Novel carbon-based materials mixing different hybridization kinds

In the last twenty years, carbon-based materials and nanostructures have gained more and more popularity. Driven by the breakthrough-discovery and the synthesis of fullerenes, nanotubes, graphene and carbynes, also the search for new exotic carbon allotropes attracted increasing attention in the scientific community, also in view of applications.

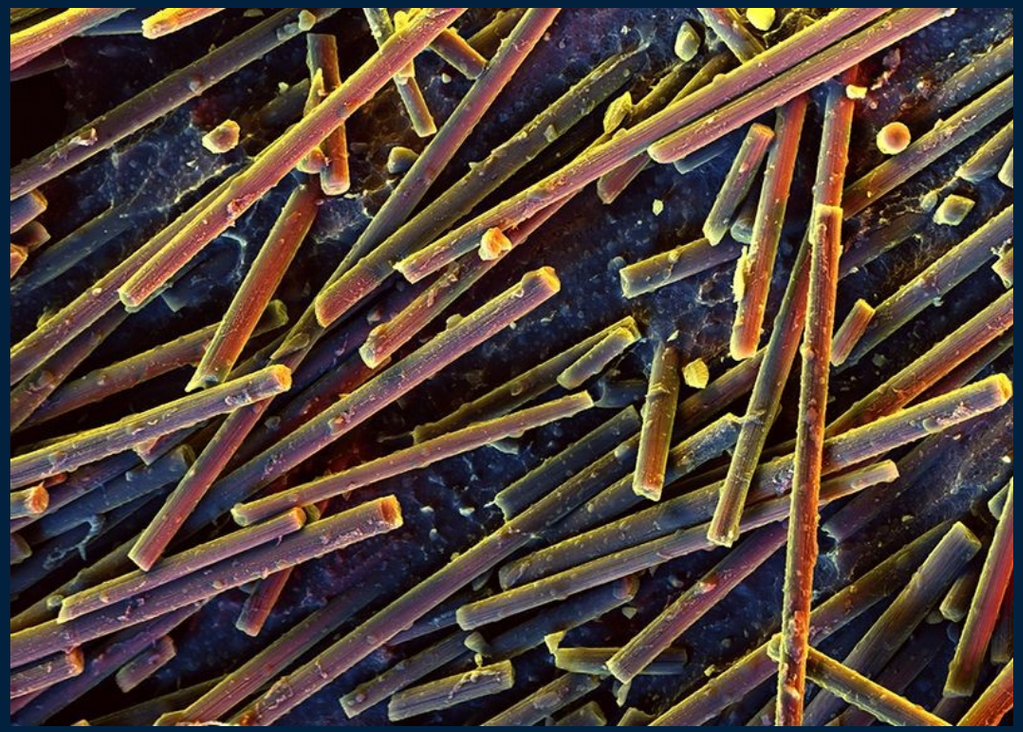
This thesis focuses on the construction and the investigation of three novel crystalline allotropes of carbon, all mixing different orbital hybridizations. We have employed state-of-the-art numerical simulations to investigate structural, electronic, and mechanical properties of the three structures.

Two of these allotropes, novamene and protomene, combine sp^2 and sp^3 hybridizations and exhibit a semiconductor character in their lowest-energy Peierls-dimerized configuration. Both structures show transitions towards a metallic state at a relatively small energy cost.

The third allotrope, zayedene, mixes *sp*, in the form of a linear chain, and sp^3 providing an enclosing cage. This structure exhibits a clear metallic character due to the dangling bonds inside the cavity. We predict characteristic high-frequency vibrations associated with *sp* chain stretching modes. We also investigate the thermodynamic stability of zayedene at standard conditions.

Finally we suggest how hundreds of different allotropes can be built from the simple ones investigated.

F. Delodovici Novel carbon-based materials mixing different hybridization kinds



Novel carbon-based materials mixing different hybridization kinds

Francesco Delodovici

Physics, Astrophysics and Applied Physics PhD School
Università degli Studi di Milano

UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di Dottorato in Fisica, Astrofisica e Fisica Applicata

Dipartimento di Fisica

Corso di Dottorato in Fisica, Astrofisica e Fisica Applicata

Ciclo XXXII

Novel carbon-based materials mixing different hybridization kinds

Settore Scientifico Disciplinare FIS/03

Supervisor: Professor G.ONIDA

Supervisor: Professor N.MANINI

Ph.D. thesis of:

Francesco Delodovici

Anno Accademico 2018/2019

Commission of the final examination:

External Member:
M. Bernasconi

External Member:
C. Molteni

Internal Member:
R. Martinazzo

Final examination:

Date 11-21-2019

Università degli Studi di Milano, Dipartimento di Fisica “Aldo Pontremoli”, Milano, Italy

Cover illustration:

Carbon nanotubes: coloured scanning electron micrograph (SEM) of carbon nanotubes in a POM matrix.

Author:

Stefan Diller - Science photo library

MIUR subjects:

FIS/03 - CHIM/02

PACS: 81.05.Zx

Contents

CHAPTER 1

Introduction

The quest for new materials

During the last decades materials prediction and characterization has gained increasing popularity in the scientific community. Several areas of physics received significant boost from this quest and new physics was discovered. Supported by a rich theoretical framework and driven by a number of crucial experimental discoveries in the field, the search of exotic carbon allotropes pioneered this quest. Fullerenes [1], nano-onions [2], carbon nanotubes (CNT) [3, 4] , sp linear carbon chains (SPCC) [5] and graphene [6] are cornerstones in material science and opened the route to theoretical analysis and experimental synthesis of numerous innovative structures. For instance, graphene is one of the forefathers for the physics of bi-dimensional materials which is nowadays a rich playground for several disciplines, from quantum field theory to electronic engineering. As a matter of fact, low-dimensional systems including graphene and SPCCs intrigued physicists for long times to the point of being addressed and largely theoretically characterized long before their synthesis [7–9]. The experimental realization of graphene prompted the synthesis of several 2-dimensional, or quasi 2-dimensional materials, such as: hexagonal boron nitride, molybdenum disulfide, black phosphorus [10, 11]. Even in this case some of these low-dimensional systems, such as molybdenum disulfide, had already been theoretically characterized. In a similar way, the synthesis of carbon nanotubes prompted that of inorganic nanotubes, made of boron-nitride [12] or transition-metal disulfides [13, 14]. In this context, the evolution of carbon allotropes was so important to induce Hirsch to define of the current century as the era of carbon allotropes [15]. What is common to the synthetic carbon allotropes is the contemporary presence of conjugated π electrons and of strong covalent bonding in low dimensionality: either zero, one or two-dimensions. These characteristics lead to unusual physical and chemical properties.

Among these low-dimensional systems, perhaps graphene is currently the most popular and broadly investigated one. Graphene uniqueness arises mainly form its bi-dimensional atomic honeycomb arrangement. Due to the trigonal-planar bonding geometry, graphene has a remarkable structural flexibility. σ bonds are responsible for its high mechanical resistance, ultra-high elastic modulus and ultimate tensile strength [16, 17]. At the same time, the electronic properties of graphene mainly arise from the half-filled π orbitals. Graphene is a semi-metal with Fermi level falling exactly at the so-called Dirac cones [18]. These are features in the band structures with peculiar approximately linear dispersion relation. Given the bi-dimensionality of the systems the bands become locally the intersection of two conical surfaces. The electrons wave function close to these points is described by the relativistic Dirac equation, even though their

Fermi velocity is considerably lower than the speed of light, approximately $v_F = c/300$ [19]. These electrons behaving as mass-less Dirac particles exhibit unusual behaviour with respect to ordinary electrons in metals. For instance integer [20] and fractional [21] quantum Hall effect and Klein tunneling have been observed in graphene [22]. Klein tunneling specifically allows electrons to propagate up to micro-metric distances without being scattered [6, 23]. A direct consequence is graphene's remarkable charge-carrier mobility with no need for doping [23–25]. Dirac electrons have a chiral nature, following the honeycomb decomposition in triangular sublattices. This characteristic is employed in "valleytronics" where the electron flavor becomes a controllable variable [26]. Moreover, graphene's electronic properties depends on it's mesoscopic shape. For instance, different boundaries come with different electronic character [27]: zig-zag nano-ribbons are predicted to be always metallic [28], whereas armchair nano-ribbons can be metallic or semiconducting [29]. The gap, whose size decreases with increasing the ribbon width, opens due to the anti-ferromagnetic coupling between the magnetic moments of edge atoms sitting on different sublattices. Besides the unusual electronic properties, bi-dimensional structure of graphene also comes with interesting mechanical features. For instance, it posseses a flexural vibrational mode [30] which is absent in bulk systems and it is responsible for the out-of-plane corrugation of graphene and for its superior thermal conductivity at room temperature together with the other acoustic modes. The Young's modulus of graphene, that is of the order of TPa, is also related to graphene's flexural mode, through its characteristic quadratic dispersion relation [30, 31]. Besides they are also connected through the mean-square amplitute of thermal vibration at finite temperature [32].

Graphene can be thought of as a "precursor" of the fullerenes and nanotubes. Indeed, fullerenes can be imagined starting from graphene flakes introducing pentagons defects and conveniently wrapping the flake. Instead, nanotubes can be thought as rolled graphene ribbons along a given direction. Fullerenes have been first convincingly detected in 1985 [1] after being largely debated and first-time observed several times [33]. Since then, they received significant attention from the scientific community. Specifically, the one detected by Kroto et al. is C_{60}, consisting of twelve pentagons located at the vertex of an icosahedron whose faces are composed by twenty regular hexagons. Twelve is the only possible number of pentagons, whereas the hexagons number can be higher or lower. A high number of combinations has been predicted and some of them experimentally detected, including non-spherical close cages such as C_{70} which is characterized by an ellipsoidal shape. In 2000 the smallest possible fullerene C_{20} was synthesized [34], made by 12 pentagons arranged in a tetrahedron. Different experimental techniques [35, 36] reveal a discrepancy in bond lengths in C_{60}: the bonds involving two adjacent hexagons are sistematically shorter than those involving hexagons and pentagons. This is due to the fact that C_{60} is not spherically aromatic [37], meaning that double bonds are avoided on pentagons. Thus the delocalization of electrons around the molecule is limited, contributing to the buckyballs electron affinity [38]. The geometry of C_{60} can be interpreted in terms of sp^2 hybridization, even though the curvature of fullerenes requires a significant deviation from trigonal planar shape. Bent planar configuration confers an excess of strain to the structure. This extra strain contributes to fullerenes enhanced chemical reactivity [39]. Besides appearing as isolated molecules, fullerenes can form extended patterns known as fullerites: regular arrangements of C_n connected by Van der Waals bonds with semiconductor character [40–42]. Doping or functionalization can attribute metallic or even superconductive character to fullerites [43].

Carbon nanotubes can be imagined as degenerate fullerenes, even tough they consist in tubes made of rolled graphene nanoribbons [44]. Their radius ranges from a few to tens of nanometers, while they can be up to half meter long [45]. CNTs exist in single, double-walled or multi-walled configurations. The multi CNT configurations consist of nested coaxial tubes, interacting with Van der Walls forces, with an average separation of the order of 0.3 nm. CNTs can be divided in two categories, depending on the path that encircle the tubes orthogonally to its central axis. According to this path they are named zigzag and armchair. CNTs not belonging to these two families are named chiral, and lack inversion symmetry [46]. This distinction radically affects electronic properties [47]. Indeed all armchair CNTs are metallic, whereas the zizag have a small gap at the Fermi level. Chiral CNTs show a configuration-dependent behaviour [48]. This net separation in the electronic behaviour has exceptions, because electrical properties are strongly dependent on curvature effects, particularly in small-diameter tubes [47]. Despite their immediate connection with graphene, CNTs do not show analogous to Dirac cones. This is due to the distortion from the sp^2 planar hybridization, which induces hybridization between σ and π orbitals that modifies the bands dispersion. Nonetheless valleys can be identified also in nanotubes. Indeed it is possible to divide the electronic orbits in two flavors, that roughly correspond to clockwise and counterclockwise motion around the nanotube [46]. Similarly to graphene, sp^2 covalent bonds confer to nanotube a remarkable tensile strength and elastic modulus. Such characteristics are referred to single CNTs. In case of multi-walled CNTs or bundles of CNTs weak shear interactions between adjacent tubes or shells lead to significant reduction in the effective strength [49]. A possible way to partially compensate this loss appears to be light-irradiation to induce cross-linking among the different shells or single tubes [50].

Alongside these low dimensional systems, it has largely been debated the existence of an allotrope consisting only of sp hybridized carbons. This search for sp-carbon wires dates back to the second half of the XX century [51–54]. Due to their peculiar mono-dimensional nature they have been elusive for long time, challenging chemists for half a century, causing also some inconvenient [55]. The analysis of linear sp chains received a boost following the enthusiasm for graphene synthesis, and they gained popularity for their enhanced mechanical and transport properties [56–59]. Contrary to other synthetic allotropes, stable extended SPCCs are extremely difficult to obtain even with modern laboratory techniques. Indeed, the longest chain ever reported consists of 6000 carbons, embedded in a double-walled CNT [60]. The scaffold is crucial to confer stability to the carbon wire, as confirmed by the current length-limit for the totality of isolated chains, which is approximately 40 atoms [61]. Due to their instability, isolated sp-chains have been studied only in gas-phase [62] or by means of matrix isolation at low temperature [63], or when separated by synthesis reactions byproducts to prevent chain-chain interactions [64]. On the other hand, pure carbon solid containing considerable amount of carbynoid structure can be efficiently produced [65]. The reason for SPCCs instability is their extreme reactivity towards chemical environment, whether it consists of controlled-atmosphere [66, 67] or other chemical reactants such as water and oxygenr [58, 59]. Also thermal fluctuations induce quick buckling in the chains which can lead the sp atoms to undergo transition to sp^2 or sp^3 hybridizations in order to gain stability [59]. In theory, sp-carbon chains may exist in two different states, depending on how carbon atoms link. Chains with carbon connected with all-equivalent double bonds are named cumulenes, those made of alternated single (long) and triple (short) bonds are named polyynes. Due to Peierls distortion, infinite chains, which are usually referred to as carbyne, ex-

ists in the polyyne state [58, 68]. For finite chains, *ab-initio* simulations show that such distortion induces a significant bond length alternation (BLA) only for chains longer than fifty atoms [69]. Despite being ideal systems, carbynes have been widely characterized theoretically, revealing interesting properties. Simulations result in Young modulus exceeding that of graphene [70]. High frequency phonons and ballistic thermal transport allow for remarkable thermal conductivity at room temperature [71]. Electronic properties change with the Peierls transition. Indeed, cumulene carbynes are metallic, whereas polyyne are semiconducting [72]. Furthermore, due to their one-dimensional nature, the electronic properties of finite SPCCs radically depends on the local environment, including the different end groups which can be employed to stabilize the chains [73]. For instance hydrogen-terminated SPCCs were stabilized for several months together with Ag nanoparticles when diluted in solutions. SPCCs properties appear to be tunable by means of a broad range of tools: engineering the length by means of external strain, embedding the chains within different chemical environment or within different end groups.

A further step in the search for artificial allotropes of carbon, was the synthesis of polyynes terminated with sp^2-conjugated end-groups [74]. Such systems proved to be stable in ambient conditions without need of solvents even against reaction with oxygen [75]. Moreover, they open the route to the creation of systems graphene-like, but based on sp-sp^2 combination. In such systems, the connection sp sp^2-conjugated carbons could affect the benzenic-ring conjugation. This could produce a further delocalization of π electrons from the molecule to the chain, affecting its BLA and therefore its electronic gap. Evidences emerged regarding the independence of the electronic behaviour of sp-chains with respect to the dimension of the sp^2 end-groups. Thus, to tune the electronic properties of SPCCs, the use of extended graphene-like terminations turns out to be more or less equivalent to finite-size terminations for tuning SPCCs properties. Specific end groups should be employed instead [58]. Indeed, it is now recognized that rather than the nature of the end group, the specific chemical connectivity between sp chain and the end group is responsible for tuning the chain character. Moving from phenyl-groups terminations, extended artificial bi-dimensional structures involving regular arrangements of sp-sp^2 carbons have been synthesized. Such systems, named graphynes, had been addressed years back as modification of graphene [76], but they recently gained more attention. For instance, Sun et al. [77] were able to synthesize graphyne made of four-atoms carbon-wires connected trough benzenic rings on a gold substrate starting from halogenated precursors of carbynes. This study along with others [78] addressing the possibility of mixing orbital hybridizations, set the basis for a further exploration of the family of artificial carbon allotropes.

Besides scientific interest, what makes these allotropes appealing are their numerous potential applications. For instance, chemical functionalization of fullerenes provides the possibility of combining their properties with those of other compounds. Even though a newer-generation of more efficient organic photo-voltaic devices based on non-fullerene acceptors has emerged in the latest years [79, 80], fullerenes-based devices have been widely studied for their promising electron-acceptor nature [81, 82]. Moreover, fullerenes have been largely analyzed in chemical and biological applications [83, 84] such as in inhibition process of HIV-1 replication. Nanomedicine benefits also from functionalized carbon nanotubes which could be employed as a tool to smart drug delivery [85]. CNTs have also been debated as potential water transport pumps and filters [86, 87]. Besides, their strength and flexibility suggest that they could have a significant role in nanotechnology engineering for instance as a reinforcement for polymer matrices [88].

Graphene, as well as other 2-dimensional systems, has been widely and intensely studied for digital and analog electronics [89], for flexible and stretchable electronics [90] and opto-electronics [91, 92]. Also linear sp chains have been intensively investigated due to their potential applications in molecular electronics, specifically when integrated within other carbon allotropes [58].

Nonetheless, compared to the wide variety of newly identified structures [93], relatively few breakthrough applications have been realized to date. Indeed, the depicted outstanding properties of carbon allotropes are limited by several factors. For instance, graphene suffers from the lack of a systematic and practical approach to open a band-gap [94–97] and from the interaction with any surrounding "supports", which actually affects its properties. Furthermore, graphene properties appear to depend directly on the method of production of the layers. At the moment the capability of generating extended layers with industrially sustainable techniques is limited. Nanotubes and carbynes face similar issues. For the CNTs production it is not yet clear how to control their helicity and how to dis-entangle bundled CNT matrices and then disperse them uniformly and align them in a controlled way [98, 99]. SPCCs instead do not exists in free-standing conditions as they need physical supports right from their synthesis. Besides they show a high chemical reactivity toward chemical environment [59, 65, 100]. All these limitations, as much as the quest for new physics, routed the search for materials with similar properties, but easier to produce and control compared to carbon, specifically for bi-dimensional systems. Therefore, the quest for new materials, including carbon allotropes, as for new synthesis routes and new applications, focusing on electronics, opto-electronics, quantum computing, coatings, continues nowadays.

Theoretical research, running in parallel to the experimental inquiry, greatly benefited from the evolution and the spread of computational resources. The diffusion of *ab-initio*-based codes, as much as the development of new more agile algorithms [93, 101] opened new possibilities in this area of physics. Large databases of characterized structures are available today [102, 103], with a clear tendency to develop computational environments aimed to automated high-throughput calculations and to the automatic storage and sharing of data [104, 105]. Besides the possibilities of increase the information sharing across the scientific community, the underlying idea consists in removing guesswork from materials design helping experimental research to target the most promising materials starting from computational datasets. A possible drawback of this race toward automatization is that research becomes closer and closer to informatics, and materials physics to data-mining. Among these open-access digital materials databases, the Samara carbon allotropes database (SACADA) [106] collects hundreds of theoretical and experimental polytopes of carbon. This high number of structurally non-equivalent atomic arrangements arises from the variety of possible carbon orbital hybridizations. Indeed, carbon possesses three energetically competitive different types of hybridization sp, sp^2, and sp^3 [107, 108]. The sp^3 configuration gives rise to three-dimensional networks with insulating properties along with high stiffness, as in cubic and hexagonal diamond [109]. In contrast the sp and sp^2 hybridizations can be responsible for flexible low-dimensional structures [23, 58, 88] such as carbyne and graphene, which usually come with smaller electronic interband gaps, or even metallic or semi-metallic properties. Intermediate hybridizations are quite frequent as well, as in the fullerenes and the nanotubes.

During this PhD, I focused on the definition and the characterization of new allotropes of carbon. At the basis of this work stands a paradigm regarding the possibility of mixing different carbon hybridization kinds in such a way to obtain novel periodic structural configurations. These

"hybrid" allotropes may come along with intriguing physical and chemical properties, as is the case for those investigated in this thesis. For instance, we mention the possibility of controlling a semiconducting-metal transition by tuning the temperature. For what concern spontaneously formed allotropes, the possibility of mixing hybridizations is not a-priori excluded by nature, but apparently it is restricted to non-periodic structures such as amorphous carbon [110–112]. Beyond carbon-only based structures, nature frequently provides examples of coexistence of different hybridizations of carbon orbitals in the same composite. Following this direction deep search and analysis have been performed in the past decades. Indeed, various carbon nano-aggregates have been widely analyzed theoretically and experimentally [113, 114], with the aim of understanding and controlling their growth or aggregation properties, and eventually controlling relative percentages of hybridizations. Occasionally hybrid mesoscopic crystalline structures showing different atomic coordinations were synthesized: for instance fullerites. Besides, fullerites were largely employed as precursor for more complex nanostructures: schwarzite, hollow diamonds, nanowires [115–117]. Moreover, carbon nanoparticles, nanocluster and thin films [114, 118] gained broad interest for their numerous potential applications: from enhancer of electrical conductivity in polymer materials to drug release in medicine, from reinforcement of rubbers, to aerospace applications. Lately the inquiry for unconventional carbon materials has been push forward by interesting discoveries such as that of Q-carbon [119] and graphynes structures [58, 65]. In this sense, a deeper theoretical knowledge of the physical and the chemical properties of "hybrid" complex allotropes, could provide hints for their synthesis and their further usage in suitable applications.

Thesis summary

The theoretical discovery and the characterization of three new crystalline carbon allotropes is reported in this PhD thesis. Two of them are based on mixing sp^2 and sp^3 orbital hybridizations, the third one includes linear hybridization sp combined with sp^3. Chapter 2 defines and investigates the single-ring novamene [120] and the protomene allotropes [121]. The structures are outlined and commented in detail with their energetics compared with the standard carbon allotropes. The electronic and vibrational properties are also outlined and commented. A broader class of novamene allotropes is also introduced. These allotropes exhibit the possibility to undergo a transition from metallic to semiconductor behaviour and back. Such transition is induced by the formation of dimers reducing the presence of planar hybridizations in the allotropes. Chapter 3 presents the simplest structure of the zayedene class [122]. Its energetic, structural and electric properties are introduced. It possess a well-defined metallic behaviour independently on the presence of the sp carbons. The thermodynamical stability of the linear sp-chain is analyzed starting from the high-temperature behaviour of the structure. The composition of this class and preliminary results for a few of its elements are sketched. Chapter 4 draws the conclusion and the outlook of the present research.

CHAPTER 2

Mixing sp^2 and sp^3 hybridizations

The current chapter focuses on two carbon allotropes obtained combining planar and tetrahedral hybridizations, named novamene and protomene. The former is actually a class of allotropes, each element being defined by a certain number of benzenic ring around which the basic unit is build. The allotrope analyzed hereby is the simplest element in this class: the single-ring novamene. Instead, the planar hybridized carbons in protomene do not form such a regular arrangement as benzenic ring. Thus we did not identify a class of connected allotropes. Both the structures were analyzed throuh *ab-initio* simulations for instance comparing their relative stability and their electronic behaviour to those of natural carbon allotropes.

2.1 Novamene

The allotrope we describe in the present section, is based and developed starting from the atomic structure of fullerenes. Indeed, if one considers one of the benzenic-ring of a standard C_{60} bucky-ball, with the three pentagons surrounding it one obtains the basic building block of the current allotrope [123–125]. Instead of bending the pentagons to form the "spherical" shape, one could keep them in the same plane of the benzenic ring. This basic structure has a three-fold rotational symmetry around the z axis at the center of the hexagon, assuming the ring in the xy plane. In order to construct a new allotrope, one needs to define its primitive cell. The basic structure of fullerenes is a good starting point, but it is necessary to fill the three empty volumes included between two pentagons in such a way to allow periodicity in space. One possible solution is presented in Fig. 2.1. If the carbons on the external sides of the pentagons have tetrahedral sp^3 hybridization, then is possible to fill the empty space between the pentagons without bending the structure towards a bucky-ball shape. The resulting cell is described by a threefold rotational symmetry, with the axis placed at the center of the hexagon. This identifies the underlying Bravais lattice as hexagonal. However, the periodic repetition of such a cell, along a hexagonal lattice primitive vectors led a bulk system spotted by holes, as shown in Fig. 2.2. Such cavities would lower the stability of the structure, eventually giving rise to local reconstructions modifying the regular allotrope. Thus, it is necessary to introduce two more carbon atoms per primitive cell, disposed at the middle of two out of the six cavities reported in Fig. 2.2 in order to prevent significant reconstruction events. The result of this scheme is a new allotrope of carbon, which is named "Novamene".

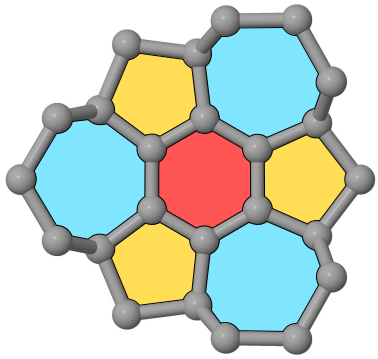


Figure 2.1: Possible arrangement of sp^2 carbon inside a network of pentagons and heptagons that prevents the formation of a C_{60} bucky-bal. This pattern keep the overall average structure flat, even though not planar. The basic planar geometry of fullerene is highlighted with red and yellow, the heptagons preventing the buckling are colored in light-blue.

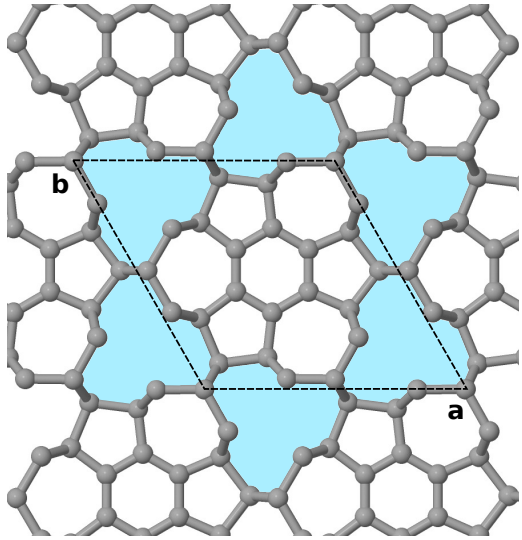


Figure 2.2: The cavities in the periodic repetition of the structure shown in Fig. 2.1 are highlighted in light blue. The primitive cell is identified with dashed black lines.

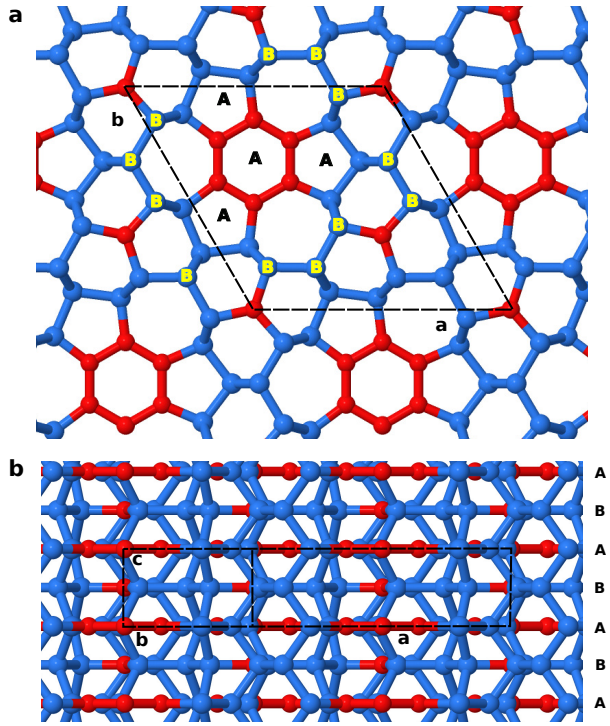


Figure 2.3: The basic structure of single-ring novamene: the primitive cell is highlighted with the dashed line. The switching carbons are colored in red, all the remnants in blue. Panel 1 reports a cut perpendicular to the **c** primitive vector, the atoms are labelled with letter A or B, according to the plane where they belong. Panel 2 reports a vertical cut in the plane of **a** and **c** primitve vectors; the planes are labelled with A and B.

2.2 The crystal structure of single-ring novamene

Figures 2.1 and 2.2 show the result of surrounding the six-carbon ring with sp^3 carbons. This sp^3 carbon arranges in 3 five-carbon rings and 3 seven-carbon rings, which assemble together. The seven-carbon rings, locally recognized as hexagonal diamond, partially sticks out of the plane giving rise to a non-planar configuration. The first model of this allotrope was constructed as a physical three-dimensional ball-and-stick model. With such a model, at first a unit cell has been identified and starting from that, the primitive unit cell and its symmetries were determined. This primitive cell is reported in Fig. 2.3: the section perpendicular to the **c** vector is reported in panel 1, and panel 2 reports a vertical section in the plane xz. Since the study of the proposed carbon allotrope starts from a qualitative model, it is important to propose physical constraints to the bond-lengths in this arrangement. As such, the proposed structure is better understood by making comparisons to known carbon allotropes, including diamond, graphite lonsdaleite, and fullerene. Indeed, the proposed structure draws on elements of each of these compounds. For instance as previously stated, the motif of single hexagonal rings of carbon surrounded by three carbon pentagons is similar to the fundamental repeating pattern in “bucky-balls”. In that structure, C-C

Number of rings	Number of combinations
1	1
2	1
3	3
4	7
5	22
7	339
9	7036
12	829987
15	110716585

Table 2.1: The number of nonequivalent combinations hexagonal rings as a function of the number of rings in the plane.

distances in the hexagonal rings surrounded by a total of six alternating pentagons and hexagons are on the order of $1.45 - 1.49$ Å [125]. In contrast, in the present novamene the grouping of 3 pentagons surrounding a central hexagonal ring remains "in plane", thanks to 7-atom carbon rings popping up from the A to the nearing B plane. For the AB interplanar distance we initially take a tentative 2.5 Å. The initial guess for the structure is built enforcing the threefold symmetry around the center of the sp^2 hexagon, and a repeated hexagonal cell with the same symmetry. Each periodically repeated cell includes 26 carbon atoms, 15 of which in the A plane (the one containing the sp^2 hexagons) plus 11 in the B plane, see Fig. 2.3. It is crucial to consider the possibility that the atoms in the B layer marked in red, hereby named "switching atoms", could move out of the plane switching from sp^2 to sp^3 and forming an extra bond with their partners in the next B layer. As this bonding occurs with the atoms in one plane moving toward the other, the two planes are not equivalent any more, with the result that the minimum repeated cell for the dimerized structure includes a ABA'B' alternation for a total of 52 atoms. This possiblity should be taken into account also for the sp^2 atoms in the benzenic ring. It is intuitive that in this case the different sp^2 carbon cannot form extra bonds indipendently one from the other. The primitive cells are reported in Fig. 2.4 and 2.5. The latter represents the configuration where all the switching atoms in red dimerize with their counterparts in previously equivalent layers. The former represents a possible halfway configuration where the benzenic ring is not distorted. Each of these structures is characterized by the same space group P-62m (#189).

2.2.1 The novamene class

Each combination of rings with the additional enveloping lonsdaleite layer leads to a different allotrope of carbon. A parallel to this allotropy was demonstrated with the fullerenes: their study began with the pursuit of identifying the structure of C_{60}. This concept grew rapidly with other allotropes such as C_{70}, carbon nano-onions, and so on. In turn, this lead to the idea and discovery of single-walled carbon nanotubes, multi-walled ones, and numerous other fullerene-type allotropes of carbon. The geometric arrangements of adjacent hexagonal rings are called fusenes [126, 127], as sketched in Fig. 2.6, and numbered in Tab. 2.1. For a quick idea of hexagonal fusenes possible nonequivalent arrangements see Fig. 2.6. This classification provides a "menu"

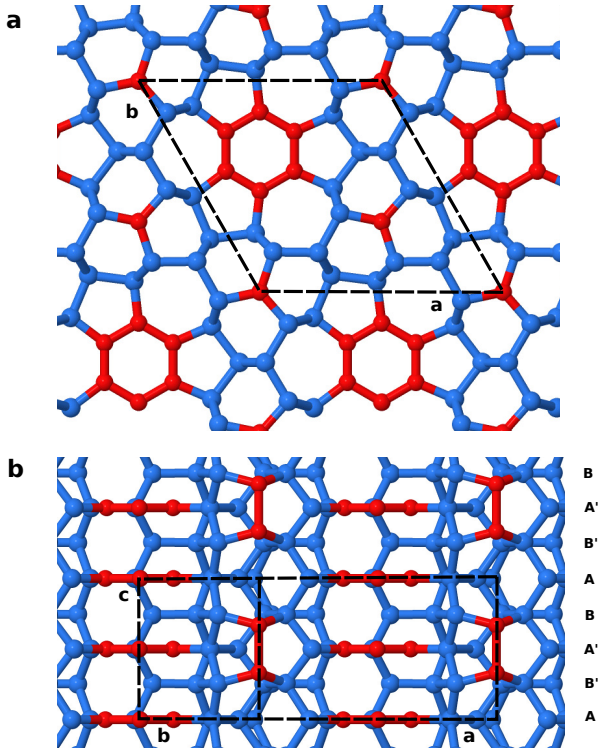


Figure 2.4: The structure of single-ring novamene where only the isolated sp^2 atoms form dimers. The primitive cell is highlighted with the dashed line. Panel 1 reports a cut perpendicular of the **c** primitive vector. Panel 2 reports a vertical cut in the plane of **a** and **c** primitve vectors; the planes are labelled with A,A' and B,B'.

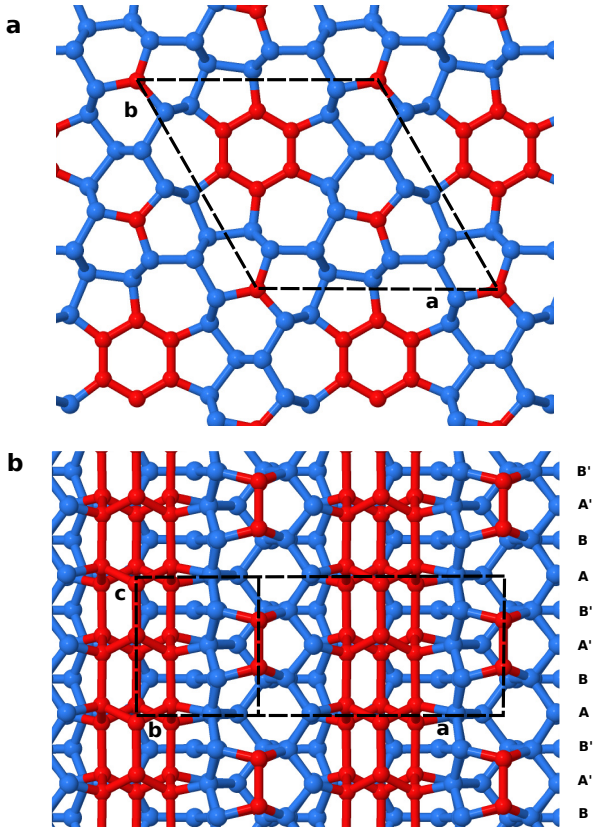


Figure 2.5: The structure of single-ring novamene with all possible dimers formed: those involving the isolated sp^2 carbons and those involving the sp^2 carbons forming the benzenic ring. The primitive cell is highlighted with the dashed line. Panel 1 reports a cut perpendicular to the **c** primitive vector. Panel 2 reports a vertical cut in the plane of **a** and **c** primitve vectors; the planes are labelled with A,A' and B,B'.

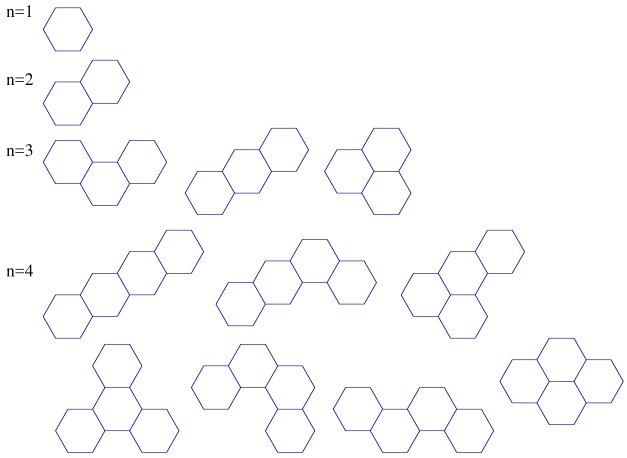


Figure 2.6: The possible nonequivalent disposition of hexagonal fusenes, which consist in the basis of the novamene class, see Tab. 2.1.

for the configurations of the hexagonal rings forming the sp^2 carbon core of the novamene compounds that we introduce in the present work. Based on the above considerations, it is intuitive that other novamene-type allotropes would also be possible. This concept expands the number of rings "encased" in hexagonal diamond, tilting the sp^2-sp^3 balance in favor of sp^2, and leading to a number of other structures.

2.3 Structural and electronic properties

To obtain a reliable equilibrium structure for the single-ring novamene, we employed state of the art DFT simulations. For this allotrope the electronic structure is described in the local density approximation (LDA) using a plane-waves basis with a kinetic energy cutoff of 408 eV, with standard ultrasoft pseudopotentials to account for the 1s core electrons of carbon [128] as implemented in the Quantum Espresso [129, 130] package. The same standard approach was adopted in similar carbon-only contexts in previous works [100, 131–137]. Starting from the tentative initial structure described in section 2.2, the atomic positions and cell parameters are relaxed until a local minimum of the total DFT energy is identified. Atomic relaxation is carried out until all all components of all forces are equal to or less than 4 pN, the residual stress less than 10 bar and the total energy converged to less than 10^{-6} eV/atom. Tab. 2.2 reports the main structural and electronic properties of three possible configurations of the single-ring novamene compared to selected carbon allotropes. The * symbol refers to experimental values. Each configuration corresponds to a different possible combination of the switching atoms present in the structure, coloured in red in Fig. 2.3, 2.4 and 2.5. The first column of Tab.2.2 describe the configuration where no one of the switching atom has an sp^2 hybridization. Thus the benzenic ring warps to bond with the vertically nearest-neighbour rings, as depicted in Fig. 2.5. The second column represents a configuration where the benzenic ring does not dimerize while the two extra sp^2 carbon do, as shown in Fig. 2.4. The third column describe the structure reported in Fig. 2.3 where no

Structural parameter	Single-ring novamene ground state	Single-ring novamene benzenic	Single-ring novamene no dimers	graphite	diamond	lonsdaleite
N_{atoms} per cell	52	52	26	4	2	4
ΔE per atom [eV]	0.203	0.227	0.269	-0.146	0.0	0.025
Band gap [eV]	3.061	0.34	0.0	0.0 0.0*	4.22 5.47*	3.03 4.5*
a,b [Å]	8.440	8.419	8.418	2.433 2.464*	3.523 3.567*	2.477 2.510*
c [Å]	4.791	4.998	2.519	5.898 6.711*	– –	4.126 4.120*
Density [kg/m^3]	3509	3381	3354	2639 2261*	3649 3516*	3638 3521*

Table 2.2: The main structural properties of three novamene configurations. The first column corresponds to the lowest-energy configuration with all dimers formed, reported in Fig. 2.5, in which all the sp^2 atoms dimerize. The second column describes the configuration reported in Fig. 2.4, where the isolated sp^2 carbon form dimers, while the benzenic ring is still undistorted, hereby named "Novamene benzenic" . The third column corresponds to the configuration with no dimers formed, shown in Fig. 2.3. The numerical values for the standard carbon allotropes are reported as a comparison. The * indicate experimental values which of course are not yet available for novamene. The difference in binding energy is computed through equation 2.1, where the reference binding energy of diamond is $E_{\mathrm{b}}^{\mathrm{LDA}} = -8.908$ eV, as computed using the LDA.

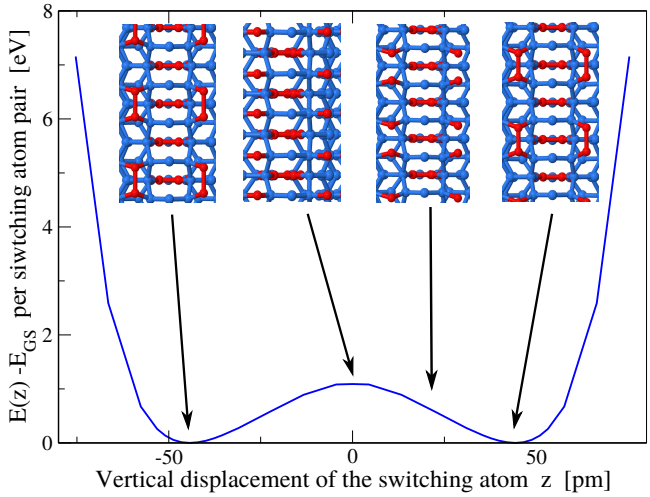


Figure 2.7: The total energy of the single-ring novamene crystal structure as a function of the (fixed) vertical position of the switching atoms, with all other atomic positions allowed to relax. The benzenic ring remains plain. Insets display the atomic configurations corresponding to the minima, the maximum and an intermediate distortion. The minima represent equivalent relaxed configurations with the dimers forming across B-B' or B' -B planes, see Fig. 2.4. The reference coordinate $z = 0$ identifies the high-symmetry non-dimerized condition where switching atoms hybridize exactly sp^2 and remain in the B planes, so that planes A = A' and B = B', and the crystal can be described in terms of a 26-atoms cell.

sp^2 atom dimerized. The second line of Tab. 2.2 reports the values of ΔE defined as:

$$\Delta E = \frac{E_{\text{novamene}}}{N_{\text{novamene}}} - \frac{E_{\text{diamond}}}{N_{\text{diamond}}}, \tag{2.1}$$

where $\mathrm{E}_{\text{novamene}}$ and E_{diamond} are respectively the total energy of the novamene configuration and that of diamond. These energies are divided by the number of atoms of the corresponding primitive cell. ΔE coincides also with the difference in binding energy for the two allotropes. The values of ΔE in Tab. 2.2 indicate that the resulting lowest-energy structure is the one with no sp^2 bonds. This ground state is just 0.203 eV higher in binding energy with respect to diamond. This configuration is likely to be characterized by sharp local energy minimum, which would guarantee its long-term stability. Simulations confirm the intuitive idea that the binding energy decreases with increasing number of formed dimers. Indeed the configuration reported in Fig. 2.4 is recognized as a meta-stable state with a binding energy 0.024 eV per atom above the ground state, or more significantly, 0.208 eV per pair of switching atoms. Instead, the simulations evaluate the total energy of the no-dimer configuration reported in Fig. 2.3 above the ground state as 0.066 eV per atom, or 0.429 eV per pairs of switching bonds (8 dimers can form at most). That is to say, the formation of the dimers involving sp^2 carbons from the benzenic ring has a different energetic cost with respect to the formation of dimers involving the lonely sp^2 carbons. Such a discrepancy in the dimers formation energy arises from the high stability of the benzenic

ring. Indeed the high stability favour the system in Fig.2.4 by lowering its binding energy towards the value of the ground state configuration. As reported in Fig. 2.7 the energy barrier separating the no dimer configuration from the novamene benzenic one, respectively shown in Fig. 2.3 and in Fig. 2.4, is relatively shallow around its ≈ 1 eV maximum value. A similar profile is expected also for the energy barrier of the switching sp^2 carbons forming the benzenic ring, even though with lower values. Increasing the temperature would certainly favor the exploration of this shallow double-well potential profiles, most likely leading to the non-dimerized phase. This energetics suggests a weak Peierls dimerization transition [138]. The electronic character of the different configurations is also reported in Tab. 2.2. DFT simulations identify the ground state configuration as a wide-gap semiconductor, while the novamene benzenic configuration as a small-gap semiconductor. Instead, when no dimer is present a metallic behaviour is detected. The tendency of LDA to overestimate binding and to underestimate energy gaps is evident in Tab. 2.2 in the comparison of the predicted properties of standard allotropes with the experimental values. The rather small difference in total energy among the various configurations suggests the possibility to switch between the ground-state distorted semiconductor to the undistorted metal and back. Such switching may be driven by temperature, light, and/or uniaxial strain. One further investigated possibility consists in an alternating bonding pattern with one of the pairs of the switching atoms binding inside the cell, and the other two atoms binding to like atoms in neighboring cells. The total energy of such a combination has been determined for the configuration with the plain benzenic ring: it is the same within the simulation resolution, and the crystal is a small-gap semiconductor in this other structure too. This suggests how in the actual crystal each individual vertical row of switching atoms can form dimers almost independently of the other rows. As a result, even with perfectly ordered dimerized lines, we can predict a weak lateral correlation (and therefore structural disorder) of these dimerized lines. Given the complexity of the analyzed structures it is rather surprising that from a thermodynamical point of view, at least at 0 K, DFT simulations predict the three single-ring novamene structures to be slightly less stable than graphite and diamond, which are defined by much simpler atomic arrangements.

Electronic properties of the single-ring novamene

For every band structure analyzed in this section, the path in the first Brillouin zone runs over high symmetry points reported in Fig. 2.8. Fig. 2.9 presents the band structure of the relaxed lowest energy configuration. The overall density of states is very much similar to that of hexagonal diamond [139]. These bands describe the energies of σ orbitals characterizing lonsdaleite. Around the Fermi level, the bands are rather flat in all the sampled directions. They retain some dispersion in the $\mathrm{k_z}$ direction especially. These simulations indicate that single ring novamene in its lowest configuration is a semiconductor characterized by a 3.061 eV indirect gap, most likely larger in reality due to the well-known tendency of the LDA to underestimate band gaps. The minimum of the conduction band is at the Γ point while the maximum of the valence band is along the Γ-A line not far away from the center of the first Brillouin zone. Such a gap value could cause single-ring novamene to have an absorption band falling in the shortest wave-length portion of visible spectra, which would led to a yellow color. Nonetheless, LDA has the tendency to underestimate band gap, therefore ground state single-ring novamene is likely to be transparent against visible light, but to have an absorption window in the ultraviolet region. Fig. 2.9, 2.10 and

2.11 reproduce the band structures of the alternative investigated configurations reported in Tab. 2.2. Figure 2.9 illustrates the electronic bands of the lowest-energy structure, reported in the first column of Tab. 2.2 and in Fig. 2.5. This is recognised as a wide-gap semiconductor (3.061 eV).

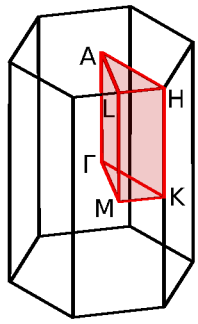


Figure 2.8: The first Brillouin zone of an hexagonal lattice. The irreducible zone is colored with shaded red, the red line show the adopted path connecting the high symmetry points in the region.

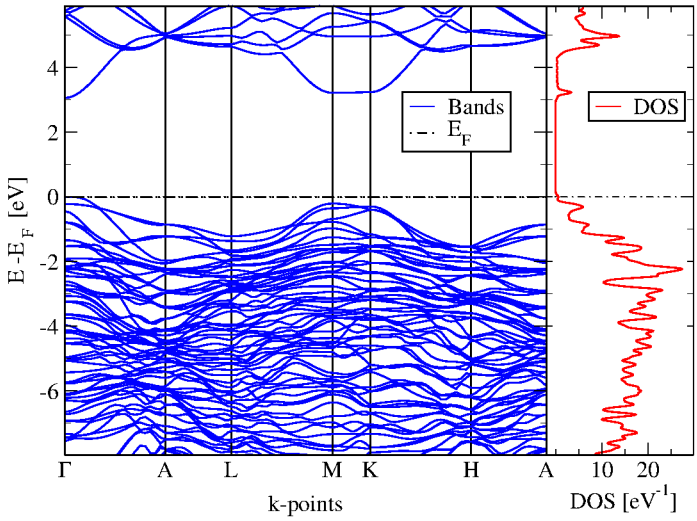


Figure 2.9: The electronic states of ground-state single-ring novamene in an energy region around the top of the valence band (dashed line). Left: the DFT-LDA Kohn-Sham band structure along the path in the first Brillouin zone highlighted in Fig. 2.8. Right: the density of electronic states of these bands.

The overall band structure is moderatly flat: the dispersive character is concentrated along those portion of the path parallel to the k_z direction: namely the Γ-A, L-M, and K-H segments reported in Fig. 2.8. This corresponds to a certain localization around the σ bonds forming the structure, which is a little less intense for those bonds directed along the **c** primitive vector, parallel to the **z** axis. This character is going to be accentuated as dimers break towards sp^2 hybridization. Figure 2.10 illustrates the electronic states of the structure with the benzenic ring undistorted. This is identified as a small-gap semiconductor (0.34 eV) with flat bands surrounding the Fermi

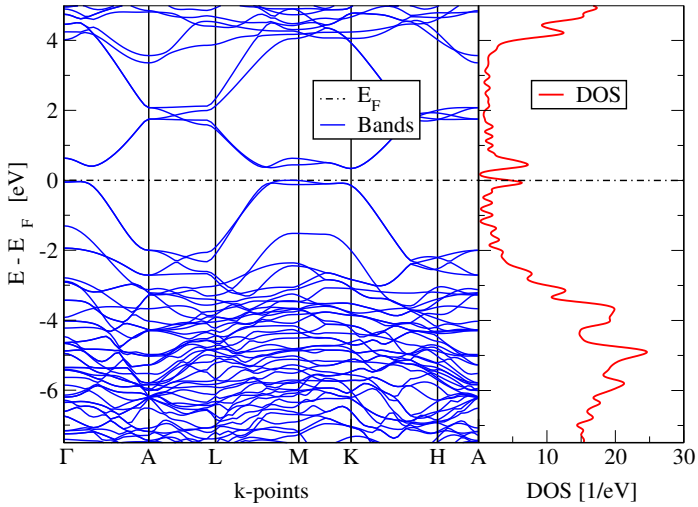


Figure 2.10: The electronic states of single-ring novamene-benzenic configuration in an energy region around the top of the valence band (dashed line). Left: the DFT-LDA Kohn-Sham band structure along the path in the first Brillouin zone highlighted in Fig. 2.8. Right: the density of electronic states of these bands.

level. These bands describe the energies of π orbitals localized mostly on the hexagonal rings. As for the ground state configuration, they are dispersive mostly in the k_z direction, while they are quite flat in the in-plane directions, indicating strong localization around the sp^2 rings, the vertically-dispersed π bonds and "insulation" provided by the surrounding sp^3 carbons. This flatness determines peaks in the density of states near the band edges. The density of state reported on the right panel is analogous to the hexagonal diamond one, except for the bands near the Fermi energy, which move in very close, and remain separated by a quite small gap. This could originate strong light absorption in the infrared immediately above the gap energy. The anisotropy of these novamene structures could provide directional guides for excitons generated by absorbed light in solar-energy harvesting applications [140]. Fig. 2.11 reports the band structure of the no-dimer combination, which is clearly metallic. The bands are particularly dispersive on the vertical portion of the path and rather flat in the horizontal sections. As a consequence, the density of states close the Fermi level diminish significantly towards zero. The bands crosses the Fermi level in the k_z direction, revealing an increasing delocalization of the π orbitals that are localized on the benzenic ring. Even though the proposed allotrope has not been isolated yet, the DFT relaxation provides enough structural information regarding the equilibrium geometry to generate a theoretical X-ray diffraction (XRD) pattern, shown in Fig. 2.12. While the XRD pattern of the proposed single-ring novamene structure exhibits several overlapping occurrences with experimental peaks of diamond, graphite, and lonsdaleite, the larger cell generates numerous additional peaks in between the main peaks of the more conventional allotropes.

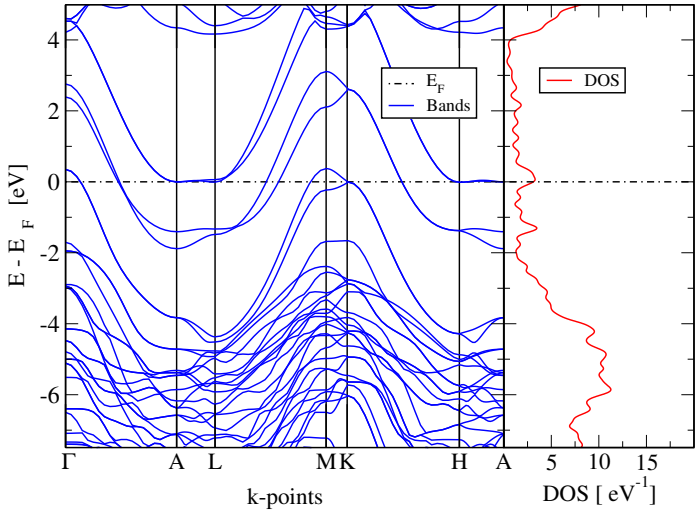


Figure 2.11: The electronic states of the no-dimer single-ring novamene. Left: the DFT-LDA Kohn-Sham band structure along the path in the first Brillouin zone highlighted in Fig. 2.8. Right: the density of electronic states of these bands.

2.4 Further investigations

The results of this chapter have been published in 2017 [120]. Following this publication, single-ring novamene has also been investigated by other authors. Oliveria et al. in 2018 [141] studied its elastic properties and its fracture mechanisms in tensile/elongation regimes by means of *ab-initio* and fully atomistic molecular dynamics simulations, using the ReaxFF force field within LAMMPS [142]. According to their work single-ring novamene presents ultimate strength values around ≈ 100 [GPa], lower than other carbon allotropes (≈ 150 GPa for armchair graphene and ≈ 148 Gpa for diamond), but it has the highest ultimate strain along the z-direction $\approx 22.5\%$ ($\approx 21.5\%$ for armchair graphene and 18.9% for diamond). Its Young's modulus is ≈ 600 GPa to be compared with ≈ 1300 Gpa and ≈ 1270 Gpa for diamond and graphene. Their results indicate a significantly anisotropic behavior, with a recurrent sudden drop of the stress values with no clear evidence of a plastic region, which indicates a brittle material. Moreover, the authors point out as how the allotrope anisotropic behaviour is present also in the fracture dynamics. Indeed the fracture originating from deformed heptagons and pentagons propagates in different way along the x and y directions. The major anisotropy is present along the z direction, characterized by the highest strength before failure. Finally, one interesting feature reported in [141] is the formation of multiple and long carbon linear chains in the final fracture stages. This effect has been theoretically predicted and observed experimentally in other carbon-based nanostructures [137].

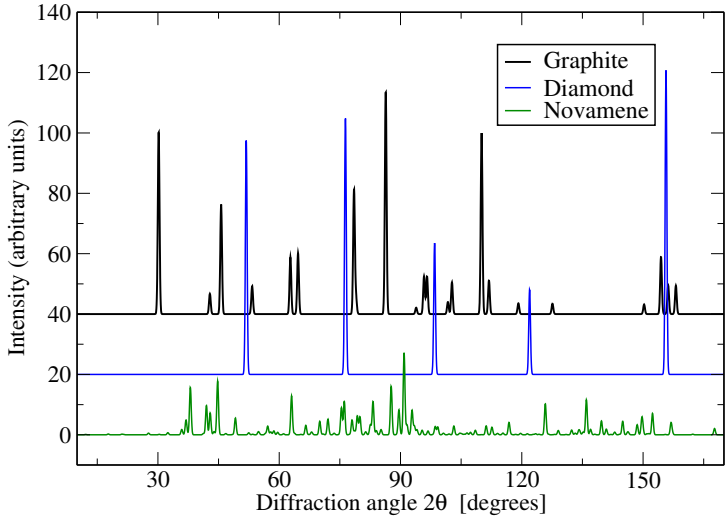


Figure 2.12: Simulated XRD pattern of novamene (green) in its DFT-LDA ground-state configuration, compared to the diffraction patterns of graphite (black) and diamond (blue), shifted upward for better visibility. The patterns are computed for the standard radiation wavelength $\lambda = 154$ pm of the Cu K_α line.

2.5 Protomene

The second allotrope investigated in this thesis is named protomene. It shares some general features with the novamenes of the previous section, with some significant novelties. The underlying idea is analogous: mixing sp^2 and sp^3 hybridization kinds to form a periodic structure. Also in this case the majority of the atoms share a tetrahedral hybridization: 6 atoms out of 24 against 2 out of 26 in novamene, can adopt a perfectly planar sp^2 geometry. In principle, the sp^2 atoms can move out of the plane to build comparably weak bonds with partner atoms in the next vertically stacked lattice cell. Similarly to novamene one can expect this extra bond formation to lower the total energy by approximately 1 eV per bond.

2.6 The crystal structure of protomene

As described for novamene, the structure of protomene has threefold symmetry axes, which are compatible with crystallization in a hexagonal lattice. Starting from an initial conjecture for the structure we adapted it to a proper crystal repetition symmetry, which is reported in Fig. 2.13. This figure identifies the correct minimum primitive cell before any dimer formation: its primitive cell, containing 24 atoms, is sketched by a dashed line. The sp^2 carbon atoms, labelled from 1 to 12, are colored in red, the sp^3 in blue. The group of 4 atoms labeled 1, 3, 5, and 7 sits around the corner of the cell, the two more labeled 9 and 11 isolated within the cell. The atoms labeled with even-number share equivalent positions in the primitive cell below. This can be better understood from panel b of Fig. 2.13, which reports a side cut of the repeated primitive cells. The main novelty with respect to the allotrope class analyzed in the previous section consists in the absence of any benzenic ring. The sp^2 atoms are distributed all over the primitive cell:

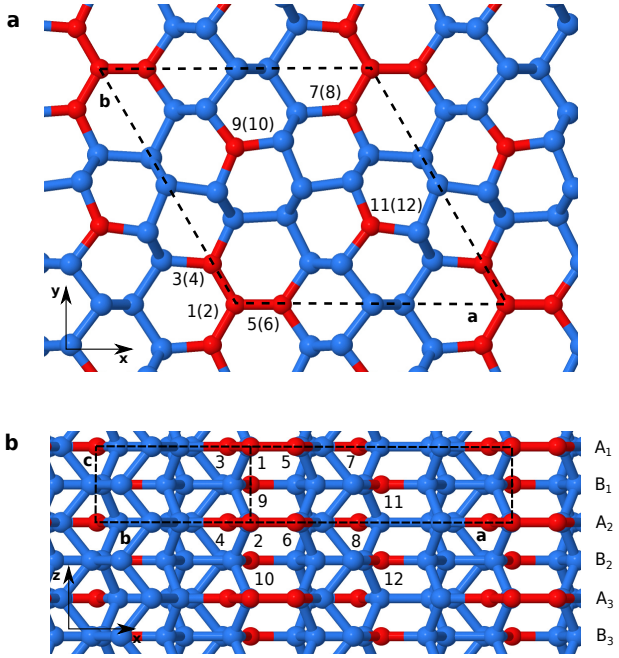


Figure 2.13: A ball-stick model of the structure of protomene, relaxed in its no-dimer configuration. **(a)** A view down the threefold-symmetry $\hat{z}$ axis; **(b)** a side view in the $\hat{y}$ direction. The hexagonal primitive cell is highlighted by black dashed lines and contains 24 atoms. The primitive vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are indicated. The "switching" sp^2 carbon atoms are highlighted in red. In this structure these switching atoms are all in their no-dimer (higher-energy) configuration: atoms $1, 3, 5, 7$ are equivalent to $2, 4, 6, 8$ and they are placed in successive A planes of adjacent cells; atoms $9, 11$ are equivalent to $10, 12$ and they occupy successive B planes, which stand at intermediate heights between A-type planes.

two of them bound with sp^3 carbons, instead the one placed at the edges of the highlighted primitive cell in Fig. 2.13 bond together forming group of four atoms. This arrangement could possibly be the starting point for an extension of this single allotrope towards its specific class. Indeed with some creativity one could imagine to enlarge this sp^2-only areas to more complex shapes, forming new and more intricate configurations. One interesting solution could be to imagine extended sp^2 area, surrounding the primitive cell reported in Fig. 2.13, which should be modified accordingly. Such a class of allotropes would be reminiscent of novamene. The labelled sp^2 atoms in Fig. 2.13 have the possibility to form further bonds with similar atoms in adjacent crystalline planes. Since dimer formation usually leads to energy lowering, among all possible realizations of interplane dimers we expect that the lowest-energy state, the favored state, will be the one with as many interplane dimers as possible. When any number of dimers is formed, successive A or B planes are not equivalent any more, with the result that the $\hat{z}$-directed $\boldsymbol{c}$ primitive vector doubles in length, implying a primitive-cell doubling, as illustrated in the comparison of figs. 2.13 and 2.14. In principle, one could conceive extended patterns of dimer formation in the horizontal xy plane. This would lead to a cell-multiplication to form a supercell with primitive vectors being multiples the $\boldsymbol{a}$ and $\boldsymbol{b}$. Dimers-pattern extended along the $\hat{c}$ direction are also possible. The number of possible patterns of cross-plane dimers, and of those extended in the three-dimensional volume would increase exponentially with the number of non-equivalent cells involved in the pattern. Each of this patterned configurations is higher in energy with respect to the most stable one. Thus, one could imagine to produce each of them, exciting the ground state combination by mean of an external perturbation, such as an electromagnetic field with the suitable frequency. For sake of simplicity, due to the rapid growth of the cell, in this work we do not consider such extended combinations. Instead we consider only all the possible dimer patterns within one horizontal primitive cell, combining two 24-atoms cells stack on top of each other to a full 48-atoms primitive cell. The central pairs (9-10 and 11-12) can form independently from each other. Indeed they are structurally rather distant and the distortion induced by the formation of a single dimer remains localized to its nearest neighbours. As a consequence, the configurations, as in fig. 2.14, with both central dimers formed around plane A_2 of fig. 2.13 should have very similar cohesive energy compared to the structurally non-equivalent configuration with alternate central dimers, namely one dimer formed around the central A_2 plane and two others around the A_1 and A_3 planes in fig. 2.13. On the other hand, vertical displacements of the four pairs of corner atoms are strongly correlated with one another, since these atoms are connected by direct in-plane bonds: the formation of one dimer will drive the formation of all three others. Concerning the corner atoms, it turns out that the most stable configuration is the one with alternating dimers: if atoms 1 and 2 bond together across the B_1 plane, then the neighbouring sp^2 atoms $(3-6)$ will tend to form bonds across the B_0 and B_2 planes of vertically adjacent cells, as shown in panel b of Fig. 2.13. All other possible combinations are likely to be associated with a less favorable cohesive energy. This is due to the tetrahedral conformation of the sp^3 hybridization, which minimizes the total energy of the electronic configuration. Any displacement from this minimum rise the energy of the configuration. Consider, e.g., the configuration with all corner dimers formed across the same B plane: this is a rather unstable combination, involving three 4-atoms rings with highly deformed bond geometries. Having fixed the corner sp^2 carbons, three possible combinations exist for the central dimers. Indeed, one could find alternatively the formation of: no central dimer, two dimers across the same B plane, or two dimers across alternating planes.

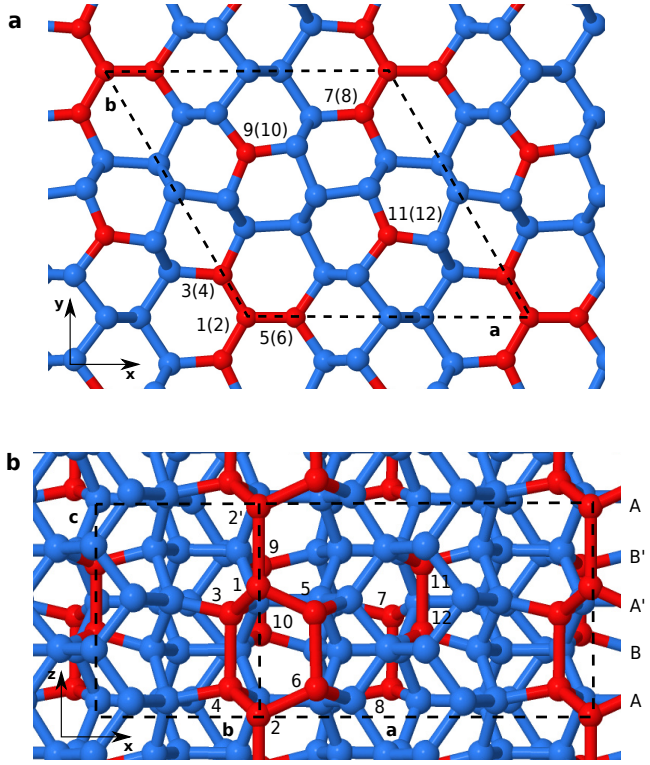


Figure 2.14: The completely relaxed ground-state configuration of protomene, with the energetically most convenient dimerization pattern. **(a)** top view down the $\hat{z}$ direction; **(b)** side view in the $\hat{y}$ direction. The **c** primitive vector is twice as long compared to the no-dimer configuration because A' and B' planes are no longer translationally equivalent to A and B planes. The primitive cell is highlighted with a black dashed line.

The combination with a single central dimer formed is not a stable configuration as is. Indeed the formation of a single dimer breaks one of the lattice symmetries and induces the remaining pair of central sp^2 atoms to bond. This is also valid for the corner dimers. But in that case, apart from the symmetry breaking, local distortion of bonding geometry has a direct role in driving the change in the coordination of the neighbouring sp^2 carbons. In general, the space group describing the symmetries of different configurations will depend on the dimers pattern: specifically, the 48-atom-cell structure displayed in fig. 2.14 is characterized by the space group P-31m (#157).

2.7 Structural and electronic properties

To characterize the different equilibrium configurations and their electronic properties, we employed density functional theory (DFT) simulations, with a well-established approach [100, 132, 137, 143–151]. For each configuration we compared the results obtained within local density approximation (LDA) and within generalized gradient approximation [152, 153] in the Perdew-Burke-Erzoff (PBE) [154] formulation. We employed ultrasoft pseudo-potentials to account for core electrons as implemented in Quantum Espresso [129]. For both functionals the wave functions are expanded on a plain-waves basis with a kinetic-energy cutoff of 408 eV, which guarantees the convergence of the total energy of a simple carbon structure (diamond) within 10^{-3} eV. We sampled the Brillouin zone with a $10 \times 10 \times 20$ Monkhorst-Pack k-point grid [155] for each configuration showing at least one dimer formed. Instead, due to its higher fisrt Brillouin zone, to sample the no-dimer combination we employed a $10 \times 10 \times 30$ Monkhorst-Pack k-point grid. This choices guarantee a mesh whose density is at least 10 points$\times$nm^3 for every analyzed dimer-pattern. We performed simulations to fully relax the nuclei positions allowing the cell primitive vectors to change with a conjugated-gradient method. The starting point is an initial rough configuration built from the analysis of the structure symmetry as described in Sect. 2.6. Atomic relaxation is performed until the following convergence conditions are reached: all components of all forces must be smaller than 4 pN and the total-energy difference between two successive relaxation steps must be smaller than 10^{-3} eV. Tab. 2.3 summarizes the simulation results on the two main relaxed structures of protomene: the no-dimer configuration of fig. 2.13, and the ground state, namely the configuration with the largest possible number of dimers formed, as displayed in fig. 2.14. The same quantities for the case of novamene are shown for comparison. The difference in binding energies relative to diamond is reported in such a way to have an absolute reference of the stability of these allotropes. ΔE is defined by equation 2.1 with the appropriate protomene quantities: Where the reference binding energy of diamond is $E_{\mathrm{b}}^{\mathrm{LDA}} = -8.908$ eV ($E_{\mathrm{b}}^{\mathrm{PBE}} = -8.252$ eV), as computed using the LDA or (parenthesized) the PBE exchange and correlation functional. The LDA-DFT simulations result in a binding energy per atom of -8.708 eV for the ground state and -8.636 eV for the no-dimer configuration, only 2.2% and 3% less stable than diamond respectively, while novamene is 2.6% less stable than diamond [120]. These low values are quite surprising given the much deeper complexity of the reported configurations. As can be seen from Tab. 2.3, the PBE functional corresponds to LDA in predicting the protomene ground state to be slightly more stable than novamene. The ground state configuration is clearly identified by the one with dimers formed by $9-10$ and $11-12$ atoms across the same A' plane (as in fig. 2.14), but the alternative possibility of bonds across alternating $A-A'$ planes is essentially degenerate within 1 meV per atom. Although cross-plane bonding occurs independently,

Structural parameter	protomene no-dimer LDA(PBE)	protomene ground state LDA(PBE)	novamene ground state LDA(PBE)	diamond LDA (PBE)	graphite LDA (PBE)
N_{atoms} per cell	24	48	52	2	4
N_{dimers} corner	–	4	–	–	–
N_{dimers} central	–	2	2	–	–
ΔE per atom [eV]	0.271 (0.188)	0.199 (0.132)	0.227 (0.1350)	0 (0)	−0.146 (−0.318)
band gap [eV]	0.000 (0.000)	3.380 (1.274)	0.336 (0.371)	4.220 (4.445)	0.000 (0.000)
a,b [Å]	8.072 (8.166)	8.074 (8.157)	8.419 (8.510)	3.523 (3.570)	2.433 (2.458)
c [Å]	2.472 (2.509)	4.828 (4.977)	4.998 (5.090)	3.523 (3.570)	5.898 (6.444)
density [kg m^{-3}]	3432 (3303)	3512 (3338)	3381 (3248)	3649 (3504)	2639 (2366)

Table 2.3: The main structural properties of two protomene configurations, the no-dimer one and the dimerized ground state, compared with those of novamene, of fcc diamond and of graphite. The difference in binding energy is computed through equation 2.1, where the reference binding energy of diamond is $E_{\mathrm{b}}^{\mathrm{LDA}} = -8.908$ eV ($E_{\mathrm{b}}^{\mathrm{PBE}} = -8.252$ eV), as computed using the LDA or (parenthesized) the PBE exchange and correlation functional.

Corner dimers	0	0	4	4
Internal dimers	0	2	0	2
N_{atoms}	24	48	48	48
ΔE [eV]	0.271	0.228	0.264	0.199
band gap [eV]	0.000	0.000	0.000	3.380

Table 2.4: The energetical properties (computed with LDA) of the four inequivalent protomene configurations, the no-dimer one and the dimerized ground state compared to two intermediate combination with 2 and 4 dimers formed out of 6. The difference in binding energy is computed through equation 2.1, where the reference binding energy of diamond is $E_{\text{b}}^{\text{LDA}} = -8.908$ eV ($E_{\text{b}}^{\text{PBE}} = -8.252$ eV), as computed using the LDA or (parenthesized) the PBE exchange and correlation functional.

the formation of bonds between B and B' planes favors the formation of bonds of the other kind by reducing slightly the lattice spacing in the $\hat{z}$ direction. Indeed, as reported in Tab. 2.3, c lattice spacing of the ground-state structure is significantly shorter than twice the c spacing of the no-dimer 24-atom-cell structure. Structures with an intermediate number of dimers, as reported in Tab. 2.4, exhibit intermediate values of the total energy. As said, the formation of one single dimer in the corner-group induces the formation of the other. Indeed, the third column in Tab. 2.4 represents the fully relaxed state for all the possible combination among the corner dimers. The DFT simulation provides enough information to evaluate the X-ray powder diffraction pattern (XRD) of protomene in its ground-state configuration: fully relaxed atomic positions and lattice primitive vectors. This pattern is displayed in fig. 2.15, and compared with the XRD patterns of novamene and diamond. This very simple analysis of the XRD pattern is intended as theoretical benchmark for future experimental attempts to isolate and detect protomene samples. These patterns are obtained for an incident X-ray beam with $\lambda = 154$ pm. The diffracted intensity is distributed, as in novamene, among numerous small-intensity peaks, while diamond, with its quite small cell in real space, has widely spaced reciprocal vectors, generating few strong peaks.

Natural tiling analysis

It is interesting to examine the structural properties of the new allotrope with classification techniques based on a network-topology analysis. To this purpose, we used face symbols to describe the so-called natural tilings of protomene. Comprehensive guides for the detailed meaning of these topology concepts can be found in literature, for instance in articles [156–158]. Here we give just a very introductory definition of some basic concepts. In short, a face symbol characterizing a tile, has the form $[A^{\alpha}.B^{\beta}\ldots]$. The exponents α, β represent the number of faces whose perimeter is a ring with A, B sides. According to the cited literature, a *ring* is a cycle that is not the sum of two smaller cycles, while a *strong ring* cannot be decomposed in any number of smaller cycles. Obviously a ring is made by the bonds connecting the atoms of a system. The union face-to-face of different tiles represents the *net* composed by vertices (atoms) and edges (bonds). If several tiles are present together, the natural tiling is the sum of all single face symbols. To count the prevalence of each one of the face symbols in the crystal structure, stoichiometric coefficient

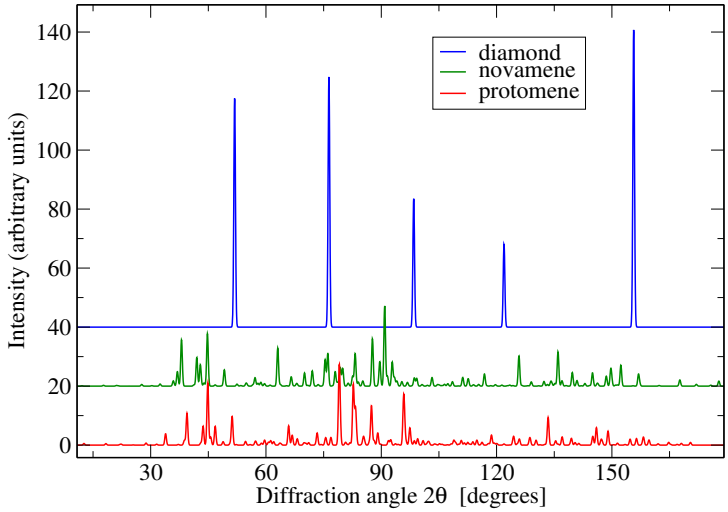


Figure 2.15: Simulated XRD pattern of protomene (red) in its DFT-LDA ground-state configuration, compared to the diffraction patterns of novamene (green) and diamond (blue), shifted upward for better visibility. The patterns are computed for the standard radiation wavelength $\lambda = 154$ pm of the Cu K_α line.

are introduced before the actual description of the tile. For example: $X[A^\alpha.B^\beta] + Y[C^\gamma.D^\delta]$ is the natural tiling of a structure with two different tilings appearing in a $X : Y$ ratio. We made use of the ToposPro software [159] to identify the natural tiling of the protomene ground structure, obtaining:

$$13[6^3] + 3[5^2.6^2] + 6[6^4] + 3[6^2.7^2] + 6[5^2.6^3] + 3[6^4.7^2] + 3[5^2.6^5] + 2[6^9].$$

Fig. 2.16 displays the individual tiles forming protomene. Each color corresponds to a different class of tiles. For instance, the thirteen red tiles in Fig. 2.16 are part of the $13[6^3]$ cage. Since the single tiling characterizing the diamond structure is $[6^4]$, we deduce that protomene includes 15% of diamond cages. On the other hand, we exclude the presence of regions topologically equivalent to hexagonal diamond (lonsdaleite), classified $[6^3] + [6^5]$, due to the lack of $[6^5]$ in the protomene natural tiling. The obtained topological classification can be useful as an example to search this structure in a database of previously published structures, such as SACADA [106, 160]. Since the natural tiling is unique, this search is particolarly useful to make sure that protomene had not been developed before.

Electronic properties

Fig. 2.17 reports the electron energy bands and the corresponding density of states for the ground-state configuration of protomene. As reported in Tab. 2.3, the dimer formation involving all initially sp^2 atoms produces a gap opening in the band structure at the Fermi level. The band structure depicted in fig. 2.17 predicts protomene to be a wide-bandgap semiconductor. The DFT-LDA gap amplitude is estimated at 3.38 eV. However the DFT-PBE simulation estimates

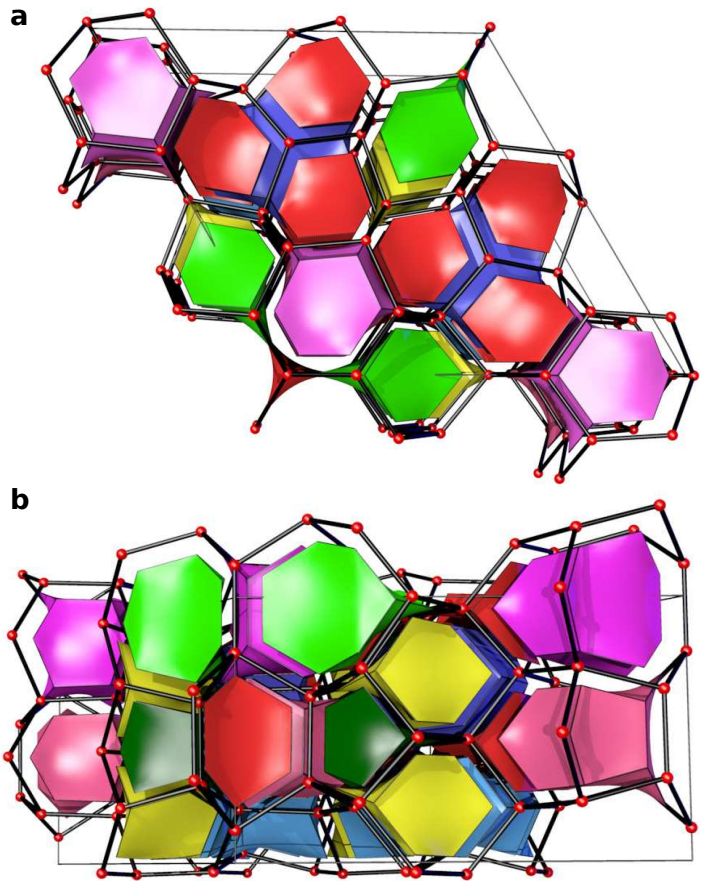


Figure 2.16: Natural tiling for the protomene structure: panel **a** top view down from the $\hat{z}$ direction; panel **b** side view in the $\hat{y}$ direction. Different classes of tiles are labeled by different colors: $13[6^3]$ red, $3[5^2.6^2]$ dark green, $6[6^4]$ yellow, $3[6^2.7^2]$ light-green, $6[5^2.6^3]$ light-blue, $3[6^4.7^2]$ purple, $3[5^2.6^5]$ pink, $2[6^9]$ blue.

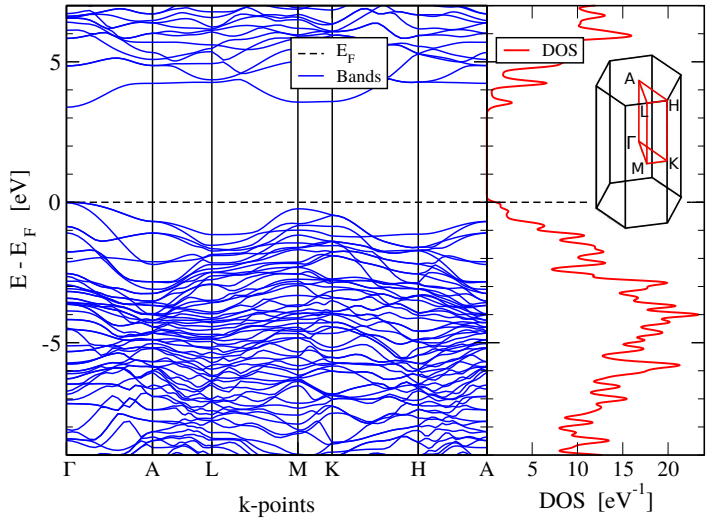


Figure 2.17: The electronic states of protomene in an energy region around the top of the valence band (dashed line). Left: the DFT-LDA Kohn-Sham band structure along the path in the first Brillouin zone highlighted in the inset at the right side. Right: the density of electronic states of these bands.

the same gap at less than half this energy. The reason for this difference lays in the well-known tendency of DFT-LDA to overbind: the LDA computed interplane dimers are probably shorter and more strongly bonded than in reality, thus generating a larger bonding-antibonding separation, which reflects in the protomene band gap. On the other hand, the DFT-LDA method is known to underestimate band gaps. The PBE functional may sometimes capture the electronic band gap slightly more accurate than LDA, but its tendency to underbind determines a quite small bonding-antibonding splitting associated with weaker interplane dimerization. As a result, a gap estimation in the 3 eV region is most likely realistic. One interesting feature of the protomene band structure is the presence of a direct gap, locate at the center of the first Brillouin zone . Despite the discussed uncertainty about the precise amplitude of the protomene gap, it is significantly larger than in novamene ($0.3-0.4$ eV), but clearly smaller than in diamond (where, however, the gap is indirect). The density of electronic states is reported in fig. 2.17. Despite the rather flat nature of certain regions of the band structure, the overall density of states is fairly smooth. The gap between the valence band and the conduction band is the only gap in this energy region, and the density of states changes from relatively small near the gap itself to a rather large few electron-volts away from it, in both the valence and the conduction regions. Fig. 2.18 reports the analogous electronic properties of the no-dimer configuration, where the primitive cell consists of 6 sp^2 hybridized carbon plus 18 sp^3 ones. The path in the first Brillouin zone is the same shown in Fig. 2.17. As reported in Tab. 2.3, the total absence of dimers corresponds to having 6 electrons per primitive cell that are unpaired. As such, this combination is predicted to have a metallic behaviour, as the electronic states crossing the Fermi energy level confirm. Four bands can be distinguished crossing the Fermi level along those portion of the path parallel to the **c** primitive vector, revealing a prevalent metallic behaviour in such a direction. Besides, the high dispersion

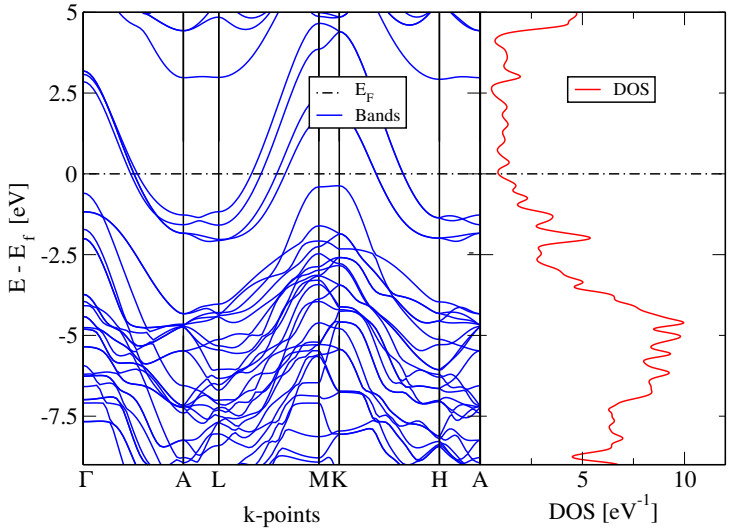


Figure 2.18: The electronic states of metallic protomene in an energy region around the Fermi level (dashed line). Left: the DFT-LDA Kohn-Sham band structure along the path in the first Brillouin zone highlighted in the inset at the right side in fig. 2.17 . Right: the density of electronic states of these bands.

of the bands along those portion of the path parallel to k_z corresponds to a certain delocalization of the corresponding orbitals.

2.8 The vibrational spectrum

In order to simulate the vibrational properties of the ground state configuration of protomene, we make use of density-functional theory in its perturbative expansion (DFPT), as implemented in Quantum Espresso [129, 130]. Besides the cited articles, for detailed information see Appendix B. Basically in DFPT the vibrational frequencies are obtained through diagonalization of the dynamical matrix, connected to the force constants matrix. The construction of such matrices can be reduced to the estimate of the response of the electrons density to a displacement of the ions (the perturbation). This response function is computed within linear response theory employing the perturbed Kohn-Sham orbital and the unperturbed ones. In this analysis the phonons frequencies are computed along the same k-points path as for the electron energy bands. This is a standard path for hexagonal Brillouin zone, as it goes through all the high symmetry point in the irreducible zone. The frequencies are computed separately at each $\vec{k}$ point, in order to use efficiently the available computational resources. The overall 24 points are distributed as follows along the path: 5 in each long segment ($\Gamma \rightarrow$A, L$\rightarrow$M, K$\rightarrow$H), three for the other portions. Fig. 2.19 displays the phonon frequencies along such a path in k-space. The acoustic branches are clearly visible as they depart from zero frequency at Γ; the transverse acoustic branch, the lower of the two, is twofold degenerate along the $\Gamma-$A high-symmetry line. At higher frequency, mostly in the $400 - 1400$ cm^{-1} range, the numerous optical phonon branches cluster and intersect with each other. The optical branches are numerous accordingly to the high number of atoms form-

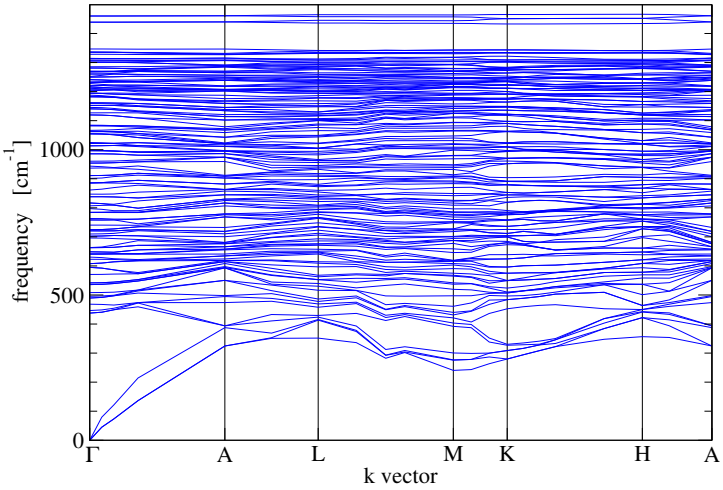


Figure 2.19: Phonon dispersion computed along the same Brillouin-zone path as drawn for the electron bands in Fig. 2.17.

ing the primitive cell of protomene: following the $3N - 3$ rule, N being the number of atoms in the primitive cell, one obtains 141 branches. Near 1500 cm^{-1} and separated from the great majority of phonon branches by a small gap (~ 90 cm^{-1}), we find two optical phonons, each of them twofold-degenerate along the $\Gamma-$A line. We analyze in particular the atomic displacement corresponding to these optical modes close to Γ. Fig. 2.20 reports the eigenvectors of the dynamical matrix computed by means of Quantum Espresso. Each arrow represents the direction of the displacement in selected modes. The length of each arrow is directly proportional to the maximum atomic displacement associated with the selected vibration mode. Note that each arrow is magnified by a factor 10^2 and only those components larger than 0.01 Åare shown. Each panel represents a specific normal mode. Panel **a** and **b** represent the eigenvectors of the 139^{th} and the 140^{th} modes as representative of the optical modes below the gap in Fig.2.19. Panel **c**, **d**, **e**, **f** represent the 4 high-frequency optical modes above the gap. Panel **f** shows one of the acustic modes as a comparison. The 4 high-frequency modes are characterized by localized vibraion as they only involves the dimerized 9 and 11 carbons from picture 2.14 and their nearest neighbours, whereas below the gap the modes involve also different carbons. From the analysis of the eigenvectors one can recognized the modes above the gap as mainly stretching modes of the short nearly-horizontal bonds connecting the atoms labeled 9-11 from Fig. 2.14 forming the central dimers with one of those surrounding them. This feature of the phonon spectrum could be considered as a signature for experimental identification. Overall, the computed vibrational spectrum is consistent with a quite congested Raman spectrum, probably not too much different from the "D band" of amorphous carbon [161]. This band is associated with the irregular (the D stands for disorder) distribution of atoms. It is centered at 1350 cm^{-1}. Such a D peak is not present in perfect crystals, for instance pristine graphite, which is instead characterized by the so called "G" band at approximately $1540-1600$ cm^{-1}. If limitations in the graphite domain shows up, for instance induced by grain boundaries, imperfections, adatoms, then the D-peak would be present with an intesity proportional to the "degree" of disorder.

2.9 Further investigations

The results of this section have been published in 2018 [121]. Starting from this work, Oliveira *et al.* [162] investigated the mechanical properties of protomene and their temperature dependence, by means of fully atomistic reactive (ReaxFF potential) molecular dynamics and DFT (GGA/PBE functional) simulations. Their results suggest that similarly to novamene, protomene is a mechanically anisotropic system against tensile deformations. At room temperature, the predicted ultimate strength is ≈ 100 GPa and the Young's modulus is ≈ 600 Pa. Other carbon allotropes have higher values, but protomene still performs better than other materials, such as some ceramics, silicon, steel under tensile stress. Like for novamene, protomene exhibits the highest ultimate strain along the z-direction $\approx 24.7\%$ and a sudden failure along each Cartesian direction. The temperature dependence of the mechanical properties has been studied from 10 up to 900 K. At the ultimate strength tends to decrease along all the directions by some 10%, especially in the xy plane. Also the ultimate strain is affected by the temperature: the most considerable variations are observed for the z-direction, with a decreasing of 77.0% in its value through the analyzed range. The Young modulus instead is not very sensitive to temperature increase with variation of the order of 10%. The authors point out how the complex deformation mechanism of protomene is associated with changes in the relative number of sp^3 to sp^2 bonds. Indeed, the increase of temperature causes an almost linear increase in the ratio between the number of sp^2 and sp^3 carbons, at a rate of $\approx 0.04\%$ for each 1 K of increase. As shown in the previous sections, these changes have important consequences because they are associated with structural changes, besides inducing a transition from a semiconductor to a metallic structure. The results of Oliveria *et al.* show that this transition could be obtained through temperature variations applied to the system, or by uniaxial mechanical stretching parallel to protomene $\hat{c}$ primitive vector, or through a simultaneous combination of both. Temperature gradients would induce the formation of alternating dimer-non dimer areas within novamene sample.

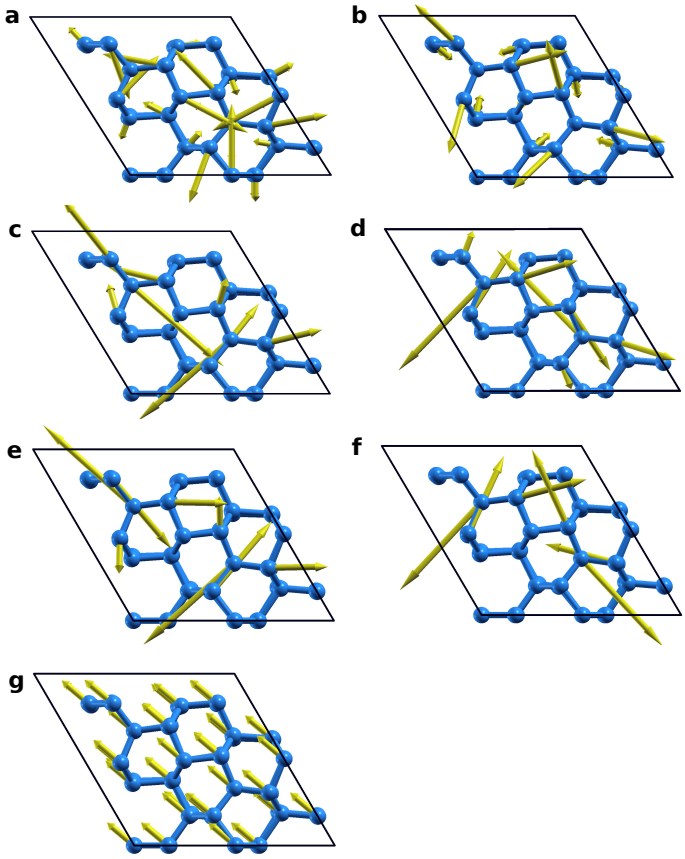


Figure 2.20: Each panel reports the eigenvectors of the dynamical matrix for a specific normal mode, in the limit of long wavelength perturbations. Panel **a** and **b** represent the eigenvectors of the 139^{th} and the 140^{th} modes; Panel **c**, **d**, **e**, **f** represent the 4 high-frequency optical modes above the gap; panel **f** shows one of the acoustic modes as a comparison.

CHAPTER 3

Mixing sp and sp^3 hybridizations

3.1 Zayedene

In this chapter we explore the possibility of combining sp and sp^3 hybridized carbon. Specifically, we consider sp-carbon chains (SPCCs) included in hexagonal diamond cavities. Due to their highly unsaturated configuration, SPCCs tend to react easily with other carbon forming extra bonds, switching from sp to sp^2 or sp^3 hybridization. As such, in order to find a SPCC inside a cavity its should be sufficiently large. On one hand, the cavity should be long enough to accommodate the chain without too much strain and, on the other hand, wide enough to keep it isolated and prevent spontaneous reaction with the atoms on the cage surface and the consequent deterioration of the sp-hybrization. To fabricate such a cavity in bulk carbon, we could equally well consider standard cubic diamond, or any other sp^3 allotrope. Due to its hexagonal structure, the choice falls on lonsdaleite: its **c** primitive vector suggests a natural direction for the cavity long axis. Thus, the analyzed systems are constructed as follow: bulk hexagonal diamond spotted with an array of cavities, each of which holding one or several SPCCs. In principle random distributions of cavities could be present in a sample, but here for simplicity we consider periodic arrays. The irregular case would indicate a defective structure rather then a new carbon allotrope. The name to indicate these allotropes should be "zayedene". This name comes as a tribute to Zayed Bin Sultan al Nahyan, the former sheikh of the United Arab Emirates, where this structure was first imagined. Likewise novamene in section 2.1, it would be more appropriate to define zayedenes as a class of allotropes. Indeed, a large number of periodic cavity distributions can be imagined. One could classify different structures on the basis of the shape of the cavity and as a function of the number and the orientation of the chains. The number of possible combinations scales as a function of this three degrees of freedom. Clearly the simplest system to imagine is based on cylindrical cavities. Fig. 3.1 reports one of the simplest of such structures: the cylindrical cavity have the symmetry axis parallel to the primitive cell **c** vector, with a single SPCC aligned alon this axis. Fig. 3.2 reports more possible structures belonging to this class: each panel but the last one reports a different possible orientation of the chain in the same cylindrical cavity reported in Fig. 3.1. The last panel is a horizontal section of the system; possible non-equivalent bonding sites located at the upper/lower cylinder faces are labelled $\mathrm{S_i}$, the atoms on the cavity axis are labelled C.

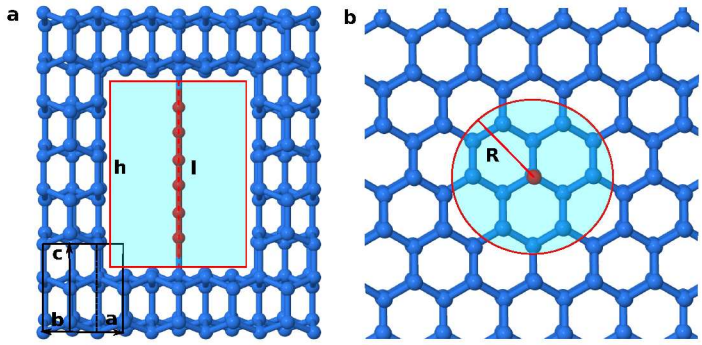


Figure 3.1: An example of the proposed structure combining sp (red) and sp^3 (blue) carbon. **a**: A side cut in the $\hat{x} - \hat{z}$ plane; the primitive cell of bulk lonsdaleite is highlighted in black. **b**: A top cut in the $\hat{x} - \hat{y}$ plane, showing the circular base of the cylinder (radius $r = 3$ Å) used to cut out the lonsdaleite atoms and construct the cavity. The $N = 6$ SPCC binds to carbons of the sp^3 cage where dangling bonds stick out inside the cavity.

3.2 Single SPCC in cylindrical cavities: the model

A periodic arrangement of SPCCs could involve several (possibly different) cavities per unit cell. The number of atoms N_{tot} in the primitive cell increases roughly proportionally to the number of cavity it includes. In such condition, even for simple configurations, such as two cylindrical cavities next to each other, N_{tot} would rapidly approach the computational limits. Therefore, for sake of simplicity, we analyzed a system defined by one single cavity per primitive cell, with a single SPCC included. From a practical point of view, the primitive cell of this polytrope is assembled by repeating the lonsdaleite primitive cell a number of times along its primitive vectors and successively removing the atoms falling inside the cavity volume. With the spirit of considering simple systems, the cavity is chosen to be highly symmetric: a cylinder with the symmetry axis parallel to the **c** primitive vector of lonsdaleite. All the atoms included in this volume are then removed. The axis of the cylinder is taken in a way that the bottom and top C atoms at the centre stick out toward the inside of the cavity. Fig. 3.1 reports the described system: panel **a** shows a side-cut of the structure through the cylindrical cavity. Panel **b** displays its mid-cavity horizontal section. The cavity is identified by the shady light-blue area; the primitive cell of lonsdaleite is defined by the black lines. The SPCC is placed in the most symmetric configuration, coincident with the cylinder axis. In this scheme, the maximum length of a chain included in the cavity with the smallest possible strain is related to the height h of the cylinder. Indeed, a chain much shorter than h would be too strongly strained to bind to both top and bottom sides of the cavity without breaking. On the contrary, a chain significantly longer than h would suffer compressive strain, thus buckling [148] and potentially even recombining with the lonsdaleite cage. Clearly this latter eventuality depends sensitively on the cavity radius and the atomic-scale roughness of the cavity inner surfaces. Table 3.1 resumes the relation among the lengths of the SPCCs and the height of the cage: m is the number of lonsdaleite c units defining the cavity height $h = m \cdot c + b$; where $c = 4.12$ Å is the length of the **c** primitive vector in Fig. 3.1; $b = 1.54$ Å is the bond length in lonsdaleite; N is the number of atoms in the SPCC [163, 164].

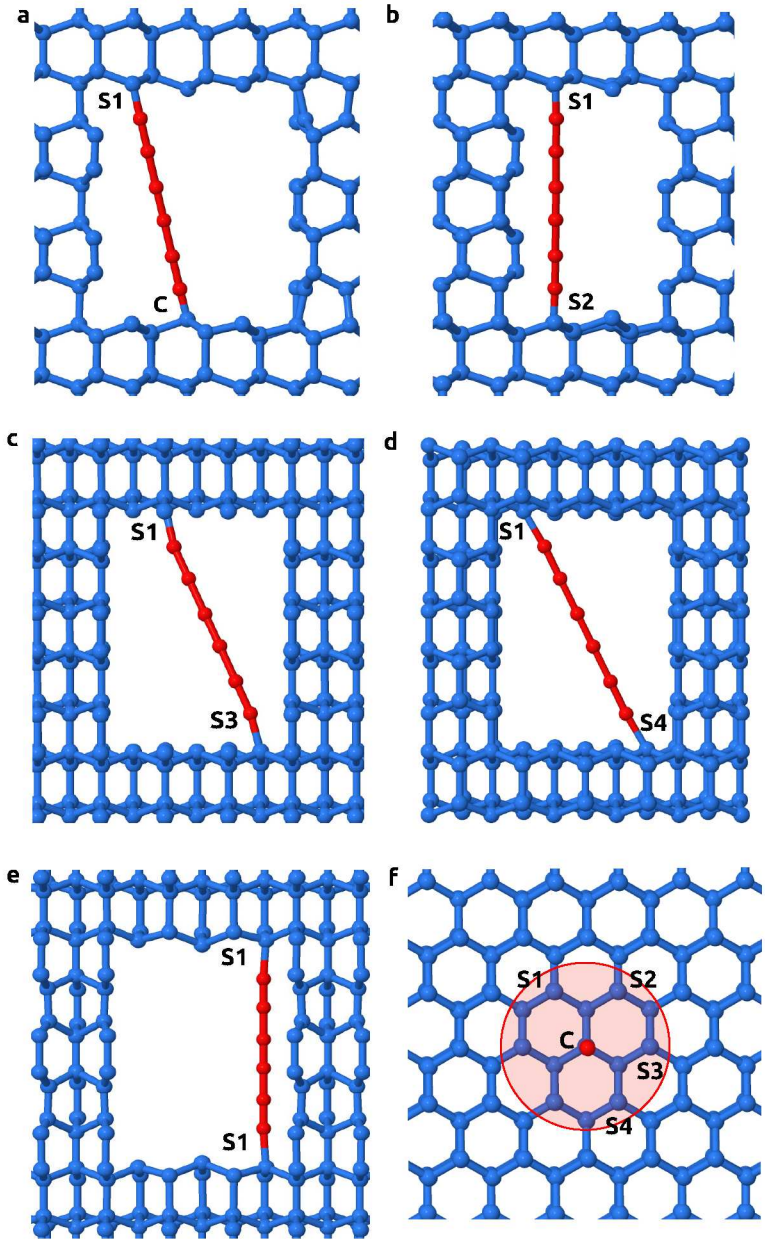


Figure 3.2: Different possible SPCC configurations. The SPCC atoms, are bound to the carbons of the cage that stick into the cavity. Figures from **a** to **e** represent different possible orientation of the chain in the cavity; panel **f** is a horizontal section of the system; the possible non-equivalent bonding sites are labelled with S_i, the atom at the center of the cavity is labelled with C.

m	h [Å]	N	l [Å]	ϵ
1	5.66	4	3.87	46 %
2	9.78	6	8.95	9.3 %
3	13.90	10	14.03	-0.9 %
4	18.02	12	16.57	8.7 %
5	22.14	16	21.65	2.2 %
6	26.26	20	26.73	-1.7 %
7	30.38	22	29.27	3.7 %
8	34.50	26	34.35	0.4 %
9	38.62	28	36.89	4.9 %
10	42.74	32	41.97	1.9 %

Table 3.1: A comparison between the height h Å of the cavity and the unstrained length of the SPCC l Å ; m is the number of lonsdaleite units defining the cavity and N is the number of atoms in the chain, see Fig 3.1; ϵ is the percentage strain, i.e. the difference in lengths normalized to the SPCC length l.

Since the carbons in the cage are sp^3, the extremal atoms of the SPCC are bound with the cage through a single bond. As a consequence, we expect the carbon chain to display a non-negligible bond-length alternation (BLA). That is, to be in a polyynic state. Then, the length of SPCC can be written as $l = N \cdot (b_t + b_s) + b_t$; where $b_t = 1.21$ Å and $b_s = 1.33$ Å are the length of the triple bond and of the single bond [100, 148, 163]. In principle, some qualitative information about the stress state of the SPCC can be extracted by the a comparison of its length l with the significant dimensions of the cavity (e.g. in our case is the length of the cavity in the symmetry axis direction) reported in table 3.1. The 26-atoms chain is the one that fits best its corresponding cage: the strain $\epsilon = (h - l)/l$ is as small as $\approx 0.4\%$. Among the first three configurations, C_6 (6 atom chain) has $\epsilon \approx 9.3\%$, C_{10} has $\epsilon \approx -0.9\%$, C_{12} has $\epsilon \approx 8.7\%$. Then, we can infer that C_6 and C_{12} experience a significant tensile strain whereas C_{10} is subjected to a small compressive strain. The volume of the cylinder is crucial to define the number of atoms N_{tot} in the primitive cell. Such volume $\mathrm{V_{cage}}$ considered as a function of the radius R of the cylinder scales more rapidly than $\mathrm{V_{cyl}} = \pi h R^2$. The reverse would lead to a value of R where holes start to appear in the cage. The thickness of the cage is another central parameter to describe the number of atoms as a function of R. Since a thick cage would make the structure more reminiscent to impure lonsdaleite rather than to a new allotrope, hereby we consider the thinnest meaningful cage. This corresponds to a single vertical slice of bulk lonsdaleite separating two cavities belonging to adjacent primitive cells. This choice is also motivated by the rapid scaling of the number of atoms with the increase of the primitive cell dimensions. The role of the radius R in the simulated model structures can be defined more precisely introducing the quantities m_{c} and m_{a}. They describe the repeated lonsdaleite units in the vertical **c** stacking and in the horizontal xy plane, respectively. To prevent a cavity in one repeated block to get too close (or even worst to merge with) neighboring cavities, the following inequalities must be respected:

$$m_c \geq m + 1 \tag{3.1}$$

$$m_a\, a > 2\, R + t_{\mathrm{cage}}\,, \tag{3.2}$$

where $a = 2.51$ Å is the lonsdaleite lattice spacing in the xy plane; and $t_{\rm cage} = 0.72$ Åis the minimum acceptable thickness of the cage, which is the length of a single C-C bond projected in the horizontal plane. It corresponds to the thickness of a zig-zag shaped portion of lonsdaleite. At the same time, R must exceed several Å in order to allow a longer SPCC to bend or vibrate like a string, without a rapid re-bind with the cavity inner surface. Given these conditions, the model involves a total number of atoms

$$N_{\rm tot} = N + 4\, m_{\rm a}^{\,2}\, m_{\rm c} - N_{\rm cavity}\,, \tag{3.3}$$

enclosed in a periodically repeated cell of volume $m_{\rm a}^2\, m_{\rm c}\, a^2\, c\sqrt{3}/2$, with N being the number of atoms in the SPCC. In Eq. (3.3), the number of atoms removed from the cavity cylindrical volume is

$$N_{\rm cavity} \simeq 8 \times 3^{-1/2}\, \pi\, m\, R^2 a^{-2} c^{-2}\,, \tag{3.4}$$

where the $\simeq$ symbol indicates that this estimation needs to be rounded to a near integer, depending on which atoms near the cylindrical surface are counted in or out. Tab. 3.2 reports the different possible combinations of the quantities defining the different zayedenes along with the number of atoms characterizing the structures. With 4 valence electrons per carbon atom, the total number of electrons per simulation cell amounts to $4\, N_{\rm tot}$.

m_a	m_c	R [Å]	m	N	$N_{\rm tot}$	$N_{\rm cavity}$
4		3			210	52
5		4			330	76
5	3	5	2	6	306	100
6		6			416	160
7		7			552	232
4		3			252	78
5		4			396	114
5	4	5	3	10	360	144
6		6			490	240
7		7			666	348
4		3			292	104
5		4			460	152
5	5	5	4	12	412	200
6		6			556	320
7		7			728	464

Table 3.2: The parameters used to build the nine different configurations. m_a and m_c are the number of repeated lonsdaleite unit cell along the $\mathbf{a}, \mathbf{c}$ primitive vectors; R is the radius of the cavity, m is the number of lonsdaleite c units defining the cavity height, and N is the number of atoms in the SPCC, see Fig. (3.1). $N_{\rm tot}$ is the total number of atoms in the allotrope primitive cell according to Eq. (3.3); $N_{\rm cavity}$ is the number of the removed carbons according to Eq.(3.4).

3.3 Numerical Simulations

Due to the relatively large number of atoms forming the primitive cell, it is worth using a numerical method capable to handle such large systems without being too computationally demanding. *Ab-initio* methods would be the state of the art, given their accuracy in the reproduction of electronic ground-state of molecular and periodic systems. But usually it is required to build and diagonalize a $\mathrm{N_{el}} \times \mathrm{N_{el}}$ Hamiltonian matrix several times during a run, and the state of the art algorithms scales as $\mathrm{N_{el}^3}$, with N being the number of electrons in the system. In these systems $\mathrm{N_{el}} \approx 10^2 - 10^3$. This means that the relaxation of these system could become extremely demanding from a computational point of view. As a consequence, for a qualitative analysis on the energetics of the class of allotropes and its dependence on the size of the cavity and the length of the chain, we employed tight-binding (TB). Whereas to specifically characterize the electronic properties of few specific structures in this class we employed DFT in the version implemented in Siesta [165]. This choice arises for three main reasons. First, tight-binding does not involve by default a self consistent updating of the Hamiltonian; this speeds up significantly the calculations. In this sense TB can be seen as a "quick and dirty" method, not being so far from the exact quantum description of DFT, retaining a wave-function description of the electronic states, but still relying on a previously prepared parametrization of the matrices elements necessary to describe the system. Second, TB exists in several different formulation including one derived as zero-order perturbation to DFT formalism (DFTB) [166], see paragraph 1.2. At least in the formalism this DFTB formulation retains some of the features of DFT such as the existence of some terms in the Hamiltonian related to human's ignorance on those part of the many-body problem representing electronic correlation. Furthermore, since DFTB is a perturbative approach, one could in principle add more and more terms in the expansion until the required precision is reached. The zero order is a pure TB method, the inclusion of the first-order correction aims to adjust the independent-electron picture of TB, etc. Third, TB relies on an externally and previously produced parametrization, which can be targeted to reproduce determined quantities and highly refined.

Validation of the tight-binding model

We employed a TB parameterization which has been reported to reproduce structural and mechanical properties of sp^3 carbon quite accurately [167, 168]. Nontheless, it is worth to validate the force field also on a one-dimensional systems, such as SPCCs which are part of the systems under analysis. For this purpose, we compare the dependence of the total energy of a $\mathrm{C_{12}}$ SPCC isolated in vacuum on the length of its central bond, from a shorter-than equilibrium spacing, until full decomposition into 2 $\mathrm{C_6}$. The SPCC is not interacting with its periodic replica: the $\hat{c}$ vector of the box is much longer than the SPCC length. After detaching the two halves of the $\mathrm{C_{12}}$ around its center, all the atoms are let free to relax to their equilibrium positions apart from the two forming the central bond, whose distance is kept fixed. As a reference, we take DFT-GGA as implemented in SIESTA (Perdew-Burke-Ernzerhof functional [128], DZP basis)[165]. In Fig. 3.3, we compare the resulting *ab-initio* adiabatic potential with the homologous TB curves, as obtained by means of the DFTB+ implementation [169]. The energies are reported in eV and the displacement in pm. The energies are shifted by twice the total energy of a $\mathrm{C_6}$ chain, in this way the 0 eV level in the graph corresponds to the dissociation energy computed with each method. Three curves are

shown: the solid line represents *ab-initio* results; the dashed line corresponds to TB energies; the dot-dashed line reports TB energies computed with second order fluctuation in electron density included in the Hamiltonian, which is now computed self-consistently (SCC) [166]. The aim of including a self-consistent cycle is to improve the independent particle picture typical of TB. In dftb method specifically, the portion of the Hamiltonian which is computed self consistently contains the exchange-correlation terms of the Coulomb interaction. Which means, it take account of our ignorance on the electron correlations. Fig. 3.3 reports a general qualitative agreement between TB and DFT descriptions for this 1D system. Nonetheless few remarkable differences can be identified. Both SCC-TB and pure TB tend to underestimate the bond energy by $\simeq 2$ eV and to overestimate the optimum length for the central bond: $\simeq 1.3$ Å against $\simeq 1.25$ Å from the ab-initio method. Moreover, the predicted bond length alternation (BLA) assumes different values when computed with the different methods. BLA is computed as follows for relaxed SPCCs:

$$BLA = \sum_{i=1}^{(N-1)/2} \left[\frac{d_{2i}}{N_{\text{even}}} - \frac{d_{2i-1}}{N_{\text{odd}}} \right] \tag{3.5}$$

where:

$$d_{2i} = |\vec{r}_{2i+1} - \vec{r}_{2i}| \qquad N_{\text{even}} = \lfloor \frac{N}{2} - 1 \rceil$$

$$d_{2i-1} = |\vec{r}_{2i} - \vec{r}_{2i-1}| \qquad N_{\text{odd}} = \lfloor \frac{N}{2} + 1 \rceil$$

N represents the number of atoms, so $N-1$ is the number of bonds; $\vec{r}_i$ the position of the i-th atom. Then, the BLA is simply the difference in the average length of even/odd alternating bonds. Its value obtained through DFT is 0.1 Å, whereas tight binding BLA is 0.4 Åwhen relaxed with SCC approximation and twice as much without it. This is probably a consequence of the poor performances of the employed TB Hamiltonian when finite-size effects in Peierls distortion appears. It was reported [58, 69] that finite-size effects decrease the degree of alternation in bond lengths for short SPCCs. As a consequence, the BLA is predicted to increase with the number of atoms in the chains, and to be more significant for the central bonds of the chain. In zayedene, the carbon in the cage shows sp^3 hybridization forcing the SPCC to be in a well defined polyyne state. So this incapability of TB to reproduce DFT prediction is not likely to be relevant. Finally, TB fails for high positive strain of the central bond. DFT-GGA give rise to the expected attractive profile in Fig. 3.3, whereas the TB model shows a level crossing to a dissociative state above 225 pm . Indeed, Fig. 3.3 indicates that the asymptotic limit is incorrect. The reason is an unphysical charge transfer between the two C_6 sections which leads, at high separation, to a total energy lower than twice that of a single isolated C_6. SCC corrections smoothen out the level crossing through renormalization of the Mulliken charges. Suppressing the unphysical charge transfer leads to a symmetric charge distribution between the two C_6 and allows for a correct description of the asymptotic energetics. The incorrect dissociative regime above 230 pm remains. Due to the unsaturated carbons, the molecular C_{12} dissociation studied here should be considered as a worst-case scenario. In the calculations we carried out, the bond saturation of all sp carbon prevents the formation of such unphysical charged states, leading to far better reliability of the TB model, even in its no-SCC version, which we adopted throughout. The only significant problem in the adopted TB model stands in its spontaneous tendency to convert the (001) surface layer of lonsdaleite from

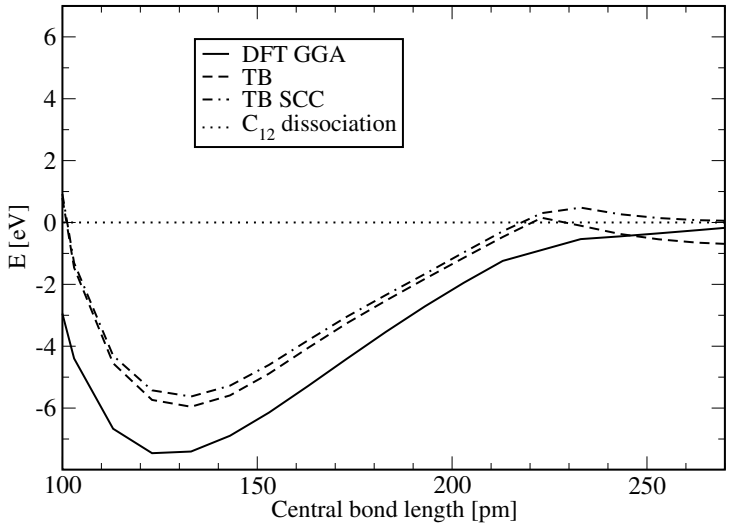


Figure 3.3: Comparison of the binding energy curves of a C_{12} chain $\rightarrow$ 2 C_6 chains, as a function of the length of the central bond (with all other bonds being left free to relax) obtained with the different electronic-structure models considered in the present thesis. Solid line: reference calculation, the accurate DFT-GGA obtained by means of the SIESTA package; dashed line: one-shot TB, no charge self consistency; dot-dashed line: self-consistent charge TB (TB-SCC). For all curves, the zero reference energy (dotted) is taken as twice the energy of an isolated C_6 chain.

sp^3 to sp^2. This can be problematic in case of larger cavities, as will become clear in the next section.

TB energetic analysis: symmetric SPCC orientations

In principle SPCCs with different number of atoms can be found in wide variety of orientations. Hereby we focus on configurations similar to the one reported in Fig. 3.1: straight chains included in cylindrical cavities with the SPCC placed along the symmetry axis of the cylinder. We analyze three different chain lengths, each inserted in five cavities with different radius: fifteen prototype configurations overall. The defining parameters according to the model equations 3.1 are listed in Table 3.2. Once the cavity is defined and the SPCC included, the system is optimized through tight-binding conjugated-gradient (CG) relaxation as implemented in DFTB+ [169]. The maximum force component has to be lower than $5 \cdot 10^{-6}$ eV/Å before convergence is achieved. Both atomic positions and primitive vector are optimized. Table 3.3 summarizes the main results of these relaxations. First, consider the difference in the total adiabatic energy per atom of each structure, relative to that of perfect bulk lonsdaleite, defined as:

$$\Delta E = \frac{E_{\rm tot}}{N_{\rm tot}} - \frac{E_{\rm bulk}}{N_{\rm bulk}}\,. \tag{3.6}$$

The differences in binding energy range from 0.3 to $\approx$ 0.5 eV per atom, apparently with a visible tendency to increase as the size of the cavity increases, involving more cut bonds. The non-monotonic trend of ΔE as a function of $N_{\rm surf}$ (the number of atoms forming the internal cavity

surface) is due to the different cage thickness characterizing the different combinations. Indeed as one can infer from table 3.2, a larger cavity radius does not necessarily require a corresponding increase in the lateral side of the supercell represented by $\mathrm{m_a}$. If it does not, then the cage will be thinner, the imagines of the cavity closer to each other and the overall system is less reminiscent of a defective bulk system, which would be more stable than any element in this types of structure. This is basically the reason why the difference in binding energy of those configurations having $R = 5$ Åis systematically larger than to the others. This result suggests that the dominant factor

N	R [Å]	a [Å]	c [Å]	ρ [kg/m^3]	ΔE [eV]	N_{surf}	E_{cost} [eV]	E_{gain} [eV]
	3	10.06	16.58	2884.0	0.35	38	3.64	3.70
	4	12.56	16.63	2917.4	0.33	50	3.81	3.23
6	5	12.584	16.590	2682.4	0.42	60	3.86	3.91
	6	15.051	16.611	2539.8	0.40	86	3.00	-3.53
	7	17.594	16.553	2480.9	0.42	110	3.62	-15.97
	3	10.07	20.82	2749.7	0.40	56	3.22	3.15
	4	12.54	20.84	2783.6	0.37	68	3.69	2.67
10	5	12.596	20.810	2511.0	0.48	84	3.82	3.48
	6	15.069	20.815	2387.4	0.44	101	3.87	-3.81
	7	17.493	20.836	2375.6	0.46	128	4.18	-7.74
	3	10.07	24.84	2669.5	0.43	74	3.03	2.71
	4	12.57	24.9	2691.2	0.40	98	3.26	2.26
12	5	12.606	24.861	2401.7	0.51	108	3.46	1.39
	6	15.000	25.009	2274.8	0.45	146	3.19	1.45
	7	17.687	25.001	2142.9	0.43	182	3.27	-8.88

Table 3.3: The energetics of the relaxed configurations: N fixes the length of the C_N chain, R is the cavity radius, and the other structural parameters are as listed in Table 3.2. ρ is the average density of the allotrope, systematically smaller than that of pure lonsdaleite, which is $\rho \simeq 3548.9$ kg/m^3 within the same DFTB model, to be compared with the experimental value $\rho \simeq 3510$ kg/m^3. ΔE is the difference in total energy per atom between the considered structure and bulk lonsdaleite. E_{cost}, Eq. (3.7), is the energy required to cut the N_{surf} bonds, thus creating the cylindrical cavity. E_{gain} is on half of the energy variation associated to the chain bonding to the inner cavity walls.

in the lower stability of these allotropes is the energy cost

$$E_{\mathrm{cost}} = \frac{E_{\mathrm{cage}} + E_{\mathrm{cyl}} - E_{\mathrm{bulk}}}{N_{\mathrm{surf}}} \tag{3.7}$$

required to cut the bonds and dig the appropriate cylindrical cavity out of bulk lonsdaleite. The three total adiabatic energies at the right hand side of Eq. (3.7), have the following meaning: E_{cage} is the energy of the lonsdaleite cage surrounding the cavity without SPCC; E_{cyl} is the energy of the cylindrical portion of lonsdaleite that has been removed to create the cavity; and E_{bulk} is the energy of a portion of bulk lonsdaleite with the number of atoms $N_{\mathrm{bulk}} = N_{\mathrm{cage}} + N_{\mathrm{cyl}}$. All these energies refer to fully-relaxed configurations. N_{surf}, also reported in Table 3.3, is the number of cut bonds at the surface of the cylinder. Table 3.3 reports this energy per cut bond, which, as

binding sites	$C-S1$	$S1-S2$	$S1-S3$	$S1-S4$	$S1-S1$
δE [meV]	0.0	3.3	-11.9	-15.6	-8.3

Table 3.4: The difference in the total adiabatic energy per atom of each structure, relative to the C-C structure with both sides of the SPCC attached to the central C binding sites, as in Fig. 3.1. $\delta E = (E_S - E_{C-C})/N$, where the $\mathrm{S_i}$ labels identify the relevant pair of binding sites, see Fig. 3.4.

expected, is approximately constant as a function of the cavity size. Of course the number of cut bonds $N_{\rm surf}$ increases as the cavity size is increased, resulting in a progressive deterioration of the overall stability of the allotrope, observed in the increasing excess energy ΔE above. On the other hand, the bonds of the SPCC with the cage are beneficial to the overall stability. We quantify this stabilizing contribution with

$$E_{\rm gain} = \frac{E_{\rm cage} + E_{\rm free\ C_N} - E_{\rm tot}}{2}\,, \tag{3.8}$$

namely the bonding energy (per bond) that is gained when attaching the SPCC inside the cavity. Here $E_{\rm free\ C_N}$ is the energy of an unsaturated C_N SPCC isolated in vacuum. The quantity $E_{\rm gain}$ depends on the stress state of the SPCCs, arising from the strain ϵ reported in Table 3.1. Indeed, according to that table, the C_6 and C_{12} SPCCs are strained by $\sim 9\%$, while for C_{10} buckles under a $\sim -1\%$ compressive strain. In this comparison, large strains ϵ decrease the bonding energy of the chain, as occurs for the C_6 and C_{12} configurations, due to the additional elastic-deformation contribution. The unexpected dependence of $E_{\rm gain}$ on the cavity radius R occurs as a consequence of the tendency of the surface layers of hexagonal carbon to graphitize and correspondingly bulge out inside the cavity, thus pressing the SPCC into buckling. However, recent research [170] reported lonsdaleite to possess a remarkable graphitization resistance mechanism which is partially responsible for its high hardness. Therefore the graphitization characterizing the larger cavities is an non-physical feature of the adopted TB model. The values obtained for the larger cavities are therefore strongly biased by this feature.

3.3.1 Other structures: shifted and rotated SPCC

Each of the 15 prototypes analyzed in the previous section include a SPCC that is initially parallel to the **c** of lonsdaleite along the cylinder axis. The configurations reported in Fig. 3.2 loose this symmetry. This possibility introduces extra degrees of freedom in the definition of the class elements: whether the SPCC is simply shifted or also rotated, this group consists of an extension of the previous class. We carried out full relaxation in TB also for these systems, with the same level of theory employed in section 3.3. Panels a-e of Figure 3.4 report side views of the resulting relaxed configurations, labelled according to the initial binding sites. Table 3.4 reports the difference δE in energy per atom, relative to the center-center (C-C) configuration described in the first line of tables 3.2 and 3.3. The arrangements reported in Fig. 3.4 a,b , $S1-C$ and $S1-S2$, are energetically almost equivalent to the $C-C$ one. In these configurations a consistent section of the SPCC maintains its sp-character. The other three arrangements ($S1-S3$, $S1-S4$ and the vertical $S1-S2$) are energetically favorable. As can be see in the corresponding panels,

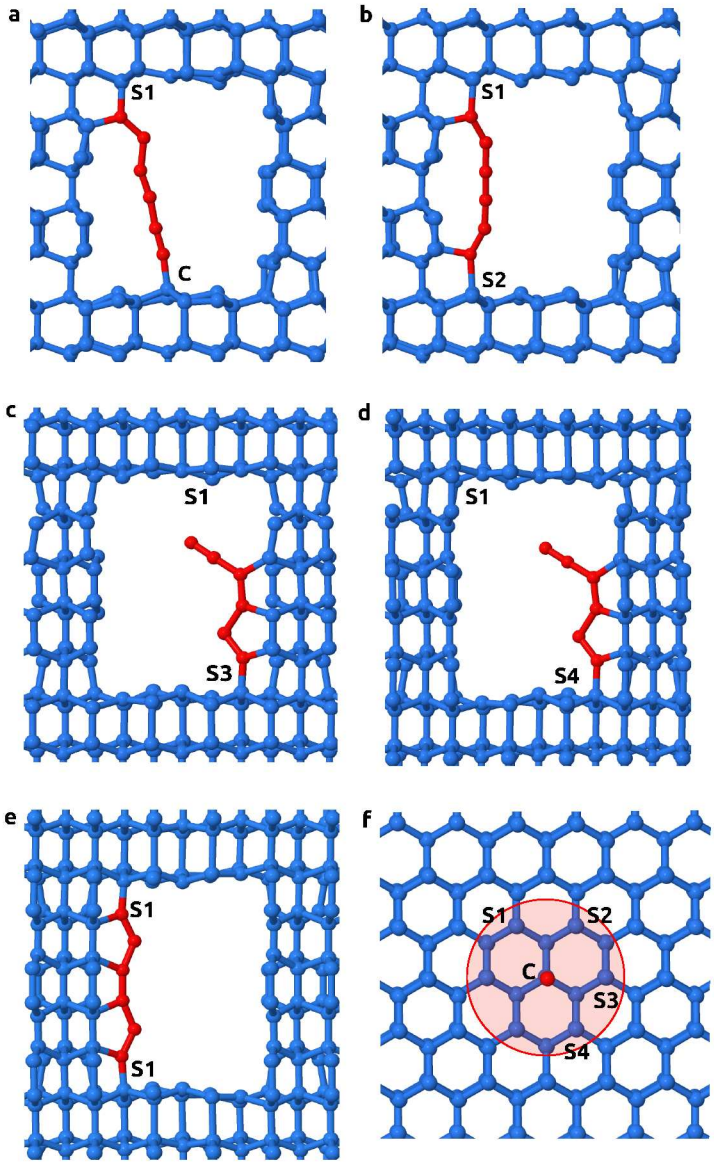


Figure 3.4: A few extra possible relaxed arrangements of the C_6 SPCC in a lonsdaleite-based cage obtained by means of total-energy relaxations starting from the initial conditions of Fig.3.2. Panels **a-e** report TB-relaxed configurations.

this is obviously a consequence of the reconstruction of the chain with the cavity side. Such a reconstruction doesn't produce proper lonsdaleite. Indeed frustrated arrangements including pentagonal carbon rings originates from the TB relaxation. The presence of these frustrated atomic dispositions prevent a further lowering in the adiabatic energy. Generally speaking, the energy gain increases proportionally with the number of atoms rebinding properly to lonsdaleite. The main message of these relaxations is that SPCCs are indeed unstable against re-conversion to sp^2 or sp^3 carbon. As a consequence, in order to prevent this re-conversion, one needs to place the SPCCs near the central axis in cavities of sufficiently large lateral size.

3.4 High-temperature stability

From the analysis of Sect. 3.3, one can conclude that the SPCCs bound to the inside surface of a suitably sized cavity in lonsdaleite correspond to local minima of the adiabatic potential energy of carbon, therefore to metastable allotropes. Given a sufficiently long time, these metastable allotropes would eventually reconvert to one of the more stable ones. Clearly, one interesting issue is to understand whether a sample of this allotropes would remain stable at ordinary pressure and temperature conditions. Likely, the decomposition of the *sp*-chain would be the first step in this re-conversion process. The conversion of sp carbon into sp^3 carbon, and concluded by an intricate sequence of crystallization steps. Besides, it would be interesting to understand if the three-dimensional cavity surrounding the SPCC could limit its high reactivity. Indeed *sp* bonding easily tends to interact with external, e.g. atmospheric, reactant which enables carbon to switch to a more stable (saturated) electronic configuration [58]. Specifically, oxygen atoms are responsible for SPCCs gradual destruction whereas helium, hydrogen and nitrogen do not chemically interact with the *sp* chains [59]. The reported lifetime of SPCCs range from few seconds [171] to several hours [59]. These wide range depends both on the thermodynamic conditions under which the experiments are carried out and on the nature of the system analyzed. We focus here on the thermal stability of the simplest configuration among those described in the previous sections: a C_6 included in a cylindrical cavity having a R $= 3\mathring{A}$, and being coincident with the cavity symmetry axis. The number of atoms in the primitive cell is 210 including the SPCC. The room-T stability is studied following a well-established method. The kinetics of the SPCC breaking reaction at high temperature and standard pressure is simulated by means of tight-binding molecular dynamics (TBMD). From these high-T MD simulations one can extrapolate the average lifetime of the chain under standard conditions: direct investigation of properties at room temperature would be computationally overwhelming. The molecular dynamics is performed using the Velocity-Verlet algorithm as implemented in DFTB+ [169], with the force field given by the TB Hamiltonian (TBMD, see [172]).

3.4.1 Equilibration

To equilibrate the system we employed the Nose-Hoover thermostat [173] and the Berendsen barostat [174]. In the Nose-Hoover thermostat method the external bath is represented as one virtual particle coupled with the system. The coupling is defined in such a way to correct the kinetic energy and therefore the temperature during the evolution toward the target temperature. This method becomes more effective if one introduces chains of particles instead of a single

one to represent the bath. This method take care of the limit in the one-particle formulation which predicts canonical distribution only for ergodic systems [175]. In the present work, the coupling factor and the length of the chain representing the thermostat are chosen through a variational approach as those values that minimize the fluctuations in T. Similarly to the Nose-Hoover case, in the Berendsen barostat method the target pressure is obtained by coupling the system to an external bath. The internal virial is modified to adjust the pressure to the needed value; that is to say, an appropriate correcting term is added to the equation of motion describing particles position. The coupling factor is directly proportional to the isothermal compressibility of the system and is chosen to minimize pressure fluctuations. The system is initially equilibrated at 1500 K and 1 bar; the simulations run for 5 ps plus $\approx$ 17 ps for temperature and pressure equilibration respectively, on a total number of $\approx 11 \cdot 10^4$ MD steps, with time step of $\tau = 0.5$ fs. From this run we selected 7 groups of 10 uncorrelated configurations each. The weak correlation is is guaranteed by the time-distance separating any of these two configurations, longer than the decay-time of the velocity auto-correlation function computed from the equilibration at 1500 K. This decay-time is evaluated to $\simeq 0.02$ ps. Each configurations group is then equilibrated to a different temperature: $\mathrm{T} = 1500, 1800, 2100, 2400, 2550, 2700$ K . The seventh group is equilibrated to $\mathrm{T} = 3000$ K with the aid of an intermediate equilibration step at 2550 K. This intermediate step is necessary to prevent possible breaking of the SPCC caused by the rapid temperature variation.

3.4.2 MD and analysis

In order to prevent the Berendsen barostat to artificially displace SPCC atoms in unstable positions closer to the breaking condition of the chain, the average lifetime of the SPCC is computed in the microcanonical fixed-volume (NVE) ensemble. The choice of the energy-conserving ensemble is acceptable since the size of the system allows the temperature to be well-defined against fluctuations, that scales as the inverse of the number of particles. The average lifetime is defined as the time that one has to wait on average in order for the population to be reduced to one third of its initial value. This definition accounts for any breaking reaction of the chain, no matter what kind of failure mechanism drives it. The SPCC-breaking event is detected by means of a marker l which represents the length of the longest bond in the chain:

$$l = \max_{i=0\ldots N} |\mathbf{r}_i - \mathbf{r}_{i+1}| \ . \tag{3.9}$$

In this definition, for $i = 1, \ldots, N$, the $\mathbf{r}_i$ is the position of the i-th atom of the C_N SPCC. $\mathbf{r}_0$ and $\mathbf{r}_{N+1}$ represent the positions of the two cage atoms bound to SPCC extrema. At first, the SPCC is defined to be broken when the marker l overcomes a threshold chosen to be 2.1 Å, which is above the breaking point of the dimer suggested by the adopted TB parametrization, and is approximately where such parametrization fails, see Fig. 3.3. Fig. 3.5 displays the evolution of this marker at $T = 2100$ K as a function of the simulation time for four out of the ten independent configuration employed to extrapolate the average lifetime. As reported by Fig. 3.6, this definition leads to a non-monotonic trend of the unbroken chains population over time. The presence of the upward steps induces some uncertainty during the fit procedure which propagates through the extrapolation procedure resulting in an unphysical decay rate. In order to reduce the sensitivity to these instantaneous fluctuations, the istantaneous value of l has been replaced by its average $\langle l \rangle$

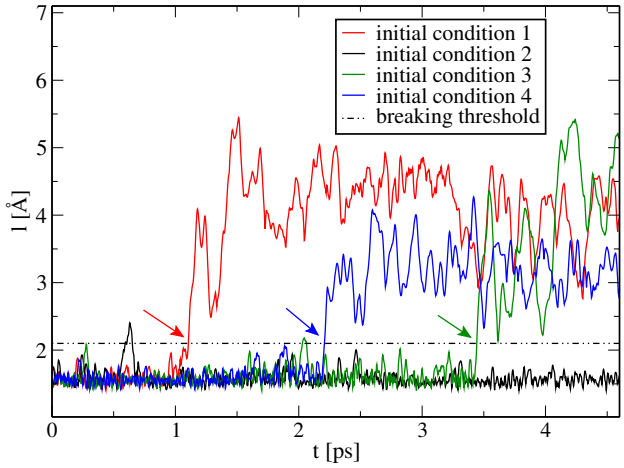


Figure 3.5: Four examples (out of 10) of the time dependence of the longest SPCC bond length l, defined in Eq. (3.9). Each curve reports an independent microcanonical simulation carried out starting with uncorrelated snapshots previously equilibrated at $T = 2100$ K. Arrows indicate the definitive rupture times.

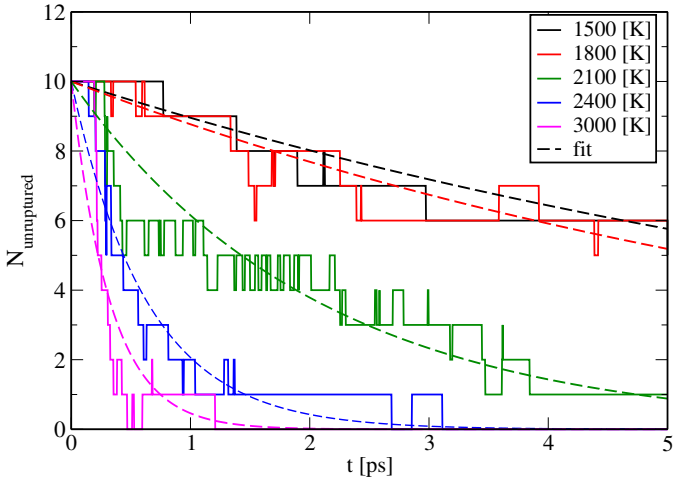


Figure 3.6: Solid lines: The time dependence of the population of unbroken SPCCs obtained executing, for each temperature T, 10 different numerical simulations starting from independent initial conditions. For sake of clarity, 5 populations out of the 7 considered are reported. A downward step is marked each time a SPCC breaks. Dashed curves: the exponential fit of each population decay.

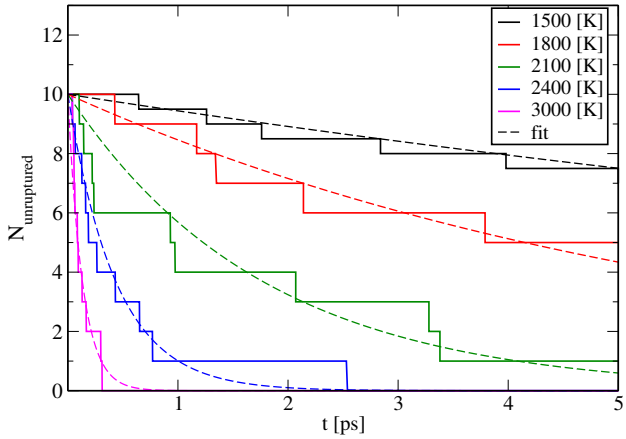


Figure 3.7: Solid lines: the time dependence of the population of unbroken SPCCs obtained executing, for each temperature T, 10 independent numerical simulations starting from uncorrelated initial conditions. For sake of clarity, 5 populations out of the 7 considered are reported. The $T = 1500$ K case involves 20 independent simulations, for improved accuracy. The downward steps in every curve correspond to a rupture of the SPCC, described by the time-average of the quantity l defined in eq. 3.9. Dashed curves: the exponential fit of each population decay.

over a 100 fs time interval. Therefore we define the SPCC to be broken when this time-average $\langle l \rangle$ exceeds 2.1 Å. Also in this case this threshold corresponds approximately to the validity range of the adopted TB model and to the C-C bond inflection energy, see section 3.3 and in particular Fig. 3.3. The breaking time is taken as the first time step for which the $\langle l \rangle > 2.1$ Åcondition is verified. Fig. 3.5 illustrates the evolution of the longest bond length as a function of the simulation time for four out of the ten simulations carried out for $T = 2100$ K. Note that, in Fig. 3.5, the spike of the black curve crossing the threshold now is not recognized as a rupture event, thanks to averaging l: for the black-curve sample, the rupture is recognized to occur at a later time, not visible in this figure. At a fixed T, each simulation run independently from the others, therefore the population of unbroken SPCC is expected to decay exponentially with time. With such a new definition of breaking, the resulting population of entire chains is represented by a stepwise decreasing function,Fig. 3.7, starting from the initial number of independent samples, and decreasing by one each time a SPCC breaks. This stepwise decreasing function can be fit with an exponential function: $N_{\mathrm{unruptured}} = N_0 \exp(-t/\tau)$ (dashed curves in Fig. 3.7). Fig. 3.7 reports how higher temperature simulations exhibit a more rapid decay of the unbroken SPCC population compared to the colder ones. 5 ps TBMD simulation runs are sufficient to evaluate this decay for all considered temperatures, except the lowest one, $T = 1500$ K, where, for a better accuracy we both triple the time duration to 15 ps, and double the number of independent samples to 20 (reporting the properly normalized population in Fig. 3.7). In this way, all simulations last for a time longer than or comparable with the relevant decay time τ. Fig. 3.7 shows these decays and fits for the seven temperatures considered. According to Arrhenius formula, the temperature dependence of the decay rate τ^{-1} is described by logarithmic equation:

$$\log(\tau^{-1}) = \log(A) - \frac{E_{\mathrm{b}}}{k_{\mathrm{B}} T} = a - \frac{T_{\mathrm{b}}}{T} \,. \tag{3.10}$$

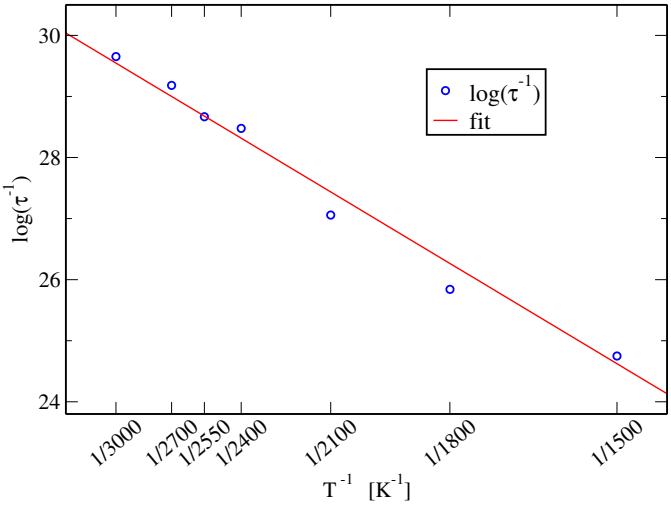


Figure 3.8: Points: the inverse lifetime τ^{-1} of the SPCC obtained from the fits of the decaying population of unruptured chains, as a function of the inverse temperature. The error bars are too small ($\simeq 0.01\%$) to be visible on the reported scale. Solid line: an Arrehenis fit of the lifetimes.

coeff	value	standard deviation	std percentage
$a = \log(A)$ [log(s^{-1})]	34.471	± 0.006	0.2%
$T_b = E_b/k_B$ [K]	14771	± 121	0.8%

Table 3.5: Best-fit values and standard deviation σ on the coefficients in the Arrhenius Eq. (3.10). The last column reports the relative uncertainties. The fit correlation coefficient $C_{a,T_b} = 0.972$.

Fig. 3.8 reports the decay rates τ^{-1} resulting from exponential fits of the stepwise decaying populations with errorbars representing the fit uncertainties. A weighted linear fit interpolates through these data, with the weights being the inverse of the squared uncertainties on τ^{-1} obtained from the population fits. The fit has been performed with the Scipy python package. Table 3.5 reports the resulting fitting coefficients with their standard deviations and percentage errors. The aim of this fit is to obtain some quantitative prediction for the energy barrier $E_\mathrm{b} = k_\mathrm{B} T_\mathrm{b}$ against SPCC breaking, and the thermal attempt rate A for this process. The resulting energy barrier is $E_b = (1.273 \pm 0.001)$ eV. By means of an extrapolation to 300 K, one obtains the desired estimate for the room-temperature decay rate. Inserting the obtained parameters in the Arrhenius equation one finds:

$$k_\mathrm{room} = A e^{[-\frac{E_\mathrm{b}}{k_B T_\mathrm{room}}]} \simeq (3.86 \pm 0.35) \times 10^{-7}\ \mathrm{s}^{-1}. \tag{3.11}$$

which corresponds to an average life-time τ of the order of 1 month. Following the interpretation of the Arrhenius equation we predict that in an hypothetical sample of zayedene, of this or some analogous class, the number of entire SPCC would undergo breaking reaction with a rate expressed by eq. 3.11. That is to say, the number of entire SPCC would be reduced approximately to one third in time corresponding to τ. In section 3.3 we described the limit of the employed TB parametrization, specifically the failure in the description of the long range interactions among

the carbon atoms in *sp* systems. It is straightforward that the numerical results reported in the current section are not to be intended as quantitative estimates of the the properties of these allotrope. From fig. 3.3 the barrier against breaking appears to be lower in the TB model compared to the more accurate DFT. Therefore a longer room-temperature lifetime is to be expected in reality compared to the rest of our analysis. From a semi-quantitative point of view, our estimation agrees well with the observation of the slow decay of the fraction of sp carbon and its reconversion into nanostructured sp^2-sp^3 carbon: the SPCC fraction decays steadily over the course of months, if kept at room temperature under vacuum or in an inert-gas atmosphere [65, 148, 176–180].

3.5 Vibrational density of states

Interesting information can be extracted from the correlation function of specific dynamical or static quantities of a system. For instance, the fluctuation-dissipation theorem connects the response of a system to external perturbation to microscopical fluctuations that characterizes its equilibrium state. The Green-Kubo relations [181] describe how such responses are relaxed towards equilibrium connecting transport coefficients to time-integrated correlation function of suitable dynamical quantities. The dynamical structure factor is a clear example of such connections. In this case we are interested in studying the vibration frequency spectrum of our configuration, because the vibration frequencies of the SPCCs can be used as a fingerprint to experimentally detect their presence when not directly detectable, for instance with Raman spectroscopy [58]. It can be shown [182, 183] that the density of vibration frequencies can be extracted from simulation trajectories through the following equation:

$$\rho(\omega) = \frac{1}{3N_{\rm at}k_{\rm B}T}\int_{-\infty}^{\infty}\frac{\sum_{i=1}^{N_{\rm at}}\langle \mathbf{v}_i(t_0)\cdot\mathbf{v}_i(t_0+t)\rangle}{\sum_{i=1}^{N_{\rm at}}\langle \mathbf{v}_i(t_0)\cdot\mathbf{v}_i(t_0)\rangle}\, e^{i\,\omega\,t}\,dt\,, \tag{3.12}$$

We computed the correlations among velocities from a 1 ps long TBMD simulation, using as a starting point previously equilibrated runsat $T = 30$ K. The duration of the simulations is chosen to give a resolution in the frequencies of the order of $\Delta\nu \simeq 1/t_{max} \simeq 1$ THz. The trajectories are recorded with a time interval sufficiently short to safely cover even the highest possible vibrational frequencies Given that the highest optical frequencies in Diamond are $\simeq 1300$ cm^{-1}, and those of SPCC are reported around $1800 - 2300$ cm^{-1} [58] we adopt $\tau = 0.5$ fs. Then, the maximum frequency sampled will be $\nu_{max} \simeq 1/\tau \simeq 6\cdot 10^4$ cm^{-1}, safely exceeding the expected frequency ranges. Fig. 3.9 reports the density of states of frequencies averaged over 10 different simulations at $T = 30$ K. In this way different regions of the configuration space are sampled to excite uniformly the normal modes. The atomic initial positions in every simulation are displaced adding random Gaussian displacement from the equilibrated configurations, with $\sigma = 0.05$ Å. It is important to have small displacements around equilibrium position in such a way to remain in the validity limit of the harmonic approximation. If not, the density of vibration frequency would include anharmonic effects, described by phonon interactions (in quantum mechanics language: three body annihilations or decays, at the pure anharmonic order) with corresponding variation in the phonons DOS profiles. Fig. 3.9 reports the density of normal-mode frequencies for three different structures: in red the zayedene of the first row of Table 3.2 repeated $2\times2\times2$ times along its primitive vectors, in blue the corresponding lonsdaleite cage with the empty cylindrical cavities,

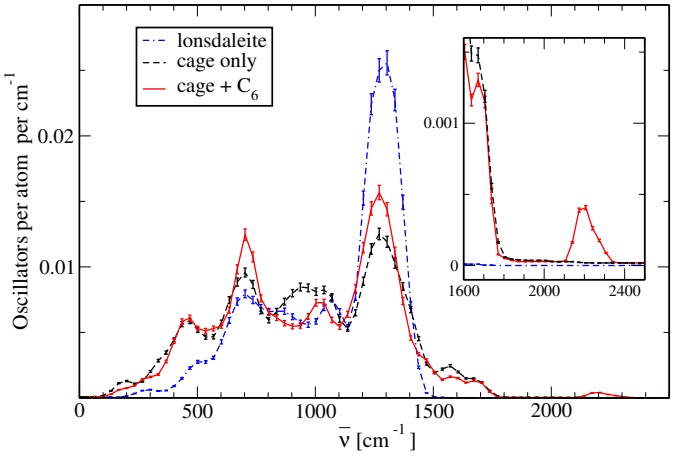


Figure 3.9: The phonon densities of states extrapolated from the velocity auto-correlation function. The area is normalized to 3, which is the number of oscillators per atom. According to the literature the group of high frequencies around $2200[\ \mathrm{cm}^{-1}]$ represents the modes localized mainly on the chain.

and in black pure lonsdaleite represented by a $7\times7\times8$ repetition of its primitive cell. The profile of the blue and the red curves is qualitatively similar. The main difference consists in the small peak around $2200\,\mathrm{cm}^{-1}$ which is missing in the DOS profile from the configurations without the SPCC. We therefore point it as a fingerprint for the presence of the chain in the cavity. This could be directly verified with an analysis of the modes corresponding to the frequencies in the peak, but this is going to be part of future developments. In this thesis we simply observe that this explanation is in agreement with what reported in the literature [58, 184–186]. The effect of the cavity is to shift the vibrational frequencies down to smaller values compared to pure lonsdaleite. This effect is more evident for the long wave-length modes which become softer with respect to those of pure hexagonal carbon. This correspond to a loss of rigidity that can be expected for bulk systems when spotted by an array of cavities.

3.6 Electronic properties

This section illustrates the electronic properties of the simplest possible configuration among those described in section 3.2, specifically in Tab. 3.1: C_6 chain included in its corresponding cylindrical cage with radius $R = 3$ Å. The resulting complete primitive cell is made by 210 atoms, including the 6 *sp* atoms. The symmetry group is trigonal while for bulk lonsdaleite is hexagonal. Specifically, as obtained through Findsym tool [187] the group is number 156, or P3m1 in the Hermann-Mauguin notation, which consists in 6 symmetry operations.

3.6.1 TB vs DFT description

The state of the art for the analysis of electronic band structures is density functional theory. Nonetheless, due to its preferable “agility”, TB retains some utility in this sense, specifically when intricate systems are addressed, see appendix C.1. For coherence reason with the methods

employed so far for the analysis of zayedene, and due to the high number of electron present in the system, we considered the possibility to employ TB to compute the band structure of zayedene. The feasibility of this route critically depends on the performances of TB parametrization to reproduce electronic properties. Therefore, it is indispensable to compare the bands obtained with ab-initio and TB simulations for a reference simple system. For sake of simplicity this reference is going to be pure lonsdaleite.

Siesta and localized basis

The possibility of obtain a poor description of electronic bands with TB and the consequent need for DFT, make necessary to understand which DFT implementation employ. Given the high number of atoms involved in zayedene and the presence of the cavity, using a plan-wave based code such as Quantum Espresso could turn into cumbersome simulations. Indeed, an accurate expansion of the electronic wave functions localized on the internal cavity surface would require a high number of wave vectors. Instead, a suitable choice consists in employing a localized atomic-orbital based code, such as Siesta, see appendix A.2. In principle at the same parallelization level, localized-basis should be less computationally demanding with respect to plane-waves basis, giving qualitative comparable results [165]. Several options for the orbital basis are given within the framework of Siesta, with different level of complexity. It's worth analyzing how the different basis perform with respect to plane-waves DFT. In this way, one has a benchmark for the level of precision obtained by using a certain basis. One possible way, is a variational study to establish which basis best approximate plane-waves DFT results for a specific quantity. For instance we chose the shorter lattice spacing a of pure carbon lonsdaleite. We focus on two standard pseudo-atomic orbitals (PAO) bases as implemented in Siesta: DZP and SZP. SZP employs a single ζ function to approximate the radial part of the orbital, adding polarization orbitals constructed from perturbation theory. The DZP basis employs two ζ functions plus polarization orbitals. Normally, both basis are subjected to an orbital-confining cutoff radius. In the Siesta implementation, this radius can be determined defining the excitation energy of the PAO given to the confinement to a finite-range. Fig 3.10 reports the total energy per primitive cell of lonsdaleite as a function of the lattice step a for SZP and DZP with different confinement energy values. For each choice of the basis, self-consistent simulations with fixed cell and fixed atomic positions are performed for different lonsdaleite a lattice-spacing. The coordinates of the atoms in the primitive cell are scaled accordingly. A 3x3x4 Monkhorst-Pack k-point mesh is employed. For the exchange correlation energy, we employed GGA through the Perdew-Burke-Ernzerhof functional [128]. As shown in Fig.3.10, the DZP basis with a confinement energy of 272 meV (green line) minimizes the distance of the minimum from to the plane-wave result obtained with Quantum Espresso [129] with a converged kinetic-energy cutoff (100 Ry), shown in red.

Lonsdaleite band structure

Fig .3.11 reports a comparison between the band structure of pure lonsdaleite computed with TB and DFT as implemented in Siesta with DZP basis with 272 meV as confinement energy. The inadequacy of TB parametrization employed in the previous sections to reproduce electronic properties of carbon stands out. Indeed, while the occupied Kohn-Sham states are reproduced with an overall qualitative good agreement by the two methods, the energies of the empty states

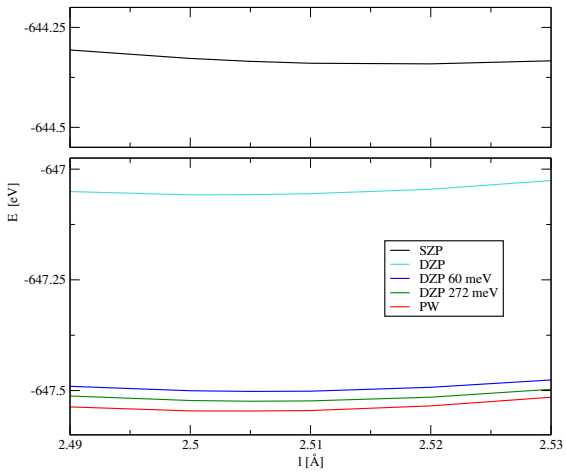


Figure 3.10: The total energy per primitive cell of bulk lonsdaleite as a function of its lattice spacing a. The different lines correspond to different localized basis. The red line correspond to the results obtained with Quantum Espresso with a highly converged kinetic energy cutoff.

are largely unsatisfactory. For instance, TB predicts a direct band gap located at the center of the first Brillouin zone, whereas DFT predicts an indirect gap, with the maximum of the valence band in Γ and the minimum of conducting band in K. Further more, TB by far overestimates the ab-initio value: 6.6 eV against 3.4 eV predicted by Siesta, which is smaller than the corresponding value for cubic diamond, as it should be [139, 163]. Therefore cheap TB description has to be abandoned in favour of more reliable ab-initio method.

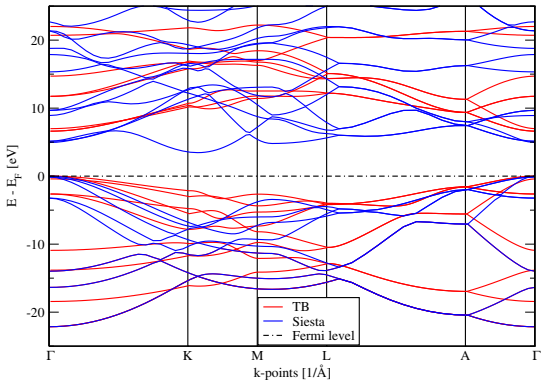


Figure 3.11: The band structure of the pure hexagonal diamond along the path reported on the abscissa axis. The Fermi level is represented by the black dashed line. The blue bands are computed with Siesta, the red with TB. The qualitative agreement of filled bands can be appreciated, whereas empty TB bands dot not coincide with DFT ones.

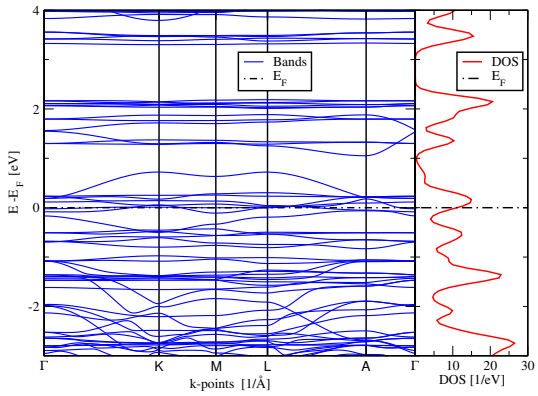


Figure 3.12: The band structure of the C_6 zayedene evaluated along the path reported in Fig. 2.8. The flat bands around E_F (black dot-dashed line) are given mainly by the dangling bonds inside the cavity, with little SPCC involvement.

3.6.2 Bands, DOS and PDOS

Once TB has been discarded and the best basis kind for DFT calculations is identified, see section 3.6.1, zayedene structure is relaxed by means of CG algorithm as implemented in Siesta [165] . The primitive vectors and the atomic positions are both allowed to vary. The forces are converged below 10^{-4} eV/Å , with a 3x3x4 Monkhorst-Pack k-point mesh. For the exchange correlation energy, we consider a generalized gradient approximation (GGA), the Perdew-Burke-Ernzerhof functional [128]. Bands and DOS are computed for the completely relaxed structure. The left panel in Fig. 3.12 reports the electronic band structure obtained on the path shown in Fig. 2.8. The right panel in Fig. 3.12 reports the corresponding DOS. The metallic character of the allotrope is evident in both figures. Several flat bands are present around the Fermi level, arising from the dangling bonds inside the cage. Specifically, a more careful analysis reveals that two bands cross the Fermi level. To verify whether also the SPCC contributes to the metallicity of the structure or not, we evaluate the projected density of states (PDOS) on the SPCC and on the atomic orbitals of the cage. The PDOS $g_\mu(E)\,dE$ counts the number of Kohn-Sham states having energy in the range E and $E + dE$, weighted by their overlap with a given atomic orbital μ:

$$g_\mu(E) = \frac{1}{N_{\mathbf{k}}} \sum_{i}^{\text{bands}} \sum_{\mathbf{k}} \sum_{\nu} c^*_{\nu,i}(\mathbf{k}) c_{\mu,i}(\mathbf{k}) S_{\nu,\mu}(\mathbf{k}) \delta(E - E_i(\mathbf{k}))\,. \tag{3.13}$$

Here $c_{\mu,\mathrm{i}}(\mathbf{k})$ is the coefficient of the projection of the $\psi_\mathrm{i}(\mathbf{k})$ Kohn-Sham wave function with eigenenergy $E_\mathrm{i}(\mathbf{k})$ on the μ-th atomic orbital; $S_{\nu,\mu}(\mathbf{k})$ is the superposition matrix of the atomic basis. We replace the delta function with a Gaussian with width $\sigma = 0.15$ eV. For the DOS calculations we employ a $15 \times 15 \times 15$ grid of $\mathbf{k}$ points, after verifying that convergence has been achieved. The DOS of a specific set α of atoms is then simply the sum of the PDOS relative to every orbital of every atom in that set: $g_\alpha(E) = \sum_{\mu \in \alpha} g_\mu(E)$. For instance, if the set α coincides

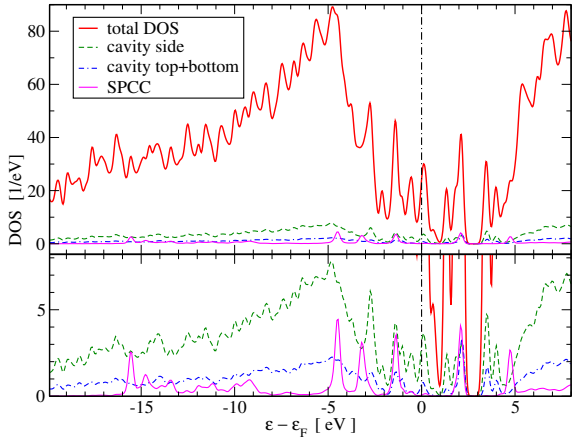


Figure 3.13: The total DOS of C_6 zayedene with the partial contributions arising from the SPCC and the cage. Those coming from the horizontal portion and the side of the cage are highlighted with different colors. The Fermi level is represented by the black dot-dashed line. The inset clarifies that the metallicity is given by the dangling bonds, not by the SPCC.

with the entire primitive cell, then $g_\alpha(E)$ coincides with the total DOS $g(E)$. Given the small fractional number of atoms in the SPCC, their contribution to the total DOS is relatively small. In particular, around the Fermi level E_F there is little or no contribution of the SPCC orbitals. On the contrary, the contributions of the atoms forming the cavity surface are dominant around E_F. This decomposition proves that the metallic behavior of this allotrope arises due to the cage dangling bonds, and not because of the presence of the SPCC. Accordingly, even though these results hold for this specific zayedene, if one computed the band structure of any similar crystal, one would find a qualitatively similar situation: extremely flat bands crossing the Fermi level associated to the cavity dangling bonds.

CHAPTER 4

Conclusions

In this thesis we focused on the development and the characterization of three specific exotic carbon allotropes. Two out of the three allotropes analyzed are the simplest elements in potentially rich classes of analogous structures. The fundamental concept of this thesis is the possibility of having periodic structures made of pure carbon where different bonds hybridization coexists. Such a possibility is not a-priori excluded by nature, indeed amorphous carbon is spontaneously formed (coal for instance) where sp^2-sp^3 bonding are simultaneously present with variable percentages. Nonetheless, apart from amorphous carbon, this is not so common: peculiar allotropes such as glassy carbon, have been recognized to be formed just by sp^2 bonding [110]. On the other hand, given their high complexity, the synthesis of these specific polytropes would require a technological level which is probably beyond the current state-of-the-art laboratory techniques.

The first class introduced, named novamene, is made of sp^3 carbon doped with sp^2 with specific benzenic ring substructures. The variety of the arrangements of these substructures determines the complexity of the class. Probably novamene with extended sp^2 areas enclosed inside sp^3 have higher possibility of being synthesized. By ab-initio studies we prove the relative stability of the simplest of these compounds, single-ring novamene, and its possibility to undergo a metal-semiconductor transition when external perturbation are applied, whether they are electromagnetic, mechanical, or involving temperature or pressure. The structural information (space group, lattice parameters, and atomic positions) provided the necessary starting ground for the successive investigations of the mechanical properties of the structure. The main challenge ahead is the synthesis of actual crystals of novamene, and their experimental characterization. Recently, J. Narayan, and A. Bhaumik [188] synthesized a novel phase of carbon by nanosecond laser melting and quenching carbon from the super under-cooled state in the form of thin films or filaments. They called this material "Q-carbon", and showed that it has an amorphous structure. The Raman spectrum of Q-carbon exhibits a very large fraction of sp^3 ($75\% - 85\%$) bonded carbon from the relative intensity of the diamond peak at 1333 cm^{-1} and the sp^2-related peaks at 1140 cm^{-1} and 1580 cm^{-1}. Even though this is pure speculation, this single-ring novamene has an sp^3 fraction (77%) that is comparable. Following this intriguing conjecture, one could imagine that a detailed structural analysis could prove that Q-carbon is a disordered arrangement containing a considerable percentage of single-ring novamene, plus hexagonal carbon, plus possibly multiple-ringed novamene. After all, also amorphous carbon is not purely amorphous but it is spotted by locally ordered atomic arrangements dispersed in a matrix of disordered carbons. A further step will be the exploration of phonon spectra and thermodynamical stability as a function of pressure and temperature. Besides, deeper electronic characterization of more complex elements in the

novamene class could be interesting. Also electrical transport properties and the optical properties of these compounds should be investigated in depth. In principle, in addition to carbon, novamene-type structures could be relevant even for silicon and germanium.

The second allotrope investigated is a variation of novamene without the benzenic rings. In analogy with novamene we name it protomene. Also in this case we characterized the basic energetic and electronic properties of the allotrope. From DFT, protomene qualifies as a new potentially-useful direct-gap semiconductor, with band gap close to well known semiconductors such as GaN [189]. As a consequence, protomene may possess similar semiconducting properties, which would enable it to have applications for high-power, high-temperature and high-frequency electronic devices with high breakdown voltages. Due to the gap amplitude near the blue end of the visible spectrum, protomene may find applications in opto-electronic components, e.g. to blue- or UV-light generating light emitting diodes, or as a UV filter in optics. Furthermore, the well defined direct gap suggests that potentially protomene may have more applications in semiconductors devices than carbon nanotubes (CNT) and graphene in terms of energy band gap. Indeed, one obstacle to overcome in fabrication of CNTs is controlling whether the CNT is metallic and semiconducting. Protomene instead is clearly semiconducting up to a transition temperature. One relevant aspect to consider is that thermal expansion in protomene is likely to play against the inter-plane bonding. On the contrary, it would favour the no-dimer metallic configuration. Therefore, as temperature is raised, a structural phase transition is likely to occur from the low-temperature semiconducting 48-atoms cell structure of fig. 2.14 to the high-temperature metallic phase characterized by the 24-atoms cell structure of fig. 2.13. As this transition is approached, the band gap would close rapidly, much faster than the slow decay due to thermal expansion in diamond and silicon. Accordingly, this phase transition would provide a sensitive temperature-controlled optical filter and electric-switching applications. Moreover, given the increasing number of sp^2 bonds, this phase transition would undermine the hardness of protomene in its insulating phase [162]. Therefore, protomene, as much as novamene, is expected to show somewhat lowering hardness as the temperature is raised. The gap in the phonon optical frequencies is to be considered as a fingerprint of this allotrope, detectable through Raman spectroscopy or neutron scattering. Besides, this gap could lead to protomene potential applications in surface-acoustic-wave based devices, acoustic filtering or mirroring. The metal-semiconductor transition induced for instance by rising the temperature is a common feature of the two allotropes analyzed in Chapter 2 and it is purely dictated by the presence of the sp^2 "switching atoms". Considering its potential applications, this is probably the most intriguing feature of the analyzed allotropes.

Finally, the zayedene allotrope class consists in a mixture of sp and sp^3 hybridized carbon. sp hybridization of carbon creates linear acetylenic chains, which in these allotropes are included in surrounding sp^3 cages. The arbitrariness in the definition of the cavity and the surrounding cage, on the number and on the orientation of the SPCCs makes this class of allotropes potentially rich. One further degree of freedom is the choice of the sp^3 system forming the cage: in principle we could define these allotropes digging cavities in any sp^3 bulk arrangements, hexagonal or cubic diamond etc. Using a TB model we analyze from a qualitative point of view the structural and binding properties of several different combinations of SPCCs with different lengths and cylindrical cavities with different radius. The analysis of high-temperature kinetics of the chain-breaking reaction for the simplest structure allows to predict the room-temperature lifetime of the order of months, similar to what is observed experimentally for sp-sp^2 combinations [58]. This is probably

a lower limit to the thermodynamic stability of zayedene. Through TBMD we identify a high-frequency peak in the vibration spectrum around 2200 cm^{-1} which is recognized as the fingerprint of the SPCC. Such a peak should be visible in IR or Raman spectroscopy. Its intensity should also allow one to estimate the concentration of cavities with SPCCs in the sample. The metallic character of the simplest allotrope is revealed by accurate ab-initio simulations. The obtained flat band sub-structures are associated with the dangling bonds of the cavity surface whereas the SPCC orbitals do not contribute significantly to the metallic character. The presence of very narrow bands at the Fermi level, suggests that zayedene structures, if synthesized, could provide an interesting playground for materials science, where the interplay of electron-electron correlation, electron-phonon coupling, and disorder may lead to several competing metallic, insulating, and possibly even superconductive states, in a potentially rich phase diagram. For these reasons, it would be quite interesting to synthesize and characterize one or more of these allotropes in the lab. This same strategy of inserting SPCCs in suitable cavities could turn useful even outside the carbon-only concept. More general scenarios worth pursuing as well, include the insertion or growth of SPCCs inside cavities e.g. in silicon or germanium clathrates [190–192] which could be the subject of further investigation.

APPENDIX A

A.1 The many-body problem

The behaviour of electrons, ions and of all constituents of matter is described by complex objects called wave-functions, hereby labelled with the Greek letter Ψ. The degree of complexity of these objects greatly depends on the nature of the system. Indeed, their shape as long as their time evolution is defined by the Hamiltonian operator $\hat{H}$ describing the system through the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\ \Psi \tag{A.1}$$

in such a way that, generally speaking, Ψ is a function of all the degrees of freedom of the system. For instance, calling $\mathbf{x}_j = (\mathbf{r}_j, \sigma_j)$ a vector whose component are the spacial coordinates and the spin of the j-particle, then

$$\Psi = \Psi(\mathbf{x}_1, \mathbf{x}_1, \dots \mathbf{x}_N, t). \tag{A.2}$$

If the system is made by non-interacting components, or if these interact with a common external potential, $\hat{H}$ assumes a simple form, being the sum of single-particle Hamiltonians. In such a case, the wave function Ψ is reduced to the symmetric or anti-symmetric products of single-particle wave functions. For instance, the global wave function describing a system of independent identical Fermions is a Slater determinant:

$$\Psi(\mathbf{x}_1, \mathbf{x}_1, \dots \mathbf{x}_N, t) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_1(\mathbf{x}_2) & \dots & \phi_1(\mathbf{x}_N) \\ \phi_2(\mathbf{x}_1) & \phi_2(\mathbf{x}_2) & \dots & \phi_2(\mathbf{x}_N) \\ \vdots & \vdots & & \vdots \\ \phi_N(\mathbf{x}_1) & \phi_N(\mathbf{x}_2) & \dots & \phi_N(\mathbf{x}_N) \end{vmatrix} \tag{A.3}$$

Unfortunately, due to the complex interacting nature of ordinary matter, one can not rely on this independent particle picture to describe real systems. As a consequence, the Hamiltonian cannot be separated in a sum of identical terms and the wave function has a complicated system-dependent shape. The Hamiltonian describing a system of N_e interacting electrons and N_n interacting ions is:

$$\hat{H} = \hat{T}_\mathrm{e} + \hat{T}_\mathrm{n} + \hat{V}_{\mathrm{e-e}} + \hat{V}_{\mathrm{e-n}} + \hat{V}_{\mathrm{n-n}} \tag{A.4}$$

where $\hat{T}_{\rm e}$ and $\hat{T}_{\rm n}$ are respectively the electronic and the nuclei kinetic terms defined as

$$\hat{T}_e = \sum_i \frac{\mathbf{p}_i^2}{2\,m_{\rm e}} \qquad \hat{T}_n = \sum_I \frac{\mathbf{P}_I^2}{2\,m_{\rm n}} \tag{A.5}$$

and the $\hat{V}$ terms represent the Coulomb interactions among the constituent, defined as:

$$\hat{V}_{e-e} = \sum_{i<j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \qquad \hat{V}_{n-n} = \sum_{I<J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} \qquad \hat{V}_{n-e} = -\sum_{I,j} \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_j|}. \tag{A.6}$$

A universal solution for eq.(A.1) for many-body systems described by Hamiltonian (A.4) is unknown. Actually, analytic solutions have been determined only for very simple systems, such as the hydrogen atom or the ideal particle gas. Thus, the analysis of system showing more complex features is usually performed with numerical methods. One of the earliest attempt to address many-electrons problems is the Hartree method. The electron are assumed as independent particles with no requirement on the symmetry of the collective wave function. As a consequence, this method fails in describing a standard model used to describe the role of electrons interaction in solids such as the homogeneous electron gas. The simple requirement of the wave function anti-symmetry led to the Hartree-Fock (HF) method, which provides an analytic treatment of the jellium model. Moreover, HF ensures fairly accurate treatment of atomic and small molecules energetic, but it is not suitable for the description of many electron systems such as bulk solids. What is common to the failure of Hartree and HF methods is the crude approximation of the wave function which is based on an independent particle picture. Within this picture all the correlations among the electrons are neglected.

A.1.1 The Born-Oppenheimer approximation

The Hamiltonian (A.4) describes a portion of matter coupling the electrons and the nuclei dynamics through the $\hat{V}_{\rm n-e}$. Therefore, the wave function describing such system, is a function of both electronic and ionic degrees of freedom:

$$\Psi = \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_{N_e}, \mathbf{R}_1, \mathbf{R}_2, \ldots \mathbf{R}_{N_I}, t). \tag{A.7}$$

From now on in the current section, the letter $\mathbf{r}$ and $\mathbf{R}$ represent the whole set of electrons and ionic coordinates. An intuitive way to simplify the description of this intricate behaviour consists in assuming Ψ to have a separable form. This approximation, named after Born and Oppenheimer [193, 194], relies on the fact that the nuclei are much more massive than the electrons, hence on the time-scale of electronic motion they are nearly fixed. Thus the electronic component of the separated wave function have a simple parametric dependency on the specific ionic configuration. Specifically, the total wave function in eq. (A.7) can be expanded[1] on the eigenfunctions of the electronic Hamiltonian:

$$\hat{H}_{\rm el} = \hat{T}_{\rm e} + \hat{V}_{\rm e-e} + \hat{V}_{\rm e-n} + \hat{V}_{\rm n-n} \tag{A.8}$$

[1] To the extent of Born-Oppenheimer approximation validity, accurate solutions can be obtained using only one or a few terms.

where the nuclei kinetic term $\hat{T}_{\rm n}$ has been neglected. The secular problem reads:

$$\hat{H}_{\rm el}\phi_i^{\rm el}(\mathbf{r};\mathbf{R}) = \mu_i\phi_i^{\rm el}(\mathbf{r};\mathbf{R}) \tag{A.9}$$

where the eigenfunctions $\{\phi_i^{\rm el}(\mathbf{r};\mathbf{R})\}$ define the $i^{\rm th}$ many-electron states clamped to the nuclei configuration represented with $\mathbf{R}$, having an Hamiltonian eigenstate $\mu_{\rm i}$. Therefore the total wave function reads:

$$\Psi = \sum_i \phi_i^{\rm el}(\mathbf{r};\mathbf{R})\chi_i(\mathbf{R}) \tag{A.10}$$

where the $\chi_i(\mathbf{R})$ corresponds to the nuclei wave-function. A first-order Born-Oppenheimer approximation holds also for a single contribution in the decomposition of the total wave function. With the aid of eq. (A.10), the secular problem for the total Hamiltonian can be written, after some manipulations, as:

$$\big[\frac{1}{2m_{\rm n}}\nabla^2_{\mathbf{R}} + (\mu_j - E)\big]\chi_j(\mathbf{R}) - \frac{1}{2m_{\rm n}}\sum_i \big[2\tau_{ji}\cdot\nabla_{\mathbf{R}} + \tau'_{ji}]\chi_i(\mathbf{R}) = 0 \tag{A.11}$$

where the equation was projected onto the $j^{\rm th}$ electronic state, and the electronic degrees of freedom were integrated out. For simplicity the nuclei are taken with the same mass $m_{\rm n}$.

The matrices τ, τ' are the non-adiabatic couplings describing the mutual dependency of the electronic and the nuclei equations. Their matrix elements are defined as:

$$\tau_{ji} = \frac{1}{2m_{\rm n}}\int \phi_j^{\rm el\,*}(\mathbf{r},\mathbf{R})\nabla_{\mathbf{R}}\phi_i^{\rm el}(\mathbf{r},\mathbf{R})d\mathbf{r} \tag{A.12}$$

and

$$\tau'_{ji} = \frac{1}{2m_{\rm n}}\int \phi_j^{el\,*}(\mathbf{r},\mathbf{R})\nabla^2_{\mathbf{R}}\phi_i^{\rm el}(\mathbf{r},\mathbf{R})d\mathbf{r}. \tag{A.13}$$

When these matrix elements are negligible, or zero, the electron and the nuclei dynamics are uncoupled. As a consequence the electrons motion follow immediately the motions of the nuclei during their time evolution. This condition corresponds to adiabatic surfaces in the ionic configurations space that never cross each other[2]. When two electronic state come very close in energy, a condition named avoided crossing, or even exactly degenerate in energy, the adiabatic approximation fails. Indeed, in that case, two adiabatic potential surfaces would become very close, or in contact with each other at some point in the configuration space, making the so-called vibronic-coupling τ crucial to define the correct coupled dynamics of ions and electrons. Molecular systems where the Jahn-Teller effect takes place can be thought as an example of the limits of Born-Oppenheimer validity. The adiabatic approximation is a cornerstone for most of the available DFT codes today. Indeed, the electronic Hamiltonian is solved assuming fixed nuclei configuration given as an input, and the ions repulsive energy is added at the end of the calculations in order to compute the total energy $E_{\rm tot}$ of the system. Such $E_{\rm tot}$ coincide with

[2]After some manipulation, the dependence on τ' in equation (A.11) can be removed. In this formulation the equation of nuclei assume a compact shape which emphasizes the connection of τ to Berry's phase. The dependence on the Berry's connection could be removed through suitable unitary transformation of the nuclei wave function $\chi(\mathbf{R})$, usually named diabatic transformation. The disadvantage in this case would be the loss of the diagonal shape of the μ matrix, thus one would lose the visual description of adiabatic surfaces, one per electronic state.

the eigenvalue E in equation (A.11) where the τ are neglected.

A.2 Density functional theory

The state of the art for the description of electronic properties in solid state and molecular physics is density functional theory (DFT). The focus of the quantum mechanical description of matter is transposed from the global wave function to the much simpler ground-state electron density. The convenience is straightforward: the global wave function in eq. (A.2) depends on every degree of freedom of the system, the single-electron density just on three spatial coordinates. Indeed, the electron density can be written as:

$$\rho(\mathbf{r}) = \langle\Psi|\hat{\rho}(\mathbf{r})|\Psi\rangle = \tag{A.14}$$

$$= N \sum_{\sigma,\sigma_2,\ldots,\sigma_N} \int \left|\Psi(\mathbf{r}\sigma, \mathbf{r}_2\sigma_2, \mathbf{r}_3\sigma_3, \ldots, \mathbf{r}_N\sigma_N)\right|^2 d\mathbf{r}_2 d\mathbf{r}_3 \ldots d\mathbf{r}_N \tag{A.15}$$

where $\hat{\rho}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$. Note that the spin dependency is included in the density, therefore all possible spin components are comprised. Basically this choice allows to map the many-electron problem to a self-consistent-field one-particle problem for the ground-state properties. This is a huge simplification which allowed DFT to become probably the most popular simulation technique for first principle study of molecular and solids ground-state properties. What is even more surprising is that DFT is an exact theory for the ground-state energy. This exactness is extended to ground-state structural properties and with a reasonable degree of approximation to a large number of quantities directly connected to ground-state energy when suitable perturbations are included (i.e. polarizability, elastic constant, magnetic susceptibility etc.).

A.2.1 Density functional formalism

The first theoretical formulation of DFT appeared in 1964 in a paper from Hohenberg and Kohn [195]. In their work, the authors proved that the connection between the external potential included in a many-body Hamiltonian and the corresponding electronic ground-state density is a one-to-one map. The work is developed under the hypothesis of a static external potential and for non-degenerate ground-states. The theory has been shown to be extendable beyond these initial assumptions. Therefore both time-dependent external potentials and degenerate ground-states can be described with an analogous formalism [196]. The main results of Hohenberg and Kohn can be summarized in the following statements:

1. the ground-state energy of the many-body system can be written as a functional of the electron density

$$E_{v_{\text{ext}}}[\rho] = \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d(\mathbf{r}) + F[\rho] \tag{A.16}$$

   where the second functional F does not depend on v_{ext} and therefore is called a universal functional of the density;

2. $E_{v_{\text{ext}}}[\rho(\mathbf{r})]$ is minimum for the ground-state electron density:

$$E_{v_{\text{ext}}}[\rho_{gs}] < E_{v_{\text{ext}}}[\rho] \qquad \forall \rho \neq \rho_{gs}. \tag{A.17}$$

The demonstration of these propositions, which is a reductio ad absurdum and employs the Ritz variational principle, can be easily found in textbooks and in literature [195, 197, 198]. According to the second statement the ground-state density minimizes the energy functional in eq. (A.16). Thus the following equation holds:

$$\frac{\delta}{\delta\rho(\mathbf{r})}\Big\{E_{v_{\rm ext}}[\rho(\mathbf{r})] - \lambda\Big[\int \rho(\mathbf{r})d\mathbf{r} - N\Big]\Big\}\Big|_{\rho_{\rm GS}(\mathbf{r})} = 0 \tag{A.18}$$

where the number of electrons is fixed to N and λ is the required Lagrange multiplier. It is important to stress that the Hohenberg and Kohn theorem do not give any information on the shape of $F[\rho]$. If it was a simple functional of the electron density then it would be relatively straightforward to determine the ground-state density and energy, but generally speaking the universal functional is unknown. Indeed from eq. (A.16) it is clear that $F[\rho] = T[\rho] + E_{\rm e-e}[\rho] = T[\rho] + E_{\rm Ha}[\rho] + E_{\rm xc}[\rho]$, where T is the kinetic-energy, $E_{\rm Ha}$ is the mean-field interaction energy while $E_{\rm xc}$ is the exchange-correlation contribution to interaction energy i.e. those interaction contribution not included in a mean-field term. Thus, DFT formalism suffer from our ignorance about electrons interaction beyond the mean-field Hartree term: the universal functional inherit the ignorance that we had on the many-body wave function. As a consequence at this level DFT formalism is exact but useless from a practical point of view. One first possible attempt to overcome this issue is the Thomas-Fermi method [199, 200], where the electrons are treated as an ideal gas subjected to the ionic potential. The kinetic-energy is then computed per unit volume resulting in term scaling with $\rho^{\frac{5}{3}}(\mathbf{r})$ with $\rho(\mathbf{r})$ being the electron density. This model is valuable since it is a starting point for building a practical orbital-free density functional technique, but led to poor quantitative results for atomic and molecular calculation. Specifically, Teller [201] reported the failure of Thomas-Fermi model to predict correct molecular bonding energy: this is predicted to be higher than those of separated atoms. This inaccuracy is mainly due to the absence of exchange and correlations terms in the energy functional. The exchange term was later on included by Dirac, but also in this case the model led to poor predictions due to the crudeness of the kinetic-energy expression. All these attempt to retain an orbital-free density functional have been circumvented by Kohn-Sham DFT formulation which will be outlined in the next section.

A.2.2 Kohn-Sham formulation

Kohn and Sham introduced a powerful formalism [202] that became the cornerstone of modern practical use of DFT. The drawback of this formulation consists in the introduction of a basis set of wave functions, because one of the purposes of DFT is exactly to get rid of wave functions in the many-body problems. Nonetheless, the exceptional simplicity of the basis introduced by Kohn and Sham led to a valuable environment for a practical development of DFT.

Kohn-Sham equations

In addition to the real interacting electrons system, Kohn and Sham introduce an auxiliary non-interacting identical electron system having the same density of the real one and the same number of particles N. The auxiliary system is therefore described by a Slater determinant as defined in eq.(A.3), built with single-particle wave functions $\{\phi_{\rm i}(\mathbf{r})\}_{\rm i=1}^{\rm N}$ named Kohn-Sham orbitals. Thus,

the electron density is simplified from eq. (A.14) to the following:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} f_i \big|\phi_i(\mathbf{r})\big|^2 , \tag{A.19}$$

where

$$f_i = \frac{1}{e^{\beta(\epsilon_i - E_{\mathrm{F}})} + 1}$$

is the average occupation number of the single-particle state with energy ϵ_{i} according to the Fermi-Dirac statistics. Moreover, the single-electron basis enables also the electronic kinetic term to assume a very simple shape:

$$\langle T \rangle_{\mathrm{KS}} = -\frac{\hbar^2}{2m_{\mathrm{e}}} \sum_{i}^{N} f_i \int \phi_i^*(\mathbf{r}) \nabla_i^2 \phi_i(\mathbf{r}) d\mathbf{r} \tag{A.20}$$

where E_{F} is the Fermi energy of the system and $\beta = (k_{\mathrm{B}}T)^{-1}$ with T the temperature of the Fermions system. The introduction of the Kohn-Sham orbitals allows one to isolate all the complexities of the many-body wave function into one single term of the energy functional expressed in eq. (A.16): the exchange-correlation component. Indeed the energy functional can be written as:

$$E_{v_{\mathrm{ext}}}[\rho] = \int v_{\mathrm{ext}}(\mathbf{r}) \rho(\mathbf{r}) d(\mathbf{r}) + T_{\mathrm{KS}}[\rho] + \frac{1}{2} \int \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{\mathrm{xc}}[\rho] \tag{A.21}$$

where the exchange-correlation functional is defined as:

$$E_{\mathrm{xc}}[\rho] = T[\rho] - T_{\mathrm{KS}}[\rho] + E_{\mathrm{e-e}}[\rho] - E_{\mathrm{Ha}}[\rho]. \tag{A.22}$$

If the variational principle as written in eq. (A.17) is expressed in terms of the Kohn-Sham orbitals, a set of N single-particle equations are derived, which are formally equivalent to the Hartree-Fock equations. Indeed through eq. (A.19):

$$\frac{\delta}{\delta \rho(\mathbf{r})} \frac{\delta \rho(\mathbf{r})}{\delta \phi_i^*(\mathbf{r})} = \frac{\delta}{\delta \phi_i^*(\mathbf{r})} \tag{A.23}$$

and eq. (A.17) becomes:

$$\frac{\delta}{\delta \phi_i^*(\mathbf{r})} \big\{ E_{v_{\mathrm{ext}}}[\rho(\mathbf{r})] - \sum_{k,l} \lambda_{k,l} \big[\int \phi_k^*(\mathbf{r}) \phi_l(\mathbf{r}) d\mathbf{r} - \delta_{k,l} \big] \big\} =$$

$$\big[-\frac{\hbar^2}{2m} \nabla_i^2 + v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{Ha}}(\mathbf{r}) + v_{\mathrm{xc}}(\mathbf{r}) \big] \phi_i(\mathbf{r}) - \sum_j \lambda_{i,j} \phi_j(\mathbf{r}) = 0.$$

Through a suitable unitary transformation which preserves the electron density and the sum of the eigenvalues, one obtain a set of single-particle equation, whose behaviour is governed by a

self-consistent effective Hamiltonian:

$$H_{\text{eff}}\phi_i(\mathbf{r}) = \big[-\frac{\hbar^2}{2m}\nabla_i^2 + v_{\text{eff}}(\mathbf{r})\big]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}), \tag{A.24}$$

where

$$v_{\text{eff}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + v_{\text{Ha}}(\mathbf{r}) + v_{\text{xc}}(\mathbf{r}) \tag{A.25}$$

$$v_{\text{Ha}}(\mathbf{r}) = \frac{\delta E_{\text{Ha}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \tag{A.26}$$

$$v_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}. \tag{A.27}$$

Eq. (A.24) represents the single-particle Kohn-Sham equation for the i^{th} orbital. These are very similar to the HF equations introduced in sect. A.1. The main difference is the exchange-correlation term defined in eq. (A.27). The exact exchange interaction is explicitly considered in HF through the non-local Fock potential involving the density matrix, favouring parallel spin order. Notably, the Fock potential removes exactly the self-interaction spurious contribution included in the classic Hartree term. Instead, in Kohn-Sham equations the exchange is incorporated in the local v_{xc} term, which account also for the correlations among electrons.

Once the Kohn-Sham orbitals and eigenvalues ϵ_i in eq. (A.24) have been determined, the exact total ground-state energy of the system reads:

$$E_{v_{\text{ext}}} = \sum_i f_i\epsilon_i - \frac{1}{2}\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' - \int v_{\text{xc}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + E_{\text{xc}}[\rho(\mathbf{r})]. \tag{A.28}$$

The second term in the right member is needed to remove the double-counting effects in eq. (A.24), arising from the functional derivative in eq. (A.26). The last two terms in the right member are needed since no theoretical justification exists to state that E_{xc} is a linear functional of electrons density.

Kohn-Sham eigenvalues

Despite the Kohn-Sham equations provide an exact description of the ground-state electrons density and total energy, the Lagrange multipliers ϵ_i do not have an a-priori physical meaning. Nonetheless, they are usually considered as the orbital energies characterizing the electronic states. Indeed, theoretical results exist which attribute certain physical meaning to the eigenenergies. As stated by Koopmans' theorem [203], in HF under the hypothesis of frozen orbitals i.e. the orbitals of the ion are identical to those of the neutral molecule, the Lagrange multiplier of the highest occupied orbital corresponds to the first ionization energy when the self-consistent HF potential vanishes at infinite distance. In Kohn-Sham equations, the presence of the correlation term in (A.28) prevent the Koopmans' theorem to be valid. Instead the Janak's theorem holds [204]. The theorem states that the variation of the ground-state total energy in eq. (A.28) with

respect to an orbital occupation coincides to the Kohn-Sham eigenvalue of that orbital:

$$\frac{\partial E_{v_{\text{ext}}}}{\partial f_i} = \epsilon_i.$$

Thus, the ground-state total energy is a continuous function of the occupation number, and this function is piece-wise linear. The slope of the segments are the eigenvalues ϵ_{i}. In case on exact DFT, it can be shown that these slopes are discontinuous with the occupancy f_i. This reflects the non-analytic behaviour of the exact functional. Janak's theorem attributes a physical meaning to the single Kohn-Sham eigenenergies. For instance the difference in total energies of an $N+1$ electrons system with that of an N-electrons one is

$$E_{N+1} - E_N = \int_0^1 \frac{\partial E_{v_{\text{ext}}}}{\partial f_N} df_N$$

which is interpreted as the first ionization energy of a system, and could be evaluated with approximate methods such as the Slater transition state method [204, 205]. The f_N in the last equation is the occupancy of the highest occupied Kohn-Sham state.

Besides Janak's theorem, a DFT version of the Koopmans' theorem exists which ensures that the eigenvalue of the uppermost occupied Kohn-Sham orbital equals the exact ionization potential with opposite sign. Almbladh and von Barth [206] compare the asymptotic behaviour of DFT electron density (which is exact) to the one derived directly from the many-body Hamiltonian. The main difference is recognized in the exponent determining the decay into vacuum. This is related to the highest occupied eigenvalue for DFT density and to the first ionization potential for the second.

DFT has been widely used to characterize the electronic behaviour of materials through its prediction on the band gap. Besides a dependency on the choice of the approximated exchange-correlation functional, which is discussed in section (A.2.3), a fundamental issue exists regarding the computed gap within DFT. Indeed, the fundamental gap defined as the difference of the electron affinity and the first ionization energy [207], has often been reported to be underestimated [208, 209]. The reason stands in its definition [210]:

$$\begin{aligned} E_{\text{gap}}^{\text{fund}} &= I_N - A_N = \\ &= \epsilon_{N+1}(N+1) - \epsilon_{N+1}(N) + \epsilon_{N+1}(N) - \epsilon_N(N) = \\ &= \Delta_{\text{xc}} + E_{\text{gap}}^{\text{DFT}}. \end{aligned}$$

where $E_{\text{gap}}^{\text{DFT}} = \epsilon_{N+1}(N) - \epsilon_N(N)$ being $\epsilon_i(N)$ the i^{th} Kohn-Sham eigenvalue of a system with N electrons. The quantity Δ_{xc} is connected to the discontinuity in the derivative of the exchange-correlation potential. Therefore, DFT is somehow by construction underestimating the band gap. It is actually debated the possibility of this "natural" underestimation of not being a drawback, but an effect following the capability of DFT to approximately account for exchange-correlation holes [211]. Indeed it is physically expected to be an approximation to the excitation energy if electrons and holes are close rather than an approximation to the fundamental gap.

As a ground-state theory, DFT in its reported formulation is not developed to investigate excited states. Extensions have been proposed and developed during the years to deal with external

perturbation, in order to have a theoretical method to compare and analyze experimental results. For instance, time-dependent DFT (TD-DFT) relaxes the hypothesis of stationary states at the base of classic DFT [212]. TD-DFT shares some similarity with Hohenberg-Kohn formalism for the ground state even though no variational principle holds for time-dependent states. This extension evolved towards practical application mainly in the field of spectroscopic excitation and real-time evolution in non-perturbative fields.

A.2.3 Exchange-correlation functionals

Kohn-Sham formulation of DFT provides a practical framework to face complex many-electron problems, which would be otherwise unmanageable. Nonetheless, the complexity of the many-body problem is not entirely eliminated. Indeed, our ignorance of the exact wave-function in eq. (A.2) is now reflected on the exact shape of the exchange-correlation functional defined in eq. (A.22). Thus, several approximations has been proposed and employed during the years: the local-density approximation (LDA) and general gradient approximation (GGA) are probably the most widely employed. These approximations are often represented as the first steps of the Jacob's ladder, which points towards the exchange-correlation heaven-exactness (and complexity) [213].

Local-density approximation is probably the most popular one. This is particularly suitable for system having a slowly varying spatial density. Indeed, the exchange-correlation functional is built from the exchange-correlation energy of an homogeneous interacting electron gas with density $\rho(\mathbf{r})$:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})] = \int \epsilon_{\mathrm{xc}}^{\mathrm{heg}}(\rho(\mathbf{r}))\rho(\mathbf{r})d\mathbf{r} \tag{A.29}$$

where $\epsilon_{\mathrm{xc}}^{\mathrm{heg}}(\rho(\mathbf{r})) = \epsilon_{x}^{\mathrm{heg}}(\rho(\mathbf{r})) + \epsilon_{c}^{\mathrm{heg}}(\rho(\mathbf{r}))$. The first term is known in its exact form the analysis of the homogeneous gas, the second one can be analytically known only in its extreme density limits. The behaviour for intermediate values is based on interpolations of quantum Monte Carlo results performed by Ceperley *et al.* [214]. Despite its simplicity, LDA provides a good description of bulk materials. This is due to its exactness for at least one very peculiar system, the homogeneous gas, which make LDA capable to respect the correct sum rule to the exchange-correlation hole. That is, there is a total electronic charge of one electron excluded from the neighbourhood of the electron. Nonetheless being a local approximation make LDA incapable of correct the nonphysical electronic self-repulsion arising from the Hartree potential in eq. (A.25). This results in systematic inaccuracy in the prediction of specific quantities. Specifically, underestimates of band gaps and over-bind of electrons. Such an nonphysical term is compensated exactly in HF by the Fock operator which naturally remove the interaction of one electron with itself. However, this precise compensation results in an overestimate of the band gap, because the exchange term do not cancel the self-interaction for virtual orbitals.

The next step in the Jacobi's ladder is the general-gradient approximation which relax the strict locality requirement of LDA. For those systems whose electronic density exhibits significant variations in space, LDA is no longer a suitable approximation. A natural extension consists in including the gradient of the density in the description. In this way, the exchange-correlation functional acquires a semi-local character. Unluckily, low-order density gradient expansions are useless since the results are worse than LDA. Indeed at this order the sum rules for exchange-

correlation holes is no longer satisfied. Practically, generalized expressions are used, which have been shown to usually improve results over LDA [208]. Such generalized-gradient approximation functionals read:

$$E_{\rm xc}^{\rm GGA}[\rho(\mathbf{r})] = \int \epsilon_{\rm xc}(\rho(\mathbf{r}),\left|\nabla\rho(\mathbf{r})\right|)\rho(\mathbf{r})d\mathbf{r} \tag{A.30}$$

These functionals are usually created by fitting a wide number of materials, while requiring the kernel to respect some constrains such as to recover LDA kernel for homogeneous system, or having some known asymptotic behaviour. The Perdew-Burke-Ernzerhof functional comes from the second route, requiring, among other properties, the correct behaviour of the functional for low-limit density gradient. Both LDA and GGA admit extensions for spin-polarized system. The electron density dependence is split in the two spin-component: $\epsilon_{\rm xc}(\rho^{\uparrow},\rho^{\downarrow}) = \epsilon_x(\rho^{\uparrow},\rho^{\downarrow}) + \epsilon_c(\rho^{\uparrow},\rho^{\downarrow})$, where following the work of Oliver and Perdew [215] the paramagnetic exchange term can be written in term of the exact HF exchange term as: $\epsilon_x(\rho^{\uparrow},\rho^{\downarrow}) = \frac{1}{2}[\epsilon_x(\rho^{\uparrow},\rho^{\uparrow}) + \epsilon_x(\rho^{\downarrow},\rho^{\downarrow})]$.

One further extension of GGA is meta-GGA. A dependency on the non-interacting kinetic-energy density is used as input to the functional as well as the electron density and its gradient. The spin-independent form of this functional is thus:

$$E_{\rm xc}^{\rm MGGA}[\rho(\mathbf{r})] = \int \epsilon_{\rm xc}(\rho(\mathbf{r}),\left|\nabla\rho(\mathbf{r})\right|,\tau_{\rm KS}(\mathbf{r}))\rho(\mathbf{r})d\mathbf{r} \tag{A.31}$$

where $\tau_{\rm KS}(\mathbf{r})) = \sum_i \frac{1}{2}\left|\nabla\phi_i(\mathbf{r})\right|^2$. This functional allows for a more accurate description of band gaps. However the derivation of $v_{\rm xc}$ would require to calculation the derivative of the kinetic-energy density, which increases the computational complexity.

In order to circumvent the spurious self-interactions effect arising in the Hartree term, one possible solution is to include the exact Fock operator in the functional. Based on this idea, hybrid functionals have been arranged: some amount of the exact exchange is combined with given local or semi-local density functional approximation. This is done in the following way:

$$E_{\rm xc}^{\rm hybrid} = \alpha E_{\rm xc}^{\rm Fock} + (1-\alpha)E_x^{\rm DFT} \tag{A.32}$$

with α running from 0 to 1 determines the percentage of exact exchange to introduce. It's physical meaning is connected through the quasi-particle equation to the system dielectric function. Since the term E_x^{Fock} is built from the Kohn-Sham orbitals, and it is non local, Hybrid functional calculations could be expensive. Different form of hybrid functionals exist, based on different recipes: HSE06 addresses metal systems correcting short-range terms of PBE functional; PBE0 substitute a specific fraction of PBE exchange; B3LYP is based on a mixture of LDA and GGA functionals.

A.3 DFT in practice

The standard approach to solve Kohn-Sham equations consists of numerical self-consistent methods. The following procedure forms the basis common to modern available ab-initio codes. The first step is the choice of a starting point for the electronic density of the system. With this tentative density, the effective potential of entering Kohn-Sham equations is built. Here is where the approximation on the exchange-correlation functional enters. Then, the equations are solved

and the eigenvalues and the orbitals are obtained. Actually, following a long-standing approach in condensed matter, the differential equations are solved through the diagonalization of the corresponding secular-problem matrix. The new orbitals define a new density which is mixed with those stemming from a number of previous cycles. The code then returns to the calculation of the Hamiltonian and start again until convergence is reached, according to some criterion. The criterion(s) are usually connected to the relative change of some physical quantity between consecutive steps, usually the electronic density or density matrix elements. Once the convergence criterion is met the calculation is completed and the Kohn-Sham equations are considered to be numerically solved self-consistently. This is a general outline of the self-consistent procedure common to ab-initio codes.

The very starting point of an ab-initio simulation consists in the definition of the basis used to expand the orbitals. The most common distinction is between localized basis and delocalized basis sets. The expansion on a wave-function basis is truncated to obtain a computationally accessible set of equations. With this in mind, the reliability of the simulations depends directly on how many terms are employed to expand the orbitals. Which is usually determined by the convergence of extensive physical quantities.

A.3.1 Plane-waves basis

Bloch's theorem provides a basis for the expansion of electrons wave function which naturally describe their delocalized nature when subjected to a periodic potential. Bloch's wave functions are lattice-periodic functions modulated by plane waves: $\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u(\mathbf{r})$. Following this idea, one possible basis set to express Kohn-Sham orbitals in bulk systems is the plane wave (PW) basis. Such a basis make very easy to control the convergence quality of simulations. Indeed, better convergence can be achieved directly increasing the number of plane waves employed in calculations. This is usually driven by one cutoff parameter: the kinetic-energy related to the plane-wave state. Therefore, the PW basis is defined by:

$$\langle \mathbf{r}|\mathbf{k}+\mathbf{G}\rangle \propto \frac{1}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \qquad \frac{\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2 \leq E_{cut} \tag{A.33}$$

where $\mathbf{G}$ is a reciprocal-lattice vector, $\mathbf{k}$ is a wave vector in the first Brillouin zone, Ω is the primitive-cell volume and E_{cut} is the cutoff on the kinetic-energy of PW. Given a PW basis $\{\varphi_{\mathbf{k}}(\mathbf{r})\}$, the i^{th} Kohn-Sham orbitals can be expanded as:

$$\phi_i(\mathbf{r}) = \sum_G c_{i,\mathbf{k}+\mathbf{G}}\varphi_{\mathbf{k}+\mathbf{G}}(\mathbf{r}).$$

Thus, the secular problem corresponding to Kohn-Sham equation (A.24) for the i^{th} orbital becomes:

$$\sum_{\mathbf{G}'}[H_{\mathrm{KS}}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') - \epsilon_i\delta_{\mathbf{G},\mathbf{G}'}]c_{i,\mathbf{k}+\mathbf{G}'} = 0. \tag{A.34}$$

The matrix $H_{\mathrm{KS}}(\mathbf{k})$ is an $N \times N$ matrix, being N the number of basis elements. The expansion over PW basis provides many advantages: the elements of the basis are orthonormal and periodic by construction, they are unbiased because there is no freedom in choosing PWs and there is no dependence on the ions coordinates. Besides the matrix elements of the Hamiltonian assume

relatively simple form. Finally and importantly to achieve a better convergence of the results it is simply necessary to raise the cutoff. Drawbacks exists: supercells are required to simulate non-periodic systems, surfaces or defective lattices for instance, and fast-varying densities are not easily represented with plane waves. Specifically, capturing the rapid variation of the valence electrons wave-function close to the nuclear position requires a stunning number of plane waves. This specific issue led to the development of pseudo-potentials, introduced in section (A.3.3).

A.3.2 DFT with localized atomic orbitals

Localized basis sets are one possible alternative to plane waves. Contrary to delocalized basis set, different flavours can be selected: linear combination of atomic orbitals (LCAO), Gaussian-type orbitals, linearized Muffin-Tin orbitals etc. the SIESTA code [165], which has been employed in chapter 3.1 of the present thesis, makes use of the LCAO approximation. This means that the Kohn-Sham orbitals are expanded as:

$$\phi_i(\mathbf{r}) = \sum_{\nu} c_{i,\nu}\varphi_\nu(\mathbf{r})$$

where $\varphi_\nu(\mathbf{r})$ is an atomic orbital describing the single-electron state ν. Taking the atomic orbitals $\varphi_\nu(\mathbf{r})$ as Bloch-type functions, allow to deal with periodic systems. The secular problem connected to Kohn-Sham equations is now:

$$\sum_{\nu}[H^{\mathrm{KS}}_{\mu,\nu}(\mathbf{k}) - \epsilon_i(\mathbf{k})S_{\mu,\nu}(\mathbf{k})]c_{i,\mu}(\mathbf{k}) = \mathbf{0}. \tag{A.35}$$

which is already diagonal in the crystal momentum-space, indeed no $\mathbf{k} \neq \mathbf{k}'$ term contributes. The matrix $S_{\mu,\nu}(\mathbf{k})$ is the overlap matrix, which would be the identity matrix if the basis was orthogonal.

The orbitals are localized in space in such a way that they vanish beyond a certain radius $\mathrm{r_c}$. With this approximation the number of matrix elements to be computed and stored is significantly reduced. This is needed specifically for order-N method [165] which rely heavily on the sparsity of the Hamiltonian and overlap matrices. Being N the number of electrons, order-N method is an N-scaling method alternative to the matrix diagonalization which is the main bottleneck for DFT simulations in terms of computational effort since it normally scales as N^3. Unfortunately, the order-N method in SIESTA is only implemented for well-defined gap bulk systems, which is not the case for the system analyzed in chapter 3.1. Inside the cutoff radius, the atomic basis orbitals are the products of a numerical radial function times a spherical harmonic. Generally speaking, several orbitals could share the same angular dependence, but different radial dependence. This is one of the main degree of freedom in the definition of basis: the number of radial functions ζ employed to define the orbital radial profile inside the cutoff radius. Single-ζ orbital basis could be used. In this case the radial profile is a angular-momentum dependent eigenfunction of a pseudo-atom (described with a pseudo-potential) within a spherical box, for an energy ε chosen so that the first node occurs at the desired cutoff radius $r_{\mathrm{l,c}}$. As an alternative, multiple-ζ basis are introduced: two or more radial wave functions for a given angular momentum are computed within a "split-valence" scheme [216] and are successively mixed to retain the zero value beyond $r_{\mathrm{l,c}}$. Furthermore, to obtain well converged results, it is generally necessary to include polariza-

tion orbitals, to account for the wavefunction deformation induced by bond formation. To obtain these extra orbitals the pseudo-atom is perturbed with an external electric field. The orbitals resulting from the first-order perturbation theory are added to the basis [165]. Conventionally the different basis are indicated with acronyms: single-ζ plus polarization-orbitals is SZP, double-ζ plus polarization-orbital basis is DZP etc. Given its good balance between well converged results and reasonable computational cost, DZP is usually named the standard basis.

A.3.3 Pseudo-potentials

The possibility of employing plane-wave basis drove the introduction of pseudo-potentials which rapidly become a common ingredient for ab-initio calculations. This is a further approximation introduced in *ab-initio* simulations, which aims to simplify the practical treatment of electrons interactions while preserving the accuracy of DFT calculations.

The necessity emerges when the core electrons wave functions are considered in the expansion on the plane wave basis[3]. Indeed, given the true orbitals orthogonality, the higher angular momentum wave functions shows strong oscillations near the potential-energy singularities at the ionic core sites. Such feature requires a huge amount of plane wave components to be captured. Thus a large amount of terms have to be computed and stored to build the electron density and the hamiltonian in Kohn-Sham equations. However, only valence electrons are significant to the physics of chemical bonding and in general to the physical properties of the system. On the contrary, the wave functions of core electrons are mostly unaffected by the chemical environment. As a consequence, core electrons can be frozen out and ignored during the calculations. A simple justification of the nature of pseudo-potential is given in the so-called orthogonal plane waves method, developed by Herring in 1940 [210, 218]. The wave functions of valence electrons is decomposed in a smooth term and in a rapidly oscillating term. In order to make the smooth term orthogonal to the core states, the second component is expanded in terms of the such core electrons wave functions:

$$|\phi_i^v\rangle = |\tilde{\phi}_i^v\rangle + \sum_k a_k |\phi_k^c\rangle. \tag{A.36}$$

The coefficient of the expansion a_{k} are readily found by projecting the valence state over the k^{th} core state:

$$a_k = -\langle \phi_k^c | \tilde{\phi}_k^v \rangle. \tag{A.37}$$

Inserting the two previous definitions in the Schrödinger equation for the i^{th} valence electron gives a first definition of pseudo-potential:

$$\hat{H}|\phi_i^v\rangle = \hat{H}\big[|\tilde{\phi}_i^v\rangle + \sum_k a_k |\phi_k^c\rangle\big] = \epsilon_i |\phi_i^v\rangle \tag{A.38}$$

gives:

$$[\hat{H} + \hat{V}_{nl}]|\tilde{\phi}_i^v\rangle = [\hat{T} + \hat{V}_{\mathrm{eff}}]|\tilde{\phi}_i^v\rangle = \epsilon_i |\tilde{\phi}_i^v\rangle. \tag{A.39}$$

where a non-local extra-potential is added to the original hamiltonian, so that $\hat{V}_{\mathrm{eff}} = \hat{V} + \hat{V}_{nl}$,

[3]Despite there is no strict necessity to use them, pseudo-potentials can also be employed even within local-basis context. In SIESTA for instance they are employed in the norm-conserving Troullier-Martins parametrization [217], to get rid of the core electrons and to allow for the expansion of a smooth (pseudo)charge density on a uniform spatial grid.

defined as follows:

$$\hat{V}_{nl} = \sum_j (\epsilon_i^v - \epsilon_j^c)|\phi_j^c\rangle\langle\phi_j^c|. \tag{A.40}$$

The term V_{eff} is a pseudo-potential. With such an effective hamiltonian, the smooth component of the valence orbitals are degenerate in energy with the exact orbitals. Actually, as stated by Phillips and Kleinman [219] the non local potential in equation (A.39) is a repulsive term which soften the action of the original Coulomb potential.

One practical way of implement pseudo-potential formalism is to write them in a fully non-local form. This can be done for instance with the formalism introduced by Kleinman in [220] where the pseudo-potential reads:

$$\hat{V}_{\text{PS}} = \hat{V}_{\text{local}} + \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{m=+l} \frac{|\delta V_l \phi_{lm}\rangle\langle\phi_{lm}\delta V_l|}{\langle\phi_{lm}|\delta V_l|\phi_{lm}\rangle} \tag{A.41}$$

The ϕ_{lm} are the atomic pseudo wave functions (including angular term) for the reference state $|lm\rangle$ and $\delta V_l(r) = V_l(r) - V_{\text{local}}(r)$. The local term $V_{\text{local}}(r)$ is in principle arbitrary, but it has a long-ranged behaviour reminiscent of Coulomb potential: $-\frac{Z_v e^2}{r}$. Usually, $V_l(r) = V_{\text{local}}(r)$ for $r > r_{\text{core}}$ irrespective of the value of the angular momentum l.

Operatively, a pseudo-potential is obtained with following procedure. First, through the radial Schrödinger equation, one obtains the atomic valence orbitals of a specific element. With these orbitals one creates a set of node-less pseudo wave functions. The radial Schrödinger equation for these node-less (pseudo)orbitals can now be inverted to build the pseudo-potentials $V_l(r)$ for each angular momentum l. Finally the, Hartree and exchange-correlation contributions due to the valence electron density built from the pseudo-orbitals, are subtracted from the total pseudo-potential. This procedure is called "unscreening" and is needed to avoids to double counting the exchange-correlation contributions by the valence electrons when the pseudo-potential is employed in effective simulations. Furthermore, since the exchange-correlation potential is not a-priori a linear functional of the density, namely $V_{\text{xc}}[\rho_{\text{PS}} + \rho_{\text{c}}] \neq V_{\text{xc}}[\rho_{\text{PS}}] + V_{\text{xc}}[\rho_{\text{c}}]$, a simple subtraction of the contribution by $V_{\text{xc}}[\rho_{\text{PS}}]$ is usually not sufficient and the so-called non-linear core corrections are needed [210].

When the pseudo-orbitals are computed, some extra conditions may be required. For instance, generally speaking by construction each pseudo-orbitals shares the same orbital energy of the corresponding all-electron orbital $\phi_{lm}^{\text{ae}}(r)$, has to be node-less, it coincides with the all-electron orbital for $r > r_{\text{cutoff}}$. Further more, norm-conservation can be required:

$$\int_{r<r_{\text{cutoff}}} \left|\phi_{lm}^{\text{ps}}(r)\right|^2 r^2 dr = \int_{r<r_{\text{cutoff}}} \left|\phi_{lm}^{\text{ae}}(r)\right|^2 r^2 dr$$

The core radius r_{cutoff} is approximately at the outermost maximum of the wavefunction. These pseudo-potentials have wide transferability, but could become hard to employ, meaning that high cutoff in the plane-wave basis are required. This translate into larger CPU and RAM requirements. To overcome this possible hardness and obtain even more simplified and quick pseudo-potentials one could relax the condition over the conservation of the norm. These sorts of potentials are named ultra-soft. The cutoff radius r_{c} can be increased without loss of transferability, but a drawback appears. Since the orbital norm is no longer conserved, a so-called augmentation charge

must be included in the electronic density at each step of the self-consistent procedure. Such extra-charge need also to be included in forces routine when calculated within the Hellman-Feynman theorem [221].

APPENDIX B

B.1 *Ab-initio* calculation of phonons

Phonons in crystals are collective excitations corresponding to atomic vibrational modes. These vibrations propagate with a wave-vector $\mathbf{q}$ and are characterized by vibrational frequencies $\omega(\mathbf{q})$ defining the periodic displacements of the atoms around their equilibrium positions. The wave-vector $\mathbf{q}$ is the equivalent of the Bloch vector for the electronic states. Therefore all non-equivalent $\mathbf{q}$ are localized inside the first Brillouin zone (1BZ), i.e. the unit cell of the reciprocal lattice. Similarly to electrons, phonon frequencies fold in the 1BZ forming "bands". A system with N atoms in the unit cell is characterized by $3N$ frequencies for each $\mathbf{q}$. Among these, 3 are acoustic modes involving coherent vibrations of all the atoms in the primitive cell, the other $3N-3$ are the optical modes involving relative atomic motion inside the primitive cell. To describe the vibrational properties of a crystal one introduces the dynamical matrix $D(\mathbf{q})$, which is the mass-reduced Fourier transform of the inter-atomic force constant matrix. The $D(\mathbf{q})$ matrix can be derived under the harmonic approximation when expanding the crystal adiabatic energy around atomic equilibrium positions for small atomic displacements. In this case, "small" refers to interatomic distances. The dynamical-matrix eigenvalues ω^2 are the squared phonon frequencies while the atomic displacements are related to its eigenvectors ϵ [222–224]:

$$\sum_{\alpha}[D_{\alpha,\beta}(\mathbf{q}) - \omega^2\delta_{\alpha,\beta}]\epsilon_{\beta} = 0. \tag{B.1}$$

where Greek letters α, β are collective indices: they represent the Cartesian coordinates of all the atoms in a primitive cell. Thus the indices range from 1 to $3N$ and the dynamical matrix is $3N \times 3N$. According to harmonic approximation, phonons corresponds to stationary states. They can not account for several physical phenomena such as thermal expansion or thermal conductivity limitations in insulators. To describe these effects one needs to go beyond the harmonic approximation and to include higher-order terms in the adiabatic energy expansion.

To obtain crystal vibration frequencies within DFT framework, one can rely mainly on two techniques: the frozen phonons and perturbative expansion of DFT. In the frozen phonons method, the inter-atomic force constant matrix is built directly through an explicit evaluation of its elements. It is later diagonalized at a generic $\mathbf{q}$. Thus finite, periodic, displacements of a few atoms in an otherwise perfect crystal at equilibrium are introduced, and the forces acting on the displaced atoms are computed. To obtain the frequencies of a lattice vibrations with generic $\mathbf{q}$, this method needs to work with proper supercells whose linear dimensions must be therefore at least

of the order of $\frac{2\pi}{q}$. Since the number of atoms grows rapidly with the size of the supercell, the application of this technique are usually limited to zone-center phonon modes analysis or along selected high-symmetry path. Besides, complex materials are not advisable. The main advantage of the frozen-phonon approach is that it does not require any specialized computer code. This technique can indeed be implemented using any standard total energy and force code. The main drawback is the unfavorable scaling of the computational load with the number of atoms in the primitive cell and with the range R of the inter-atomic force constants. Indeed supercells needs linear dimensions higher than R, including a number of atoms $N_{sc} \propto R^3$. Thus, given the cubic scaling of self-consistent DFT calculations with the number of atom considered, a full frozen phonons calculation will scale as $3NR^9$. Obviously for increasing system complexity such calculation becomes rapidly overwhelming for ordinary computational resources. As a consequence, complex systems require techniques which can perform better than frozen phonons. Perturbative expansion of DFT, also known as DFPT, reduces the overall computational effort computing the dynamical matrix through linear-response theory. Besides, it includes the possibility of treat polar materials, which is something not directly provided by frozen phonons [224]. DFPT is desirable for complex systems: having no necessity for a supercell the scaling with the number of atoms is favourable with respect to frozen phonons. For instance, consider the calculation of inter-atomic force constants through grid-based DFPT. The wave vectors in the Brillouin need a spacing of the order of the inverse of the range of the inter-atomic force constants: $\Delta q \propto \frac{2\pi}{R}$. Hence the number of $\mathbf{q}$ points in such a grid is of the order of R^3. The calculation of each column of the matrix has a computational cost of the order of N^3 and the number of such columns is $3N$. Therefore, the cost for the calculation of the complete inter-atomic force constant matrix, thus of complete phonon dispersion, is of the order of $3N^4R^3$.

B.2 Linear response: density functional perturbation theory

In the framework of Born-Oppenheimer approximation, where the lattice and electrons have uncoupled dynamics, the ground-state energy obtained by DFT coincide with the energy of a system of interacting electrons moving in the field of fixed nuclei. Within this approximation, one define the adiabatic-energy surface as:

$$E(\{\mathbf{R}\}) = \mu_i^{\mathrm{el}} + V_{\mathrm{n-n}}(\{\mathbf{R}\}) \tag{B.2}$$

where $V_{\mathrm{n-n}}(\{\mathbf{R}\})$ is the Coulomb energy ion-ion repulsion defined in eq. (A.6). with μ_i^{el} defined in eq. (A.9) where i labels an electronic eigenstate (in the context of Kohn-Sham DFT it is defined through (A.28)). In the adiabatic approximation context, given an electronic state, the adiabatic-energy surfaces governs ionic motion. The equilibrium geometry of the system is the one where each atom do not experience any net resulting force, represented by the global minimum of the adiabatic potential energy surface (B.2). Local minima correspond to metastable states. For small deviation from this condition, the harmonic approximation allows the adiabatic energy surface to be treated as a paraboloid. In this context it can be shown that the dynamical matrix can be obtained through linear response theory, with the perturbing potential being a variation in the external nuclear potential.

The Feynman-Hellman theorem states that the first derivative of the eigenvalues of a Hamiltonian, H_λ that depends on a parameter λ, is given by the expectation value of the derivative of the Hamiltonian:

$$\frac{\partial E_\lambda}{\partial \lambda} = \langle \Psi | \frac{\partial H_\lambda}{\partial \lambda} | \Psi \rangle.$$

In the current case the parameter λ coincides with one, or several, atomic coordinates, and the theorem holds due to Born-Oppenheimer approximation because the ions coordinates enter the electronic wave functions only as parameters[1]. Thus, within the adiabatic approximation context the force acting on the I^{th} nucleus in the electronic ground state is:

$$\mathbf{F}_I = -\frac{\partial E_{\mathbf{R}}}{\partial \mathbf{R}_I} = -\int d\mathbf{r} \rho_{\mathbf{R}}(\mathbf{r}) \frac{\partial V_{\mathbf{R}}(\mathbf{r})}{\partial \mathbf{R}_I} - \frac{\partial E_n n(\mathbf{R})(\mathbf{r})}{\partial \mathbf{R}_I} \tag{B.3}$$

where $\rho_{\mathbf{R}}(\mathbf{r})$ is the electronic density, $V_{\mathbf{R}}(\mathbf{r})$ is the external potential of DFT representing the Coulomb interaction among nuclei and electrons and $E_N(\mathbf{R})$ is defined in eq. (A.6). The subscript $\mathbf{R}$ labels the parametric dependency on the fixed ionic configuration.

Deriving eq. (B.3), one obtains the inter-atomic force constant matrix, which is by construction the Hessian matrix of the adiabatic energy with respect to nuclei coordinates:

$$\frac{\partial^2 E_{\mathbf{R}}}{\partial \mathbf{R}_I \partial \mathbf{R}_J} = \int \rho_{\mathbf{R}}(\mathbf{r}) \frac{\partial V_{\mathbf{R}}(\mathbf{r})}{\partial \mathbf{R}_I \partial \mathbf{R}_J} d\mathbf{r} + \int \frac{\partial \rho_{\mathbf{R}}(\mathbf{r})}{\partial \mathbf{R}_J} \frac{\partial V_{\mathbf{R}}(\mathbf{r})}{\partial \mathbf{R}_I} d\mathbf{r} + \frac{\partial^2 E_n n(\mathbf{R})}{\partial \mathbf{R}_I \partial \mathbf{R}_J}. \tag{B.4}$$

The terms on the right member of this equation require the calculation of the ground-state electron charge density $\rho(\mathbf{r})$ as well as of its linear response[2] to a distortion of the nuclear geometry $\partial \rho_{\mathbf{R}}(\mathbf{r}) / \partial \mathbf{R}_J$. Thus, once the ground state $\rho(\mathbf{r})$ is known, one is left to compute the linear response of the electron-density to the ions displacement [224, 226]. This linear response can be obtained in terms of the perturbed Kohn-Sham orbitals in a self consistent way, through first-order standard perturbation theory [227]. Being H_{scf} the unperturbed hamiltonian in eq. (A.24), $|\phi_i\rangle$, ϵ_i the unperturbed i^{th} Kohn-Sham orbital and its eigenvalue respectively, then the perturbed orbital $|\delta\phi_i\rangle$ is defined by the equation:

$$[H_{\text{scf}} - \epsilon_i] |\delta\phi_i\rangle = -(\delta v_{\text{scf}} - \delta\epsilon_i) |\phi_i\rangle. \tag{B.5}$$

where δv_{scf} is the first order correction to the self-consistent potential of eq. (A.25):

$$\delta v_{\text{scf}} = \delta v_{\text{ext}} + \int \frac{\delta\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta v_{\text{xc}}}{\delta\rho(\mathbf{r})} \delta\rho(\mathbf{r}) \tag{B.6}$$

and $\delta\epsilon_i = \langle \phi_i | \delta v_{\text{scf}} | \phi_i \rangle$ is the first order correction to the i^{th} eigenvalue. In the context of solid state physics eq. (B.5) is also known as Sternheimer equation [228]. The self-consistent character of these equations is evident in the dependency of the perturbed potential on the corrections to the electronic density. Such corrections are defined in terms of the perturbed orbitals, which are the

[1] In DFT implementation, if a plane wave basis is employed to expand the Kohn-Sham orbitals then the theorem is still valid, but it is no longer true in case of a LCAO approximation, because the atomic orbitals depends explicitly on the ions coordinates and pulay forces appears

[2] This result is understood in terms of the so-called $2n+1$ theorem [225], which basically states that the $2n+1$ order derivative of an hamiltonian eigenvalue, can be expressed in term of the previous derivative up to order n.

solutions of eq. (B.5), through:

$$\delta\rho(\mathbf{r}) = 2\Re\sum_{j=1}^{N_{\rm el}} \phi_j^*(\mathbf{r})\delta\phi_j(\mathbf{r}) \tag{B.7}$$

Thus an iterative procedure is needed to solve equation (B.5). Once convergence is achieved, one is able to compute the linear response of electrons density to the displacement of the ions, and the hessian matrix in (B.4) and the dynamical matrix. The secular problem of eq. (B.1) is now recovered and the phonons eigenvalues can be obtained for a generic wave-vector $\mathbf{q}$.

One of the main advantages of DFPT, compared to non-perturbative methods, is that within DFPT the response to the different Fourier components of a perturbation are decoupled. That is to say, the response of the system on perturbation with a certain wavelength is independent from the response to different wavelengths perturbations. Indeed eq. (B.5), after some manipulation, can be decomposed into Fourier components in such a way to deal with lattice-periodic functions only. This feature prevents the use of supercells to calculate phonon frequencies at arbitrary wave vectors $\mathbf{q}$ with workload that is essentially independent of the phonon wavelength.

B.3 DFPT in practice

Plane-waves based codes such as Abinit [229] or Quantum espresso [129] compute the phonon frequencies by means of density functional perturbation theory. The approach coincides with the one described in the previous section. Thus, the charge response to lattice distortion of definite wave-vector $\mathbf{q}$ is computed starting from the electronic structure of the regular crystal, obtained from a conventional DFT self-consistent calculation. A different response in the electron density have to be calculated for each independent wave vector. The DFPT implementation in Quantum Espresso allows for two routes for the calculation of phonons dispersion, either with or without a $\mathbf{q}$ grid. If the grid is employed, the dynamical matrix is computed on each $\mathbf{q}$ vector. Then, the inter-atomic force constant matrix is built through a Fourier interpolation. Finally, the knowledge of this approximated force constant matrix allow for the dynamical matrix to be computed and diagonalized at a generic $\mathbf{q}$ vector. Otherwise, one can perform a direct calculation of the dynamical matrix at each specified $\mathbf{q}$ wave vector along the required path in the first Brillouin zone. The $3N$ frequencies are then extracted by diagonalization of the matrix at each $\mathbf{q}$.

In Quantum-Espresso implementation of the ground-state energy calculation, the $\mathbf{k}$-points grid employed to sample the first Brillouin zone, i.e. to solve Kohn-Sham equations, is not entirely used to compute the various physical quantities. Instead, calling G the group of crystal symmetries, only a subset of such $\mathbf{k}$ points is considered: only the wave vectors which are nonequivalent under the operations of G are employed. Thus, only those $\mathbf{k}$ points enclosed in the so called irreducible Brillouin zone of the lattice are needed. The physical quantities that are defined by means of first Brillouin zone integration are obtained averaging over its value on this $\mathbf{k}$-point subset after having imposed the symmetries of the crystal. In general, regardless of the details of the systems, time-reversal symmetry is assumed so $\mathbf{k}$ and $-\mathbf{k}$ are considered as equivalent. Thus the inversion symmetry on the grid is always present, and the sums can always be performed on half of the $\mathbf{k}$ points of the mesh. Similarly, in a phonon calculation the linear response equations are not computed on each $\mathbf{k}$ wave vector on the grid. This implies that, the Sternheimer equations (B.5)

are not solved on each element of the grid. Instead, only the $\mathbf{k}$ vectors which are nonequivalent under rotations of the group $G_{\mathbf{q}}$ are considered. The group $G_{\mathbf{q}}$ is called the small group of $\mathbf{q}$ and includes all the rotations of the symmetry group G which leaves unchanged $\mathbf{q}$ or transform it into an equivalent $\mathbf{q} + \mathbf{G}'$ ($\mathbf{G}'$ in this case is a primitive reciprocal lattice vector). The density response function $\partial\rho(\mathbf{r})/\partial\mathbf{R}_{\mathbf{q}}$ is obtained by summing on this $\mathbf{k}$-point subset and then, by imposing the symmetries of $G_{\mathbf{q}}$. The denominator of the response function indicates a perturbation of the ions coordinates characterized by a wave vector $\mathbf{q}$. As for ground-state DFT, due of time reversal symmetry, the response is not needed for both $\mathbf{q}$ and $-\mathbf{q}$. Of course in case of lack of symmetries, the $\mathbf{k}$ points are all needed. Once the density response function has been determined at a certain $\mathbf{q}$, one is able to determine the dynamical matrix $D(\mathbf{q})$. Moreover, one can exploit the lattice symmetries once more. Indeed, let's define the subset of $\{\tilde{\mathbf{q}}\}$ obtained rotating the $\mathbf{q}$ with the rotations of the group G as the star of $\mathbf{q}$. Thus, every vector in the star of $\mathbf{q}$ yields to the same electron-density response function: $\partial\rho(\mathbf{r})/\partial\mathbf{R}_{\mathbf{q}} = \partial\rho(\mathbf{r})/\partial\mathbf{R}_{\tilde{\mathbf{q}}}$.

APPENDIX C

C.1 Density functional tight-binding

Despite its great success, DFT computational scaling limits the size of the treatable systems. Thus semi-empirical and empirical methods remains valid for the analysis of space and time scales not achievable by DFT. One empirical method which is useful in many ways is density functional tight-binding (DFTB) [166]. It is a tight binding formulation which can be derived from a Taylor-series expansion of the Kohn-Sham density functional total energy of eq. (A.28) (plus the ion-ion repulsion energy) around a reference density $\rho(\mathbf{r})$. The reference density is assumed to be that of a free and neutral atomic system. Clearly this choice does not minimize the DFT energy functional. But one could assume that it neighbors the true minimizing density: $\rho_{\min}(\mathbf{r}) = \rho_{\mathbf{0}}(\mathbf{r}) + \delta\rho(\mathbf{r})$ with $\delta\rho(\mathbf{r})$ assumed to be small. The second-order expansion of DFT total energy in fluctuation $\delta\rho(\mathbf{r})$ reads:

$$E[\delta\rho] = E_{\mathrm{BS}}[\delta\rho] + E_{\mathrm{Coul}}[\delta\rho] + E_{\mathrm{rep}} \tag{C.1}$$

where:

$$E_{\mathrm{BS}}[\delta\rho] = \sum_a f_a \langle\psi_a|H[\rho_0]|\psi_a\rangle \tag{C.2}$$

corresponds to the evaluation of the effective Kohn-Sham Hamiltonian built with the unperturbed atomic density ρ_0 over the single particle state $|\psi_{\mathrm{a}}\rangle$. Eq. (C.2) describes the band structure of this atomic neutral system. The term:

$$E_{\mathrm{Coul}}[\delta\rho] = \frac{1}{2}\int\int \Big(\frac{\delta^2 E_{\mathrm{xc}}[\rho_0]}{\delta\rho\,\delta\rho'} + \frac{1}{|\mathbf{r}-\mathbf{r}'|}\Big) d\mathbf{r}d\mathbf{r}' \tag{C.3}$$

represents the energy from charge fluctuations. It is mainly due to Coulomb interaction but contains also xc-contributions. The last term:

$$E_{\mathrm{Rep}} = -\frac{1}{2}\int V_{\mathrm{Ha}}[\rho_0](\mathbf{r})\rho_0(\mathbf{r})d\mathbf{r} + E_{\mathrm{xc}}[\rho_0(\mathbf{r})] + E_{\mathrm{II}} - \int V_{\mathrm{xc}}[\rho_0](\mathbf{r})\rho_0(\mathbf{r})d\mathbf{r} \tag{C.4}$$

is collectively named the repulsive energy due to the presence of the ion-ion repulsion term defined in eq. (A.6), but it contains more quantities from the Taylor expansion. The repulsive term is usually reduced to a sum over different atoms

$$E_{\mathrm{rep}} = \sum_{I<J} V_{\mathrm{rep}}^{IJ}(R_{IJ}) \tag{C.5}$$

depending on the relative distance R_{IJ} of the I and J atoms. The Coulomb term can be written in terms of the Mulliken charges [230], which represent the induced charge on the atom I, thus an extra electrons population on the atom I, as

$$E_{\text{Coul}} = \sum_{I<J} \gamma(R_{IJ}) \Delta q_I \Delta q_J . \tag{C.6}$$

The $\gamma(R_{IJ})$ represents the distance-dependent charge-charge interaction connected to the chemical hardness $U = IE - AE$ where IE is the ionization energy and EA is the electron affinity.

Following a tight-binding approach, the energy in eq. (C.1) is expanded on a minimal atomic orbital basis $|\psi_\mu\rangle$. Minimality means having only one radial function for each angular momentum state. Thus:

$$|\psi_a\rangle = \sum_\mu c_\mu^a |\phi_\mu\rangle. \tag{C.7}$$

After simple manipulations, the energy in equation (C.1) becomes:

$$E = \sum_a f_a \sum_{\mu\nu} c_\mu^{a*} c_\nu^a H_{\mu\nu}^0 + \sum_{I<J} \gamma_{IJ}(R_{IJ}) \Delta q_I \Delta q_J + \sum_{I<J} V_{\text{rep}}^{IJ}(R_{IJ}). \tag{C.8}$$

with the Mulliken charge expressed as:

$$\Delta q_I = \sum_a f_a \sum_{\mu \in I} \sum_\nu \frac{1}{2} \big(c_\mu^{a*} c_\nu^a + c.c. \big) S_{\mu\nu}. \tag{C.9}$$

f_a is the occupation of the a^{th} single particle state, $S_{\mu\nu}$ is the overlap matrix of the atomic orbitals basis. The minimum of this expression is obtained by variation of $\delta(E - \sum_a \epsilon_a \langle \psi_a | \psi_a \rangle)$, where ϵ_a are undetermined Lagrange multipliers, constraining the wave function norms. Thus:

$$\sum_\nu c_\nu^a (H_{\mu\nu} - \epsilon_a S_{\mu\nu}) = 0 \tag{C.10}$$

is the secular equation to be solved, where the Hamiltonian matrix element, with I and J labelling different atoms, is:

$$H_{\mu\nu} = H_{\mu\nu}^0 + \frac{1}{2} S_{\mu\nu} \sum_k (\gamma_{Ik} + \gamma_{Jk}) \Delta q_k \qquad \mu \in I, \nu \in J \, . \tag{C.11}$$

Of course molecular systems can be perfectly described with this atomic orbital formalism. For periodic systems the Kohn-Sham orbitals are expanded on a Bloch waves basis instead of a pure localized basis:

$$\phi_\mu(\mathbf{k}, \mathbf{r}) = \frac{1}{\sqrt{N}} \sum_R e^{i\mathbf{k}\cdot\mathbf{R}} \phi_\mu(\mathbf{r} - \mathbf{R}) \tag{C.12}$$

where $\mathbf{R}$ is a lattice vector. The Kohn-Sham orbitals, hence, being expressed as linear combinations on this basis, are themselves Bloch waves. In either cases, the atomic orbitals should not be taken from free atoms, as they would be to diffuse. To this purpose, they are defined by an

Hamiltonian with an additional artificial confinement potential, which reads:

$$H_{\mathrm{pseudo}} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r} + V_{\mathrm{Ha}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r}) + V_{\mathrm{conf}}(r). \tag{C.13}$$

Typically, the confining potential is a spherical power potential. Other form has been employed, such as the Woods-Saxon potential [231]. In this way the orbital diffuse tails are cut off and their overlap are reduced. As a consequence the sparsity of the Hamiltonian matrix increases, helping its diagonalization.

This TB formulation naturally incorporate a pure independent-electron approach and its extension including electron-electron interaction through Mulliken charges. The independent electrons picture is obtain neglecting the term depending on the extra-charges Δq_k in the Hamiltonian (C.11). If the correction are included, then the construction of the Hamiltonian is treated self-consistently. From an initial guess of the Mulliken charge population one obtains the second term at the right member of eq. (C.11). Once the matrix has been assembled new coefficient $\{c^{\mathrm{a}}_{\mu}\}$ are obtained from the secular problem in eq. (C.10) and new Mulliken population is built from (C.9). The procedure is repeated until certain convergence condition is reached. In both cases, the tight-binding spirit is adopted by accepting the matrix elements $H^0_{\mu\nu}$ as the principal parameters of the method. This means that these matrices elements are previously computed and stored in a parameters file read every time the matrices have to be constructed. The repulsive term in the Hamiltonian has to be previously computed too. Usually, it is designed in such a way to adjust the pseudo-atomic description of some reference systems against *ab-initio* predictions or empirical results.

C.2 Tight-binding parametrization

Three objects in DFTB are parametrized before the simulations: the Hamiltonian matrix, the overlap matrix and the repulsive potential. Usually, the method to build these objects consists in the minimization of some objective functions against a set of parameters introduced in the Hamiltonian and in the repulsive term. There is no standard recepy for such parametric description, on the contrary a certain freedom is possible. The objective function is usually the difference between some physical quantity computed with the current parametrization and *ab-initio* methods, or with respect to some empirically known quantity. The minimization procedure occurs through global minimum search algorithms, such as particle swarm, simulated annealing, genetic algorithms etc. [232]. The objective function is defined according to what the tight-binding aims to reproduce. For instance it will be the difference in binding energy if energetic stability is to be analyzed. Moreover, several functions could be employed. The difference in binding energy could be supported by a forces comparison if geometrical relaxation is foreseen, or the band structure could be also addressed, in order to obtain a parametrization capable to reproduce the electronic properties of the analyzed system as well. One widely employed workflow to prepare the parametrization consists in starting with the optimization of the Hamiltonian and overlap matrices, followed by the fit of the repulsive term in (C.8) to best approximate *ab-initio* results. Some finer method includes a second round, where the matrices elements are adjusted in such a way to reproduce the first-round final results. The repulsive part is newly fitted to *ab-initio* results [232].

The elements $H_{\mu\nu}$ parametrized are described more in detail in what follows. The on-diagonal

elements of the Hamiltonian $H^0_{\mu\mu}$ are the on site energies of the pseudo-atomic orbitals for the different angular momentum, computed within DFT formalism. The effective potential of the Kohn-Sham equations does not contains the additional confinement potential. Besides, only one-center integrals are retained in the calculation of these elements. I.e. the crystal field contribution from different atoms is neglected. Thus, calling v^A_{eff} the effective potential arising from the nucleus A, the matrix element reads:

$$H^0_{\mu\mu} = \int \phi^*_\mu(\mathbf{r}) \big[-\frac{\hbar^2}{2m}\nabla^2 + v^A_{\text{eff}} \big] \phi_\mu(\mathbf{r}) d\mathbf{r}. \tag{C.14}$$

All the terms including an effective potential v^B_{eff} arising from a $B \neq A$ are neglected. The on-diagonal elements correspond to the on-site energies for the state labeled with μ. Being on-site, these elements are stored and read as they are whenever required during the simulations. Instead, the off-diagonal elements of the Hamiltonian are computed with DFT within a two-center approximation: one assumes the integral with three different localized centers, to be small enough to be neglected. In this case, the matrix element reads:

$$H^0_{\mu\nu} = \int \phi^*_\mu(\mathbf{r}) \big[-\frac{\hbar^2}{2m}\nabla^2 + v^A_{\text{eff}} + v^B_{\text{eff}} \big] \phi_\nu(\mathbf{r}) d\mathbf{r}. \tag{C.15}$$

where $\mu \in A$ and $\nu \in B$. The off-diagonal elements are computed through this equation and are stored in the form of tables as a function of the distance among the nuclei A and B, which coincide with the orbitals center. However, the calculation of these off-diagonal matrix elements, as for the overlap matrix, is simplified by the Slater-Koster transformations: only few integrals have to be calculated numerically for each pair of orbitals which are later combined to assemble all the others.

The repulsive potential is built to fit the results of the Hamiltonian parametrization to the DFT ones. Several procedures exists in order to build these pairwise potentials. One possible method consists in evaluating the repulsive potential as the difference in the total energy computed with DFT and DFTB according to eq. (C.2) as a function of the pairwise atomic distance over a group of target structures:

$$V_{\text{rep}}(R) = \langle E_{\text{DFT}}(R) - E_{\text{BS}}(R) \rangle. \tag{C.16}$$

The brackets correspond to averaging over the ensemble of the target structures. The points $\{R, V_{\text{rep}}(R)\}$ are then fitted with spline functions to ensure continuity of the derivatives up to a certain order. Another possible solution consists in expressing each pairwise repulsive potential as a linear combination of cutoff polynomials. This is a faster solution giving potentials with comparable quality to the spline method.

C.2.1 Slater-Koster integrals

The elements of the matrices are not entirely computed from scratch to prepare the parametrization. Instead, only specific integrals, representing the dependence of the matrices elements on the orbital centers distance, are needed to complete the matrices. Slater-Koster transformation rules allow every $H_{\mu\nu}$ or $S_{\mu\nu}$ involving s, p or d orbitals, to be written in term of one to three simple integrals named Slater-Koster integrals. These simpler integrals are taken among a group of 10,

labeled $dd\sigma, dd\pi, dd\delta, pd\sigma, pd\pi, pp\sigma, pp\pi, sd\sigma, sp\sigma, ss\sigma$, depending on the atomic orbitals μ, ν involved and on the value of l of the molecular orbitals considered σ, π, δ. For instance, using the letter τ to label these 10 integrals, an overlap matrix element namely reads:

$$S_{\mu\nu} = \sum_{\tau} c_{\tau} S_{\mu\nu}(\tau) \tag{C.17}$$

The matrices elements are hence computed according to this law, and stored in the form of tables as a function of the distance among the nuclei, which coincide with the orbitals center. The number of columns forming the table is twice the number of nonequivalent orbitals combinations: one for the Hamiltonian one for the overlap matrix. For instance, the integral table for a homonuclear system consisting only of carbon will have 20 columns: ten for the Hamiltonian $H_{dd\sigma}, H_{dd\pi}, H_{dd\delta}, H_{pd\sigma}, H_{pd\pi}, H_{pp\sigma}, H_{pp\pi}, H_{sd\sigma}, H_{sp\sigma}, H_{ss\sigma}$ and the equivalent for the overlap matrix. Clearly, each column involving d orbitals will be a column of zeros.

C.2.2 Tight-binding molecular dynamics

Regular tight-binding total energy does not require a self-consistent construction, since the Hamiltonian and the repulsive energy are read by the parameters file. As such, tight-binding is a preferential framework to perform molecular dynamic simulations for systems which are not accessible by *ab-initio* molecular dynamics. Indeed, Car-Parriniello molecular dynamics provides a powerful framework to describe the inter-atomic interactions through self-consistent first-principles calculations in connection with the time-evolution of the system. The electronic and atomic levels are described, but information on the dependency of materials properties on the temperature and pressure can be extracted. Although, the method is computationally too expensive for many purposes in materials research [233]. Tight-binding based molecular dynamics is a compromise between the effectiveness of classical molecular dynamics, and the quantum description, exact but cumbersome, of *ab-initio* molecular dynamics. The quantum mechanics description is retained through the LCAO expansion, and the force field is obtained by the total energy expressed in equation (C.8) through its gradient with respect to the atomic positions. The force field cannot be prepared before the simulations as it depends directly on the atomic coordinates of each specific configuration during the time evolution. It is computed at every step of advancement of the dynamics of the nuclei and it is employed to determine the successive one.

APPENDIX D

D.1 Supplementary materials

The atomic coordinates of the investigated allotropes are given as supplementary contents of the publications produced throughout this PhD.

Specifically, the Cartesian coordinates of single-ring novamene configuration reported in Fig.2.4 can be found as .xyz file at
https://www.sciencedirect.com/science/article/pii/S2405844016310386#upi0005.

The fractional coordinates of the atomic non-equivalent positions of protomene can be found at https://www.sciencedirect.com/science/article/pii/S0008622317310692 as .cif file.

The fractional atomic coordinates of zayedene can be found at https://pubs.rsc.org/en/content/articlelanding/2019/cp/c9cp03978c#!divAbstract as .cif file.

List of Figures

List of Tables

Bibliography

1. Kroto, H., Heath, J., O'Brien, S., *et al.* C_{60}: Buckminsterfullerene. *Nature* **318,** 162 (1985).
2. Sano, N., Wang, H., Chhowalla, M., *et al.* Nanotechnology: Synthesis of carbon'onions' in water. *Nature* **414,** 506 (2001).
3. Iijima, S. & Ichihashi, T. Single-shell carbon nanotubes of 1-nm diameter. *Nature* **363,** 603 (1993).
4. Bethune, D., Kiang, C., De Vries, M., *et al.* Cobalt-catalysed growth of carbon nanotubes with single-atomic-layer walls. *Nature* **363,** 605 (1993).
5. Yamada, K., Kunishige, H. & Sawaoka, A. Formation process of carbyne produced by shock compression. *Sci. Nat.* **78,** 450 (1991).
6. Novoselov, K. S., Geim, A. K., Morozov, S. V., *et al.* Electric Field Effect in Atomically Thin Carbon Films. *Science* **306,** 666 (2004).
7. Wallace, P. R. The Band Theory of Graphite. *Phys. Rev.* **71,** 622 (1947).
8. McClure, J. W. Diamagnetism of Graphite. *Phys. Rev.* **104,** 666 (1956).
9. Geim, A. & Novoselov, K. *Nat. Mater.* **6,** 183 (2007).
10. Xu, M., Liang, T., Shi, M., *et al.* Graphene-like two-dimensional materials. *Chemi. Rev.* **113,** 3766 (2013).
11. Hirsch, A. & Hauke, F. Post-Graphene 2D Chemistry: The Emerging Field of Molybdenum Disulfide and Black Phosphorus Functionalization. *Angew. Chem.* **57,** 4338 (2018).
12. Chopra, N., Luyken, R., Cherrey, K., *et al.* Boron nitride nanotubes. *Science* **269,** 966 (1995).
13. Nath, M., Govindaraj, A. & Rao, C. N. R. Simple synthesis of MoS2 and WS2 nanotubes. *Adv. Mater.* **13,** 283 (2001).
14. Na, S., Kim, S., Hong, W., *et al.* Fabrication of TiO2 nanotubes by using electrodeposited ZnO nanorod template and their application to hybrid solar cells. *Electrochim. Acta* **53,** 2560 (2008).
15. Hirsch, A. The era of carbon allotropes. *Nat. Mater.* **9,** 868 (2010).

16. Liu, F., Ming, P. & Li, J. Ab initio calculation of ideal strength and phonon instability of graphene under tension. *Phys. Rev. B* **76,** 064120 (2007).

17. Papageorgiou, D., Kinloch, I. & Young, R. Mechanical properties of graphene and graphene-based nanocomposites. *Prog. Mat. Sci.* **90,** 75 (2017).

18. Wang, J., Deng, S., Liu, Z., *et al.* The rare two-dimensional materials with Dirac cones. *Natl. Sci. Rev.* **2,** 22 (2015).

19. Katsnelson, M., Novoselov, K. & Geim, A. Chiral tunnelling and the Klein paradox in graphene. *Nature phys.* **2,** 620 (2006).

20. Novoselov, K., Geim, A., Morozov, S., *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438,** 197 (2005).

21. Bolotin, K., Ghahari, F., Shulman, M., *et al.* Observation of the fractional quantum Hall effect in graphene. *Nature* **462,** 196 (2009).

22. Robinson, T. On Klein tunneling in graphene. *Am. J. Phys.* **80,** 141 (2012).

23. Castro Neto, A. H., Guinea, F., Peres, N. M. R., *et al.* The electronic properties of graphene. *Rev. Mod. Phys.* **81,** 109 (2009).

24. Morozov, S. V., Novoselov, K. S., Katsnelson, M. I., *et al.* Giant Intrinsic Carrier Mobilities in Graphene and Its Bilayer. *Phys. Rev. Lett.* **100,** 016602 (2008).

25. Chen, J.-H., Jang, C., Xiao, S., *et al.* Intrinsic and extrinsic performance limits of graphene devices on SiO2. *Nat. Nanotechnol.* **3,** 206 (2008).

26. Rycerz, A., Tworzydło, J. & Beenakker, C. Valley filter and valley valve in graphene. *Nature Phys.* **3,** 172 (2007).

27. Han, M., Özyilmaz, B., Zhang, Y., *et al.* Energy Band-Gap Engineering of Graphene Nanoribbons. *Phys. Rev. Lett.* **98,** 206805 (2007).

28. Shemella, P., Zhang, Y., Mailman, M., *et al.* Energy gaps in zero-dimensional graphene nanoribbons. *Appl. Phys. Lett.* **91,** 042101 (2007).

29. Barone, V., Hod, O. & Scuseria, G. Electronic structure and stability of semiconducting graphene nanoribbons. *Nano lett.* **6,** 2748 (2006).

30. Jiang, J., Wang, B., Wang, J., *et al.* A review on the flexural mode of graphene: lattice dynamics, thermal conduction, thermal expansion, elasticity and nanomechanical resonance. *J. Phys. Condens. Matter* **27,** 083001 (2015).

31. Landau, L. D., Kosevich, A., Pitaevskii, L. P., *et al. Theory of elasticity* (Butterworth-Heinemann, 1986).

32. Jiang, J., Wang, J. & Li, B. Young's modulus of graphene: a molecular dynamics study. *Phys. Rev. B* **80,** 113405 (2009).

33. Dinadayalane, T. & Leszczynski, J. in *Handbook of Computational Chemistry* 793–867 (Springer Netherlands, Dordrecht, 2012).

34. Prinzbach, H., Weiler, A., Landenberger, P., *et al.* Gas-phase production and photoelectron spectroscopy of the smallest fullerene, C 20. *Nature* **407,** 60 (2000).

35. Hedberg, K., Hedberg, L., Bethune, D., *et al.* Bond lengths in free molecules of buckminsterfullerene, C60, from gas-phase electron diffraction. *Science* **254,** 410 (1991).

36. Yannoni, C., Bernier, P., Bethune, D., *et al.* NMR determination of the bond lengths in C60. *J. Am. Chem. Soc.* **113,** 3190 (1991).

37. Hirsch, A., Chen, Z. & Jiao, H. Spherical Aromaticity in Ih Symmetrical Fullerenes: The 2(N+1)2 Rule. *Angew. Chem. Int. Ed.* **39,** 3915 (2000).

38. Mitsumoto, R., Araki, T., Ito, E., *et al.* Electronic structures and chemical bonding of fluorinated fullerenes studied by NEXAFS, UPS, and vacuum-UV absorption spectroscopies. *J. Phys. Chem.* **102,** 552 (1998).

39. Haddon, R. C. Chemistry of the Fullerenes: The Manifestation of Strain in a Class of Continuous Aromatic Molecules. *Science* **261,** 1545 (1993).

40. Heiney, P., Fischer, J., McGhie, A., *et al.* Orientational ordering transition in solid C_{60}. *Physical Review Letters* **66,** 2911 (1991).

41. Tingaev, M. & Belenkov, E. *Carbon materials formed by polymerization of C20 and C24 fullerites* in *J. Phys. Conf. Ser.* **1124** (2018), 022011.

42. Núñez-Regueiro, M., Marques, L., Hodeau, J., *et al.* Polymerized Fullerite Structures. *Phys. Rev. Lett.* **74,** 278 (1995).

43. Iwasa, Y. Superconductivity: Revelations of the fullerenes. *Nature* **466,** 191 (2010).

44. Ibrahim, K. Carbon nanotubes-properties and applications: a review. *Carbon Lett.* **14,** 131 (2013).

45. Zhang, R., Zhang, Y., Zhang, Q., *et al.* Growth of half-meter long carbon nanotubes based on Schulz–Flory distribution. *Acs Nano* **7,** 6156 (2013).

46. Laird, E., Kuemmeth, F., Steele, G., *et al.* Quantum transport in carbon nanotubes. *Rev. Mod. Phys.* **87,** 703 (3 2015).

47. Lu, X. & Chen, Z. Curved pi-conjugation, aromaticity, and the related chemistry of small fullerenes (¡ C60) and single-walled carbon nanotubes. *Chem. Rev.* **105,** 3643 (2005).

48. L, J., Lu, J., Lin, X., *et al.* The electronic properties of chiral carbon nanotubes. *Comput. Mater. Sci.* **129,** 290 (2017).

49. Filleter, T., Bernal, R., Li, S., *et al.* Ultrahigh strength and stiffness in cross-linked hierarchical carbon nanotube bundles. *Adv. Mater.* **23,** 2855 (2011).

50. Peng, B., Locascio, M., Zapol, P., *et al.* Measurements of near-ultimate strength for multiwalled carbon nanotubes and irradiation-induced crosslinking improvements. *Nat. Nanotechnol.* **3,** 626 (2008).

51. Shirakawa, H., Louis, E., MacDiarmid, A., *et al.* Synthesis of electrically conducting organic polymers: halogen derivatives of polyacetylene CH. *J. Chem. Soc., Chem. Commun.* 578 (1977).

52. A., W.-G. Carbon: A New View of Its High-Temperature Behavior. *Science* **200,** 763 (1978).

53. Kudryavtsev, Y. in *Carbyne and Carbynoid Structures* 1–6 (Springer, 1999).

54. Heimann, R., Evsyukov, S. & Koga, Y. Carbon allotropes: a suggested classification scheme based on valence orbital hybridization. *Carbon* **35,** 1654–1658 (1997).

55. Baughman, R. Dangerously Seeking Linear Carbon. *Science* **312,** 1009 (2006).

56. *Polyynes: Synthesis, Properties, and Applications* (ed Cataldo, F.) (CRC, Taylor&Francis, London, 2005).

57. Bogana, M. & Colombo, L. *Appl. Phys. A* **86,** 275 (2007).

58. Casari, C. & Milani, A. Carbyne: from the elusive allotrope to stable carbon atom wires. *MRS Commun.* **8,** 207 (2018).

59. Casari, C. S., Li Bassi, A., Ravagnan, L., *et al.* Chemical and thermal stability of carbyne-like structures in cluster-assembled carbon films. *Phys. Rev. B* **69,** 075422 (2004).

60. Shi, L., Rohringer, P., Suenaga, K., *et al.* Confined linear carbon chains as a route to bulk carbyne. *Nature materials* **15,** 634 (2016).

61. Chalifoux, W. & Tykwinski, R. Synthesis of polyynes to model the sp-carbon allotrope carbyne. *Nature Chem.* **2,** 967 (2010).

62. Lagow, R., Kampa, J., Wei, H., *et al.* Synthesis of linear acetylenic carbon: the" sp" carbon allotrope. *Science* **267,** 362 (1995).

63. Szczepanski, J., Fuller, J., Ekern, S., *et al.* Electronic absorption and resonance Raman spectra of large linear carbon clusters isolated in solid argon. *Spectrochim. Acta Part A* **57,** 775 (2001).

64. Kastner, J., Kuzmany, H., Kavan, L., *et al.* Reductive preparation of carbyne with high yield. An in situ Raman scattering study. *Macromolecules* **28,** 344 (1995).

65. Ravagnan, L., Siviero, F., Lenardi, C., *et al.* Cluster-Beam Deposition and in situ Characterization of Carbyne-Rich Carbon Films. *Phys. Rev. Lett.* **89,** 285506 (2002).

66. Heimann, R., Evsyukov, S. & Kavan, L. *Carbyne and carbynoid structures* (Springer Science & Business Media, 1999).

67. Cinquanta, E., Manini, N., Ravagnan, L., *et al.* Oxidation of carbynes: Signatures in infrared spectra. *J. Chem. Phys.* **140,** 244708–244714 (2014).

68. Peierls, R. E. *Quantum Theory of Solids* p. 23. (Oxford University Press, London, 1955).

69. Yang, S. & Kertesz, M. Linear Cn clusters: Are they acetylenic or cumulenic? *J. Phys. Chem. A* **112,** 146 (2008).

70. Liu, M., Artyukhov, V. I., Lee, H., *et al.* Carbyne from first principles: chain of C atoms, a nanorod or a nanorope. *ACS Nano* **7,** 10075 (2013).

71. Wang, M. & Lin, S. Ballistic thermal transport in carbyne and cumulene with micron-scale spectral acoustic phonon mean free path. *Sci. Rep.* **5,** 18122 (2015).

72. Shi, L., Rohringer, P., Wanko, M., *et al.* Electronic band gaps of confined linear carbon chains ranging from polyyne to carbyne. *Phys. Rev. Materials* **1,** 075601 (2017).

73. Zhu, Y., Bai, H. & Huang, Y. Electronic Property Modulation of One-Dimensional Extended Graphdiyne Nanowires from a First-Principle Crystal Orbital View. *ChemistryOpen* **5,** 78 (2016).

74. Cataldo, F., Ursini, O., Angelini, G., *et al.* Simple synthesis of α, ω-diarylpolyynes part 1: Diphenylpolyynes. *J. Macromol. Sci. A* **47,** 739–746 (2010).

75. Cataldo, F., Ursini, O., Milani, A., *et al.* One-pot synthesis and characterization of polyynes end-capped by biphenyl groups (α, ω-biphenylpolyynes). *Carbon* **126,** 232 (2018).

76. Baughman, R., Eckhardt, H. & Kertesz, M. Structure-property predictions for new planar forms of carbon: Layered phases containing sp 2 and sp atoms. *The Journal of chemical physics* **87,** 6687 (1987).

77. Sun, Q., Cai, L., Ma, H., *et al.* Dehalogenative homocoupling of terminal alkynyl bromides on Au (111): incorporation of acetylenic scaffolding into surface nanostructures. *ACS nano* **10,** 7023 (2016).

78. Sun, Q., Tran, B., Cai, L., *et al.* On-Surface Formation of Cumulene by Dehalogenative Homocoupling of Alkenyl gem-Dibromides. *Angew. Chem.* **129,** 12333 (2017).

79. Hou, J., Inganäs, O., Friend, R. H., *et al.* Organic solar cells based on non-fullerene acceptors. *Nat. Mater.* **17,** 119 (2018).

80. Cheng, P., Li, G., Zhan, X., *et al.* Next-generation organic photovoltaics based on non-fullerene acceptors. *Nat. Photonics* **12,** 131 (2018).

81. Dennler, G., Scharber, M. C. & Brabec, C. J. Polymer-fullerene bulk-heterojunction solar cells. *Adv. Mater.* **21,** 1323 (2009).

82. Ross, R. B., Cardona, C. M., Guldi, D. M., *et al.* Endohedral fullerenes for organic photovoltaic devices. *Nat. Mater.* **8,** 208 (2009).

83. Martinez, Z. S., Castro, E., Seong, C., *et al.* Fullerene Derivatives Strongly Inhibit HIV-1 Replication by Affecting Virus Maturation without Impairing Protease Activity. *Antimicrob. Agents Chemother.* **60,** 5731 (2016).

84. Friedman, S. H., De Camp, D. L., Sijbesma, R. P., *et al.* Inhibition of the HIV-1 protease by fullerene derivatives: model building studies and experimental verification. *J. Am. Chem. Soc.* **115,** 6506 (1993).

85. Bianco, A., Kostarelos, K. & Prato, M. Applications of carbon nanotubes in drug delivery. *Curr. Opin. Cell Biol.* **9,** 674 (2005).

86. Duan, W. & Wang, Q. Water transport with a carbon nanotube pump. *ACS nano* **4,** 2338 (2010).

87. Das, R., Ali, M., Hamid, S., *et al.* Carbon nanotube membranes for water purification: a bright future in water desalination. *Desalination* **336,** 97 (2014).

88. Eatemadi, A., Daraee, H., Karimkhanloo, H., *et al.* Carbon nanotubes: properties, synthesis, purification, and medical applications. *Nanoscale Res. Lett.* **9,** 393 (2014).

89. Fiori, G., Bonaccorso, F., Iannaccone, G., *et al.* Electronics based on two-dimensional materials. *Nat. Nanotechnol.* **9,** 768 (2014).

90. Jang, H., Park, Y. J., Chen, X., *et al.* Graphene-based flexible and stretchable electronics. *Adv. Mater.* **28,** 4184 (2016).

91. Bonaccorso, F., Sun, Z., Hasan, T., *et al.* Graphene photonics and optoelectronics. *Nat. Photonics* **4,** 611 (2010).

92. Li, X., Tao, L., Chen, Z., *et al.* Graphene and related two-dimensional materials: Structure-property relationships for electronics and optoelectronics. *Appl. Phys. Rev.* **4,** 021306 (2017).

93. Deringer, V. L., Csányi, G. & Proserpio, D. M. Extracting Crystal Chemistry from Amorphous Carbon Structures. *ChemPhysChem* **18,** 873–877 (2017).

94. Jung, J., DaSilva, A. M., MacDonald, A. H., *et al.* Origin of band gaps in graphene on hexagonal boron nitride. *Nat. Commun.* **6,** 6308 (2015).

95. Moreno, C., Vilas-Varela, M., Kretz, B., *et al.* Bottom-up synthesis of multifunctional nanoporous graphene. *Science* **360,** 199 (2018).

96. Wang, X., Yang, X., Wang, B., *et al.* Significant band gap induced by uniaxial strain in graphene/blue phosphorene bilayer. *Carbon* **130,** 120 (2018).

97. Hu, Y., Xie, P., De Corato, M., *et al.* Bandgap Engineering of Graphene Nanoribbons by Control over Structural Distortion. *J. Am. Chem. Soc.* **140,** 7803 (2018).

98. Jia, X. & Wei, F. in *Single-Walled Carbon Nanotubes: Preparation, Properties and Applications* 299 (Springer International Publishing, 2019).

99. Gupta, T. & Kumar, S. in *Carbon Nanotube-Reinforced Polymers* 61 (Elsevier, 2018).

100. Ravagnan, L., Manini, N., Cinquanta, E., *et al.* Effect of Axial Torsion on sp Carbon Atomic Wires. *Phys. Rev. Lett.* **102,** 245502 (2009).

101. Schmidt, J., Shi, J., Borlido, P., *et al.* Predicting the thermodynamic stability of solids combining density functional theory and machine learning. *Chem. Mater.* **29,** 5090 (2017).

102. Jain, A., Ong, S. P., Hautier, G., *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials* **1,** 011002 (2013).

103. Kirklin, S., Saal, J. E., Meredig, B., *et al.* The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT Sawaoka energies. *Npj Comput. Mat.* **1,** 15010 (2015).

104. Curtarolo, S., Hart, G. L., Nardelli, M. B., *et al.* The high-throughput highway to computational materials design. *Nat. Mater.* **12,** 191 (2013).

105. Pizzi, G., Cepellotti, A., Sabatini, R., *et al.* AiiDA: automated interactive infrastructure and database for computational science. *Comput. Mater. Sci.* **111,** 218 (2016).

106. Hoffmann, R., Kabanov, A. A., Golov, A. A., *et al.* Homo citans and carbon allotropes: for an ethics of citation. *Angew. Chem. Int. Ed.* **55,** 10962 (2016).

107. Pauling, L. The nature of the chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. *J. Am. Chem. Soc.* **53,** 1367 (1931).

108. Maksić, Z. Symmetry, hybridization and bonding in molecules. *Comput. Math. Appl.* **12,** 697 (1986).

109. Salehpour, M. R. & Satpathy, S. Comparison of electron bands of hexagonal and cubic diamond. *Phys. Rev. B* **41,** 3048 (1990).

110. Harris, P. J. F. Fullerene-related structure of commercial glassy carbons. *Philos. Mag.* **84,** 3159 (2004).

111. Ho, M. & Lau, A.-T. in *Fillers and Reinforcements for Advanced Nanocomposites* (eds Dong, Y., Umer, R. & Lau, A. K.-T.) 309 (Woodhead Publishing, 2015).

112. Ravagnan, L., Piseri, P., Bruzzi, M., *et al.* Influence of Cumulenic Chains on the Vibrational and Electronic Properties of $sp-sp^2$ Amorphous Carbon. *Phys. Rev. Lett.* **98,** 216103 (21 May 2007).

113. Benedek, G., Milani, P. & Ralchenko, V. G. *Nanostructured Carbon for Advanced Applications: Proceedings of the NATO Advanced Study Institute on Nanostructured Carbon for Advanced Applications Erice, Sicily, Italy July 19–31, 2000* (Springer Science & Business Media, 2001).

114. Cassagnau, P. in *Rheology of Non-Spherical Particle Suspensions* 59 (Elsevier, 2015).

115. Milani, P. & Manfredini, M. Laser-Induced Coalescence of C_{60} in the Solid State. *J. Phys. Chem.* **99,** 16119 (1995).

116. Meinardi, F., Paleari, A., Manfredini, M., *et al.* Raman characterization of amorphous carbon produced by fullerite laser-induced transformation. *Solid State Commun.* **93,** 335 (1995).

117. Milani, P., Ferretti, M., Parisini, A., *et al.* Synthesis of carbon nano and meso-structures by laser-induced coalescence of fullerenes. *Carbon* **36,** 495 (1998).

118. Street Jr, K., Miyoshi, K. & Vander Wal, R. in *Superlubricity* 311 (Elsevier, 2007).

119. Narayan, J., Bhaumik, A., Gupta, S., *et al.* Progress in Q-carbon and related materials with extraordinary properties. *Mater. Res. Lett.* **6,** 353 (2018).

120. Burchfield, L. A., Fahim, M. A., Wittman, R. S., *et al.* Novamene: A new class of carbon allotropes. *Heliyon* **3,** 242 (2017).

121. Delodovici, F., Manini, N., Wittman, R., *et al.* Protomene: A new carbon allotrope. *Carbon* **126,** 574 (2018).

122. Delodovici, F., Choi, D., Al Fahim, M., *et al.* Carbon sp chains in diamond nanocavities. *Phys. Chem. Chem. Phys.* **21,** 21814 (2019).

123. Manini, N. & Tosatti, E. in *Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials: Vol. 2* (eds Kadish, K. & Ruoff, R.) 1017 (The Electrochemical Society, Pennington, NJ, 1995).

124. Manini, N. & Tosatti, E. in *Fullerene-Related Materials* (ed Margadonna, S.) chap. 6 (Springer, Berlin, 2014, 2011).

125. Dorset, D. L. & McCourt, M. P. Disorder and the molecular packing of C_60 buckminsterfullerene: a direct electron-crystallographic analysis. *Acta Crystallogr. A* **50,** 344 (1994).

126. Brinkmann, G., Caporossi, G. & Hansen, P. A constructive enumeration of fusenes and benzenoids. *J. Algorithms* **45,** 155 (2002).

127. Brinkmann, G., Caporossi, G. & Hansen, P. A Survey and New Results on Computer Enumeration of Polyhex and Fusene Hydrocarbons. *J. Chem. Inf. Comput. Sci* **43.** PMID: 12767142, 842 (2003).

128. Favot, F. & Dal Corso, A. Phonon dispersions: Performance of the generalized gradient approximation. *Phys. Rev. B* **60,** 11427 (1999).

129. Giannozzi, P., Baroni, S., Bonini, N., *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21,** 395502 (2009).

130. Baroni, S., de Gironcoli, S., Dal Corso, A., *et al.* Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **73,** 515–562 (2001).

131. Manini, N., Dal Corso, A., Fabrizio, M., *et al.* Electron-vibration coupling constants in positively charged fullerene. *Philos. Mag. B* **81,** 793 (2001).

132. Lüders, M., Bordoni, A., Manini, N., *et al.* Coulomb couplings in positively charged fullerene. *Philos. Mag. Part B* **82,** 1611 (2002).

133. Onida, G., Manini, N., Ravagnan, L., *et al. Phys. Status Solidi B* **247,** 2017 (2010).

134. Cataldo, F., Ravagnan, L., Cinquanta, E., *et al. J. Phys. Chem. B* **114,** 14834 (2010).

135. Cinquanta, E., Ravagnan, L., Castelli, I. E., *et al. J. Chem. Phys.* **135,** 194501 (2011).

136. Castelli, I. E. *Structural and Magnetic Properties of sp-Hybridized Carbon* diploma thesis. MA thesis (University Milan, 2010).

137. Castelli, I. E., Ferri, N., Onida, G., *et al.* Carbon sp chains in graphene nanoholes. *J. Phys. Condens. Matter* **24,** 104019 (2012).

138. Grüner, G. The dynamics of charge-density waves. *Rev. Mod. Phys.* **60,** 1129 (1988).

139. A., D. & C., P. Electronic structure and optical properties of Si, Ge and diamond in the lonsdaleite phase. *J. Phys. Condens. Matter* **26,** 045801 (2014).

140. Zhang, Y., Maddox, S. & Manteghi, S. Verification of Class B S-N curve for fatigue design of steel forgings. *Int. J. Fatigue* **92,** 246 (2016).

141. Oliveira, E., da Silva Autreto P.A, Woellner, C., *et al.* On the mechanical properties of novamene: A fully atomistic molecular dynamics and DFT investigation. *Carbon* **139,** 782 (2018).

142. Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **117,** 1 (1995).

143. Kawai, T., Miyamoto, Y., O. Sugino, O., *et al.* Graphitic ribbons without hydrogen-termination: Electronic structures and stabilities. *Phys. Rev. B* **62,** R16349 (2000).

144. Liu, Y., Jones, R., Zhao, X., *et al.* Carbon species confined inside carbon nanotubes: a density functional study. *Phys. Rev. B* **68,** 125 (2003).

145. Wassmann, T., Seitsonen, A. P., Saitta, A. M., *et al.* Structure, stability, edge states, and aromaticity of graphene ribbons. *Phys. Rev. Lett.* **101,** 096 (2008).

146. Okada, S. Energetics of nanoscale graphene ribbons: Edge geometries and electronic structures. *Phys. Rev. B* **77,** 041408 (2008).

147. Jr., E. H., Pontes, R., Fazzio, A., *et al.* Formation of atomic carbon chains from graphene nanoribbons. *Phys. Rev. B* **81,** 201406 (2010).

148. Castelli, I. E., Salvestrini, P. & Manini, N. Mechanical properties of carbynes investigated by ab initio total-energy calculations. *Phys. Rev. B* **85,** 214–110 (2012).

149. Zanolli, Z., Onida, G. & Charlier, J. Quantum Spin Transport in Carbon Chains. *ACS Nano* **4,** 5174 (2010).

150. Cinquanta, E., Manini, N., Ravagnan, L., *et al.* Oxidation of carbynes: Signatures in infrared spectra. *J. Chem. Phys.* **140,** 244708–244714 (2014).

151. Esser, M., Esser, A. A., Proserpio, D. M., *et al.* Bonding analyses of unconventional carbon allotropes. *Carbon* **121,** 154 (2017).

152. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38,** 3098 (1988).

153. Perdew, J. P., Chevary, J. A., Vosko, S. H., *et al.* Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **46,** 6671 (1992).

154. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77,** 3865 (1996).

155. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13,** 5188 (1976).

156. Blatov, V. A., O'Keeffe, M. & Proserpio, D. M. Vertex-, face-, point-, Schläfli-, and Delaney-symbols in nets, polyhedra and tilings: recommended terminology. *CrystEngComm* **12,** 44 (2010).

157. Anurova, N. A., Blatov, V. A., Ilyushin, G. D., *et al.* Natural tilings for zeolite-type frameworks. *J. Phys. Chem. C* **114,** 10160 (2010).

158. Baburin, I. A., Proserpio, D. M., Saleev, V. A., *et al.* From zeolite nets to sp 3 carbon allotropes: a topology-based multiscale theoretical study. *Phys. Chem. Chem. Phys.* **17,** 1332 (2015).

159. Blatov, V. A., Shevchenko, A. P. & Proserpio, D. M. Applied topological analysis of crystal structures with the program package ToposPro. *Crys. Growth & Des.* **14,** 3576 (2014).

160. Proserpio, D. M., Golov, A. A. & Kabanov, A. A. *the "Samara Carbon Allotrope Database" - SACADA* (`http://sacada.sctms.ru/`, 2012).

161. Marton, M., Vojs, M., Zdravecká, E., *et al.* Raman spectroscopy of amorphous carbon prepared by pulsed arc discharge in various gas mixtures. *J. Spectrosc.* **2013,** 467079 (2012).

162. Oliveira, E., da Silva Autreto, P., Woellner, C., *et al.* On the mechanical properties of protomene: A theoretical investigation. *Comput. Mat. Sci.* **161,** 190 (2019).

163. Bundy, F. & Kasper, J. Hexagonal Diamond-A New Form of Carbon. *J. Chem. Phys.* **46,** 3437 (1967).

164. He, H., Sekine, T. & Kobayashi, T. Direct transformation of cubic diamond to hexagonal diamond. *Appl. Phys. Lett.* **81,** 610–612 (2002).

165. Soler, J. M., Artacho, E., D., J., *et al.* The SIESTA method for ab initio order-N materials simulation. *J. Phys. Condens. Matter* **14,** 2745 (2002).

166. Koskinen, P. & Mäkinen, V. Density-functional tight-binding for beginners. *Comput. Mater. Sci.* **47,** 237 (2009).

167. Köhler, C. & Frauenheim, T. Molecular dynamics simulations of CFx (x=2,3) molecules at Si3N4 and SiO2 surfaces. *Surf. Sci.* **600,** 453 (2006).

168. Rauls, E., Elsner, J., Gutierrez, R., *et al.* Stoichiometric and non-stoichiometric (1010) and (1120) surfaces in 2H-SiC: a theoretical study. *Solid State Commun.* **111,** 459 (1999).

169. Aradi, B., Hourahine, B. & Frauenheim, T. DFTB+, a sparse matrix-based implementation of the DFTB method. *J. Phys. Chem. A* **111,** 5678 (2007).

170. Ma, X., Shi, L., He, X., *et al.* Graphitization resistance determines super hardness of lonsdaleite, nanotwinned and nanopolycrystalline diamond. *Carbon* **133,** 69 (2018).

171. Cretu, O., Botello-Mendez, A., Janowska, I., *et al.* Electrical Transport Measured in Atomic Carbon Chains. *Nano Lett.* **13.** PMID: 23879314, 3487 (2013).

172. Laasonen, K. & Nieminen, R. Molecular dynamics using the tight-binding approximation. *J. Phys. Condens. Matter* **2,** 1509 (1990).

173. Frenkel, D. & Smit, B. *Understanding Molecular Simulation. From Algorithms to Applications* (Academic, London, 1996).

174. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., *et al.* Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684 (1984).

175. Martyna, G. J., Klein, M. L. & Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **97,** 2635–2643 (1992).

176. Ravagnan, L., Manini, N., Cinquanta, E., *et al. Phys. Rev. Lett.* **102,** 245502 (2009).

177. Weimer, M., Hieringer, W., Della Sala, F., *et al.* Electronic and optical properties of functionalized carbon chains with the localized Hartree–Fock and conventional Kohn–Sham methods. *Chem. Phys.* **309,** 77–87 (2005).

178. Innocenti, F., Milani, A. & Castiglioni, C. Can Raman spectroscopy detect cumulenic structures of linear carbon chains? *J. Raman Spectrosc.* **41,** 226 (2010).

179. Ravagnan, L., Mazza, T., Bongiorno, G., *et al.* sp hybridization in free carbon nanoparticles–presence and stability observed by near edge X-ray absorption fine structure spectroscopy. *Chem. Commun.* **47,** 2952 (2011).

180. Zhang, G. P., Fang, X. W., Yao, Y. X., *et al.* Electronic structure and transport of a carbon chain between graphene nanoribbon leads. *J. Phys. Condens. Matter* **23,** 025302 (2011).

181. Kubo, R. Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems. *J. Phys. Soc. Jpn* **12,** 570 (1957).

182. Lee, C., Vanderbilt, D., Laasonen, K., *et al.* Ab initio studies on the structural and dynamical properties of ice. *Phys. Rev. B* **47,** 4863 (1993).

183. Meyer, R. & Comtesse, D. Vibrational density of states of silicon nanoparticles. *Phys. Rev. B* **83,** 014301 (2011).

184. Agarwal, N., Lucotti, A., Fazzi, D., *et al.* Structure and chain polarization of long polyynes investigated with infrared and Raman spectroscopy. *J. Raman Spectrosc.* **44,** 1398 (2013).

185. Milani, A., Tommasini, M., Russo, V., *et al.* Raman spectroscopy as a tool to investigate the structure and electronic properties of carbon-atom wires. *Beilstein J. Nanotechnol.* **6,** 480 (2015).

186. Casari, C., Tommasini, M., Tykwinski, R., *et al.* Carbon-atom wires: 1-D systems with tunable properties. *Nanoscale* **8,** 4414 (2016).

187. Stokes, H. T. & Hatch, D. M. *FINDSYM*: program for identifying the space-group symmetry of a crystal. *J. Appl. Crystallogr.* **38,** 237 (2005).

188. Narayan, J. & Bhaumik, A. Novel phase of carbon, ferromagnetism, and conversion into diamond. *J. Appl. Phys.* **118,** 215303 (2015).

189. Miwa, K. & Fukumoto, A. First-principles calculation of the structural, electronic, and vibrational properties of gallium nitride and aluminum nitride. *Phys. Rev. B* **48,** 7897 (1993).

190. Kawaji, H., Horie, H., Yamanaka, S., *et al.* Superconductivity in the Silicon Clathrate Compound $(Na, Ba)_x Si_{46}$. *Phys. Rev. Lett.* **74,** 1427 (1995).

191. Yamanaka, S. Silicon clathrates and carbon analogs: high pressure synthesis, structure, and superconductivity. *Dalton Trans.* **39,** 1901 (2010).

192. Dolyniuk, J., Owens-Baird, B., Wang, J., *et al.* Clathrate thermoelectrics. *Mater. Sci. Eng. R* **108,** 1 (2016).

193. Born, M. & Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **389,** 457 (1927).

194. Griffiths, D. J. & Schroeter, D. F. *Introduction to Quantum Mechanics* 3rd ed. (Cambridge University Press, 2018).

195. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev. B* **136,** 864 (1964).

196. Gross, E. & Kohn, W. in *Density Functional Theory of Many-Fermion Systems* 255 (Academic Press, 1990).

197. Cohen, M. L. & Louie, S. G. *Fundamentals of Condensed Matter Physics* (Cambridge University Press, 2016).

198. Grosso, G. & Parravicini, G. *Solid State Physics* (Elsevier Science, 2000).

199. Thomas, L. H. The calculation of atomic fields. *Math. Proc. Camb. Philos. Soc.* **23,** 542 (1927).

200. Fermi, E. Un metodo statistico per la determinazione di alcune proprietá dell'atomo. *Rend. Accad. Naz. Lincei* **6,** 602 (1927).

201. Teller, E. On the Stability of Molecules in the Thomas-Fermi Theory. *Rev. Mod. Phys.* **34,** 627 (1962).

202. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **140,** A1133 (1965).

203. Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica* **1,** 104 (1934).

204. Janak, J. F. Proof that $\frac{\partial E}{\partial n_i} = \epsilon$ in density-functional theory. *Phys. Rev. B* **18,** 7165 (1978).

205. Slater, J. C. in *Advances in quantum chemistry* 1–92 (Elsevier, 1972).

206. Almbladh, C.-O. & von Barth, U. Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues. *Phys. Rev. B* **31,** 3231 (1985).

207. Bredas, J.-L. Mind the gap! *Mater. Horiz.* **1,** 17 (2014).

208. Perdew, J. P. Accurate Density Functional for the Energy: Real-Space Cutoff of the Gradient Expansion for the Exchange Hole. *Phys. Rev. Lett.* **55,** 1665 (1985).

209. Perdew, J. P., Yang, W., Burke, K., *et al.* Understanding band gaps of solids in generalized Kohn–Sham theory. *Proc. Natl. Acad. Sci.* **114,** 2801 (2017).

210. Moritz, M. *Theory of Ultrafast Electron Transfer from Localized Quantum States at Surfaces* PhD thesis (2019).

211. Baerends, E. J., Gritsenko, O. V. & van Meer, R. The Kohn–Sham gap, the fundamental gap and the optical gap: the physical meaning of occupied and virtual Kohn–Sham orbital energies. *Phys. Chem. Chem. Phys.* **15,** 16408 (2013).

212. Runge, E. & Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **52,** 997 (1984).

213. Sousa, S. F., Fernandes, P. A. & Ramos, M. J. General Performance of Density Functionals. *J. Phys. Chem. A* **111,** 10439 (2007).

214. Ceperley, D. M. & Alder, B. J. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.* **45,** 566 (1980).

215. Oliver, G. L. & Perdew, J. P. Spin-density gradient expansion for the kinetic energy. *Phys. Rev. A* **20,** 397 (1979).

216. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7,** 3297 (2005).

217. Troullier, N. & Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **43,** 1993 (1991).

218. Herring, C. A New Method for Calculating Wave Functions in Crystals. *Phys. Rev.* **57,** 1169 (1940).

219. Phillips, J. C. & Kleinman, L. New Method for Calculating Wave Functions in Crystals and Molecules. *Phys. Rev.* **116,** 287 (1959).

220. Kleinman, L. & Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **48,** 1425 (1982).

221. Feynman, R. P. Forces in Molecules. *Phys. Rev.* **56,** 340 (1939).

222. Ashcroft, N. & Mermin, M. *Solid State Physics* (Holt-Saunders, Philadelphia, 1976).

223. Gonze, X. & Lee, C. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B* **55,** 10355 (1997).

224. Baroni, S., de Gironcoli, S., Dal Corso, A., *et al.* Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **73,** 515 (2001).

225. Parr, R. G. & Snyder, L. C. Perturbation Theory Calculations. *J. Chem. Phys.* **35,** 1898 (1961).

226. De Cicco, P. D., Johnson, F. A. & Smith, R. A. The quantum theory of lattice dynamics. IV. *Proc. Royal Soc. A. Math. Phys.* **310,** 111 (1969).

227. Sakurai, J. *Modern Quantum Mechanics* (Benjamin, Menlo Park, Calif., 1985).

228. Sternheimer, R. M. Electronic Polarizabilities of Ions from the Hartree-Fock Wave Functions. *Phys. Rev.* **96,** 951 (1954).

229. Gonze, X., Boulanger, P. & Côté, M. Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure. *Ann. Phys.* **523,** 168 (2010).

230. Mulliken, R. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. *J. Chem. Phys.* **23,** 1833 (1955).

231. Woods, R. D. & Saxon, D. S. Diffuse Surface Optical Model for Nucleon-Nuclei Scattering. *Phys. Rev.* **95,** 577 (1954).

232. Huran, A. W., Steigemann, C., Frauenheim, T., *et al.* Efficient Automatized Density-Functional Tight-Binding Parametrizations: Application to Group IV Elements. *J. Chem. Theory Comput.* **146,** 2947 (2018).

233. Wang, C. Z. & Ho, K. M. Tight-binding molecular dynamics for materials simulations. *J. Comput. Aided Mater.* **3,** 139 (1996).