Politecnico di Milano

School of Industrial and Information Engineering

Department of Chemistry, Materials and Chemical Engineering "Giulio Natta"

Master of Science in Materials Engineering and Nanotechnology

Computational study of the energy gap of benzenoid hydrocarbons through topological descriptors

Master Thesis of
Noemi Carol D’Ausilio
Student ID: 987322

Supervisor:
Professor Matteo Maria Saverio Tommasini
Co-supervisor:
Gianluca Serra

Academic Year 2022-2023

Abstract

The comprehension of the link between the topological properties and the electronic properties of materials is of primary interest. In this Thesis, the investigation regarded the HOMO-LUMO gap of benzenoid hydrocarbons, which are a class of organic molecules made of fused benzene rings.

The energy gap is an electronic property of molecules relevant to the determination of their reactivity and their optical characteristics. In the case of benzenoid hydrocarbons, properties such as aromaticity are influenced by the topological arrangement of the rings [1]. Therefore, for these compounds, it holds significance investigating the energy gap through topological descriptors within the framework of chemical graph theory.

Ciosłowski attempted to describe through topological descriptors the energy gap computed by Hückel theory for a set of benzenoid hydrocarbons with a number of rings between one and eight [2]. In the initial chapter, I adopt the same topological parameters introduced by Ciosłowski to predict the energy gap calculated by Density Functional Theory (DFT) for selected benzenoid hydrocarbons with a number of rings between one and six rings. Here, I give particular attention to non-planar molecules. Subsequently, I extend this approach to benzenoid hydrocarbons with more than eight rings, considering additional topological descriptors, to enhance the accuracy of HOMO-LUMO gap predictions for molecules with more than twenty rings. The correlation between the energy gap calculated using Hückel theory and the one estimated through DFT calculations is examined through the analysis of average carbon-carbon bond lengths. A robust correlation between the energy gap estimated with Hückel theory and the one calculated with DFT is demonstrated across all the analyzed classes of benzenoid hydrocarbons. Lastly, I bring back the concept of isoarithmicity observed by Balaban [3] for a subset of benzenoid hydrocarbons and I apply it to the HOMO-LUMO gap calculated with the Hückel theory. I also verified its validity for the HOMO-LUMO gap calculated with the DFT.

The results presented in this Thesis showcase the utility of concepts from chemical graph theory in predicting DFT-calculated properties with reduced computational effort.

Keywords: Benzenoid hydrocarbons, DFT, Hückel theory, energy gap, HOMO-LUMO gap.

Sommario

La comprensione del legame tra le caratteristiche topologiche e le proprietà elettroniche dei materiali è di primaria importanza. In questa Tesi, l'analisi è focalizzata sul gap HOMO-LUMO di idrocarburi benzenoidi aromatici, che sono una classe di composti organici costituiti da anelli di benzene condensati.

Il gap energetico è una proprietà elettronica delle molecole rilevante per lo studio della loro reattività e delle loro proprietà ottiche. Nel caso di idrocarburi benzenoidi, proprietà come l'aromaticità sono influenzate dalla disposizione degli anelli nella catena [1]. Quindi, nell'ambito della teoria chimica dei grafi, è pertinente investigare il gap energetico in termini di descrittori topologici.

Ciosłowski ha espresso tramite descrittori topologici il gap energetico calcolato con la teoria di Hückel per una serie di idrocarburi benzenoidi con un numero di anelli compreso tra uno e otto [2]. Nel capitolo iniziale, gli stessi descrittori topologici sono usati per predire l'HOMO-LUMO gap calcolato con la teoria del funzionale della densità (*Density Functional Theory*, DFT) per una selezione di idrocarburi benzenoidi con un numero di anelli compreso tra uno e sei, con particolare attenzione alle molecole non planari. Quindi, questo approccio è stato esteso alle molecole con più di otto anelli, considerando ulteriori descrittori topologici per migliorare l'accuratezza delle previsioni del gap HOMO-LUMO per molecole con più di venti anelli.

La correlazione tra il gap calcolato con la teoria di Hückel e quello calcolato con la DFT è esaminata tramite l'analisi della lunghezza media del legame carbonio-carbonio. Per tutte le classi di idrocarburi benzenoidi analizzate, è dimostrata una robusta correlazione tra il gap energetico stimato con la teoria di Hückel e quello calcolato con la DFT. Infine, riprendiamo il concetto di isoaritmicità, osservato da Balaban [3] per un sottoinsieme di idrocarburi benzenoidi e precedentemente applicato per il gap HOMO-LUMO calcolato con la teoria di Hückel e ne verifichiamo la validità anche per il gap HOMO-LUMO calcolato con la DFT.

I risultati presentati nella Tesi mostrano l'utilità dei concetti di teoria chimica dei grafi nello studio di proprietà calcolate tramite DFT, al fine di ridurre lo sforzo computazionale.

Keywords: idrocarburi benzenoidi, DFT, teoria di Hückel , gap HOMO-LUMO, gap energetico.

Contents

Abstract ii

Sommario iii

List of Figures viii

List of Tables xiii

1 Introduction 1
- 1.1 Polycyclic aromatic hydrocarbons 1
- 1.2 Density functional theory 4
- 1.3 Chemical graph theory 6

2 Methods 8
- 2.1 Computational methods 8
  - 2.1.1 The Born-Oppenheimer approximation 9
  - 2.1.2 Mean field approximation 10
  - 2.1.3 Density functional theory 10
- 2.2 Notions of graph theory 11
  - 2.2.1 Computation of the number of Kekulé structures 17
- 2.3 Hückel theory 21
- 2.4 Computational tools 23

3 Energy gap of BHs 24
- 3.1 Ciosłowski formula 24
- 3.2 Comparison between Hückel theory and DFT 38

4 Isoarithmicity 52

5 Conclusion 58

References 61

A Appendix 75

List of Figures

1.1 Layers and the corresponding energy level diagrams in OSCs and OLEDs . 2

1.2 Three representative structures of different PAHs. 2

1.3 Number of papers that contain the word "DFT" in their keyword list, title, or abstract from 2004 to 2022. Data from Web of Science. 4

2.1 Equivalent representations of pyrene 12

2.2 The molecular graph and the dual graph of naphthalene. Grey circles represent the vertices of the molecular graph, black segments the edges of the molecular graph, purple circles are the vertices of the dual graph and purple segments are the edges of the dual graph. 13

2.3 Particular perimeter configurations in BHs. In Panel **(a)**, **(b)**, **(c)**, the bay, cove, and fjord conformation are represented. In each panel, the thick red segments highlight the illustrated perimeter conformation, Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph, purple circles are the vertices of the dual graph and purple segments are the edges of the dual graph. 14

2.4 On the left, the molecular graph of a catacondensed BH associated to a tetrahex. On the right, the molecular graph of a pericondensed BH associated to a tetrahex. 14

2.5 Particular types of ring and their notation. In the first row, from left to right, L_1 and L_2 types of ring and their adjacent rings. On the second row, from left to right, A_2 and A_3 rings and their adjacent rings. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph, and purple segments are the edges of the dual graph. 15

2.6 The three perfect matches of the only possible dihex. 16

2.7 The three Kekulé structures of naphtalene. 17

2.8 Example of David-Gordon algorithm 18

2.9 The molecular graph of a BH equivalent to a prolate rectangle $P(p, q)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph. 18

2.10 The molecular graph of a BH equivalent to a parallelogram $L(p,q)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph. 19

2.11 The molecular graph of a BH equivalent to an oblate rectangle $O(p,q)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph. 20

2.12 The molecular graph of a BH equivalent to a hexagon $H(p,q,r)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph. 21

3.1 In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap estimated by the Ciosłowski formula χ_{HL} in equation 3.1 for BHs associated to n-hexes with $1 \leq n \leq 6$. The red line is the bisector of the first and third quadrants. 25

3.2 In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the DFT theory and the HOMO-LUMO gap estimated by the Ciosłowski formula χ_{HL} in equation 3.2 for BHs associated to n-hexes with $1 \leq n \leq 6$. The red line is the bisector of the first and third quadrants. 25

3.3 Planar and non-planar molecules with their relative error. 27

3.4 In the first row, isodensity surface plot at the value $0.0223a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent to parallelogram $L(2,3)$. In the second row, isodensity surface plot at the value $0.0116a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent to parallelogram $L(7,7)$. . . 29

3.5 In the first row, isodensity surface plot at the value $0.0115a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent to oblate rectangles $O(3,2)$. In the second row, isodensity surface plot at the value $0.0179a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent oblate rectangles $O(6,5)$. 30

3.6 In the first row, isodensity surface plot at the value $0.0177a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent to prolate rectangles $P(3,3)$. In the second row, isodensity surface plot at the value $0.0129a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent to prolate rectangles $P(5,5)$. 31

3.7 In the first row, isodensity surface plot at the value $0.0164a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent of hexagon $H(2,3,3)$. In the second row, isodensity surface plot at the value $0.0112a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent of hexagon $H(5,5,5)$. 32

3.8 In Panel **(a)**, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with DFT and the HOMO-LUMO gap estimated by the Ciosłowski formula χ_{HL} in equation 3.3 for BHs associated to parallelograms in green diamonds, hexagons in green asterisks, prolate rectangles in blue circles and oblate rectangles in blue asterisks. The red line is the bisector of the first and third quadrants. In Panel **(b)**, the relative error related to BHs associated to parallelograms in green diamonds, hexagons in green asterisks, prolate rectangles in blue circles and oblate rectangles in blue asterisks estimated by equation 3.3. 34

3.9 In blue circles, $K^{2/N}$ over p for parallelogram $L(p,p)$, with p going from 2 to 38. 36

3.10 At the top, molecular graphs of BHs associated to oblate rectangle $O(2,2)$ on the left and oblate rectangle $O(6,4)$ on the right are displayed. The investigated outer path is highlighted in blue, while the inner path is marked in green. At the bottom, the C-C bond alternation plot along the outer path and the inner path of the molecule is illustrated. 46

3.11 At the top, molecular graphs of BHs of BHs associated to parallelogram $L(2,3)$ on the left and parallelogram $L(6,7)$ on the right are displayed. The investigated outer path is highlighted in blue, while the inner path is marked in green. At the bottom, the C-C bond alternation plot along the outer path and the inner path of the molecule is illustrated. 47

3.12 At the top, molecular graphs of BHs associated to prolate $P(3,2)$ on the left and parallelogram $P(6,4)$ on the right are displayed. The investigated outer path is highlighted in blue, while the inner path is marked in green. At the bottom, the C-C bond alternation plot along the outer path and the inner path of the molecule is illustrated. 48
3.13 In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to parallelograms. 49
3.14 In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to oblate rectangles. The optimized linear fitting coefficient is 5.1093. 49
3.15 In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to prolate rectangles. The optimized linear fitting coefficient is 4.822. 50
3.16 In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to hexagon. The optimized linear fitting coefficient is 4.903. 50
4.1 The molecular graphs of two isoarithmic molecules. The two unbranched catacondensed BHs are associated to the same dual graphs and the same "LA" sequences but their dualist graphs are different (see in Appendix A **4.03** and **4.04**). Grey circles represent the vertices of the molecular graph, black segments the edges of the molecular graph, purple circles are the vertices of the dual graph and purple segments are the edges of the dual graph. 52

List of Tables

3.1 Comparison between the correlation coefficients of the Ciosłowski formula χ_{HL} in equation 3.1 and HOMO-LUMO gap calculated with Hückel theory and the correlation coefficients of the modified formula χ_{HL} in equation 3.2 and HOMO-LUMO gap calculated with DFT theory. 26

3.2 Correlation coefficients of the HOMO-LUMO gap calculated with DFT for all the analyzed molecules equivalent to prolate rectangles, oblate rectangles, parallelograms, and hexagons with the inverse of the perimeter, the inverse of the number of carbon atoms, the ratio between perimeter and the number of internal edges, and the ratio between the perimeter and the number of carbon atoms. 37

3.3 Correlation coefficients of the HOMO-LUMO gap calculated with DFT for molecules with more than 20 rings equivalent to prolate rectangles, oblate rectangles, parallelograms, and hexagons with the inverse of the perimeter, the inverse of the number of carbon atoms, the ratio between perimeter and the number of internal edges, and the ratio between the perimeter and the number of carbon atoms. 38

3.4 In the second column, the average mean value of the bond length for each molecule associated to a parallelogram. In the third column, the maximum difference in absolute value between the average value of the bond lengths and the bond lengths themselves. 40

3.5 In the second column, the average mean value of the bond length for each molecule associated to a oblate rectangle. In the third column, the maximum difference in absolute value between the average value of the bond length and the bond lengths themselves. 42

3.6 In the second column, the average mean value of the bond length for each molecule associated to a prolate rectangle. In the third column, the maximum difference in absolute value between the average value of the bond lengths and the bond lengths themselves. 43

3.7 In the second column, the average mean value of the bond length for each molecule associated to a hexagon. In the third column, the maximum difference in absolute value between the average value of the bond lengths and the bond lengths themselves. 44

4.1 HOMO-LUMO gap evaluated at the Hückel theory level of unbranched catacondensed BHs of four, five, and six rings. We group them by "*LA*" sequences, reported in the first column. The fourth column contains the mean value of the HOMO-LUMO gap of each isoarithmic BH. The last column shows the difference in absolute value between the HOMO-LUMO gap of each group of isoarithmic BH. 54

4.2 HOMO-LUMO gap evaluated at the DFT level of unbranched catacondensed BHs of four, five, and six rings. We group them by "*LA*" sequences, reported in the first column. The fourth column contains the mean value of the HOMO-LUMO gap of each isoarithmic BH. The last column shows the difference in absolute value between the HOMO-LUMO gap of each group of isoarithmic BH. 56

A.1 For each polyhex, in the first row he molecular graphs associate to the monohex [**1.01**], dihex [**2.01**], trihexes [**3.01**-**3.03**], tetrahexes [**4.01**-**4.05**], pentahexes [**5.01**-**5.12**] and hexahexes [**6.01**-**6.36**] are reported. Grey circles represent the vertices of the molecular graph, black segments the edges of the molecular graph, purple circles the vertices of the dual graph, and purple segments the edges of the dual graph. For each polyhex, an ID is assigned in the second row as outlined in Section 2.2. For each polyhex, the third row features the chain sequence for unbranched catacondensed BHs by recognizing the connectivity of each ring using the notation introduced in Section 2.2. 79

1| Introduction

1.1. Polycyclic aromatic hydrocarbons

According to Goals for Sustainable Developtment, global warming must not exceed 1.5 °C above pre-industrial levels, thus emission of carbon dioxide needs to be reduced by 45% by 2030 and reach net zero by 2050 [4]. Research in the latest decades has focused on organic semiconductors to cut emissions. These materials lead to new flexible [5–8], stretchable [9–12], low-weight [8, 13, 14], and low-cost fabrication over large-areas devices [8, 13–15], such as organic solar cells (OSCs) [16–21], organic field effect transistors (OFETs) [15, 22–27], and organic light emitting diodes (OLEDs) [16, 28–31]. In every type of device, the pairing between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) is of primary importance. Their difference is defined as the HOMO-LUMO gap.

OSCs are third-generation solar cells characterized by a donor and an acceptor [32], responsible for harvesting light and generating excitons, sandwiched between two transport layers and two electrodes. The good pairing between the LUMO of the donor and the HOMO of the acceptor, as illustrated in Figure 1.1, allows to accomplish higher absorption, lower voltage losses, and lower molecular instability when exposed to humidity, light, or higher temperatures [33].

OFETs are three-terminal (source, drain, and gate) devices made of an organic active layer and a substrate devoted to the switch of current or signal. The compatibility of the Fermi energy level of the metal of the electrode with the HOMO level of the organic active layer in p-type OFETs (or with the LUMO level for n-type OFETs) enables the transport of the electron [34].

OLEDs are a mainstream light source made of a multilayer structure of organic thin film sandwiched between transparent conductive oxides and metallic cathodes, as shown in Figure 1.1. The injection of charges from the electrodes to the organic semiconductors permits radiative recombination, for example fluorescence [35] or phosphorescence [36], and so light emission. The alignment of the HOMO of emitting layer and the LUMO of the electron buffering layer leads to high efficiencies [36].

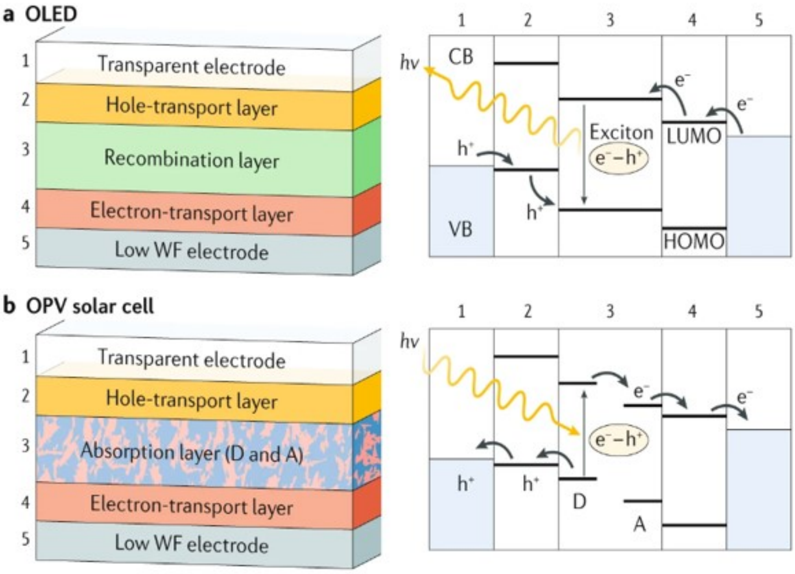

a OLED
1
2
3
4
5
Transparent electrode
Hole-transport layer
Recombination layer
Electron-transport layer
Low WF electrode
hv
CB
VB
h+
h+
Exciton
e−–h+
e−
e−
LUMO
HOMO
b OPV solar cell
Transparent electrode
Hole-transport layer
Absorption layer (D and A)
Electron-transport layer
Low WF electrode
e−–h+
D
A

Figure 1.1: On the left side, a scheme of the layers in OSCs and OLEDs. On the right side, the corresponding energy level diagram, underlining the importance of HOMO and LUMO. Reproduced from [37].

Polycyclic aromatic hydrocarbons (PAHs) and nanographenes (NGs) are ideal candidates for organic devices. PAHs and NGs are molecules whose defining feature is the systematic escalation of benzene rings along one, two, or three dimensions [38]. Three representative examples are shown in Figure 1.2.

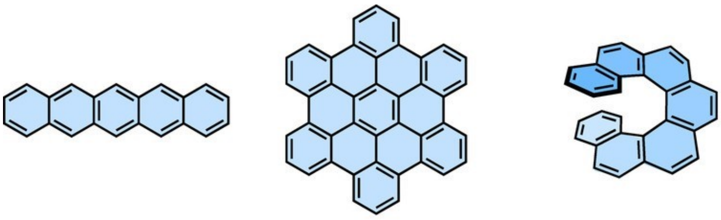


Figure 1.2: Three representative structures of different PAHs. From left to right, a one dimensional acene, a two dimensional NG and a three dimensional [n]helicene. Reproduced from [39].

They are synthesized through organic chemistry methods [40], starting from precursors, and obtaining the final compound through ring-closing, Diels-Alder polymerization and cycloaddition, and transition-metal catalyzed polycondensation [38].

PAHs and NGs are characterized by a π-conjugated structure resulting in strong intermolecular interaction and favorable solid stacking [41]. These molecules are chemically and thermally stable due to their high carbon content [42].

The HOMO-LUMO gap of PAHs and NGs is tuneable because they are easily functionalizable by replacing hydrogen atoms with heteroatoms or benzene rings with thiophene ones [43]. The gap also depends on their size and warping in non planar structures. For example, increasing the size of NGs results in a variation in the HOMO-LUMO gap, which in turn leads to a change in the position of the absorption bands [38]. For NGs, deviating from planarity allows the control of the band gap and the increase of processability [44]. The addition of a functional group by itself leads to a control of the curvature with a consequent decrease of the energy gap, as it happens for substituted ovalene [45], and better hole/electron transport material with high drift mobility for practical applications, as in contorted PAHs imides [46].

PAHs show a wide absorption spectrum with strong absorption in the visible light region due to their large π-conjugated structure, resulting in high performance in OSCs [41]. For instance, antranthrene derivatives as p-type material show a power conversion efficiency of 3.04% when coupled with a functionalized fullerene as an acceptor. Large PAHs have been tested in OSCs in combination with fullerene-based molecules and a power conversion efficiency value of 2.9% has been obtained. Their problems are solubility, the absence of chemical handles for further functionalization, and a missing marked donor or acceptor character [42].

Planar PAHs, such as pentacene and rubrene, show excellent p-type OFET behavior, with charge mobility greater than 10 $cm^2\ V^{-1}\ s^{-1}$, but also non planar molecules as contorted hexabenzocoronene can be further functionalized to achieve hole mobility up to 0.49 $cm^2\ V^{-1}\ s^{-1}$ [47]. Molecules based on coronene, properly functionalized with chlorine, fluorine, or tetraalkoxy substitution, show increased mobility, and in the last case, in the presence of carbon nanotubes as point contacts, efficient charge transport was realized and I_{on}/I_{off} current was detected [41].

In optoelectronics, large PAHs of dimensions comprised between 1 nm and 10 nm can be seen as quasi-zero-dimensional graphene quantum dots [40]. Some materials such as functionalized ovalene in polystyrene matrix [45] or PAHs doped with nitrogen are characterized by spontaneous emission. The latter shows a photoluminescence quantum yield of up to 86%, while the best-performing compound possessing an electron-withdrawing CF_3 group [48] shows an efficiency of 12% so they are suitable candidates for OLEDs or lasers. Also, PAH derivatives could constitute an ideal electron transport material [49].

1.2. Density functional theory

Density functional theory (DFT) is used to investigate the electronic properties of molecular and condensed systems. It relies on Hohenberg-Kohn theorems, which demonstrate that the electron density $\rho(x, y, z)$ is a key quantity to evaluate the ground state energy of a many-body system [50]. Subsequently, calculations are greatly simplified since the electron density depends just on three spatial variables, so it is a theory especially suitable for computer calculation. As reported in Figure 1.3, the amount of research related to DFT steadily increased over the years. Indeed, the seminal articles by Hohenberg, Sham, and Kohn are among the most cited papers of all time [51].

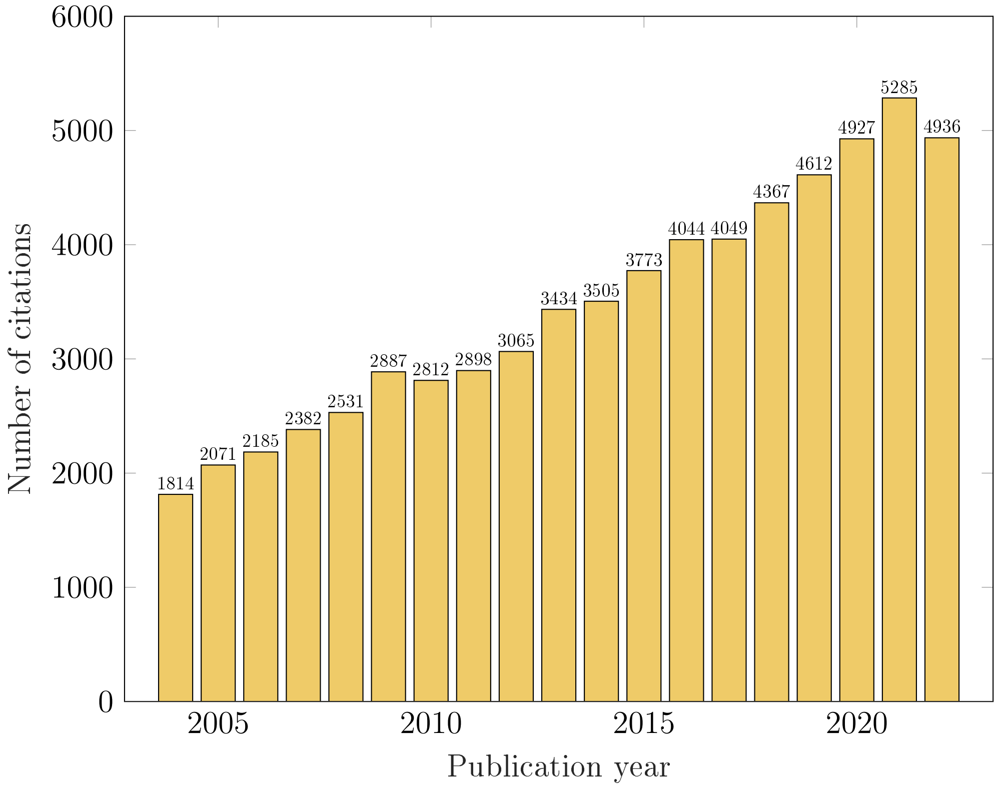


Figure 1.3: Number of papers that contain the word "DFT" in their keyword list, title, or abstract from 2004 to 2022. Data from Web of Science.

DFT is a widely used method in material science because it can accurately estimate the energies and the geometries of many quantum systems of practical interest [52]. We will now analyze some classes of materials studied through DFT, such as transition metals and conjugated polymers.

Transition metals are decisive for technological and medical applications. Studying their peculiar structural, optical, and magnetic properties with computational techniques is compelling to compare them with experimental evidences or to understand their working mechanisms. DFT methods can help in understanding the adsorption mechanism of molecules on transition metal particles [53].They can be used in nanomedicine as a contrasting agent, if they bind a ligand,. In that case, it is relevant to model metal-ligand interactions [54]. The variety of DFT methods that properly consider spin effects and evaluate excited state energies allows us to do that [53]. Furthermore, machine learning algorithms combined with the DFT-calculated energy band gap of functionalized transition metals help the design process of novel hybrid perovskite solar cells [55].

Moving to conjugated polymers, we recognize their importance in the field of organic electronics and, as seen before, a crucial parameter to consider is the energy band gap. The comparison of experimental and DFT results allows us to study the influence of the presence of donors or acceptors [56]. Furthermore, through computational techniques, it is possible to calculate the energy gaps of polyyne, polyacetylene, and polythiophene up to a certain length of the chain [57, 58]. However, polymers are also characterized by intramolecular interactions and intermolecular non-covalent bonds and one of the main DFT limitations is modeling such interactions. Improved functionals and methods are addressing this issue [59, 60].

DFT is also extensively applied to study properties of PAHs of different dimensionality, given their importance as functional materials. DFT simulation can help study their energy levels, which are relevant in determining their charge transport properties, useful in OFETs [61], and their behavior as materials for OSCs [62]. DFT serves to assess also other chemical properties, such as the electronic affinity, which finds confirmation in experimental data for linear PAHs [63], and isomerization energy, which regards reactions that lead one isomer to another and is estimated through more sophisticated DFT methods [64]. Moreover, DFT is efficient in the evaluation of Clar's aromaticity [65], a key property of PAHs, related to the shared π-electrons among the rings. Various aromaticity indices, accurately computed through DFT, provide insights into this phenomenon [66], especially in the presence of long-term interactions [67–69].

1.3. Chemical graph theory

Chemical graph theory is the application of graph theory to chemistry [3]. Different types of graphs can be obtained. For example, if these descriptors illustrate molecular topology we speak of molecular or constitutional graphs. Instead, when they are used to show the intermediate steps of a reaction, we can call them reaction graphs [3]. Here, we discuss some possible applications of graph theory, specifically those related to constitutional graphs, for different classes of molecules, to show the validity of this approach.

For instance, chemical graph theory is decisive in determining quantitative structural-property correlation. A possible technological application of this type of study is the discovery of new molecules and this is especially helpful in the design of new drugs [70]. Furthermore, it may lead to the discovery of novel chemical and physical properties of materials and to the prediction of already known material properties given the topological characteristics of the molecule itself [71].

One of the most important historical applications of chemical graph theory dates back to 1947 when Wiener studied the boiling point of different alkane chains using a linear formula containing two indices. These indices can be interpreted as physical quantities, respectively as the compactness of the molecule and the measurement of intramolecular attraction forces, so they are connected to the constitutional graph of the analyzed molecules. Indeed, these parameters consider the variation of the boiling temperature when in the presence of isomers. The error estimated between the experimental analysis and the theoretical prediction is less than 1°C [72]. A widely investigated property of polymers is the glass transition temperature, given its influence on processability and technological application [73]. A linear fitting considering some topological indices, such as Randić-Kier-Hall connectivity indices [74–76] and the Wiener index [72], leads to a coefficient of determination of the glass transition temperature with respect to the molecular mass of the polymers higher than 0.9, with a maximum residual error in the order of 15% compared to experimental results. The equivalence between chemical and graph theoretical descriptors has also been shown. This approach extended also to optical properties such as the refractive index [77].

Chemical graph theory provides information for a better understanding of drugs' mechanisms and also allows the description of the complex interaction between proteins [78]. In the case of viruses, it can furnish predictive methods for the interaction between the host and the virus [79]. For example, SARS-CoV-2 protein-protein interaction was studied

with this method, helping the development of drugs [80]. Additionally, host-virus interaction was modeled in the same way, also considering the presence of the vaccine [81].
PAHs are studied through graph theory and the first efforts were devolved to the assessment of aromaticity [82]. A graph theory-based aromaticity index related to the number of Kekulé structures is defined [83]. The bond length, which is correlated strictly to aromaticity, was also predicted with no significant error compared with experimental data thanks to topological indices such as carbon atoms in the surroundings of the bond and the total number of hexagons [84]. Furthermore, the rationalization of PAHs through graph theory facilitates the enumeration of isomers and so the comparison of properties as the HOMO-LUMO gap of nanoflakes [85].
A topological descriptor that accounts for the delocalization of electrons in the structure can estimate the electronic affinity, the ionization potential, and the excitation energy with a level of accuracy of 0.1 eV [86]. This evaluation allows at first instance to avoid heavy calculations associated with DFT methods. However, if quantum-chemical calculations are performed, the descriptor can be extended also to thiophenes [86, 87]. Nevertheless, this approach requires the determination of four parameters that need to be assigned by hand or through an automatized system [88] to increasingly large PAH structures, considering each time their molecular graph.

2| Methods

2.1. Computational methods

In this Thesis, electronic properties of molecules are analyzed through computational methods, in particular DFT calculations. A brief introduction to quantum mechanics and molecular mechanics, considering the main approximations, is provided by Jensen [89]. Schrödinger time-independent equation describes stationary states of a quantum system.

$$H\ \Psi = E\ \Psi \qquad \text{where} H = \frac{\hat{p}^2}{2m} + V(\mathbf{r}) = -\frac{\hbar}{2m}\nabla^2 + V(\mathbf{r}) \tag{2.1}$$

The equation 2.1 is an eigenvalue equation, where the Hamiltonian operator H is the sum of the kinetic operator and potential operator V and is dependent on coordinates $\mathbf{r}$ and the momentum $\mathbf{p}$ of a single non-relativistic particle of mass m, and E is the total energy of the quantum particle described by the wave function Ψ. The energy of the particle E is the eigenvalue of the Hamiltonian operator H, and the wave function Ψ is the corresponding eigenfunction. In the context of molecules and atoms, there are different stationary states. We considered only the lowest one in energy, which is the ground state.

For a molecule with N electrons of mass m_i associated to a set of coordinates $\mathbf{r}$ and M nuclei of mass M_I, atomic mass Z_I associated to a set of coordinates $\mathbf{R}$, the kinetic energy operator T is the sum of all the kinetic terms of each particle.

$$\begin{aligned}
H &= T_{\mathrm{e}} + T_{\mathrm{n}} + V_{\mathrm{ee}} + V_{\mathrm{en}} + V_{\mathrm{nn}} \\
T_{\mathrm{e}} &= \sum_{i=1}^{N} -\frac{\hbar}{2m_i}\nabla_i^2 \\
T_{\mathrm{n}} &= \sum_{i=1}^{M} -\frac{\hbar}{2M_I}\nabla_I^2
\end{aligned}$$

The potential energy component V is the Coulomb interaction between each pair of charged entities, and these are nuclei-electron attraction V_{en}, electron-electron repulsion

V_{ee}, and nuclei-nuclei repulsion V_{nn}.

$$V_{\mathrm{ee}} = \frac{e^2}{4\pi\epsilon_0}\sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.3a}$$

$$V_{\mathrm{en}} = -\frac{e^2}{4\pi\epsilon_0}\sum_{i,I}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \tag{2.3b}$$

$$V_{\mathrm{nn}} = \frac{e^2}{4\pi\epsilon_0}\sum_{I<J}\frac{Z_I Z_J}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{2.3c}$$

$$\tag{2.3d}$$

2.1.1. The Born-Oppenheimer approximation

The Schrödinger equation for stationary states has an exact solution for hydrogenoid systems, however, it is unfeasible to solve analytically the Schrödinger equation for many-electron problems. Hence, a numerical solution is proposed, introducing some approximations. The Born-Oppenheimer approximation assumes that the nuclei are thousands of times heavier than electrons ($\mathrm{m_n}/\mathrm{m_e} \gg 2000$). As a result, the velocity of electrons is much higher compared to the velocity of the nuclei. Therefore, nuclei are considered fixed bodies for electrons. So, for more than two particles, it is possible to simplify the molecular model by separating nuclear and electronic motions. For a given quantum state described by the wave function $\Psi(\mathbf{r})$, we consider a nuclear wave function that describes the quantum state of nuclei and an electronic wave function expressed as a parametric function of the degree of freedom of the nuclei ($\mathbf{R}$).

$$\Psi(\mathbf{r}, \mathbf{R}) = \psi_{\mathrm{e}}(\mathbf{r}_i|\mathbf{R})\ \phi_{\mathrm{n}}(\mathbf{R})$$

So the total Hamiltonian is made of purely nuclear terms, which are the nuclei-nuclei interaction and the kinetic energy term of nuclei, and the electronic Hamiltonian. The electronic Hamiltonian describes the motion of electrons considering the wave function and the energy parametric to $\mathbf{R}$.

$$H_{\mathrm{e}}(\mathbf{R})\ \psi_{\mathrm{e}}(r_1, \ldots, r_N|\mathbf{R}) = \varepsilon_{\mathrm{e}}(\mathbf{R})\ \psi_{\mathrm{e}}(r_1, \ldots, r_N|\mathbf{R}) \tag{2.2a}$$

$$H_{\mathrm{e}} = T_{\mathrm{e}} + V_{\mathrm{ee}} + V_{\mathrm{en}} \tag{2.2b}$$

Solving this equation for the electronic wave function will produce the electronic energy that together with V_{nn} gives the effective nuclear Hamiltonian.

$$H_{\mathrm{n}} = T_{\mathrm{n}} + V_{\mathrm{nn}} + \varepsilon_{\mathrm{e}}$$

This Hamiltonian is used in the Schrödinger equation for nuclear motion, which provides the vibrational, rotational, and translational modes of nuclei.

2.1.2. Mean field approximation

The expression of the electronic Hamiltonian in equation 2.2a still contains the electron-electron repulsion term that is not written as a single-body operator as we see in equation 2.3a. So, we introduce another approximation, which is the mean field approximation. Each electron i interacts with the mean field given by the interaction of all the other electrons i,j.

$$V_{\mathrm{ee}} = \sum_{i<j} v_{\mathrm{ee}}(i,j) \rightarrow \sum_{i} v_{\mathrm{eff}}(i)$$

The Hartree-Fock approach is an example of mean field approximation [90].

2.1.3. Density functional theory

The Hohenberg-Kohn theorems [50] were the foundation of the Kohn-Sham approach [91], which helped to approximate the solution of the electronic many-body problem more effectively. The first Hohenberg-Kohn theorem states that there is a unique functional of the density that describes the external potential and the second Hohenberg-Kohn theorem proves that the functional of the ground state density determines the ground state energy of the system. The density functional is not univocally determined. In the Kohn-Sham approach, the approximated density functional writes:

$$\rho(\mathbf{r}) = \sum_{i=0}^{N} |\phi_i(\mathbf{r})|^2$$

so it is expressed in terms of Kohn-Sham orbitals. The new expression of the Hamiltonian is

$$\left(-\frac{\hbar}{2m}\nabla^2 + v_{\mathrm{eff}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i \ \phi_i(\mathbf{r})$$
$$v_{\mathrm{eff}}(\mathbf{r}) = v(\mathbf{r}) + \left[\int \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}\, dx\right] + v_{\mathrm{XC}}(\mathbf{r})$$
$$v_{\mathrm{XC}}(\mathbf{r}) = \frac{\partial E_{\mathrm{XC}}[\rho]}{\partial \rho}$$

where the effective potential contains all the electronic terms in terms of electron density except the exchange-correlation functional. This energy term does not have a classical equivalent, since it arises from the antisymmetry of the quantum wave function and the correlation between electrons. Different functionals approximate this expression in terms of the integral of spin density and their gradients. One of them is the hybrid B3LYP functional [92, 93]. Basis sets are sets of functions used to describe the wave function of electrons in molecules or atoms. Basis sets are larger and more accurate depending on the number of primitive Gaussians used to approximate the core and valence orbitals. Polarization functions account for the electron density distribution around the molecule, while diffuse functions describe the distribution of the wave function, especially in regions away from the atomic nuclei. The basis sets adopted are 6-31G(d,p), 6-311G(d,p), and def2TVZP, in order of increasing accuracy.

2.2. Notions of graph theory

PAHs are a class of aromatic compounds made of fused aromatic rings possibly hosting heteroatoms in the ring or carrying atoms as substituents. Benzenoid hydrocarbons (BHs) are defined as condensed polycyclic unsaturated fully conjugated hydrocarbons composed exclusively of six-membered rings [94]. Graph theory is a convenient tool for the examination of BHs. We start introducing some mathematical definitions [94].

Definition 2.1. A graph $G = G(V, E)$ is a couple of sets such that $V = \{v_i\}$ is a set of vertices or nodes, and $E = \{\{v_i, v_j\}\}$ is a set of pairs of vertices called edges.

Two vertices connected by an edge are said to be adjacent to each other. Conversely, two non-adjacent vertices are connected by a sequence of edges, forming a path denoted as P. When an edge connects a vertex, the vertex is defined as covered. The degree of a node is the number of edges connected to a node. A graph can be classified depending on the connectivity between edges and nodes.

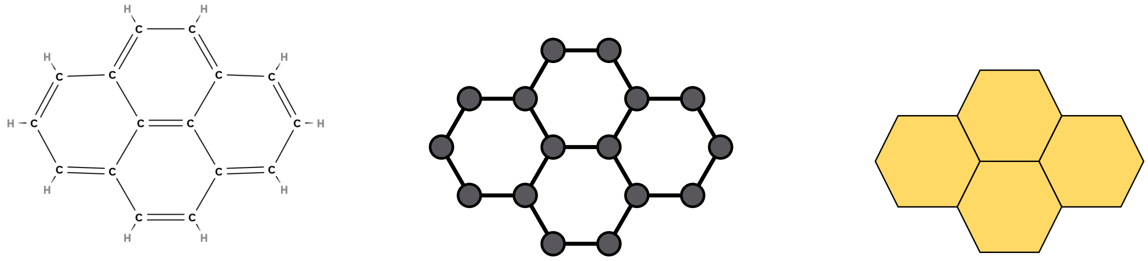
- A graph is simple if there are no edges starting and ending in the same node and with a single edge between a pair of connected vertices.
- A graph is connected if there is a path between whatever vertices v_1 and v_n.
- A graph is planar if edges intercept only at nodes. The portion of plane enclosed by edge and vertices in a planar graph is defined as face F.

Graph enumeration is the problem of counting all the obtainable graphs given a set of constraints.

We will use the graph enumeration of polyhexes. However, polyhexes are plane figures constructed by joining together hexagons. So, they are polygons with a precise shape, different from graphs that are characterized only by the connectivity between their vertices. We can still define polyhexes in terms of graph theory. For a given polyhex, we can choose an equivalence class of graphs with sets of E and V that are the same up to re-enumeration and represent the connectivity of the polyhex. This chosen equivalence class encapsulates all possible graphical representations of graphs that can be transformed into the shape of the designated polyhex. Thus, we define polyhex as a graph because we build a one-to-one connection between some graphs that can be drawn as a particular polyhex and that particular polyhex. So, in graph theory, polyhexes are a class of simple, planar, connected graphs where subunits of edge and vertices forming regular hexagon faces are joined together. We call trihexes the polyhexes with three hexagonal rings, tetrahexes the ones with four hexagonal rings and so on. In this thesis, trihexes are indicated as $\mathbf{3.01}, \mathbf{3.02}, \ldots, \mathbf{3.}\boldsymbol{n}$. Similarly, we refer to tetrahexes are indicated as $\mathbf{4.01}, \mathbf{4.02}, \ldots, \mathbf{4.}\boldsymbol{n}$, and so on. Polyhexes have been enumerated and are tabulated [95].

A molecular graph represents a molecular structure in terms of graph theory, meaning each atom is a vertex and each bond is an edge. In a hydrogen-depleted molecular graph, hydrogen atoms and the related bonds are deleted from the original molecular structure. The molecular graph of BHs corresponds to a graph where C atoms are nodes and the C-C σ bonds are the edges. An example is shown in Figure 2.1.

We define polyhexes as the discrete mathematical equivalent of the hydrogen-depleted molecular graph of BHs to simplify the enumeration of all the BHs with a prescribed number of rings. From now on, we will use the terms 'graph', 'molecular graph', or 'polyhex' interchangeably to refer to both polyhexes and the hydrogen-depleted molecular graph of BHs. In BH graphs, the degree of a node is always two or three.



(a) Molecular formula of pyrene

(b) Molecular graph of pyrene

(c) Polyhex associated to pyrene

Figure 2.1: Equivalent representations of pyrene

We introduce here also the concept of adjacency for two hexagons. Hexagons are adjacent if they have two common vertices, so they share an edge that we call internal edge. In BH graphs, we define the perimeter as the shortest closed path connecting nodes without considering the internal edges. The edges composing the perimeter are referred to as external edges. Since the representation of all the polyhexes of a given number of rings can be expensive, dual graphs are now introduced.

Definition 2.2. The dual graph $D(r, u)$ of a BH is a graph where a node u is placed at the center of the face F of each hexagon and an edge r connects the node u to the node t of an adjacent hexagon.

An example is shown in Figure 2.2.

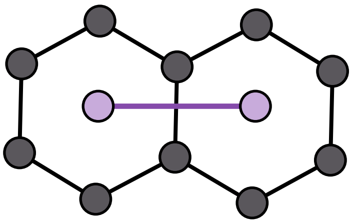


Figure 2.2: The molecular graph and the dual graph of naphthalene. Grey circles represent the vertices of the molecular graph, black segments the edges of the molecular graph, purple circles are the vertices of the dual graph and purple segments are the edges of the dual graph.

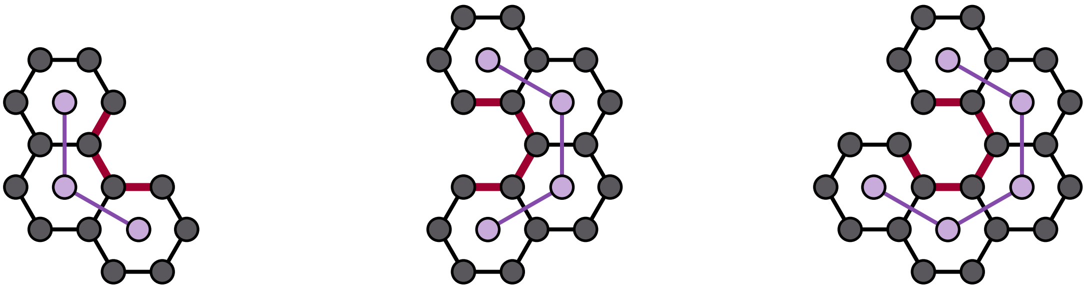
However, the dual graph does not identify univocally a BH [3].

Definition 2.3. The dualist graph is a dual graph where the angle between each pair of edges is conserved, so that the real connectivity of the molecule is respected.

There are particular perimeter configurations that can be highlighted through the dualist graph, represented in Figure 2.3

- The bay is a path of external edges whose corresponding dualist graph is a connection of three nodes with an angle of 120°.
- The cove is a path of external edges whose corresponding dualist graph is a connection of four nodes with two angles of 120°.
- The fjord is a path of external edges whose corresponding dualist graph is a connection of five nodes with three angles of 120°.

Coves and fjords may lead to distortion from planarity due to steric hindrance of hydrogens [96].



(a) Bay (b) Cove (c) Fjord

Figure 2.3: Particular perimeter configurations in BHs. In Panel **(a)**, **(b)**, **(c)**, the bay, cove, and fjord conformation are represented. In each panel, the thick red segments highlight the illustrated perimeter conformation, Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph, purple circles are the vertices of the dual graph and purple segments are the edges of the dual graph.

Furthermore, we distinguish between catacondensed and pericondensed BHs. Pericondensed BHs have at least one vertex of degree three. Catacondensed BHs have only vertices of degree two. An example is shown in Figure 2.4.

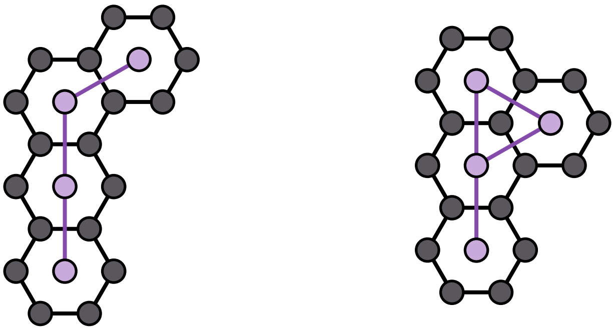


Figure 2.4: On the left, the molecular graph of a catacondensed BH associated to a tetrahex. On the right, the molecular graph of a pericondensed BH associated to a tetrahex.

Each BH is characterized by a particular pattern of connections between its hexagonal rings. It is worthwhile to introduce a notation to identify it. In the "LA" sequence, for each ring, the capital letter specifies the connection between that ring and the adjacent ones. Indeed, L and A specify the angle formed by the edges departing from the corresponding vertex of that ring in the dual graph. The number at the subscript specifies the number of rings adjacent to that one. For catacondensed BHs we give the following definitions:

- L_1 are extremal rings, connected only to another ring.
- L_2 are rings connected to two other rings. The edges connected to their dual graph form an angle of 180°.
- A_2 are rings connected to two other rings. The edges connected to their dual graph form an angle of 120°.
- A_3 are rings connected to three other rings. The edges connected to their dual graph form an angle of 60°.

An example is shown in Figure 2.5.

We distinguish branched and unbranched catacondensed BHs. Branched catacondensed BHs have at least one A_3 ring. Regarding unbranched catacondensed BHs, we will simplify the previous notation as L and A since the extremal rings are all L_1 and only L_2 and A_2 rings are possible. We will use the sequence of L and A to classify BHs, particularly the small ones [94]. All the branched and unbranched catacondensed BHs associated to n-hexes with $1 \leq n \leq 6$ are listed

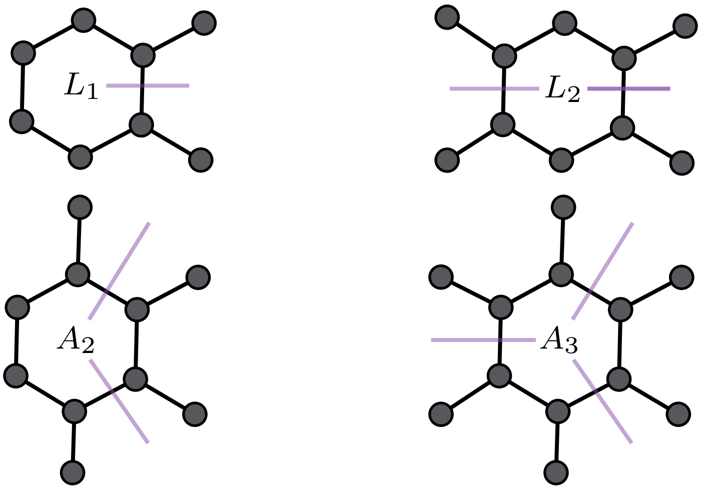


Figure 2.5: Particular types of ring and their notation. In the first row, from left to right, L_1 and L_2 types of ring and their adjacent rings. On the second row, from left to right, A_2 and A_3 rings and their adjacent rings. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph, and purple segments are the edges of the dual graph.

Additionally, matrices can also be employed to represent the connectivity of a graph. The adjacency matrix **A** of a graph is a square matrix where the rows and the columns of the matrix are the nodes of the graph and the entries of the table indicate the presence of an

edge connecting two vertices. For simple graphs, the entries in the adjacency matrix are 1 when there is an edge and 0 in its absence.
Matches (see below) are related to the connectivity of the graph.

Definition 2.4. A match is a set of mutually independent edges, so they are all non-adjacent and share no vertices. A perfect match is a set of matches where all the vertices are covered.

An example is shown in Figure 2.6. Up to now, we have analyzed the connectivity of a graph, but it is feasible to identify properties that are not related to the abstract representation of the graph.

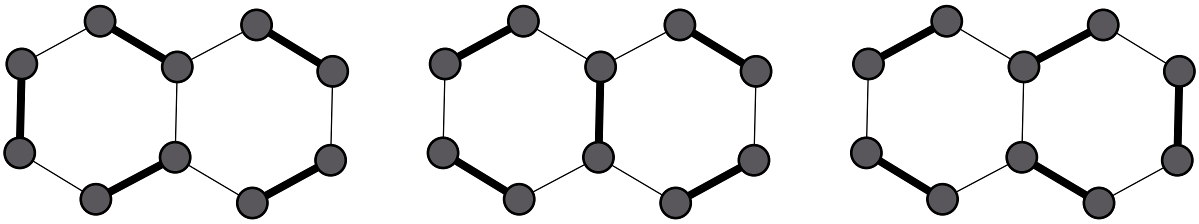


Figure 2.6: The three perfect matches of the only possible dihex.

Definition 2.5. A graph invariant is a property independent on the structure of the graph itself.

The β index is the ratio between the number of edges over the number of vertices of a graph and is an example of graph invariant.
Coming back to the molecular graph of BHs, it is possible to introduce chemical definitions that have also a meaning in graph theory. We investigate Kekulé structures. A Kekulé structure is a simplified representation of the distribution of π electrons in a π-conjugated molecule, highlighting the alternation between single and double bonds. More precisely, it is one of the possible Lewis structures that contribute to the resonance effect in the valence bond theory. An example is provided in Figure 2.7.
The perfect matching is the mathematical equivalent of Kekulé structures, that means that any Kekulé structure of a BH corresponds to a perfect matching of its molecular graph. Consequently, the number of Kekulé structures is the number of perfect matchings of its molecular graph [97]. It is also a graph invariant.

Not all BHs possess a Kekulé structure, for example, triangulene [98, 99]. A graph theory tool serviceable to determine whether the BH has a Kekulé structure is graph coloring.

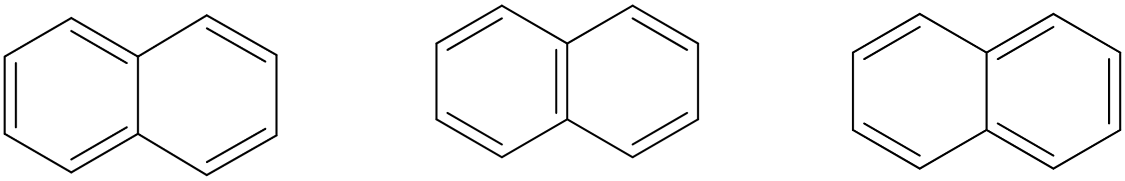


Figure 2.7: The three Kekulé structures of naphtalene.

The coloring of nodes consists of the different identification of adjacent vertices by assigning them different colors, for example, black and white. We define the excess of color Δ as the difference between the black and white vertices.

$$\Delta = |n_{\text{black}} - n_{\text{white}}|$$

It is demonstrated that molecular graphs with at least one Kekulé structures always have $\Delta = 0$. BHs graph with a color excess different from zero are non-Kekuléan, so they do not possess a Kekulé structures. These molecules are less stable and difficult to synthesize. However, the existence of triangulane and other diradicals was demonstrated [100, 101]. Non-Kekuléan molecules are tabulated [97].

2.2.1. Computation of the number of Kekulé structures

The number of Kekulé structures may be calculated with a pen-and-paper approach but for some classes it is computed with algorithms.

Catacondensed BHs. The number of Kekulé structures for these BHs is determined through the Davidson-Gordon algorithm [102], which is built upon three rules.

- Starting from the first hexagon with the number 2 (benzene has two Kekulé structures), we add one to the following hexagon until the first kink.
- On the hexagon immediately following the kink, we report the sum of the numbers of the last two hexagons.
- For any other linear hexagon, we continue adding the number of the hexagon before the latest kink.

  The last number is the number of Kekulé structures of that BHs.

An example is shown in Figure 2.8. For pericondensed BHs with a number of rings from one to six, the number of Kekulé structures is tabulated [97].

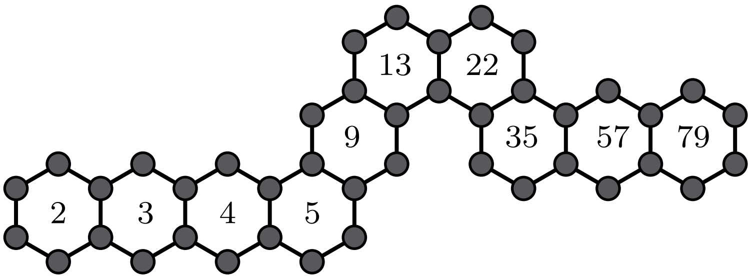


Figure 2.8: The molecular graph of an unbranched catacondensed BH whose number of Kekulé structure is calculated with the David-Gordon algorithm. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph. Starting from left to right, inside each ring, we label the number of Kekulé structures up to that ring. In the rightmost ring, we read the total number of Kekulé structures for the molecule. Reproduced from [102].

Some classes of polyhexes are defined by repetition along a direction in the plane of a specific sequence of hexagons. Here we list four of them to which molecules are associated. **Prolate rectangle**. Prolate rectangles $P(p, q)$ are polyhexes characterized by a profile of size p with zig-zag edges and the other of size q with armchair edges, as shown in Figure 2.9. They are associated to BHs named peri-Acenes [103, 104].

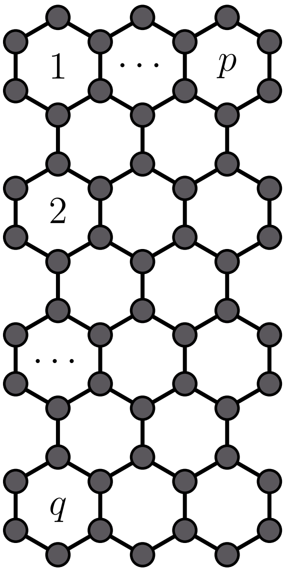


Figure 2.9: The molecular graph of a BH equivalent to a prolate rectangle $P(p, q)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph.

For prolate rectangles $P(p,q)$, the number of Kekulé structures is determined through the formula [97]:

$$K\{P(p,q)\} = (p \ + 1)^q \tag{2.4}$$

Parallelograms. Parallelograms $L(p,q)$ are polyhexes characterized by zig-zag edges. Their size is expressed in terms of p and q, where p and q are the number of rings on each side of the parallelogram, as we see in Figure 2.10. They are associated to BHs named peri-Acenoacenes [105–107].

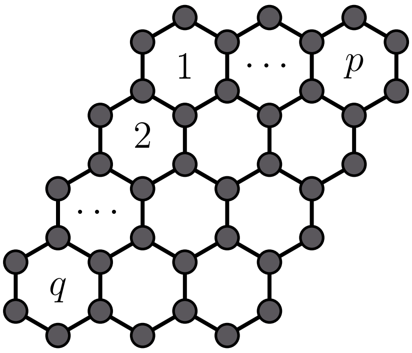


Figure 2.10: The molecular graph of a BH equivalent to a parallelogram $L(p,q)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph.

For parallelograms $L(p,q)$, the number of Kekulé structures is determined through the formula [97]:

$$K\{L(p,q)\} = \binom{p+q}{q} \tag{2.5}$$

Oblates. Oblate rectangles $O(p,q)$ are characterized by a profile of size p with zig-zag edges and the other of size q with armchair edges, as shown in Figure 2.11.

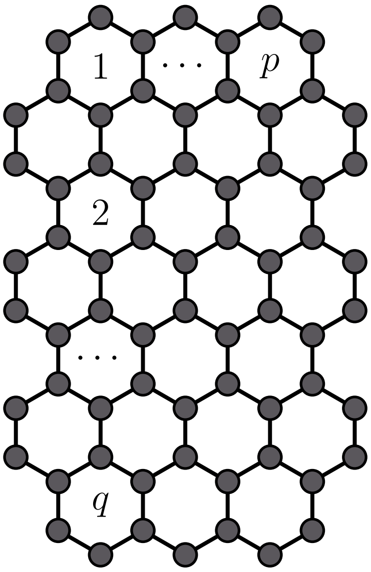


Figure 2.11: The molecular graph of a BH equivalent to an oblate rectangle $O(p,q)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph.

For oblate rectangles $O(p,q)$, the number of Kekulé structures is not determined in terms of an explicit expression in p and q [97].

$$
\begin{aligned}
K\{O(1,q)\} &= q+1 \\
K\{O(2,q)\} &= \frac{1}{12}(q+1)(q+2)^2(q+3) \\
K\{O(3,q)\} &= \frac{1}{120}(q+1)(q+2)^3(q+3)(q^2+4q+5) \\
K\{O(4,q)\} &= \frac{1}{20160}(q+1)(q+2)^4(q+3)(17q^4+136q^3+439q^2+668q+420) \\
K\{O(5,q)\} &= \frac{1}{362880}(q+1)(q+2)^5(q+3)(31q^6+372q^5+1942q^4+5616q^3+9511q^2 \\
&\quad +8988q+3780) \\
K\{O(6,q)\} &= \frac{1}{79833600}(q+1)(q+2)^6(q+3)(691q^8+11056q^7+79788q^6+338320q^5 \\
&\quad +921759q^4+1654264q^3+1915562q^2+1315560q+415800)
\end{aligned}
$$

Hexagons. Hexagons $H(p,q,r)$ are characterized by a profile of size p, q and r with zig-zag edges, as shown in Figure 2.12.

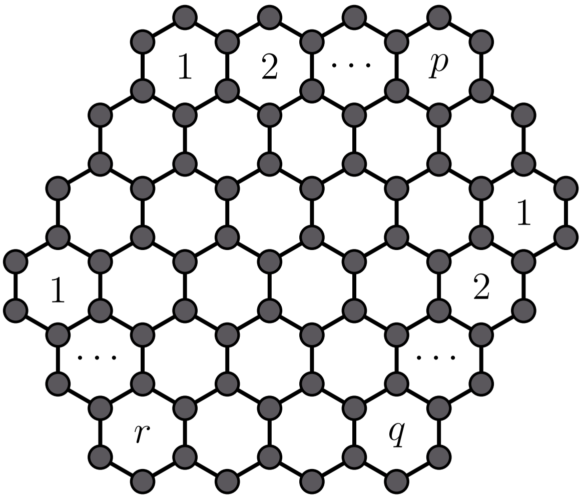


Figure 2.12: The molecular graph of a BH equivalent to a hexagon $H(p,q,r)$. Grey circles represent the vertices of the molecular graph, black segments are the edges of the molecular graph.

For hexagons $H(p,q,r)$, the number of Kekulé structures is determined through the formula [97]:

$$K\{H(p,q,r)\} = \prod_{i=0}^{q-1} \frac{\binom{p+r+i}{r}}{\binom{r+i}{r}} \tag{2.6}$$

2.3. Hückel theory

The Hückel molecular orbital theory, which from now on we will call Hückel theory, is another example of mean field approximation, firstly applied by Hückel on the small π-conjugated hydrocarbons [108–111] and then extended to larger molecules as polymers and carbon nanotubes [112, 113]. The main approximation introduced by Hückel is that the Hamiltonian related to the π molecular orbitals in conjugated molecules can be separated from the Hamiltonian related to the σ molecular orbital due to symmetry reasons. Indeed, we know that planar π-conjugated molecules have a horizontal plane of symmetry and that the π orbitals are odd with respect to reflection about that plane, while σ orbitals are even. This is equivalent to stating that the matrix elements of the Hamiltonian where there is the product of antisymmetric orbitals are null and that it is possible to simplify the Hückel matrix into two distinct blocks. These symmetry conditions are not valid for non-planar molecules, thus the extended Hückel theory is used [114]. The k^{th} π orbital is the linear combination of the N p_z atomic orbitals:

$$\pi^k = \sum_{\mu=1}^{N} c_\mu^k \, p_z^\mu$$

So, we rewrite the electronic Schrödinger equation into a matrix form:

$$\mathbf{H} \, \mathbf{c}_k = \epsilon_k \, \mathbf{S} \, \mathbf{c}_k$$

where:

$$H_{ij} = \int p_z^i \; H \; p_z^j \, d\tau$$
$$S_{ij} = \int p_z^i \; p_z^j \, d\tau$$

S is the overlap matrix. We now simplify this expression introducing two hypotheses:

- the p_z orbitals are orthonormal;

$$S_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

- the only interaction is between first neighbors, meaning i and j are bonded atoms, so all the other Hamiltonian terms are null.

$$H_{ii} = \int p_z^i \; H \; p_z^j \, d\tau = \alpha$$
$$H_{ij} = \int p_z^i \; H \; p_z^j \, d\tau = \beta$$

α is the parameter that describes the energy of an electron on a single orbital, meanwhile, β is the interaction term between two p_z orbitals. It is different from the β introduced in Section 2.2, that is related to graph theory. Considering this approximation, the Hückel matrix can be rewritten in terms of the adjacency matrix **A** of the graph of the molecule:

$$\mathbf{H} = \alpha \, \mathbf{I} + \beta \, \mathbf{A}$$

The molecular orbitals are determined by solving the following equation:

$$\det|(-\epsilon_k + \alpha) \, \mathbf{I} + \beta \, \mathbf{A}| = 0$$

so the HOMO-LUMO gap is the difference between the central eigenvalues of the adjacency matrix (without the hydrogen atoms) in units of β.

2.4. Computational tools

In this Thesis, molecular graphs were built through Matlab (version 2019b) and the molecules were visualized with Avogadro (an open-source molecular builder and visualization tool, Version 1.0.2 [115]) and Vesta (Version 3 [116]).

All quantum chemical calculations have been carried out with Gaussian09, a computational chemistry software initially released in 1970 by John Pople and his research group [117].

AConsidering DFT calculations, geometry optimization was performed through B3LYP/6-311G(d,p) [118–120] and all the molecules have reached their minimum. Considering previous literature, the basis set chosen for BHs associated to n-hexes with $1 \leq n \leq 6$ is 6-311G(d,p) [118–120]. The basis sets chosen for molecules associated to prolate rectangles are 6-311G(d,p) [121, 122], 6-31G(d,p) [123, 124], def2TVZP [125–127]. The HOMO-LUMO gap related to the analyzed molecules calculated with various basis sets exhibits minimal variation.

For simplicity, we opt to analyze all other molecules, including molecules associated to oblate rectangles, parallelograms, and hexagons, using the 6-311G(d,p) basis set.

The analyzed BHs are all the Kekuléan singlet-ground states associated to:

- n-hexes $1 \leq n \leq 6$;
- prolate rectangle $P(p, q)$ with $2 \leq p, q \leq 6$;
- oblate rectangle $O(p, q)$ with $1 \leq p, q \leq 6$;
- parallelograms $L(p, q)$ with $2 \leq p, q \leq 7$;
- hexagons $H(3, 3, r)$ with $2 \leq r \leq 7$, $H(4, q, r)$ with $3 \leq q \leq 4$ and $3 \leq r \leq 6$, $H(5, 5, r)$ with $5 \leq r \leq 6$.

3| Energy gap of BHs

3.1. Ciosłowski formula

In the ‘80s, Ciosłowski wrote a formula to describe the HOMO-LUMO gap χ_{HL} calculated with the Hückel theory of BHs from one to eight rings through topological descriptors [2]. We want to prove the validity of the formula also for the HOMO-LUMO gap calculated at the DFT level. Before examining the calculations, we recall here the Ciosłowski formula:

$$\chi_{HL} = 2\left(-\ 2.90611\ (2M/N)^{1/2} +\ 3.91744\ K^{2/N}\right) \tag{3.1}$$

The formula contains the number of carbon-carbon bonds, referred to as M, the number of carbon atoms, defined as N, and the number of Kekulé structures, named K. The first term containing $(2M/N)^{1/2}$ can be rewritten as $2\beta^{1/2}$. The β term allows the formula to distinguish between pericondensed BHs and catacondensed BHs since pericondensed BHs with the same number of hexagonal rings as catacondensed BHs are characterized by higher connectivity, given their higher M. Instead, all the unbranched catacondensed BHs with the same number of benzenoid rings are featured by the same M and N. The second term of the expression contains $K^{2/N}$, which is recognized as a stability index. The correlation coefficient of this formula for the dataset analyzed by Ciosłowski is $r = 0.927$ with a root mean square error of 0.06 [2].

Here we replicate the Ciosłowski work for a limited set of BHs molecules going from one to six rings. We show the fitting of the formula in equation 3.1 in Figure 3.1. We observe the correlation coefficient r is 0.9391, higher than the one reported in the original paper because the coefficients are optimized for all BHs up to 8 rings while our dataset includes BHs up to six rings.

We now apply a modified Ciosłowski formula, with the coefficients adjusted to predict the DFT-calculated energy gap of BHs associated to n-hexes with $1 \leq n \leq 6$. The coefficients are optimized using the least square curve fit.

$$\chi_{HL} = -\ 18.13518\ (2M/N)^{1/2} +\ 25.27383\ K^{2/N} \tag{3.2}$$

We observe the correlation coefficient is $r = 0.9540$. From Figure 3.2, the maximum relative error in absolute value is 0.25, where the error is the ratio of the difference

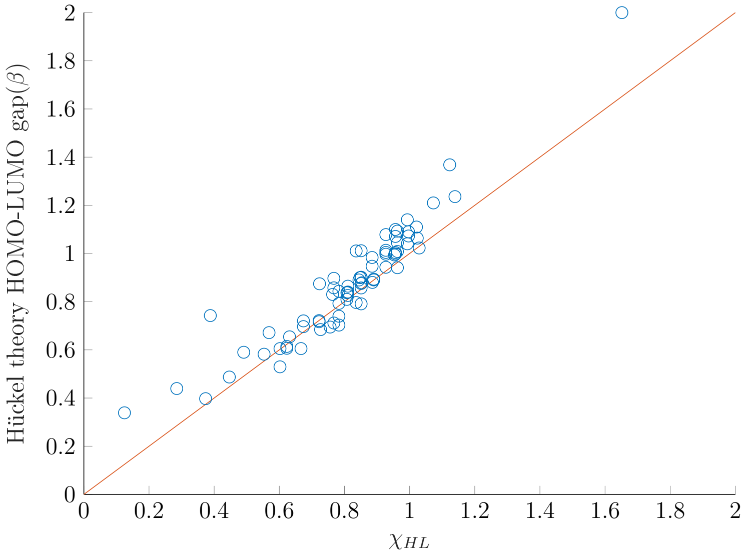


Figure 3.1: In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap estimated by the Ciosłowski formula χ_{HL} in equation 3.1 for BHs associated to n-hexes with $1 \leq n \leq 6$. The red line is the bisector of the first and third quadrants.

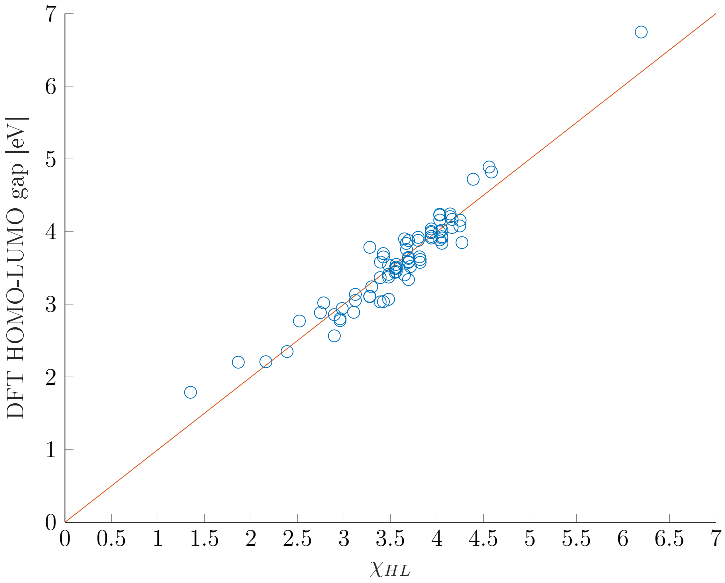


Figure 3.2: In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the DFT theory and the HOMO-LUMO gap estimated by the Ciosłowski formula χ_{HL} in equation 3.2 for BHs associated to n-hexes with $1 \leq n \leq 6$. The red line is the bisector of the first and third quadrants.

n-hexes with $1 \leq n \leq 6$		
χ_{HL}(eq 3.1)	Hückel theory HOMO-LUMO gap	$r = 0.9391$
χ_{HL}(eq 3.2)	DFT HOMO-LUMO gap	$r = 0.9540$

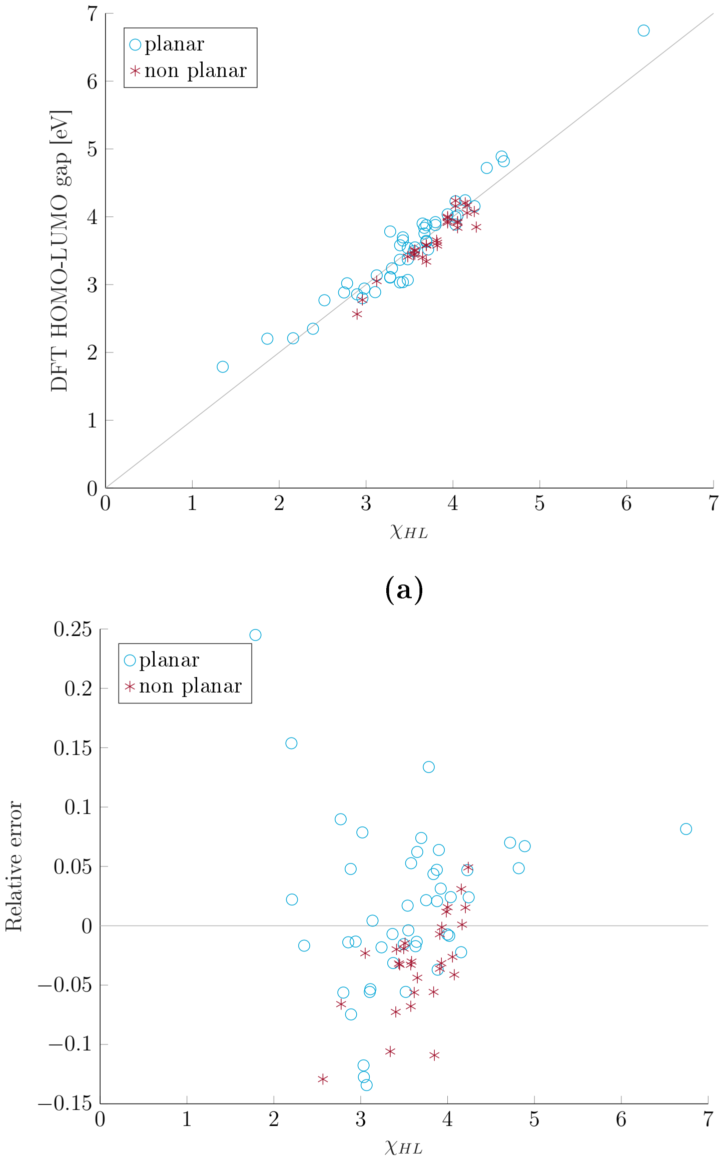
Table 3.1: Comparison between the correlation coefficients of the Ciosłowski formula χ_{HL} in equation 3.1 and HOMO-LUMO gap calculated with Hückel theory and the correlation coefficients of the modified formula χ_{HL} in equation 3.2 and HOMO-LUMO gap calculated with DFT theory.

between the DFT-calculated HOMO-LUMO gap and the estimated HOMO-LUMO gap χ_{HL} over the DFT-calculated HOMO-LUMO gap.

So, the modified Ciosłowski formula can predict the HOMO-LUMO gap of the considered set of BHs calculated at the DFT level.

In Table 3.1, we report the correlation coefficients of the computed HOMO-LUMO gap and the predicted HOMO-lUMO for molecules associated to n-hexes with a number of rings between one and six.

Effect of planarity. As mentioned in Section 2.2, coves and fjords result in non-planarity in BHs. We want to investigate whether the distortion from the planarity of the molecules leads to a sensible misfit of the formula.



(a)

(b)

Figure 3.3: In Panel **(a)**, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the DFT theory and the HOMO-LUMO gap estimated by the Ciosłowski formula χ_{HL} in equation 3.2 of planar (blue circles) and non planar (red stars) of BHs with a number of rings between one to six rings. The grey line is the bisector of the first and third quadrants. In Panel **(b)**, the relative error related to planar (blue circles) and non planar (red stars) BHs over the HOMO-LUMO gap estimated by equation 3.2.

In Figure 3.3, we observe that the maximum relative error in absolute value for the non-planar molecules is 0.15 and it is comparable to the maximum relative error in absolute value for planar molecules.

The HOMO-LUMO gap does not depend on the planarity of the BHs, but only on the topology of the molecule, namely the number of Kekulé structures.

Thus, the adjusted parameters of the Ciosłowski formula predict the electronic properties of BHs with a number of rings between one and six also considering the HOMO-LUMO gap estimated by DFT, which is more accurate than Hückel theory as it accounts for the actual geometry of the molecule and not just its graph.

Extending the Ciosłowski formula. We cease the enumeration of all the possible structures given a certain number of rings because the algorithm to determine all the n-hexes is nontrivial and their number increases exponentially. Indeed, for seven rings the number of polyhexes that is possible to enumerate is 333, but for nine or ten rings they become respectively 6572 and 30490 [95]. So, more than analyzing a big dataset of small molecules, we want to test the applicability of the Ciosłowski formula to molecules with more than ten rings. With that goal in mind, it is more feasible to consider classes of highly symmetric BHs rather than n-hexes. So, we choose classes of planar molecules that are easier to build by increasing their topological parameters, and for which a simple expression to evaluate the number of Kekulé structures is recognized, as underlined in Section 2.2.1.

We consider molecules corresponding to prolate rectangles, oblate rectangles, parallelograms, and hexagons.

Before going on with the calculation, we want to study the frontier orbitals of the so-chosen BHs. By observing the molecular orbitals in Figures 3.4, 3.5, 3.6, 3.7 we see that for compact molecules, characterized by a small p/q ratio, the HOMO and LUMO orbitals distribute themselves around the carbon atoms in all the molecules. For larger molecules featuring both longer armchairs and zig-zag profiles, the HOMO and LUMO orbital are more localized at the edge states of their zig-zag edges. Up to the molecule associated to the hexagon $H(5, 5, 5)$, the frontier orbitals calculated at the DFT level are not localized at the edges.

Looking at the frontier orbitals, we hypothesize that the HOMO-LUMO gap can be predicted by parameters related to the topology of the entire molecules, such as β and $K^{2/N}$, for BHs of limited dimension. So, we extend the Ciosłowski formula to estimate the DFT-calculated HOMO-LUMO gap of the BHs of the mentioned classes with a number of rings ≥ 8 but we limit ourselves to molecules with ≤ 20 rings.

The modified expression of the Ciosłowski formula, with coefficients optimized to predict the DFT calculated HOMO-LUMO gap of the molecules associated to prolate rectangles,

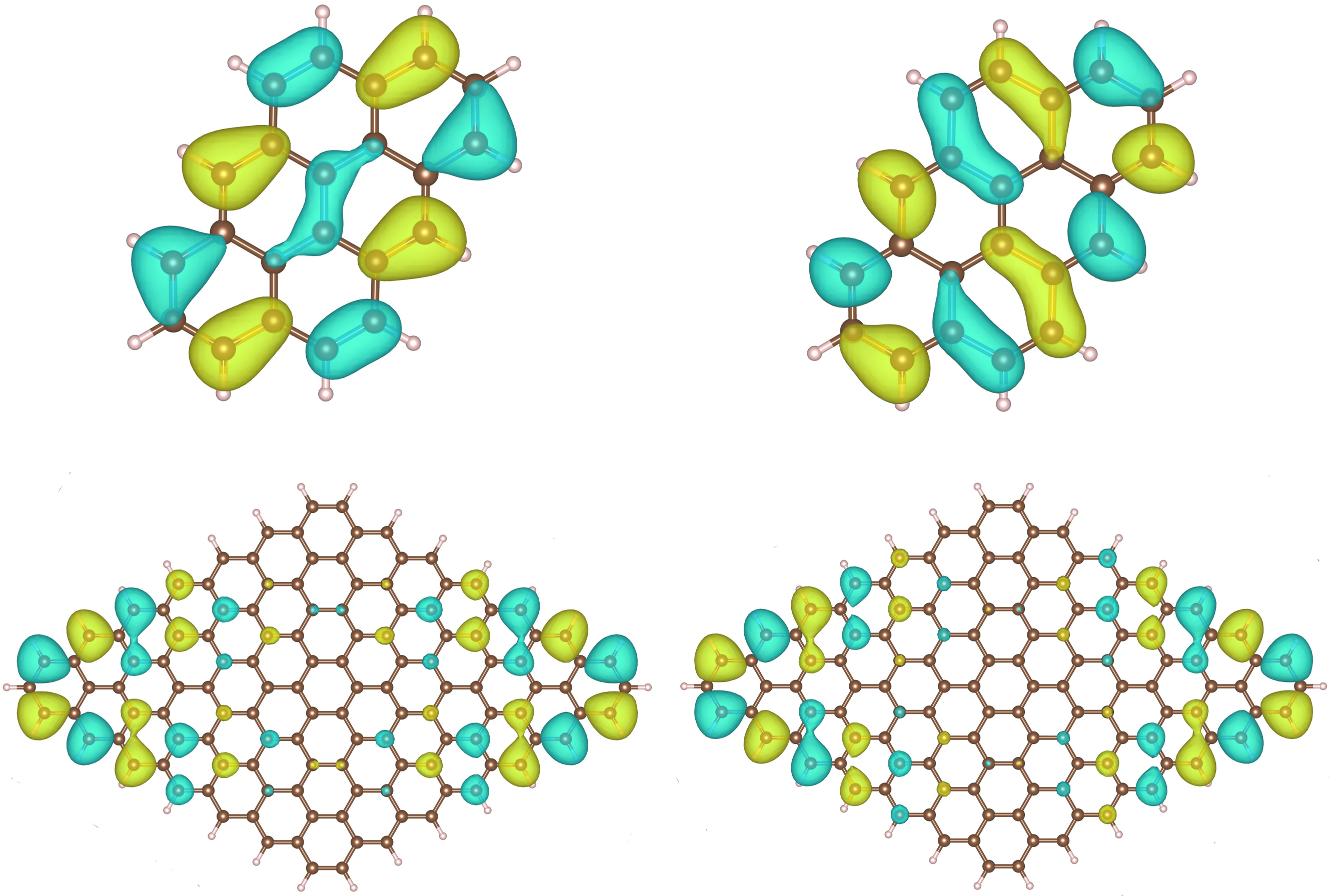


Figure 3.4: In the first row, isodensity surface plot at the value $0.0223a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent to parallelogram $L(2,3)$. In the second row, isodensity surface plot at the value $0.0116a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent to parallelogram $L(7,7)$.

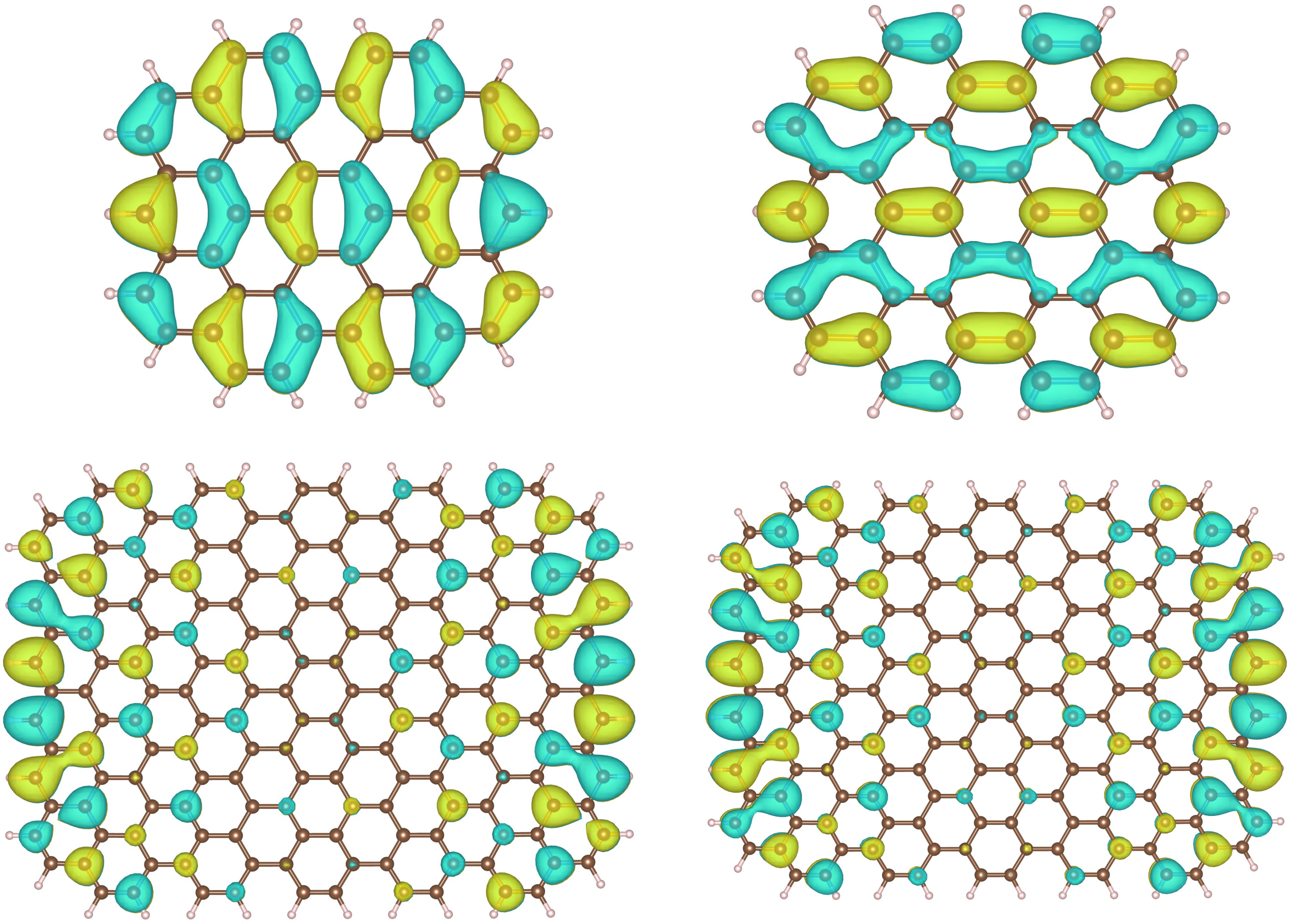


Figure 3.5: In the first row, isodensity surface plot at the value $0.0115a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent to oblate rectangles $O(3,2)$. In the second row, isodensity surface plot at the value $0.0179a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent oblate rectangles $O(6,5)$.

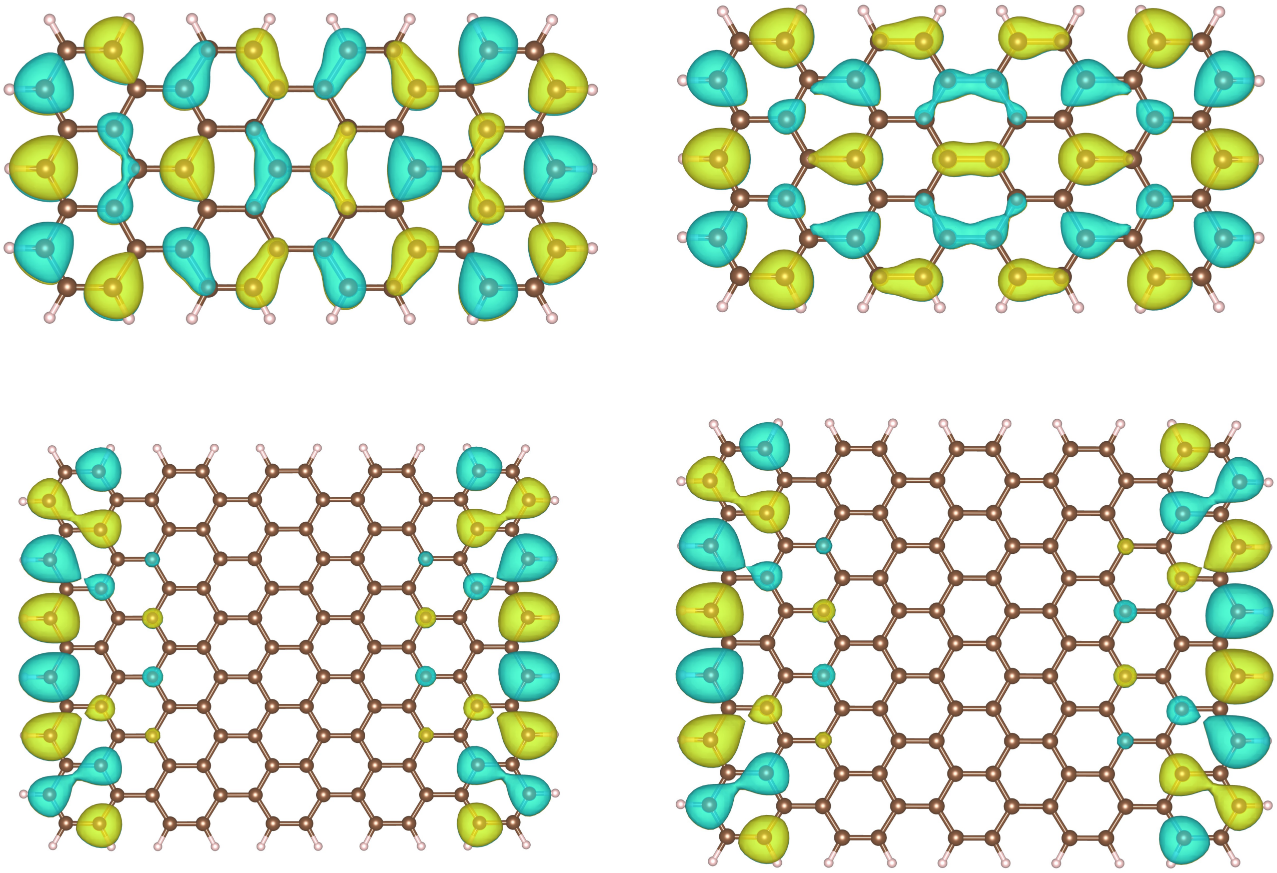


Figure 3.6: In the first row, isodensity surface plot at the value $0.0177 a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent to prolate rectangles $P(3,3)$. In the second row, isodensity surface plot at the value $0.0129 a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent to prolate rectangles $P(5,5)$.

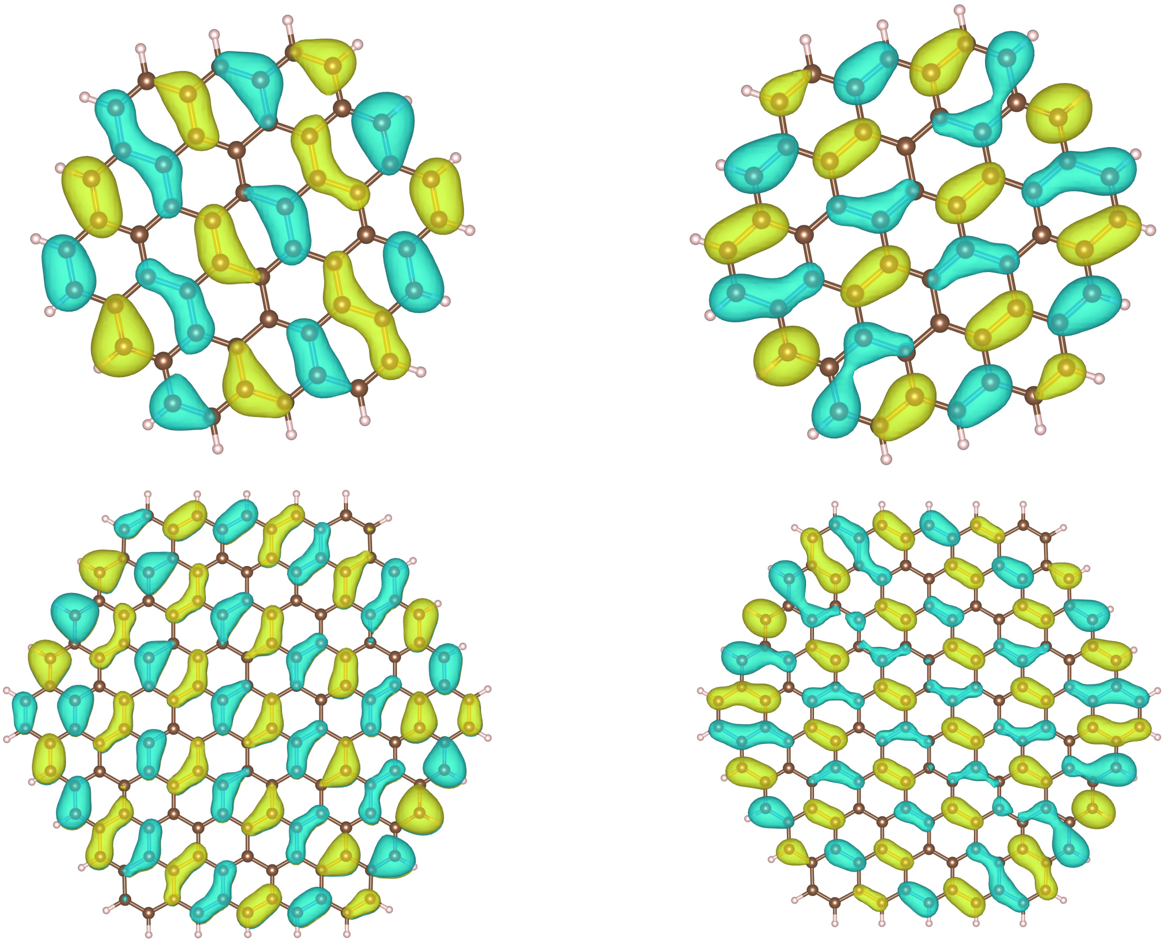


Figure 3.7: In the first row, isodensity surface plot at the value $0.0164a_0^{-3/2}$ (a_0:Bohr radius) of the HOMO on the left and of the LUMO on the right of the molecule equivalent of hexagon $H(2,3,3)$. In the second row, isodensity surface plot at the value $0.0112a_0^{-3/2}$ of the HOMO on the left and of the LUMO on the right of the molecule equivalent of hexagon $H(5,5,5)$.

oblate rectangles, parallelograms, and hexagons with less than 20 rings is:

$$\chi_{HL} = -\ 11.72999\ (2M/N)^{1/2} +\ 16.75529\ K^{2/N} \tag{3.3}$$

The correlation coefficient r is 0.8794. Furthermore, we notice that the stability term $K^{2/N}$ is not a valid topological indicator for molecules associated to prolate rectangles since if we express N in terms of p and q, we obtain:

$$N = q\ (2(p-2)+12) = 2\ (pq+4q) \tag{3.4}$$

Substituting in the equation 2.4, we get:

$$K^{2/N} = (p+1)^{\frac{2q}{2\ (pq+4q)}} = (p+1)^{\frac{1}{(p+4)}} \tag{3.5}$$

We will now assess the correlation of the newly modified χ_{HL} in equation 3.3 with the DFT calculated energy gap by examining its performance within each class, as depicted in Figure 3.8.

For molecules associated to prolate rectangles, in Figure 3.8b, the plotted relative error increases for molecules with smaller HOMO-LUMO gap, so associated to prolate rectangles with higher q. This was predicted by equation 3.5.

For molecules associated to parallelograms, in Figure 3.8a, we observe that χ_{HL} is consistently underestimated compared to the DFT data, with a relative error around 0.20, as shown in Figure 3.8b.

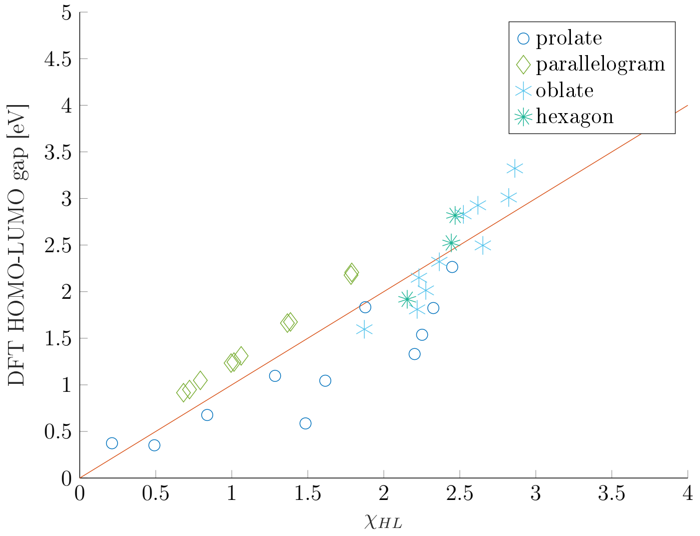
For molecules associated to oblate rectangles, in Figure 3.8b, the maximum relative error in absolute value is 0.25.

For molecules associated to hexagons, in Figure 3.8b, the maximum relative error in absolute value is 0.12.

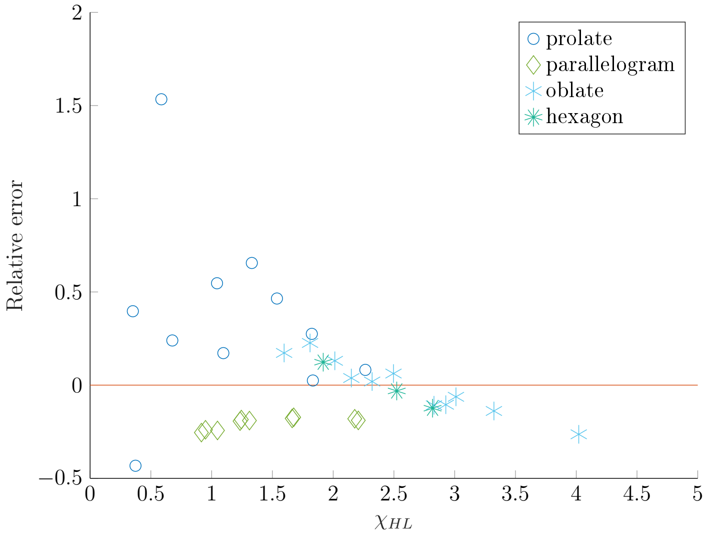
Thus, the parameters in the Ciosłowski formula can describe the HOMO-LUMO gap calculated by DFT for molecules with a limited number of rings, optimizing its coefficients. So, we write an ultimate modified version to account for molecules associated to n-hexes with $1 \leq n \leq 6$ and oblate rectangles, parallelograms, and hexagons with less than 20 rings.

$$\chi_{HL} = \ -16.87575\ (2M/N)^{1/2} +\ 23.67559\ K^{2/N} \tag{3.6}$$

The correlation coefficient r is 0.9531. However, as we have seen for molecules associated to prolate rectangles in equation 3.5, $K^{2/N}$ is not an effective index. As we have deducted from the images of the frontier orbitals in Figures 3.4, 3.5, 3.6, 3.7, the topological parameters in the Ciosłowski formula do not properly describe the energy gap for bigger molecules. So, new topological parameters should be analyzed.



(a)



(b)

Figure 3.8: In Panel **(a)**, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with DFT and the HOMO-LUMO gap estimated by the Ciosłowski formula χ_{HL} in equation 3.3 for BHs associated to parallelograms in green diamonds, hexagons in green asterisks, prolate rectangles in blue circles and oblate rectangles in blue asterisks. The red line is the bisector of the first and third quadrants. In Panel **(b)**, the relative error related to BHs associated to parallelograms in green diamonds, hexagons in green asterisks, prolate rectangles in blue circles and oblate rectangles in blue asterisks estimated by equation 3.3.

Limit for the graphene. We started from molecules with a number of rings between one and six. Then we extended our research to larger molecules. Thus, it appears natural to investigate whether the Ciosłowski formula is applicable at infinity. So graphene, which is an infinite two-dimensional layer, is our limit case.

We start analyzing the terms of the Ciosłowski formula. We begin from $(2M/N)^{1/2}$. The ratio between C-C bond and C carbon atoms M/N is constant in graphene. So we can study it from the graphene unit cell. In the graphene unit cell, the number of carbon atoms is 2 and the number of carbon-carbon bonds is 3. We can substitute the determined values in the expression:

$$\chi_{HL} = A\ (2M/N)^{1/2} +\ B\ K^{2/N} = 0 \tag{3.7}$$

So, the connectivity term becomes a constant and a general condition on the coefficients of the Ciosłowski formula can be derived.

$$K^{2/N} = -\frac{A}{B}\sqrt{3} \tag{3.8}$$

We now look at the stability term. K goes to infinity for graphene, as N, so we have an indeterminate form for $K^{2/N}$.

We study the case of the parallelogram $L(p,p)$. Looking at the molecule, we derive the number of carbon atoms N as a function of p.

$$N(L(p,q)) = 2(pq + p + q) \tag{3.9}$$

$$N(L(p,p)) = 2(p^2 + 2p) \tag{3.10}$$

Then, we evaluate the limit for molecules associated to a parallelogram $P(p,p)$ of the expression of the number of Kekulé structure as shown in equation 2.5, deriving:

$$\lim_{p\to\infty} K\{L(p,p)\} = \lim_{p\to\infty} \left[\frac{2p!}{p!p!}\right]^{\frac{1}{p^2+2p}} \tag{3.11}$$

Using Stirling's approximation[128], we get:

$$\lim_{p\to\infty} K\{L(p,p)\}^{2/N} = \lim_{p\to\infty} \left[\frac{\sqrt{2\pi 2p}\left(\frac{2p}{e}\right)^{2p}}{\sqrt{2\pi p}\left(\frac{p}{e}\right)^{p}\sqrt{2\pi p}\left(\frac{p}{e}\right)^{p}}\right]^{\frac{1}{p^2+2p}}$$

$$\begin{aligned}\lim_{p\to\infty} K\{L(p,p)\}^{2/N} &= \lim_{p\to\infty} \left[\frac{\sqrt{2}\left(\frac{2p}{e}\right)^{2p}}{\sqrt{2\pi p}\left(\frac{p}{e}\right)^{2p}}\right]^{\frac{1}{p^2+2p}} \\ &= \lim_{p\to\infty} \left[\frac{2^{2p}}{\sqrt{\pi p}}\right]^{\frac{1}{p^2+2p}}\end{aligned}$$

Exploiting the properties of limits, we get at the numerator

$$\lim_{p\to\infty} 2^{\frac{2p}{p^2+2p}} = \lim_{p\to\infty} 2^{\frac{2}{p+2}} = 1$$

While at the denominator

$$\lim_{p\to\infty} \pi^{\frac{1}{2(p^2+2p)}} p^{\frac{1}{2(p^2+2p)}} = \lim_{p\to\infty} p^{\frac{1}{2(p^2+2p)}}$$

Using the formula $\alpha^{\beta} = e^{\beta\ ln\alpha}$ and applying L'Hopitals rule [128], we get:

$$\lim_{p\to\infty} \frac{1}{e^{\frac{ln(p)}{2p^2+4p}}} = 1$$

The behaviour of the function is shown in Figure 3.9.

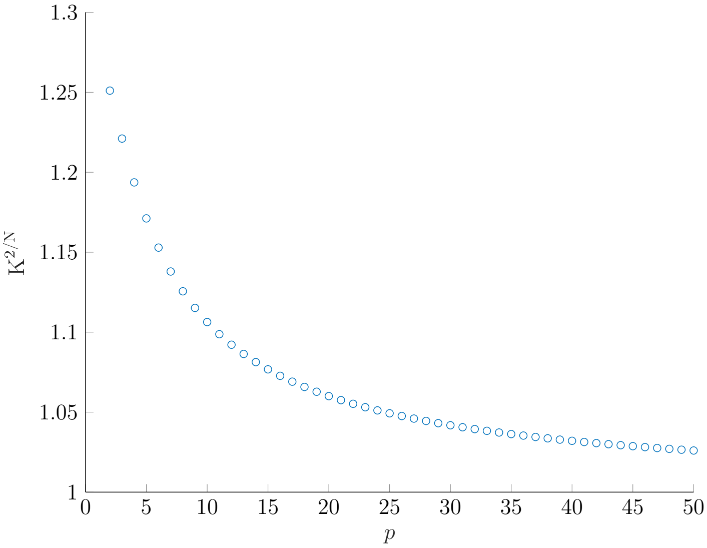


Figure 3.9: In blue circles, $K^{2/N}$ over p for parallelogram $L(p,p)$, with p going from 2 to 38.

As a result, the Ciosłowski formula predicts the behavior of the natural limit of BH and GNs that is the two-dimensional infinite layer of graphene. Indeed, extending molecules associated to parallelograms $L(p,p)$ to the infinite, we determine

$$K^{2/N} = -\frac{A}{B}\sqrt{3} = 1$$

So, in the end we get a relation between the two coefficient A and B that can take into account the zero gap of graphene.

$$\frac{A}{B} = -\frac{1}{\sqrt{3}}$$

However, to confirm this result, the derivation should be extended to other classes of BHs. But, as stated in Section 2.2.1, not all BHs are characterized by an expression of K in terms of p and q.

Other topological parameters. Given the limited applicability of the Ciosłowski formula for classes of BHs associated to polyhexes such as prolate rectangles, we look for new indices that are easier to calculate than the number of Kekulé structures but still have a sufficiently high correlation coefficient. The investigated coefficients are:

- the inverse of the perimeter P, namely the inverse of the number of external edges
- the inverse of the number of carbon atoms N
- the ratio between P and the number of internal edges M_{int}
- the ratio between P and N.

We can consider the $1/P$ and $1/N$ topological descriptor as strictly dependent on the dimension of the molecule, while P/M_{int} and P/N can be seen as a ratio between the perimeter of the molecule and its area, so they look more like of a ratio between the peripherical region and the connectivity of the entire molecule.

	$1/P$	$1/N$	P/M_{int}	P/N
prolate rectangles	0.9210	0.8264	0.9720	0.8779
oblate rectangles	0.7881	0.7766	0.5972	0.7170
parallelograms	0.9951	0.9979	0.9544	0.9044
hexagons	0.7249	0.7491	0.6821	0.6628

Table 3.2: Correlation coefficients of the HOMO-LUMO gap calculated with DFT for all the analyzed molecules equivalent to prolate rectangles, oblate rectangles, parallelograms, and hexagons with the inverse of the perimeter, the inverse of the number of carbon atoms, the ratio between perimeter and the number of internal edges, and the ratio between the perimeter and the number of carbon atoms.

From Table 3.2, we observe that the P/M_{int} has a correlation coefficient smaller than 0.7 for molecules associated to oblate rectangles and hexagons. $1/P$ and $1/N$ instead are more reliable indices, since their value is close to 1 both for molecules associated to parallelograms and prolate rectangles and higher than the other topological parameters analyzed for the remaining two classes.

	$1/P$	$1/N$	P/M_{int}	P/N
prolate rectangles	0.9310	0.9246	0.8844	0.8283
oblate rectangles	0.7942	0.6936	0.6574	0.5510
parallelograms	0.9938	0.9964	0.9890	0.9673
hexagons	0.6426	0.7531	0.5005	0.4862

Table 3.3: Correlation coefficients of the HOMO-LUMO gap calculated with DFT for molecules with more than 20 rings equivalent to prolate rectangles, oblate rectangles, parallelograms, and hexagons with the inverse of the perimeter, the inverse of the number of carbon atoms, the ratio between perimeter and the number of internal edges, and the ratio between the perimeter and the number of carbon atoms.

As we observe in Table 3.3, for molecules associated to prolate or oblate rectangles, we see that $1/P$ performs the best as a topological parameter in terms of correlation coefficients. This is not true for molecules related to parallelograms and hexagons. In those cases, it is $1/N$ that gives the higher correlation coefficient especially for hexagons. For all the molecules, parameters involving the ratio between the perimeter and an indicator regarding the composition of the entire molecules, such as M_{int} or N, do not correlate as well as the others. So, we can infer that for molecules with more than 20 rings, the topological descriptors which strictly increase with the dimension of the molecule correlate with th DFT calculated HOMO-LUMO gap better than descriptors that express the ratio between the characteristics of the peripherical region over the connectivity of the entire molecule.

3.2. Comparison between Hückel theory and DFT

Analysis of bond lengths. Our goal is to investigate the correlation between the Hückel theory and DFT for various classes of BHs. The Hückel theory is an accurate approximation when the C-C bond lengths among first neighbors are nearly equal. That means the variation of C-C bond length in the molecule must be minimal, with limited differences between single and double bonds. So, we analyze the bond length alternation for molecules associated to parallelograms, prolate rectangles, oblate rectangles, and hexagon. We observe from Tables 3.4, 3.5, 3.6, 3.7 that the maximum deviation in the absolute value of the bond lengths from their average is 0.0741 and is reported for molecule associated to hexagon $H(4, 3, 2)$. We compare this deviation to the bond length alternation of the polyacetylene, which is known to have single and double bonds alternating in the structure and for which the Hückel theory does not effectively predict the energy levels. In that case, the bond length alternation is 0.0600 Å [129]. So, the Hückel theory may not accurately predict the HOMO-LUMO gap of the analyzed molecules, since the bond length alternation is not negligible. However, for simplicity, we do not adoperate the extended Hückel theory but we expect the Hückel theory to embed this error. We continue by analyzing the variation of the average bond lengths within each class.

Parallelograms		
(p,q)	Average bond lengths [Å]	Maximum difference from average [Å]
$(2,2)$	1.4087	0.0503
$(2,3)$	1.4115	0.0568
$(2,4)$	1.4129	0.0593
$(2,5)$	1.4140	0.0608
$(2,6)$	1.4146	0.0616
$(2,7)$	1.4151	0.0621
$(3,3)$	1.4136	0.0609
$(3,4)$	1.4149	0.0629
$(3,5)$	1.4157	0.0638
$(3,6)$	1.4163	0.0646
$(3,7)$	1.4167	0.0651
$(4,4)$	1.4159	0.0641
$(4,5)$	1.4167	0.0650
$(4,6)$	1.4171	0.0658
$(4,7)$	1.4175	0.0660
$(5,5)$	1.4173	0.0656
$(5,6)$	1.4178	0.0659
$(5,7)$	1.4181	0.0666
$(6,6)$	1.4182	0.0664
$(6,7)$	1.4185	0.0667

Table 3.4: In the second column, the average mean value of the bond length for each molecule associated to a parallelogram. In the third column, the maximum difference in absolute value between the average value of the bond lengths and the bond lengths themselves.

Oblate rectangles		
(p,q)	Average bond length [Å]	Maximum difference from average [Å]
$(1,2)$	1.4087	0.0503
$(1,3)$	1.4123	0.0539
$(1,4)$	1.4137	0.0547
$(1,5)$	1.4145	0.0551
$(1,6)$	1.4150	0.0555
$(2,2)$	1.4118	0.0425
$(2,3)$	1.4145	0.0477
$(2,4)$	1.4157	0.0512
$(2,5)$	1.4163	0.0519
$(2,6)$	1.4168	0.0525
$(3,2)$	1.4136	0.0518
$(3,3)$	1.4159	0.0564
$(3,4)$	1.4169	0.0576
$(3,5)$	1.4175	0.0581
$(3,6)$	1.4178	0.0584
$(4,2)$	1.4148	0.0559
$(4,3)$	1.4168	0.0597
$(4,4)$	1.4177	0.0608
$(4,5)$	1.4182	0.0613
$(4,6)$	1.4185	0.0613
$(5,2)$	1.4156	0.0580
$(5,3)$	1.4174	0.0608
$(5,4)$	1.4182	0.0614
$(5,5)$	1.4186	0.0616
$(5,6)$	1.4189	0.0615

Table 3.5: Continued on next page

Table 3.5: Continued from previous page

Oblate rectangles		
(p, q)	Average bond lengths [Å]	Maximum difference from average [Å]
$(6, 2)$	1.4161	0.0590
$(6, 3)$	1.4178	0.0611
$(6, 4)$	1.4185	0.0612
$(6, 5)$	1.4189	0.0613
$(6, 6)$	1.4192	0.0610

Table 3.5: In the second column, the average mean value of the bond length for each molecule associated to a oblate rectangle. In the third column, the maximum difference in absolute value between the average value of the bond length and the bond lengths themselves.

Prolate rectangles		
(p, q)	Average bond lengths [Å]	Maximum difference from average [Å]
$(2, 2)$	1.4113	0.0644
$(2, 3)$	1.4135	0.0573
$(2, 4)$	1.4146	0.0547
$(2, 5)$	1.4151	0.0537
$(2, 6)$	1.4154	0.0530
$(3, 2)$	1.4141	0.0572
$(3, 3)$	1.4159	0.0456
$(3, 4)$	1.4166	0.0419
$(3, 5)$	1.4170	0.0480
$(3, 6)$	1.4173	0.0507
$(4, 2)$	1.4154	0.0557
$(4, 3)$	1.4168	0.0428
$(4, 4)$	1.4175	0.0445
$(4, 5)$	1.4179	0.0479

Table 3.6: Continued on next page

Table 3.6: Continued from previous page

Prolate rectangles		
ID of the molecule	Average bond length [Å]	Maximum difference from average [Å]
$(4,6)$	1.4182	0.0486
$(5,2)$	1.4161	0.0554
$(5,3)$	1.4174	0.0438
$(5,4)$	1.4181	0.0424
$(5,5)$	1.4185	0.0457
$(5,6)$	1.4188	0.0468
$(6,2)$	1.4163	0.0524
$(6,3)$	1.4179	0.0445
$(6,4)$	1.4185	0.0421
$(6,5)$	1.4189	0.0461
$(6,6)$	1.4191	0.0484
$(7,2)$	1.4167	0.0534
$(7,3)$	1.4182	0.0445
$(7,4)$	1.4188	0.0424
$(7,5)$	1.4192	0.0473
$(7,6)$	1.4192	0.0596

Table 3.6: In the second column, the average mean value of the bond length for each molecule associated to a prolate rectangle. In the third column, the maximum difference in absolute value between the average value of the bond lengths and the bond lengths themselves.

Hexagons		
ID of the molecule	Average bond lengths [Å]	Maximum difference from average [Å]
$(3,3,2)$	1.4150	0.0587
$(3,3,3)$	1.4161	0.0562
$(3,3,4)$	1.4173	0.0620

Table 3.7: Continued on next page

Table 3.7: Continued from previous page

Hexagons		
(p, q, r)	Average bond length [Å]	Maximum difference from average [Å]
$(3, 3, 5)$	1.4177	0.0629
$(4, 3, 2)$	1.4160	0.0617
$(4, 3, 3)$	1.4168	0.0601
$(4, 3, 4)$	1.4174	0.0631
$(4, 3, 5)$	1.4179	0.0644
$(4, 3, 6)$	1.4178	0.0741
$(4, 4, 2)$	1.4167	0.0637
$(4, 4, 3)$	1.4174	0.0631
$(4, 4, 4)$	1.4179	0.0621
$(4, 4, 5)$	1.4183	0.0640
$(4, 4, 6)$	1.4186	0.0648
$(5, 5, 5)$	1.4189	0.0651

Table 3.7: In the second column, the average mean value of the bond length for each molecule associated to a hexagon. In the third column, the maximum difference in absolute value between the average value of the bond lengths and the bond lengths themselves.

We now analyze the variation of the average bond lengths within each class. We observe for all the molecules analyzed within the class of parallelograms, oblate rectangles, prolate rectangles, and hexagons that the average bond length increases with p, q, and r parameters. Indeed, the average bond length increases going from the average carbon-carbon length of the benzene (1.400 Å) toward the average carbon-carbon bond length of graphene (1.425 Å) [130].

Notably, when we hold the parameter p constant, we observe that the maximum difference between the average bond lengths and individual bond lengths tends to increase with the parameter q for all molecules, except those linked to prolate rectangles. This deviation is attributed to the fact that molecules associated to prolate rectangles have fewer atoms in the core region compared to the other classes of molecules of analogous dimensionality. Consequently, for molecules associated to prolate rectangles, the average bond length shows less consistent growth, resulting in a decrease in the measured difference between the average bond length and the bond lengths observed in the peripheral region [130]. We now compare the bond length alternation in smaller and bigger molecules for each class considering a specific path to see whether the bond length alternation is localized in particular regions. In the examined molecules in Figures 3.10, 3.11, 3.12, the bond length alternation is more pronounced in the peripheral region than in the inner core. Consequently, the distortion of the molecule is more localized in the extremal regions compared to the inner core for both smaller and larger molecules, as observed in the past for selected large PAHs [131]. Shifting the focus to the outer paths of larger molecules, differences in bond length alternation are noted, particularly in prolate and oblate rectangles featuring both zig-zag and armchair edges. We observe a more evident increase in bond length alternation in the armchair region compared to the zig-zag region.

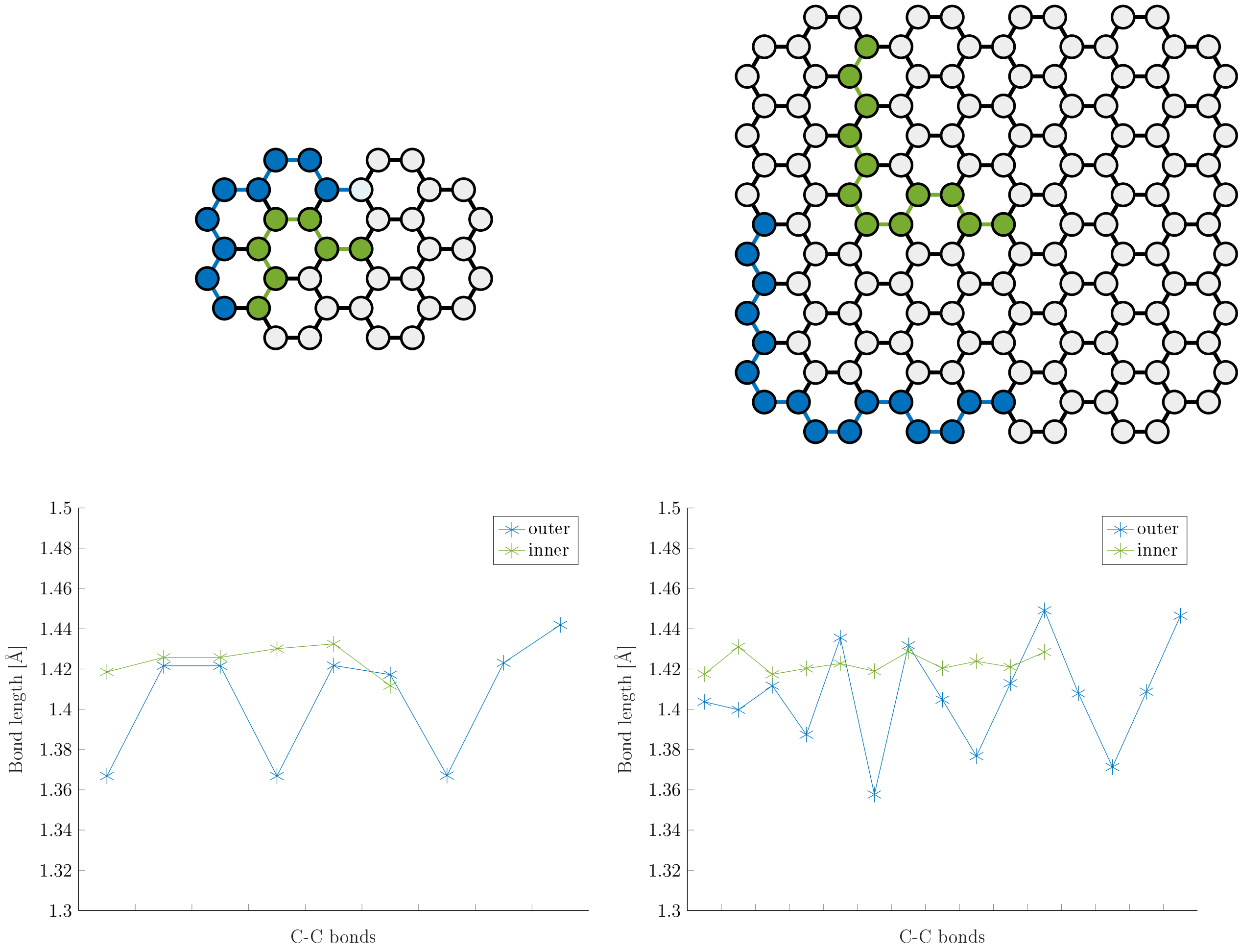


Figure 3.10: At the top, molecular graphs of BHs associated to oblate rectangle $O(2,2)$ on the left and oblate rectangle $O(6,4)$ on the right are displayed. The investigated outer path is highlighted in blue, while the inner path is marked in green. At the bottom, the C-C bond alternation plot along the outer path and the inner path of the molecule is illustrated.

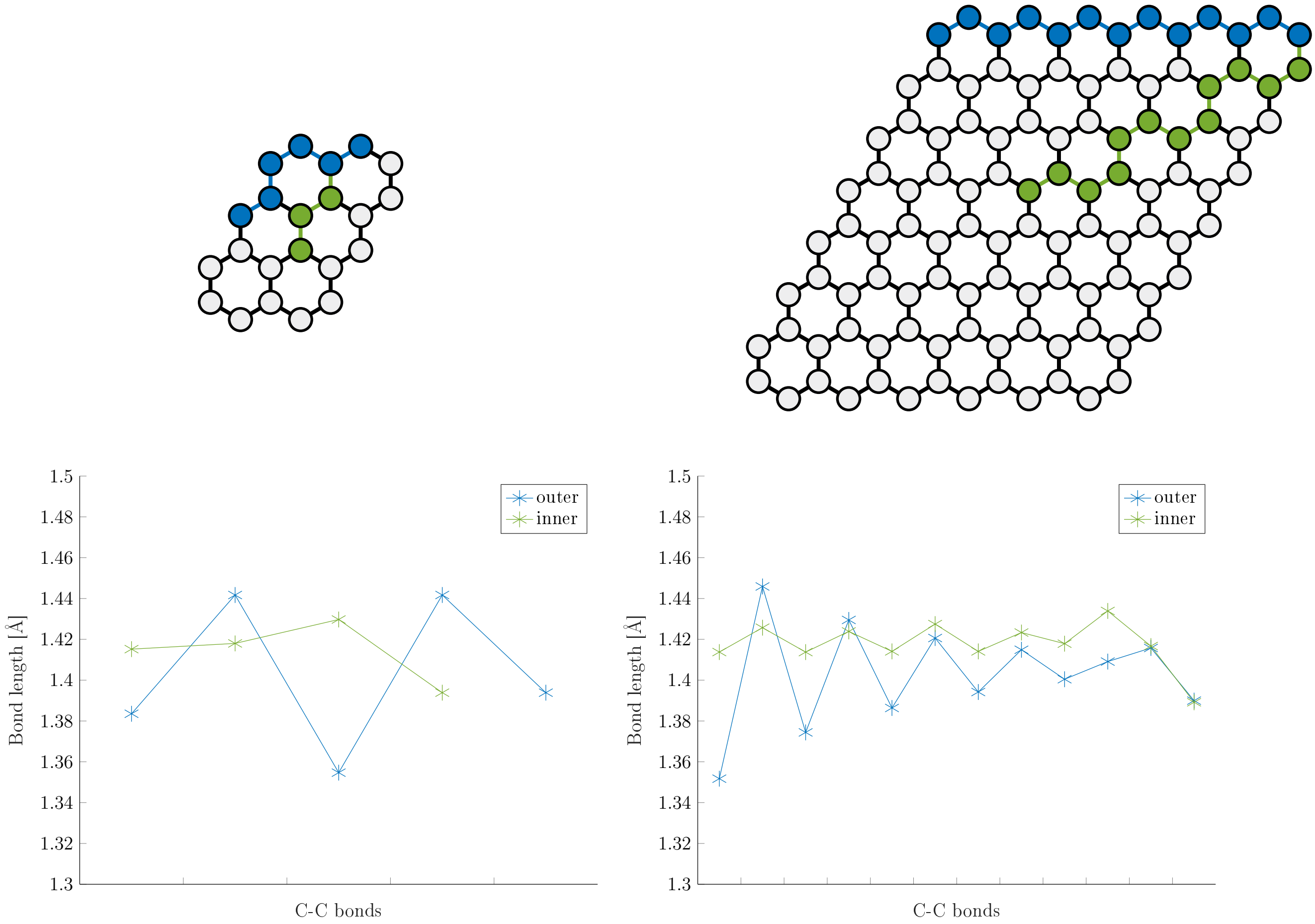


Figure 3.11: At the top, molecular graphs of BHs of BHs associated to parallelogram $L(2,3)$ on the left and parallelogram $L(6,7)$ on the right are displayed. The investigated outer path is highlighted in blue, while the inner path is marked in green. At the bottom, the C-C bond alternation plot along the outer path and the inner path of the molecule is illustrated.

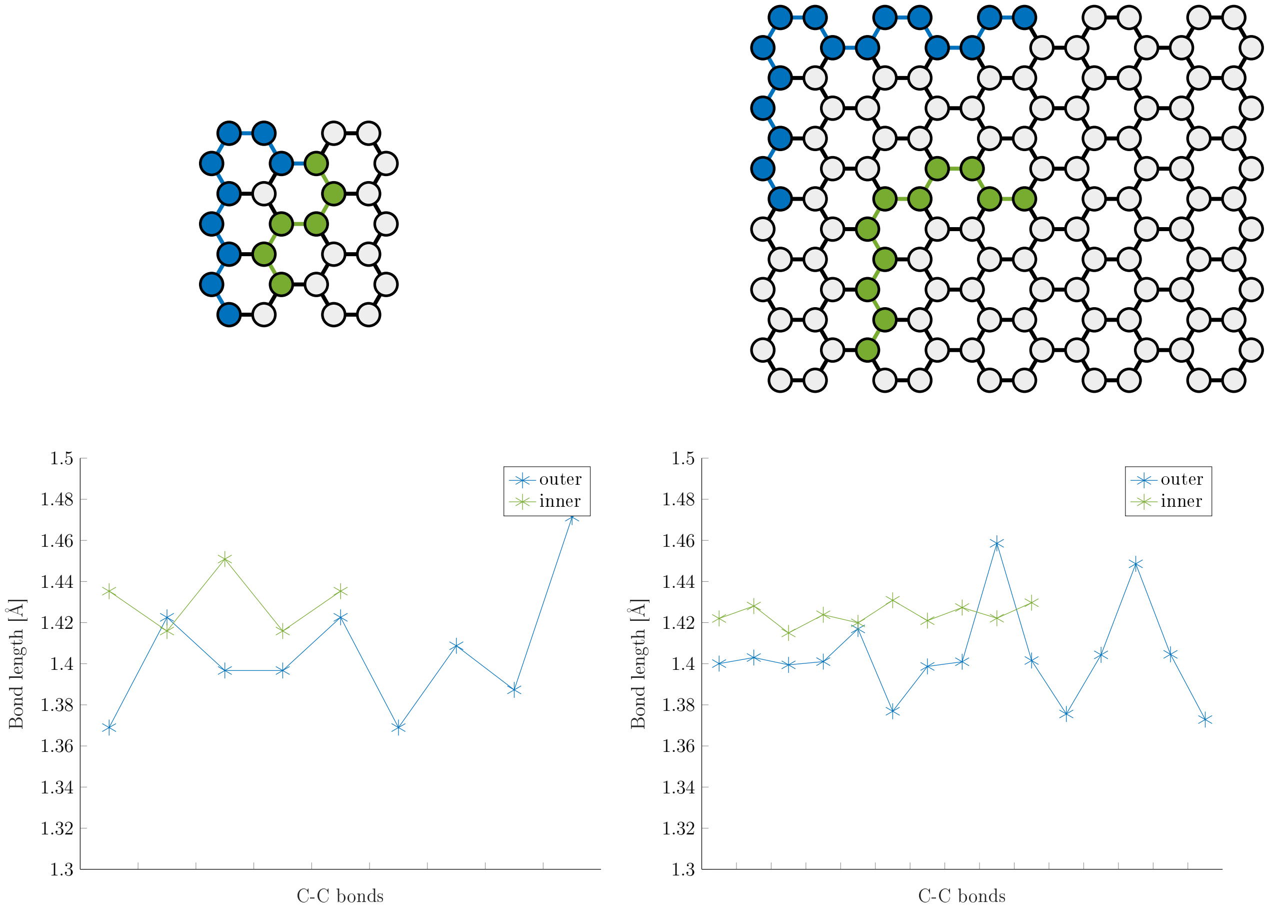


Figure 3.12: At the top, molecular graphs of BHs associated to prolate $P(3,2)$ on the left and parallelogram $P(6,4)$ on the right are displayed. The investigated outer path is highlighted in blue, while the inner path is marked in green. At the bottom, the C-C bond alternation plot along the outer path and the inner path of the molecule is illustrated.

For molecules related to parallelograms, characterized only by zig-zag edges we see the bond length alternation is still localized along specific edges.
A comparison of bond length alternation shown in Figure 3.10, 3.11, 3.12, with the isodensity surfaces in Figures 3.4, 3.5, 3.6 reveals that the localization of the HOMO and LUMO orbitals aligns with regions exhibiting smaller bond length alternation. This alignment is evident in the zig-zag profiles of molecules associated to oblate rectangles and prolate rectangles, as well as in the edges with less marked bond length alternation in molecules associated to parallelogram.
Correlation between the energy gap calculated with DFT and with Hückel theory. We want to investigate the presence of a linear correlation between the HOMO-LUMO gap calculated with the DFT method and the energy gaps calculated with Hückel theory by investigating the presence of a linear correlation between the so-calculated

HOMO–LUMO gap for all the different classes.

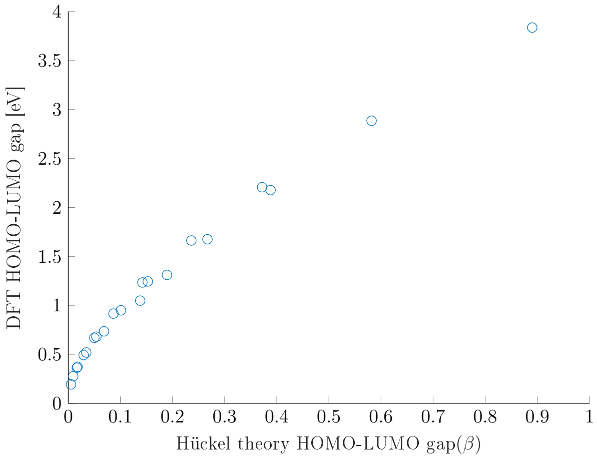


Figure 3.13: In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to parallelograms.

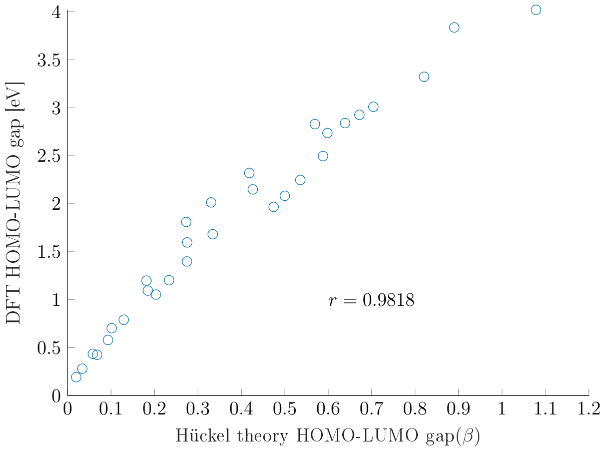


Figure 3.14: In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to oblate rectangles. The optimized linear fitting coefficient is 5.1093.

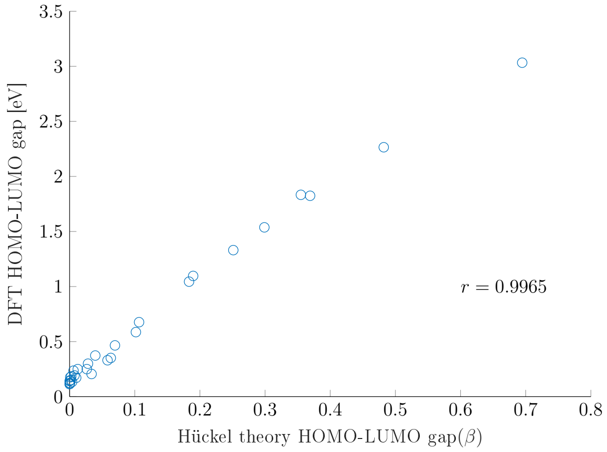


Figure 3.15: In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to prolate rectangles. The optimized linear fitting coefficient is 4.822.

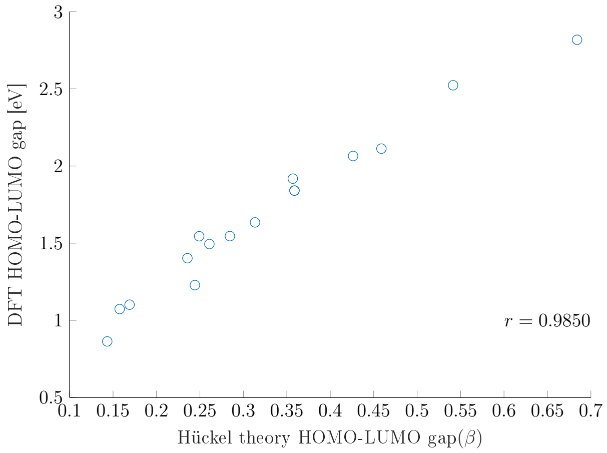


Figure 3.16: In blue circles, the scatter plot showing the relation between the HOMO-LUMO gap evaluated with the Hückel theory and the HOMO-LUMO gap calculated with DFT for molecules associated to hexagon. The optimized linear fitting coefficient is 4.903.

We observe in Figure 3.13 that for molecules associated to parallelogram the HOMO-LUMO gap calculated with Hückel theory is underestimated compared to the HOMO-LUMO gap calculated with DFT and the trend is not linear. This difference is due to the instability of the parallelogram associated with the exclusive presence of zig-zag edges, that lead to an open shell configuration, characterized by unpaired electrons [132]. For molecules associated to oblate rectangles, prolate rectangles, and hexagons, as we see in Figures 3.14, 3.15, 3.16 instead, the linear trend of the plot is more marked, so choosing the right β in the Hückel theory leads to an estimation of the HOMO-LUMO gap close to DFT results. From the data collected in the previous Section, we also associate the deviation from the linear correlation of the HOMO-LUMO gap calculated with Hückel theory and the HOMO-LUMO gap calculated with DFT to the bond length alternation, which is not negligible.

4| Isoarithmicity

Unbranched catacondensed BHs characterized by the same "LA" sequence, as discussed in Section 2.2, are referred to as isoarithmic, as we see in Figure 4.1 [3].
We now look closer at unbranched catacondensed BHs and study their HOMO-LUMO gap calculated at the DFT level. Considering Hückel theory, the energy gap of each isoarithmic group of molecules is comparable [133]. We want to validate the accuracy of this assertion also for the HOMO-LUMO gap calculated at the DFT level.

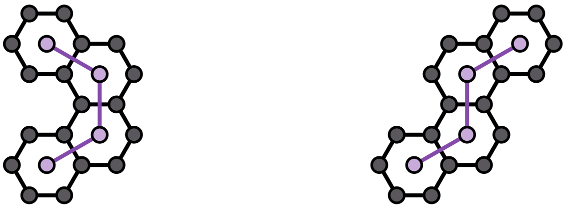


Figure 4.1: The molecular graphs of two isoarithmic molecules. The two unbranched catacondensed BHs are associated to the same dual graphs and the same "LA" sequences but their dualist graphs are different (see in Appendix A **4.03** and **4.04**). Grey circles represent the vertices of the molecular graph, black segments the edges of the molecular graph, purple circles are the vertices of the dual graph and purple segments are the edges of the dual graph.

Sequence	ID of the molecule	Hückel HOMO-LUMO gap (β units)	Mean value of the HOMO-LUMO gap (β units)	Difference from the mean value (β units)
$LLLL$	**4.01**	0.5899	0.5899	0.0000
$LALL$	**4.02**	0.9046	0.9046	0.0000
$LAAL$	**4.03**	1.1352	1.0878	0.0475
	4.04	1.0403		
$LLLLL$	**5.01**	0.4394	0.4394	0.0000
$LALLL$	**5.02**	0.6541	0.6541	0.0000

Table 4.1: Continued on next page

Table 4.1: Continued from previous page

Sequence	**ID of the molecule**	**Hückel HOMO-LUMO gap (β units)**	**Mean value of the HOMO-LUMO gap (β units)**	**Difference from the mean value (β units)**
LAALL	**5.04**	0.8096	0.8234	0.0138
	5.08	0.8372		
LLALL	**5.03**	0.8743	0.8743	0.0000
LALAL	**5.07**	0.9470	0.9653	0.0183
	5.05	0.9835		
LAAAL	**5.10**	1.0038	1.0518	0.0543
	5.09	1.0997		0.0416
	5.11	1.0709		0.0128
LLLLLL	**6.01**	0.3387	0.3387	0.0000
LALLLL	**6.02**	0.4872	0.4872	0.0000
LAALLL	**6.03**	0.6066	0.6104	0.0038
	6.05	0.6142		
LLALLL	**6.04**	0.6715	0.6715	0.0000
LLAALL	**6.10**	0.6953	0.7078	0.0125
	6.12	0.7203		
LALLAL	**6.07**	0.7168	0.7191	0.0023
	6.06	0.7213		
LAAALL	**6.14**	0.8258	0.8422	0.0164
	6.13	0.8387		0.0035
	6.16	0.8400		0.0022
	6.15	0.8643		0.0221
LAALAL	**6.18**	0.8571	0.8838	0.0267
	6.19	0.8755		0.0083
	6.20	0.9011		0.0174
	6.17	0.9013		0.0176
LALALL	**6.08**	0.8575	0.8772	0.0197
	6.11	0.8969		

Table 4.1: Continued on next page

Table 4.1: Continued from previous page

Sequence	ID of the molecule	Hückel HOMO-LUMO gap (β units)	Mean value of the HOMO-LUMO gap (β units)	Difference from the mean value (β units)
LAAAAL	**6.21**	0.9428	1.0072	0.0644
	6.24	0.9969		0.0103
	6.22	1.0130		0.0058
	6.09	1.0790		0.0718
	6.09	1.0790		0.0028

Table 4.1: HOMO-LUMO gap evaluated at the Hückel theory level of unbranched cata-condensed BHs of four, five, and six rings. We group them by "*LA*" sequences, reported in the first column. The fourth column contains the mean value of the HOMO-LUMO gap of each isoarithmic BH. The last column shows the difference in absolute value between the HOMO-LUMO gap of each group of isoarithmic BH.

Sequence	ID of the molecule	DFT HOMO-LUMO gap (eV)	Mean value of the HOMO-LUMO gap (eV)	Difference from the mean value (eV)
LLLL	**4.01**	2.7693	2.7693	0.0000
LALL	**4.02**	3.7530	3.7530	0.0000
LAAL	**4.03**	4.2052	4.2240	0.01878
	4.04	4.2428		
LLLLL	**5.01**	2.7693	2.7693	0.0000
LALLL	**5.02**	2.9413	2.9413	0.0000
LAALL	**5.08**	3.4392	3.4669	0.0276
	5.04	3.4945		
LLALL	**5.10**	3.7835	3.7835	0.0000
LALAL	**5.07**	3.8787	3.8998	0.0211
	5.05	3.9209		

Table 4.2: Continued on next page

Table 4.2: Continued from previous page

Sequence	ID of the molecule	DFT HOMO-LUMO gap (eV)	Mean value of the HOMO-LUMO gap (eV)	Difference from the mean value (eV)
LAAAL	**5.06**	4.1579	4.2078	0.0499
	5.10	4.2276		0.0198
	5.09	4.2379		0.0301
LLLLLL	**6.01**	1.7881	1.7881	0.0000
LALLLL	**6.02**	2.3492	2.3492	0.0000
LAALLL	**6.05**	2.7745	2.7873	0.0128
	6.03	2.7745		
LLALLL	**6.04**	3.0191	3.0191	0.0000
LLAALL	**6.12**	3.0518	3.0938	0.0420
	6.10	3.1358		
LALLAL	**6.07**	3.1043	3.1077	0.0034
	6.06	3.1111		
LAAALL	**6.16**	3.4488	3.5007	0.0331
	6.13	3.4950		0.0239
	6.15	3.5105		0.0218
	6.14	3.5105		0.0352
LAALAL	**6.20**	3.5761	3.6093	0.0331
	6.17	3.5854		0.0239
	6.18	3.6311		0.0218
	6.19	3.6444		0.0352
LALALL	**6.08**	3.6507	3.6738	0.0231
	6.11	3.6969		0.0231

Table 4.2: Continued on next page

Table 4.2: Continued from previous page

Sequence	**ID of the molecule**	**DFT HOMO-LUMO gap (eV)**	**Mean value of the HOMO-LUMO gap (eV)**	**Difference from the mean value (eV)**
LAAAAL	**6.09**	3.9108	3.9734	0.0625
	6.24	3.9334		0.040
	6.23	3.9854		0.0120
	6.22	4.0012		0.0278
	6.21	4.0360		0.0626

Table 4.2: HOMO-LUMO gap evaluated at the DFT level of unbranched catacondensed BHs of four, five, and six rings. We group them by "LA" sequences, reported in the first column. The fourth column contains the mean value of the HOMO-LUMO gap of each isoarithmic BH. The last column shows the difference in absolute value between the HOMO-LUMO gap of each group of isoarithmic BH.

Results. Both in Tables 4.1 and 4.2, molecules with the smallest HOMO-LUMO gap are the ones characterized by the sequence $L \dots L$, with zig-zag edges. Molecules with the largest HOMO-LUMO gap are the ones characterized by the sequence $LA \dots AL$, characterized by armchair edges [1]. Between these two extreme values, the HOMO-LUMO gap does not simply increase depending on the number of A rings, but it also depends on the concatenation in every single chain.

In Table 4.1 we see that the maximum difference between the energy gap and the mean value of the corresponding isoarithmic group is $0.0711|\beta|$ for molecule **6.09**. In Table 4.2 we notice that the maximum deviation of the energy gap from the mean value of the corresponding isoarithmic group is 0.0626 eV for molecule **6.09**. So, generally, the mean value of each isoarithmic group correctly approximates the HOMO-LUMO gap of the single molecule both at the Hückel level and at the DFT level.

Discussion. Unbranched catacondensed BHs with the same chain sequence have almost the same HOMO-LUMO gap calculated at the DFT level. So, we can infer that the HOMO-LUMO gap can be described just by looking at the chain sequence of the molecules without considering the characteristics of its external edges. More precisely, fjords and coves which lead to non-planar structures do not affect the HOMO-LUMO gap calculated at the DFT level. Also, at the DFT level, the HOMO-LUMO gap of an unbranched catacondensed BH can be correctly described by considering a graph invariant that distinguishes between L and A rings and accounts for their connectivity. The number of Kekulé structures K serves the scope. Isoarithmic BHs have the same number of Kekulé structures [3]. Furthermore, from the Davidson-Gordon algorithm exposed in Section 2.2.1, the counting of the number of Kekulé structure changes in a polyhex depending on the position of the kinks along the chain. Thus, the number of Kekulé structures correctly embodies the concept of isoarithmicity and allows us to distinguish between different isoarithmic groups. Also for the HOMO-LUMO gap calculated at the DFT value, in the Ciosłowski formula, the $K^{2/N}$ stability index accounts for all these aspects.

5| Conclusion

The study of electronic properties through topological descriptors can lead to massive simplification in calculation. Focusing on the HOMO-LUMO gap of benzenoid hydrocarbons, we find in the first Chapter that the two topological parameters used by Ciosłowski to estimate the Hückel theory HOMO-LUMO gap of molecules associated to n-hexes with $1 \leq n \leq 8$ allow to estimate also DFT-calculated HOMO-LUMO gap of molecules associated to n-hexes with $1 \leq n \leq 6$.

An adjusted version of the formula correctly predicts the DFT HOMO-LUMO gap both for planar and non-planar molecules.

By looking at the molecular orbital of BHs associated to different types of polyhexes of representative dimension, we hypothesize that the already mentioned parameters can estimate the DFT HOMO-LUMO gap of molecules associated to polyhexes with a number of rings between eight and twenty rings (bondaries included). The coefficients of the Ciosłowski formula are adjusted to do so, and a new modified version of the Ciosłowski formula is derived.

The main limitations of the parameters considered by Ciosłowski is that the DFT calculated HOMO-LUMO gap of molecules associated to specific classes, namely prolate rectangles, and the DFT calculated HOMO-LUMO gap of bigger molecules are not well approximated. So, we look for other topological parameters. For molecules with more than twenty rings characterized both by armchairs and zig-zag edges, the topological parameter showing the highest correlation coefficient is the inverse of the number of carbon-carbon bonds in the peripherical region, $1/P$. For molecules associated to parallelograms and hexagons characterized only by zig-zag edges, the inverse of the number of carbon atoms $(1/N)$ works the best.

Studying the bond length alternation of the molecules, we find that the average bond length for molecules associated to polyhexes of all the classes grows with the increasing of the parameters p, q, r. Furthermore, the deviation from the average bond length is localized in correspondence of the zig-zag edges for molecules associated to prolate and oblate rectangles and in correspondence of specific edges for molecules associated to parallelograms. Moreover, in polyhexes the localization of the HOMO and LUMO orbital is in correspondence with the region showing smaller bond length alternation.

In Chapter 2, we observe that the HOMO-LUMO gap calculated with the DFT is almost the same for each subclass of isoarithmic molecules, namely unbranched catacondensed molecules characterized by the same ring concatenation. So, the HOMO-LUMO gap does not depend much on the planarity of the molecule or its external edge, but more on its topology.

Perspectives for future work

It would be interesting to estimate the HOMO-LUMO gap through a simpler parameter connected just to the topology of the molecule rather than a more complex parameter such as the number of Kekulé structures. The main limitation of this parameter is that combinatorial formulas are not extendable to a generic BH, while a topological parameter associated to the structure will do the job.

Exploring PAHs could provide insights into how substituents or the presence of non-hexagonal rings impact the count of Kekulé structures, and whether the Ciosłowski parameters remain effective in predicting the HOMO-LUMO gap. Additionally, examining other topological parameters in this modified context would be worthwhile.

Investigating the applicability of the Ciosłowski formula to highly non-planar geometries would be of interest to determine potential limitations in that direction. In a broader sense, new computational and data analysis tools, such as machine learning algorithms, will lead for sure to a more efficient application of chemical graph theory to the study of materials properties other than the HOMO-LUMO gap, since now it will be possible to increase the dataset of molecule analyzed and to use an approach which is not just pen-and-paper.

References

[1] Yosadara Ruiz-Morales. Homo- lumo gap as an index of molecular size and structure for polycyclic aromatic hydrocarbons (pahs) and asphaltenes: A theoretical study. i. *The Journal of Physical Chemistry A*, 106(46):11283–11308, 2002.

[2] J Cioslowski. A unified theory of the stability of benzenoid hydrocarbons. *International journal of quantum chemistry*, 31(4):581–590, 1987.

[3] Alexandru T Balaban. Applications of graph theory in chemistry. *Journal of chemical information and computer sciences*, 25(3):334–343, 1985.

[4] United Nations. Sustainable developtment goals, 2023.

[5] Gregor Schwartz, Benjamin C.-K. Tee, Jianguo Mei, Anthony L. Appleton, Do Hwan Kim, Huiliang Wang, and Zhenan Bao. Flexible polymer transistors with high pressure sensitivity for application in electronic skin and health monitoring. *Nature Communications*, 4(1), May 2013.

[6] Jiajun Song, Hong Liu, Zeyu Zhao, Peng Lin, and Feng Yan. Flexible organic transistors for biosensing: Devices and applications. *Advanced Materials*, July 2023.

[7] Rui-Peng Xu, Yan-Qing Li, and Jian-Xin Tang. Recent advances in flexible organic light-emitting diodes. *Journal of Materials Chemistry C*, 4(39):9116–9142, 2016.

[8] Kenjiro Fukuda, Kilho Yu, and Takao Someya. The future of flexible organic solar cells. *Advanced Energy Materials*, 10(25), May 2020.

[9] Darren J. Lipomi, Benjamin C.-K. Tee, Michael Vosgueritchian, and Zhenan Bao. Stretchable organic solar cells. *Advanced Materials*, 23(15):1771–1775, February 2011.

[10] Zhi Hong Chen, Rui Fang, Wei Li, and Jianguo Guan. Stretchable transparent conductors: from micro/macromechanics to applications. *Advanced Materials*, 31(35), June 2019.

[11] Yan Qian, Xinwen Zhang, Linghai Xie, Dianpeng Qi, Bevita K. Chandran, Xiaodong Chen, and Wei Huang. Stretchable organic semiconductor devices. *Advanced Materials*, 28(42):9243–9265, August 2016.

[12] Xi Fan, Wanyi Nie, Hsinhan Tsai, Naixiang Wang, Huihui Huang, Yajun Cheng, Rongjiang Wen, Liujia Ma, Feng Yan, and Yonggao Xia. PEDOT:PSS for flexible and stretchable electronics: Modifications, strategies, and applications. *Advanced Science*, 6(19), July 2019.

[13] Martin Kaltenbrunner, Matthew S. White, Eric D. Głowacki, Tsuyoshi Sekitani, Takao Someya, Niyazi Serdar Sariciftci, and Siegfried Bauer. Ultrathin and lightweight organic solar cells with high flexibility. *Nature Communications*, 3(1), April 2012.

[14] Xiaheng Huang, Yue Qu, Dejiu Fan, Jongchan Kim, and Stephen R. Forrest. Ultrathin, lightweight and flexible organic light-emitting devices with a high light outcoupling efficiency. *Organic Electronics*, 69:297–300, June 2019.

[15] C.D. Dimitrakopoulos and P.R.L. Malenfant. Organic thin film transistors for large area electronics. *Advanced Materials*, 14(2):99–117, January 2002.

[16] C. W. Tang. Two-layer organic photovoltaic cell. *Applied Physics Letters*, 48(2):183–185, January 1986.

[17] C. W. Tang and S. A. VanSlyke. Organic electroluminescent diodes. *Applied Physics Letters*, 51(12):913–915, September 1987.

[18] N. S. Sariciftci, L. Smilowitz, A. J. Heeger, and F. Wudl. Photoinduced electron transfer from a conducting polymer to buckminsterfullerene. *Science*, 258(5087):1474–1476, November 1992.

[19] G. Yu, J. Gao, J. C. Hummelen, F. Wudl, and A. J. Heeger. Polymer photovoltaic cells: Enhanced efficiencies via a network of internal donor-acceptor heterojunctions. *Science*, 270(5243):1789–1791, December 1995.

[20] C. J. Brabec, N. S. Sariciftci, and J. C. Hummelen. Plastic solar cells. *Advanced Functional Materials*, 11(1):15–26, February 2001.

[21] Lin X. Chen. Organic solar cells: Recent progress and challenges. *ACS Energy Letters*, 4(10):2537–2539, October 2019.

[22] A. Tsumura, H. Koezuka, and T. Ando. Macromolecular electronic device: Field-effect transistor with a polythiophene thin film. *Applied Physics Letters*, 49(18):1210–1212, November 1986.

[23] John A. Rogers, Zhenan Bao, Kirk Baldwin, Ananth Dodabalapur, Brian Crone, V. R. Raju, Valerie Kuck, Howard Katz, Karl Amundson, Jay Ewing, and Paul Drzaic. Paper-like electronic displays: Large-area rubber-stamped plastic sheets of electronics and microencapsulated electrophoretic inks. *Proceedings of the National Academy of Sciences*, 98(9):4835–4840, April 2001.

[24] H. E. Katz. Organic molecular solids as thin film transistor semiconductors. *Journal of Materials Chemistry*, 7(3):369–376, 1997.

[25] Sheng-Yi Yang, Yang-Kun Qu, Liang-Sheng Liao, Zuo-Quan Jiang, and Shuit-Tong Lee. Research progress of intramolecular π-stacked small molecules for device applications. *Advanced Materials*, 34(22):2104125, 2022.

[26] Henning Sirringhaus. 25th anniversary article: Organic field-effect transistors: The path beyond amorphous silicon. *Advanced Materials*, 26(9):1319–1335, January 2014.

[27] Francis Garnier. Organic-based electronics a la carte. *Accounts of chemical research*, 32(3):209–215, 1999.

[28] J. H. Burroughes, D. D. C. Bradley, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burns, and A. B. Holmes. Light-emitting diodes based on conjugated polymers. *Nature*, 347(6293):539–541, October 1990.

[29] Arno Kraft, Andrew C. Grimsdale, and Andrew B. Holmes. Electroluminescent conjugated polymers - seeing polymers in a new light. *Angewandte Chemie International Edition*, 37(4):402–428, March 1998.

[30] Ullrich Mitschke and Peter Bäuerle. The electroluminescence of organic materials. *Journal of Materials Chemistry*, 10(7):1471–1507, 2000.

[31] Kwon-Hyeon Kim and Jang-Joo Kim. Origin and control of orientation of phosphorescent and TADF dyes for high-efficiency OLEDs. *Advanced Materials*, 30(42), April 2018.

[32] Luyao Lu, Tianyue Zheng, Qinghe Wu, Alexander M Schneider, Donglin Zhao, and Luping Yu. Recent advances in bulk heterojunction polymer solar cells. *Chemical reviews*, 115(23):12666–12731, 2015.

[33] Dou Luo, Woongsik Jang, Dickson D Babu, Min Soo Kim, Dong Hwan Wang, and Aung Ko Ko Kyaw. Recent progress in organic solar cells based on non-fullerene acceptors: materials to devices. *Journal of Materials Chemistry A*, 10(7):3255–3295, 2022.

[34] Hongliang Chen, Weining Zhang, Mingliang Li, Gen He, and Xuefeng Guo. Interface engineering in organic field-effect transistors: principles, applications, and perspectives. *Chemical reviews*, 120(5):2879–2949, 2020.

[35] Jinouk Song, Hyeonwoo Lee, Eun Gyo Jeong, Kyung Cheol Choi, and Seunghyup Yoo. Organic light-emitting diodes: pushing toward the limits and beyond. *Advanced Materials*, 32(35):1907539, 2020.

[36] Martin Pfeiffer, Stephen R Forrest, Karl Leo, and Mark E Thompson. Electrophosphorescent p–i–n organic light-emitting devices for very-high-efficiency flat-panel displays. *Advanced Materials*, 14(22):1633–1636, 2002.

[37] Mats Fahlman, Simone Fabiano, Viktor Gueskine, Daniel Simon, Magnus Berggren, and Xavier Crispin. Interfaces in organic electronics. *Nature Reviews Materials*, 4(10):627–650, 2019.

[38] Yanwei Gu, Zijie Qiu, and Klaus Müllen. Nanographenes and graphene nanoribbons as multitalents of present and future materials science. *Journal of the American Chemical Society*, 144(26):11499–11524, 2022.

[39] Prince Ravat. Carbo [n] helicenes restricted to enantiomerize: an insight into the design process of configurationally stable functional chiral pahs. *Chemistry–A European Journal*, 27(12):3957–3967, 2021.

[40] Giuseppe Maria Paternò, Goudappagouda, Qiang Chen, Guglielmo Lanzani, Francesco Scotognella, and Akimitsu Narita. Large polycyclic aromatic hydrocarbons as graphene quantum dots: from synthesis to spectroscopy and photonics. *Advanced Optical Materials*, 9(23):2100508, 2021.

[41] Qingbin Li, Yihan Zhang, Ziyi Xie, Yonggang Zhen, Wenping Hu, and Huanli Dong. Polycyclic aromatic hydrocarbon-based organic semiconductors: ring-closing synthesis and optoelectronic properties. *Journal of Materials Chemistry C*, 10(7):2411–2430, 2022.

[42] Cyril Aumaitre and Jean-François Morin. Polycyclic aromatic hydrocarbons as potential building blocks for organic solar cells. *The Chemical Record*, 19(6):1142–1154, 2019.

[43] Chin-Ti Chen. Evolution of red organic light-emitting diodes: materials and devices. *Chemistry of Materials*, 16(23):4389–4400, 2004.

[44] Yunbin Hu, Peng Xie, Marzio De Corato, Alice Ruini, Shen Zhao, Felix Meggendorfer, Lasse Arnt Straasø, Loic Rondin, Patrick Simon, Juan Li, et al. Bandgap engineering of graphene nanoribbons by control over structural distortion. *Journal of the American Chemical Society*, 140(25):7803–7809, 2018.

[45] Giuseppe M Paternò, Qiang Chen, Xiao-Ye Wang, Junzhi Liu, Silvia G Motti, Annamaria Petrozza, Xinliang Feng, Guglielmo Lanzani, Klaus Müllen, Akimitsu Narita, et al. Synthesis of dibenzo [hi, st] ovalene and its amplified spontaneous emission in a polystyrene matrix. *Angewandte Chemie International Edition*, 56(24):6753–6757, 2017.

[46] Jingyun Tan, Guanghui Zhang, Congwu Ge, Jun Liu, Long Zhou, Chao Liu, Xike Gao, Akimitsu Narita, Yingping Zou, and Yunbin Hu. Electron-deficient contorted polycyclic aromatic hydrocarbon via one-pot annulative π-extension of perylene diimide. *Organic Letters*, 24(12):2414–2419, 2022.

[47] Kalishankar Bhattacharyya, Titas Kumar Mukhopadhyay, and Ayan Datta. Controlling electronic effects and intermolecular packing in contorted polyaromatic hydrocarbons (c-pahs): towards high mobility field effect transistors. *Physical Chemistry Chemical Physics*, 18(22):14886–14893, 2016.

[48] Jakub Wagner, Paola Zimmermann Crocomo, Michał Andrzej Kochman, Adam Kubas, Przemysław Data, and Marcin Lindner. Modular nitrogen-doped concave polycyclic aromatic hydrocarbons for high-performance organic light-emitting diodes with tunable emission mechanisms. *Angewandte Chemie*, 134(27):e202202232, 2022.

[49] Dongdong Zhang and Lian Duan. Polycyclic aromatic hydrocarbon derivatives toward ideal electron-transporting materials for organic light-emitting diodes. *The Journal of Physical Chemistry Letters*, 10(10):2528–2537, 2019.

[50] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

[51] Richard Van Noorden, Brendan Maher, and Regina Nuzzo. The top 100 papers. *Nature News*, 514(7524):550, 2014.

[52] Sharon Hammes-Schiffer. A conundrum for density functional theory. *Science*, 355(6320):28–29, 2017.

[53] Christopher J Cramer and Donald G Truhlar. Density functional theory for transition metals and transition metal chemistry. *Physical Chemistry Chemical Physics*, 11(46):10757–10816, 2009.

[54] Laura Riccardi, Vito Genna, and Marco De Vivo. Metal–ligand interactions in drug design. *Nature Reviews Chemistry*, 2(7):100–112, 2018.

[55] Shuaihua Lu, Qionghua Zhou, Yixin Ouyang, Yilv Guo, Qiang Li, and Jinlan Wang. Accelerated discovery of stable lead-free hybrid organic-inorganic perovskites via machine learning. *Nature communications*, 9(1):3405, 2018.

[56] Bong-Gi Kim, Xiao Ma, Chelsea Chen, Yutaka Ie, Elizabeth W Coir, Hossein Hashemi, Yoshio Aso, Peter F Green, John Kieffer, and Jinsang Kim. Energy level modulation of homo, lumo, and band-gap in conjugated polymers for organic photovoltaic applications. *Advanced Functional Materials*, 23(4):439–445, 2013.

[57] Shujiang Yang, Pavel Olishevski, and Miklos Kertesz. Bandgap calculations for conjugated polymers. *Synthetic Metals*, 141(1-2):171–177, 2004.

[58] Shujiang Yang and Miklos Kertesz. Bond length alternation and energy band gap of polyyne. *The Journal of Physical Chemistry A*, 110(31):9771–9774, 2006.

[59] Ahmed Dkhissi, Jean Marie Ducéré, Ralf Blossey, and Claude Pouchan. Can the hybrid meta gga and dft-d methods describe the stacking interactions in conjugated polymers? *Journal of computational chemistry*, 30(8):1179–1184, 2009.

[60] Bin Liu, Dario Rocca, He Yan, and Ding Pan. Beyond conformational control: effects of noncovalent interactions on molecular electronic properties of conjugated polymers. *JACS Au*, 1(12):2182–2187, 2021.

[61] Thao P Nguyen, Ji Hoon Shim, and Jin Yong Lee. Density functional theory studies of hole mobility in picene and pentacene crystals. *The Journal of Physical Chemistry C*, 119(21):11301–11310, 2015.

[62] Marcel Kastler, Jochen Schmidt, Wojciech Pisula, Daniel Sebastiani, and Klaus Müllen. From armchair to zigzag peripheries in nanographenes. *Journal of the American Chemical Society*, 128(29):9526–9534, 2006.

[63] Jonathan C Rienstra-Kiracofe, Christopher J Barden, Shawn T Brown, and Henry F Schaefer. Electron affinities of polycyclic aromatic hydrocarbons. *The Journal of Physical Chemistry A*, 105(3):524–528, 2001.

[64] Amir Karton. How reliable is dft in predicting relative energies of polycyclic aromatic hydrocarbon isomers? comparison of functionals from different rungs of jacob's ladder. *Journal of Computational Chemistry*, 38(6):370–382, 2017.

[65] Eric Clar. The aromatic sextet. In *Mobile source emissions including policyclic organic species*, pages 49–58. Springer, 1983.

[66] Miquel Solà. Forty years of clar's aromatic π-sextet rule. *Frontiers in chemistry*, page 22, 2013.

[67] Frank De Proft and Paul Geerlings. Conceptual and computational dft in the study of aromaticity. *Chemical reviews*, 101(5):1451–1464, 2001.

[68] Robert Kalescky, Elfi Kraka, and Dieter Cremer. Description of aromaticity with the help of vibrational spectroscopy: Anthracene and phenanthrene. *The Journal of Physical Chemistry A*, 118(1):223–237, 2014.

[69] Dariusz W Szczepanik, Miquel Sola, Tadeusz M Krygowski, Halina Szatylowicz, Marcin Andrzejak, Barbara Pawełek, Justyna Dominikowska, Mercedes Kukułka, and Karol Dyduch. Aromaticity of acenes: the model of migrating π-circuits. *Physical Chemistry Chemical Physics*, 20(19):13430–13436, 2018.

[70] Ramón García-Domenech, Jorge Gálvez, Jesus V de Julián-Ortiz, and Lionello Pogliani. Some new trends in chemical graph theory. *Chemical Reviews*, 108(3):1127–1169, 2008.

[71] Alan R Katritzky, Minati Kuanar, Svetoslav Slavov, C Dennis Hall, Mati Karelson, Iiris Kahn, and Dimitar A Dobchev. Quantitative correlation of physical and chemical properties with chemical structure: utility for prediction. *Chemical reviews*, 110(10):5714–5789, 2010.

[72] Harry Wiener. Structural determination of paraffin boiling points. *Journal of the American chemical society*, 69(1):17–20, 1947.

[73] Tu Le, V Chandana Epa, Frank R Burden, and David A Winkler. Quantitative structure–property relationship modeling of diverse materials properties. *Chemical reviews*, 112(5):2889–2919, 2012.

[74] Lemont Burwell Kier and Lowell H Hall. Molecular connectivity in structure-activity analysis. 1986.

[75] Lemont Burwell Kier and Lowell H Hall. Molecular structure description: the electrotopological state. 1999.

[76] Lowell H Hall, Brian Mohney, and Lemont B Kier. The electrotopological state: structure information at the atomic level for molecular graphs. *Journal of chemical information and computer sciences*, 31(1):76–82, 1991.

[77] R García-Domenech and JV de Julián-Ortiz. Prediction of indices of refraction and glass transition temperatures of linear polymers by using graph theoretical indices. *The Journal of Physical Chemistry B*, 106(6):1501–1507, 2002.

[78] Kaifu Gao, Rui Wang, Jiahui Chen, Limei Cheng, Jaclyn Frishcosy, Yuta Huzumi, Yuchi Qiu, Tom Schluckbier, Xiaoqi Wei, and Guo-Wei Wei. Methodology-centered review of molecular modeling, simulation, and prediction of sars-cov-2. *Chemical Reviews*, 122(13):11287–11368, 2022.

[79] Gregory F Albery, Daniel J Becker, Liam Brierley, Cara E Brook, Rebecca C Christofferson, Lily E Cohen, Tad A Dallas, Evan A Eskew, Anna Fagre, Maxwell J Farrell, et al. The science of the host–virus network. *Nature microbiology*, 6(12):1483–1492, 2021.

[80] Hakimeh Khojasteh, Alireza Khanteymoori, and Mohammad Hossein Olyaee. Comparing protein–protein interaction networks of sars-cov-2 and (h1n1) influenza using topological features. *Scientific reports*, 12(1):5867, 2022.

[81] Felipe Torres, Miguel Kiwi, and Ivan K Schuller. The impact of the suppression of highly connected protein interactions on the corona virus infection. *Scientific Reports*, 12(1):9188, 2022.

[82] Milan Randić. Aromaticity of polycyclic conjugated hydrocarbons. *Chemical Reviews*, 103(9):3449–3606, 2003.

[83] Marcos Mandado, María J González-Moa, and Ricardo A Mosquera. Chemical graph theory and n-center electron delocalization indices: A study on polycyclic aromatic hydrocarbons. *Journal of Computational Chemistry*, 28(10):1625–1633, 2007.

[84] Rudolf Kiralj and Márcia MC Ferreira. Predicting bond lengths in planar benzenoid polycyclic aromatic hydrocarbons: a chemometric approach. *Journal of chemical information and computer sciences*, 42(3):508–523, 2002.

[85] A Kuc, T Heine, and G Seifert. Structural and electronic properties of graphene nanoflakes. *Physical Review B*, 81(8):085430, 2010.

[86] Lam H Nguyen and Thanh N Truong. Quantitative structure–property relationships for the electronic properties of polycyclic aromatic hydrocarbons. *ACS omega*, 3(8):8913–8922, 2018.

[87] Lam H Nguyen, Tuan H Nguyen, and Thanh N Truong. Quantum mechanical-based quantitative structure–property relationships for electronic properties of two large classes of organic semiconductor materials: Polycyclic aromatic hydrocarbons and thienoacenes. *ACS omega*, 4(4):7516–7523, 2019.

[88] Tuan H Nguyen, Lam H Nguyen, and Thanh N Truong. Application of machine learning in developing quantitative structure–property relationship for electronic properties of polyaromatic compounds. *ACS omega*, 7(26):22879–22888, 2022.

[89] Frank Jensen. *Introduction to computational chemistry*. John wiley & sons, 2017.

[90] John C Slater. A simplification of the hartree-fock method. *Physical review*, 81(3):385, 1951.

[91] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

[92] Axel D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, 38(6):3098, 1988.

[93] Chengteh Lee, Weitao Yang, and Robert G Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.

[94] Ivan Gutman and Sven J Cyvin. *Introduction to the theory of benzenoid hydrocarbons*. Springer Science & Business Media, 2012.

[95] The on-line encyclopedia of integer sequences. A000228: Number of hexagonal polyominoes (or hexagonal polyforms, or planar polyhexes) with n cells., 2023.

[96] R Rieger and Klaus Müllen. Forever young: polycyclic aromatic hydrocarbons as model cases for structural and optical studies. *Journal of Physical Organic Chemistry*, 23(4):315–325, 2010.

[97] Sven J Cyvin and Ivan Gutman. *Kekulé structures in benzenoid hydrocarbons*, volume 46. Springer Science & Business Media, 2013.

[98] E Clar and DG Stewart. Aromatic hydrocarbons. lxv. triangulene derivatives. *Journal of the American Chemical Society*, 75(11):2667–2672, 1953.

[99] E Clar and DG Stewart. Aromatic hydrocarbons. lxviii. triangulene derivatives. part ii. *Journal of the American Chemical Society*, 76(13):3504–3507, 1954.

[100] NC Yang and AJ Castro. Synthesis of a stable biradical. *Journal of the American Chemical Society*, 82(23):6208–6208, 1960.

[101] Niko Pavliček, Anish Mistry, Zsolt Majzik, Nikolaj Moll, Gerhard Meyer, David J Fox, and Leo Gross. Synthesis and characterization of triangulene. *Nature Nanotechnology*, 12(4):308–311, 2017.

[102] M Gordon and WHT Davison. Theory of resonance topology of fully aromatic hydrocarbons. i. *The Journal of Chemical Physics*, 20(3):428–435, 1952.

[103] Jun-Jian Shen, Yi Han, Shaoqiang Dong, Hoa Phan, Tun Seng Herng, Tingting Xu, Jun Ding, and Chunyan Chi. A stable [4, 3] peri-acene diradicaloid: Synthesis, structure, and electronic properties. *Angewandte Chemie International Edition*, 60(9):4464–4469, 2021.

[104] MR Ajayakumar, Ji Ma, and Xinliang Feng. π-extended peri-acenes: Recent progress in synthesis and characterization. *European Journal of Organic Chemistry*, 2022(13):e202101428, 2022.

[105] Yanwei Gu, Y Gopalakrishna Tullimilli, Jiaqi Feng, Hoa Phan, Wangdong Zeng, and Jishan Wu. peri-acenoacenes. *Chemical Communications*, 55(39):5567–5570, 2019.

[106] Marco Franceschini, Martina Crosta, Ruben R Ferreira, Daniele Poletto, Nicola Demitri, J Patrick Zobel, Leticia González, and Davide Bonifazi. peri-acenoacene ribbons with zigzag bn-doped peripheries. *Journal of the American Chemical Society*, 144(47):21470–21484, 2022.

[107] Alicia Omist, Gaetano Ricci, Amel Derradji, AJ Pérez-Jiménez, Emilio San-Fabián, Yoann Olivier, and Juan-Carlos Sancho-Garcia. peri-acenoacene molecules: tuning of the singlet and triplet excitation energies by modifying their radical character. *Physical Chemistry Chemical Physics*, 23(41):24016–24028, 2021.

[108] Erich Hückel. Quantum-theoretical contributions to the benzene problem. i. the electron configuration of benzene and related compounds. *Z. physik*, 70(3-4):204–286, 1931.

[109] Erich Hückel. Quanstentheoretische beiträge zum benzolproblem: Ii. quantentheorie der induzierten polaritäten. *Zeitschrift für Physik*, 72(5-6):310–337, 1931.

[110] Erich Hückel. Quantentheoretische beiträge zum problem der aromatischen und ungesättigten verbindungen. iii. *Zeitschrift für Physik*, 76:628–648, 1932.

[111] Erich Hückel. Die freien radikale der organischen chemie: Quantentheoretische beiträge zum problem der aromatischen und ungesättigten verbindungen. iv. *Zeitschrift für Physik*, 83:632–668, 1933.

[112] Kyoko Nakada, Mitsutaka Fujita, Gene Dresselhaus, and Mildred S Dresselhaus. Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Physical Review B*, 54(24):17954, 1996.

[113] Eugenio Di Donato, Matteo Tommasini, Chiara Castiglioni, and Giuseppe Zerbi. Assignment of the g+ and g- raman bands of metallic and semiconducting carbon nanotubes based on a common valence force field. *Physical Review B*, 74(18):184306, 2006.

[114] Roald Hoffmann. An extended hückel theory. i. hydrocarbons. *The Journal of Chemical Physics*, 39(6):1397–1412, 1963.

[115] Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, 4(1):1–17, 2012.

[116] Koichi Momma and Fujio Izumi. Vesta: a three-dimensional visualization system for electronic and structural analysis. *Journal of Applied crystallography*, 41(3):653–658, 2008.

[117] Micheal J. et. al Frisch. Gaussian09.

[118] Frank De Proft and Paul Geerlings. Relative hardness as a measure of aromaticity. *Physical Chemistry Chemical Physics*, 6(2):242–248, 2004.

[119] Alberto Modelli, Laura Mussoni, and Daniele Fabbri. Electron affinities of polycyclic aromatic hydrocarbons by means of b3lyp/6-31+ g* calculations. *The Journal of Physical Chemistry A*, 110(20):6482–6486, 2006.

[120] Angiras Menon, Jochen AH Dreyer, Jacob W Martin, Jethro Akroyd, John Robertson, and Markus Kraft. Optical band gap of cross-linked, curved, and radical polyaromatic hydrocarbons. *Physical Chemistry Chemical Physics*, 21(29):16240–16251, 2019.

[121] Xiao-Ye Wang, Xuelin Yao, Akimitsu Narita, and Klaus Müllen. Heteroatom-doped nanographenes with structural precision. *Accounts of Chemical Research*, 52(9):2491–2505, 2019.

[122] Shane J Goettl, Lotefa B Tuli, Andrew M Turner, Yahaira Reyes, A Hasan Howlader, Stanislaw F Wnuk, Patrick Hemberger, Alexander M Mebel, and Ralf I Kaiser. Gas-phase synthesis of coronene through stepwise directed ring annulation. *Journal of the American Chemical Society*, 145(28):15443–15455, 2023.

[123] Yanwei Gu, Victor Vega-Mayoral, Saül Garcia-Orrit, Dieter Schollmeyer, Akimitsu Narita, Juan Cabanillas-González, Zijie Qiu, and Klaus Müllen. Cove-edged hexa-peri-hexabenzo-bis-peri-octacene: Molecular conformations and amplified spontaneous emission. *Angewandte Chemie International Edition*, 61(18):e202201088, 2022.

[124] Yanwei Gu, Xiaojin Wu, Tullimilli Y Gopalakrishna, Hoa Phan, and Jishan Wu. Graphene-like molecules with four zigzag edges. *Angewandte Chemie International Edition*, 57(22):6541–6545, 2018.

[125] Sabine Seifert, Kazutaka Shoyama, David Schmidt, and Frank Würthner. An electron-poor c64 nanographene by palladium-catalyzed cascade c- c bond formation: One-pot synthesis and single-crystal structure analysis. *Angewandte Chemie*, 128(22):6500–6505, 2016.

[126] Prabhleen Kaur and Md Ehesan Ali. The influence of the radicaloid character of polyaromatic hydrocarbon couplers on magnetic exchange interactions. *Physical Chemistry Chemical Physics*, 24(21):13094–13101, 2022.

[127] Iago Pozo, Zsolt Majzik, Niko Pavliček, Manuel Melle-Franco, Enrique Guitián, Diego Peña, Leo Gross, and Dolores Pérez. Revisiting kekulene: synthesis and single-molecule imaging. *Journal of the American Chemical Society*, 141(39):15488–15493, 2019.

[128] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55. US Government printing office, 1968.

[129] JC Sancho-García and AJ Pérez-Jiménez. Improved accuracy with medium cost computational methods for the evaluation of bond length alternation of increasingly long oligoacetylenes. *Physical Chemistry Chemical Physics*, 9(44):5874–5879, 2007.

[130] Erin M Adkins and J Houston Miller. Towards a taxonomy of topology for polynuclear aromatic hydrocarbons: linking electronic and molecular structure. *Physical Chemistry Chemical Physics*, 19(41):28458–28469, 2017.

[131] Fabrizia Negri, Chiara Castiglioni, Matteo Tommasini, and Giuseppe Zerbi. A computational study of the raman spectra of large polycyclic aromatic hydrocarbons: Toward molecularly defined subunits of graphite. *The Journal of Physical Chemistry A*, 106(14):3306–3317, 2002.

[132] Zhe Sun, Qun Ye, Chunyan Chi, and Jishan Wu. Low band gap polycyclic hydrocarbons: from closed-shell near infrared dyes and semiconductors to open-shell radicals. *Chemical Society Reviews*, 41(23):7857–7889, 2012.

[133] Alexandru T Balaban. Chemical graph xxviii. a new topological index for catafusenes: L-transform of their three-digit codes. *Revue Roumaine de Chimie*, 22(45), 1977.

A| Appendix

In this Appendix, we list all the branched and unbranched catacondensed molecular graphs of BHs or polyhexes from one to six rings.

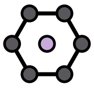
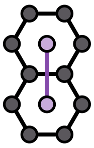
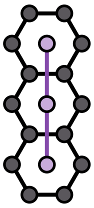
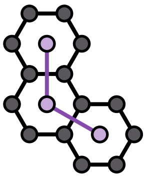
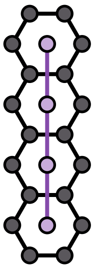
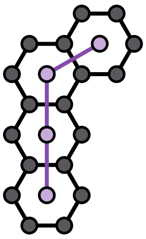
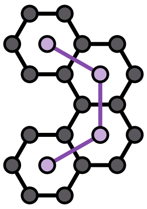
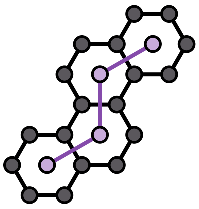
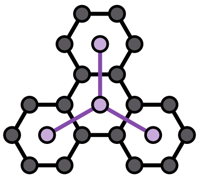
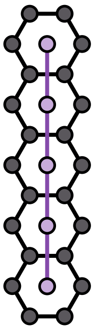
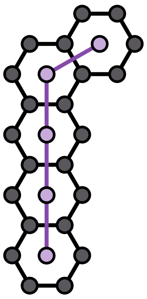
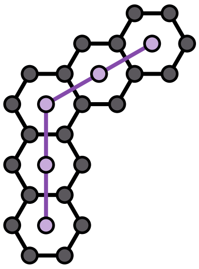
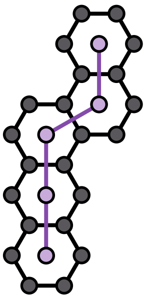
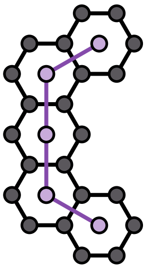
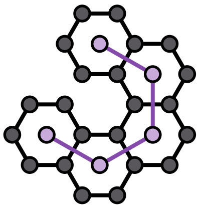
				
1.01	**2.01**	**3.01**	**3.02**	**4.01**
-	*LL*	*LLL*	*LAL*	*LLLL*
				
4.02	**4.03**	**4.04**	**4.05**	**5.01**
LALL	*LAAL*	*LAAL*	-	*LLLLL*
				
5.02	**5.03**	**5.04**	**5.05**	**5.06**
LALLL	*LLALL*	*LAALL*	*LALAL*	*LAAAL*

Table A.1: Continued on next page

Table A.1: Continued from previous page

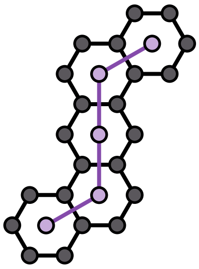
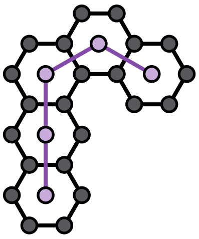
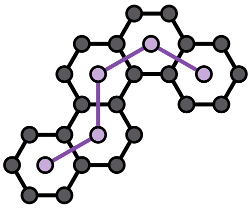
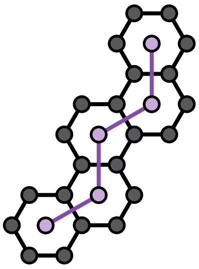
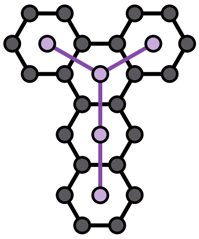
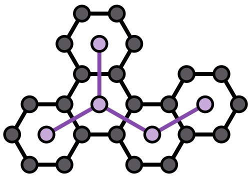
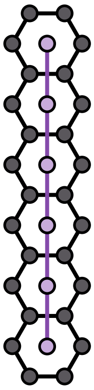
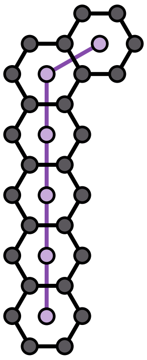
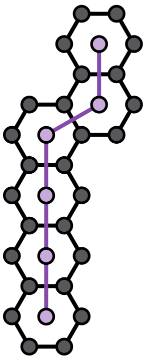
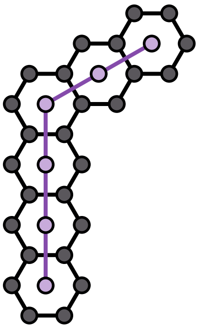
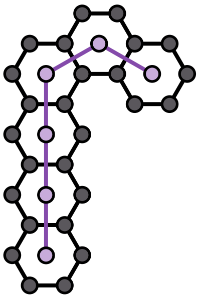
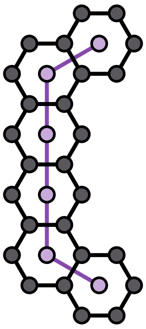
			
5.07	**5.08**	**5.09**	**5.10**
LALAL	*LAALL*	*LAAAL*	*LAAAL*
			
5.11	**5.12**	**6.01**	**6.02**
-	-	*LLLLLL*	*LALLLL*
			
6.03	**6.04**	**6.05**	**6.06**
LAALLL	*LLALLL*	*LAALLL*	*LALLAL*

Table A.1: Continued on next page

Table A.1: Continued from previous page

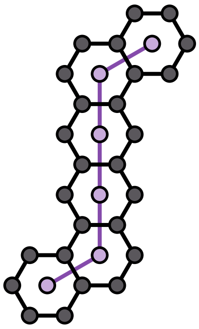
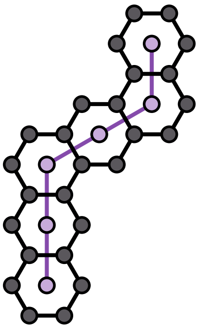
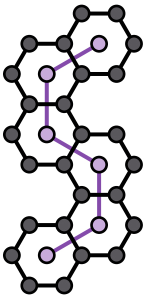
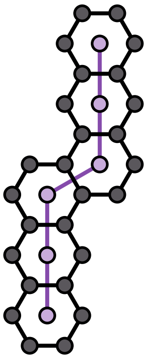
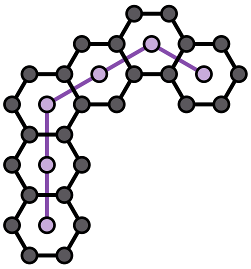
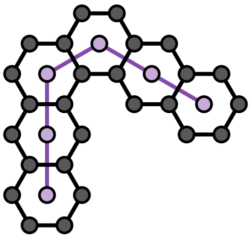
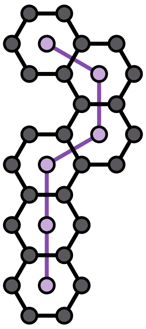
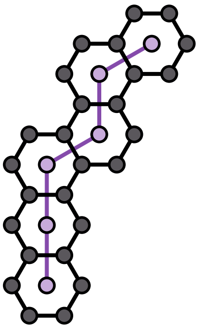
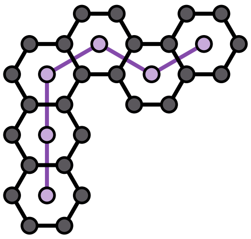
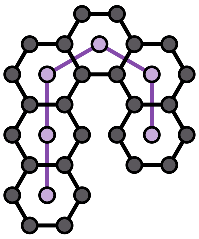
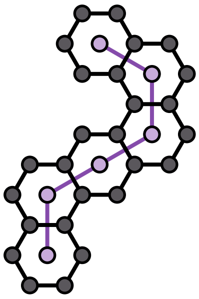
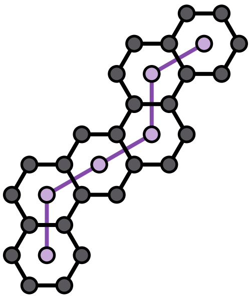
			
6.07	**6.08**	**6.09**	**6.10**
LALLAL	*LALALL*	*LAAAAL*	*LLAALL*
			
6.11	**6.12**	**6.13**	**6.14**
LALALL	*LLAALL*	*LAAALL*	*LAAALL*
			
6.15	**6.16**	**6.17**	**6.18**
LAAALL	*LAAALL*	*LAALAL*	*LAALAL*

Table A.1: Continued on next page

Table A.1: Continued from previous page

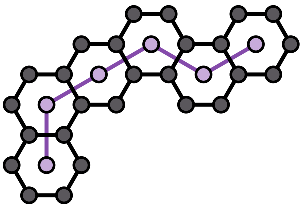
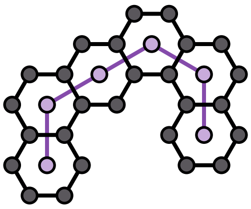
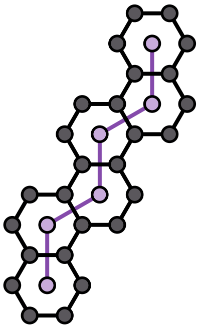
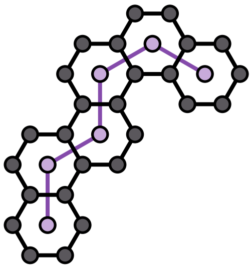
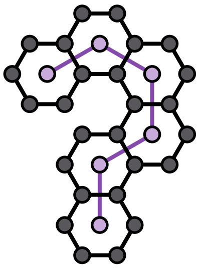
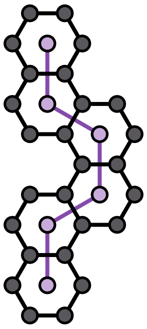
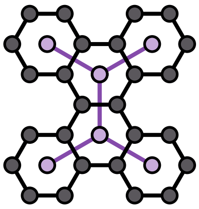
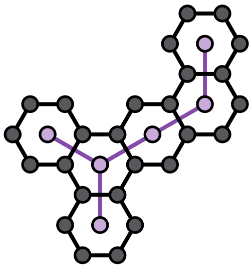
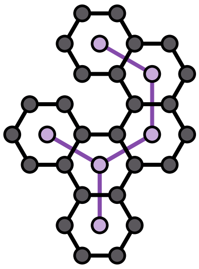
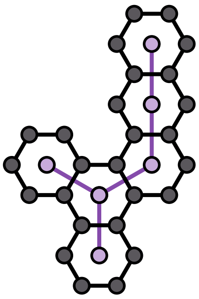
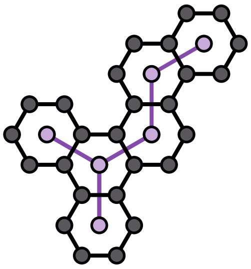
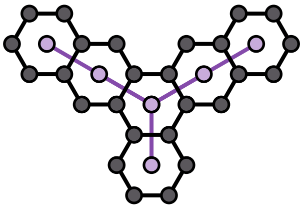
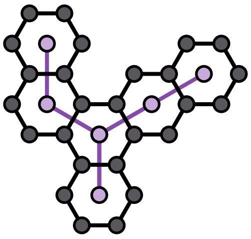
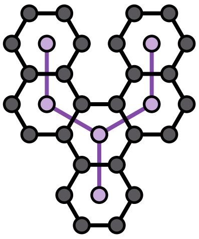
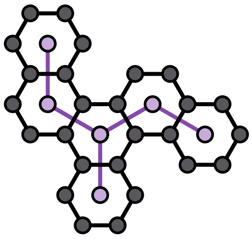
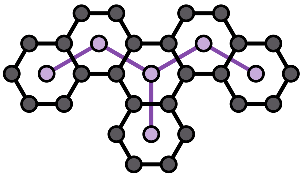
			
6.19	**6.20**	**6.21**	**6.22**
LAALAL	*LALAAL*	*LAAAAL*	*LAAAAL*
			
6.23	**6.24**	**6.25**	**6.26**
LAAAAL	*LAAAAL*	-	-
			
6.27	**6.28**	**6.29**	**6.30**
-	-	-	-

Table A.1: Continued on next page

Table A.1: Continued from previous page

			
6.31	**6.32**	**6.33**	**6.34**
-	-	-	-

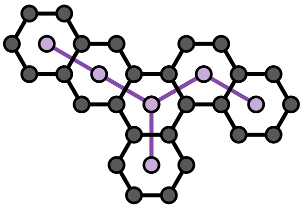
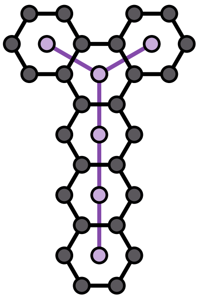
	
6.35	**6.36**
-	-

Table A.1: For each polyhex, in the first row he molecular graphs associate to the monohex [**1.01**], dihex [**2.01**], trihexes [**3.01-3.03**], tetrahexes [**4.01-4.05**], pentahexes [**5.01-5.12**] and hexahexes [**6.01-6.36**] are reported. Grey circles represent the vertices of the molecular graph, black segments the edges of the molecular graph, purple circles the vertices of the dual graph, and purple segments the edges of the dual graph. For each polyhex, an ID is assigned in the second row as outlined in Section 2.2. For each polyhex, the third row features the chain sequence for unbranched catacondensed BHs by recognizing the connectivity of each ring using the notation introduced in Section 2.2.