider any edge zw in G_2 and assume without loss of generality that $d_z \leq d_w$. Let G' be the graph obtained from G by replacing the edges xy and zw by the edges xz and yw. Then, with $i = d_y$, $j = d_z$ and $k = d_w$, we have

$$\mathtt{AG}(G')-\mathtt{AG}(G) \geq \min_{i=3,4} \min_{j=1,2,3,4} \min_{k=j,\ldots,4} (c_{4,j} + c_{i,k} - c_{4,i} - c_{j,k}) = 0.$$

  Moreover, all cases where $\mathtt{AG}(G') = \mathtt{AG}(G)$ have $d_x = d_w$ or/and $d_y = d_z$. Hence, G' has a smaller number of connected components than G, while $n_i(G) = n_i(G')$ for $0 \leq i \leq 4$ and $x_{i,j}(G) = x_{i,j}(G')$ for $1 \leq i \leq j \leq 4$. It follows that G' is also extremal, and either both of G and G' belong to $\mathcal{G}_{n,m}$, or none of them. If G' is connected, then we know from Theorem 34 that G' (and hence also G) belongs to $\mathcal{G}_{n,m}$, which means that G is not a counterexample to the theorem, a contradiction. If G' is not connected, then G is not a counterexample to the theorem with the smallest number of connected components, a contradiction.

Note that if G_1 belongs to $\mathcal{G}_{N_1,M_1}$, then the cycle C contains two adjacent vertices of degree 4 (since vertices of degree 1 do not belong to a cycle and there is at most one vertex of degree 2 or 3 in G_1). Also, the 8 graphs $H_{5,6}$, $H_{5,7}$, $H_{5,8}$, $H_{5,9}$, $H_{6,6}$, $H_{6,7}$, $H_{6,8}$, $H_{6,9}$, in Figures 7.3 and 7.4 which have a cycle and an edge linking a vertex of degree 4 to a vertex of degree at least 3, have such an edge in a cycle. Hence, $x_{3,4}(G_i) + x_{4,4}(G_i) = 0$ for all connected components G_i of G with a cycle.

Suppose now that G_2 contains an edge xy with $d_x = 4$ and $d_y \geq 3$. Consider any edge zw on C and assume without loss of generality that $d_z \leq d_w$. Let G' be the graph obtained from G by replacing the edges xy and zw by the edges xz and yw. Then, with $i = d_y$, $j = d_z$ and $k = d_w$, we have

$$\mathtt{AG}(G')-\mathtt{AG}(G) \geq \min_{i=3,4} \min_{j=2,3,4} \min_{k=j,\ldots,4} (c_{4,j} + c_{i,k} - c_{4,i} - c_{j,k}) = 0.$$

As above, the only cases where $\mathtt{AG}(G') = \mathtt{AG}(G)$ have $d_x = d_w$ or/and $d_y = d_z$. Hence, either G is not a counterexample to the theorem, or it is not a counterexample with the smallest number of connected components, a contradiction.

Hence, we know that $x_{3,4}(G)+x_{4,4}(G) = 0$. We now prove that no connected component of G can have more than 9 vertices. So assume G has a connected component H of order $N \geq 10$ and size M. We know from Theorem 34 that there are two possible cases:

- if H is one of the 22 graphs in Figures 7.3 and 7.4, then $H \simeq H_{10,9}$, which implies $x_{3,4}(G) \geq x_{3,4}(H) = 2$, a contradiction.

- if H belongs to $\mathcal{G}_{N,M}$, then $n_3(H) = 0$ else $x_{3,4}(G) \geq x_{3,4}(H) = 3$, Hence,

$$x_{4,4}(H) \geq m - n_1(H) - 2 \geq \left\lceil m - \frac{4n - 2m}{3} - 2 \right\rceil = \left\lceil \frac{5m - 4n - 6}{3} \right\rceil.$$

  The four cases here below show that $x_{3,4}(G) + x_{4,4}(G) \geq x_{3,4}(H) + x_{4,4}(H) \geq 1$, a contradiction.

  - If $M = N - 1 = 9$ then $x_{3,4}(H) = 2$.
  - If $M = N - 1 = 10$, then $x_{4,4}(H) = 2$.
  - If $M = N - 1 \geq 11$, then $x_{4,4}(H) \geq \lceil \frac{N-11}{3} \rceil \geq 1$.
  - If $M \geq N \geq 10$ then $x_{4,4}(H) \geq \frac{N-6}{3} > 1$.

We thus know that $3 \leq N_1 \leq 9$, $N_1 \leq M_1 \leq \min\{2N_1, \frac{N_1(N_1-1)}{2}\}$ and G_1 has no edge linking a vertex of degree 4 to a vertex of degree at least 3. It is easy to check that there are exactly 7 such graphs, namely, $H_{3,3}$, $H_{4,4}$, $H_{4,5}$, $H_{4,6}$, $H_{5,5}$, $H_{7,8}$ and $H_{8,8}$ (see Figures 7.3 and 7.4). Also, $1 \leq N_2 \leq 9$, $N_2 - 1 \leq M_2 \leq \min\{2N_2, \frac{N_2(N_2-1)}{2}\}$ and G_2 has no edge linking a vertex of degree 4 to a vertex of degree at least 3. There are only 14 such graphs, namely, the 7 graphs mentioned above, and $H_{1,0}$, $H_{2,1}$, $H_{3,2}$, $H_{4,3}$, $G_{5,4}$, $H_{6,5}$ and $G_{9,8}$ (see Figures 7.2, 7.3 and 7.4).

Let $i_{n,m}$ be equal to $\mathtt{AG}(G)$, where G is any connected extremal chemical graph of order n and size m. It is easy to check by enumeration that $i_{N_1,M_1} + i_{N_2,M_2} < i_{N_1+N_2,M_1+M_2}$ for the 7 pairs (N_1, M_1) and the 14 pairs (N_2, M_2). Hence by removing G_1 and G_2 and replacing these two connected components of G by a connected extremal chemical graph of order N_1+N_2 and size M_1+M_2, one gets a graph G' with $\mathtt{AG}(G) < \mathtt{AG}(G')$, which means that G is not extremal, a contradiction. ■

Corollary 31 shows that all graphs in $\mathcal{G}_{n,m}$ have the same AG value, which means that they are all extremal if (n, m) is not a pair appearing in Figures 7.3 or 7.4. Hence, putting together Property 32 and Theorems 34 and 38, we get the following characterization of extremal chemical graphs.

Theorem (Characterization of extremal chemical graphs)**.** A chemical graph G of order n and size $m \geq n - 1$ is extremal if and only if G is one of the 22 graphs in Figures 7.3 and 7.4 or G belongs to $\mathcal{G}_{n,m}$.

We indicate in Table 7.3 the number of connected and non-connected extremal chemical graphs of order n and size m for $1 \leq n \leq 14$ and $n - 1 \leq m \leq \min\{2n, \frac{n(n-1)}{2}\}$. For example, for $n = 12$ and $m = 11$, we see that there

are exactly one connected and one non-connected extremal chemical graph and these two graphs are represented at the bottom of Figure 7.2. The 22 pairs (n, m) for which $H_{n,m}$ is the only extremal graph are shown in gray boxes. We observe that the number of connected extremal chemical graphs grows exponentially, but not in a monotonic way.

The proof of Theorem 38 shows that if a non-connected chemical graph G contains a cycle, then there is a chemical graph G' having fewer connected components than G and such that $\mathtt{AG}(G) \leq \mathtt{AG}(G')$. This leads to the following corollary.

Corollary 39. *For all n and m with $0 \leq m \leq n-2$ there is a chemical forest G^* which is a disjoint union of extremal chemical trees and such that $\mathtt{AG}(G) \leq \mathtt{AG}(G^*)$ for all chemical graphs G of order n and size m.*

Table 7.3: Number of connected and non-connected extremal chemical graphs of order n and size m for $1 \leq n \leq 14$ and $n-1 \leq m \leq \min\{2n, \frac{n(n-1)}{2}\}$

m \ n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	1 0													
1		1 0												
2			1 0											
3			1 0	1 0										
4				1 0	1 0									
5				1 0	1 0	1 0								
6				1 0	1 0	1 0	1 0							
7					1 0	1 0	1 0	1 0						
8					1 0	1 0	1 0	1 0	1 0					
9					1 0	1 0	1 0	1 0	1 0	1 0				
10					1 0	1 0	1 0	1 0	1 0	2 0	1 0			
11						1 0	1 0	2 0	3 0	1 0	1 0	1 1		
12						1 0	2 0	4 0	2 0	4 0	6 0	2 0	1 0	
13							2 0	3 0	10 0	12 0	4 0	5 1	7 1	2 1
14							2 0	8 0	17 0	8 1	21 1	23 1	5 1	3 1
15								7 0	9 0	47 0	58 1	14 1	27 2	27 3
16								6 0	37 0	77 0	31 1	113 2	111 4	18 2
17									28 0	35 0	249 0	303 3	59 4	159 11
18									16 0	198 0	399 0	134 2	684 8	625 20
19										126 0	154 0	1550 1	1786 9	298 11
20										59 1	1246 1	2395 1	707 7	4620 40
21											719 1	845 1	10801 4	11855 36
22											265 1	8789 3	16433 6	4399 20
23												4721 3	5440 4	83399 19
24												1544 3	68804 12	125829 28
25													35678 11	40399 14
26													10778 8	590342 55
27														300361 45
28														88168 25

7.6 Conclusion

We have determined a sharp upper bound on the value of the arithmetic-geometric index of chemical graphs of order n and size $m \geq n - 1$, and we have characterized the chemical graphs that reach the bound. This allows, for example, to characterize extremal chemical trees as well as extremal unicyclic or bicyclic chemical graphs. For $m \leq n - 2$, we have shown that there is an extremal chemical graph or order n and size m which is a disjoint union of extremal chemical trees.

Chapter 8

Extremal graphs of maximum degree at most 3 for 33 degree-based topological indices

This chapter extends the techniques described in Chapter 7 to study 33 degree-based topological indices. It is based on our article [21]; in fact, most of this chapter follows that article, apart from minor changes for coherence. Section 8.1 introduces the topic and outlines the remainder of the chapter.

8.1 Introduction

In this chapter, we focus on the second definition of chemical graphs (where the maximum degree is at most 3) and explore the bounds on topological indices of such graphs.

In this chapter, we consider 33 degree-based topological indices that we found in the literature (see Section 8.2) and whose extremal properties have given rise to scientific publications [3–7, 36, 39, 41, 45–47, 56, 57, 59–61, 78, 81, 89, 93, 102, 103, 109, 110, 113, 146, 149, 152, 153, 173, 185, 192, 199, 209, 220, 222, 225]. In the same spirit as Gutman's words, which we mentioned in Chapter 6, we can wonder whether these indices are very different from each other. We provide a partial answer by analyzing the extremal properties of these indices. We use the word "partial" for several reasons. First, we are only interested in the *extremal* properties of topological indices and it could therefore be that various indices are distinguished by other properties of interest to chemists. Second,

we only deal with chemical graphs of maximum degree at most 3. Finally, the list of topological indices studied in this article is not exhaustive, although we have tried to consider the most cited and studied in the scientific literature. Our conclusions will be clear: five families of chemical graphs are sufficient to characterize the vast majority of extremal chemical graphs of degree-based topological indices.

An edge with endpoints u and v of degree $d_u = i$ and $d_v = j$ is called an (i, j)-edge and is denoted uv. We denote x_{ij} the number of (i, j)-edges in G while n_i is the number of vertices in G of degree i. This was previously denoted by $x_{ij}(G)$ in Chapter 7, we change the notation to fit the paper as it was published.

In the next section, we give a precise definition of the chemical graphs considered in this chapter and we give the list of 33 topological indices whose extremal properties are analyzed. Section 8.3 is dedicated to defining five families of chemical graphs which are sufficient to characterize the extremal graphs for a large majority of degree-based topological indices. Tools used in our proofs are given in Section 8.4, and a characterization of extremal chemical graphs for the 33 topological indices is given in Section 8.5.

8.2 Preliminaries

As mentioned in the previous section, we are interested in connected graphs of maximum degree at most 3. To avoid border effects, we will not consider small or dense graphs which have only few possible x_{ij} values. This is now explained in detail.

There are only 10 connected graphs of order n with $1 \leq n \leq 4$, Six of them, namely K_1,K_2,K_3, K_4, P_3 and the diamond (K_4 minus an edge), are the only ones having their order and size. They therefore maximize and minimize any topological index of their order and size. The two pairs (n, m) with $n \leq 4$ that have different chemical graphs of order n and size m are P_4 and the star with 3 branches for $(n, m) = (4, 3)$ and C_4 and a triangle plus a pending vertex for $(n, m) = (4, 4)$. By restricting ourselves to connected graphs of maximum degree at most 3, it is not difficult to show that there are 10 such graphs of order $n = 5$ and 29 ones of order $n = 6$. These can be obtained using *PHOEG* [66], *House of Graphs* [28] or *Nauty's geng* [161]. Hence, given any topological index, it is easy to determine which chemical graph of order $n \leq 6$ has maximum or minimum value. From now on, we will therefore only consider connected graphs G of order at least 7, which implies $x_{11} = 0$ and $2 \leq \Delta(G) \leq 3$.

Definition 2. A *degree-based topological index* is any function f of the form

$$f(x_{12}, x_{13}, x_{22}, x_{23}, x_{33}) = c_{12}x_{12} + c_{13}x_{13} + c_{22}x_{22} + c_{23}x_{23} + c_{33}x_{33},$$

where every c_{ij} is a real number.

By abuse of notation, for a graph G, we will write $f(G)$ instead of $f(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$, where x_{ij} is the number of (i,j)-edges in G. For example, the Randić index (see Section 8.1) is the degree-based topological index with $c_{ij} = \frac{1}{\sqrt{ij}}$.

Let's focus now on dense graphs. Since we restrict ourselves to graphs G of maximum degree at most 3, the size m of such graphs is at most $\frac{3n}{2}$: if $m = \frac{3n}{2}$, then $x_{33} = m$ (G is 3-regular); if $m = \frac{3n-1}{2}$, then $x_{23} = 2$ and $x_{33} = m - 2$; if $m = \frac{3n-2}{2}$, then there are three possible cases:

- $x_{13} = 1$ and $x_{33} = m - 1$;
- $x_{23} = 4$ and $x_{33} = m - 4$;
- $x_{22} = 1$, $x_{23} = 2$ and $x_{33} = m - 3$.

Hence, given a pair (n, m) with $m \geq \frac{3n-2}{2}$, and given any degree-based topological index f, it is not difficult to determine the x_{ij} values of the connected graphs of order n, size m and maximum degree at most 3 which maximize or minimize f.

From now on, when we talk about chemical graphs, we assume that we are not in the above extreme cases (i.e., very small or very dense graphs). More precisely, here is the definition of the chemical graphs studied in this chapter.

Definition 3. A *chemical graph* is a connected graph of order $n \geq 7$, size $m \leq \frac{3n-3}{2}$ and maximum degree at most 3.

It is important to specify here that although the results that we demonstrate are valid for chemical graphs as defined above, it is possible that these results are also valid for some connected graphs of maximum degree at most 3 and of order $n < 7$ or size $m > \frac{3n-3}{2}$.

We found in the literature 33 degree-based topological indices. They are described in Table 8.1. Most of them, namely 28, appear in [100], the exceptions being ABSC which appears in [7], AG-GA, which appears in [209] and lnZagreb1, lnZagreb2 and lnZagreb3 which can be found in [185]. We are interested in the extremal properties of these indices. More precisely, given a topological index f, we aim to characterize the chemical graphs that maximize f and those that minimize f. For the 33 indices of Table 8.1, this gives potentially 66 families of chemical graphs. As will be shown, 5 families (instead of 58) are sufficient to characterize the extremal chemical graphs of 29 of the 33 topological indices.

Definition 4. Given a degree-based topological index f defined by c_{ij} values, its complement denoted $\overline{f}$ is the degree-based topological index defined by $-c_{ij}$ values.

Table 8.1: 33 Degree-based topological indices

Name	Short name	c_{ij}
Atom-bond connectivity index	ABC	$\sqrt{\frac{i+j-2}{ij}}$
Atom-bond sum-connectivity index	ABSC	$\sqrt{\frac{i+j-2}{i+j}}$
Albertson index	Albertson	$\lvert i-j\rvert$
Arithmetic-geometric index	AG	$\frac{i+j}{2\sqrt{ij}}$
Difference between AG and GA	AG-GA	$\frac{i+j}{2\sqrt{ij}} - \frac{2\sqrt{ij}}{i+j}$
Extended index	Extended	$\frac{1}{2}(\frac{i}{j} + \frac{j}{i})$
Forgotten index	Forgotten	$i^2 + j^2$
Geometric-arithmetic index	GA	$\frac{2\sqrt{ij}}{i+j}$
First Gourava index	Gourava1	$i + j + ij$
Second Gourava index	Gourava2	$(i + j)ij$
First hyper-Gourava index	hGourava1	$(i + j + ij)^2$
Second hyper-Gourava index	hGourava2	$((i + j)ij)^2$
Gourava sum-connectivity index	GouravaSC	$\frac{1}{\sqrt{i+j+ij}}$
Gourava product-connectivity index	GouravaPC	$\sqrt{ij(i + j)}$
Harmonic index	Harmonic	$\frac{2}{i+j}$
Inverse degree index	InvDeg	$i^{-2} + j^{-2}$
Inverse sum of degree index	InvSumDeg	$\frac{ij}{i+j}$
Randić index	Randić	$\frac{1}{\sqrt{ij}}$
Reciprocal Randić index	rRandić	$\sqrt{ij}$
Sigma index	Sigma	$(i - j)^2$
Sombor index	Sombor	$\sqrt{i^2 + j^2}$
Reduced Sombor index	rSombor	$\sqrt{(i - 1)^2 + (j - 1)^2}$
Sum connectivity index	SumConn	$\frac{1}{\sqrt{i+j}}$
Reciprocal sum connectivity index	rSumConn	$\sqrt{i + j}$
First Zagreb index	Zagreb1	$i + j$
Second Zagreb index	Zagreb2	ij
Augmented Zagreb index	aZagreb	$(\frac{ij}{i+j-2})^3$
First hyper-Zagreb index	hZagreb1	$(i + j)^2$
Second hyper-Zagreb index	hZagreb2	$(ij)^2$
Nat. log. of the mult. sum Zagreb index	lnZagreb1	$\ln(i + j)$
Nat. log. of the first mult. Zagreb index	lnZagreb2	$2(\frac{\ln(i)}{i} + \frac{\ln(j)}{j})$
Nat. log. of the second mult. Zagreb index	lnZagreb3	$\ln(i) + \ln(j)$
Modified first Zagreb index	mZagreb	$i^{-3} + j^{-3}$

Determining chemical graphs with the *minimum* value for f is thus equivalent to determining chemical graphs with the *maximum* value for $\overline{f}$. In the subsequent proofs, we always aim to maximize the value of a topological index in Table 8.1 or its complement.

Definition 5. A chemical graph G is *extremal* for a degree-based topological index f if it maximizes f or $\overline{f}$ over all chemical graphs of the same order and size as G.

8.3 Five families of chemical graphs

A chemical graph is characterized by five x_{ij} values, namely, $x_{12}, x_{13}, x_{22}, x_{23}$ and x_{33}. We therefore have:

$$n_1 = x_{12} + x_{13} \tag{8.1}$$

$$n_2 = \frac{x_{12} + 2x_{22} + x_{23}}{2} \tag{8.2}$$

$$n_3 = \frac{x_{13} + x_{23} + 2x_{33}}{3} \tag{8.3}$$

$$n = n_1 + n_2 + n_3 = \frac{3}{2}x_{12} + \frac{4}{3}x_{13} + x_{22} + \frac{5}{6}x_{23} + \frac{2}{3}x_{33} \tag{8.4}$$

$$m = x_{12} + x_{13} + x_{22} + x_{23} + x_{33}. \tag{8.5}$$

We now define five families F_1, F_2, F_3, F_4 and F_5 of chemical graphs. As will be shown, these are sufficient to characterize the extremal chemical graphs of 29 topological indices. For illustration, examples of chemical graphs belonging to these families are given in Figure 8.1.

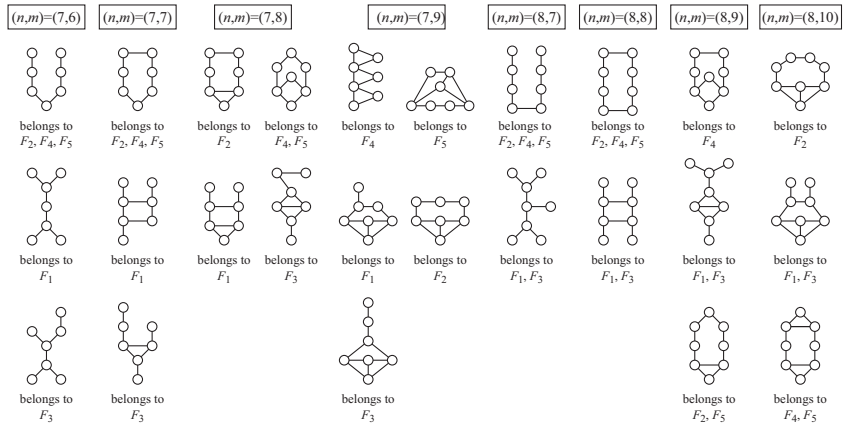


Figure 8.1: Examples of chemical graphs or order $n \in \{7, 8\}$ and size $m \leq n+2$ belonging to at least one of the families $F_1, \ldots, F_5$.

Definition 6. F_1 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
0	$\frac{3n-2m}{2}$	0	0	$\frac{4m-3n}{2}$	if n if even
0	$\frac{3n-2m-1}{2}$	0	2	$\frac{4m-3n-3}{2}$	if n if odd

Definition 7. F_2 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
2	0	$m-2$	0	0	if $m = n-1$
0	0	m	0	0	if $m = n$
0	0	$m-5$	4	1	if $m = n+1$
0	0	$3n-2m-1$	2	$3m-3n-1$	if $n+1 < m \leq \frac{3n-3}{2}$

Definition 8. F_3 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
0	$\frac{3n-2m}{2}$	0	0	$\frac{4m-3n}{2}$	if n if even
1	$\frac{3n-2m-3}{2}$	0	1	$\frac{4m-3n-1}{2}$	if n if odd

Definition 9. F_4 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
2	0	$m-2$	0	0	if $m = n-1$
0	0	$6n-5m$	$6m-6n$	0	if $n \leq m < \frac{6n}{5}$
0	0	0	$6n-4m$	$5m-6n$	if $\frac{6n}{5} \leq m \leq \frac{3n-3}{2}$

Definition 10. F_5 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
2	0	$m-2$	0	0	if $m = n-1$
0	0	m	0	0	if $m = n$
0	0	$m-6$	6	0	if $m = n+1$
0	0	$m-5$	4	1	
0	0	a	$6n-4m-2a$	$5m-6n+a$	if $n+1 < m \leq \frac{3n-3}{2}$

where a is any integer such that $\max\{0, 6n-5m\} \leq a \leq 3n-2m-1$ when $n+1<m \leq \frac{3n-3}{2}$.

It is not difficult to show that for every x_{ij} values of the five families defined above, there is at least one chemical graph having x_{ij} (i,j)-edges. This can be proved in several ways. The first approach is to use the necessary and sufficient conditions provided in Hansen *et al.* [108] for the existence of a simple connected graph with given x_{ij} values. These conditions for chemical graphs can be written as follows:

$$x_{33} \leq \tfrac{n_3(n_3-1)}{2} \quad \text{if } n_3 = 1, 2 \text{ or } 3, \tag{8.6}$$
$$x_{22} \leq \tfrac{n_2(n_2-1)}{2} \quad \text{if } n_2 = 1 \text{ or } 2, \tag{8.7}$$
$$x_{23} \leq n_2 n_3 \quad \text{if } n_2 = 1 \text{ or } 2 \text{ and } n_3 = 1, \tag{8.8}$$
$$x_{23} \geq \delta(n_2) + \delta(n_3) - 1, \tag{8.9}$$
$$x_{23} + x_{33} \geq n_3 + \delta(n_2) - 1, \tag{8.10}$$
$$x_{22} + x_{23} \geq n_2 + \delta(n_3) - 1, \tag{8.11}$$
$$x_{22} + x_{23} + x_{33} \geq n_2 + n_3 - 1. \tag{8.12}$$

where

$$\delta(x) = \begin{cases} 1 & \text{if } x \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Condition (8.12) is equivalent to $m - x_{12} - x_{13} \geq n - x_{12} - x_{13} - 1$, which is equivalent to $m \geq n - 1$. In summary, given a pair (n, m) of integers such that $n \geq 7$ and $n - 1 \leq m \leq \frac{3n-3}{2}$, and given x_{ij} values that satisfy conditions (8.1)-(8.5), we can state that there is a chemical graph of order n and size m with x_{ij} (i,j)-edges if and only if conditions (8.6)-(8.11) are satisfied. This is now illustrated with family F_1.

Given x_{ij} values as in Definition 6, Equations (8.1), (8.2) and (8.3) give

$$\begin{cases} n_1 &= \frac{3n-2m-(n \bmod 2)}{2}, \\ n_2 &= n \bmod 2, \\ n_3 &= \frac{2m-n-(n \bmod 2)}{2}. \end{cases}$$

Clearly, $n = n_1 + n_2 + n_3$ and $m = x_{12} + x_{13} + x_{22} + x_{23} + x_{33}$, which means that conditions (8.1)-(8.5) are satisfied. Let's now prove that Constraints (8.6)-(8.11) are also satisfied. Note first that $n \geq 7$ implies $m \geq n - 1 \geq 6$. Since $3n_3 = m + x_{33}$, we have $n_3 \geq 2$.

- If $n_3 = 2$ then $6 = m + x_{33} \geq 6 + x_{33}$ implies $x_{33} < 1 = \frac{n_3(n_3-1)}{2}$; if $n_3 = 3$, then $9 = m + x_{33} \geq 6 + x_{33}$ implies $x_{33} \leq 3 = \frac{n_3(n_3-1)}{2}$. Hence, Constraint (8.6) is satisfied.
- Since $x_{22} = 0 \leq \frac{n_2(n_2-1)}{2}$ for $n_2 = 1$ and 2, Constraint (8.7) is satisfied.
- As noted above, $n_3 \geq 2$ which implies that there is no Constraint (8.8).
- If n is even, then $x_{23} = n_2 = \delta(n_2) = 0$. Therefore,

  - $x_{23} = 0 \geq \delta(n_2) + \delta(n_3) - 1$;
  - Since $m \geq n - 1$, we have $2x_{33} = 4m - 3n \geq m - 3$. Hence, $m - 3 + x_{33} \leq 3x_{33}$ which implies $x_{23} + x_{33} = x_{33} \geq \frac{m+x_{33}-3}{3} = n_3 - 1 = n_3 + \delta(n_2) - 1$;
  - $x_{22} + x_{23} \geq n_2 + \delta(n_3) - 1$.

  Hence, Constraints (8.9)-(8.11) are satisfied.
- If n is odd, then, $x_{23} = 2$ and $n_2 = \delta(n_2) = 1$. Therefore,
  - $x_{23} = 2 > \delta(n_2) + \delta(n_3) - 1$.
  - Since $m \geq n - 1$, we have $2x_{33} = 4m - 3n - 3 \geq m - 6$. Hence, $m - 6 + x_{33} \leq 3x_{33}$ which implies $x_{23} + x_{33} = 2 + x_{33} \geq \frac{m+x_{33}}{3} = n_3 = n_3 + \delta(n_2) - 1$.
  - $x_{22} + x_{23} = 2 > n_2 + \delta(n_3) - 1$.

  Hence, Constraints (8.9)-(8.11) are satisfied.

Another way of proving that a chemical graph exists for given x_{ij} values is to give an explicit construction for such a graph. For family F_1, for an even order $n \geq 7$ and for $m \geq n$, one can for example consider the following construction (a similar one can be given for odd values of n and for $m = n-1$):

1. Construct a cycle on vertices $v_1, v_2, \ldots, v_{m-\frac{n}{2}}$, with edges $v_i v_{i+1}$ $(1 \leq i \leq m - \frac{n}{2} - 1)$ and $v_1 v_{m-\frac{n}{2}}$.
2. Add a matching with the $m - n$ edges $v_i v_{\lceil \frac{2m-n}{4} \rceil + i}$ $(1 < i \leq m - n)$. Let W be the set of endpoints of these edges.
3. For each $v_i \notin W$, add a pending vertex w_i adjacent to v_i.

The resulting graph belongs to F_1. Indeed, every v_i has degree 3 and every w_i has degree 1. We thus have $x_{12} = x_{22} = x_{23} = 0$. Moreover, $x_{13} = m - \frac{n}{2} - |W| = \frac{3n-2m}{2}$ and $x_{33} = m - \frac{n}{2} + |W| = \frac{4m-3n}{2}$.

In summary, it is tedious but easy to check that given x_{ij} values of one of the five graph families defined above, there is at least one chemical graph with x_{ij} (i, j)-edges. Therefore, from now on, we assume that this is true for the five families $F_1, \ldots, F_5$.

8.4 Tools used to characterize extremal chemical graphs

Given a set of x_{ij} values, we consider transformations which generate x'_{ij} values having specific properties.

Definition 11. Let $A = (a_{12}, a_{13}, a_{22}, a_{23}, a_{33})$ be a vector with integer coefficients.

- For any integer k, the (A, k)-transform of a vector $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$ is the vector $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ such that $x'_{ij} = x_{ij} + k a_{ij}$.

- We say that A is (n, m)-preserving if it satisfies the two following equations:

$$\frac{3}{2}a_{12} + \frac{4}{3}a_{13} + a_{22} + \frac{5}{6}a_{23} + \frac{2}{3}a_{33} = 0; \tag{8.13}$$

$$a_{12} + a_{13} + a_{22} + a_{23} + a_{33} = 0. \tag{8.14}$$

The idea behind these definitions is that if $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ is the (A, k)-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$ and if A is (n, m)-preserving, then

$$\frac{3}{2}x_{12} + \frac{4}{3}x_{13} + x_{22} + \frac{5}{6}x_{23} + \frac{2}{3}x_{33} = \frac{3}{2}x'_{12} + \frac{4}{3}x'_{13} + x'_{22} + \frac{5}{6}x'_{23} + \frac{2}{3}x'_{33} \qquad \text{and}$$

$$x_{12} + x_{13} + x_{22} + x_{23} + x_{33} = x'_{12} + x'_{13} + x'_{22} + x'_{23} + x'_{33}.$$

Hence, the values of n and m derived from Equations (8.4) and (8.5) are the same, whether calculated using x_{ij} or x'_{ij} values.

Let G be a chemical graph of order n and size m that maximizes the value of a topological index f over all chemical graphs of same order and same size as G. We now study the impact on the x_{ij} values of G if some of the four following values are strictly positive:

$$\begin{aligned}
V_1 &= c_{13} - c_{22} + \min_{i=1,2,3}(c_{3i} - c_{2i}) + \min_{j=2,3}(c_{3j} - c_{2j}),\\
V_2 &= c_{13} - c_{12} + \min_{i=2,3}(c_{2i} - c_{3i}),\\
V_3 &= c_{22} - c_{13} + \min_{i=1,2,3}(c_{2i} - c_{3i}) + \min_{j=2,3}(c_{2j} - c_{3j}),\\
V_4 &= 2c_{22} - c_{12} - c_{23} + 2\min_{i=1,2,3}(c_{2i} - c_{3i}).
\end{aligned}$$

Lemma 40. *Let f be a degree-based topological index such that $V_1 > 0$. A chemical graph G that maximizes f over all chemical graphs of the same order and size as G has no (2,2)-edge.*

Proof. Assume G contains an edge uv with both endpoints of degree 2.

- If u and v have no common neighbor, then let x be the second neighbor of u and let y be the second neighbor of v. At least one of x, y, say y, has degree at least two, else G has order 4. We can then obtain a chemical graph G' by replacing ux by vx. Let $i \in \{1, 2, 3\}$ be the degree of x and $j \in \{2, 3\}$ the degree of y. The graph G' contradicts the maximality of G since

  $$f(G') = f(G) + c_{13} - c_{22} + (c_{3i} - c_{2i}) + (c_{3j} - c_{2j}) \geq f(G) + V_1 > f(G).$$

- If u and v have a common neighbor w, then w has degree 3, else G has order $n = 3$. Also, the third neighbor x of w has degree at least two, else G has order $n = 4$.

– If x has degree 2, then let y be its second neighbor and let $i \in \{1, 2, 3\}$ be the degree of y. We can obtain a chemical graph G' by replacing uv by vx. Then G' contradicts the maximality of G since

$$\begin{aligned} f(G') &= f(G) + c_{13} - c_{22} + (c_{3i} - c_{2i}) + (c_{33} - c_{23}) \\ &\geq f(G) + V_1 > f(G). \end{aligned}$$

– If x has degree 3, then let y and z be the two other neighbors of x. We can obtain a chemical graph G' by replacing xy and xz by uy and vz. Then G' contradicts the maximality of G since

$$f(G') = f(G) + c_{13} - c_{22} + 2(c_{33} - c_{23}) \geq f(G) + V_1 > f(G). \quad \blacksquare$$

Lemma 41. *Let f be a degree-based topological index such that $V_1 > 0$ and $V_2 > 0$. A chemical graph G that maximizes f over all chemical graphs of the same order and size as G has no (1,2)-edge and no (2,2)-edge.*

Proof. We already know from Lemma 40 that G has no (2,2)-edge. Let uv be an edge in G with u of degree 1 and v of degree 2. Let w be the second neighbor of v. Note that w has degree 3, else G has order $n = 3$. Let x and y be the two other neighbors of w. At least one of them, say x has degree $i \geq 2$, else G has order $n = 5$. We can obtain a chemical graph G' by replacing uv and wx by uw and vx. Then G' contradicts the maximality of G since

$$f(G') = f(G) + c_{13} - c_{12} + (c_{2i} - c_{3i}) \geq f(G) + V_2 > f(G). \quad \blacksquare$$

Lemma 42. *Let f be a degree-based topological index such that $V_3 > 0$. A chemical graph G that maximizes f over all chemical graphs of the same order and size as G has no (1,3)-edge.*

Proof. Let uv be an edge with u of degree 1 and v of degree 3, and let x and y be the two other neighbors of v. At least one of x and y, say y has degree at least two, else G has order $n = 4$. We can then obtain a chemical graph G' by replacing vy by uy. Let i be the degree of x and j the degree of y. Then G' contradicts the maximality of G since

$$f(G') = f(G) + c_{22} - c_{13} + (c_{2i} - c_{3i}) + (c_{2j} - c_{3j}) \geq f(G) + V_3 > f(G). \quad \blacksquare$$

Lemma 43. *Let f be a degree-based topological index such that $V_4 > 0$. If a chemical graph G maximizes f over all chemical graphs of the same order and size as G then either $m = n - 1$ and G is P_n or $m \geq n$ and G has no (1,2)-edge.*

Proof. Let uv be an edge with u of degree 1 and v of degree 2. If G is not P_n, then there is a vertex w of degree 3 in G such that v and w are linked by a chain in which all internal vertices have degree 2 in G. Let x and y be the two neighbors of w that are not on the chain. We can obtain a chemical graph G' by replacing wx by ux. Let i be the degree of x and j the degree of y. Then G' contradicts the maximality of G since

$f(G')=f(G)+2c_{22}-c_{12}-c_{23}+(c_{2i}-c_{3i})+(c_{2j}-c_{3j})\geq f(G)+V_4>f(G)$. ■

8.5 Characterization of extremal graphs

We first characterize the extremal graphs of 29 of the 33 degree-based topological indices of Table 8.1. The proofs involve the following values:

$$\begin{aligned} V_5 &= c_{13} - 4c_{23} + 3c_{33},\\ V_6 &= c_{22} - 2c_{23} + c_{33},\\ V_7 &= c_{12} - c_{13} - c_{23} + c_{33},\\ V_8 &= -2c_{12} + 3c_{13} - 2c_{23} + c_{33}. \end{aligned}$$

8.5.1 Five graph families for 29 topological indices

We first show that the chemical graphs in F_1 maximize all degree-based topological indices such that V_1, V_2 and V_5 are strictly positive.

Theorem 44. *Let f be a degree-based topological index such that $V_1 > 0$, $V_2 > 0$ and $V_5 > 0$. A chemical graph G maximizes f over all chemical graphs of the same order and size as G if and only if $G \in F_1$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. As shown in Lemma 41, $V_1 > 0$ and $V_2 > 0$ imply $x_{12} = x_{22} = 0$. Hence, $2n_2 = x_{23}$, which means that x_{23} is even.

If $x_{23} \leq 2$ then $n_2 = n \bmod 2$ (since the number $n_1 + n_3$ of odd degree vertices is even), which implies $x_{23} = 2(n \bmod 2)$. Equations (8.4) and (8.5) then give $x_{13} = \frac{3n-2m-(n \bmod 2)}{2}$ and $x_{33} = \frac{4m-3n-3(n \bmod 2)}{2}$, which means that G belongs to F_1.

If $x_{23} \geq 4$, then consider the vector $A = (0, 1, 0, -4, 3)$ associated with V_5 and let $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ be the $(A, \lfloor \frac{x_{23}}{4} \rfloor)$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$. We then have $x'_{12} = x'_{22} = 0$ and $x'_{23} \leq 2$. Since A is (n,m)-preserving, we conclude as above that $x'_{13} = \frac{3n-2m-(n \bmod 2)}{2}$ and $x'_{33} = \frac{4m-3n-3(n \bmod 2)}{2}$. Let G' be a graph in F_1 having exactly x'_{ij} (i,j)-edges. The maximality of G is contradicted by G' since

$f(G') = f(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33}) = f(G) + V_5\lfloor\frac{x_{23}}{4}\rfloor > f(G)$. ■

Theorem 45. *Let f be a degree-based topological index such that $V_3 > 0$, $V_4 > 0$ and $V_6 > 0$. A chemical graph G maximizes f over all chemical graphs of the same order and size as G if and only if $G \in F_2$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. As shown by Lemmas 42 and 43, $V_3 > 0$ and $V_4 > 0$ imply $G = \mathsf{P}_n$ or $n_1 = 0$. Hence, if $m = n-1$, then $G = \mathsf{P}_n$ (since trees have vertices of degree 1), which means that $x_{12} = 2$, $x_{22} = m-2$ and $G \in F_2$. So assume $n \le m < \frac{3n-3}{2}$, which implies that $n_1 = 0$ and x_{23} is even.

- If $m = n$, Equations (8.4) and (8.5) give $x_{23} = x_{33} = 0$, which implies $x_{22} = m$ and $G \in F_2$.
- If $m = n+1$, Equations (8.4) and (8.5) give $x_{23} + 2x_{33} = 6$, which implies $n_3 = 2$ and $x_{33} \le 1$. Hence, there are only two possibilities:
  - $x_{22} = m-6$, $x_{23} = 6$,$x_{33} = 0$, or
  - $x_{22} = m-5$, $x_{23} = 4$,$x_{33} = 1$.

  Since $c_{22}-2c_{23}+c_{33}=V_6 > 0$, the second solution has a larger value, which implies $G \in F_2$.
- If $n+1 < m < \frac{3n-3}{2}$ and $x_{23} = 2$, then Equations (8.4) and (8.5) give $x_{22} = 3n-2m-1$ and $x_{33} = 3m-3n-1$, which implies $G \in F_2$. So assume $x_{23} \ge 4$, consider the vector $A = (0,0,1,-2,1)$ associated with V_6 and let $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ be the $(A, \frac{x_{23}-2}{2})$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$. Hence, $x'_{12} = x'_{13} = 0$ and $x'_{23} = 2$. Since A is (n,m)-preserving, we conclude as above that $x'_{22} = 3n-2m-1$ and $x'_{33} = 3m-3n-1$. Consider any chemical graph G' in F_2 having exactly x'_{ij} (i,j)-edges. The maximality of G is contradicted by G' since

  $f(G') = f(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33}) = f(G) + V_6(\frac{x_{23}-2}{2}) > f(G)$. ■

Corollary 46. *$F_1 \cup F_2$ is the set of extremal chemical graphs for the 13 degree-based topological indices ABSC, AG, AG-GA, Extended, GA, GouravaSC, Harmonic, Randić, Sombor, rSombor, SumConn, rSumConn and lnZagreb1.*

Proof. It is easy to check that

- V_1, V_2 and V_5 are strictly positive for ABSC, AG, AG-GA, Extended, rSumConn, Sombor, rSombor, lnZagreb1, $\overline{\text{GA}}$, $\overline{\text{GouravaSC}}$, $\overline{\text{Harmonic}}$, $\overline{\text{Randić}}$ and $\overline{\text{SumConn}}$, which implies that F_1 is the set of chemical graphs which maximize ABSC, AG, AG-GA, Extended, rSumConn, Sombor, rSombor, lnZagreb1 and minimize GA, GouravaSC, Harmonic, Randić and SumConn.

- V_3, V_4 and V_6 are strictly positive for $\overline{\text{ABSC}}$, $\overline{\text{AG}}$, $\overline{\text{AG-GA}}$, $\overline{\text{Extended}}$, $\overline{\text{Sombor}}$, $\overline{\text{rSombor}}$, $\overline{\text{rSumConn}}$, $\overline{\text{lnZagreb1}}$, GA, GouravaSC, Harmonic, Randić and SumConn, which implies that F_2 is the set of chemical graphs which minimize ABSC, AG, AG-GA, Extended, Sombor, rSombor, rSumConn, lnZagreb1 and maximize GA, GouravaSC, Harmonic, Randić and SumConn. ■

We now characterize the degree-based topological indices f for which the chemical graphs in F_3 and F_4 maximize f.

Theorem 47. *Let f be a degree-based topological index such that $V_1>0, V_6>0$, $V_7>0$ and $V_8>0$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_3$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. As shown by Lemma 40, $V_1 > 0$ implies $x_{22} = 0$. Let W_{33} be the set of vertices of degree 2 in G with two neighbors of degree 3, and let W_{13} be the set of vertices of degree 2 in G with one neighbor of degree 1 and the other of degree 3. Since $n > 3$, we have $n_2 = |W_{33}| + |W_{13}|$.

- If $n_2 = 0$, then $x_{12} = x_{22} = x_{23} = 0$ and Equations (8.4) and (8.5) give $x_{13} = \frac{3n-2m}{2}$ and $x_{33} = \frac{4m-3n}{2}$, which implies that G belongs to F_3.
- If $n_2 = 1$, then let v be the vertex of degree 2 in G.
  - if $v \in W_{13}$ then $x_{12} = x_{23} = 1, x_{22} = 0$ and Equations (8.4) and (8.5) give $x_{13} = \frac{3n-2m-3}{2}$ and $x_{33} = \frac{4m-3n-1}{2}$;
  - if $v \in W_{33}$ then $x_{12} = x_{22} = 0, x_{23} = 2$ and Equations (8.4) and (8.5) give $x_{13} = \frac{3n-2m-1}{2}$ and $x_{33} = \frac{4m-3n-3}{2}$.

  Since $c_{12} - c_{13} - c_{23} + c_{33} = V_7 > 0$, the first case has a larger value $f(G)$, which implies that G belongs to F_3.
- If $n_2 > 1$, then the two neighbors of each vertex in W_{33} are adjacent. Indeed, if a vertex $v \in W_{33}$ has two non-adjacent neighbors u_1 and u_2, then consider any other vertex w of degree 2 and let u_3 be one of its neighbors: by replacing vu_1, vu_2, wu_3 by vw, vu_3 and u_1u_2, we get a chemical graph G' which contradicts the maximality of G since

$$f(G') = f(G) + c_{22} - 2c_{23} + c_{33} = f(G) + V_6 > f(G).$$

  Let us now show that $|W_{33}| \leq 1$. Assume by contradiction that W_{33} contains at least two vertices v_1 and v_2. Since $n > 4$, there are two non-adjacent vertices u_1 and u_2 such that u_1 is adjacent to v_1 but not to v_2, while u_2 is adjacent to v_2 but not to v_1. Let w_1 be the second neighbor of v_1, and let G' be the chemical graph obtained from G by

replacing v_1w_1 and v_2u_2 by v_1u_2 and v_2w_1. Then G' has $n_2 > 1$ vertices of degree 2 and one of them, namely v_1, has two non-adjacent neighbors u_1, u_2 of degree 3. We have shown above that this implies that G' does not maximize f while $f(G') = f(G)$, a contradiction.
Hence, $|W_{33}| \leq 1$, which implies $|W_{13}| \geq 1$. So let v be a vertex in W_{13}, let u be another vertex of degree 2, let w be the neighbor of v of degree 1, and let G' be the chemical graph obtained from G by replacing vw by uw:

- if $u \in W_{13}$ then $f(G') = f(G) - 2c_{12} + 3c_{13} - 2c_{23} + c_{33} = f(G) + V_8 > f(G)$;
- if $u \in W_{33}$ then $f(G') = f(G) - c_{12} + 2c_{13} - 3c_{23} + 2c_{33} = f(G) + V_7 + V_8 > f(G)$.

In both cases, G' contradicts the maximality of G. ■

Theorem 48. *Let f be a degree-based topological index such that $V_3>0$, $V_4>0$ and $V_6<0$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_4$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. As shown by Lemmas 42 and 43, $V_3 > 0$ and $V_4 > 0$ imply $G = \mathsf{P}_n$ or $n_1 = 0$. Hence, if $m = n - 1$, then $G = \mathsf{P}_n$, which means that $x_{12} = 2$, $x_{22} = m - 2$ and $G \in F_4$. So assume $m > n - 1$, which implies $n_1 = 0$. Equations (8.4) and (8.5) give $x_{23} = 6n - 4m - 2x_{22}$ and $x_{33} = 5m - 6n + x_{22}$. Hence, $x_{22} \geq \max\{0, 6n - 5m\}$.

- If $x_{22} = 0$, then $x_{23} = 6n - 4m$, $x_{33} = 5m - 6n$ and $G \in F_4$.
- If $x_{22} = 6n - 5m > 0$, then $x_{23} = 6m - 6n$, $x_{33} = 0$ and $G \in F_4$.
- If $x_{22} > 0$ and $x_{22} \neq 6n - 5m$, then $x_{33} > 0$. Consider the vector $A = (0, 0, -1, 2, -1)$ associated with $-V_6$ and let $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ be the $(A, \min\{x_{22}, x_{22} - 6n + 5m\})$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$. Hence, $x'_{22} = \max\{0, 6n - 5m\}$ and $x'_{12} = x'_{13} = 0$. Since A is (n, m)-preserving we conclude as above that there is a chemical graph G' in F_4 having exactly x'_{ij} (i,j)-edges. Then G' contradicts the maximality of G since

$$\begin{aligned} f(G') &= f(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33}) \\ &= f(G) - V_6 \min\{x_{22}, x_{22} - 6n + 5m\} > f(G). \end{aligned}$$

■

Corollary 49. *$F_3 \cup F_4$ is the set of extremal chemical graphs for the 10 degree-based topological indices Gourava1, Gourava2, hGourava1, hGourava2, GouravaPC, InvSumDeg, rRandić, Zagreb2, hZagreb1, hZagreb2.*

Proof. It is easy to check that

- $V_1{>}0$, $V_6{>}0$, $V_7{>}0$ and $V_8{>}0$ for Gourava1, Gourava2, hGourava1, hGourava2, GouravaPC, InvSumDeg, rRandić, Zagreb2, hZagreb1, hZagreb2, which implies that F_3 is the set of chemical graphs which maximize these topological indices.
- $V_3{>}0, V_4{>}0$ and $V_6{<}0$ for $\overline{\text{Gourava1}}$, $\overline{\text{Gourava2}}$, $\overline{\text{hGourava1}}$, $\overline{\text{hGourava2}}$, $\overline{\text{GouravaPC}}$, $\overline{\text{InvSumDeg}}$, $\overline{\text{rRandić}}$, $\overline{\text{Zagreb2}}$, $\overline{\text{hZagreb1}}$, $\overline{\text{hZagreb2}}$, which implies that F_4 is the set of chemical graphs which minimize the 10 topological indices. ■

Note that $F_1 \cup F_3$ is the set of chemical graphs with x_{ij} (i,j)-edges such that

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
0	$\frac{3n-2m}{2}$	0	0	$\frac{4m-3n}{2}$	if n if even
1	$\frac{3n-2m-3}{2}$	0	1	$\frac{4m-3n-1}{2}$	if n is odd
0	$\frac{3n-2m-1}{2}$	0	2	$\frac{4m-3n-3}{2}$	

Theorem 50. *Let f be a degree-based topological index such that $V_1{>}0$, $V_5{>}0$ and $V_7{=}0$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_1 \cup F_3$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. Note that the two possibilities for the x_{ij} values when n is odd give the same value $f(G)$ since $c_{12} - c_{13} - c_{23} + c_{33} = V_7 = 0$. As shown by Lemma 40, $V_1 > 0$ implies $x_{22} = 0$. Moreover, $n > 3$ implies $x_{12} \leq x_{23}$, and x_{12} and x_{23} have the same parity.

1. If $x_{12} = 0$ then, as shown in Theorem 44, $G \in F_1$, else there is a graph $G' \in F_1$ so that $f(G') > f(G)$.
2. If $x_{12} = 1$ then
   - if $x_{23} = 1$ then Equations (8.4) and (8.5) give $x_{13} = \frac{3n-2m-3}{2}$ and $x_{33} = \frac{4m-3n-1}{2}$, which implies that G belongs to F_3.
   - if $x_{23} \geq 3$, then consider the vector $A = (-1, 2, 0, -3, 2)$ associated with $V_5{-}V_7$ and let $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ be the $(A,1)$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$. Note that $x'_{12} = x'_{22} = 0$ and A is (n,m)-preserving. Hence, we have shown in case 1. that there is a graph $G' \in F_1$ which contradicts the maximality of G since

$$\begin{aligned} f(G') &\geq f(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33}) \\ &= f(G) + V_5 - V_7 = f(G) + V_5 > f(G). \end{aligned}$$

3. If $x_{12} \geq 2$ then $x_{23} \geq x_{12} \geq 2$. Consider the vector $A = (-2, 3, 0, -2, 1)$ associated with $V_5 - 2V_7$ and let $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ be the $(A, \lfloor \frac{x_{12}}{2} \rfloor)$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$. Since $x'_{12} = x'_{22} = 0$ and A is (n, m)-preserving, we have shown in case 1. that there is a graph $G' \in F_1$ which contradicts the maximality of G since

$$f(G') \geq f(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33}) = f(G) + \lfloor \frac{x_{12}}{2} \rfloor (V_5 - 2V_7)$$
$$= f(G) + \lfloor \frac{x_{12}}{2} \rfloor V_5 > f(G).$$ ■

Theorem 51. *Let f be a degree-based topological index such that $V_3>0$, $V_4>0$ and $V_6=0$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_5$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i, j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. Note first that the two possibilities for the x_{ij} values when $m = n + 1$ have the same value $f(G)$ since $c_{22} - 2c_{23} + c_{33} = V_6 = 0$. For the same reason, all solutions for $n + 1 < m \leq \frac{3n-3}{2}$ have the same value.

As shown by Lemmas 42 and 43, $V_3 > 0$ and $V_4 > 0$ imply $G = \mathsf{P}_n$ or $n_1 = 0$. Hence, if $m = n - 1$, then $G = \mathsf{P}_n$, which means that $x_{12} = 2$, $x_{22} = m - 2$ and $G \in F_5$. So assume $n \leq m < \frac{3n-3}{2}$, which implies that $n_1 = 0$ and x_{23} is even.

- If $m = n$, Equations (8.4) and (8.5) give $x_{23} = x_{33} = 0$ and $x_{22} = m$, which implies $G \in F_5$.
- If $m = n+1$, Equations (8.4) and (8.5) give $x_{23}+2x_{33} = 6$, which implies $n_3 = 2$ and $x_{33} \leq 1$. Hence, there are only two possibilities:
  - $x_{22} = m - 6$, $x_{23} = 6$,$x_{33} = 0$, or
  - $x_{22} = m - 5$, $x_{23} = 4$,$x_{33} = 1$,

  and both imply $G \in F_5$.
- If $n + 1 < m \leq \frac{3n-3}{2}$, then $x_{23} \geq 2$ and Equations (8.4) and (8.5) give $x_{22} = a$, $x_{23} = 6n - 4m - 2a$ and $x_{33} = 5m - 6n + a$. Since $x_{33} \geq 0$ and $x_{23} \geq 2$, we have $\max\{0, 6n - 5m\} \leq a \leq 3n - 2m - 1$, which implies $G \in F_5$. ■

Corollary 52. *$F_1 \cup F_3 \cup F_5$ is the set of extremal chemical graphs for the topological indices Forgotten, InvDeg, Zagreb1, lnZagreb2, lnZagreb3 and mZagreb.*

Proof. It is easy to check that

- $V_1>0$, $V_5>0$ and $V_7=0$ for Forgotten, InvDeg, Zagreb1, $\overline{\text{lnZagreb2}}$, lnZagreb3 and mZagreb which implies that $F_1 \cup F_3$ is the set of chemical graphs which maximize Forgotten, InvDeg, Zagreb1 and mZagreb and minimize lnZagreb2.

- $V_3{>}0, V_4{>}0$ and $V_6{=}0$ for $\overline{\text{Forgotten}}$, $\overline{\text{InvDeg}}$, $\overline{\text{Zagreb1}}$, lnZagreb2, $\overline{\text{lnZagreb3}}$ and $\overline{\text{mZagreb}}$, which implies that F_5 is the set of chemical graphs which minimize Forgotten, InvDeg, Zagreb1, mZagreb and maximize lnZagreb2. ■

8.5.2 Additional families of extremal chemical graphs

As proved in the previous section, the five families $F_1, \ldots, F_5$ are sufficient to characterize all extremal graphs of 29 topological indices. However, some degree-based topological indices have extremal chemical graphs that do not belong to any of the five families. We give here four examples. More precisely, we characterize the extremal chemical graphs of the topological indices ABC, Albertson, Sigma, and aZagreb. For this purpose, we define new families $F_6, \ldots, F_{11}$ of chemical graphs characterized by x_{ij} values. Here again, as explained in Section 8.3, it is easy to check that given x_{ij} values of one of the graph families, there is at least one chemical graph with x_{ij} (i,j)-edges. Examples of chemical graphs belonging to at least one of the families $F_6, \ldots, F_{11}$, but to none of the families $F_1, \ldots, F_5$ are given in figure 8.2.

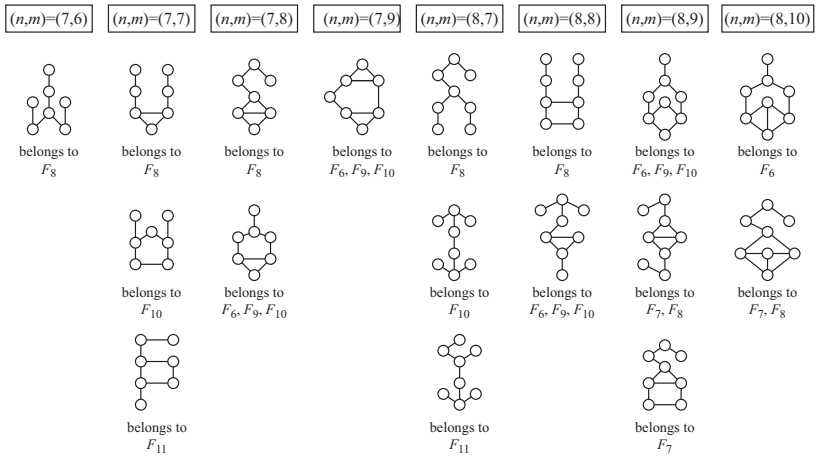


Figure 8.2: Examples of chemical graphs or order $n{\in}\{7,8\}$ and size $m{\leq}n{+}2$ belonging to at least one of the families $F_6, \ldots, F_{11}$ but to none of the families $F_1, \ldots, F_5$.

Definition 12. F_6 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}
0	a	0	$6n-4m-4a$	$5m-6n+3a$

where a is any integer such that $\max\{0, \lceil \frac{6n-5m}{3} \rceil\} \leq a \leq \lfloor \frac{3n-2m}{2} \rfloor$.

Theorem 53. *Let f be a degree-based topological index such that $V_1>0$, $V_2>0$ and $V_5=0$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_6$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. As shown in Lemma 41, $V_1 > 0$ and $V_2 > 0$ imply $x_{12} = x_{22} = 0$. Hence, it follows from Equations (8.4) and (8.5) that $x_{13} = a$, $x_{23} = 6n - 4m - 4a$ and $x_{33} = 5m - 6n + 3a$. Since $x_{23} \geq 0$ and $x_{33} \geq 0$, we have $\max\{0, \lceil \frac{6n-5m}{3} \rceil\} \leq a \leq \lfloor \frac{3n-2m}{2} \rfloor$. All possible solutions with the various values of a have the same value since $c_{13} - 4c_{23} + 3c_{33} = V_5 = 0$. ■

Definition 13. F_7 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

$(n \geq 8)$

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
2	0	$m-2$	0	0	if $m = n-1$
0	0	m	0	0	if $m = n$
0	0	$m-5$	4	1	
1	0	$m-7$	3	3	if $m = n+1$
2	0	$m-9$	2	5	
0	0	$3n-2m-1$	2	$3m-3n-1$	if $n+1<m\leq\frac{3n-3}{2}$
1	0	$3n-2m-3$	1	$3m-3n+1$	

Theorem 54. *Let f be a degree-based topological index such that $V_3>0$, $V_6>0$ and $V_5+V_7=2V_6$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_7$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. Note that the three possible cases for $m = n + 1$ have the same value since $c_{12} - 2c_{22} - c_{23} + 2c_{33} = V_5 + V_7 - 2V_6 = 0$. For the same reason, the two possibilities for $n + 1 < m < \frac{3n-3}{3}$ have the same value.

As shown by Lemma 42, $V_3 > 0$ implies $x_{13} = 0$. In what follows, for two integer a and b such that $(a,b) \neq (x_{12}, x_{23})$, we say that G is (a,b)-dominated if $a \leq x_{12}$, $b-a \leq x_{23} - x_{12}$, and there is a chemical graph of order n and size m which has a $(1,2)$-edges, b $(2,3)$-edges, and no $(1,3)$-edge. In such a case consider the (n,m)-preserving vector $A = (0,0,1,-2,1)$ associated with V_6 and let $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ be the $(A, \frac{x_{23}-x_{12}+a-b}{2})$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$. We thus have $x_{12} = x'_{12}$ and $x'_{23} = x_{12} + b - a$. Let $A' = (-1,0,3,-1,-1)$ be the (n,m)-preserving vector associated with

$3V_6 - V_5 - V_7$ and let $(x''_{12}, x''_{13}, x''_{22}, x''_{23}, x''_{33})$ be the $(A', x_{12} - a)$-transform of $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$. We now have $x''_{12} = a$ and $x''_{23} = b$. Let G' be a graph with x''_{ij} (i, j)-edges. Note that if $x_{12} = a$, then $x_{23} - x_{12} + a - b > 0$. Hence, G' contradicts the maximality of G since $3V_6 - V_5 - V_7 = V_6 > 0$ and

$$\begin{aligned} f(G') &= f(x''_{12}, x''_{13}, x''_{22}, x''_{23}, x''_{33}) = f(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33}) + (x_{12} - a)V_6 \\ &= f(G) + (\frac{x_{23} - x_{12} + a - b}{2})V_6 + (x_{12} - a)V_6 > f(G). \end{aligned}$$

Let us now analyze the possible values for m and use Equations (8.4) and (8.5) to derive x_{ij} values.

- If $m = n - 1$, then assume $x_{23} > 0$. It follows that $3 \leq x_{12} \leq x_{23}$. Hence, $x_{12} = x_{23} = 3$, else G is $(3, 3)$-dominated. We therefore have $x_{22} = m - 6$ and $x_{33} = 0$ and $f(\mathsf{P}_n) - f(G) = -c_{12} + 4c_{22} - 3c_{23} = 4V_6 - V_5 - V_7 = 2V_6 > 0$, which contradicts the maximality of G. Hence $x_{23} = 0$ which implies $G = \mathsf{P}_n \in F_7$.
- if $m = n$, then $x_{12} = x_{23} = 0$, else G is $(0, 0)$-dominated. We therefore have $x_{22} = m$ and $x_{33} = 0$ which implies $G = \mathsf{C}_n \in F_7$.
- If $m = n + 1$, then
  - if $x_{12} = 0$, then $x_{23} \geq 4$. Hence, $x_{23} = 4$ else G is $(0, 4)$-dominated. We therefore have $x_{22} = m - 5$ and $x_{33} = 1$, meaning that $G \in F_7$.
  - if $x_{12} = 1$, then $x_{23} \geq 3$. Hence, $x_{23} = 3$ else G is $(1, 3)$-dominated. We therefore have $x_{22} = m - 7$ and $x_{33} = 3$, meaning that $G \in F_7$.
  - if $x_{12} = 2$, then $x_{23} \geq 2$. Hence, $x_{23} = 2$ else G is $(2, 2)$-dominated. We therefore have $x_{22} = m - 9$ and $x_{33} = 5$, meaning that $G \in F_7$.
  - if $x_{12} > 2$, then $x_{23} > 2$, which means that G is $(2, 2)$-dominated, which contradicts the maximality of G.
- If $n + 1 < m \leq \frac{3n-3}{2}$ then
  - if $x_{12} = 0$, then $x_{23} \geq 2$. Hence, $x_{23} = 2$, else G is $(0, 2)$-dominated. We therefore have $x_{22} = 3n - 2m - 1$ and $x_{33} = 3m - 3n - 1$, meaning that $G \in F_7$.
  - if $x_{12} = 1$, then $x_{23} \geq 1$. Hence, $x_{23} = 1$, else G is $(1, 1)$-dominated. We therefore have $x_{22} = 3n - 2m - 3$ and $x_{33} = 3m - 3n + 1$, meaning that $G \in F_7$.
  - if $x_{12} \geq 2$, then $x_{23} \geq 2$. Hence, G is $(1, 1)$-dominated, which contradicts the maximality of G. ■

Corollary 55. *$F_6 \cup F_7$ is the set of extremal chemical graphs for the topological index Sigma.*

Proof. It is easy to check that $V_1 > 0$, $V_2 > 0$ and $V_5 = 0$ for Sigma, while $V_3 > 0$, $V_6 > 0$ and $V_5 + V_7 = 2V_6$ for $\overline{\text{Sigma}}$, which means that F_6 is the set of chemical graphs which maximize Sigma, while F_7 is the set of chemical graphs which minimize Sigma. ■

Definition 14. F_8 is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
2	0	$m-2$	0	0	if $m+1=n \in \{7,8,9\}$
3	0	$m-6$	3	0	
2	0	$m-7$	4	1	if $m=n \in \{7,8\}$
1	0	1	3	3	if $n=7$ and $m=8$
$\lfloor \frac{3n-2m}{3} \rfloor$	0	$m \bmod 3$	$\lfloor \frac{3n-2m}{3} \rfloor$	$\lfloor \frac{7m-6n}{3} \rfloor$	otherwise

Theorem 56. *Let f be a degree-based topological index such that $V_3>0$, $V_6>0$ and $V_5+V_7=4V_6$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_8$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. As shown by Lemma 42, $V_3 > 0$ implies $x_{13} = 0$.

If $m+1 = n \in \{7,8,9\}$, then there are only two possibilities: either $x_{12} = 2$ and $x_{22} = m-2$, or $x_{12} = 3$, $x_{23} = 3$ and $x_{22} = m-6$. Since $c_{12}-4c_{22}+3c_{23} = V_5+V_7-4V_6 = 0$, we deduce that both cases correspond to an optimal graph G that belongs to F_8. We now assume $m \geq n$ or $m+1 = n \geq 10$. If $n_3 = 0$, then $G = \mathsf{P}_n$ and $m+1 = n \geq 10$ or $G = \mathsf{C}_n$ and $n = m$.

- If $m+1 = n \geq 10$, then the graph G' with $x_{12} = x_{23} = 4, x_{33} = 1$ and $x_{22} = m-9$ has value $f(G') = f(G) + 2c_{12} - 7c_{22} + 4c_{23} + c_{33} = f(G)+2V_5+2V_7-7V_6 = f(G)+V_6 > 0$, which contradicts the maximality of G.
- If $n = m$, then the graph G' with $x_{12} = 2, x_{23} = 4$, $x_{33} = 1$ and $x_{22} = m-7$ has value $f(G') = f(G) + 2c_{12} - 7c_{22} + 4c_{23} + c_{33} = f(G)+2V_5+2V_7-7V_6 = f(G)+V_6 > 0$, which contradicts the maximality of G.

If $n_3 > 0$, then consider the following 4 cases:

- If $m = n = 7$, then there are only two possibilities: either $x_{12} = 2$, $x_{23} = 4$ and $x_{33} = 1$, or $x_{12} = 1$, $x_{23} = 3$ and $x_{22} = 3$. Since $c_{12}+c_{23}+c_{33}-3c_{22} = V_5 + V_7 - 3V_6 = V_6 > 0$, we deduce that the first solution is the best, which means that $G \in F_8$.

- If $m = n = 8$, then there are only three possibilities: either $x_{12} = 2$, $x_{22} = 1$, $x_{23} = 4$ and $x_{33} = 1$, or $x_{12} = 1$, $x_{23} = 3$ and $x_{22} = 4$, or $x_{12} = 2$ and $x_{23} = 6$. As in the previous case, the first solution is better than the second. Also, since $c_{22} - 2c_{23} + c_{33} = V_6 > 0$, we deduce that the first solution is better than the third one, which implies $G \in F_8$.
- If $n = 7$ and $m = 8$, there are four possibilities:
  - $x_{22} = 2$ and $x_{23} = 6$;
  - $x_{22} = 3$, $x_{23} = 4$ and $x_{33} = 1$;
  - $x_{12} = 1$, $x_{23} = 5$ and $x_{33} = 2$;
  - $x_{12} = x_{22} = 1$, $x_{23} = x_{33} = 3$.

  The fourth is better than the first since $c_{12} - c_{22} - 3c_{23} + 3c_{33} = V_5 + V_7 - V_6 = 3V_6 > 0$. It is better than the second since $c_{12} - 2c_{22} - c_{23} + 2c_{33} = V_5 + V_7 - 2V_6 = 2V_6 > 0$. It is better than the third since $c_{22} - 2c_{23} + c_{33} = V_6 > 0$. Hence, $G \in F_8$.
- For the remaining case where $n \in \{7, 8, 9\}$ and $m \geq 9$, or $n \geq 10$, consider the two (n, m)-preserving vectors $A = (0, 0, 1, -2, 1)$ and $A' = (1, 0, -3, 1, 1)$ associated with V_6 and $V_5 + V_7 - 3V_6$, respectively. Let $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$ be the $(A, \frac{x_{23}-x_{12}}{2})$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$. Note that $x'_{12} = x'_{23}$ and $x'_{13} = 0$. Let $(x''_{12}, x''_{13}, x''_{22}, x''_{23}, x''_{33})$ be the $(A', \lfloor \frac{x'_{22}}{3} \rfloor)$-transform of $(x'_{12}, x'_{13}, x'_{22}, x'_{23}, x'_{33})$. Note that $x''_{13} = 0$, $x''_{12} = x''_{23}$ and $x''_{22} \leq 2$. Equations (8.4) and (8.5) then give $x''_{22} = m \bmod 3$, $x''_{12} = x''_{23} = \lfloor \frac{3n-2m}{3} \rfloor$, and $x''_{33} = \lfloor \frac{7m-6n}{3} \rfloor$. Consider any graph G' in F_8 with x''_{ij} (i, j)-edges. We then have

$$\begin{aligned} f(G') &= f(G) + \frac{x_{23} - x_{12}}{2} V_6 + \lfloor \frac{x'_{22}}{3} \rfloor (V_5 + V_7 - 3V_6) \\ &= f(G) + (\frac{x_{23} - x_{12}}{2} + \lfloor \frac{x'_{22}}{3} \rfloor) V_6. \end{aligned}$$

  If $x_{23} - x_{12} > 0$, or $x_{23} - x_{12} = 0$ and $x_{22} = x'_{22} > 2$,then $f(G') > f(G)$, which contradicts the maximality of G. Hence, we can choose G' equal to G, which implies $G \in F_8$. ■

Definition 15. F_9 is the set of chemical graphs with the following numbers x_{ij} of (i, j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
0	$\frac{6n-5m+(2m \bmod 3)}{3}$	0	$\frac{8m-6n-4(2m \bmod 3)}{3}$	$2m \mod 3$	if $n-1 \leq m \leq \frac{6n+2}{5}$
0	0	0	$6n - 4m$	$5m - 6n$	if $\frac{6n+3}{5} \leq m \leq \frac{3n-3}{2}$

Theorem 57. *Let f be a degree-based topological index such that $V_1 > 0$, $V_2 > 0$ and $V_5 < 0$. A chemical graph G maximizes f over all graphs of the same order and size as G if and only if $G \in F_9$.*

Proof. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. As shown by Lemma 41, $V_1 > 0$ and $V_2 > 0$ imply $x_{12} = x_{22} = 0$.

- If $x_{33} \leq 2$, Equations (8.4) and (8.5) give $x_{33} = (2m \bmod 3)$, $x_{13} = \frac{6n-5m+(2m \bmod 3)}{3}$ and $x_{23} = \frac{8m-6n-4(2m \bmod 3)}{3}$. Since $x_{13} \geq 0$, we have $6n - 5m \geq -2$, which implies $G \in F_9$.
- If $x_{33} \geq 3$ then
  - if $x_{13} = 0$, Equations (8.4) and (8.5) give $x_{23} = 6n - 4m$ and $x_{33} = 5m - 6n$. Since $x_{33} \geq 3$, we have $m \geq \frac{6n+3}{5}$, which implies $G \in F_9$.
  - if $x_{13}>0$, consider the (n,m)-preserving vector $A=(0,1,0,-4,3)$ associated with V_5, let $a = \min\{x_{13}, \lfloor \frac{x_{33}}{3} \rfloor\}$, and let ($x'_{12}$, x'_{13}, x'_{22}, x'_{23}, x'_{33}) be the $(A,-a)$-transform of (x_{12}, x_{13}, x_{22}, x_{23}, x_{33}). Then either $x'_{33} \leq 2$, or $x'_{33} \geq 3$ and $x_{13} = 0$. In both cases, we have seen that there is a graph $G' \in F_9$ with x'_{ij} (i,j)-edges. We therefore have $f(G') = f(G) - aV_5 > f(G)$, which contradicts the maximality of G. ■

Corollary 58. *$F_8 \cup F_9$ is the set of extremal chemical graphs for the topological index ABC.*

Proof. It is easy to check that $V_3 > 0$, $V_6 > 0$ and $V_5 + V_7 = 4V_6$ for $\overline{\text{ABC}}$, while $V_1 > 0$, $V_2 > 0$ and $V_5 < 0$ for ABC, which means that F_8 is the set of chemical graphs which minimize ABC, while F_9 is the set of chemical graphs which maximize ABC. ■

Theorem 59. *A chemical graph G maximizes the f=aZagreb topological index over all graphs of the same order and size as G if and only if $G \in F_8$, except in two cases where the x_{ij} values of G are as follows: if n=7 and m=8 then x_{12}=x_{13}=x_{23}=1, $x_{22} = 0$ and x_{33}=5; if n=8 and m=8 then x_{12}=x_{23}=2, x_{13}=1, $x_{22} = 0$ and x_{33}=3.*

Proof. The aZagreb topological index is defined by c_{ij}=$(\frac{ij}{i+j-2})^3$ (see Table 8.1). Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m.

Consider the three (n,m)-preserving vectors A_1, A_2, A_3 associated with the three following strictly positive values W_1, W_2, W_3:

- A_1=$(1,-1,0,-1,1)$ is associated with W_1=$c_{12}-c_{13}-c_{23}+c_{33} \approx 8.01$;
- A_2=$(1,-1,-1,1,0)$ is associated with W_2=$c_{12}-c_{13}-c_{22}+c_{23} \approx 4.62$;
- A_3=$(2,-3,0,2,-1)$ is associated with W_3=$2c_{12}-3c_{13}+2c_{23}-c_{33} \approx 10.48$.

Let

- $(x^1_{12},x^1_{13},x^1_{22},x^1_{23},x^1_{33})$ be the $(A_1, \max\{0, \min\{x_{13}, \frac{x_{23}-x_{12}}{2}\}\})$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$;
- $(x^2_{12}, x^2_{13}, x^2_{22}, x^2_{23}, x^2_{33})$ be the $(A_2, \min\{x^1_{13}, x^1_{22}\})$-transform of $(x^1_{12}, x^1_{13}, x^1_{22}, x^1_{23}, x^1_{33})$;
- $(x^3_{12}, x^3_{13}, x^3_{22}, x^3_{23}, x^3_{33})$ be the $(A_3, \lfloor \frac{x^2_{13}}{3} \rfloor)$-transform of $(x^2_{12}, x^2_{13}, x^2_{22}, x^2_{23}, x^2_{33})$.

Note that if $x^2_{13} > 0$, then $x^2_{22} = 0$ and $x^2_{12} = x^2_{23}$, which implies $x^2_{33} > 0$, else G has order $n \leq 6$. We then have $x^3_{13} \leq 2$ and if $x^3_{13} = 0$, then $x^3_{12} = x^3_{23}$ and $x^3_{22} = 0$. There are therefore only 3 possible cases for which we can derive the x^3_{ij} values using Equations (8.4) and (8.5):

(1) if $x^3_{13} = 0$, then $G \in F_8$. Indeed the proof of Theorem 56 uses the fact that $V_3 > 0$ only to show that $x_{13} = 0$, and it is easy to check that $V_6 > 0$ and $V_5 + V_7 = 4V_6$ for the aZagreb topological index. Therefore,
   - if $x_{13} = 0$, then all x_{ij} values are equal to the x^3_{ij} values and as in Theorem 56, we conclude that $G \in F_8$;
   - if $x_{13} > 0$, then as in Theorem 56, we know that there exists a graph $G' \in F_8$ of order n and size m which contradicts the maximality of G since

$$\begin{aligned} f(G') \geq & f(x^3_{12}, x^3_{13}, x^3_{22}, x^3_{23}, x^3_{33}) \\ = & f(G) + \max\left\{0, \min\{x_{13}, \frac{x_{23}-x_{12}}{2}\}\right\} W_1 \\ & + \min\{x^1_{13}, x^1_{22}\} W_2 + \lfloor \frac{x^2_{13}}{3} \rfloor W_3 \\ > & f(G). \end{aligned}$$

(2) if $x^3_{13} = 1$, $x^3_{12} = x^3_{23}$ and x^3_{22}=0, then Equations (8.4) and (8.5) give $x^3_{12} = x^3_{23} = \frac{3n-2m-2}{3}$ and $x^3_{33} = \frac{7m-6n+1}{3}$, which implies $m \bmod 3 = 2$.

(3) if $x^3_{13} = 2$, $x^3_{12} = x^3_{23}$ and x^3_{22}=0, then Equations (8.4) and (8.5) give $x^3_{12} = x^3_{23} = \frac{3n-2m-4}{3}$ and $x^3_{33} = \frac{7m-6n+2}{3}$, which implies $m \bmod 3 = 1$.

Let us analyze the situation according to the value of $m \bmod 3$:

- if $m \bmod 3 = 0$ then $G \in F_8$ (since we are in Case (1));
- if $m \bmod 3 = 1$, then
  - if $m \geq 10$ or $m + 1 = n = 8$, then Case (1) is better than Case (3) since $c_{12} - 2c_{13} + c_{22} + c_{23} - c_{33} \approx 5.85 > 0$. Hence, $G \in F_8$.
  - if $m = n = 7$, then Case (1) is better than Case (3) since $c_{12} - 2c_{13} + 3c_{23} - 2c_{33} \approx 2.46 > 0$. Hence $G \in F_8$.
- if $m \bmod 3 = 2$,
  - if $m \geq 11$ or $m + 1 = n = 9$ then Case (1) is better than Case (2) since $-c_{13} + 2c_{22} - c_{33} \approx 1.23 > 0$. Hence $G \in F_8$.

– if $m = 8$ and $n \in \{7, 8\}$ then Case (2) is better than Case (1) since $c_{13} - c_{22} - 2c_{23} + 2c_{33} \approx 2.15 > 0$. Moreover, the x_{ij} values of G are equal to the x^3_{ij} values else the graph G' with x^3_{ij} (i, j)-edges is such that $f(G') > f(G)$. Hence, $x_{12} = x_{13} = x_{23} = 1$ and $x_{33} = 5$ if $n = 7$ and $x_{12} = x_{23} = 2$, $x_{13} = 1$ and $x_{33} = 3$ if $n = 8$. ■

Definition 16. F_{10} is the set of chemical graphs with the following numbers x_{ij} of (i, j)-edges:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
0	$\frac{6n-5m}{3}$	0	$\frac{8m-6n}{3}$	0	if $n-1 \leq m \leq \frac{6n-2}{5}$ and $m$mod 3=0
0	$\frac{6n-5m-1}{3}$	1	$\frac{8m-6n-2}{3}$	0	if $n-1 \leq m \leq \frac{6n-2}{5}$ and $m$mod 3=1
0	$\frac{6n-5m+1}{3}$	0	$\frac{8m-6n-4}{3}$	1	if $n-1 \leq m \leq \frac{6n-2}{5}$ and $m$mod 3=2
0	0	1	$m-1$	0	if $m = \frac{6n-1}{5}$
0	0	0	$6n-4m$	$5m-6n$	if $\frac{6n}{5} \leq m \leq \frac{3n-3}{2}$

Theorem 60. *A chemical graph G minimizes the f=aZagreb topological index over all graphs of the same order and size as G if and only if $G \in F_{10}$.*

Proof. As in the previous theorem, we have $c_{ij}=(\frac{ij}{i+j-2})^3$. Let G be a chemical graph of order n, size m and with x_{ij} (i, j)-edges. Assume that it minimizes f over all chemical graphs of order n and size m. Note that $G \neq \mathsf{P}_n$ since for $n = m - 1$, the chemical graph G' with $x'_{12} = 1$, $x'_{13} = 2$, $x'_{23} = 1$ and $x'_{22} = m-4$ would have value $f(G') = f(G)-c_{12}+2c_{13}-2c_{22}+c_{23} \approx f(G)-9.25$. Hence, $x_{12} \leq x_{23}$.

Consider the five (n, m)-preserving vectors A_1, A_2, A_3, A_4, A_5 associated with the five following strictly negative values W_1, W_2, W_3, W_4, W_5:

- A_1=$(-1, 1, 1, -1, 0)$ is associated with W_1=$-c_{12}+c_{13}+c_{22}-c_{23}\approx-4.62$;
- A_2=$(0, 1, -2, 0, 1)$ is associated with $W_2 = c_{13} - 2c_{22} + c_{33} \approx -1.23$;
- A_3=$(0, -1, 0, 4, -3)$ is associated with W_3=$-c_{13}+4c_{23}-3c_{33}\approx-5.54$;
- A_4=$(0, 0, -1, 2, -1)$ is associated with W_4=$-c_{22}+2c_{23}-c_{33}\approx-3.39$;
- A_5=$(0, -1, 1, 2, -2)$ is associated with W_5=$-c_{13}+c_{22}+2c_{23}-2c_{33}\approx-2.15$.

Let

- $(x^1_{12}, x^1_{13}, x^1_{22}, x^1_{23}, x^1_{33})$ be the (A_1, x_{12})-transform of $(x_{12},x_{13},x_{22},x_{23},x_{33})$;
- $(x^2_{12}, x^2_{13}, x^2_{22}, x^2_{23}, x^2_{33})$ be the $(A_2, \lfloor \frac{x^1_{22}}{2} \rfloor)$-transform of $(x^1_{12},x^1_{13},x^1_{22},x^1_{23},x^1_{33})$;
- $(x^3_{12}, x^3_{13}, x^3_{22}, x^3_{23}, x^3_{33})$ be the $(A_3, \min\{\lfloor \frac{x^2_{33}}{3} \rfloor, x^2_{13}\})$-transform of $(x^2_{12},x^2_{13}, x^2_{22},x^2_{23},x^2_{33})$;

- $(x^4_{12}, x^4_{13}, x^4_{22}, x^4_{23}, x^4_{33})$ be the $(A_4, \min\{x^3_{22}, x^3_{33}\})$-transform of $(x^3_{12}, x^3_{13}, x^3_{22}, x^3_{23}, x^3_{33})$;
- $(x^5_{12}, x^5_{13}, x^5_{22}, x^5_{23}, x^5_{33})$ be the $(A_5, \min\{x^4_{13}, \lfloor \frac{x^4_{33}}{2} \rfloor\})$-transform of $(x^4_{12}, x^4_{13}, x^4_{22}, x^4_{23}, x^4_{33})$.

We then have $x^5_{12} = 0, x^5_{22} \leq 1$, $x^5_{13} = 0$ or $x^5_{33} \leq 1$, and $x^4_{22} + x^4_{33} \leq 1$. There are therefore only 5 possible cases for which we can derive the x^5_{ij} values using Equations (8.4) and (8.5):

- if $x^5_{22} = x^5_{33} = 0$ and $x^5_{13} \geq 1$, then $x^5_{13} = \frac{6n-5m}{3}$ and $x^5_{23} = \frac{8m-6n}{3}$, which implies $6n - 5m \geq 3$ and $m \bmod 3 = 0$;
- if $x^5_{22} = 1, x^5_{33} = 0$ and $x^5_{13} \geq 1$, then $x^5_{13} = \frac{6n-5m-1}{3}$ and $x^5_{23} = \frac{8m-6n-2}{3}$, which implies $6n - 5m \geq 4$ and $m \bmod 3 = 1$;
- if $x^5_{22} = 0, x^5_{33} = 1$ and $x^5_{13} \geq 1$, then $x^5_{13} = \frac{6n-5m+1}{3}$ and $x^5_{23} = \frac{8m-6n-4}{3}$, which implies $6n - 5m \geq 2$ and $m \bmod 3 = 2$;
- if $x^5_{22} = x^5_{13} = 0$ and $x^5_{33} \geq 0$, then $x^5_{33} = 5m - 6n$ and $x^5_{23} = 6n - 4m$, which implies $6n - 5m \leq 0$;
- if $x^5_{22} = 1$ and $x^5_{13} = x^5_{33} = 0$, then $x^5_{33} = 5m-6n+1$ and $x^5_{23} = 6n-4m-2$, which implies $6n - 5m = 1$.

Hence all possible x^5_{ij} values correspond to those in F_{10}. So, let G' be a graph with x^5_{ij} (i,j)-edges. We have

$$\begin{aligned} f(G') =& f(G) + x_{12}W_1 + \lfloor \frac{x^1_{22}}{2} \rfloor W_2 + \min\{\lfloor \frac{x^2_{33}}{3} \rfloor, x^2_{13}\} W_3 \\ &+ \min\{x^3_{22}, x^3_{33}\} W_4 + \min\{x^4_{13}, \lfloor \frac{x^4_{33}}{2} \rfloor\} W_5. \end{aligned}$$

If G does not belong to F_{10} then at least one of the five values x_{12}, $\lfloor \frac{x^1_{22}}{2} \rfloor$, $\min\{\lfloor \frac{x^2_{33}}{3} \rfloor, x^2_{13}\}$, $\min\{x^3_{22}, x^3_{33}\}$ and $\min\{x^4_{13}, \lfloor \frac{x^4_{33}}{2} \rfloor\}$ is strictly positive, which implies $f(G') < f(G)$, a contradiction. ■

Definition 17. Family F_{11} is obtained from F_{10} by adding the following possible values:

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
1	$\frac{6n-5m-4}{3}$	0	$\frac{8m-6n+1}{3}$	0	if $n-1 \leq m \leq \frac{6n-2}{5}$ and $m \bmod 3 = 1$
0	$\frac{6n-5m+2}{3}$	0	$\frac{8m-6n-8}{3}$	2	
0	1	0	$m-3$	2	if $m = \frac{6n-1}{5}$

Theorem 61. *A chemical graph G maximizes the f=Albertson topological index over all graphs of the same order and size as G if and only if $G \in F_{11}$.*

Proof. The Albertson topological index is defined by $c_{ij} = |i - j|$ (see Table 8.1). Let G be a chemical graph of order n, size m and with x_{ij} (i, j)-edges. Assume that it maximizes f over all chemical graphs of order n and size m. Note that the two possible cases for $m = \frac{6n-1}{5}$ have the same value since $c_{13} - c_{22} - 2c_{23} + 2c_{33} = 0$. Also, the case in F_{10} for $m \leq \frac{6n-2}{5}$ and $m\text{mod } 3 = 1$ has the same value as the two new possibilities in F_{11} since $c_{12} - c_{13} - c_{22} + c_{23} = c_{13} - c_{22} - 2c_{23} + 2c_{33} = 0$. Moreover, $G \neq \mathsf{P}_n$ since for $n = m - 1$, the chemical graph G' with $x'_{12} = 1$, $x'_{13} = 2$, $x'_{23} = 1$ and $x'_{22} = m-4$ would have value $f(G') = f(G) - c_{12} + 2c_{13} - 2c_{22} + c_{23} = f(G) + 4$. Hence, $x_{12} \leq x_{23}$.
Consider the five following (n, m)-preserving vectors $A_1, A_2, A_3, A_4, A_5, A_6$:

- $A_1 = (-2, 3, 0, -2, 1)$ is associated with $-2c_{12} + 3c_{13} - 2c_{23} + c_{33} = 2$;
- A_2=$(-1, 2, -1, -1, 1)$ is associated with $-c_{12}$+$2c_{13}$−c_{22}−c_{23}+c_{33}=2;
- $A_3 = (0, 1, -2, 0, 1)$ is associated with $c_{13} - 2c_{22} + c_{33} = 2$;
- $A_4 = (0, 0, -1, 2, -1)$ is associated with $-c_{22} + 2c_{23} - c_{33} = 2$;
- $A_5 = (-1, 1, 0, 1, -1)$ is associated with $-c_{12} + c_{13} + c_{23} - c_{33} = 2$;
- $A_6 = (0, -1, 0, 4, -3)$ is associated with $-c_{13} + 4c_{23} - 3c_{33} = 2$.

Let

- $(x^1_{12}, x^1_{13}, x^1_{22}, x^1_{23}, x^1_{33})$ be the $(A_1, \lfloor \frac{x_{12}}{2} \rfloor)$-transform of $(x_{12}, x_{13}, x_{22}, x_{23}, x_{33})$;
- $(x^2_{12}, x^2_{13}, x^2_{22}, x^2_{23}, x^2_{33})$ be the $(A_2, \min\{x^1_{12}, x^1_{22}\})$-transform of $(x^1_{12}, x^1_{13}, x^1_{22}, x^1_{23}, x^1_{33})$;
- $(x^3_{12}, x^3_{13}, x^3_{22}, x^3_{23}, x^3_{33})$ be the $(A_3, \lfloor \frac{x^2_{22}}{2} \rfloor)$-transform of $(x^2_{12}, x^2_{13}, x^2_{22}, x^2_{23}, x^2_{33})$;
- $(x^4_{12}, x^4_{13}, x^4_{22}, x^4_{23}, x^4_{33})$ be the $(A_4, \min\{x^3_{22}, x^3_{33}\})$-transform of $(x^3_{12}, x^3_{13}, x^3_{22}, x^3_{23}, x^3_{33})$;
- $(x^5_{12}, x^5_{13}, x^5_{22}, x^5_{23}, x^5_{33})$ be the $(A_5, \min\{x^4_{12}, x^4_{33}\})$-transform of $(x^4_{12}, x^4_{13}, x^4_{22}, x^4_{23}, x^4_{33})$;
- $(x^6_{12}, x^6_{13}, x^6_{22}, x^6_{23}, x^6_{33})$ be the $(A_6, \min\{x^5_{13}, \lfloor \frac{x^5_{33}}{3} \rfloor\})$-transform of $(x^5_{12}, x^5_{13}, x^5_{22}, x^5_{23}, x^5_{33})$.

We then have $x^6_{12} + x^6_{22} \leq 1$, $\min\{x^6_{12}, x^6_{33}\}$=0, $\min\{x^6_{22}, x^6_{33}\}$=0, and $x^6_{13} = 0$ or $x^6_{33} \leq 2$. Hence, there are 6 possible cases for which we can derive the x^6_{ij} values using Equations (8.4) and (8.5):

- if $x^6_{12} = 0$, $x^6_{22} = 1$ and $x^6_{33} = 0$, then $x^6_{13} = \frac{6n-5m-1}{3}$ and $x^6_{23} = \frac{8m-6n-2}{3}$, which implies $6n - 5m \geq 1$ and $m\text{mod } 3 = 1$;
- if $x^6_{12} = 1, x^6_{22} = 0$ and $x^6_{33} = 0$, then $x^6_{13} = \frac{6n-5m-4}{3}$ and $x^6_{23} = \frac{8m-6n+1}{3}$, which implies $6n - 5m \geq 4$ and $m\text{mod } 3 = 1$;

- if $x_{12}^6 = x_{22}^6 = x_{13}^6 = 0$ then $x_{23}^6 = 6n - 4m$ and $x_{33}^6 = 5m - 6n$, which implies $6n - 5m \leq 0$;
- if $x_{12}^6 = x_{22}^6 = x_{33}^6 = 0$ and $x_{13}^6 \geq 1$, then $x_{13}^6 = \frac{6n-5m}{3}$ and $x_{23}^6 = \frac{8m-6n}{3}$, which implies $6n - 5m \geq 3$ and $m \bmod 3 = 0$;
- if $x_{12}^6 = x_{22}^6 = 0$, $x_{33}^6 = 1$ and $x_{13}^6 \geq 1$, then $x_{13}^6 = \frac{6n-5m+1}{3}$ and $x_{23}^6 = \frac{8m-6n-4}{3}$, which implies $6n - 5m \geq 2$ and $m \bmod 3 = 2$;
- if $x_{12}^6 = x_{22}^6 = 0$, $x_{33}^6 = 2$ and $x_{13}^6 \geq 1$, then $x_{13}^6 = \frac{6n-5m+2}{3}$ and $x_{23}^6 = \frac{8m-6n-8}{3}$, which implies $6n - 5m \geq 1$ and $m \bmod 3 = 1$.

Hence all possible x_{ij}^6 values correspond to those in F_{11}. So, let G' be a graph with x_{ij}^6 (i,j)-edges. We have

$$\begin{aligned} f(G') =& f(G) + 2\left(\lfloor\frac{x_{12}}{2}\rfloor + \min\{x_{12}^1, x_{22}^1\}\right) \\ &+ 2\left(\lfloor\frac{x_{22}^2}{2}\rfloor + \min\{x_{22}^3, x_{33}^3\} + \min\{x_{12}^4, x_{33}^4\} + \min\{x_{13}^5, \lfloor\frac{x_{33}^5}{3}\rfloor\}\right) \end{aligned}$$

If G does not belong to F_{11} then at least one of the six values $\lfloor\frac{x_{12}}{2}\rfloor$, $\min\{x_{12}^1, x_{22}^1\}$, $\lfloor\frac{x_{22}^2}{2}\rfloor$, $\min\{x_{22}^3, x_{33}^3\}$, $\min\{x_{12}^4, x_{33}^4\}$, $\min\{x_{13}^5, \lfloor\frac{x_{33}^5}{3}\rfloor\}$ is strictly positive, which implies $f(G') > f(G)$, a contradiction. ■

Definition 18. F_{12} is the set of chemical graphs with the following numbers x_{ij} of (i,j)-edges:

$(n \geq 8)$

x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	
2	0	$m-2$	0	0	if $m = n-1$
0	0	m	0	0	if $m = n$
2	0	$m-9$	2	5	
1	1	$m-8$	1	5	
1	0	$m-7$	3	3	if $m = n+1$
0	1	$m-6$	2	3	
0	0	$m-5$	4	1	
0	0	$3n-2m-1$	2	$3m-3n-1$	if $n+1 < m \leq \frac{3n-3}{2}$
1	0	$3n-2m-3$	1	$3m-3n+1$	

Theorem 62. *A chemical graph G minimizes the f=Albertson topological index over all graphs of the same order and size as G if and only if $G \in F_{12}$.*

Proof. As in the previous theorem, we have $c_{ij} = |i - j|$. Let G be a chemical graph of order n, size m and with x_{ij} (i,j)-edges. Assume that it minimizes f over all chemical graphs of order n and size m. If $m = n - 1$, then $G = \mathsf{P}_n$ since $f(\mathsf{P}_n) = 2$ while $f(G) \geq 6$ if $n_3 > 0$. Also, if $m = n$, then $G = \mathsf{C}_n$ since

$f(\mathsf{C}_n) = 0$ while $f(G) > 0$ if $n_1 + n_3 > 0$ and $m < \frac{3n}{2}$. Hence, in these cases, we have $G \in F_{12}$.

Assume now $m \geq n + 1$. We thus have $x_{12} \leq x_{23}$ and $f(G)$ is an even number at least equal to 2 (since $m < \frac{3n}{2}$). To reach value 2, there are only three possibilities:

- if $x_{13} = 1$ and $x_{12} = x_{23} = 0$, then $x_{22} = 0$ and Equations (8.4) and (8.5) give $x_{22} = 3n - 2m - 2$ and $x_{33} = 3m - 3n + 1$ which implies $m = \frac{3n-2}{2}$, a contradiction.
- if $x_{12} = x_{23} = 1$ and $x_{13} = 0$, then Equations (8.4) and (8.5) give $x_{22} = 3n - 2m - 3$ and $x_{33} = 3m - 3n + 1$, which implies $m > n + 1$ (else $x_{33} = 4$ and $n_3 = 3$ imply $x_{33} = 4 > 3 = \frac{n_3((n_3-1)}{2})$) and $G \in F_{12}$.
- if $x_{12} = 0$, $x_{23} = 2$ and $x_{13} = 0$, then Equations (8.4) and (8.5) give $x_{22} = 3n - 2m - 1$ and $x_{33} = 3m - 3n - 1$, which implies $m > n + 1$ (else $x_{33} = 1$ and $n_3 = 1$ imply $x_{33} = 1 > 0 = \frac{n_3((n_3-1)}{2})$) and $G \in F_{12}$.

Hence, if $m > n + 1$ then $G \in F_{12}$. The remaining case is $m = n + 1$ for which $f(G) \geq 4$. There are only six possibilities to reach the minimum value 4:

- if $x_{12} = x_{23} = 2$ and $x_{13} = 0$, then $x_{22} = m - 9$ and $x_{33} = 5$.
- if $x_{12} = x_{23} = 1$ and $x_{13} = 1$, then $x_{22} = m - 8$ and $x_{33} = 5$.
- if $x_{12} = 1$, $x_{23} = 3$ and $x_{13} = 0$, then $x_{22} = m - 7$ and $x_{33} = 3$.
- if $x_{12} = 0$, $x_{23} = 2$ and $x_{13} = 1$, then $x_{22} = m - 6$ and $x_{33} = 3$.
- if $x_{12} = 0$, $x_{23} = 4$ and $x_{13} = 0$, then $x_{22} = m - 5$ and $x_{33} = 1$.
- if $x_{12} = x_{23} = 0$ and $x_{13} = 2$, then $x_{22} = 0$ and $x_{33} = m - 2$, which implies $n = 6$.

Since G is of order $n \geq 7$, we have $G \in F_{12}$. ■

8.6 Conclusion

Many topological indices have been proposed to study the chemical properties of molecules, and many papers focus on extremal graphs for these indices, each paper dealing with a particular index. We have shown that many of these topological indices have the same extremal properties in the sense that the chemical graphs that maximize or minimize the values of these indices are often the same. Thus, for example, for 29 of these indices, one might expect 58 classes of extremal chemical graphs, while 5 families are sufficient to describe them all. Also, for another example, chemical graphs of even order n for which $x_{13} = \frac{3n-2m}{2}$, $x_{33} = \frac{4m-3n}{2}$ and $x_{12} = x_{22} = x_{23} = 0$ are extremal for 29 topological indices (since these graphs belong to $F_1 \cap F_3$).

Most of the characterizations we have given for extremal graphs are based on a set of 8 values $V_1, \ldots, V_8$. If new topological indices are proposed, it is therefore easy to check whether they have the same extremal properties of the indices studied in this chapter. Note that some degree-based topological indices that we have not analyzed in this chapter do not have any of the stated properties that allow us to characterize their extremal chemical graphs. For example, the reduced reciprocal Randić index (rrRandić) mentioned in [103] and defined by $c_{ij} = \sqrt{(i-1)(j-1)}$ is such that V_1, V_2, V_3 and V_4 are strictly negative. An analysis similar to those performed in Section 8.5 easily shows that the set of extremal chemical graphs of order $n \geq 10$ for the aZagreb index is strictly contained in that for rrRandić.

We believe we have provided a tool to quickly test whether a new topological index has the same extremal properties as many existing indices, thereby offering a partial criterion in response to Ivan Gutman's statement.

Chapter 9

Complete polyhedral description of chemical graphs of maximum degree at most 3

After our work on the different families of degree-based topological indices in Chapter 8, this chapter further generalizes our results. In particular, this chapter gives a complete polyhedral description of chemical graphs with maximum degree at most 3. Section 9.1 introduces the topic and outlines the structure of the chapter. These results also led to the development of ChemicHull, a tool that makes them more accessible; we mention it in this chapter and describe it in more detail in Chapter 11.

The mathematical details are presented in a first article [75], to which we contributed. This first article is not included as its contents are intended for inclusion in the author's thesis. This thesis instead relies on a second paper [22], which was written with chemists as the intended audience, focusing on how the results can be applied without presenting full mathematical details. We also contributed to this second article.

9.1 Introduction

Throughout this chapter, we restrict our attention to chemical graphs with a maximum degree of at most three.

Our approach relies on a polyhedral representation of chemical graphs. Specifically, we associate each chemical graph with a point in a three-dimensional space defined by the numbers of edges classified according to the degrees of their endpoints. A degree-based topological index then corresponds to a linear

function over this space. This polyhedral framework leads to two major consequences. First, the extremal values of any degree-based topological index can be determined by optimizing a linear function over the associated polyhedron, a straightforward task once the polyhedral representation is known. Second, the structure of the polyhedron reveals that only a small number of families of chemical graphs can be extremal, regardless of the chosen index. This observation provides a unified explanation for the recurrent appearance of the same few families as extremal graphs across different indices and implies that certain chemical graphs can never be extremal for any degree-based topological index.

The structure of the chapter is as follows. In Section 9.2, we outline the relevant background from chemical graph theory and geometry, and introduce the proposed polyhedral methodology. Section 9.3 characterizes all families of extremal chemical graphs, irrespective of the particular degree-based topological index under study. Section 9.4 presents examples demonstrating how our approach can be used to recover known results from the literature in a simple and systematic way. Finally, Section 9.5 concludes the chapter with a summary and a discussion of directions for future research.

Throughout the chapter, we have taken particular care to ensure that the exposition is both accessible and self-contained. Numerous examples and reminders of key concepts are provided, with special attention given to graph classes of central importance in chemical graph theory, such as trees and unicyclic graphs. Our goal is to make the proposed framework transparent and readily applicable to researchers studying extremal problems on degree-based topological indices.

9.2 Basics and notations

In what follows, T_n is the tricyclic graph obtained by adding two adjacent vertices linked to the endpoints of a path P_{n-2}. For illustration, P_5, C_6 and T_7 are shown in Figure 9.1.

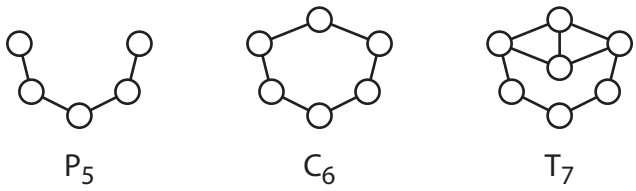


Figure 9.1: Examples of graphs P_n, C_n and T_n.

9.2.1 Chemical graphs

A chemical graph is a simple connected undirected graph in which vertices represent atoms and edges represent the chemical bonds between them. The set of chemical graphs of order n, size m, and maximum degree at most Δ is denoted by $\mathcal{G}_\Delta(n, m)$. For instance, $\mathcal{G}_4(n, n-1)$ represents the set of chemical trees with maximum degree not exceeding 4. Throughout this work, we restrict our attention to chemical graphs with maximum degree at most 3. Accordingly, unless explicitly noted, any chemical graph of order n and size m will be understood to belong to $\mathcal{G}_3(n, m)$.

For a graph G and integers i, j with $1 \leq i \leq j$, an ij-edge in G is an edge whose endpoints have degree i and j. We denote by m_{ij} the number of ij-edges in G. This was previously denoted by $x_{ij}(G)$ in Chapter 7 and by x_{ij} in Chapter 8, we change the notation to fit the paper as it was published. We denote by n_i the number of vertices of degree i in G.

The only chemical graphs of order 1 and 2 are the paths P_1 and P_2, respectively. Accordingly, to exclude trivial instances, we restrict our attention to chemical graphs with a minimum of three vertices, which gives

$$2 \leq n - 1 \leq m \leq \min\left\{\left\lfloor \frac{3n}{2} \right\rfloor, \frac{n(n-1)}{2}\right\}. \tag{9.1}$$

Since chemical graphs are connected, Inequalities (9.1) imply $m_{11} = 0$. Hence, a chemical graph can have at most 5 nonzero m_{ij} values, namely $m_{12}, m_{13}, m_{22}, m_{23}$ and m_{33}. One can readily verify that

$$n_1 = m_{12} + m_{13} \tag{9.2}$$

$$n_2 = \frac{m_{12} + 2m_{22} + m_{23}}{2} \tag{9.3}$$

$$n_3 = \frac{m_{13} + m_{23} + 2m_{33}}{3} \tag{9.4}$$

$$n = n_1 + n_2 + n_3 = \frac{3}{2}m_{12} + \frac{4}{3}m_{13} + m_{22} + \frac{5}{6}m_{23} + \frac{2}{3}m_{33} \tag{9.5}$$

$$m = m_{12} + m_{13} + m_{22} + m_{23} + m_{33}. \tag{9.6}$$

As emphasized in the introduction, our goal is to study the extremal properties of degree-based topological indices of chemical graphs. These indices are computed as the sum of edge weights, where the weight of an edge vw is determined by a formula involving the degrees of v and w.

Definition 19. A *degree-based topological index* I for a chemical graph $G \in \mathcal{G}_3(n, m)$ with $n \geq 3$ is a function of the form

$$I(G) = c_{12}m_{12} + c_{13}m_{13} + c_{22}m_{22} + c_{23}m_{23} + c_{33}m_{33},$$

where every c_{ij} is a real number, and m_{ij} is the number of ij-edges in G.

Note that given n, m and three of the five m_{ij} values, the remaining two can be determined. For example, if m_{12}, m_{13} and m_{33} are known, it follows from Equations (9.5) and (9.6) that

$$m_{22} = 6n - 5m - 4m_{12} - 3m_{13} + m_{33} \tag{9.7}$$
$$m_{23} = 6m - 6n + 3m_{12} + 2m_{13} - 2m_{33}. \tag{9.8}$$

In this chapter, for a pair (n, m) satisfying (9.1), we consider tuples (m_{12}, m_{13}, m_{33}) for which at least one corresponding chemical graph exists.

Definition 20. A point (m_{12}, m_{13}, m_{33}) is said to be *realizable* for a pair (n, m) if there exists a graph $G \in \mathcal{G}_3(n, m)$ of order n and size m containing m_{12} 12-edges, m_{13} 13-edges and m_{33} 33-edges.

9.2.2 Basic notions in geometry

To study the extremal properties of chemical graphs of order n and size m, we consider the convex hull of all realizable points for a given pair (n, m). We briefly recall the definition of a convex hull along with two equivalent representations.

Definition 21. Let $S = \{x_1, \ldots, x_p\}$ be a finite set of p points in $\mathbb{R}^k$. The *convex hull* of S, denoted $\mathrm{conv}(S)$, is the set of all convex combinations of the points in S; explicitly,

$$\mathrm{conv}(S) = \left\{ \sum_{i=1}^{p} \lambda_i x_i \;\middle|\; \lambda_i \geq 0,\ \sum_{i=1}^{p} \lambda_i = 1 \right\}.$$

An *extreme point* of a polyhedron is a point that cannot be expressed as the inner point of any line segment within the polyhedron. Since S is finite, its convex hull forms a *polytope* P which can be expressed in the following two equivalent ways:

1. The *V-representation* of P is the subset $S' \subseteq S$ consisting of its extreme points, meaning that P is the convex hull of S'; that is, $\mathrm{conv}(S) = P = \mathrm{conv}(S')$.

2. The *H-representation* of P is the set of linear inequalities corresponding to the intersection of halfspaces. Explicitly, $\mathrm{conv}(S) = P = \{x \in \mathbb{R}k \mid Ax \leq b\}$ for some matrix A and vector b. This system is *minimal* in the sense that removing any inequality strictly enlarges the corresponding region in $\mathbb{R}k$.

Thus, the polytope $P = \text{conv}(S)$ can be seen either as the smallest convex set containing its points or as the intersection of all half-spaces that include S. A *facet-defining inequality* specifies a *facet* of P, i.e., a face of maximal dimension $k - 1$. Some valid inequalities, however, may be redundant: they hold for all points of P without defining new facets. Facet-defining inequalities thus provide a minimal and exact description of the polytope via its bounding hyperplanes

Example 1. Let $S = \{(0,0), (1,0), (1,1), (1,2), (2,0), (2,1), (3,0), (3,1)\}$ be a set of points in $\mathbb{R}2$. The gray polytope in Figure 9.2 corresponds to $\text{conv}(S)$, its V-representation is $\{(0, 0), (1, 2), (3, 0), (3, 1)\}$ and its H-representation consists of the following four inequalities:

$$x_1 \leq 3, \tag{9.9}$$
$$-x_2 \leq 0, \tag{9.10}$$
$$x_1 + 2x_2 \leq 5, \tag{9.11}$$
$$x_1 - 2x_2 \leq 0. \tag{9.12}$$

The inequality $-x_1 + 2x_2 \leq 4$ is valid for all points of S but is redundant (the inequality defines a half-plane lying below the dotted line in Figure 9.2).

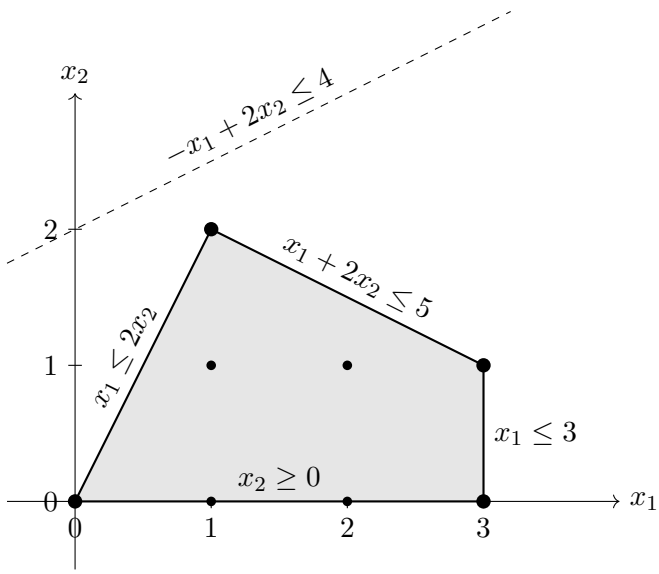


Figure 9.2: A 2D polytope illustrating facet-defining inequalities and a valid (but redundant) inequality (dashed line).

Let S denote the set of all feasible solutions of an integer optimization problem. A point is an *interior point* of the polyhedron conv(S) if it strictly satisfies every inequality in the H-representation of conv(S). Standard results from integer programming (see, for example [216]) show that a linear objective function never attains its optimum over S at an interior point of conv(S). Instead, its optimum over S is achieved at least at one extreme point of conv(S). Hence, knowing the V-representation of conv(S) is sufficient to identify at least one optimal solution.

As an illustration, in Example 1, if $f(x_1, x_2) = x_2$, the maximum occurs only at $(1, 2)$, while the minimum is reached at all points on the facet $x_2 \geq 0$, namely $\{(0, 0), (1, 0), (2, 0), (3, 0)\}$ with $(0, 0)$ and $(3, 0)$ being extreme points. The points $(1, 1)$ and $(2, 1)$ are interior to conv(S) and do not yield an optimum over S for any linear objective function.

9.2.3 Polytopes of chemical graphs

Definition 20 maps each chemical graph in $\mathcal{G}_3(n, m)$ to a point in $\mathbb{R}^3$. We now define the polytope induced by all points in $\mathbb{R}^3$ that are realizable for a given pair (n, m).

Definition 22. Let n and m be two integers satisfying (9.1). The associated polytope $\mathcal{P}_{n,m}$ is the convex hull of all points (m_{12}, m_{13}, m_{33}) that are realizable for (n, m).

According to this definition, we consider a *full-dimensional* polytope $\mathcal{P}_{n,m}$ to have dimension 3. Note that a single point in $\mathcal{P}_{n,m}$ may correspond to multiple chemical graphs. For instance, the point $(0, 4, 4)$ in $\mathcal{P}_{8,8}$ corresponds, by Definitions 20 and 22 and Equations (9.7) and (9.8), to all chemical graphs of order $n = 8$, size $m = 8$, and with $m_{12} = 0$, $m_{13} = 4$, $m_{22} = m_{23} = 0$ and $m_{33} = 4$. There are two non-isomorphic chemical graphs[1] corresponding to these values, shown in Figure 9.3. As another example, for any even n, the point $(0, 0, m = \frac{3n}{2})$ corresponds to all cubic graphs of order n.

Figure 9.3: Two non-isomorphic chemical graphs with coordinates $(0, 4, 4)$ in $\mathcal{P}_{8,8}$.

[1]The list of non-isomorphic graphs with these values is exhaustive, as computed with the help of the discovery system PHOEG [66].

Although a given point may correspond to more than one chemical graph, the knowledge of $\mathcal{P}_{n,m}$ is sufficient to bound the invariants that can be expressed as linear functions of m_{12}, m_{13} and m_{33}. As noted earlier, it is enough to focus on the vertices or the facets of the polytope. This applies in particular to degree-based topological indices. Indeed, for a degree-based topological index I defined by coefficients c_{ij}, we have, by Definition 19 and Equations (9.7) and (9.8):

$$\begin{aligned} I(G) =& c_{12}m_{12} + c_{13}m_{13} + c_{22}m_{22} + c_{23}m_{23} + c_{33}m_{33} \\ =& c_{12}m_{12} + c_{13}m_{13} + c_{22}(6n - 5m - 4m_{12} - 3m_{13} + m_{33}) \\ & + c_{23}(6m - 6n + 3m_{12} + 2m_{13} - 2m_{33}) + c_{33}m_{33} \\ =& (c_{12} - 4c_{22} + 3c_{23})m_{12} + (c_{13} - 3c_{22} + 2c_{23})m_{13} \\ & + (c_{22} - 2c_{23} + c_{33})m_{33} + (6n - 5m)c_{22} + (6m - 6n)c_{23}. \end{aligned}$$

Definition 23. Let I be a degree-based topological index defined by c_{ij} values. We define

- $\widehat{I}(m_{12}, m_{13}, m_{33}) = c'_{12}m_{12} + c'_{13}m_{13} + c'_{33}m_{33}$, where
  - $c'_{12} = c_{12} - 4c_{22} + 3c_{23}$,
  - $c'_{13} = c_{13} - 3c_{22} + 2c_{23}$,
  - $c'_{33} = c_{22} - 2c_{23} + c_{33}$,
- $C_I(n, m) = (6n - 5m)c_{22} + (6m - 6n)c_{23}$.

Note that for a chemical graph G of order n, size m, and with m_{ij} ij-edges, we have $I(G) = \widehat{I}(m_{12}, m_{13}, m_{33}) + C_I(n, m)$. Hence, maximizing or minimizing $I(G)$ over all graphs of order n and size m is equivalent to maximizing or minimizing the linear function $\widehat{I}(m_{12}, m_{13}, m_{33})$. The optimal value of $I(G)$ is then obtained by adding $C_I(n, m)$ to the previously determined optimum.

Example 2. Maximizing or minimizing the Randić index R for chemical graphs of order n and size m reduces to optimizing the linear function

$$\begin{aligned} \widehat{R}(m_{12}, m_{13}, m_{33}) &= \frac{\sqrt{2} + \sqrt{6} - 4}{2} m_{12} + \frac{2\sqrt{3} + 2\sqrt{6} - 9}{6} m_{13} + \frac{5 - 2\sqrt{6}}{6} m_{33} \\ &\simeq -0.068 m_{12} - 0.106 m_{13} + 0.016 m_{33}. \end{aligned}$$

Once the optimal values m^*_{12}, m^*_{13} and m^*_{33} are determined, one can simply add

$$\begin{aligned} C_R(n, m) &= \frac{6n - 5m}{2} + \frac{6m - 6n}{\sqrt{6}}, \\ &= (3 - \sqrt{6})n + \left(\sqrt{6} - \frac{5}{2}\right) m, \end{aligned}$$

$$\simeq 0.55n - 0.05m,$$

to $\widehat{R}(m_{12}^*, m_{13}^*, m_{33}^*)$ to get the maximum (resp. minimum) value of R.

By abuse of notation, we denote by $I(p)$ the value of $I(G)$ for any graph G associated with a realizable point $p = (m_{12}, m_{13}, m_{33})$.

As indicated in Section 9.2, since $\widehat{I}$ is linear, it is sufficient to know the V-representation of $\mathcal{P}_{n,m}$ in order to maximize or minimize $\widehat{I}$. That is what is done in the following section.

9.3 Extreme points

Our recent chapter [75] provides a complete description of all possible polytopes $\mathcal{P}_{n,m}$ for any pair (n, m) satisfying (9.1). In this section, we present the principal results of that study from the viewpoint of their applicability to chemistry, particularly for researchers interested in determining the chemical graphs that maximize or minimize a given degree-based topological index. Accordingly, we concentrate on the identification of the extreme points of these polytopes for a fixed pair (n, m). Remarkably, the number of such points remains very small, even when n and m are arbitrarily large, never exceeding 16.

Let's first focus on the polytopes $\mathcal{P}_{n,m}$ corresponding to the pairs (n, m) that satisfy the following condition:

$$\max\{12, n-1\} \leq m \leq \left\lfloor \frac{3n-3}{2} \right\rfloor. \tag{9.13}$$

As proved in [75], for any such polytope $\mathcal{P}_{n,m}$, its set of extreme points is a subset of the 21 points whose coordinates are given in Table 9.1. Table 9.2 lists the extreme points of these polytopes all of which are full-dimensional. Note that, in all cases shown in Table 9.2, the extreme points are identical for both even and odd values of n; the only exception is the point V9, which appears exclusively when n is odd.

The extreme points of the remaining polytopes that do not satisfy condition (9.13) (i.e., the polytopes for chemical graphs with at least two but fewer than twelve edges, or with more than $\left\lfloor \frac{3n-3}{2} \right\rfloor$ edges) are given in Table 9.3.

Example 3. Suppose we are interested in determining the optimal values of a degree-based topological index I for chemical graphs with 8 vertices and 8 edges. As indicated in Table 9.3, the polytope $\mathcal{P}_{8,8}$ contains only seven extreme points, namely, $(0, 0, 0)$, $(0, 2, 0)$, $(0, 3, 1)$, $(0, 4, 4)$, $(2, 0, 0)$, $(2, 0, 1)$, and $(2, 1, 3)$. To determine the optimal values of the index I, it is sufficient to evaluate $\widehat{I}(p)$ for these seven points p. For example, if I is the Randić

Table 9.1: Twenty-one points (m_{12}, m_{13}, m_{33})

Id	m_{12}	m_{13}	m_{33}
V1	0	0	0
V2	2	0	0
V3	0	0	1
V4	0	0	$5m-6n$
V5	$\frac{6n-5m-3((m-2n) \bmod 4)}{4}$	$(m-2n) \bmod 4$	0
V6	$\frac{6n-5m+(m-2n) \bmod 4}{4}$	0	$(m-2n) \bmod 4$
V7	$\frac{6n-5m-(2n-m) \bmod 4}{4}$	0	0
V8	0	$\frac{3n-2m-n \bmod 2}{2}$	$\frac{4m-3n-3(n \bmod 2)}{2}$
V9	1	$\frac{3n-2m-3}{2}$	$\frac{4m-3n-1}{2}$
V10	$3m-3n-2$	$3m-3n-2$	$6m-6n-1$
V11	$3m-3n-1$	0	$6m-6n-1$
V12	0	$3m-3n-2$	$6m-6n-3$
V13	0	$\frac{6n-5m-m \bmod 3}{3}$	0
V14	$m \bmod 3$	$\frac{6n-5m-4(m \bmod 3)}{3}$	0
V15	0	$\frac{6n-5m+(2m) \bmod 3}{3}$	$(2m) \bmod 3$
V16	$\frac{3n-2m-m \bmod 3}{3}$	0	$\frac{7m-6n-4(m \bmod 3)}{3}$
V17	$\frac{3n-2m-2((2m) \bmod 3)}{3}$	$(2m) \bmod 3$	$\frac{7m-6n+(2m) \bmod 3}{3}$
V18	$\frac{3n-2m-m \bmod 3}{3}$	0	$\frac{7m-6n-m \bmod 3}{3}$
V19	0	$3n-3m+1$	0
V20	0	0	$3m-3n-1$
V21	1	0	$3m-3n+1$

index R, the formula in Example 2 gives $\widehat{R}(0,0,0) = 0$, $\widehat{R}(0,2,0) \simeq -0.2123$, $\widehat{R}(0,3,1) \simeq -0.3016$, $\widehat{R}(0,4,4) \simeq -0.3573$, $\widehat{R}(2,0,0) \simeq -0.1362$, $\widehat{R}(2,0,1) \simeq -0.1194$, and $\widehat{R}(2,1,3) \simeq -0.1919$.

These observations show that the chemical graphs maximizing the Randić index all correspond to the point $(0,0,0)$, while those minimizing it correspond to the point $(0,4,4)$. One can readily verify that the unique graph with 8 vertices and 8 edges associated with $(0,0,0)$ is the cycle C_8. Likewise, the two graphs depicted in Figure 9.3 are precisely those that achieve the minimum Randić index among all chemical graphs on 8 vertices with 8 edges. In contrast, the graph shown in Figure 9.4 corresponds to the point $(1,1,1)$, which neither

Table 9.2: Extreme points for $\mathcal{P}_{n,m}$ when the pair (n, m) satisfies condition (9.13)

$m = n - 1$

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	n mod 2
	x			x	x	x	x					x	x	x	x	x	x	x			0
	x			x	x	x	x	x				x	x	x	x	x	x	x			1

$m = n$

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	n mod 2
x				x	x	x	x					x	x	x	x	x	x				0
x				x	x	x	x	x				x	x	x	x	x	x				1

$m = n + 1$

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	n mod 2
x		x		x	x	x	x		x	x	x	x	x	x	x	x	x				0
x		x		x	x	x	x	x	x	x	x	x	x	x	x	x	x				1

$n + 1 < m < \frac{6n}{5}$

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	n mod 2	$6n-5m \in \{1,2,5\}$
x				x	x	x	x					x	x	x	x	x	x		x	x	0	
x					x	x	x					x		x	x	x	x		x	x	0	✓
x				x	x	x	x	x				x	x	x	x	x	x		x	x	1	
x					x	x	x	x				x		x	x	x	x		x	x	1	✓

$\frac{6n}{5} \leq m \leq \lfloor \frac{3n-3}{2} \rfloor$

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	n mod 2
			x				x								x	x	x		x	x	0
			x				x	x							x	x	x		x	x	1

maximizes nor minimizes the Randić index. Note that this point lies in the interior of $\mathcal{P}_{8,8}$, since it strictly satisfies all inequalities in its H-representation (see Figure 11.1). Consequently, the graph in Figure 9.4 is not extremal in $\mathcal{G}_3(8, 8)$, not only for the Randić index, but for any degree-based topological index.

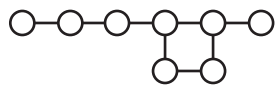


Figure 9.4: A graph in $\mathcal{G}_3(8, 8)$ that is not extremal with respect to any degree-based topological index.

Example 4 (Chemical trees, unicyclic and bicyclic chemical graphs)**.** Table 9.2 shows that, for any even number $n \geq 13$, no matter how large, the polytope $\mathcal{P}_{n,n-1}$ corresponding to chemical trees of order n has only 12 extreme points. A thirteenth extreme point $\text{V9} = \left(1, \frac{n-1}{2}, \frac{n-5}{2}\right)$ must be included when n is odd. Similarly, the polytopes corresponding to chemical unicyclic ($m = n$) and bicyclic ($m = n + 1$) graphs of even order n and size m at least 12 have 11 and 15 extreme points, respectively, with V9 added in each case if n is odd. Thus, as illustrated in Section 9.4, it is sufficient to evaluate only a small number of points to determine a chemical tree, unicyclic, or bicyclic chemical graph that optimizes a degree-based topological index.

Table 9.3: Extreme points for $\mathcal{P}_{n,m}$ when the pair (n, m) does not satisfy condition (9.13)

n	m	extreme points
3	2	(2,0,0)
3	3	(0,0,0)
4	3	(0,3,0), (2,0,0)
4	4	(0,0,0), (0,1,0)
4	5	(0,0,1)
5	4	(1,2,0), (2,0,0)
5	5	(0,0,0), (0,1,0), (0,2,1), (1,0,0)
5	6	(0,0,0), (0,0,1), (0,1,3)
6	5	(0,4,1), (1,2,0), (2,0,0), (2,1,0)
6	6	(0,0,0), (0,2,0), (0,3,3), (1,0,0), (1,1,1)
6	7	(0,0,0), (0,0,1), (0,1,2), (0,2,5), (1,0,3)
7	6	(0,4,0), (1,3,1), (2,0,0), (3,0,0)
7	7	(0,0,0), (0,2,0), (0,3,2), (1,0,0), (1,1,0), (1,2,3), (2,0,1)
7	8	(0,0,0), (0,0,1), (0,1,1), (0,1,3), (0,2,4), (1,0,2), (1,0,3), (1,1,5)
7	9	(0,0,3), (0,0,5), (0,1,6), (1,0,7)
8	7	(0,4,0), (0,5,2), (1,3,0), (2,0,0), (2,2,1), (3,0,0)
8	8	(0,0,0), (0,2,0), (0,3,1), (0,4,4), (2,0,0), (2,0,1), (2,1,3)
8	9	(0,0,0), (0,0,1), (0,1,0), (0,1,3), (0,3,6), (1,0,1), (1,1,5), (2,0,5)
8	10	(0,0,2), (0,0,5), (0,2,8), (1,0,6), (1,0,7)
9	8	(0,4,0), (0,5,1), (1,4,2), (2,0,0), (2,2,0), (3,0,0), (3,1,1)
9	9	(0,0,0), (0,3,0), (0,4,3), (1,3,4), (2,0,0), (3,0,3)
9	10	(0,0,0), (0,0,1), (0,1,0), (0,1,3), (0,2,2), (0,3,5), (1,0,0), (1,1,5), (1,2,6), (2,0,4), (2,0,5)
9	11	(0,0,1), (0,0,5), (0,2,7), (1,0,5), (1,0,7), (1,1,8)
10	9	(0,4,0), (0,5,0), (0,6,3), (2,0,0), (3,0,0), (3,1,0), (4,0,1)
10	10	(0,0,0), (0,3,0), (0,4,2), (0,5,5), (1,2,0), (2,0,0), (2,2,4), (3,0,2), (3,0,3)
10	11	(0,0,0), (0,0,1), (0,1,0), (0,1,3), (0,2,1), (0,4,7), (1,0,0), (1,1,5), (2,0,3), (2,0,5), (2,1,6)
11	10	(0,4,0), (0,5,0), (0,6,2), (1,4,0), (1,5,3), (2,0,0), (3,2,2), (4,0,0), (4,0,1)
11	11	(0,0,0), (0,3,0), (0,4,1), (0,5,4), (1,4,5), (2,0,0), (2,1,0), (3,0,1), (3,0,3), (3,1,4)
12	11	(0,4,0), (0,5,0), (0,6,1), (0,7,4), (2,0,0), (2,3,0), (4,0,0), (4,0,1), (4,1,2)
even ≥ 4	$\frac{3n}{2}$	(0,0,m)
odd ≥ 5	$\frac{3n-1}{2}$	(0,0,$m-2$)
even ≥ 6	$\frac{3n-2}{2}$	(0,0,$m-4$), (0,0,$m-3$), (0,1,$m-1$)

9.4 Examples of ChemicHull applications

In this section, we reproduce established results from the literature to demonstrate the effectiveness of the approach based on the extreme points of the polytopes $\mathcal{P}_{n,m}$. We focus on the following degree-based topological indices:

- the generalized Randić index R_α is defined by $c_{ij} = (ij)^\alpha$ [147]. When $\alpha = -\frac{1}{2}$, this is the standard Randić index introduced in 1975 [183];
- as mentioned in [100], the Sombor index, denoted SO, is defined by $c_{ij} = \sqrt{i^2 + j^2}$, while the reduced Sombor index, denoted SO_{red}, is defined by $c_{ij} = \sqrt{(i-1)^2 + (j-1)^2}$;
- the atom-bond sum-connectivity index, denoted ABS and introduced in [7] is defined by $c_{ij} = \sqrt{\frac{i+j-2}{i+j}}$.

Recall that for a degree-based topological index I, we have $\widehat{I}(m_{12}, m_{13}, m_{33})$

$= c'_{12}m_{12} + c'_{13}m_{13} + c'_{33}m_{33}$, where

- $c'_{12} = c_{12} - 4c_{22} + 3c_{23}$,
- $c'_{13} = c_{13} - 3c_{22} + 2c_{23}$,
- $c'_{33} = c_{22} - 2c_{23} + c_{33}$.

The c'_{ij} values for the above topological indices are given in Table 9.4. The examples presented in the following subsections cover a broad spectrum of cases, as the optimal graphs in $\mathcal{G}_3(n, m)$ correspond to V1, V2, V3, V18 or V20, depending on the chosen topological index and the condition relating m to n. The reader interested in more examples is invited to consult [21], in which the extremal graphs in $\mathcal{G}_3(n, m)$ for 33 topological indices are described.

Table 9.4: Some values associated with five topological indices

	$R_{-\frac{1}{2}}$	R_{-1}	SO	SO_{red}	ABS
c_{ij}	$\frac{1}{\sqrt{ij}}$	$\frac{1}{ij}$	$\sqrt{i^2+j^2}$	$\sqrt{(i-1)^2+(j-1)^2}$	$\sqrt{\frac{i+j-2}{i+j}}$
c'_{12}	$\simeq -0.068$	0	$\simeq 1.739$	$\simeq 2.051$	$\simeq 0.073$
c'_{13}	$\simeq -0.106$	$-\frac{1}{12}$	$\simeq 1.888$	$\simeq 2.229$	$\simeq 0.135$
c'_{33}	$\simeq 0.017$	$\frac{1}{36}$	$\simeq -0.140$	$\simeq -0.229$	$\simeq -0.026$

9.4.1 Conditions for paths to be optimal chemical trees

The following results provide sufficient conditions on a degree-based topological index for P_n to be an optimal graph in $\mathcal{G}_3(n, n-1)$.

Theorem 63. *If I is a degree-based topological index such that $c'_{13} < c'_{12} < -c'_{33} < 0$, then P_n is the only tree in $\mathcal{G}_3(n, n-1)$ that maximizes I.*

Proof. Let I be a degree-based topological index such that $c'_{13} < c'_{12} < -c'_{33} < 0$. We have to prove that V2 $= (2, 0, 0)$ is the unique point that maximizes $\widehat{I}$. We know from Table 9.2 that the candidate points to maximize $\widehat{I}$ over all chemical trees in $\mathcal{G}_3(n, n-1)$ of order at least 13 are V2, V5, V6, V7, V8, V13, V14, V15, V16, V17, V18 and V19 as well as V9 if n is odd. Note that

- $\widehat{I}$(V5) and $\widehat{I}$(V7) are at most $\lceil \frac{n-4}{4} \rceil c'_{12} < 2c'_{12}$;
- $\widehat{I}$(V6) $\leq \lceil \frac{n+5}{4} \rceil c'_{12} + 3c'_{33} \leq 5c'_{12} + 3c'_{33} < 2c'_{12}$;
- $\widehat{I}$(V8) and $\widehat{I}$(V9) are at most $\lceil \frac{n-1}{2} \rceil c'_{13} + \lfloor \frac{n-4}{2} \rfloor c'_{33} \leq (\lceil \frac{n-1}{2} \rceil - \lfloor \frac{n-4}{2} \rfloor) c'_{13} = 2c'_{13} < 2c'_{12}$;
- $\widehat{I}$(V13), $\widehat{I}$(V14) and $\widehat{I}$(V15) are at most $\lceil \frac{n-3}{3} \rceil c'_{13} + 2c'_{33} \leq 4c'_{13} + 2c'_{33} < 2c'_{13} < 2c'_{12}$;

- $\widehat{I}$(V16), $\widehat{I}$(V17) and $\widehat{I}$(V18) are at most $\lceil \frac{n+2}{3} \rceil c'_{12} + \lfloor \frac{n-5}{3} \rfloor c'_{33} < (\lceil \frac{n+2}{3} \rceil - \lfloor \frac{n-5}{3} \rfloor) c'_{12} = 3c'_{12} < 2c'_{12}$;

Since $\widehat{I}(\mathrm{V2}) = 2c'_{12}$, we conclude that for $n \geq 13$, the path is the only chemical tree in $\mathcal{G}_3(n, n-1)$ that maximizes the I index. For $3 \leq n \leq 12$, it is easy to verify from Table 9.3 that (2,0,0) (that is, the path) maximizes $\widehat{I}$. ■

Given that minimizing $\widehat{I}$ is equivalent to maximizing $-\widehat{I}$, we can derive the following corollary.

Corollary 64. *If I is a degree-based topological index such that $c'_{13} > c'_{12} > -c'_{33} > 0$, then P_n is the only tree in $\mathcal{G}_3(n, n-1)$ that minimizes I.*

It is shown in [33] and [105] that P_n maximizes $R_{-\frac{1}{2}}$ in $\mathcal{G}_4(n, n-1)$. Hence, this is also true for chemical trees of maximum degree at most 3. This is a direct corollary of Theorem 63. Indeed, as shown in Table 9.4, for this index, we have

$$c'_{13} \simeq -0.106 < c'_{12} \simeq -0.068 < -c'_{33} = -0.017 < 0.$$

As another example, the authors of [151] have characterized the 14 trees in $\mathcal{G}_4(n, n-1)$ with smallest Sombor and reduced Sombor indices. In particular, they have shown that P_n has minimum value. Hence, this is also true for chemical trees of maximum degree at most 3, and this is a direct consequence of Corollary 64. Indeed, as shown in Table 9.4, for the SO index, we have

$$c'_{13} \simeq 1.888 > c'_{12} \simeq 1.739 > -c'_{33} \simeq 0.140 > 0,$$

while for SO_{red}, we have

$$c'_{13} \simeq 2.229 > c'_{12} \simeq 2.051 > -c'_{33} \simeq 0.229 > 0.$$

As a further example, consider the ABS index. The authors of [226] show, among other results, that P_n minimizes this index among all chemical trees in $\mathcal{G}_4(n, n-1)$. Hence, this is also true for trees of maximum degree at most 3, and this follows directly from Corollary 64. Indeed, as shown in Table 9.4, for the ABS index, we have

$$c'_{13} \simeq 0.135 > c'_{12} \simeq 0.073 > -c'_{33} \simeq 0.026 > 0.$$

9.4.2 Conditions for cycles to be optimal unicyclic chemical graphs

The following results provide sufficient conditions on a degree-based topological index for C_n to be an optimal graph in $\mathcal{G}_3(n, n)$.

Theorem 65. *If I is a degree-based topological index such that* $\max\{0, -c'_{33}\} < \min\{c'_{12}, c'_{13}\}$*, then* C_n *is the only graph in* $\mathcal{G}_3(n,n)$ *that minimizes* I*.*

Proof. Let I be a degree-based topological index such that $\max\{0, -c'_{33}\} < \min\{c'_{12}, c'_{13}\}$. We have to prove that $\mathrm{V1} = (0,0,0)$ is the unique point that minimizes $\widehat{I}$. We know from Table 9.2 that the candidate points to minimize $\widehat{I}$ over all graphs in $\mathcal{G}_3(n,n)$ of order at least 12 are V1, V5, V6, V7, V8, V13, V14, V15, V16, V17 and V18 as well as V9 if n is odd. It is easy to observe that $\widehat{I}(\mathrm{V}i)$ is strictly positive for all these points, except for V1 for which $\widehat{I}(\mathrm{V1}) = 0$. For smaller values of n, one can check from Table 9.3 that $(0,0,0)$ is always the point with minimum I index value. We therefore conclude that for $n \geq 3$, C_n is the only graph in $\mathcal{G}_3(n,n)$ that minimizes the I index. ■

Given that maximizing $\widehat{I}$ is equivalent to minimizing $-\widehat{I}$, this leads to the following corollary.

Corollary 66. *If I is a degree-based topological index such that* $\max\{c'_{12}, c'_{13}\} < \min\{0, -c'_{33}\}$*, then* C_n *is the only chemical graph in* $\mathcal{G}_3(n,n)$ *that maximizes* I*.*

It is shown in [34] that C_n maximizes $R_{-\frac{1}{2}}$ over all graphs in $\mathcal{G}_4(n,n)$. Hence, this is also true for the graphs in $\mathcal{G}_3(n,n)$. This follows from Corollary 66. Indeed, as shown in Table 9.4, for the $R_{-\frac{1}{2}}$ index, we have

$$\max\{c'_{12}, c'_{13}\} = c'_{12} \simeq -0.068 < -c'_{33} \simeq -0.017 < 0.$$

As a second example, it is shown in [226] that C_n minimizes the ABS index in $\mathcal{G}_4(n,n)$. Hence, this is also true for graphs in $\mathcal{G}_3(n,n)$, and this is a direct consequence of Theorem 65. Indeed, for this index, we have

$$\min\{c'_{12}, c'_{13}\} = c'_{12} \simeq 0.073 > -c'_{33} \simeq 0.026 > 0.$$

9.4.3 Trees that maximize R_{-1} in $\mathcal{G}_3(n, n-1)$

The problem of maximizing R_{-1} over all graphs in $\mathcal{G}_3(n, n-1)$ is solved in [184]. We can easily prove the same result with our polyhedral approach. Indeed, assuming $m = n-1$, we get from Table 9.3 that the points that maximize $\widehat{R}_{-1}$ for $3 \leq n \leq 12$ are

- $(2,0,0) = \mathrm{V2}$ if $n \in \{3,4,5,6\}$,
- $(2,0,0) = \mathrm{V2}$ and $(3,0,0) = \mathrm{V18}$ if $n \in \{7,8,9\}$, and
- $(4,0,1) = \mathrm{V18}$ if $n \in \{10,11,12\}$.

For $n \geq 13$, we know from Table 9.2 that the only candidate points to maximize $\widehat{R}_{-1}$ for graphs in $\mathcal{G}_3(n, n-1)$ are V2, V5, V6, V7, V8, V13, V14, V15, V16, V17, V18 and V19 as well as V9 if n is odd. Clearly, the only candidates Vi for which $\widehat{R}_{-1}(\mathrm{V}i)$ can be strictly positive are V6, V16, V17 and V18, and it is easy to observe that

- $\widehat{R}_{-1}(\mathrm{V16}) = \widehat{R}_{-1}(\mathrm{V17}) = \widehat{R}_{-1}(\mathrm{V18})$ when $n \mod 3 = 1$ and both $\widehat{R}_{-1}(\mathrm{V16})$ and $\widehat{R}_{-1}(\mathrm{V17})$ are strictly smaller than $\widehat{R}_{-1}(\mathrm{V18})$ otherwise;
- $\widehat{R}_{-1}(\mathrm{V6}) = \widehat{R}_{-1}(\mathrm{V18})$ if $n = 13$ or 16, and $\widehat{R}_{-1}(\mathrm{V6}) < \widehat{R}_{-1}(\mathrm{V18})$ otherwise.

Hence, in all cases, $\widehat{R}_{-1}(\mathrm{V}i)$ is maximized with V18 when $n \geq 13$. In summary, V2 maximizes $\widehat{R}_{-1}$ if $n < 7$ while V18 maximizes $\widehat{R}_{-1}$ if $n \geq 7$.

The value $R_{-1}(V)$ of the R_{-1} index of a point V corresponding to a chemical tree of order n is $R_{-1}(V) = \widehat{R}_{-1}(V) + C_{R_{-1}}(n, m)$ where

$$C_{R_{-1}}(n, m) = (6n - 5m)c_{22} + (6m - 6n)c_{23} = \frac{n+5}{4} - 1 = \frac{n+1}{4}.$$

Since $R_{-1}(\mathrm{V2}) = \frac{n+1}{4}$ and

$$R_{-1}(\mathrm{V18}) = \frac{1}{36}\left\lfloor \frac{n-7}{3} \right\rfloor + \frac{n+1}{4} = \frac{7n}{27} + \frac{20 - (n+2) \bmod 3}{108},$$

we can conclude that given any chemical tree T in $\mathcal{G}_3(n, n-1)$, as shown in [184], we have

$$R_{-1}(T) \leq \begin{cases} \frac{n+1}{4} & \text{if } n \in \{3, 4, 5, 6\}, \\ \frac{7n}{27} + \frac{1}{6} & \text{if } n \geq 7 \text{ and } n \equiv 0 \bmod 3, \\ \frac{7n}{27} + \frac{5}{27} & \text{if } n \geq 7 \text{ and } n \equiv 1 \bmod 3, \\ \frac{7n}{27} + \frac{19}{108} & \text{if } n \geq 7 \text{ and } n \equiv 2 \bmod 3. \end{cases}$$

9.4.4 More optimal values for the $R_{-\frac{1}{2}}$ and ABS indices

Let's consider bicyclic graphs. It is proved in [34] that over all graphs in $\mathcal{G}_4(n, n+1)$, V3 maximizes $R_{-\frac{1}{2}}$, and in [226] that the same point minimizes the ABS index. Hence, this is also true when the maximum degree is at most 3, and this can be proved with our polyhedral approach. Indeed, for bicyclic chemical graphs with at least 11 vertices, Table 9.2 shows that the candidate points to achieve an optimal value are V1, V3, V5, V6, V7, V8, V10, V11, V12, V13, V14, V15, V16, V17 and V18 as well as V9 if n is odd.

- It is easy to observe that $\widehat{R}_{-\frac{1}{2}}(\mathrm{V}i)$ is less than or equal to zero for all these points, except for V3 for which $\widehat{R}_{-\frac{1}{2}}(\mathrm{V3}) \simeq 0.017$. For smaller

values of n, it is easy to check from Table 9.3 that $(0,0,1)$ is always the point with maximum $R_{-\frac{1}{2}}$ index.

- It is easy to observe that $\widehat{ABS}(\mathrm{V}i)$ is non-negative for all these points, except for V3 for which $\widehat{ABS}(\mathrm{V3}) \simeq -0.026$. For smaller values of n, it is easy to check from Table 9.3 that $(0,0,1)$ is always the point with minimum ABS index.

Let's now consider tricyclic graphs. It is argued in [34] (without proof) that the graph that maximizes $R_{-\frac{1}{2}}$ over all graphs in $\mathcal{G}_4(n, n+2)$ has 4 edges with both endpoints of degree 3 and no vertex of degree 1 or 4. This corresponds to the point $(0,0,4)$ in $\mathcal{G}_3(n, n+2)$. This is however wrong. Indeed, we prove here below that the tricyclic graph T_n (i.e., V20 $= (0,0,5)$) is the unique graph in $\mathcal{G}_3(n, n+2)$ that maximizes $R_{-\frac{1}{2}}$ in $\mathcal{G}_3(n, n+2)$:

- For graphs in $\mathcal{G}_3(n, n+2)$ with at least 11 vertices, Table 9.2 shows that the candidate points to achieve the maximum value are V1, V6, V7, V8, V13, V15, V16, V17, V18, V20 and V21, as well as V5 and V14 if $n \notin \{11, 12, 15\}$ and V9 if n is odd (since $n+1 < n+2 < \frac{6n}{5}$). It is easy to observe that $\widehat{R}_{-\frac{1}{2}}(\mathrm{V}i)$ is non-positive for all these points, except $\widehat{R}_{-\frac{1}{2}}(\mathrm{V21}) = \widehat{R}_{-\frac{1}{2}}(1,0,7) = c'_{12} + 7c'_{33} \simeq 0.05$ and $\widehat{R}_{-\frac{1}{2}}(\mathrm{V20}) = \widehat{R}_{-\frac{1}{2}}(0,0,5) \simeq 0.08$.

- For $n = 10$, we have $m = n+2 = \frac{6n}{5}$, and Table 9.2 shows that the candidate points are V4, V8, V16, V17, V18, V20 and V21, as well as V9 if n is odd. Since V4 $=$ V1 $= (0,0,0)$, we have observed in the previous case that V20 $= (0,0,5)$ maximizes $R_{-\frac{1}{2}}$.

- For $5 \leq n \leq 9$, it is easy to check from Table 9.3 that $(0,0,5)$ is also always the point with maximum $R_{-\frac{1}{2}}$ index.

We therefore conclude that for $n \geq 5$, V20 maximizes the $R_{-\frac{1}{2}}$ index over all tricyclic chemical graphs.

Let's conclude this section with the graphs in $\mathcal{G}_3(n, n+2)$ that minimize the ABS index. It is proved in [226] that for a maximum degree at most 4, the optimal tricyclic graph is V20. Hence this is also true when the maximum degree is at most 3 which we now prove with our polyhedral approach. Recall that the candidate points to achieve the minimum value are V1, V6, V7, V8, V13, V15, V16, V17, V18, V20 and V21, as well as V5 and V14 if $n \notin \{11, 12, 15\}$ and V9 if n is odd. Note that $\widehat{ABS}(\mathrm{V20}) \simeq -0.128$ and $\widehat{ABS}(\mathrm{V21}) \simeq -0.106$,

and it is easy to observe that $\widehat{ABS}(\mathrm{V}i)$ is non-negative for all the other points, except in the following cases:

- if $n = 10$, $\widehat{ABS}(\mathrm{V17})$=$\widehat{ABS}(\mathrm{V18})$=$\widehat{ABS}(2,0,8) = 2c'_{12} + 8c'_{33} > 5c'_{33}$;
- if $n = 11$, $\widehat{ABS}(\mathrm{V6}) = \widehat{ABS}(1,0,3) = c'_{12} + 3c'_{33} > 5c'_{33}$, $\widehat{ABS}(\mathrm{V16}) = \widehat{ABS}(2,0,7) = 2c'_{12} + 7c'_{33} > 5c'_{33}$, and $\widehat{ABS}(\mathrm{V18}) = \widehat{ABS}(2,0,8) = 2c'_{12} + 8c'_{33} > 5c'_{33}$;
- if $n = 12$, $\widehat{ABS}(\mathrm{V16})$=$\widehat{ABS}(2,0,6) = 2c'_{12} + 6c'_{33} > 5c'_{33}$, and $\widehat{ABS}(\mathrm{V18})$=$\widehat{ABS}(2,0,8) = 2c'_{12} + 8c'_{33} > 5c'_{33}$;
- if $n = 13$, $\widehat{ABS}(\mathrm{V16})$=$\widehat{ABS}(\mathrm{V17})$=$\widehat{ABS}(\mathrm{V18})$=$\widehat{ABS}(3,0,9) = 3c'_{12} + 9c'_{33} > 5c'_{33}$;
- if $n = 14$ or 15, $\widehat{ABS}(\mathrm{V18})$=$\widehat{ABS}(3,0,9) = 3c'_{12} + 9c'_{33} > 5c'_{33}$.

We deduce that for $n \geq 10$, V20 minimizes the ABS index over all graphs in $\mathcal{G}_3(n, n+2)$. For $5 \leq n \leq 9$, it is easy to check from Table 9.3 that $(0,0,5)$ is also always the point with minimum ABS index.

9.5 Concluding remarks and future work

In this study, we have shown that the polyhedral approach offers a powerful and systematic framework for identifying families of chemical graphs that are extremal with respect to degree-based topological indices. Beyond reproducing many established results from the literature, this method also allows for their verification and, when required, correction, as exemplified by our discovery of a counterexample to a previously reported extremal result involving the Randić index.

To facilitate the exploration and application of these results, we have developed an online tool called ChemicHull. It offers an intuitive interface for visualizing extremal chemical graphs and analyzing optimal values of degree-based topological indices. This tool is further described in Chapter 11.

Moreover, our findings indicate that some graphs can never serve as extreme points under any degree-based topological index. Therefore, if a molecule of interest in extremal chemistry is not represented among the extremal points identified in this study, it must instead be described by a topological index beyond the degree-based class.

An important direction for future work is to extend these results to chemical graphs with a maximum degree of at most 4. Preliminary investigations, however, suggest that such an extension will be considerably more complex, both combinatorially and computationally.

Part III

Tools for Extremal & Chemical Graph Theory

Chapter 10

PHOEG Helps Obtaining Extremal Graphs

This chapter is based on an article currently in preparation describing the latest version of PHOEG, an online tool for Extremal Graph Theory. PHOEG enables the exploration of user-defined two-dimensional invariant spaces and serves both educational and research purposes. Section 10.1 presents an overview of PHOEG and details the organization of this chapter.

10.1 Introduction

Extremal Graph Theory studies bounds on graph invariants, which are values preserved under isomorphism, such as order, size, chromatic number, or average distance. These problems are typically subject to constraints like fixing the value of some invariants. Finding these bounds, and the extremal graphs that realize them, is challenging: the sheer number of graphs grows extremely rapidly, exceeding a hundred billion for graphs of order 12 alone.

To tackle this complexity, several computational tools have been developed over the years. Early efforts include Graph by Cvetkovic et al. [48] and its successor newGRAPH by Brankov et al. [198]. Graffiti [80], developed by Fajtlowicz in 1988, used heuristics and pre-computed data to generate thousands of conjectures automatically. AutoGraphiX [35], by Caporossi and Hansen in 2000, applied variable neighborhood search to identify extremal graph candidates. Concurrently, databases like House of Graphs [28, 44, 65] emerged, offering a searchable repository of interesting graphs allowing researchers to query specific invariant values or search for counterexamples. These tools have proven valuable, but they share a common characteristic: they are not exhaustive. Whether through heuristics, sampling, or selective search, they

sacrifice exactness for scalability.

GraPHedron [163], introduced by Mélot in 2008, took a deliberately different stance: rather than approximating, it committed to an exact approach. It embedded all graphs up to a given order into an invariant space and computed their convex hull, interpreting facets as inequalities between invariants and vertices as extremal graphs. Historically, GraPHedron was instrumental in the early days of House of Graphs, as it was used to compute and introduce the very first graphs into the database. However, unlike some of its contemporaries, GraPHedron itself was never updated - motivating the development of its successor, PHOEG (a recursive acronym for PHOEG Helps to Obtain Extremal Graphs).

First introduced in 2019 [66], PHOEG embraces this same goal of exactness. The trade-off is explicit and assumed: PHOEG operates on all pairwise non-isomorphic graphs up to order 10, the practical frontier imposed by combinatorial growth. It combines a database of graphs enabling fast queries and computations with a proof-assistance module called TransProof. Since its initial release, we have significantly expanded PHOEG to democratize its access (`https://phoeg.umons.ac.be`). Through extensive use in both research and educational settings, we refined the tool and developed a comprehensive web interface and API.

This chapter presents these modernizations and illustrates them through a detailed walkthrough of the platform. Section 10.2 briefly recalls the underlying geometrical approach PHOEG is built on. Section 10.3 details the new web interface functionalities and the technologies used to develop them. Section 10.4 serves as a mini-survey of the extremal results obtained over the years using this geometric approach, and illustrates PHOEG's practical applications in research through concrete examples. Section 10.5 details its use in educational settings. Finally, in Section 10.6 we conclude and discuss future directions.

10.2 Geometrical approach

This section explains the geometrical approach underlying PHOEG and introduced in GraPHedron [163].

As stated by Hansen et al. [107], many results in extremal graph theory are expressed in terms of inequalities between graph invariants. This observation guided the design of GraPHedron and remains the core principle in PHOEG.

The geometrical approach used in PHOEG places simple, undirected graphs into a two-dimensional space where each dimension corresponds to a specific graph invariant. Each graph of a given order is represented as a point whose coordinates are the values of the two chosen invariants evaluated on that graph.

It is important to note that a single point in this space may represent several non-isomorphic graphs, as they can share the exact same values for the chosen pair of invariants. The convex hull of all these points forms a convex polytope. Recall that the *convex hull* of a finite set of points in the plane is the smallest convex set enclosing them, and that its boundary consists of *facets*, the edges of the polygon in two dimensions. Consequently, each facet of this convex polytope corresponds to a linear inequality between the invariants (typically of the form $aX + bY \leq c$), and the vertices represent the extremal graphs that tightly achieve these bounds.

An example of such a polytope is shown in Figure 10.1. The two invariants defining the space in this example are the *chromatic number* (the minimum number of colors needed to color the vertices such that no two adjacent vertices share the same color) for the X-axis and the *clique number* (the number of vertices in the largest complete subgraph) for the Y-axis. The points plotted represent all graphs of order 7. To illustrate that a single point can represent several non-isomorphic graphs, consider the coordinates in this figure: the point $(1, 1)$ corresponds uniquely to the empty graph on 7 vertices, and the point $(7, 7)$ corresponds uniquely to the complete graph on 7 vertices. In contrast, the point $(2, 2)$ represents 87 non-isomorphic graphs of order 7, a set that notably includes all 11 non-isomorphic trees on 7 vertices. Finally, this polytope makes an important inequality visually immediate: the chromatic number $\chi(G)$ of a graph G is always greater than or equal to its clique number $\omega(G)$, that is the well-known relation $\chi(G) \geq \omega(G)$.

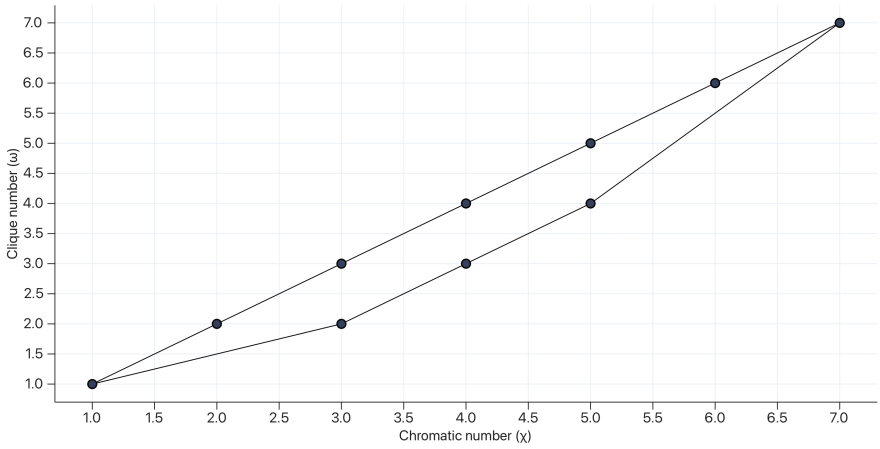


Figure 10.1: Convex polytope formed by all graphs of order 7, with the chromatic number and the clique number used as dimensions

Beyond observing inequalities for specific orders, the ultimate goal of this geometrical approach is to obtain a complete, parameterized polyhedral description valid for any graph order n. When such a general description can be formulated and mathematically proven, it yields the optimal set of linear inequalities governing the relationship between the two invariants. Indeed, by the very definition of a convex hull, its facets constitute the minimal set of bounding linear inequalities required to fully describe that relationship. This complete characterization has been successfully achieved for several pairs of invariants. For instance, it was established for the *diameter* (the maximum eccentricity among all vertices) and the size of the graph in [163], as well as for the Fibonacci index (the total number of stable sets, where a stable set is a subset of pairwise non-adjacent vertices) and the *independence number* (or stability number, that is the size of the largest stable set) in both general and connected graphs in [31]. Similarly, the complete polyhedral description for the independence number and the size of connected graphs was proven in [40]. Notably, this latter result formally resolved an open problem posed by Ore in 1962 [176], which serves as a connected variant of the classical Turán's theorem [206].

10.3 Web interface and API

This section details the current architecture and features of the PHOEG platform. Section 10.3.1 describes the underlying software stack. Section 10.3.2 provides a comprehensive overview of the web interface functionalities, while Section 10.3.3 presents the standalone API available for direct database access.

10.3.1 Technologies

While our research team has utilized the core PHOEG database for years, we recognized the need to make this resource accessible to the broader community, especially to researchers unfamiliar with formulating complex SQL queries. This objective drove the development of a dedicated web interface.

The platform follows a modern, three-tier architecture to ensure a clear separation of concerns: a frontend for user interaction, a backend Application Programming Interface (API) for business logic and query processing, and a relational database for reliable storage.

The frontend is built in TypeScript with React [165], leveraging its component-based architecture for efficient state management and dynamic UI updates. To ensure a clean, responsive, and accessible design, the interface heavily relies on Chakra UI [2], a robust component library.

The backend consists of a RESTful API developed in Rust [202] using the *rocket* framework [18]. This API acts as an intermediary, translating user interactions from the frontend into optimized database queries. Crucially, this API is also exposed for programmatic access independently of the web interface, as further detailed in Section 10.3.3.

The data layer is powered by a PostgreSQL relational database [201]. It stores the invariant values for about 34 million pairwise non-isomorphic graphs including all graphs up to order 10, along with other sparser families of graphs such as trees or chemical graphs (note that orders higher than 10 are not yet available within the interface). To guarantee that each graph is represented uniquely, the database relies on canonical forms computed via McKay's Nauty [162]. Currently, the database encompasses approximately fifty numeric invariants and a dozen Boolean ones. We continuously expand this collection as new research needs arise, and we actively encourage users to suggest the addition of new, relevant invariants.

10.3.2 Functionalities of the website

This section details the core features of the web interface, publicly accessible at `https://phoeg.umons.ac.be`. The user workflow is logically divided into two main stages: defining the invariant space (Section 10.3.2) and visually exploring the resulting polytopes and graphs (Section 10.3.2).

Problem definition

Upon accessing the homepage, the user's first task is to configure the mathematical problem, as shown in Figure 10.2. The only mandatory step is defining the two-dimensional invariant space by selecting the invariants for the X and Y axes. The input fields feature auto-completion to assist the user in navigating the extensive list of available invariants (see Figure 10.3a). Once these axes are defined, the resulting polytope is dynamically generated and displayed (see Section 10.3.2).

To refine the exploration, the user can incrementally add three types of optional parameters:

- **Coloration:** Applies a continuous color gradient to the polytope points based on the value of a third selected invariant. The user defines a bi-color gradient representing the minimum and maximum values of this invariant (Figure 10.3b).

- **Highlighting:** Isolates and distinctly colors points (and their corresponding drawn graphs) that strictly match a specific target value for

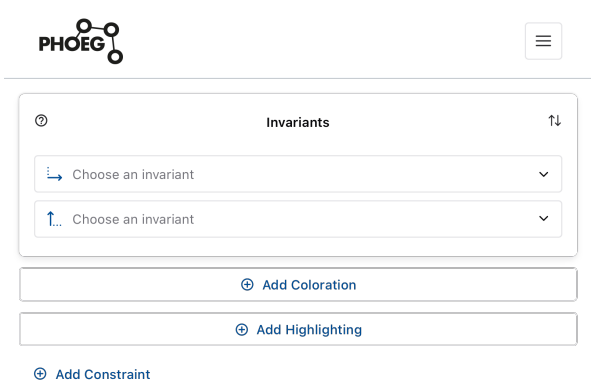
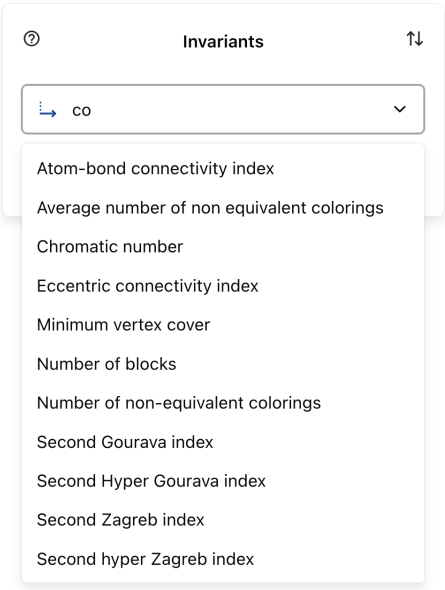
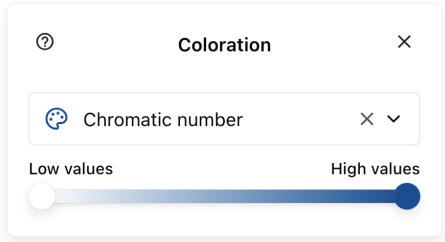


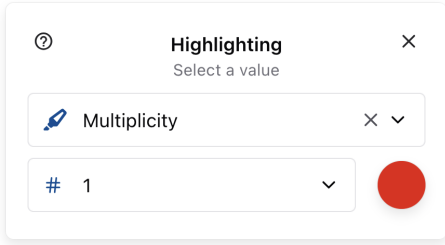
Figure 10.2: Basic problem definition



(a) Invariant input field for the X-axis



(b) Defining the coloration

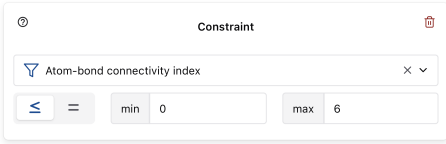


(c) Defining the highlighting

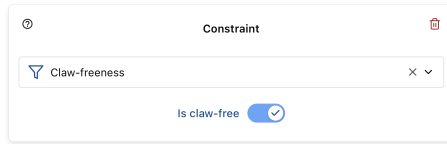
Figure 10.3: Overview of the parameter input boxes in the Problem Definition panel.

a chosen invariant (see for example Figure 10.3c where the points that represent graphs with an average number of non-equivalent colorings of 2 (and drawn graphs meeting the same criterion) will be colored red).

- **Constraints:** Acts as a dataset filter, restricting the visualized graphs to those satisfying specific Boolean conditions (true/false) or numerical bounds (greater than, less than, or equal to a target value). Figures 10.4a and 10.4b illustrate these constraint definition boxes.



(a) Numerical constraint



(b) Boolean constraint

Figure 10.4: Examples of constraint definition boxes.

Examples of practical applications of coloration and highlighting are provided in the next section. Finally, it is worth noting a special option available within the coloration and highlighting menu: *Multiplicity*. While not a mathematical graph invariant per se, multiplicity acts as a powerful visual tool by mapping the color scale (or highlight) to the density of the points—specifically, the exact number of pairwise non-isomorphic graphs residing at a single (x, y) coordinate. This allows users to instantly distinguish visually between rare (e.g., points with a multiplicity of 1) and dense hubs where numerous graphs overlap.

Problem display

Once the problem is defined, the exploration interface is populated (Figure 10.5). The user can filter and reorder the visualization of the polytopes by associated graph orders using toggle buttons at the top left (Figure 10.6). Polytopes from order 2 to 10 are available, with order 6 displayed by default.

The interface is divided into two highly interactive panels. To ensure optimal usability across any screen size, users can manually adjust the relative width of the left and right panels, with all resulting polytopes and graph drawings scaling automatically to fit the available space:

Polytope Exploration (Left Panel): This section displays and allows to scroll the polytopes for the selected orders, constructed from the appoach described in Section 10.2. The axes reflect the chosen invariants, and metadata (such as the number of points and vertices) is provided. Users can navigate the space by zooming, panning, or synchronizing axes across different orders for direct scale comparison. Interactivity is a core feature: clicking a specific point selects it, while clicking a facet automatically selects all its incident points.

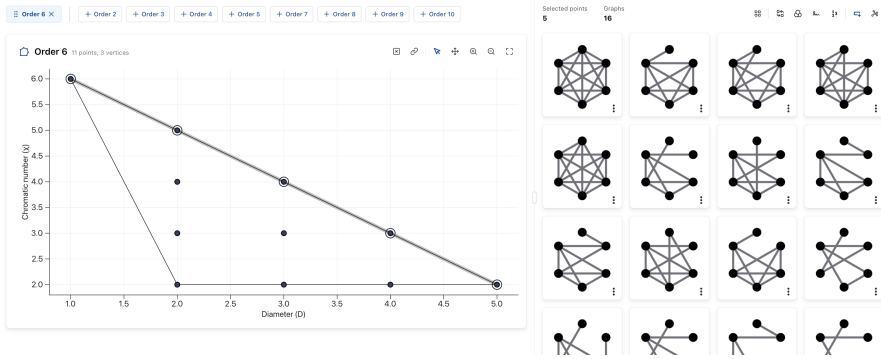


Figure 10.5: Problem display interface

Figure 10.6: Order selection

Graph Visualization (Right Panel): This section dynamically renders the graphs corresponding to the points selected in the left panel. To handle large datasets, graphs are lazy-loaded upon scrolling. The graph drawing module offers extensive customization:

- **Layout and Rendering:** Users can drag vertices manually or apply one of 8 automated layout algorithms (e.g., *circle* as seen in Figure 10.5 or *cose*, as seen in Figure 10.7). If necessary, the user can drag vertices around to move them.
- **Graph Data:** By default, only graph drawings are shown, but the user can toggle the display to reveal graph signatures, invariant values (Figure 10.8), or render complement graphs (where edges and non-edges are inverted and displayed in green, see Figure 10.9).
- **Visual Cues:** Vertex colors can be mapped to their respective degrees for immediate structural insight. For example, the graphs drawn in Figure 10.7 have their vertices colored depending on their degree (degree 1 in yellow, degree 2 in red, etc.).

These visual settings can be applied globally to all displayed graphs or individually via a context menu below each graph.

Furthermore, if a *Coloration* or *Highlighting* was defined in the previous step, it is visually applied in both panels. For instance, Figure 10.10 illustrates a

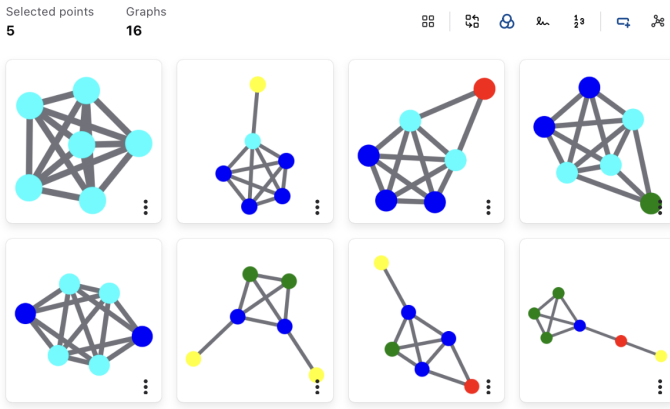


Figure 10.7: Layout using the *cose* algorithm and with a degree-based coloration of the vertices

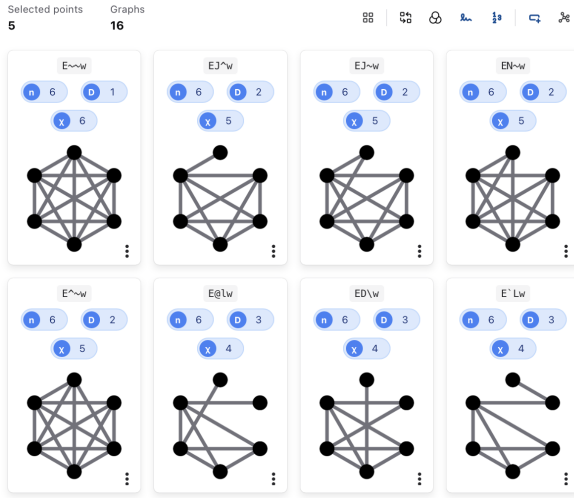


Figure 10.8: Graph rendering with signature and invariant annotations

coloration based on the *maximum degree* (the highest number of edges incident to any single vertex in the graph), applying a gradient from white to blue across the points of the polytope. As established in Section 10.2, a single geometric coordinate can represent multiple non-isomorphic graphs. If these underlying graphs yield different values for the invariant used for coloration, the tool cannot assign a single color; instead, it renders the point as a diamond

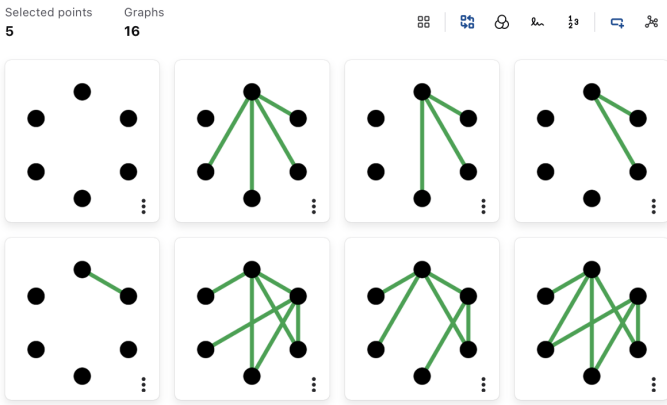


Figure 10.9: Complement graphs rendering

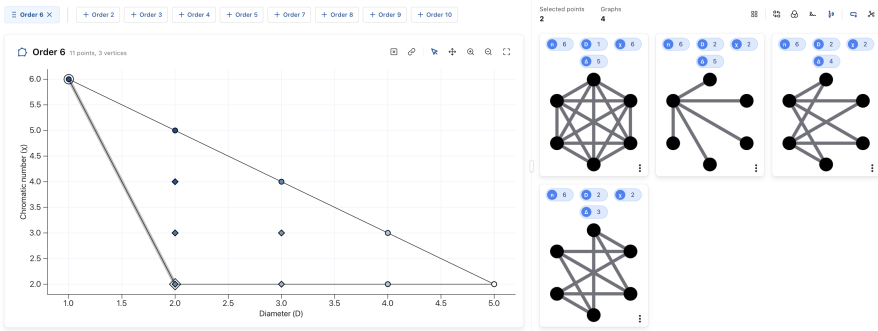


Figure 10.10: Coloration based on the maximum degree, with a gradient from white to blue

to alert the user of this variance, a feature clearly visible in the same figure. To properly understand Figure 10.11, we first introduce two notions. The *eccentricity* $\varepsilon(v)$ of a vertex v is defined as the greatest distance between v and any other vertex in the graph. The *Eccentric Connectivity Index* (ECI) of a graph G, denoted by $\xi^c(G)$, is the sum over all vertices of the product of their degree and their eccentricity, that is, $\xi^c(G) = \sum_{v \in V(G)} d(v)\,\varepsilon(v)$. Figure 10.11 demonstrates a highlighting rule applied to the polytope of graphs of order 8, plotting the eccentric connectivity index against the size. Specifically, points corresponding to graphs with a diameter of 2 are colored red. This feature allows to immediately identify that graphs of diameter 2 are clustered in the upper region of the polytope. The highlighting is also applied to the right panel,

where the corresponding graph drawings are displayed with a red background. In Figure 10.11, the complete graph on 8 vertices is the only non-highlighted graph since its diameter is 1.

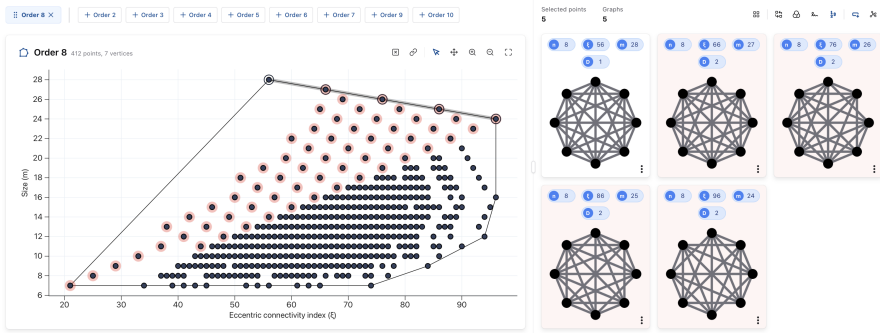


Figure 10.11: Highlighting graphs with a diameter of 2

Finally, a notable feature of the web interface is that most of the session state is encoded in the URL. This encompasses not only the full problem definition (selected invariants, applied constraints, coloration, and highlighting parameters) but also most of the current state of the display: the selected graph orders, the points selected in the polytopes, whether axis synchronization is enabled, and the active graph visualization options. This has several practical implications: the interface is inherently refresh-resistant, as reloading the page faithfully restores the exact state of the session. It also makes problems trivially shareable. To streamline this process, the interface features a dedicated **Share Configuration** button located in the problem definition panel. Clicking this button automatically copies the complete, state-encoded URL to the user's clipboard. Anyone opening this link is brought to the exact same analytical view, making it an ideal tool for collaborative research, saving progress, or linking directly to a specific dataset from an academic paper. For instance, if you follow this link, you will get the exact configuration used in Figure 10.11.

10.3.3 API

Beyond the web interface, PHOEG exposes its core functionalities through a standalone RESTful API. This programmatic access is designed for researchers who wish to integrate PHOEG's extensive database into their own computational workflows, scripts, or automated pipelines without manual interaction with the UI.

The API serves as a bridge to the PostgreSQL database, allowing users

to perform complex queries using standard HTTP requests. Data is typically exchanged in JSON format, ensuring compatibility with most modern programming languages such as Python or R. Key capabilities of the API include:

- **Data Retrieval:** Fetching specific invariant values for a given graph identified by its canonical signature.
- **Advanced Filtering:** Retrieving the individual graphs lying on specific points of the polytope. Given a pair of invariants defining the two axes, a graph order, and a list of (x, y) coordinates, it returns all graphs of that order whose invariant values match those coordinates. The result set can be further narrowed using range constraints on any invariant, and optionally enriched with values of additional invariants. A coloration invariant can also be specified, whose value is then attached to each returned graph to support the coloration feature of the interface.
- **Exporting:** Obtaining graph data in the *graph6* standard format [160] for further local analysis or for use with other graph theory software.

The current API is documented at `https://phoeg.umons.ac.be/phoeg/api/docs`.

10.4 Using PHOEG in research

PHOEG can be utilized in research across various contexts. It serves as an excellent starting point for the preliminary inspection of a problem, helping researchers to build intuition. Crucially, it assists in formulating new conjectures, as detailed in Section 10.4.1, in discovering counterexamples to existing ones, as demonstrated in Section 10.4.2, and even provides assistance for formal proofs, as discussed in Section 10.4.3.

10.4.1 Finding conjectures

The visual and geometric nature of PHOEG acts as a powerful catalyst for forming new mathematical hypotheses. By visualizing the invariant space and explicitly identifying the graphs located on the bounding facets, researchers can gather critical insights even when restricted to graphs of small orders. This empirical observation often makes it surprisingly straightforward to deduce the structural properties of extremal graphs for any arbitrary order n, and subsequently, to generalize the corresponding inequalities. In practice, when computing and displaying polytopes for successive graph orders, it is remarkably

common to observe that their overall geometric shapes remain structurally consistent. These recurring polyhedral patterns immediately capture the observer's attention, significantly easing the extrapolation of vertex coordinates and the formulation of general, parameterized conjectures.

An illustration of this phenomenon is provided in Figure 10.12, which displays the polytopes for graphs of orders 6, 7, and 8, mapping the independence number α against the size m. The consistent geometric shape across these successive orders is immediately apparent. By interacting with the lower boundary of these polytopes - representing the problem of minimizing the number of edges for a given independence number - a user can easily inspect the corresponding extremal graphs. Such an examination quickly reveals that these vertices consistently correspond to Turán graphs (formed by disjoint unions of cliques of balanced sizes), perfectly aligning with the well-known result established by Turán [206].

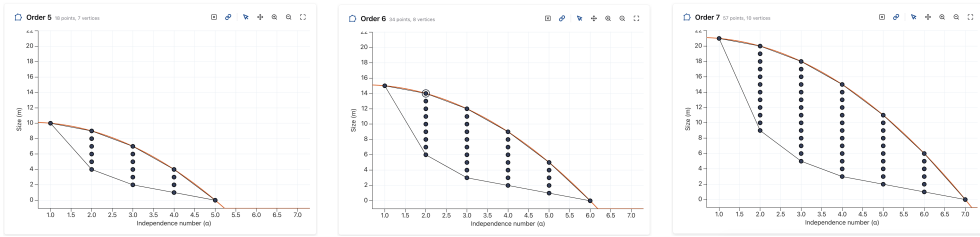


Figure 10.12: Independence number against the number of edges for graphs of order 6, 7, and 8.

To further support hypothesis formulation and testing, PHOEG includes a *Plot Function* feature. This tool allows users to superimpose custom algebraic curves onto the two-dimensional space by defining a formula that relates the X- and Y-axis invariants, optionally incorporating the graph order n. As an illustration, while the lower boundary of the polytopes in Figure 10.12 is populated by Turán graphs, the upper boundary represents the maximization of the size m for a fixed independence number α. Using this plotting feature, we overlaid the function $m = \frac{1}{2}\left(n^2 - n - \alpha^2 + \alpha\right)$ (in red in Figure 10.12). This curve perfectly aligns with the upper extremal vertices, which correspond to *complete split graphs*—graphs whose vertex set can be partitioned into a clique and an independent set, with every vertex in the clique adjacent to every vertex in the independent set.

This core methodology has a long track record of driving research. Its predecessor, GraPHedron, was used internally by our research team starting in 2005, before being progressively replaced by PHOEG in the mid-2010s.

When the geometrically derived conjectures are successfully proven, they contribute directly to the broader field of extremal graph theory. Specifically, the hypotheses generated and refined through GraPHedron and PHOEG have led to several peer-reviewed publications. Notable examples include results on the independence number [30, 31, 40], the number of non-equivalent colorings [1, 117], the Eccentric Connectivity Index [67, 111], an analysis of the approximation factor of a maximal matching heuristic [37], extremal properties in chemical graph theory [21, 114], as well as invariants concerning the average number of colorings and the average number of matchings [115, 118, 119]. A comprehensive and continuously updated list of all related publications is maintained on the PHOEG website.

10.4.2 Finding counterexamples

Beyond original discovery, PHOEG is a highly effective platform for verifying existing mathematical claims. If a known conjecture is formulated in a way that is compatible with our two-dimensional geometric approach - typically involving bounds or relationships between invariants - it can be systematically tested within the interface. By generating the relevant polytope, researchers can quickly gather empirical support for the conjecture, refine its proposed bounds, or outright refute it by uncovering unexpected points that violate the hypothesized inequalities.

To give a concrete example, in 2014, Zhang et al. [221] proposed a conjecture regarding the Eccentric Connectivity Index (ECI). In this context, they defined

$$d_{n,m} = \left\lfloor \frac{2n + 1 - \sqrt{17 + 8(m - n)}}{2} \right\rfloor$$

and $E_{n,m}$ as the graph obtained from a clique K_{n-d-1} and a path $P_{d+1} = v_0 v_1 \dots v_d$ by joining each vertex of the clique to both v_d and v_{d-1}, and by joining $m - n + 1 - \binom{n-d}{2}$ vertices of the clique to v_{d-2}. The conjecture is as follows:

Conjecture 67 (Zhang, Liu and Zhou, [221])**.** *Let $d_{n,m} \geq 3$. Then $E_{n,m}$ is the unique graph with maximal eccentric connectivity index among all connected graphs with n vertices and m edges.*

The authors prove that the conjecture is true for certain values of n and m. However, with the help of PHOEG, we can find a counterexample with little effort. Figure 10.13 shows this counterexample. The selected point corresponds to $n = 7$ and $m = 15$, which gives $d_{7,15} = 3$. This point corresponds to two graphs that have an ECI of 65. The graph on the right is $E_{7,15}$, but the other

graph is not isomorphic to $E_{7,15}$, which shows that $E_{7,15}$ is not the unique graph with maximal eccentric connectivity index among all connected graphs with 7 vertices and 15 edges.

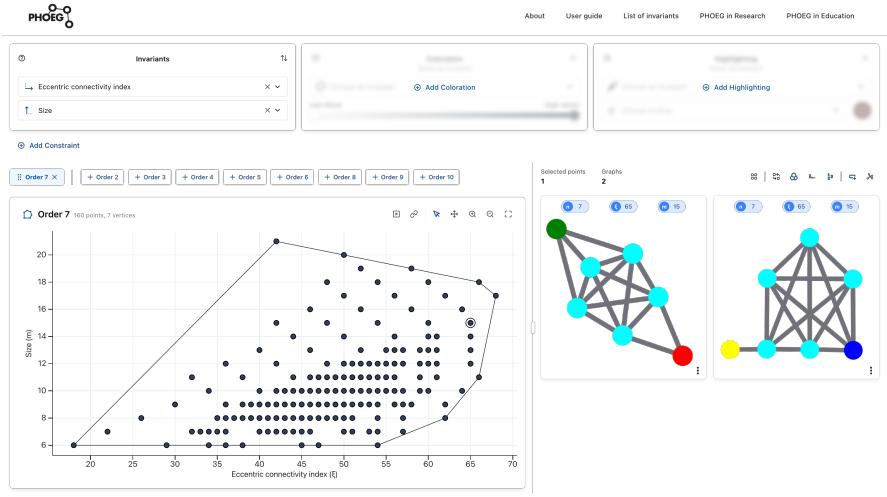


Figure 10.13: Counterexample for Zhang's conjecture for $n = 7, m = 15, d = 3$

10.4.3 Help for proofs

Beyond formulating conjectures, visually observing extremal points and their corresponding graph structures can provide critical intuition for formal demonstrations, particularly for proofs by transformation. In this mathematical approach, a graph is modified step-by-step to reach an extremal configuration while strictly increasing or decreasing a given invariant.

To support this methodology, a dedicated computational module named TransProof was developed as part of Devillez's PhD thesis [64]. Although TransProof is not currently available within the public PHOEG web interface, its mechanics are deeply tied to the project's ecosystem. The primary function of TransProof is to compute a *metagraph* for a specified set of graphs and structural transformations. In this metagraph, vertices represent the graphs themselves, and directed arcs represent the applied transformations.

Once computed, this metagraph is stored in a database, allowing researchers to run complex queries to test and refine their proof strategies. For instance, a user can query the database to check for counterexamples to a proposed proof idea by ensuring the absence of non-improving arcs. Furthermore, the

flexibility of the database query language allows for more advanced topological checks, such as verifying whether every non-extremal graph possesses at least one strictly improving arc, or identifying if certain non-extremal graphs cannot be improved by the considered set of transformations.

10.5 Using PHOEG in education

Beyond its applications in research, PHOEG is actively utilized for educational purposes at both the University of Mons (UMONS) and Polytechnique Montréal, specifically within the context of advanced graph theory courses.

To facilitate the onboarding process for students and new researchers alike, the platform includes a comprehensive suite of built-in interactive tutorials. Accessible directly from the interface's help section, these step-by-step guides allow users to familiarize themselves with the tool at their own pace. Through hands-on, guided sandboxes, users learn the core geometric concepts, how to apply dataset constraints and colorations, and how to effectively navigate the polytope and graph panels before tackling more complex, unguided assignments.

For instance, at UMONS, in the *Graphs & Artificial Intelligence* course, once students have completed these introductory tutorials, the tool is further explored through a practical session where students are invited to use PHOEG's web interface to explore and solve classical problems in extremal graph theory.

The first problem assigned to the students is based on a fictitious police investigation involving Al Capone, presented as a direct continuation of a problem introduced earlier in the course. In that earlier problem, an inspector had met seven mafiosi: John Gotti, Joe Adonis, Albert Anastasia, Liborio Bellomo, Tommaso Buscetta, Giuseppe Calicchio, and Al Capone. Each of them declared the number of people they had business relations with, except Al Capone, who remained deliberately vague and only admitted to having 'more than 3' such relations. Using the handshaking lemma, students had previously established that the sum of all degrees must be even, which forces Al Capone to have exactly 5 business relations, yielding the full degree sequence $2, 3, 3, 3, 4, 4, 5$.

The investigation has since progressed: the police have almost proven that Al Capone is guilty of murder, but he claims he was with Albert Anastasia at the time of the crime, providing himself an alibi. A new piece of evidence has also emerged: the largest group of people with no business relations between them is of size 2. In graph-theoretical terms, this clue is meant to inform them that the independence number of the graph representing this situation is of size 2. The objective of the riddle is to determine whether Al Capone (the vertex of degree 5) could have a business relationship with Albert Anastasia

(the vertex of degree 2).

By applying these constraints in PHOEG, students discover that only a single graph matches the entire description. Because this unique graph contains exactly one vertex of degree 5 and one vertex of degree 2, students can directly inspect its structure to see that no edge connects these two specific vertices. They can thereby conclude that Al Capone and Albert Anastasia do not share a business relationship.

In a subsequent exercise, students are asked to use PHOEG to hypothesize an upper bound on the *chromatic number* based on the *maximum degree* for *connected graphs*. The pedagogical goal is to guide them toward empirically rediscovering Brooks' theorem [29], or at least closely approaching it. This foundational theorem states that for any connected graph that is neither a *complete graph* nor an *odd cycle*, its chromatic number is at most its maximum degree.

A final exercise asks students to explore general graphs of a given size m that minimize the number of *non-equivalent colorings*—that is, valid vertex colorings that are genuinely distinct and cannot be transformed into one another simply by permuting the color labels. Unlike the previous problems, accurately formulating the conjecture for these presumed extremal graphs is significantly more intricate, primarily because their formal description involves triangular numbers. Students are informed that a conjecture regarding these specific graphs was initially identified in 2014 with the assistance of PHOEG. However, despite continuous efforts to prove or disprove it, it remains an open problem today. Pedagogically, this exercise is extremely valuable. It demonstrates to students the open-ended nature of scientific research and perfectly illustrates a fundamental reality of extremal graph theory: while computational tools like PHOEG can make the empirical identification of a conjecture surprisingly accessible, providing its formal mathematical proof can remain a formidable challenge.

10.6 Conclusion

This chapter presented PHOEG, an interactive online tool designed to support both research and education in extremal graph theory. Built around an exact geometrical approach that embeds graphs into a two-dimensional invariant space and computes their convex hulls, PHOEG empowers users to visually explore relationships between graph invariants, identify extremal graphs, and formulate conjectured inequalities.

We detailed the platform's modern architecture, including its comprehensive web interface — featuring problem definition, interactive polytope visualization,

and highly customizable graph rendering — and its standalone RESTful API for programmatic database access. Furthermore, we demonstrated PHOEG's practical impact in two primary domains. In research, we illustrated its capacity to facilitate the discovery of new conjectures, swiftly identify counterexamples, and provide structural intuition for formal mathematical proofs, effectively serving as a short survey of several results enabled by this geometric approach. In education, we highlighted its successful integration into advanced graph theory courses at the University of Mons and Polytechnique Montréal, where students actively engage with the tool to empirically rediscover classical theorems and grapple with open mathematical problems.

Looking ahead, we plan to continuously expand the underlying database by computing higher graph orders and integrating additional invariants, actively welcoming suggestions from the scientific community. In particular, restricting to specific graph classes such as trees or chemical graphs, for which the combinatorial explosion is significantly more manageable, would allow us to push the database to higher orders. Ultimately, we hope that PHOEG will continue to evolve as an accessible and invaluable resource for both researchers and educators in the field of graph theory.

Chapter 11

ChemicHull

As reported in [75] and explained in Chapter 9, a total of 96 polytopes exist, each characterized by its own distinct set of facets. An online tool called ChemicHull [23] has been developed to allow the visualization of these 96 polytopes. This chapter presents the tool and its capabilities.

As indicated in Section 9.2.2, a polytope can be represented either by its extreme points (V-representation) or by its facets (H-representation). It is worth noting that, although two polytopes may share the same set of extreme points, their facets can differ. For instance, as observed in Section 9.3, all chemical trees with an even number $n \geq 13$ of vertices share the same set of extreme points. However, depending on the value of n mod 12, their facets may vary. For instance, when $n \bmod 12 = 2$, the polytope $\mathcal{P}_{n,n-1}$ has ten facets, whereas when $n \bmod 12 = 4$, it has only seven. This difference arises because several extreme points coincide in the latter case; indeed, when $n \bmod 12 = 4$, we have V13 = V14 = V15 and V16 = V17 = V18. The two polytopes (with $n \bmod 12 = 2$ or 4) share six facets, namely,

$$\begin{aligned} m_{12} &\geq 0, \\ m_{13} &\geq 0, \\ m_{33} &\geq 0, \\ -4m_{12} - 4m_{13} + m_{33} &\geq -n - 5, \\ (n-4)m_{12} + \left(\frac{n}{2} - 2\right) m_{13} + \left(3 - \frac{n}{2}\right) m_{33} &\geq 2n - 8, \\ (n-7)m_{12} + (n-6)m_{13} + (4-n)m_{33} &\geq 2n - 14. \end{aligned}$$

Moreover, when $n \bmod 12 = 2$, $\mathcal{P}_{n,n-1}$ possesses the following four additional facets.

$$-2m_{12} - m_{13} \geq \frac{-2(n+1)}{3},$$

$$-m_{12} + m_{33} \geq \frac{-n-2}{4},$$
$$-2m_{12} - 2m_{13} + m_{33} \geq \frac{-2(n+4)}{3},$$
$$(2n-16)m_{12} + (2n-13)m_{13} + (10-2n)m_{33} \geq 4n - 32,$$

whereas for $n \bmod 12 = 4$, $\mathcal{P}_{n,n-1}$ possesses the following additional facet:

$$-3m_{12} - 2m_{13} + m_{33} \geq \frac{-3(n+4)}{4}.$$

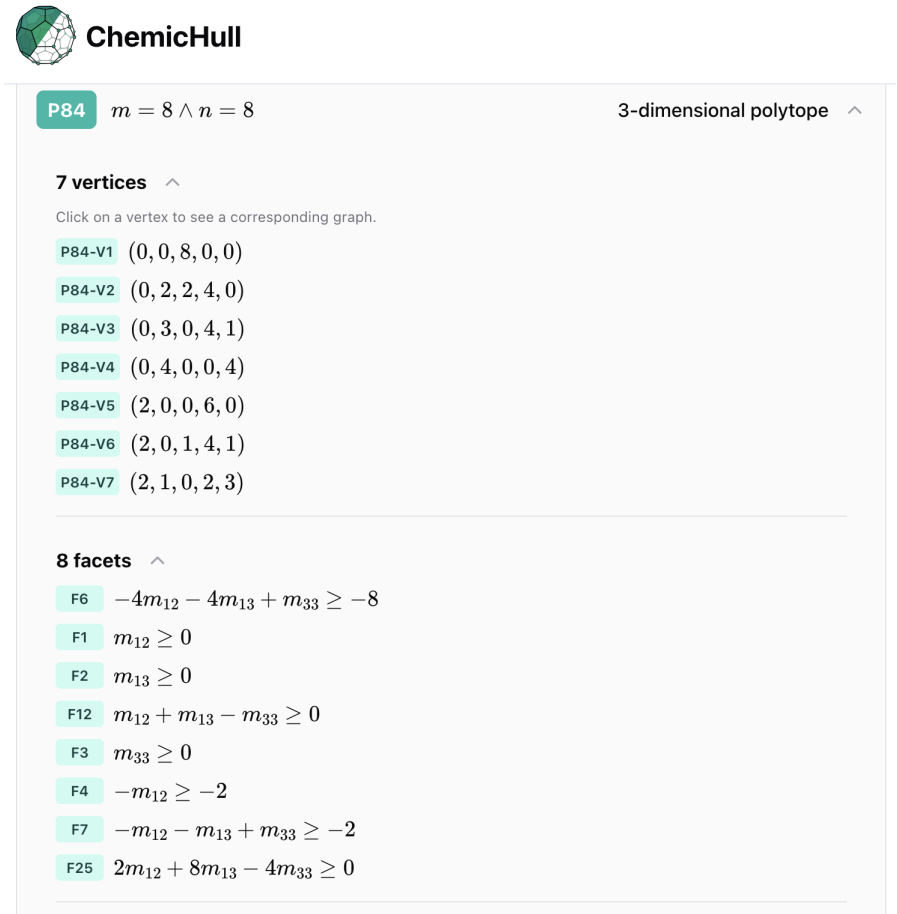


Figure 11.1: V-Representation and H-representation of $\mathcal{P}_{8,8}$ in ChemicHull.

When neither n nor m is specified, ChemicHull returns the complete set of the 96 polytopes. It is possible to filter these polytopes by assigning specific

values to n and/or m, or by using formulas involving these parameters. For example, leaving n free while specifying $m = n - 1$ yields the twelve polytopes corresponding to chemical trees. Once n and m are specified, ChemicHull provides the V-Representation and the H-representation of the corresponding polytope. As an illustration, Figure 11.1 shows the 7 extreme points and the 8 facets of $\mathcal{P}_{8,8}$. It should be noted that, unlike in Chapter 9 and in this chapter, the coordinates are provided directly in terms of the five values m_{12}, m_{13}, m_{22}, m_{23}, and m_{33} (in this order), eliminating the need for the user to calculate m_{22} and m_{23}.

In addition to listing the extreme points and facets, the interface provides a three-dimensional graphical representation of the polytopes (see Figure 11.2).

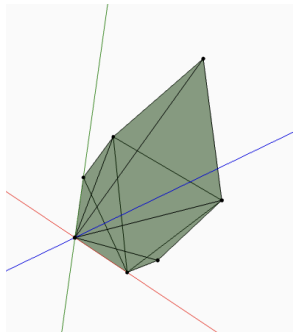


Figure 11.2: A three-dimensional graphical representation of the polytope $\mathcal{P}_{8,8}$

The interface also allows users to specify a degree-based topological index with ease. For example, Figure 11.3 demonstrates how to define the Randić index.

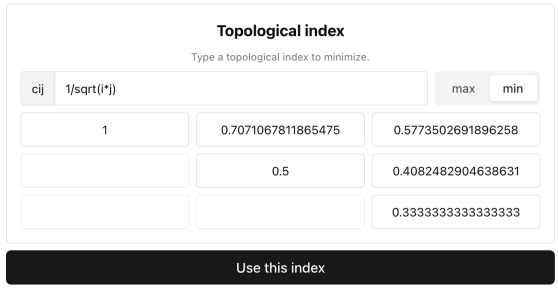


Figure 11.3: Definition of a degree-based topological index (in this case, the Randić index).

Once a degree-based topological index is defined, ChemicHull can display the coordinates of the chemical graphs that maximize or minimize this index (see Figure 11.4).

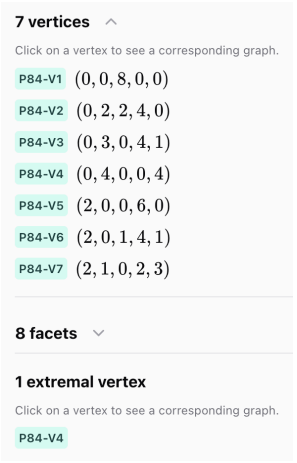


Figure 11.4: Extreme point (P84-V4) that minimizes the Randić index in $\mathcal{P}_{8,8}$.

Furthermore, users can visualize any graph corresponding to a selected point by simply clicking on it (see Figure 11.5). The construction of a graph from the coordinates m_{ij} follows the algorithm described in [108].

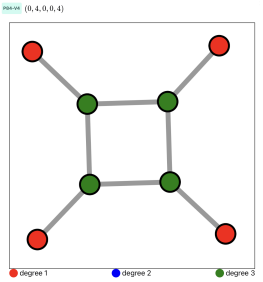


Figure 11.5: A graph corresponding to the extreme point P84-V4=$(0, 4, 0, 0, 4)$ in $\mathcal{P}_{8,8}$.

Chapter 12

Conclusion

This chapter serves as a conclusion to this thesis. In the introduction, we put forward three sets of objectives. These objectives yielded significant scholarly output, including five published papers [21, 114, 115, 118, 119] and three papers currently under review [22, 24, 75], as well as two software tools made available to the scientific community via web interfaces [22, 24]. In this conclusion, we remind the reader of these results. Each set of results of course corresponds to a chapter in this thesis. Section 12.1 summarizes this thesis, while Section 12.2 outlines perspectives for future work.

12.1 Summary

We established several bounds on $\mathcal{A}(G)$, the average number of colors among the non-equivalent colorings of a graph G. It is straightforward to show that $\mathcal{A}(G) \leq \mathcal{A}(\mathsf{K}_n) = n$ for every graph of order n, where equality holds if and only if $G \simeq \mathsf{K}_n$. Consequently, n constitutes the best possible upper bound on $\mathcal{A}(G)$ for graphs of order n.

Regarding the average number of colors in the non-equivalent colorings of a graph, we conjecture that the optimal lower bound on $\mathcal{A}(G)$ for a graph G of order n is $\mathcal{A}(\overline{\mathsf{K}}_n) = \frac{\mathsf{B}_{n+1}-\mathsf{B}_n}{\mathsf{B}_n}$. Despite the apparent simplicity of this expression, our results show that proving this conjecture lies beyond the reach of basic approaches such as sequential edge removal, suggesting that fundamentally different techniques may be required. We further refined this conjecture by proposing lower bounds related to the chromatic number $\chi(G)$ and to the maximum degree $\Delta(G)$ of G. We have thus stated three conjectures, which we have shown to be true for chordal graphs and for graphs with maximum degree at most 2.

We have given a general upper bound on $\mathcal{A}(G)$ that is valid for all graphs

G, as well as sharper bounds for graphs of order n whose maximum degree satisfies $\Delta(G) \in 1, 2, n-2$.

Regarding the average size of maximal matchings in a graph G, we have determined $\lim_{n\to\infty} \mathcal{I}(G_n)$ for many families $\{G_n\}_{n\geq 0}$ of graphs, showing that intermediate values between $\frac{1}{2}$ and 1 are attained. We proposed a general technique that can be applied when the number of maximal matchings in G_n linearly depends on the number of maximal matchings in graphs $G_{n'}$ of the same family, with $n' < n$. This technique allowed us to recalculate known values of $\lim_{n\to\infty} \mathcal{I}(G_n)$ which were obtained by other authors using different concepts such as (bivariate) generating functions.

We have determined a sharp upper bound on the arithmetic-geometric index of chemical graphs of order n and size $m \geq n-1$, and we have characterized the chemical graphs that reach the bound. This allows, for example, to characterize extremal chemical trees as well as extremal unicyclic or bicyclic chemical graphs. For $m \leq n-2$, we have shown that there is an extremal chemical graph of order n and size m which is a disjoint union of extremal chemical trees.

We have shown that for 29 degree-based indices out of 33, 5 families of extremal chemical graphs are sufficient to describe them all. Most of the characterizations we have given for extremal graphs are based on a set of 8 values $V_1, \ldots, V_8$. If new topological indices are proposed, it is therefore easy to check whether they have the same extremal properties of the studied indices.

We have shown that the polyhedral approach offers a powerful and systematic framework for identifying families of chemical graphs that are extremal with respect to degree-based topological indices. Beyond reproducing many established results from the literature, this method also allows for their verification and, when required, correction, as exemplified by our discovery of a counterexample to a previously reported extremal result involving the Randić index. Moreover, our findings indicate that some graphs can never serve as extreme points under any degree-based topological index. Therefore, if a molecule of interest in extremal chemistry is not represented among the extremal points identified in this study, it must instead be described by a topological index beyond the degree-based class.

We have extended PHOEG, an online tool for Extremal Graph Theory, by developing an accessible web interface, replacing the previous reliance on direct SQL queries to the database. The resulting platform is now publicly available.

We have developed and presented ChemicHull, an online tool that enables researchers to easily explore and apply the results presented in this work. The tool offers an intuitive interface for visualizing extremal chemical graphs and analyzing optimal values of degree-based topological indices.

12.2 Perspectives

For the lower bounds on the number of colors in the non-equivalent colorings of a graph, we have stated three open problems that, although proven true for chordal graphs and for graphs with maximum degree at most 2, are yet to be proved.

On the upper-bound side, the problem of identifying extremal graphs with maximum degree in $\{3, \ldots, n-3\}$ remains open. For $\Delta(G) \in 1, n-2, n-1$, graphs maximizing $\mathcal{A}(G)$ are known to have a highly regular structure, being isomorphic to $\left\lfloor \frac{n}{\Delta(G)+1} \right\rfloor \mathsf{K}_{\Delta(G)+1} \cup \mathsf{K}_{n \bmod (\Delta(G)+1)}$, whereas this description already fails for $\Delta(G) = 2$. It is therefore natural to ask whether this construction continues to be extremal for intermediate values of the maximum degree.

Our computational investigation, based on an exhaustive enumeration of all graphs on at most 12 vertices using *PHOEG* [66], shows that the proposed construction does not hold for a single example. Among more than 165 billion graphs, we identified a unique graph of order $n \leq 12$ with $\Delta(G) \neq 2$, namely $\overline{\mathsf{C}}_6 \cup \mathsf{K}_4$, for which the above construction is not optimal. Indeed, $\mathcal{A}(\overline{\mathsf{C}}_6 \cup \mathsf{K}_4) = 5.979 > 5.967 = \mathcal{A}(2\mathsf{K}_4 \cup \mathsf{K}_2)$, showing that $2\mathsf{K}_4 \cup \mathsf{K}_2$ does not maximize $\mathcal{A}(G)$ among graphs of order 10 with maximum degree 3.

Regarding the average size of maximal matchings, we proposed an open problem concerning the class $\boldsymbol{T}_{2n}$ of trees of even order. Dyer and Frieze [76] conjectured that $\lim_{n\to\infty} \min_{T \in \boldsymbol{T}_{2n}} \mathcal{I}^{DF}(T) = \lim_{n\to\infty} \mathcal{I}^{DF}(\widetilde{\mathsf{P}}_n)$, that is, they propose that the asymptotic worst case for $\mathcal{I}^{DF}$ over $\boldsymbol{T}_{2n}$ is attained by the trees obtained from a path P_n by attaching a single leaf to each of its vertices. In contrast, we showed that $\lim_{n\to\infty} \min_{T \in \boldsymbol{T}_{2n}} \mathcal{I}(T) < \lim_{n\to\infty} \mathcal{I}(\widetilde{\mathsf{P}}_n)$, indicating that the trees minimizing $\mathcal{I}$ over $\boldsymbol{T}_{2n}$ are distinct from those conjectured to be extremal for $\mathcal{I}^{DF}$. Identifying the structure of such extremal trees for $\mathcal{I}$ therefore appears to be an interesting direction for further research.

Our work in chemical graph theory has culminated in our complete polyhedral description of chemical graphs of maximum degree at most 3. A natural direction for further research is to extend this description to chemical graphs of maximum degree at most 4. Preliminary investigations, however, indicate that such an extension would involve substantially greater combinatorial and computational complexity. It would of course be interesting to see this work used in future research regarding degree-based topological indices.

As for the tools we have developed, there are of course many things we could still improve. For PHOEG, our recent redevelopment from scratch was actually specifically for this objective. Indeed, a clean slate allows us to more easily implement new features. Of course, we are also always open to feedback

from users, whether to add more invariants to the database, or to implement new features.

For ChemicHull, although the main functionalities are implemented, there are always features to be improved or implemented. We are currently still pursuing this avenue of work, so there will be opportunities for ideas there. We have also presented our work in conferences and will continue to do so. This should provide the opportunity for chemists to offer feedback. For instance, when multiple graphs correspond to the same (m_{12},m_{13},m_{33}) point, we currently display only one representative; we plan to extend this to allow users to browse multiple such graphs. We will also be sure to update ChemicHull if and when discoveries are made for a maximum degree of 4.

Alphabetical Index

Bibliography

[1] Absil, R., Camby, E., Hertz, A., and Mélot, H. A sharp lower bound on the number of non-equivalent colorings of graphs of order n and maximum degree $n-3$. *Discrete Applied Mathematics 234* (2018), 3–11. Special Issue on the Ninth International Colloquium on Graphs and Optimization (GO IX), 2014.

[2] Adebayo, S., and Chakra UI Team. Chakra UI: component library for React applications. `https://chakra-ui.com/`, 2026. Accessed: March 27, 2026.

[3] Ali, A., Albalahi, A. M., Alanazi, A. M., Bhatti, A. A., and Hamza, A. E. On the maximum sigma index of k-cyclic graphs. *Discrete Applied Mathematics 325* (2023), 58–62.

[4] Ali, A., Dimitrov, D., Du, Z., and Ishfaq, F. On the extremal graphs for general sum-connectivity index (χ^{α}) with given cyclomatic number when $\alpha > 1$. *Discrete Applied Mathematics 257* (2019), 19–30.

[5] Ali, A., Elumalai, S., and Mansour, T. On the symmetric division deg index of molecular graphs. *MATCH Communications in Mathematical and in Computer Chemistry 83*, 1 (2020), 205–220.

[6] Ali, A., Furtula, B., Gutman, I., and Vukičević, D. Augmented zagreb index: extremal results and bounds. *MATCH Communications in Mathematical and in Computer Chemistry* (2021).

[7] Ali, A., Furtula, B., Redžepović, I., and Gutman, I. Atom-bond sum-connectivity index. *Journal of Mathematical Chemistry 60*, 10 (2022), 2081–2093.

[8] Aliakbarpour, M., Biswas, A. S., Gouleakis, T., Peebles, J., Rubinfeld, R., and Yodpinyanee, A. Sublinear-time algorithms for counting star subgraphs via edge sampling. *Algorithmica 80* (2018), 668–697.

[9] Alochukwu, A., and Dankelmann, P. Upper bounds on the average eccentricity of graphs of girth 6 and (c_4, c_5)-free graphs. *arXiv preprint arXiv:2004.14490* (2020).

[10] Alzer, H. On Engel's inequality for Bell numbers. *Journal of Integer Sequences 22* (2019). Article 19.7.1.

[11] Andriantiana, E. O., Razanajatovo Misanantenaina, V., and Wagner, S. The average size of independent sets of graphs. *European Journal of Mathematics 6* (2020), 561–576.

[12] Andriantiana, E. O. D., Razanajatovo Misanantenaina, V., and Wagner, S. The average size of matchings in graphs. *Graphs and Combinatorics 36*, 3 (2020), 539–560.

[13] Aouchiche, M., Hallaoui, I. E., and Hansen, P. Geometric-arithmetic index and minimum degree of connected graphs. *MATCH Communications in Mathematical and in Computer Chemistry 83*, 1 (2020), 179–188.

[14] Arnosti, N. Greedy matching in bipartite random graphs. *Stochastic Systems* (2021), 1–18.

[15] Aronson, J., Dyer, M., Frieze, A., and Suen, S. On the greedy heuristic for matchings. In *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms* (USA, 1994), SODA '94, Society for Industrial and Applied Mathematics, pp. 141–149.

[16] Aronson, J., Dyer, M., Frieze, A., and Suen, S. Randomized greedy matching. ii. *Random Structures & Algorithms 6*, 1 (1995), 55–73.

[17] Bai, Y., Zhao, B., and Zhao, P. Extremal merrifield–simmons index and hosoya index of polyphenyl chains1. *way 3*, 1 (2009), 2.

[18] Benitez, S., and Rocket Contributors. Rocket: A web framework for Rust. `https://rocket.rs/`, 2026. Accessed: March 27, 2026.

[19] Besser, B., and Poloczek, M. Greedy matching: Guarantees and limitations. *Algorithmica 77*, 1 (2017), 201–234.

[20] Bianchi, M., Cornaro, A., Palacios, J. L., and Torriero, A. Lower bounds for the geometric-arithmetic index of graphs with pendant and fully connected vertices. *Discrete Applied Mathematics 257* (2019), 53–59.

[21] Bonte, S., Devillez, G., Dusollier, V., Hertz, A., and Mélot, H. Extremal chemical graphs of maximum degree at most 3 for 33 degree-based topological indices. *MATCH Communications in Mathematical and in Computer Chemistry 95*, 1 (2026), 193–231.

[22] Bonte, S., Devillez, G., Dusollier, V., Hertz, A., Mélot, H., and Schindl, D. Chemichull: an online tool for determining extremal chemical graphs of maximum degree at most 3 for any degree-based topological indices. *arXiv preprint arXiv:2511.20249* (2025).

[23] Bonte, S., Devillez, G., Dusollier, V., and Mélot, H. ChemicHull. `https://phoeg.umons.ac.be/chemichull/`.

[24] Bonte, S., Devillez, G., Dusollier, V., and Mélot, H. Phoeg: an online tool for discovery and education in extremal graph theory, 2026.

[25] Brenti, F. Expansions of chromatic polynomials and log-concavity. *Transactions of the American Mathematical Society 332*, 2 (1992), 729–756.

[26] Brenti, F., Royle, G. F., and Wagner, D. G. Location of zeros of chromatic and related polynomials of graphs. *Canadian Journal of mathematics 46*, 1 (1994), 55–80.

[27] Brightwell, G., Janson, S., and Luczak, M. The greedy independent set in a random graph with given degrees. *Random Structures & Algorithms 51*, 4 (2017), 565–586.

[28] Brinkmann, G., Coolsaet, K., Goedgebeur, J., and Mélot, H. House of Graphs: a database of interesting graphs. *Discrete Applied Mathematics 161*, 1-2 (2013), 311–314.

[29] Brooks, R. L. On colouring the nodes of a network. In *Mathematical Proceedings of the Cambridge Philosophical Society* (1941), vol. 37, Cambridge University Press, pp. 194–197.

[30] Bruyère, V., Joret, G., and Mélot, H. Trees with given stability number and minimum number of stable sets. *Graphs and Combinatorics 28*, 2 (2012), 167–187.

[31] Bruyère, V., and Mélot, H. Fibonacci index and stability number of graphs: a polyhedral study. *Journal of combinatorial optimization 18* (2009), 207–228.

[32] Buckley, F. Distance in graphs, 1990.

[33] Caporossi, G., Gutman, I., and Hansen, P. Variable neighborhood search for extremal graphs: Iv: Chemical trees with extremal connectivity index. *Computers & Chemistry 23*, 5 (1999), 469–477.

[34] Caporossi, G., Gutman, I., Hansen, P., and Pavlović, L. Graphs with maximum connectivity index. *Computational Biology and Chemistry 27*, 1 (2003), 85–90.

[35] Caporossi, G., and Hansen, P. Variable neighborhood search for extremal graphs: 1 the autographix system. *Discrete Mathematics 212*, 1-2 (2000), 29–44.

[36] Carballosa, W., Granados, A., Méndez Bermúdez, J. A., Pestana, D., and Portilla, A. Computational properties of the arithmetic–geometric index. *Journal of Mathematical Chemistry 60*, 9 (2022), 1854–1871.

[37] Cardinal, J., Labbé, M., Langerman, S., Levy, E., and Mélot, H. A tight analysis of the maximal matching heuristic. In *International Computing and Combinatorics Conference* (2005), Springer, pp. 701–709.

[38] Cash, G. G. Relationship between the hosoya polynomial and the hyper-wiener index. *Applied mathematics letters 15*, 7 (2002), 893–895.

[39] Che, Z., and Chen, Z. Lower and upper bounds of the forgotten topological index. *MATCH Communications in Mathematical and in Computer Chemistry 76*, 3 (2016), 635–648.

[40] Christophe, J., Dewez, S., Doignon, J.-P., Fasbender, G., Grégoire, P., Huygens, D., Labbé, M., Elloumi, S., Mélot, H., and Yaman, H. Linear inequalities among graph invariants: using graphedron to uncover optimal relationships. *Networks 52*, 4 (2008), 287–298.

[41] Chu, Z.-Q., Jamil, M. K., and Javed, A. Extremal hyper zagreb index for tricyclic graphs. *International Journal of Applied Mathematics 32*, 6 (2020).

[42] Chung, F. The average distance and the independence. *Journal of Graph Theory 12* (1988), 229–235.

[43] Codding, P.W., Ed. *Structure-Based Drug Design: Experimental and Computational Approaches*. Kluwer, Dordrecht, 1998.

[44] Coolsaet, K., D'hondt, S., and Goedgebeur, J. House of Graphs 2.0: A database of interesting graphs and more. *Discrete Applied Mathematics 325* (2023), 97–107.

[45] Cruz, R., Gutman, I., and Rada, J. Sombor index of chemical graphs. *Applied Mathematics and Computation 399* (2021), 126018.

[46] Cruz, R., and Rada, J. Extremal values of the sombor index in unicyclic and bicyclic graphs. *Journal of Mathematical Chemistry 59* (2021), 1098–1116.

[47] Cui, S.-Y., Wang, W., Tian, G.-X., and Wu, B. On the arithmetic-geometric index of graphs. *MATCH Communications in Mathematical and in Computer Chemistry 85*, 1 (2021), 87–107.

[48] Cvetković, D., Kraus, L., and Simić, S. Discussing graph theory with a computer i. implementation of graph theoretic algorithms/razgovor sa kompjuterom o teoriji grafova i. implementacija algoritama teorije grafova. *Publikacije Elektrotehničkog fakulteta. Serija Matematika i fizika* (1981), 100–104.

[49] Dankelmann, P., and Entringer, R. Average distance, minimum degree, and spanning trees. *Journal of Graph Theory 33*, 1 (2000), 1–13.

[50] Dankelmann, P., Goddard, W., and Swart, C. S. The average eccentricity of a graph and its subgraphs. *Utilitas Mathematica 65*, 05 (2004), 41–51.

[51] Dankelmann, P., Mukwembi, S., and Swart, H. C. Average distance and vertex-connectivity. *Journal of Graph Theory 62*, 2 (2009), 157–177.

[52] Dankelmann, P., and Osaye, F. J. Average eccentricity, minimum degree and maximum degree in graphs. *Journal of Combinatorial Optimization 40*, 3 (2020), 697–712.

[53] Das, K., Gutman, I., and Furtula, B. Survey on geometric-arithmetic indices of graphs. *MATCH Communications in Mathematical and in Computer Chemistry 65* (01 2011), 595–644.

[54] Das, K. C. A sharp upper bound for the number of spanning trees of a graph. *Graphs and Combinatorics 23*, 6 (2007), 625–632.

[55] Das, K. C., Cevik, A. S., and Cangul, I. N. The number of spanning trees of a graph. *Journal of Inequalities and Applications 2013* (2013), 1–13.

[56] Das, K. C., Gutman, I., and Furtula, B. On atom-bond connectivity index. *Chemical Physics Letters 511*, 4 (2011), 452–454.

[57] Das, K. C., Xu, K., and Wang, J. On inverse degree and topological indices of graphs. *Filomat 30*, 8 (2016), 2111–2120.

[58] Davies, E., Jenssen, M., Perkins, W., and Roberts, B. On the average size of independent sets in triangle-free graphs. *Proceedings of the American Mathematical Society 146*, 1 (2018), 111–124.

[59] Deng, H., Balachandran, S., Ayyaswamy, S., and Venkatakrishnan, Y. On the harmonic index and the chromatic number of a graph. *Discrete Applied Mathematics 161*, 16 (2013), 2740–2744.

[60] Deng, H., Elumalai, S., and Balachandran, S. Maximum and second maximum of geometric-arithmetic index of tricyclic graphs. *MATCH Communications in Mathematical and in Computer Chemistry 79*, 2 (2018), 467–475.

[61] Deng, H., Tang, Z., and Wu, R. Molecular trees with extremal values of sombor indices. *International Journal of Quantum Chemistry 121*, 11 (2021), e26622.

[62] Devillers, J. *Neural Networks in QSAR and Drug Design.* Academic Press, London, 1996.

[63] Devillers, J., and Baladan, A., Eds. *Topological Indices and Related Descriptors in QSAR and QSPAR.* Gordon and Breach, Newark, 2000.

[64] Devillez, G. *Proofs by Transformation in Extremal Graph Theory.* PhD thesis, University of Mons, 2022.

[65] Devillez, G., D'hondt, S., and Goedgebeur, J. Knowledge management in House of Graphs. *arXiv preprint arXiv:2603.23070* (2026).

[66] Devillez, G., Hauweele, P., and Mélot, H. PHOEG Helps to Obtain Extremal Graphs. In *Operations Research Proceedings 2018 (GOR (Gesellschaft fuer Operations Research e.V.))* (sept. 12-14 2019), B. Fortz and M. Labbé, Eds., Springer, Cham, p. 251 (Paper 32).

[67] Devillez, G., Hertz, A., Mélot, H., and Hauweele, P. Minimum eccentric connectivity index for graphs with fixed order and fixed number of pendant vertices. *Yugoslav Journal of Operations Research 29*, 2 (2019), 193–202.

[68] Diestel, R. *Graph Theory*, 6th ed. Springer-Verlag, 2025.

[69] Došlić, T., Németh, L., and Podrug, L. Hosoya index of thorny polymers. *Discrete applied mathematics 343* (2024), 277–287.

[70] Došlić, T., and Short, T. Maximal matchings in polyspiro and benzenoid chains. *Applicable Analysis and Discrete Mathematics 15*, 1 (2021), 179–200.

[71] Došlić, T., and Zubac, I. Counting maximal matchings in linear polymers. *Ars Mathematica Contemporanea 11*, 2 (2015), 255–276.

[72] Doyle, J. K., and Graver, J. E. Mean distance in a graph. *Discrete Mathematics 17*, 2 (1977), 147–154.

[73] Duncan, B. Bell and Stirling numbers for disjoint unions of graphs. *Congressus Numerantium 206* (01 2010).

[74] Duncan, B., and Peele, R. B. Bell and Stirling numbers for graphs. *Journal of Integer Sequences 12* (2009). Article 09.7.1.

[75] Dusollier, V., Bonte, S., Devillez, G., Hertz, A., Mélot, H., and Schindl, D. Complete polyhedral description of chemical graphs of maximum degree at most 3. *arXiv preprint arXiv:2506.19768* (2025).

[76] Dyer, M., and Frieze, A. Randomized greedy matching. *Random Structures & Algorithms 2*, 1 (1991), 29–45.

[77] Dyer, M., Frieze, A., and Pittel, B. The average performance of the greedy matching algorithm. *The Annals of Applied Probability* (1993), 526–552.

[78] Elumalai, S., Hosamani, S. M., Mansour, T., and Rostami, M. A. More on inverse degree and topological indices of graphs. *Filomat 32*, 1 (2018), 165–178.

[79] Euler, L. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae* (1741), 128–140.

[80] Fajtlowicz, S. On conjectures of graffiti. In *Annals of Discrete Mathematics*, vol. 38. Elsevier, 1988, pp. 113–118.

[81] Falahati-Nezhad, F., Azari, M., and Došlić, T. Sharp bounds on the inverse sum indeg index. *Discrete Applied Mathematics 217* (2017), 185–195.

[82] Flajolet, P., and Sedgewick, R. *Analytic Combinatorics*. Cambridge University Press, 2009.

[83] Flory, P. J. Intramolecular reaction between neighboring substituents of vinyl polymers. *Journal of the American Chemical Society 61*, 6 (1939), 1518–1521.

[84] Fowler, P. Private communication, 2024.

[85] Fox, J., and Wei, F. On the number of cliques in graphs with a forbidden minor. *Journal of Combinatorial Theory, Series B 126* (2017), 175–197.

[86] Fox, J., and Wei, F. On the number of cliques in graphs with a forbidden subdivision or immersion. *SIAM Journal on Discrete Mathematics 34*, 4 (2020), 2556–2582.

[87] Friedgut, E., and Kahn, J. On the number of hamiltonian cycles in a tournament. *Combinatorics, Probability and Computing 14*, 5-6 (2005), 769–781.

[88] Frieze, A., Radcliffe, A. J., and Suen, S. Analysis of a simple greedy matching algorithm on random cubic graphs. *Combinatorics, Probability and Computing 4*, 1 (1995), 47–66.

[89] Furtula, B., Graovac, A., and Vukičević, D. Augmented zagreb index. *Journal of Mathematical Chemistry 48* (2010), 370–380.

[90] Galvin, D., and Marmorino, P. Counting independent sets in regular graphs with bounded independence number. *arXiv preprint arXiv:2410.19959* (2024).

[91] Galvin, D.and Thanh, D.T. Stirling numbers of forests and cycles. *Electronic Journal of Combinatorics 20* (2013). Paper P73.

[92] Gan, W., Loh, P.-S., and Sudakov, B. Maximizing the number of independent sets of a fixed size. *Combinatorics, Probability and Computing 24*, 3 (2015), 521–527.

[93] Ghorbani, M., Zangi, S., and Amraei, N. New results on symmetric division deg index. *Journal of Applied Mathematics and Computing 65*, 1-2 (2021), 161–176.

[94] Goedgebeur, J., Meersman, B., and Zamfirescu, C. Graphs with few hamiltonian cycles. *Mathematics of Computation 89*, 322 (2020), 965–991.

[95] Goel, G., and Tripathi, P. Matching with our eyes closed. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science* (2012), IEEE, pp. 718–727.

[96] Gonen, M., Ron, D., and Shavitt, Y. Counting stars and other small subgraphs in sublinear-time. *SIAM Journal on Discrete Mathematics 25*, 3 (2011), 1365–1411.

[97] Guo, X., and Gao, Y. Arithmetic-geometric spectral radius and energy of graphs. *MATCH Communications in Mathematical and in Computer Chemistry 83* (2020), 651–660.

[98] Gutman, I. Distance of thorny graphs. *Publications de l'Institut Mathématique 63(77)*, 83 (1998), 31–36.

[99] Gutman, I. Degree-based topological indices. *Croatica Chemica Acta 86* (2013), 351–361.

[100] Gutman, I. Geometric approach to degree-based topological indices: Sombor indices. *MATCH Communications in Mathematical and in Computer Chemistry 86* (2021), 11–16.

[101] Gutman, I. Relation between geometric–arithmetic and arithmetic–geometric indices. *Journal of Mathematical Chemistry 59*, 6 (2021), 1520–1525.

[102] Gutman, I., and Das, K. C. The first zagreb index 30 years after. *MATCH Communications in Mathematical and in Computer Chemistry 50*, 1 (2004), 83–92.

[103] Gutman, I., Furtula, B., and Elphick, C. Three new/old vertex-degree-based topological indices. *MATCH Communications in Mathematical and in Computer Chemistry 72* (2014), 617–632.

[104] Gutman, I., Furtula, B., and Redžepović, C. On topological indices and their reciprocals. *MATCH Communications in Mathematical and in Computer Chemistry 91* (2024), 287–297.

[105] Gutman, I., Miljković, O., Caporossi, G., and Hansen, P. Alkanes with small and large randić connectivity indices. *Chemical Physics Letters 306*, 5 (1999), 366–372.

[106] Hansch, C., Leo, A., and Hoekman, D. *Exploring QSAR, Vols. I and II.* ACS Professional Reference Book, Oxford University Press, New York, 1995.

[107] Hansen, P., Aouchiche, M., Caporossi, G., Mélot, H., and Stevanovic, D. What forms do interesting conjectures have in graph theory? *Dimacs series in discrete mathematics and theoretical computer science 69* (2005), 231.

[108] Hansen, P., Caporossi, G., Hertz, A., Aouchiche, M., Vukicević, D., and Sellal, C. Edge realizability of connected simple graphs. *MATCH Communications in Mathematical and in Computer Chemistry 78* (2017), 689–712.

[109] Hansen, P., and Mélot, H. Variable neighborhood search for extremal graphs. IX. bounding the irregularity of a graph. In *Graphs and Discovery*, vol. 69 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* Amer. Math. Soc., Providence, RI, 2005, pp. 253–264.

[110] Hansen, P., and Vukicević, D. Variable neighborhood search for extremal graphs. 23. on the randić index and the chromatic number. *Discrete Mathematics 309*, 13 (2009), 4228–4234.

[111] Hauweele, P., Hertz, A., Mélot, H., Ries, B., and Devillez, G. Maximum eccentric connectivity index for graphs with given diameter. *Discrete Applied Mathematics 268* (2019), 102–111.

[112] Haythorpe, M. On the minimum number of hamiltonian cycles in regular graphs. *Experimental mathematics 27*, 4 (2018), 426–430.

[113] Henning, M. A., and Rautenbach, D. On the irregularity of bipartite graphs. *Discrete Mathematics 307*, 11-12 (2007), 1467–1472.

[114] Hertz, A., Bonte, S., Devillez, G., Dusollier, V., Mélot, H., and Schindl, D. Extremal chemical graphs for the arithmetic-geometric index. *MATCH Communications in Mathematical and in Computer Chemistry 93*, 3 (2025), 791–818.

[115] Hertz, A., Bonte, S., Devillez, G., and Mélot, H. The average size of maximal matchings in graphs. *Journal of Combinatorial Optimization 47*, 3 (2024), 46.

[116] Hertz, A., Hertz, A., and Mélot, H. Using graph theory to derive inequalities for the bell numbers. *Journal of Integer Sequences 24* (2021). Article 21.10.6.

[117] Hertz, A., and Mélot, H. Counting the number of non-equivalent vertex colorings of a graph. *Discrete Applied Mathematics 203* (2016), 62–71.

[118] Hertz, A., Mélot, H., Bonte, S., and Devillez, G. Lower bounds and properties for the average number of colors in the non-equivalent colorings of a graph. *Discrete Applied Mathematics* (2022).

[119] Hertz, A., Mélot, H., Bonte, S., Devillez, G., and Hauweele, P. Upper bounds on the average number of colors in the non-equivalent colorings of a graph. *Graphs and Combinatorics 39*, 3 (2023), 49.

[120] Heuberger, C., and Wagner, S. G. Maximizing the number of independent subsets over trees with bounded degree. *Journal of Graph Theory 58*, 1 (2008), 49–68.

[121] Heuberger, C., and Wagner, S. G. Chemical trees minimizing energy and hosoya index. *Journal of mathematical chemistry 46* (2009), 214–230.

[122] Hierholzer, C., and Wiener, C. Über die möglichkeit, einen linienzug ohne wiederholung und ohne unterbrechung zu umfahren. *Mathematische Annalen 6*, 1 (1873), 30–32.

[123] Horák, P., and Stacho, L. A lower bound on the number of hamiltonian cycles. *Discrete Mathematics 222*, 1-3 (2000), 275–280.

[124] Hosoya, H. Topological index. a newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bulletin of The Chemical Society of Japan 44* (1971), 2332–2339.

[125] Hosoya, H. Topological index as a common tool for quantum chemistry, statistical mechanics, and graph theory. *Mathematical and computational concepts in chemistry (Dubrovnik, 1985), Ellis Horwood Ser. Math. Appl* (1986), 110–123.

[126] Hou, Y. On acyclic systems with minimal hosoya index. *Discrete Applied Mathematics 119*, 3 (2002), 251–257.

[127] Huang, Y., Shi, L., and Xu, X. The hosoya index and the merrifield–simmons index. *Journal of Mathematical Chemistry 56* (2018), 3136–3146.

[128] Huntemann, S., and McKay, N. A. Counting domineering positions. *Journal of Integer Sequences 24* (2021). Article 21.4.8.

[129] Ilić, A. On the extremal properties of the average eccentricity. *Computers & Mathematics with Applications 64*, 9 (2012), 2877–2885.

[130] Jackson, J. L., and Montroll, E. W. Free Radical Statistics. *The Journal of Chemical Physics 28*, 6 (1958), 1101–1109.

[131] Karelson, M. *Molecular Descriptors in GSAR/QSPR*. Wiley, New York, 2000.

[132] Kelmans, A. K. Comparison of graphs by their number of spanning trees. *Discrete Mathematics 16*, 3 (1976), 241–261.

[133] Kelmans, A. K. On graphs with the maximum number of spanning trees. *Random Structures & Algorithms 9*, 1-2 (1996), 177–192.

[134] Kereskényi-Balogh, Z., and Nyul, G. Stirling numbers of the second kind and Bell numbers for graphs. *Australasian Journal of Combinatorics 58* (2014), 264–274.

[135] Kier, L.B., and Hall, L.H. *Molecular Connectivity in Chemistry and Drug Research*. Academic Press, New York, 1976.

[136] Kier, L.B., and Hall, L.H. *Molecular Connectivity in Structure-Activity Analysis*. Wiley, New York, 1986.

[137] Kirchhoff, G. Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird. *Annalen der Physik 148*, 12 (1847), 497–508.

[138] Knopfmacher, A., Tichy, R. F., Wagner, S., and Ziegler, V. Graphs, partitions and fibonacci numbers. *Discrete Applied Mathematics 155*, 10 (2007), 1175–1187.

[139] Korfhage, R. R. σ-polynomials and graph coloring. *Journal of Combinatorial Theory, Series B 24*, 2 (1978), 137–153.

[140] Kouider, M., and Winkler, P. Mean distance and minimum degree. *Journal of Graph Theory 25*, 1 (1997), 95–99.

[141] Krivelevich, M., Mészáros, T., Michaeli, P., and Shikhelman, C. Greedy maximal independent sets via local limits. *Random Structures & Algorithms 64*, 4 (2024), 986–1015.

[142] Lee, C., and Oum, S.-i. Number of cliques in graphs with a forbidden subdivision. *SIAM Journal on Discrete Mathematics 29*, 4 (2015), 1999–2005.

[143] Li, G., Luo, J., Xiao, Z., and Liang, C. Mtmo: an efficient network-centric algorithm for subtree counting and enumeration. *Quantitative Biology 6*, 2 (2018), 142–154.

[144] Li, G., and Zhang, M. Sharp bounds on the arithmetic–geometric index of graphs and line graphs. *Discrete Applied Mathematics 318* (2022), 47–60.

[145] Li, H., Patterson, S., Yi, Y., and Zhang, Z. Maximizing the number of spanning trees in a connected graph. *IEEE Transactions on Information Theory 66*, 2 (2019), 1248–1260.

[146] Li, S., Wang, Z., and Zhang, M. On the extremal sombor index of trees with a given diameter. *Applied Mathematics and Computation 416* (2022), 126731.

[147] Li, X., and Gutman, I. Mathematical aspects of Randić-type molecular structure descriptors. *Croatica Chemica Acta 79* (01 2006).

[148] Li, X., Li, Z., and Wang, L. The inverse problems for some topological indices in combinatorial chemistry. *Journal of Computational Biology 10*, 1 (2003), 47–55.

[149] Li, X., and Shi, Y. A survey on the randić index. *MATCH Communications in Mathematical and in Computer Chemistry 59*, 1 (2008), 127–156.

[150] Li, X., Zhao, H., and Gutman, I. On the merrifield-simmons index of trees. *MATCH Communications in Mathematical and in Computer Chemistry 54*, 2 (2005), 389–402.

[151] Liu, H., You, L., and Huang, Y. Ordering chemical graphs by sombor indices and its applications. *MATCH Communications in Mathematical and in Computer Chemistry 87* (2022), 5–22.

[152] Liu, H., You, L., Tang, Z., and Liu, J.-B. On the reduced sombor index and its applications. *MATCH Communications in Mathematical and in Computer Chemistry 86* (2021), 729–753.

[153] Liu, J., Matejic, M., Milovanovic, E., and Milovanovic, I. Some new inequalities for the forgotten topological index and coindex of graphs. *MATCH Communications in Mathematical and in Computer Chemistry 84*, 3 (2020), 719–738.

[154] Liu, J.-B., and Daoud, S. N. Number of spanning trees in the sequence of some graphs. *Complexity 2019*, 1 (2019), 4271783.

[155] Liu, J.-B., Zhao, J., Min, J., and Cao, J. The hosoya index of graphs formed by a fractal graph. *Fractals 27*, 08 (2019), 1950135.

[156] Liu, Y., Zhuang, W., and Liang, Z. Largest hosoya index and smallest merrifield–simmons index in tricyclic graphs. *MATCH Communications in Mathematical and in Computer Chemistry 73*, 1 (2015), 195–224.

[157] Lovász, L., and Plummer, M. D. *Matching theory*, vol. 367. American Mathematical Soc., 2009.

[158] Luo, R. The maximum number of cliques in graphs without long cycles. *Journal of Combinatorial Theory, Series B 128* (2018), 219–226.

[159] Magun, J. Greedy matching algorithms, an experimental study. *Proceedings of the 1 st Workshop on Algorithm Engineering 6* (1997), 22–31.

[160] Mc Kay, B. Graph6 format. http://cs.anu.edu.au/ bdm/data/formats.html.

[161] McKay, B. D. *Nauty user's guide (version 2.4)*. Computer Science Dept., Australian National University, 2007.

[162] McKay, B. D., and Piperno, A. Practical graph isomorphism, ii. *Journal of Symbolic computation 60* (2014), 94–112.

[163] Mélot, H. Facet defining inequalities among graph invariants: the system GraPHedron. *Discrete Applied Mathematics 156* (2008), 1875 – 1891.

[164] Merrifield, R. E., and Simmons, H. E. Topological methods in chemistry. *(No Title)* (1989).

[165] Meta Platforms, Inc. React: The library for web and native user interfaces. `https://react.dev/`, 2026. Accessed: March 27, 2026.

[166] Miller, Z., and Pritikin, D. On randomized greedy matchings. *Random Structures & Algorithms 10*, 3 (1997), 353–383.

[167] Milovanović, I., Matejić, M., and Milovanović, E. Upper bounds for arithmetic-geometric index of graphs. *Scientific Publications of the State University of Novi Pazar Series A: Applied Mathematics, Informatics and mechanics 10*, 1 (2018), 49–54.

[168] Mohar, B. Eigenvalues, diameter, and mean distance in graphs. *Graphs and combinatorics 7*, 1 (1991), 53–64.

[169] Mohr, E., and Rautenbach, D. On the maximum number of maximum independent sets. *Graphs and Combinatorics 34* (2018), 1729–1740.

[170] Mukwembi, S., and Munyira, S. Degree distance and minimum degree. *Bulletin of the Australian Mathematical Society 87*, 2 (2013), 255–271.

[171] Myers, B. Number of spanning trees in a wheel. *IEEE Transactions on Circuit Theory 18*, 2 (1971), 280–282.

[172] Nikiforov, V. The number of cliques in graphs of given order and size. *Transactions of the American Mathematical Society 363*, 3 (2011), 1599–1618.

[173] Nikolić, S., Kovačević, G., Miličević, A., and Trinajstić, N. The zagreb indices 30 years after. *Croatica Chemica Acta 76*, 2 (2003), 113–124.

[174] Niu, B., Zhou, S., and Zhang, H. Extremal arithmetic–geometric index of bicyclic graphs. *Circuits, Systems, and Signal Processing* (2023), 1–22.

[175] Odlyzko, A., and Richmond, L. On the number of distinct block sizes in partitions of a set. *Journal of Combinatorial Theory, Series A 38*, 2 (1985), 170–181.

[176] Ore, O. Theory of graphs. In *Colloquium Publications* (1962), American Mathematical Society.

[177] Osaye, F. J. The average eccentricity of a graph with prescribed girth. *Discrete Mathematics 345*, 12 (2022), 113066.

[178] Pedersen, A. S., and Vestergaard, P. D. The number of independent sets in unicyclic graphs. *Discrete applied mathematics 152*, 1-3 (2005), 246–256.

[179] Pedersen, A. S. P., and Vestergaard, P. D. Bounds on the number of vertex independent sets in a graph. *Taiwanese Journal of Mathematics 10*, 6 (2006), 1575–1587.

[180] Petingi, L., Boesch, F., and Suffel, C. On the characterization of graphs with maximum number of spanning trees. *Discrete mathematics 179*, 1-3 (1998), 155–166.

[181] Poloczek, M., and Szegedy, M. Randomized greedy algorithms for the maximum matching problem with new analysis. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science* (2012), IEEE, pp. 708–717.

[182] Prodinger, H., and Tichy, R. F. Fibonacci numbers of graphs. *The Fibonacci Quarterly 20*, 1 (1982), 16–21.

[183] Randić, M. Characterization of molecular branching. *Journal of the American Chemical Society 97*, 23 (1975), 6609–6615.

[184] Rautenbach, D. A note on trees of maximum weight and restricted degrees. *Discrete mathematics 271*, 1-3 (2003), 335–342.

[185] Raza, Z., and Ali, A. Bounds on the zagreb indices for molecular (n, m)-graphs. *International Journal of Quantum Chemistry 120*, 18 (2020), e26333.

[186] Ribeiro, P., Paredes, P., Silva, M. E., Aparicio, D., and Silva, F. A survey on subgraph counting: concepts, algorithms, and applications to network motifs and graphlets. *ACM computing surveys (csur) 54*, 2 (2021), 1–36.

[187] Rodríguez, J. M., and Sigarreta, J. M. Spectral properties of geometric-arithmetic index. *Applied Mathematics and Computation 277* (2016), 142–153.

[188] Sah, A., Sawhney, M., Stoner, D., and Zhao, Y. The number of independent sets in an irregular graph. *Journal of Combinatorial Theory, Series B 138* (2019), 172–195.

[189] Samotij, W. Counting independent sets in graphs. *European journal of combinatorics 48* (2015), 5–18.

[190] Sárközy, G. N., Selkow, S. M., and Szemerédi, E. On the number of hamiltonian cycles in dirac graphs. *Discrete Mathematics 265*, 1-3 (2003), 237–250.

[191] Schank, T., and Wagner, D. Finding, counting and listing all triangles in large graphs, an experimental study. In *International workshop on experimental and efficient algorithms* (2005), Springer, pp. 606–609.

[192] Sedlar, J., Stevanović, D., and Vasilyev, A. On the inverse sum indeg index. *Discrete Applied Mathematics 184* (2015), 202–212.

[193] Shegehalli, V., and Kanabur, R. Arithmetic-geometric indices of path graph. *Journal of Mathematics and Computer Science 16* (2015), 19–24.

[194] Shegehalli, V., and Kanabur, R. Arithmetic-geometric indices of some class of graph. *Journal of Computer and Mathematical sciences 6*, 4 (2015), 194–199.

[195] Shigehalli, V., and Kanabur, R. Computation of new degree-based topological indices of graphene. *Journal of Mathematics* (2016), Article ID 4341919.

[196] Short, T., and Ash, Z. The number of maximal matchings in polyphenylene chains. *Iranian Journal of Mathematical Chemistry 10*, 4 (2019), 343–360.

[197] Sridhara, V., Kanna, M. R. R., and Indumathi, R. S. Computation of topological indices of graphene. *Journal of Nanomaterials* (2015), Article ID 969348.

[198] Stevanovic, D., and Brankov, V. An invitation to newgraph. *Rendiconti del Seminario Matematico di Messina, Serie II 25* (2003), 9.

[199] Swartz, E., and Vetrík, T. Survey on the general randić index: extremal results and bounds. *Rocky Mountain Journal of Mathematics 52*, 4 (2022), 1177–1203.

[200] Sylvester, J. J. Chemistry and algebra. *Nature 17*, 432 (1878), 284.

[201] The PostgreSQL Global Development Group. PostgreSQL: Open source relational database. `https://www.postgresql.org/`, 2026. Accessed: March 27, 2026.

[202] The Rust Project Developers. The Rust programming language. `https://www.rust-lang.org/`, 2026. Accessed: March 27, 2026.

[203] Thomassen, C. On the number of hamiltonian cycles in bipartite graphs. *Combinatorics, Probability and Computing 5*, 4 (1996), 437–442.

[204] Tichy, R. F., and Wagner, S. Extremal problems for topological indices in combinatorial chemistry. *Journal of Computational Biology 12*, 7 (2005), 1004–1013.

[205] Tinhofer, G. A probabilistic analysis of some greedy cardinality matching algorithms. *Annals of Operations Research 1*, 3 (1984), 239–254.

[206] Turán, P. Eine extremalaufgabe aus der graphentheorie. *Mat. Fiz. Lapok 48*, 436-452 (1941), 61.

[207] Van der Laan, H. *Le nombre plastique*. Brill Archive, 1997.

[208] Vestergaard, P. D., and Pedersen, A. S. An upper bound on the number of independent sets in a tree. *Ars Combinatoria 84* (2007), 85–96.

[209] Vujošević, S., Popivoda, G., Vukićević, Ž. K., Furtula, B., and Škrekovski, R. Arithmetic–geometric index and its relations with geometric–arithmetic index. *Applied Mathematics and Computation 391* (2021), 125706.

[210] Vukičević, D., and Furtula, B. Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges. *Journal of mathematical chemistry 46* (2009), 1369–1376.

[211] Vukićević, Ž. K., Vujošević, S., and Popivoda, G. Unicyclic graphs with extremal values of arithmetic–geometric index. *Discrete Applied Mathematics 302* (2021), 67–75.

[212] Wagner, S. Extremal trees with respect to hosoya index and merrifield-simmons index. *MATCH Communications in Mathematical and in Computer Chemistry 57*, 1 (2007), 221–233.

[213] Wagner, S., and Gutman, I. Maxima and minima of the hosoya index and the merrifield-simmons index: a survey of results and techniques. *Acta Applicandae Mathematicae 112* (2010), 323–346.

[214] Wang, H., and Hua, H. Unicycle graphs with extremal merrifield–simmons index. *Journal of mathematical chemistry 43*, 1 (2008), 202–209.

[215] Wang, M., Hua, H., and Wang, D. The first and second largest merrifield–simmons indices of trees with prescribed pendent vertices. *Journal of mathematical chemistry 43*, 2 (2008), 727–736.

[216] Wolsey, L. A. *Integer programming*. John Wiley & Sons, 2020.

[217] Wood, D. R. On the maximum number of cliques in a graph. *Graphs and Combinatorics 23*, 3 (2007), 337–352.

[218] Yannakakis, M., and Gavril, F. Edge dominating sets in graphs. *SIAM Journal on Applied Mathematics 38*, 3 (1980), 364–372.

[219] Yu, G., Feng, L., and Wang, D. On the average eccentricity of unicyclic graphs. *Ars Combinatoria 103* (2012), 531–537.

[220] Yuan, Y., Zhou, B., and Trinajstić, N. On geometric-arithmetic index. *Journal of Mathematical Chemistry 47*, 2 (2010), 833–841.

[221] Zhang, J.-b., Liu, Z.-z., and Zhou, B. On the maximal eccentric connectivity indices of graphs. *Applied Mathematics-A Journal of Chinese Universities 29*, 3 (2014), 374–378.

[222] Zhang, X.-M., Yang, Y., Wang, H., and Zhang, X.-D. Maximum atom-bond connectivity index with given graph parameters. *Discrete Applied Mathematics 215* (2016), 208–217.

[223] Zhang, Z., Wu, B., and Comellas, F. The number of spanning trees in apollonian networks. *Discrete applied mathematics 169* (2014), 206–213.

[224] Zheng, L., Tian, G.-X., and Cui, S.-Y. On spectral radius and energy of arithmetic-geometric matrix of graphs. *MATCH Communications in Mathematical and in Computer Chemistry 83* (2020), 635–650.

[225] Zhong, L. The harmonic index for graphs. *Applied Mathematics Letters 25*, 3 (2012), 561–566.

[226] Zuo, X., Jahanbani, A., and Shooshtari, H. On the atom-bond sum-connectivity index of chemical graphs. *Journal of Molecular Structure 1296* (2024), 136849.