



UNIVERSITY OF WARSAW

DOCTORAL THESIS

Multi-scale modeling of two-dimensional layered nanomaterials built of carbon, boron, and nitrogen atoms

Author:

Agnieszka JAMRÓZ

Supervisor:

Prof. dr hab. Jacek A.
MAJEWSKI

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

Faculty of Physics
Institute of Theoretical Physics

2022

*“There are only two mistakes one can make along the road to truth;
not going all the way, and not starting.”*

Buddha

Abstract

Multi-scale modeling of two-dimensional layered nanomaterials built of carbon, boron, and nitrogen atoms

Two-dimensional materials and their alloys remain under interest of scientists, due to their very appealing and conceivably tunable properties. $C_xB_yN_{1-x-y}$ graphene-like alloys constitute intriguing class of materials, including B- and/or N- doped graphene, C-doped hexagonal boron nitride (hBN), mixtures of graphene and hBN, or CBN structures with long-range order. Many reports on synthesis of such alloys are available, as well as numerous theoretical predictions for varied configurations of constituent C, B, N atoms. However, the deep understanding of the equilibrium morphology, stability, and phase diagrams of these interesting two-dimensional compounds is still lacking.

We address these issues in the present thesis. In the first step, we investigate the stability of the 2D ternary disordered C-B-N systems for the wide range of possible compositions, employing: (i) DFT-based Cluster Expansion technique, and (ii) direct valence force field approach for systems with thousands of atoms. We find the most energetically favourable configurations of C-B-N alloys via combination of (i) and (ii) with Monte Carlo Metropolis simulations. This allows us to avoid artificial periodicity, often introduced in *ab initio* simulations (when using small supercells). We find out that all investigated systems deviate strongly from random alloys and exhibit short-range ordering, *i.e.*, avoid creating N-N and B-B bonds in graphene lattice and have tendency to form graphene and hBN domains in ternary alloys.

We attempt to establish (in an iterative scheme) the effective Heisenberg-like hamiltonian parameters in Cluster Expansion models for graphene doped with N/B, and evaluate their quality. In order to use empirical potentials reasonably, we perform extensive tests of available VFFs including Tersoff, ReaxFF and COMB3 to find that none of them is fully satisfactory for simulations of CBN alloys. We combine and modify existing Tersoff parametrizations improving its performance.

Lastly, we show exemplary results for graphene flakes, and comment on Density of States of a chosen structures determined in Empirical Tight Binding approach.

Streszczenie

Wieloskalowe modelowanie dwuwymiarowych warstwowych materiałów zbudowanych z węgla, boru i azotu

Materiały dwuwymiarowe wykazują szereg intrygujących własności, które można w miarę łatwo modyfikować i projektować funkcjonalności w zależności od potrzeb. Rodzina układów grafeno-podobnych, zawierających węgiel, bor oraz azot, o ogólnym wzorze $C_xB_yN_{1-x-y}$ zawiera takie klasy materiałów, jak grafen domieszkowany borem lub/i azotem, heksagonalny azotek boru (hBN) domieszkowany węglem, jak również struktury CBN wykazujące porządek długiego zasięgu. W literaturze można znaleźć wiele prac dotyczących metod wytwarzania układów CBN, jak i teoretycznych przewidywań własności różnorodnych konfiguracji atomów C, B i N. Jednakże, wciąż nie mamy kompletnej wiedzy dotyczącej możliwości równowagowej syntezy tych związków, ich stabilności, czy przejść fazowych.

Te tematy poruszam w swojej pracy doktorskiej. Przeprowadzam teoretyczne badania stabilności stopów CBN dla szerokiego zakresu koncentracji poszczególnych atomów, w oparciu o metody (i) Cluster Expansion, oraz (ii) pól siłowych (VFF). Wyznaczam konfiguracje odpowiadające najniższym energiom wykorzystując podejścia (i) i (ii), w połączeniu z symulacjami Monte Carlo w układach liczących tysiące atomów. To pozwala na realistyczny opis nieperiodycznych struktur CBN. We wszystkich badanych układach, równowagowe konfiguracje odbiegają znacząco od losowych, wykazując natomiast porządek krótkiego zasięgu, tzn. wiązania N-N i B-B są niekorzystne i właściwie nie występują, zaś w trójskładnikowych stopach obserwuję tendencję do separacji na domeny grafenu i obszary hBN.

W szczególności wyznaczam parametry hamiltonianu w modelu Cluster Expansion dla grafenu domieszkowanego borem/azotem, oraz szacuję dokładność otrzymanych modeli. Seria testów dostępnych potencjałów empirycznych (Tersoff, ReaxFF i COMB3) pokazuje, że żadna z dostępnych parametryzacji nie daje całkiem zadowalających wyników w 2-wymiarowych stopach CBN. Łącząc i modyfikując różne parametryzacje potencjału Tersoffa, poprawiam jego dokładność.

Prezentuję również przykładowe wyniki dla płatków grafenowych oraz dyskutuję elektronową gęstość stanów dla wybranych struktur CBN.

Acknowledgements

First and foremost I am deeply grateful to my supervisor, prof. dr hab. Jacek Adam Majewski for his invaluable advice, continuous support, and patience during my PhD study. His knowledge and plentiful experience have encouraged me throughout my academic research. My gratitude extends to the Faculty of Physics for the funding opportunity to undertake my studies at the Institute of Theoretical Physics, University of Warsaw. I would also like to thank prof. Claudia Draxl and her team at Humboldt University in Berlin, prof. Aldo di Carlo and his team at the Tor Vergata University in Rome, prof. Paweł Hawrylak and his team at the University of Ottawa and prof. Adri van Duin and his team at the Pennsylvania State University for hosting me, sharing their knowledge and technical details of the methods they use.

This research has been supported by the National Science Centre, within PRELUDIUM grant (Contract No. UMO-2017/25/N/ST3/00660). I would like to acknowledge also the scholarship program: *Increasing the mobility of doctoral students at the University of Warsaw* that made possible many of my visits at the foreign universities.

Finally, I would like to express my gratitude to my family and friends, especially to my fiancé and little son, who was born during my PhD, for their unwavering support and belief in me.

Contents

Abstract	iii
Acknowledgements	vii
1 Motives to bother with CBN monolayer alloys	1
1.1 Chemistry of graphene and hBN based systems	3
1.2 Synthesis of CBN alloys - state of the art	5
1.3 Ordering in C-B-N alloys	6
1.4 Structural defects in C-B-N alloys	7
1.5 Theoretical results concerning CBN alloys	7
1.6 Applications of CBN structures	10
1.7 The main objective of the thesis	11
2 Short review of the long list of computational methods inspected	13
2.1 Look close - Density Functional Theory	14
2.1.1 All-electron vs. pseudopotential approach	17
2.2 Cluster Expansion	20
2.2.1 Cluster expansion formulation	20
2.2.2 Fitting the Cluster Expansion model	22
2.3 Empirical potentials	23
2.3.1 Tersoff potential	24
Extended Tersoff potential	25
2.3.2 Reactive force fields	26
ReaxFF potential	27
Charge Optimized Many Body potential	28
2.4 Schemes for description of ordering in CBN alloys	29
2.4.1 Monte Carlo Metropolis Method	29
2.4.2 Warren-Cowley Short Range Order Parameters	31
2.4.3 Definitions of Energy in heteroatomic systems	32

2.4.4	Entropy and Free energy	33
2.4.5	Empirical Tight Binding approach	35
3	First principles based simulations - Cluster Expansion approach	39
3.1	Cluster Expansion Iterative Technique	39
3.2	Reference DFT data for fitting CE models	42
3.3	Results of CE models determination	47
3.4	Results of the Cluster Expansion models for GL-N and GL-B on rigid lattice	54
4	Assesment of VFF methods for simulations of CBN alloys	59
4.1	Tersoff potential: available and developed parametrizations	59
4.2	ReaxFF: available parametrizations	62
4.3	COMB3: available parametrizations	63
4.4	Tests of the empirical potentials	64
4.5	Evaluation of VFF approaches in the view of CBN 2D alloys simulations	76
4.6	Conclusions: empirical potentials for simulations in solid state physics	83
5	Tersoff potential results on ordering in large scale CBN systems	85
5.1	Monte Carlo simulations of 2D CBN systems	86
5.1.1	Results of Monte Carlo simulations for C-B-N systems on rigid hexagonal lattice.	87
5.1.2	Entropy and Free Energy of Rigid CBN systems	93
5.2	Comparison of Tersoff and Cluster Expansion results	95
5.3	Note on the 2D simulations with relaxed geometry.	96
5.4	Density of states in CBN alloys	97
5.5	Graphene and hBN flakes	104
5.5.1	Boron doped graphene flakes	104
6	Summary and Outlook	107
A	Details of determination of the Cluster Expansion models	111
A.1	Clusters Pool	111
A.2	Evolution of CE models determination	112

List of Figures

1.1	Geometry and electronic structure of pristine graphene.	2
1.2	Comparison of the lattice parameters and band gaps for hBN-graphene and the family of nitrides (GaN, InN, AlN).	3
1.3	Triangle of possible C-B-N phases.	4
3.1	Clusters constituting initial training set (REFSET1) used for fitting the first iteration of Cluster Expansion models.	44
3.2	Mixing energies for N- and B-alloyed graphene structures constituting initial training set (REFSET1) for RIGID and RELAXED atomic positions as obtained from DFT calculations.	44
3.3	Side view of an exemplary corrugated relaxed structure from the REFSET1 for doping with (a) nitrogen (b) boron after geometry and lattice constant optimization.	46
3.4	Histograms showing the value of $R^{(i)}$ for all iterations of CE i model in all investigated classes of models.	48
3.5	Predictive power of consecutive iterations CE1-CE4 during iterative procedure of Cluster Expansion model fitting for N-doped graphene in rigid lattice.	50
3.6	Evolution of nitrogen distribution over the rigid carbon honeycomb lattice after performing MC simulations for progressing generations of Cluster Expansion models for exemplary concentrations of N. . . .	51
3.7	W-C parameters for N-N pairs extracted from minimum energy structures obtained in MC simulations with i^{th} generation of CE model on rigid lattice.	52
3.8	Average mixing energy for the GL-N and GL-B structures for two distributions of atoms: (i) random, and (ii) after Monte Carlo Metropolis simulations with Cluster Expansion models fitted to energies on rigid lattice.	55

3.9	Average entropy S_{pair} for the GL-N and GL-B structures random and after MC simulations with Cluster Expansion models fitted to energies on rigid lattice.	56
3.10	Dependence of average free energy on temperature for GL-N and GL-B alloys exhibiting short range order as obtained from MC simulations with respective CE models on rigid lattice.	57
3.11	Dependence of average free energy on temperature for random GL-N and GL-B structures calculated with respective CE models on rigid lattice.	57
4.1	Energies of X-Y dimers as a function of the bond length for X,Y = C, B, N.	66
4.2	Energy of pristine GL and hBN as a function of the lattice constant. .	68
4.3	Geometry of point defects in graphene and hBN.	71
4.4	Comparison of QE-DFT , Ters-SET2 , and Ters-SET3 energies of GL-N and GL-B structures (calculated on rigid lattice) constituting test set (4-i).	79
4.5	Comparison of QE-DFT , Ters-SET2 , and Ters-SET3 energies of GL-N and GL-B structures (calculated on relaxed lattice) constituting test set (4-ii).	80
4.6	Subset of set of ternary systems tested in (4-iii), consisting of graphene with four introduced B and N substitutions.	81
4.7	Comparison of QE-DFT , Ters-SET2 and Ters-SET3 energies (calculated on relaxed lattice) of GL-BN structures constituting subset of test set (4-iii) with 4 introduced substitutions.	81
4.8	Comparison of QE-DFT , Ters-SET2 , and Ters-SET3 energies as a function of interatomic distance for B-B molecule.	82
5.1	Evolution of nitrogen dopants distribution over the rigid honeycomb lattice.	88
5.2	Evolution of boron dopants distribution over the rigid honeycomb lattice	88
5.3	Evolution of boron and nitrogen dopants distribution over the rigid honeycomb lattice.	89
5.4	Evolution of carbon dopants distribution over the rigid hexagonal boron nitride honeycomb lattice.	89

5.5	Average energy for all investigated classes of CBN alloys, for random distributions and after MC runs.	90
5.6	Characteristic Warren-Cowley parameters for the 1 st – 4 th coordination shells in four classes of investigated CBN alloys.	91
5.7	Average configurational entropy in four classes of investigated CBN alloys before and after Monte Carlo simulations.	94
5.8	Average free energy for the lowest energy structures obtained in Monte Carlo simulations for graphene doped with nitrogen or boron.	95
5.9	Total and Projected Density of States (DOS) for GL-N ordered alloys.	99
5.10	Total and Projected Density of States (DOS) for GL-B ordered alloys.	100
5.11	Total and Projected Density of States (DOS) for $C_x(BN)_{1-x}$ alloys exhibiting short range ordering.	101
5.12	Density of States in the area of hBN band gap for $C_{0.1}(BN)_{0.9}$ alloy with various arrangements of atoms.	102
5.13	Value of the band gap for the whole range of concentration of carbon $C_x(BN)_{1-x}$	103
5.14	Top and side views of exemplary graphene and hBN flakes.	105
5.15	Boron distribution over the relaxed carbon hexagonal flake after performing Monte Carlo simulations for exemplary concentrations of B.	106
A.1	Clusters Pool used as a source of clusters for fitting of CE models considered in this thesis.	111
A.2	Predictive power of consecutive iterations CE1-CE3 during iterative procedure of Cluster Expansion model fitting for B-doped graphene in rigid lattice.	113
A.3	Evolution of boron distribution over the rigid carbon honeycomb lattice after performing MC simulations for progressing generations of Cluster Expansion models for exemplary concentrations of B.	114
A.4	W-C parameters for B-B pairs extracted from minimum energy structures obtained in MC simulations with i^{th} generation of CE model on rigid lattice.	114
A.5	Predictive power of consecutive iterations CE1-CE3 during iterative procedure of Cluster Expansion model fitting for N-doped graphene in relaxed lattice.	115

A.6 Evolution of nitrogen distribution over the relaxed carbon honeycomb lattice after performing MC simulations for progressing generations of Cluster Expansion models for exemplary concentrations of N. 116

A.7 W-C parameters for N-N pairs extracted from minimum energy structures obtained in MC simulations with i^{th} generation of CE model on relaxed lattice. 116

A.8 Predictive power of consecutive iterations CE1-CE3 during iterative procedure of Cluster Expansion model fitting for B-doped graphene in relaxed lattice. 117

A.9 Evolution of boron distribution over the relaxed carbon honeycomb lattice after performing MC simulations for progressing generations of Cluster Expansion models for exemplary concentrations of B. . . . 118

A.10 W-C parameters for B-B pairs extracted from minimum energy structures obtained in MC simulations with i^{th} generation of CE model on relaxed lattice. 118

List of Tables

3.1	Mixing alloy energies obtained from DFT calculations for GL-N and GL-B for rigid and relaxed atomic positions in clusters constituting initial training set.	46
3.2	Details of the Cluster Expansion models determined during the iterative procedure.	47
3.3	Cluster Expansion models CE1-CE4 for N-alloyed graphene on a rigid lattice.	49
4.1	Tersoff interatomic potential parameters following (Matsunaga <i>et al.</i> , 2000)	61
4.2	Extended Tersoff interatomic potential parameters.	61
4.3	Tersoff interatomic potential parameters constituting Ters-SET2 . . .	62
4.4	Equilibrium binding energy and bond length for dimer molecules built of C, B, and N atoms.	65
4.5	Equilibrium lattice constant and cohesive energy of graphene and monolayer hexagonal boron nitride.	68
4.6	Corrugation limit for GL and hBN systems.	69
4.7	Values of maximal chemical potential for C, N, and B, and BN pair energy in hBN.	72
4.8	Cohesive energies per atom, zero-formation energies, and the values of characteristic interatomic distances for various defects in graphene.	74
4.9	Cohesive energies per atom, formation energies, and the values of characteristic interatomic distances for various defects in hBN.	75
4.10	General performance of all the investigated empirical potentials in the tests presented in Sec. 4.4.	77
5.1	Parameters characterizing the Monte Carlo simulations for CBN systems on rigid hexagonal lattice.	86

- A.1 Cluster Expansion models CE1-CE3 for B-alloyed graphene on a rigid lattice. 113
- A.2 Cluster Expansion models CE1-CE3 for N-alloyed graphene on a relaxed lattice. 115
- A.3 Cluster Expansion models CE1-CE3 for B-alloyed graphene on a relaxed lattice. 117

List of Abbreviations

BW	B ragg- W illiams (approximation)
CBN	C arbon B oron N itride
CE(<i>i</i>)	(<i>i</i> th iteration of) C luster E xpansion (model)
COMB	C harge O ptimized M any B ody (potential)
CV	C ross V alidation (score)
CVD	C hemical V apor D eposition
DFT	D ensity F unctional T heory
(E)TB	(E mpirical) T ight B inding
GGA	G eneralized G radient A pproximation
hBN	hexagonal B oron N itride
hBN-C	hexagonal B oron N itride (doped with) C arbon
GL	G raphene L ayer
GL-B	G raphene L ayer doped with B oron
GL-N	G raphene L ayer doped with N itrogen
GNR	G raphene N ano R ibbon
HOMO	H ighest O ccupied M olecular O rbital
K-S	K ohn- S ham
LDA	L ocal D ensity A pproximation
LCAO	L inear C ombination of A tomc O rbitals
LUMO	L owest U noccupied M olecular O rbital
MC	M onte C arlo
MD	M olecular D ynamics
ML	M achine L earning
QE	Q uantum E spresso
PAW	P rojector A ugmented W ave (method)
ReaxFF	R eactive F orce F ield
REFSET_{<i>i</i>}	R E F erence S ET (of iteration) <i>i</i>
RMSE	R oot M ean S quare E rror

W-C SRO	Warren-Cowley Short Range Order (parameters)
VFF	Valence Force Field

Chapter 1

Motives to bother with CBN monolayer alloys

Two-dimensional structures remain under constant interest of both scientists and engineers, due to their relevance in both basic research and possible applications. The story began in 2004 at the University of Manchester, where the inaugural of this kind - well-known - graphene was first exfoliated (Novoselov *et al.*, 2004), and has kept surprising researchers with its properties since then. However, the family of reported monolayer materials grows very fast - the way of already synthesized or exfoliated 2D layered nanomaterials counts more than 100 chemical compounds (Koski *et al.*, 2013). Graphene (GL) itself has a very simple crystal structure, it consists of carbon atoms stacked in honeycomb hexagonal pattern in a single layer, as shown in Fig. 1.1(a) and (b). Consequences of such type of structure are quite fascinating. Graphene is a zero band-gap semiconductor - the valence and conduction bands touch at the K points of the Brillouin zone (so-called Dirac points), and, the dispersion relation near these points is linear, that is depicted in Fig. 1.1(c). This indicates extremely high charge carrier mobility, excellent electrical and thermal conductivity, high tensile strength and flexibility (Abergel *et al.*, 2010; Cooper *et al.*, 2012). Despite the plenitude of desirable features and possibilities of applications of pristine graphene, the lack of the band-gap makes it inappropriate for electronic device purposes, since there is no easy way to control current through such material. This is why many researches have been looking for precise and efficient way to modify graphene (*e.g.*, open the band-gap) to make it suitable for electronic applications, and at the same time avoiding big deterioration of the thermal and current conductivity, *etc.* One of the strategies is chemical doping with heteroatoms, *e.g.*,

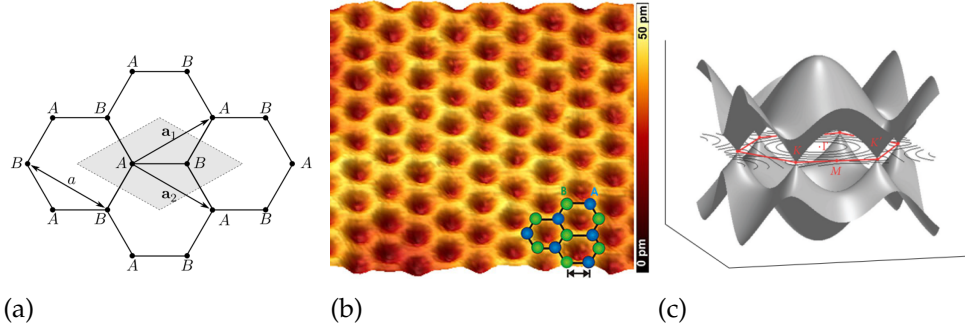


FIGURE 1.1: (a) structure of graphene, showing elementary cell containing two carbon atoms (*i.e.*, two triangular sublattices A and B) and unit cell vectors $\mathbf{a}_1$ and $\mathbf{a}_2$; (b) STM topography of atomically resolved graphene lattice (following www.nanoscience.de); (c) visualisation of valence and conduction bands of graphene (tight binding calculations, following (Cooper *et al.*, 2012)).

boron and nitrogen, that are the closest to carbon in periodic table. Such B- or/and N-doping of graphene is one of the methods to obtain monolayer material with a possibly tunable band gap. Alternatively, we can reach for and explore another member of 2D family, monolayer hexagonal boron nitride (hBN). This graphene counterpart is called sometimes *white graphene*. Structurally very similar, with each pair of C atoms replaced with B-N pair, thus making the two materials isoelectronic. However, due to the differences in electronegativity between boron and nitrogen atoms, the π electrons tend to localize around the nitrogen atomic centers, forming an insulating material, with a band gap ~ 6 eV (Elias *et al.*, 2019). Monolayer hBN was first synthesized via decomposition of borazine on the metal substrates (Nagashima *et al.*, 1995), long before the exfoliation of graphene. Since then, various methods of production *e.g.*, electron beam thinning of BN films (Meyer *et al.*, 2009), chemical vapor deposition (Sutter *et al.*, 2011) or exfoliation (C. Jin *et al.*, 2009) were developed. Substitutional doping of hBN with carbon is another option to obtain an alloy exhibiting a range of band gaps.

Bringing together above ideas, graphene and hexagonal boron nitride can be considered as building blocks of the rich and interesting family of layered CBN alloys (N. Kumar *et al.*, 2013; Song *et al.*, 2012; Yap, 2009). GL and hBN have the lattice mismatch $< 2\%$, as lateral lattice constant of GL is $2.46 \text{ \AA}$ (Cooper *et al.*, 2012) and for hBN - $2.504 \text{ \AA}$ (Lynch *et al.*, 1966). Combining the two, there is a chance to obtain structures varying their electrical properties from insulating, through semiconducting, to conducting, depending on the way of preparation, which in turn may lead

to various configuration of atoms in the alloys. Fig. 1.2 provides an illustrative comparison with the family of III-Nitrides with GaN, AlN, InN as border materials, that are commonly used in electronics. One can notice that lattice mismatch for the nitrides is much bigger, leading to problems with strain (and therefore - with stability) in those alloy materials. At the same time, CBN alloys exhibit wider range of band gaps possible to obtain, when compared to the nitrides. For that reason, layered CBN alloys are considered as obvious and promising candidates for diverse applications (as described in more detail in Sec. 1.6).

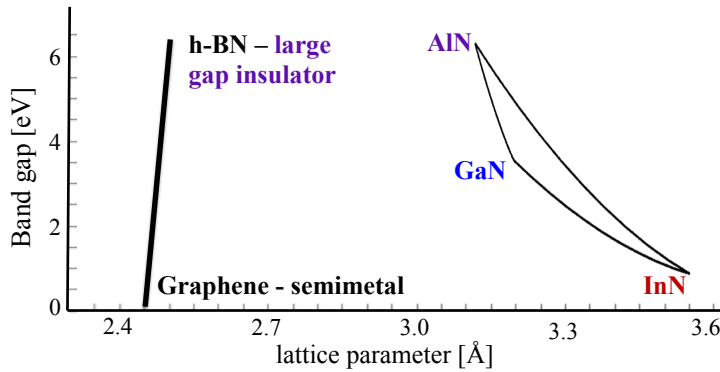


FIGURE 1.2: Comparison of the lattice parameters and band gaps for hBN-graphene and the family of nitrides (GaN, InN, AlN).

1.1 Chemistry of graphene and hBN based systems

Let's have a look at the chemistry behind CBN alloys, starting from the basic species - C, B and N. Carbon is the sixth element in the periodic table, with the electronic configuration $1s^2 2s^2 2p^2$. Two electrons fill the inner shell 1s, (close to atomic nucleus), and they do not contribute to creating interatomic bonds, while the remaining four are the valence electrons. There is a small difference in energy between the 2s and 2p levels in carbon, therefore, the electronic wave functions for these four electrons can be mixed with each other, forming so-called hybrid orbitals. Atoms in graphene exhibit sp^2 hybridization, *i.e.*, one 2s orbital is mixed with two of the 2p orbitals, let's say, p_x and p_y orbitals. The resulting orbitals lie in xy plane and make equal 120° angles. They point exactly towards the three neighbors, with which they create strong σ bonds. The fourth - p_z orbital - forms much weaker π bond with the other p_z orbitals, that is oriented out of the σ bonds plane.

The remaining species are the closest carbon neighbors: boron holds an electron less ($1s^2 2s^2 2p^1$), while nitrogen - one more ($1s^2 2s^2 2p^3$). As a consequence, B may serve as of p-type dopant of graphene, while N doping would create n-type material. Co-doping with both B and N leads to a manifold of configurations. Possible intermixing ratios as presented in triangle of CBN phases in Fig. 1.3. In our research we face only a subset of the CBN triangle: graphene alloyed with nitrogen, graphene alloyed with boron, graphene co-doped with boron and nitrogen with equal concentrations, and hexagonal boron nitride alloyed with carbon.

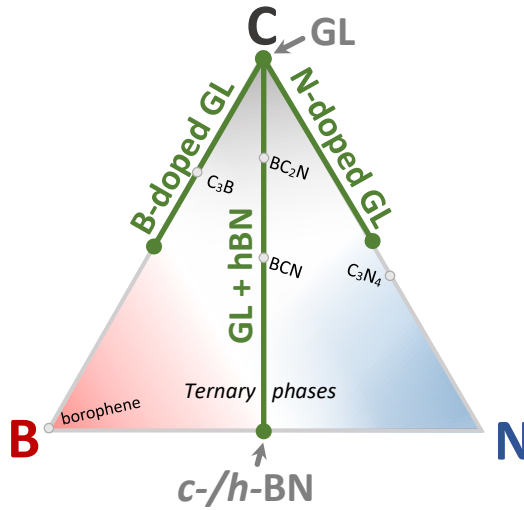


FIGURE 1.3: Triangle of possible C-B-N phases with areas embraced within this thesis marked green. Exemplary layered binary and ternary structures are pointed out.

The CBN ternary phase diagram (Fig. 1.3) is rich in stable binary and ternary phases, even so, the scientists are constantly searching for another ones. Exploring experimentally the vast amount of possible stoichiometries as well as configurations is quite difficult task, though. At this point, computer simulations can be a great help to determine possibly stable and interesting structures, that would be then synthesized in laboratory. In this thesis, we aim at finding configurations that give the most stable CBN hexagonal systems. Below, a short review is given of experimental and theoretical results concerning the C-B-N hexagonal systems, reported in the literature up-to-date.

1.2 Synthesis of CBN alloys - state of the art

Introduction of heteroatoms into graphene system modifies its charge distribution and electronic properties, leading to, *e.g.*, improved interactions with reactants. Boron-doped (GL-B), nitrogen-doped (GL-N), and co-doped structures (GL-BN) have received a lot of attention due to possibility to p- and n-dope of graphene systems. Co-doped graphene is particularly promising as it can exhibit synergistic effect of the heteroatoms. There are basically two ways to obtain doped graphene systems: direct synthesis or two-step approach, *i.e.*, synthesis of pristine graphene followed by post-synthesis doping.

In N-doped graphene, three bonding configurations of nitrogen can be found: pyridinic (N bonded to two C atoms at GL edge), pyrrolic (N being a part of 5-member ring at the edge), and graphitic N (nitrogen substitutes carbon). n-type doping effect was reported for graphitic nitrogen configuration (J. Zhang, C. Zhao, *et al.*, 2016). Synthesis methods include solid state synthesis in high temperature (Panchakarla *et al.*, 2009), electro-thermal reactions with ammonia (Y. Liu *et al.*, 2016), CVD growth (D. K. Sharma *et al.*, 2017) or low-energy implantation of N in graphene on a substrate (W. Zhao *et al.*, 2012). Direct method of assembling GL-N was reported in (Bundaleska *et al.*, 2018) with the use of microwave plasmas at atmospheric pressure conditions.

B-doping of graphene induces p-type behavior by downshifting Fermi level, also allowing for opening the band gap (Tang *et al.*, 2012). Synthesis methods include solid state synthesis in high temperature (Panchakarla *et al.*, 2009), CVD growth (D. K. Sharma *et al.*, 2017) or low-energy implantation of B (W. Zhao *et al.*, 2012). In (Tang *et al.*, 2012), application of microwave plasma to synthesize B-doped GL using trimethylboron was reported.

Simultaneous doping of graphene with B and N can further improve the quality of such alloy material, as B and N stabilize each other within graphene lattice. Synthesis methods include, *e.g.*, CVD (Bepete *et al.*, 2013), catalyst-free CVD synthesis (J. Jin *et al.*, 2014), thermal annealing (T.-W. Lin *et al.*, 2012) or copolymerization (X.-H. Li *et al.*, 2013). Co-doping of graphene with B and N leads often to formation hBN areas instead of uniform distribution of dopants due to the larger binding energies of B-N and C-C bonds compared to B-C and N-C bonds. However, (Zheng *et al.*,

2013) reported preparation of hBN-free BN-graphene by sequential incorporation of N and B into selected sites of the graphene.

On the other side C-doping can modify hBN behavior from insulating to conducting (X. Wei *et al.*, 2011). There are several ways to obtain carbon-containing hexagonal boron nitride monolayers, among them CVD (Han *et al.*, 2013), MOVPE and MBE (Mendelson *et al.*, 2021), pyrolysis of precursors in gas phase and solid state reaction (C. N. R. Rao *et al.*, 2017), topological substitution reaction converting graphene to h-BNC layers (Gong *et al.*, 2014), electron-beam mediated post-synthesis doping of boron-nitride with carbon (Krivanek *et al.*, 2010; X. Wei *et al.*, 2011).

To synthesize hexagonal carbon-boron-nitrides (h-CBN), the methane and ammonia borane can be introduced together to act as precursors for carbon and BN, respectively (Song *et al.*, 2012). In (Beniwal *et al.*, 2017) synthesis of h-CBN is held with the bis-BN cyclohexane, $B_2N_2C_2H_{12}$. B, C, and N atomic ratio can be typically modified to some extent by modifying the growth conditions, leading to samples of various homogeneity and morphology.

Comprehensive reviews of the doping techniques to obtain B- and/or N-doped graphene can be found in (Duan *et al.*, 2015; Feng *et al.*, 2020; C. Rao *et al.*, 2014), and also (Vesel *et al.*, 2020) for nitrogen doping, (R. Kumar *et al.*, 2021) for B-doping. The plethora of experimental research on h-CBN structures is reviewed, *e.g.*, in (J. Zhang, Sun, *et al.*, 2020).

1.3 Ordering in C-B-N alloys

Arrangement of dopants in both graphene and hBN depends strongly on the synthesis method, as well as the substrate used during material preparation. However, we have not encountered reports of aggregated B or N in graphene matrix. Some experimental results concerning N-doped graphene were reported, indicating that in CVD-grown samples dopants exhibit tendency to occupy only one sublattice (Lv *et al.*, 2012; Zabet-Khosousi *et al.*, 2014), while post-synthesis doping leads to a random distribution of nitrogen atoms between sublattices (Zabet-Khosousi *et al.*, 2014). For B-doping of graphene, sublattice segregation was also observed in samples after CVD growth on Co(0001) surface (Usachov *et al.*, 2016). Growth of B-GL on Cu(111) revealed in turn no doping sublattice asymmetry (L. Zhao *et al.*, 2013). The authors

of both papers pointed out possible significant influence of the substrate on positions of dopants.

In ternary CBN systems, researchers agree that there exists a driving force to cluster B and N atoms into hBN-C structures rather than leave B and N as isolated dopants distributed over the lattice (Schiros *et al.*, 2016). However hBN-free CBN alloy was successfully synthesized as reported in (Zheng *et al.*, 2013). Mixing and demixing in ternary CBN system, catalysed by the metal substrate, was studied in (Lu *et al.*, 2013).

1.4 Structural defects in C-B-N alloys

However, not only introduction of heteroatoms modifies quality and applicability of graphene/hBN/CBN. While considering theoretical crystal structure, we often neglect finite size of the system treating, *e.g.*, graphene as perfect infinite hexagonal structure, which does not seem to be a completely realistic assumption. Any kind of deviation from the perfect crystal is potentially meaningful for electronic structure, energy, or stability. Possible imperfections occurring in graphene-like structures include edges, vacancies, grain boundaries, 5-7 Throwing-Stones-Wales defects (Yang *et al.*, 2018; J. Zhang, Sun, *et al.*, 2020). Edges are always present in the experimentally investigated systems, while the other defects might appear in mutable concentration, sometimes neglectful. It is known that in graphene nanoribbons (GNRs), *i.e.*, strips of graphene with the width < 50 nm, edges significantly influence electronic structure, when compared to 2D graphene sheet. Zigzag-terminated GNRs possess states localized at edges with energies close to the Fermi level, while armchair-terminated ribbons do not exhibit such behavior (Wakabayashi, 2012). A thorough analysis on the topic of defects in graphene and graphene nanoribbons can be found, *e.g.*, in review article (Terrones *et al.*, 2012), where the subject is examined from both experimental and theoretical points of view.

1.5 Theoretical results concerning CBN alloys

There is a vast amount of computational results reported in the literature concerning CBN binary and ternary alloys, especially using *ab initio* DFT method, as they

often accompany experimental results or other computation techniques. Below, we mention several selected publications, corresponding to our area of interest, that employed the methods described and utilized in our research (they are described in following Chapter 2). Some of the results have been published during the PhD research presented in this thesis has been on the way and partly overlap with the topics we investigate.

***ab initio* methods**

- DFT study on point defects in graphene, *e.g.* (Abdullah *et al.*, 2020; Azevedo, 2006; M. Wu *et al.*, 2010).
- DFT study on point defects in hBN, *e.g.* (B. Huang *et al.*, 2012; Weston *et al.*, 2018).
- Mechanical and electronic properties of hBN nanosheets with graphene domains of various sizes and shapes under strain (Lima *et al.*, 2021).
- Structural and electronic properties of hBN/graphene interfaces in lateral heterostructures (J. Zhang, Xie, *et al.*, 2016).
- The optical absorbance and possibility of band gap engineering, phase separation, and composition fluctuations of $C_x(BN)_{1-x}$ alloys (Guilhon *et al.*, 2017).
- DFT and Monte Carlo investigation of the influence of interface geometry on phase stability and band gap engineering in BN-substituted graphene (D'Souza *et al.*, 2015).
- Theoretical proposition of a mechanism for fabrication of C-doped hBN by interaction of CO molecules with the vacancy in hBN (Z. Liu *et al.*, 2013).

Cluster Expansion

- Development of quasi-ternary Cluster Expansion (CE) model for C-BN alloys in recent (S. Du *et al.*, 2020) for $C_{2x}(BN)_{1-2x}$, in the search for enhanced solid solubility of these systems.
- CE model determined for graphene alloyed with nitrogen (Shi *et al.*, 2015), in the search for the limiting N concentration that could be introduced to graphene without destabilizing the system.
- Phase stability of CBN alloys with respect to hBN and graphene was examined by the combination of CE method and Monte Carlo simulations (Yuge, 2009).

ReaxFF

- Parametrization for carbon condensed phases (Srinivasan *et al.*, 2015) employed for Molecular Dynamics (MD) study of the thermal decomposition dynamics of a large fullerene (180 atoms).
- MD study of hydrogenated hybrid GL–hBN nanosheets for determination of their tensile properties (Dethan *et al.*, 2020).
- Simulations of thermal annealing in graphene oxide singly- and co-doped with boron and nitrogen (Bawari *et al.*, 2020).
- MD investigation of the effect of different hydroxyl regimes on the fracture and mechanical behaviour of hBN nanosheets (B. B. Sharma *et al.*, 2020).

COMB

- Calculations of chemical and physical properties of graphene oxide compared to results of other empirical potentials (Fonseca, Liang, D. Zhang, Choudhary, and Sinnott, 2016).
- MD study of graphene-titanium heterostructures, with and without substrates (Fonseca, Liang, D. Zhang, Choudhary, Phillpot, *et al.*, 2017).
- Survey of interactions of aluminium surfaces with graphene sheets in GL-aluminium nanostructures (D. Zhang *et al.*, 2019).

Tersoff potentials

- MD simulations of hybrid graphene-hBN interfaces, showing variation of thermal conductivity for stripe superlattices and BN dots in graphene (Kinaci *et al.*, 2012).
- Phase diagram determination for BN followed by MD simulations of N_2 molecule on the (100) surface of few-layer hBN.
- Introduction of extended Tersoff potential for BN based structures (Los *et al.*, 2017) followed by results of MD simulations of basic features of defect-free hBN nanosheet as well as elastic properties of defective hBN.
- Extended Tersoff potential MD study with the mechanisms of vacancy creation in monolayer and multilayer hBN under low- and high-fluence ion irradiation (Ghaderzadeh *et al.*, 2021).
- Investigation of mechanical properties of hybrid graphene-hBN sheets studied using MD with Tersoff potential (S. Zhao *et al.*, 2013).

Tight Binding approach

- Electric and transport properties calculations in lateral graphene-hBN heterostructures (Barrios-Vargas *et al.*, 2017).
- An overview of TB + DFT simulations for investigation of the role of the defects on electronic structure and transport properties of graphene (Botello-Mendez *et al.*, 2013).
- Self-consistent TB model for B and N doping in graphene (Pedersen *et al.*, 2013), assuming only differences in on-site energies between the species, yet giving some qualitative results.

1.6 Applications of CBN structures

CBN binary and ternary graphene-like alloys give us various prospects of applications, below we mention only few of them, with references to the extensive review papers.

- Doping graphene with nitrogen increases significantly adsorption of CO₂ on the graphene surface, which may be meaningful method in reduction of its excess in the atmosphere (Chandra *et al.*, 2012). N-doped graphene can be used also in the area of lithium ion batteries, supercapacitors, electrochemical devices, field-effect transistors, as revised in, *e.g.*, (Kaur *et al.*, 2018).
- B-doped graphene may serve as anode material for lithium ion batteries with super-high capacity (Z.-S. Wu *et al.*, 2011). It can be also applied in electrocatalysis, gas sensing, electrochemical sensing of photovoltaics as reviewed in, *e.g.*, (Agnoli *et al.*, 2016).
- CBN alloys may be utilized in, *e.g.* high-performance super-capacitors (Gopalakrishnan *et al.*, 2013; S. Wang *et al.*, 2018), in functional electronic and optoelectronic devices (Uddin, Majety, *et al.*, 2014), Li-Ion batteries (Bhat *et al.*, 2016), as effective and highly selective H₂ separation membranes (Ji *et al.*, 2016), or for a carbon capture (Chen *et al.*, 2018).
- B- and N-doped graphene has been investigated for the applications as removal of environmental contaminants through adsorption, catalysis, degradation, and detection (Feng *et al.*, 2020).
- Carbon impurities in hBN have been identified as the source of visible single-photon emission (Mendelson *et al.*, 2021).

1.7 The main objective of the thesis

The main purpose of this thesis is to gain deeper understanding of the mechanisms that determine structure and stability in monolayer, graphene-like materials containing carbon, boron, and nitrogen atoms. Particularly, we intend to investigate the role of the short range order in this family of materials as well as the properties resulting from varying concentrations and configurations of atoms in such CBN alloys.

Our hypothesis is that by choosing suitable concentration of constituent atoms in the CBN alloys and/or by proper functionalization one can control and modify their properties to such an extent that new physical phenomena emerge and they can be utilized in novel functional devices.

In the search for the most stable CBN alloys, we develop multi-scale theoretical scheme, based on methods and concepts well-known in condensed matter physics, which are described in the following Chapter 2. For energy calculations we employ Cluster Expansion (CE) method and Valence Force Fields (VFF). These approaches are tested and their accuracy is evaluated in Chapter 3, and Chapter 4, respectively. Throughout the thesis we also refer to multiple DFT calculations performed in small supercells, serving us as reference data. We present large-scale Monte Carlo simulations with the use of Tersoff potential in Chapter 5 in a wide range of concentrations of C, B and N atoms on hexagonal lattice, and find that all the systems in equilibrium exhibit short range order (SRO). We compare those results to respective Monte Carlo - Cluster Expansion calculations in binary alloys, and conclude that both give qualitatively the same picture of ordering. Clearly, during synthesis of CBN materials, scientists depart from equilibrium conditions, and use various substrates influencing CBN monolayer material, which explains various arrangements reported in experimental works.

Moreover, we reach out to the entropy and free energy calculations that allow us to discuss phase stability as well as possibility of obtaining random distributions of atoms in binary and ternary CBN alloys. We employ Empirical Tight Binding calculations of the Total and Projected Density of States, just to relate the electronic structure of CBN alloys with their morphology. As some kind of outline, we point out on the predictive power of the methodology developed in this study and signalize a possibility of it's usage to the problem of graphene and hBN nanoflakes.

Finally, in Chapter 6 we summarize the thesis, discuss the range of applicability of the methods employed here, and mention the emerging new applications of the tested approaches in various areas of the computational materials science.

Chapter 2

Short review of the long list of computational methods inspected

The fundamental equation of quantum mechanics is the Schrödinger equation

$$\mathcal{H} |\psi\rangle = E |\psi\rangle . \quad (2.1)$$

It allows for the most accurate description of the system of many interacting particles, *i.e.*, electrons and ions and including the influence of external fields. $\mathcal{H}$ is hamiltonian describing the interactions between the particles, and ψ - the wavefunction, that holds the information about all the $3N$ ($4N$ when we include spin) degrees of freedom of the N -particle system. In principle, it should be possible to describe the whole universe with such an equation. But, can you imagine solving it when the estimated number of atoms in the known, observable universe is between 10^{78} and 10^{82} ? Well, our universe is a pretty huge thing. Even when we limit ourselves to a small coffee table, the dimensionality of its wavefunction would overwhelm us. Furthermore, number of atoms in a *small* 20x20 nm graphene flake is roughly 30 000 atoms. It is really a lot for the quantum mechanical simulations. Therefore, other, more approximate methods are in high demand when we want to learn about the properties of *realistic* systems.

In this chapter we would like to present the tools that we have used during our research. We will start from methods for calculations of energy and geometry with the Density Functional Theory in Sec. 2.1, then zoom out and see Cluster Expansion

technique in Sec. 2.2, and zoom out further to discuss empirical Valence Force Fields in Sec. 2.3. In second part of the chapter, Sec. 2.4, we shortly introduce Monte Carlo approach that is used to determine the equilibrium alloy configuration in the case of Cluster Expansion as well as in the VFF approaches. It is followed by the definitions of Warren-Cowley Short-Range Order parameters, and various energies that we will refer to throughout the thesis, configurational entropy and free energy. We also mention Empirical Tight Binding method, which allows for computation of the density of states of the investigated structures.

2.1 Look close - Density Functional Theory

A minor, yet crucial, part of the results documented in this thesis have been obtained in the framework of Density Functional Theory (DFT) in the Hohenberg-Kohn-Sham (K-S) formulation. It is by far the most widely used method of *ab initio* calculations nowadays.

Ab initio from Latin means *from first principles, from the beginning*. The basic idea of quantum mechanics is to use only universal physical constants in the description of the system. Subsequently, all the calculated properties are based on that, not requiring higher-order parameters. The most general (non-relativistic) description of many-body interacting system consisting of atomic nuclei and electrons can be written with the Hamiltonian operator $\mathcal{H}$

$$\mathcal{H} = \mathcal{T}_e + \mathcal{T}_n + \mathcal{V}_{ee} + \mathcal{V}_{ne} + \mathcal{V}_{nn}, \quad (2.2)$$

which characterizes its total energy and time evolution. It is a sum of the operators describing: kinetic energy of the electrons $\mathcal{T}_e$, kinetic energy of the nuclei $\mathcal{T}_n$, Coulomb interaction between the electrons $\mathcal{V}_{ee}$, Coulomb interaction between the electrons and the nuclei $\mathcal{V}_{ne}$ and Coulomb interaction between the nuclei $\mathcal{V}_{nn}$. In the analysis of electrons' behavior, nuclei, much heavier than the electrons, are treated typically as fixed (Born-Oppenheimer approximation), and thus generating a constant external potential for the electrons

$$v_{ext}(\mathbf{r}) = - \sum_{I=1}^M \frac{Z_I e^2}{4\pi\epsilon_0 |\mathbf{R}_I - \mathbf{r}|}, \quad (2.3)$$

where $\mathbf{R}_I$ is position of I^{th} nucleus having charge $Z_I e$. The sum is running over M nuclei in the system. Then, the hamiltonian of N electrons in the field of M nuclei might be written as

$$\mathcal{H} = \sum_i^N \left(-\frac{\hbar^2}{2m_e} \nabla_i^2 + v_{ext}(\mathbf{r}_i) \right) + \sum_{i<j}^N \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.4)$$

where the last term describes interaction between the electrons. Exact solution of the Schrödinger equation with such a hamiltonian is not possible even for two ($N = 2$) electrons. Therefore, there has been a thorough need to simplify the problem for a system of many interacting electrons. The idea of using single-particle density of electrons $\rho(\mathbf{r})$ instead of their many-particle wavefunction $\Psi(x_1, x_2, \dots, x_N)$, where $x_i \equiv (\mathbf{r}_i, \sigma_i)$ to describe the properties of electronic systems appeared in the papers (Hohenberg *et al.*, 1964) and (Kohn *et al.*, 1965). The authors proved as masterly as simple that ρ and Ψ both carry a full information about the ground state of the system of interacting electrons in an external potential v_{ext} (which can be generated by the nuclei, and also by external electric and magnetic fields). Moreover, the total energy could be written as a functional of the density $E[\rho(\mathbf{r})]$, and it would be minimized by the density of the ground state (analogous to variational theorem from quantum mechanics). The functional can be written as a sum of the part dependent on v_{ext} and universal terms dependent only on the density ρ

$$E[\rho] = T[\rho] + E_{Coul}[\rho] + E_{XC}[\rho] + \int v_{ext}(\tau) \rho(\tau) d\tau, \quad (2.5)$$

where T is a kinetic energy of the interacting electrons, E_{Coul} is a Coulomb repulsive interaction, and E_{XC} - remaining, unknown explicitly, exchange-correlation functional. In the next, significant step, Kohn and Sham introduced the fictitious system of N non-interacting electrons that would provide the one particle density ρ_s , identical to the one particle density ρ of the N interacting electrons. The density of the non-interacting electrons can be always written as the sum of orbital densities $\phi_i^*(\mathbf{r})\phi_i(\mathbf{r})$

$$\rho_s(\mathbf{r}) = \sum_{i=1}^{occ} \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}), \quad (2.6)$$

and because of the requirement $\rho_s = \rho$, ρ_s also integrates to N

$$\int \rho_s(\mathbf{r}) d^3r = N \quad (2.7)$$

The one-particle orbitals ϕ_i can be determined through the minimization of $E[\rho]$ as defined in Eq. (2.5). As it has been shown in (Kohn *et al.*, 1965), the orbitals ϕ_i fulfill the Schrödinger equation

$$\left[\frac{\hbar^2}{2m_e} \nabla_i^2 + v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_{XC}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i, \quad (2.8)$$

where

$$v_H(\mathbf{r}) = \frac{\delta E_{Coul}[\rho]}{\delta \rho(\mathbf{r})} = \int \frac{e^2 \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \quad (2.9)$$

is the Coulomb potential of the charge distribution determined by $\rho(\mathbf{r})$, and the exchange - correlation potential

$$v_{XC} = \frac{\delta E_{XC}[\rho]}{\delta \rho(\mathbf{r})} \quad (2.10)$$

is the functional derivative of $E_{XC}[\rho]$ with respect to the density. The effective potential describing the dynamics of non-interacting electrons, so called Kohn-Sham potential $v_{K-S}(\mathbf{r}) = v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_{XC}(\mathbf{r})$ is a local potential, however, it is functionally dependent on ρ . Therefore, the equations Eq. (2.8) for ϕ_i and Equations (2.6) and (2.7) that define ρ_s must be solved iteratively in a self-consistent procedure.

Kohn-Sham realization of the DFT clearly demonstrates that the mapping of the system of N -interacting electrons onto a system of N -non-interacting particles that move in an effective potential $v_{K-S}(\mathbf{r})$ is rigorously possible. The Kohn-Sham scheme paved the road towards simple computational method for interacting electrons. However, the main obstacle of the K-S method is the fact that the exact correlation energy functional $E_C[\rho]$ being the smaller part of $E_{XC}[\rho] = E_X[\rho] + E_C[\rho]$, is not known, and must be approximated in the practical applications.

In more than 50 years since the first formulation of K-S method, a various approximate $E_C[\rho]$ functionals have been developed. Albeit the exact exchange energy functional $E_X[\rho]$ is known, in many computational schemes $E_X[\rho]$ is also approximated, just to reduce computational burden.

Nowadays, above than 50 approximations for $E_X[\rho]$ and $E_C[\rho]$ have been developed, so the detailed discussion of them is outside the scope of this thesis and could be

found in some reviews (see, *e.g.*, Becke, 2014; Mardirossian *et al.*, 2017). Typically, existing computer packages implementing K-S method provide numerous options for $E_{XC}[\rho]$. The Local Density Approximation (LDA) is the basic one and it reads

$$E_{XC}^{LDA}[\rho] = \int \rho(\mathbf{r}) \epsilon_{XC}(\rho) d^3r. \quad (2.11)$$

Here we assume that locally (*i.e.*, at a given $\mathbf{r}$) the exchange-correlation energy of the non-homogeneous electron system of density $\rho(\mathbf{r})$ is equal to the exchange-correlation energy of the homogeneous electron gas ϵ_{XC} of the same density. On the basis of E_{XC}^{LDA} , one can develop the Generalized Gradient Approximation (GGA) as follows

$$E_{XC}^{GGA}[\rho] = \int \rho(\mathbf{r}) f_{XC}(\rho, \nabla \rho) d^3r, \quad (2.12)$$

where $f_{XC}(\rho, \nabla \rho)$ is a function of the one-particle density and the gradient of this density. In the DFT calculations performed in this thesis, typically so-called PBE (Perdew *et al.*, 1996) form of the GGA functional has been used, that became kind of standard nowadays.

2.1.1 All-electron vs. pseudopotential approach

In principle, the electronic Hamiltonian of Eq. (2.4) contains all electrons of the atoms constituting a molecule, nanostructure, or solid. By all means, in periodic systems like solids, only electrons in a primitive cell should be considered. In literature, there exist several numerical codes that provide solutions of the so called *all-electron* problems in the framework of the K-S method, such as WIEN2K, *exciting!*, or FHI-aims. Employment of an all-electron approach would be essential, *e.g.*, for the studies of photoemission phenomena, where the deep states of core electrons matter. On the other hand, most of the properties of molecules and solids, including their equilibrium geometry, are determined by electronic states with energies around the Fermi level. These states originate from the valence electrons of atomic outer shells.

Since the computational burden of K-S method is of the order $\mathcal{O}(N^3)$, restricting the calculations only to valence electrons would be extremely beneficial. The Z (atomic number) electrons of a neutral atom might be divided into N_c^{at} core electrons, and N_v^{at} valence electrons ($Z = N_c^{at} + N_v^{at}$). Even if for carbon atom Z^C and N_v^C differ

only by 2 (two 1s core electrons), the differences for heavier atoms are substantial, e.g., for antimony $Z^{Sb} = 51$, while $N_v^{Sb} = 5$.

K-S scheme can be applied then, in the M -atomic system, for the $N_v = \sum_{I=1}^M N_v^I$, where N_v^I is a number of valence electrons in atom I . These valence electrons move in the field of atomic ions with positive charges diminished by the number of core electrons included into the ion $\tilde{Z}^I = Z^I - N_c^I$.

However, the interaction of valence electrons with such ions consisting of atomic nucleus and core electrons is not given by the simple Coulomb potential of Eq. (2.9). It should be rather described by an effective potential, named pseudopotential, that has to be constructed according to a well defined prescription. By all means, there is some degree of arbitrariness in 'pseudizing' the electron problem and building pseudopotentials, and there exists vast literature reporting differing procedures and the resulting pseudopotentials. The most relevant three categories are norm-conserving pseudopotentials (Hamann *et al.*, 1979), ultra-soft pseudopotentials (Vanderbilt, 1990), and Projector Augmented Waves (PAW) pseudopotentials (Blöchl, 1994). Common pseudopotential codes (e.g., SIESTA, ABINIT, Quantum Espresso, VASP) allow for calculations with the types of pseudopotentials mentioned above, and additionally provide them for Periodic Table in the numerical package.

In the Kohn-Sham approach, the sum of ionic pseudopotentials of all ions plays the role of external potential $v_{ext}(\mathbf{r}) \rightarrow V_{ion}^{pp}$, and the one particle electron density in the functionals is just the one-particle density of valence electrons (or pseudo-electrons).

For an electronic system such as molecule or solid, the choice of approximation to the exchange-correlation potential E_{XC} and the choice of pseudopotential, determines the accuracy of the calculations. The collected experience in DFT-K-S calculations shows that there are no unique, the most accurate E_{XC} functionals and pseudopotentials, and the accuracy of these two depends unfortunately on the electronic system studied (Mardirossian *et al.*, 2017). This fact should be kept in mind assessing the obtained results of the calculations. The other important issue is the precision with which we solve K-S equations, i.e., Eq. (2.8) for a given E_{XC} and pseudopotential.

Usually K-S equations for K-S orbitals ϕ_i are solved in three ways:

- ϕ_i are expanded in a set of atomic like orbitals,
- ϕ_i are expanded into set of plane waves, and
- ϕ_i are solved directly on a mesh in a real space.

For example, the first method is used in SIESTA, OPENMX, the second in ABINIT, Quantum Espresso, VASP, CPMD, CASTEP, PHASE, STATE, xTAPP, QMAS, and the third one in PARSEC, GPAW, RSDFT. Each of these methods has its advantages and drawbacks, that are described in the literature. The important point is that it is known how to control the precision of calculation through a suitable choice of the parameters.

The most commonly used codes employ plane-wave expansion of Kohn-Sham orbitals ϕ_i . It implies the periodicity of the system in all three space dimensions (it is built in the codes). In order to calculate the system lacking the periodicity in any dimension one has to place the system in the specific supercell. The lattice constant of nonphysical, artificial periodicity should be chosen in such a way that the spurious interaction between the system and its images in the neighboring cells is negligible.

The last issue we would like to address in this short discussion of the DFT based simulations is the possibility to optimize the system geometry. The calculated total Kohn-Sham energy of electrons E_{tot}^{el} is parametrically dependent on the ionic positions $\mathbf{R}_I$. The total energy of the system can be written in the form

$$E_{tot}(\mathbf{R}_I) = E_{tot}^{el}(\mathbf{R}_I) + \sum_{I < J} \frac{e^2 Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (2.13)$$

The optimized geometry is found when all the forces on ions in the system disappear, *i.e.*, when for any I

$$\frac{\delta E_{tot}(\mathbf{R}_I)}{\delta \mathbf{R}_I} = 0 \quad (2.14)$$

The procedure for finding the optimized geometry is implemented in all numerical codes mentioned above.

In this section, we have provided short and fairly general description of the *ab initio* calculations in the framework of Density Functional Theory. The details of the

computational schemes that were employed in this thesis are given and discussed in Chapters 3 to 5 where the results of DFT simulations are presented.

2.2 Cluster Expansion

Among the other methods, first-principles calculations, (*e.g.*, DFT, as described in previous section) are the most accurate for determining ground state of the systems without conducting classical experiments. Nowadays (despite the increasing computational powers available in materials science), a typical researcher has a possibility to perform simulations systems of several hundreds atoms. This limits substantially the use of *ab initio* methods for the investigation of the systems' thermodynamics. Furthermore, in order to calculate the phase diagram, adequate sampling of the configuration space is a prerequisite, which is particularly relevant in the multicomponent systems, to which present thesis is devoted. The ordering phenomena and phase transitions are directly related in such systems.

Therefore, in order to access precise thermodynamic functions, we need to reach out to the method that includes electronic effects with nearly *ab initio* accuracy, but allows for calculations of relatively large systems. With regard to this, Cluster Expansion method, especially combined with Monte Carlo simulations, can serve as a powerful tool, providing a great balance between computational efficiency and accuracy. It proved its usefulness in many studies, when CE model is mapping total or mixing energy of the systems, *e.g.*, for analyzing phase diagrams in both metallic and semiconducting alloys (Ozoliņš *et al.*, 1998; Zhou *et al.*, 2006), surfaces (Herder *et al.*, 2015), and nanoparticles (Yuge, 2011). It can be also applied in modelling of vibrational (Walle *et al.*, 2002) or transport properties (Ångqvist *et al.*, 2017).

2.2.1 Cluster expansion formulation

Cluster Expansion is aimed at describing the variation of a given property Q (very often it is the energy of the system) with the chemical configuration. It was first proposed by Meyer (Mayer *et al.*, 1941). Let's define the occupation vector $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$, of the length N - number of atoms in the system. Any site i at the position r_i can be populated with one of the M_i species. Then, σ_i is an integer number in the range $(0, \dots, M_i - 1)$. It was shown (Sanchez *et al.*, 1984) that the set of

complete and orthonormal basis functions $\Pi_\alpha(\sigma)$ can be constructed (so called *cluster functions*), and that any function of configuration can be expanded in terms of these:

$$Q(\sigma) = \sum_{\alpha} J_{\alpha} \Pi_{\alpha}(\sigma), \quad (2.15)$$

with J_{α} being expansion coefficients, that are called *Effective Cluster Interactions (ECIs)*.

Very often when one considers two component system, and one assign a set of ‘spin’ variables $\hat{S}_i$ to each of the lattice sites ($\hat{S}_i = \pm 1$), and then the property Q is considered to be energy, the equation Eq. (2.15) is written in an equivalent form resembling the shape of Ising-like hamiltonian

$$Q(\sigma) \equiv E(\sigma) = J_0 + \sum_i J_i \hat{S}_i(\sigma) + \sum_{j < i} J_{ij} \hat{S}_i(\sigma) \hat{S}_j(\sigma) + \sum_{i < j < k} J_{ijk} \hat{S}_i(\sigma) \hat{S}_j(\sigma) \hat{S}_k(\sigma) + \dots \quad (2.16)$$

for configuration σ , where J ’s are ‘interaction energies’. The first summation runs over the all lattice points in the lattice, the second summation over all pairs of lattice points, the third summation over all triplets, and so on. The primary advantage of such cluster expansion is that interaction energies J are independent on the configurations σ .

Here, by the term *cluster* we understand the aggregation of atoms that can present the effect to the given function Q . Each $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)$ tallies with a cluster, as $\alpha_i \neq 0$ only for the points belonging to the cluster. Furthermore, cluster functions can be written as a product of orthogonal *point functions* $\Theta_{M_i, \alpha_i}(\sigma_i)$, $\alpha_i, \sigma_i \in \{0, 1, \dots, M_i - 1\}$

$$\Pi_{\alpha}(\sigma) = \prod_{i=1}^N \Theta_{M_i, \alpha_i}(\sigma_i), \quad (2.17)$$

which fill the requirements: (a) $\Theta_{M_i, 0}(\sigma_i) = 1$, and (b) orthogonality condition $M^{-1} \sum_{\sigma=0}^{M-1} \Theta_{M, \alpha}(\sigma) \Theta_{M, \beta}(\sigma) = \delta_{\alpha\beta}$. Several choices of the point cluster functions can be made, *e.g.*, trigonometric functions or Chebyshev polynomials, see *e.g.*, (Sanchez *et al.*, 1984; van de Walle, 2009; Wolverton *et al.*, 1994). Note that for the clusters α, β related by the symmetry operations, corresponding *ECIs* would be equal $J_{\alpha} = J_{\beta}$. Then, *cluster correlation function* X_{α} can be defined - average of the cluster functions belonging to the orbit $O(\alpha)$ of cluster α (*i.e.*, over all cluster functions of clusters

symmetrically equivalent to α)

$$X_\alpha(\sigma) = \frac{1}{m_\alpha} \sum_{\beta \in O(\alpha)} \Pi_\beta(\sigma), \quad (2.18)$$

with m_α denoting multiplicity of the cluster. Then, the summation of Eq. (2.15) can be carried out over symmetrically inequivalent (s.i.) clusters

$$Q(\sigma) = \sum_{\alpha}^{\text{s.i.}} J_\alpha m_\alpha X_\alpha(\sigma). \quad (2.19)$$

2.2.2 Fitting the Cluster Expansion model

The main task in creating CE model is then the determination of J_α values. Luckily, the shortsightedness of the electronic and ionic interactions strongly suggests that only several of them would be non-negligible. The procedure of choosing a set of clusters that give meaningful contributions and fitting the *ECIs* values are often performed using machine learning. One needs to prepare a *training set* of N structures $\{\sigma_1, \dots, \sigma_N\}$ with, *e.g.*, energies $E_i^{\text{target}}, i \in \{1, 2, \dots, N\}$ that are known from the first principles calculations. It is important to notice that well converged *ab initio* calculations are vital to obtain CE model with a high predictive power. Optimal *ECIs* can be found by minimizing the *objective function*

$$S^2 = \frac{1}{N} \sum_{i=1}^N (E_i^{\text{target}} - E^{\text{CE}}(\sigma_i))^2 + \phi(\{J_\alpha\}), \quad (2.20)$$

where $E^{\text{CE}}(\sigma_i)$ denotes the energy of i^{th} structure from the training set calculated with a given CE model. The first term in Eq. (2.20) is a mean squared error (MSE), while the second is a penalty function term that constrains *ECIs* in order to reduce overfitting. Specific form of ϕ depends on the chosen method. There are two prevalent approaches currently employed, Bayesian framework (Mueller *et al.*, 2009), and compressed sensing (Leong *et al.*, 2019; Nelson *et al.*, 2013). The fitting accuracy of any model should be verified. Root Mean Square Error (RMSE) and Cross Validation (CV) score are typically used to evaluate the predictive power of cluster

expansion model

$$RMSE^2 = \frac{1}{N} \sum_{i=1}^N (E_i^{\text{target}} - E^{\text{CE}}(\sigma_i))^2, \quad CV^2 = \frac{1}{N_t} \sum_{i=1}^{N_t} (E_i^{\text{target}} - \hat{E}(\sigma_i))^2 \quad (2.21)$$

Here, $\hat{E}(S_i)$ refers to predicted value of S_i structure energy obtained from CE model fitted to reference training set excluding the data point σ_i , and then $N_t = N - 1$. If the resulting CE model is not satisfactory, one can perform the fitting in an iterative way (Q. Wu *et al.*, 2016).

Once the *ECIs* values are known, the standard methods of statistical mechanics can be utilized to analyze the thermodynamic properties of the systems. Coupled with Monte Carlo simulations, can be utilized for ascertaining the phase diagrams, diffusion, concentration distribution, *etc.*.

During preparation of this thesis, we have determined the CE models for C-B and C-N systems. The specific methodology utilized there and the results can be found in Chapter 3.

2.3 Empirical potentials

Now, we take a step back, to look at the wider picture. We do not see atom as a structure containing nuclei and electrons moving and interacting with each other anymore. We see atom as a whole, with defined position and momentum (and sometimes charge) at any time point, simply from the classical point of view. Let's define interatomic potential $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ that describes interactions between N atoms having positions $\mathbf{r}_1, \dots, \mathbf{r}_N$. The form of the potential may be varied, but we want it to give predictions of the energy of the systems consistent with either experimental or *ab initio* results. In general, one can express full potential as an infinite sum of n-body terms

$$\begin{aligned} V(\mathbf{r}_1, \dots, \mathbf{r}_N) = & \sum_i V_1(\mathbf{r}_i) + \sum_{i,j} V_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i,j,k} V_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \\ & + \sum_{i,j,k,\dots,n} V_n(\mathbf{r}_i, \dots, \mathbf{r}_n) + \dots, \end{aligned} \quad (2.22)$$

but for practical reasons, it is truncated at some point. The first term describes the energies of individual atoms, while the subsequent components give the interaction energy as a function of atomic positions. Usually, it is sufficient to consider only two- and three-body (sometimes four-body) terms, as further components give negligible contributions. Clearly, such an approach is less accurate than *ab initio* methods. First, it maps complicated quantum mechanical interaction onto rather simple classical potentials, and the parameters present in the formula are by definition fitted to the results obtained elsewhere. The quality of this fit determines qualitative and quantitative agreement with other calculations and experiment, as well as flexibility and transferability of a given potential. In the following paragraphs, we describe shortly three empirical formulas that we take into consideration in the present thesis. All of them belong to the family of bond-order potentials. Bond order terms reflect the idea that the strength of any chemical bond depends on the bonding environment, *i.e.*, the number of all other bonds, their lengths, and angles between them. They are meant to describe dynamical processes, *e.g.*, formation and breaking the bonds, as well as characterize multiple bonding types.

2.3.1 Tersoff potential

Tersoff potential has been widely used in solid state simulations especially for covalent systems for many years. The formula for pure elemental systems was presented in (Tersoff, 1988) and further extension for multicomponent materials was given in (Tersoff, 1989). In Jerry Tersoff's approach, potentials for individual species (historically it was Si, C, and Ge) are determined independently by fitting to respective elemental data. Then, auxiliary parameters for each pair of elements are added, for treating mixtures of these elements properly. Below, we present the full Tersoff potential formula intended for multicomponent systems, that is used in the thesis. The total energy of the system is described as a sum over two center interatomic interactions V_{ij} between atomic pairs i and j separated by the distance r_{ij}

$$V_{ij} = f_C(r_{ij}) (f_R(r_{ij}) + b_{ij}f_A(r_{ij})), \quad (2.23)$$

where

$$f_R(r_{ij}) = A_{ij}e^{-\lambda_{ij}r_{ij}}, \quad f_A(r_{ij}) = -B_{ij}e^{-\mu_{ij}r_{ij}}, \quad (2.24)$$

$$f_C(r_{ij}) = \begin{cases} 1, & r_{ij} < R_{ij} \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{\pi(r_{ij}-R_{ij})}{(S_{ij}-R_{ij})}\right), & R_{ij} < r_{ij} < S_{ij} \\ 0, & r_{ij} > S_{ij} \end{cases} \quad (2.25)$$

f_R and f_A denote repulsive and attractive interaction functions, respectively, while f_C is a cut-off function ensuring potential to be short range. R_{ij} and S_{ij} are typically chosen to include only the first shell of neighbors. The bond-order term b_{ij} describes the influence of the neighborhood on the binding strength of a pair of atoms by modifying the attractive part of the potential

$$b_{ij} = \chi_{ij} \left(1 + \zeta_{ij}^{n_i}\right)^{-\frac{1}{2n_i}}, \quad (2.26)$$

$$\zeta_{ij} = \sum_{k \neq i,j} f_C(r_{ik}) \omega_{ik} \beta_i g(\theta_{ijk}), \quad (2.27)$$

$$g(\theta_{ijk}) = 1 + \frac{c_i^2}{d_i^2} - \frac{c_i^2}{d_i^2 + (h_i - \cos(\theta_{ijk}))^2}, \quad (2.28)$$

where θ_{ijk} is the angle between the atomic pairs ij and ik bonds. Altogether, there are 13 various parameters in the Tersoff potential: A_{ij} , B_{ij} , λ_{ij} , μ_{ij} , R_{ij} , S_{ij} , χ_{ij} , n_i , ω_{ij} , β_i , c_i , d_i , and h_i . Values of the parameters are determined during the fitting procedure, with carefully chosen set of reference data. Initially, Tersoff's approach to obtain parameters for heteroatomic interactions A , B , R , S , λ and μ was by averaging values of parameters for respective elements, yet in further parametrizations for multicomponent systems individual sets for all types of $i - j$ interactions were presented as well.

The discussion concerning available parametrizations of the Tersoff potential, as well as evaluation of suitability of the various sets for our research is given in Sec. 4.1.

Extended Tersoff potential

There were many attempts to improve performance of the Tersoff potential. We will though focus on recent extended Tersoff potential (Los *et al.*, 2017), which is particularly interesting for the purpose of this thesis. This modification of original Tersoff was aimed at describing properly B, N and BN systems with regard to dependence the energy on coordination and chemical environment of the bonding. The authors

put special attention to describe correctly monolayer hexagonal boron nitride (including lattice parameter, thermal stability and elastic properties). The formula is based on Tersoff formulation, the main change was introduced in the bond order part. b_{ij} from Eq. (2.23) is replaced by the *total bond order term* B_{ij}

$$B_{ij} = \frac{1}{2}(b_{ij} + b_{ji}) + F_{ij}. \quad (2.29)$$

with correction F_{ij} proposed by the authors of the paper to have the form

$$F_{ij} = F_{env}(N_{ij}, N_{ji})F_{IJ}(N_{ij}^{\bar{I}}, N_{ji}^{\bar{J}}). \quad (2.30)$$

F_{env} is an envelope function, smoothly switching off the correction term for the BN dimer as well as for high total coordinations $N_{ij} = \sum_{k \neq i,j} f_C(r_{ik})$. $F_{IJ}(N_{ij}^{\bar{I}}, N_{ji}^{\bar{J}})$ was proposed as a correction term for Tersoff potential in the PhD thesis of Karsten Albe (Albe, 1998), as dependent on partial reduced coordination $N_{ij}^{\bar{I}} = \sum_{k \neq i,j} f_C(r_{ik})\delta_{I,K}$. ($\delta_{I,K} = 1$ if the chemical identity K of atom k is not equal to the chemical identity I of atom i , zero otherwise). The detailed description of the whole correction can be found in (Los *et al.*, 2017), as well as the results concerning point defects in hexagonal and cubic boron nitride.

2.3.2 Reactive force fields

In the present thesis we consider the two members of the reactive force fields family, namely ReaxFF and COMB3. Although developed independently, both potentials were built around the two fundamental concepts: bond-order and self-consistent charge equilibration. The idea of bond-order was introduced already in Tersoff potential, but Tersoff potential lacks the explicit electrostatic effects, which limits its applicability, *e.g.*, in some ionic or heterogeneous systems. The answer for this problem is the self-consistent charge equilibration approach. It is the scheme, the charge on each atom is determined dynamically, based on the principle of the electronegativity equalization. In the thermodynamic equilibrium, the electrochemical potential (which is the negative value of the electronegativity) is equal on all the atomic sites. There are two main methods to satisfy the above mentioned rule: the electronegativity equalization method (EEM) introduced in (Mortier *et al.*, 1986) and charge equilibration (QEq) method (Rappe *et al.*, 1991). The first was employed in

ReaxFF, while the latter - in the COMB formulation. Both allow to assign charges to atoms, that minimize the electrostatic energy of the system, while keeping the net charge constant. The difference is that EEM uses direct Coulomb interaction, whereas QEq includes shielding between the overlapping charge clouds, and by this, is more computationally demanding.

Both potentials are quite complex, and there are many empirical parameters to be carefully trained against *ab initio* or experimental data. Elaborate formulation results, however, in the ability of the potentials to deal with various phases of a given material. Yet, there is a price for this: the computational cost of such calculations is approximately two orders of magnitude higher, compared to non-reactive force fields. Nonetheless, they are still much faster than DFT computations. In the following paragraphs both potentials will be concisely introduced. For the comprehensive comparison of the two force fields, see, *e.g.*, (Y. K. Shin *et al.*, 2012).

ReaxFF potential

ReaxFF was first formulated by A. v. Duin and coworkers in 2001 (Duin *et al.*, 2001) for hydrocarbons, then further developed to the form presented below. It was the first force field with dynamic bonds and polarization effects included in its formulation. An extensive description of the history and evolution of ReaxFF can be found in the review article (Senftle *et al.*, 2016). The energy formula has the form

$$E_{system} = E_{bond} + E_{over} + E_{angle} + E_{tors} + E_{vdWaals} + E_{Coulomb} + E_{Specific}. \quad (2.31)$$

E_{bond} is associated with forming the bonds between the atoms, directly derived from the bond orders. E_{over} is a correction preventing the overcoordination, E_{angle} characterizes three-body - angle, while E_{tors} four-body - torsion interaction. $E_{vdWaals}$ describes van der Waals, and $E_{Coulomb}$ - Coulomb interaction, the last two are calculated between all atoms. $E_{Specific}$ corresponds to system specific terms, that are to be included only in the particular cases, such as C₂ corrections, hydrogen binding, lone-pair or conjugation. Full description of the potential can be found, *e.g.*, in (Russo *et al.*, 2011).

This is a complex formulation and consists of both general model parameters and atomic specific parameters, which makes determination of the accurate parametrization a demanding task. In order to model interatomic reactions with precision,

one can include in the training set not only information about minima - lowest energy structures, but also include equations of state, heats of formation for small molecules or bulk structures. Dynamic bonds creation and breaking as well as charge equilibration that are implemented in ReaxFF force field have been very promising as an upgrade from 'non-charge-including' Tersoff potential. We have tested existing parametrizations in the area of CBN hexagonal structures. Detailed description of those parametrizations can be found in Sec. 4.2, and the results of the tests in Sec. 4.4.

Charge Optimized Many Body potential

Charge Optimized Many Body potential was presented in 2007 (Yu *et al.*, 2007), and was based on the Tersoff potential to describe short-range interactions. The successive development of the formulation resulted in second and, (the latest) third generation (COMB3) announced in 2012 (Liang, Devine, *et al.*, 2012). The detailed description of the whole formalism and development of the potential throughout the generations can be found in (Liang, Shan, *et al.*, 2013). We present here the newest formulation, where the total energy of the system U_{TOT} is the sum

$$U_{TOT} = U_{Ele}(q, \mathbf{r}) + U_{short}(q, \mathbf{r}) + U_{vdW}(\mathbf{r}) + U_{corr}(i, j, k). \quad (2.32)$$

The electrostatic term U_{Ele} is further expanded into

$$U_{Ele}(q, \mathbf{r}) = U_q(q, \mathbf{r}) + U_{qq}(q, \mathbf{r}) + U_{qz}(q, \mathbf{r}) + U_{polar}(q, \mathbf{r}), \quad (2.33)$$

with $U_q(q, \mathbf{r})$ being energy to form a charge from a neutral atom (taking into account electron affinity, ionization energies, change in electronegativity and chemical hardness, as dependent on the environment (Toufar *et al.*, 1996)). $U_{qq}(q, \mathbf{r})$ describes charge-charge interactions (integration over the Streitz and Mintmire style charge density of two atoms (Streitz *et al.*, 1994)), $U_{qz}(q, \mathbf{r})$ - charge-nuclear interactions (specifically - a sum of interactions between point charge of one atom and the shielded nuclear interaction of a neighboring atom). U_{qq} and U_{qz} are calculated using Wolf charge - neutralized direct summation method (Wolf *et al.*, 1999). U_{polar} is the energy associated with the polarizability (relative tendency of a charge distribution to distort in the presence of an external field) (Stern *et al.*, 1999).

The short-range energy term U_{short} is captured by extended Tersoff-like potential

$$U_{short}(q, \mathbf{r}) = \sum_i \sum_{(j>i)} ((F_c(r_{ij})(V_R(r_{ij}, q_i, q_j) - B_{ij}V_A(r_{ij}, q_i, q_j)))) \quad (2.34)$$

The crucial difference is the bond order term B_{ij} that has much more complicated form, compared to the Tersoff potential. The van der Waals term U_{vdW} has the form of classic Lennard – Jones formalism, and the correction term U_{corr} allows for the distinction between similar phases, and increased the accuracy in ternary potentials. COMB potential has proven generally better transferability than ReaxFF, which gives hope to get reliable results with the existing parametrizations. Comparison of COMB performance to other approaches is given in 4.4.

2.4 Schemes for description of ordering in CBN alloys

This section presents shortly computation techniques and defines the properties that we use for qualitative and quantitative description of stability and ordering in CBN alloys. All of them are broadly documented in the literature, however, we include them to keep this dissertation self-contained and to facilitate further discussion of results.

2.4.1 Monte Carlo Metropolis Method

To gain a knowledge about physical properties of the system consisting of vast number of atoms, it is common to use the language of statistical physics. In this approach, one needs to calculate the average value of the proper function describing property under interest, *e.g.*, energy, pressure, magnetization, degree of chemical order. Let's consider such a function (observable) F , dependent on all positions and momenta of the atoms in the system $(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$. The mean value of F is determined with the following formula

$$\langle F(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) \rangle = \int f(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) F(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) d\tau, \quad (2.35)$$

where $f(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$ is a probability density of finding the system in a state described by a given positions and momenta of atoms. The integral is performed

over all microscopic states in which the system can be found, *i.e.*, over the whole configurational space. In the canonical (NVT) ensemble, *i.e.*, when the temperature, volume, and number of particles are constant, the probability density f equals

$$f(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) = \frac{1}{Z} \exp \left(-\frac{E(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)}{k_B T} \right), \quad (2.36)$$

where k_B is Boltzmann constant. Partition function Z that appears in above formula ensures f to be normalized. When we only investigate spacial distribution of the density probability of finding the atoms, it can be integrated over momenta degrees of freedom to the form

$$f_S(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{Z_S} \exp \left(-\frac{V(\mathbf{r}_1, \dots, \mathbf{r}_N)}{k_B T} \right), \quad (2.37)$$

that depends only on atomic positions, and the potential $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ by which the atoms are influenced. In canonical ensemble, the microstates with equal energy (potential value) exhibit uniform probability. Calculation of Eq. (2.35) is quite challenging (analytically and computationally), as multiple integral over multidimensional space is to be evaluated. Thus, instead of explicit integration one can approximate Eq. (2.35) by the sum performed over a subset of the microscopic states $\sigma = \{\sigma_i = (\mathbf{r}_1^i, \dots, \mathbf{r}_N^i) : i = 1, \dots, M\}$ and evaluate discrete average over M states

$$\langle F \rangle = \frac{1}{M} \sum_{i=1}^M F(\mathbf{r}_1^i, \dots, \mathbf{r}_N^i). \quad (2.38)$$

However, choice of the σ cannot be completely free. The states σ_i should be selected according to the probability defined in Eq. (2.37) or generally consistent with the probability depending on the chosen ensemble. Solution to this task was the algorithm presented in (Metropolis *et al.*, 1953), where the authors construct samples from any probability distribution via generating a Markov chain (Norris, 1997). This idea was further developed in (Hastings, 1970) and is known as the Metropolis-Hastings algorithm. It generates a sequence of states, in our case - configurations of atoms - in a way that the more samples are produced, the closer the distribution is to the one we desire. The procedure starts by choosing arbitrary initial state of the system $\sigma_0 = (\mathbf{r}_1^0, \dots, \mathbf{r}_N^0)$. Then, some modification of σ_i is performed, *e.g.*, by atomic shift, spin flip or (like in our simulations) by swapping random pair of atoms of different types, that leads to $\sigma_{i'}$ state. Then, the energy difference between the states

$\Delta V = V_{i'} - V_i$ is considered. For modifications leading to negative ΔV the change is always accepted. If the energy increases, *i.e.*, $\Delta V > 0$, then $\sigma_{i'}$ is accepted as a new configuration with probability $p_i = \exp\left(-\frac{\Delta V}{k_B T}\right)$. This is typically obtained via comparing p_i to a random number from between 0 and 1. If the change is accepted, $\sigma_{i'}$ becomes σ_{i+1} , otherwise $\sigma_{i+1} = \sigma_i$.

Clearly, during the simulation described above, system tends to decrease its energy, as all the modifications leading to the lower energy are approved. Therefore, it can be employed to searching equilibrium state of the systems and is a basis of stochastic optimization methods. Other techniques are known as well, *e.g.*, simulated annealing, quantum annealing, cross-entropy method, evolutionary algorithms, *etc.*. In our research we employ Monte Carlo Metropolis algorithm for determination of equilibrium distribution of dopants in graphene or hexagonal boron nitride in the NVT canonical ensemble.

2.4.2 Warren-Cowley Short Range Order Parameters

There are multiple ways of characterizing ordering in the multicomponent systems. For the quantitative description of chemical order in CBN alloys, we choose Short Range Order approach, introduced in (Cowley, 1950). Warren-Cowley parameters $\Gamma_{AB}^{(i)}$ describe the derivation from random configuration towards ordering or segregation

$$\Gamma_{AB}^{(i)} = 1 - \frac{P_{AB}^{(i)}}{c_A c_B}. \quad (2.39)$$

where c_A and c_B denote concentrations of A and B species respectively, $P_{AB}^{(i)}$ is probability of finding A-B pair in i^{th} coordination shell. The presence of short-range ordering can be decided based on the value of $\Gamma_{AB}^{(i)}$. Values close to 0 indicate fully random structure (probability to find A and B is according to the concentrations), negative values suggest SRO (A and B tend to occupy i^{th} shell), while positive values appear when atoms A and B segregate (A and B 'avoid' each other in i^{th} shell). Limiting value is 1, when no A-B pairs are found on i^{th} coordination shell.

Let's mention a few properties of W-C parameters, following (Ducastelle, 1991):

- Symmetry of the parameters: $\Gamma_{AB}^{(i)} = \Gamma_{BA}^{(i)}$ for any species A, B and coordination shell i .

- Sum rules valid for any m -component system and any atom type C : $\sum_D c_D \Gamma_{CD}^{(i)} = 0$
- In m -component system, the total number of parameters equals m^2 for any coordination shell. But, due to the symmetry and sum rules, number of independent parameters is reduced to $m(m-1)/2$. This results in only one independent $\Gamma_{AB}^{(i)}$ in binary, while three independent parameters in ternary alloys.
- One can rewrite $\Gamma_{AB}^{(i)}$ into the form convenient in calculations. $\Gamma_{AB}^{(i)}$ is related to the number of $A-B$ pairs in i^{th} coordination shell and the total number of bonds in the system

$$\Gamma_{AB}^{(i)} = 1 - \frac{(1 + \delta_{AB})}{2c_A c_B} \frac{N_{AB \text{ bonds}}}{N_{\text{all bonds}}}. \quad (2.40)$$

with $\delta_{AB} = 1$ for $A=B$ and zero otherwise.

2.4.3 Definitions of Energy in heteroatomic systems

Throughout the thesis we will refer to various definitions of energy, each providing another type of information about the system. They are shortly described below.

1. **Cohesive Energy** E_{coh} is the energy that system D gains by arranging the atoms into molecular or crystalline state. It is calculated by subtracting sum of energies of the isolated atoms X_i from the total energy of molecule/solid:

$$E_{coh}(D) = E_{tot}(D) - \sum_i n_i E_{atomic}(X_i). \quad (2.41)$$

n_i denotes number of atoms of type X_i present in the system D . For stable systems, cohesive energy is negative, meaning that it is energetically favorable to create a given molecule or crystal. $E_{coh} > 0$ indicates instability of a system, that would split into isolated atoms to lower the total energy. From the definition Eq. (2.41) we see that cohesive energy is extensive property of the system. In many cases it will be more convenient to use cohesive energy per atom.

2. **Mixing Energy** E_{mix} for alloy system composed of A and B materials with concentrations x and y , respectively is defined as follows

$$E_{mix}(A_x B_y) = \frac{E_{tot}(A_x B_y)}{N_{atoms}} - x E_{tot}^{at}(A) - y E_{tot}^{at}(B). \quad (2.42)$$

Here $E_{tot}^{at}(A)$ and $E_{tot}^{at}(B)$ denote energy per atom of A and B systems being alloyed, respectively. Value of the mixing energy describes the preference (anti-preference) of A and B systems to intermix within the lattice. Positive value of E_{mix} suggests that alloy system would rather separate into A and B phases than form a solution. Mixing energy can be defined for system composed of more than two components as well.

3. **Formation Energy of defects** E_{form} (also called enthalpy of formation) shows how much energy we need to provide to the system to create a defect (its value might be negative as well). The formulation below follows (S.-H. Wei *et al.*, 2000). For clarity, separate formulas for defects in graphene and boron nitride are presented, which are the structures we are interested in this thesis.

$$\begin{aligned} E_f(GL_d) &= E_{tot}(GL_d) - E_{tot}(GL) + n_C\mu_C + n_B\mu_B + n_N\mu_N \\ E_f(hBN_d) &= E_{tot}(hBN_d) - E_{tot}(hBN) + n_C\mu_C + n_B\mu_B + n_N\mu_N. \end{aligned} \quad (2.43)$$

$E_{tot}(GL_d)$ and $E_{tot}(hBN_d)$ refer to the total energy of GL and hBN system with defects, respectively, while $E_{tot}(GL)$ and $E_{tot}(hBN)$ are the energies of pristine GL and hBN (*i.e.*, the same shape and size but without the defects). n_B , n_N and n_C indicate the number of B, N, and C atoms, respectively, transferred from the supercell to the reservoir in forming the defect. For example, for isolated carbon impurity in hBN substituted in place of B atom, $n_B = 1$, $n_C = -1$, and $n_N = 0$. In the case of single vacancy in graphene: $n_C = 1$, while the others null. Finally, μ_C , μ_B , and μ_N denote values of the chemical potentials for C, B, and N, respectively.

2.4.4 Entropy and Free energy

Another attribute of CBN alloys investigated in this thesis is range of miscibility of the structures based on graphene and boron nitride, when a number of heteroatoms is introduced to the pristine lattice. For this purpose, we want to access **free energy of mixing** F_{mix} of the system D being a function of temperature T

$$F_{mix}(T, D) = E_{mix}(D) - TS_{conf}(D) \quad (2.44)$$

In above equation S_{conf} denotes **configurational entropy** of a given system. There are multiple ways to approach entropy calculations. One can use thermodynamic

integration following (Binder *et al.*, 2010)

$$S_{conf}(T) = \int_0^T \frac{C_{conf}(T')}{T'} dT', \quad (2.45)$$

where specific heat C_{conf} is related to the fluctuations of the energy of mixing calculated, *e.g.*, in Monte Carlo simulations

$$C_{conf}(T) = \frac{\langle E_{mix}(T)^2 \rangle - \langle E_{mix}(T) \rangle^2}{T^2}. \quad (2.46)$$

However, due to the considerable computational effort of integration method as well as specifics of the systems we investigate (that will be discussed in the second part of the thesis), we have decided to utilize more straightforward approach. The simplest way to determine entropy connected with existence of distinct species within the system is *point approximation* as described in (Bragg *et al.*, 1934) and named **Bragg-Williams entropy** S_{BW} . Its value (per lattice point) can be determined as follows

$$S_{BW} = -k_B \sum_i c_i \ln(c_i). \quad (2.47)$$

Here c_i refer to atomic concentrations comprising the system, and k_B is Boltzman constant. Eq. (2.47) assumes that atoms behave independently, and only concentrations of species contribute to entropy. However, in reality atoms interact with each other (especially in the area of solid state physics), their configurations are rather correlated. To embrace it, *pair approximation* can be introduced where pair probabilities p_{ij} of finding $i - j$ bonds are taken into account in addition to point probabilities. In the system with coordination number Z , **entropy in the pair approximation** (per atom) (Mohri, 2007) S_{pair} is given by

$$S_{pair} = k_B \left((Z-1) \sum_i c_i \ln(c_i) - \frac{Z}{2} \sum_{i,j} p_{ij} \ln(p_{ij}) \right). \quad (2.48)$$

Point and pair approximations can be extended to a further basic cluster in tetrahedron approximation as described *e.g.*, in (Ibid.).

2.4.5 Empirical Tight Binding approach

Empirical Tight Binding (TB) approximation encompasses solving the equations that take into account the electronic structure of the system, but with computational effort being just a small percentage of the first principles simulations. It is a result of significant simplifications in the quantum mechanical hamiltonian, and introduction of empirical parameters that need to be fitted. Tight binding approach can be used to calculate electronic structure, chemical bond order, or even the transport properties in large systems.

The basis idea of the TB method is expansion of the one particle states in molecules or solids in terms of Linear Combination of Atomic Orbitals (LCAO). One can assume that the solutions of the one particle Kohn-Sham equation (see Eq. (2.8)) can be searched as

$$|\phi_n\rangle = \sum_{i,\alpha} c_n^{i\alpha} |\chi_{i\alpha}\rangle, \quad (2.49)$$

where the functions $\{|\chi_{i\alpha}\rangle\}$ constitute not necessarily orthonormal set of atomic-like functions, and the pair (i, α) indicates the node index i and the type of atomic wave function (orbital) α . The nodes might be identical with the atomic positions. The solution of the Kohn-Sham equation is then reduced to solution of the generalized eigenvalue problem $H_{i\alpha,j\beta} c_n^{j\beta} = E_n S_{i\alpha,j\beta} c_n^{j\beta}$, with hamiltonian $H_{i\alpha,j\beta}$, and overlap $S_{i\alpha,j\beta}$ matrix elements, respectively, defined as follows: $H_{i\alpha,j\beta} = \langle i\alpha | \hat{H} | j\beta \rangle$ and $S_{i\alpha,j\beta} = \langle i\alpha | j\beta \rangle$.

There are numerous DFT codes on the market solving the Kohn-Sham equation employing expansion of one particle states into different sets of atomic like functions, be it analytic or numerical ones (Blum *et al.*, 2009) (for example, ORCA, SIESTA, TURBOMOLE). However, the main simplification of the Tight Binding approach is that the matrix elements of the hamiltonian, can be treated as empirical parameters (Goringe *et al.*, 1997; Papaconstantopoulos *et al.*, 2003), or fitted to the effective TB hamiltonian through finding the most localised Wannier functions (Marzari *et al.*, 2012) from Kohn-Sham solutions and calculating with so obtained WFs the matrix elements of the Kohn-Sham hamiltonian and therefore, creating effective TB hamiltonian for the problem (Q. Li *et al.*, 2021; Nakhaee *et al.*, 2020). The classic paper of Slater and Koster (Slater *et al.*, 1954) gave the origin of the extremely successful field of the empirical tight binding theory. The theory relies on the symmetry of

orbitals and intuitive properties of the 'atomic basis functions' in contrast to numerical basis functions, which could be numerically very effective but simultaneously not intuitive. It is practically impossible to point out the area of condensed matter or molecular physics, where the empirical TB method did not contribute considerably to explanation and understanding of physics (see e.g. (Harrison, 1989)). The empirical TB method is extremely well documented in the literature and has so many applications that we are not making any attempt to summarize here even the minimal number of papers. The Slater-Koster (S-K) TB method assumes a minimal basis (orbitals of angular momentum s, p, d, ...) per atom and introduces two-center approximation for hamiltonian matrix elements, so-called molecular integrals. These two-center matrix elements can be transferred for a pair of atoms from environment to environment, therefore, the specific values of the hamiltonian matrix elements depend in the first order on the transformation properties of the l -type atomic orbitals. We finish the presentation of empirical TB with the statement that practically all bonding in molecules and solids can be understood in terms of empirical TB theory (Canadell *et al.*, 2012).

The very interesting combination of the Density Functional Theory and Tight Binding method is the so called DFTB method (Density Functional Tight Binding), or DFTB+. The DFTB method is based on a second-order expansion of the Kohn-Sham total energy with respect to charge density fluctuations. The zeroth order approach would be equivalent to a common standard non-selfconsistent TB scheme (Porezag *et al.*, 1995; Seifert *et al.*, 1996). In the second order, parameter-free and readily calculable expressions based on a self-consistent redistribution of Mulliken charges for generalized hamiltonian can be introduced (Elstner *et al.*, 1998; Spiegelman *et al.*, 2020). The numerical DFTB packages are nowadays available from many sources (see dftb.org, dftbplus.org, wiki.fysik.dtu.dk) and are definitely an option for calculations when dealing with huge systems is unavoidable.

In this thesis, we employ empirical TB method for the description of the electronic structure of CBN alloys. Since it is commonly known that in graphene and hBN the states close to the Fermi energy originate from the π -type (or p_z) orbitals (C. Huang *et al.*, 2015; Ryou *et al.*, 2017), only one p_z per lattice point is included into the basis of 'atomic like' functions. Additionally, we assume that p_z orbitals centered on different lattice sites are orthogonal to each other, and the hamiltonian matrix elements are different from zero only for two orbitals being centered at two nearest neighbours. Then the first neighbor hopping is characterized by single hopping

parameter t_0 , and the tight binding hamiltonian simplifies further to the form

$$H^{TB} = \sum_i \epsilon_i |i\rangle \langle i| + t_0 \sum_{\langle i,j \rangle} |i\rangle \langle j|, \quad (2.50)$$

where $|i\rangle$ are π -states centered on hexagonal lattice sites, and ϵ_i is the on-site energy.

With above formulation of the TB hamiltonian, we apply it to the systems considered in this thesis, for the supercells consisting of N_{lat} atoms, thus hamiltonian has the size $N_{lat} \times N_{lat}$. H^{TB} can be easily diagonalized, and we can access the total Density of States (DOS) and the Projected Density of States (PDOS) on the state $|i\rangle$ following the definitions

$$\begin{aligned} D(\epsilon) &= \sum_{\alpha} \delta(\epsilon - \epsilon_{\alpha}), \\ D_i(\epsilon) &= \sum_{\alpha} |\langle i|\alpha \rangle|^2 \delta(\epsilon - \epsilon_{\alpha}), \end{aligned} \quad (2.51)$$

where the sum runs over all basis functions α . In our simulations, we substitute the Kronecker delta δ in Eq. (2.51) by the Gaussian function $\tilde{\delta}$ of a given width Σ

$$\tilde{\delta}(\epsilon - \epsilon_{\alpha}) = \frac{1}{\sqrt{2\pi}\Sigma} \exp\left(-\frac{(\epsilon - \epsilon_{\alpha})^2}{2\Sigma^2}\right). \quad (2.52)$$

Chapter 3

First principles based simulations - Cluster Expansion approach

3.1 Cluster Expansion Iterative Technique

The possibility to investigate the physics of alloys on the first-principle level entered the condensed matter physics with the advent of the Cluster Expansion (CE) Method (as described in Sec. 2.2). In the meantime, the CE is commonly employed to investigate even many component mixtures and alloys, *e.g.*, (Sobieraj *et al.*, 2020) contributing considerably to better understanding of these very important classes of materials. In this thesis, we use the Cluster Expansion (on the *ab initio* basis) method to investigate building blocks of the CBN systems, *i.e.*, binary carbon-nitrogen and carbon-boron alloys indicated hereafter as GL-N and GL-B. Such systems can be created in many fabrication processes where graphene is doped with nitrogen or boron, mostly beyond of the thermodynamic equilibrium. However, the resulting concentrations of N and B in the graphene framework are rather low. We employ CE to find out the alloy structures of the lowest total energies for the given concentrations and further to check out whether they exhibit the possible Short Range Order (SRO). At the end, we calculate the free energy and discuss phase diagrams of binary alloys. We restrict the N and B concentrations to 30%. The results obtained within the CE in this chapter will constitute some kind of reference for the further calculations of CBN alloys within the VFF method.

In the calculations presented here, we follow closely the self-consistent technique developed and described by researchers from Humboldt University (Troppenz *et*

al., 2017). Their procedure allows for iterative fitting (or tuning) of the CE model and, to large extent, controlling of the convergence of the models. It may be applied for any type of reference structures, and to any property that can be approximated as a function of atomic configurations.

Let us now list the main points illustrating the procedure leading to the generation of the best CE model. Two sets of data play the basic role in this generation scheme: Initial training set and Clusters Pool, and, of course, the DFT that is used to determine the reference energies of the structures. As in all other places in this thesis, we employ Quantum Espresso (QE) simulation package for the DFT calculations.

1. Initial training set is used to extract the basis knowledge about the interactions between atoms in an alloy. We have chosen the set consisting of 20 structures based on 4x4 graphene supercells (containing 32 atoms) with diversified configurations of nitrogen (boron) substitutions distributed over the graphene lattice up to concentration of 28.125% (which is equivalent to 9 substituted atoms), referenced further on as **REFSET1**, and depicted in Fig. 3.1.
2. DFT total energies for all structures of the initial training set are computed and constitute the data indicated by **QE-DFT**. Based on these total energies, one can determine mixing energies E_{mix}^{DFT} (as defined in Eq. (2.42)). For GL-N and GL-B, it can be written as

$$E_{mix}^{DFT}(GL_{1-x}N_x) = \frac{E_{tot}^{DFT}(GL_{1-x}N_x)}{N_{atoms}} - (1-x)E^{DFT}(GL) - xE_{at}^{DFT}(N_2),$$

$$E_{mix}^{DFT}(GL_{1-x}B_x) = \frac{E_{tot}^{DFT}(GL_{1-x}B_x)}{N_{atoms}} - (1-x)E^{DFT}(GL) - xE_{at}^{DFT}(B_\alpha),$$

where x indicates concentration of substituted nitrogen or boron atoms, N_{atoms} is the total number of atoms in the system (here equal to 32), $E_{at}^{DFT}(N_2)$ is energy per atom in nitrogen molecule, while $E_{at}^{DFT}(B_\alpha)$ denotes energy per atom of α -borophene, the most stable layered structure of pristine two-dimensional boron. Set of E_{mix}^{DFT} values constitutes our reference data for fitting Cluster Expansion model.

3. Clusters Pool is the source of possible clusters for the fitting procedure. The bigger the number and size of clusters in Clusters Pool, the more likely we are to find the ones that are relevant for the accurate description of our property. We choose a pool consisting of 44 clusters (which is depicted in supplementary Fig. A.1): 0-point cluster (empty graphene lattice), 1-point cluster (single N(B) atom

in GL), all symmetrically inequivalent 2- and, 3-point clusters, and 4-point clusters up to the radius (*i.e.*, maximum distance between the cluster atoms) equal to 3.5 Å. The Clusters Pool size is confined only due to the computational limitations. However, note that in Clusters Pool there are typically only clusters that do not exceed the certain range of atomic interactions (see Eq. (2.15)). Also, the common approach is that the CE hamiltonian does not include higher than four-body interactions.

4. Fitting Cluster Expansion model. As an engine for determining optimum CE model, we use CELL code (Rigamonti *et al.*, n.d.; Troppenz *et al.*, 2017). Based on the reference data, the optimum set of clusters (subset of the Clusters Pool) and values of respective Effective Cluster Interactions (ECIs) are determined. The quality of the model is evaluated with Root Mean Square Error (RMSE) of the CE model predictions versus reference set of DFT energies (having size N) and Cross Validation (CV) score, as described in Section 2.2.2

$$RMSE = \left(\frac{1}{N} \sum_{i=1}^N (E_{mix}^{DFT}(S_i) - E^{CE}(S_i))^2 \right)^{\frac{1}{2}}, \quad CV = \left(\frac{1}{N_t} \sum_{i=1}^{N_t} (E_{mix}^{DFT}(S_i) - \hat{E}(S_i))^2 \right)^{\frac{1}{2}}.$$

$E^{CE}(S_i)$ denotes energy of i^{th} reference structure predicted with a fitted model, while $\hat{E}(S_i)$ refers to predicted value of S_i structure energy obtained from CE model fitted to reference training set excluding the data point S_i , $N_t = N - 1$.

The procedure summarized in the last point, leads to the first order CE model and its hamiltonian, $\hat{H}_{CE}^{(1)}$. In this model, $M = 13$ clusters from the Cluster Pool provide the best fit to the DFT mixing energies for structures included in the Initial Training Set. Now we describe the scheme allowing for systematic and iterative improvement of the CE hamiltonian $\hat{H}_{CE}$.

1. Evaluation of the model. First, for every number of dopants on the 32 lattice points superlattice (*i.e.*, $k = 0-9$), one finds atomic configuration $\sigma^{(1)}(k)$ minimizing the energy of the CE hamiltonian $\hat{H}_{CE}^{(1)}$, just creating the pairs $(\sigma^{(1)}(k), E^{CE(1)}(\sigma^{(1)}(k)))$ (note that $k=0$ and $k=1$ are trivial). This minimization procedure is performed employing the Monte Carlo method (with temperature set to 10K, and performing $2 \cdot 10^5$ MC swaps). Next, for $\sigma^{(1)}(k)$ alloy configurations, one calculates DFT mixing energies, just obtaining the set $E_{mix}^{DFT}(\sigma^{(1)}(k))$. The discrepancy in the DFT and CE energies for the corresponding configurations $\sigma^{(1)}(k)$ can be determined, for example, as $R^{(1)} := \sum_{k=0}^9 |E_{mix}^{DFT}(\sigma^{(1)}(k)) -$

$E^{CE(1)}(\sigma^{(1)}(k))$. If this residual exceeds the required value, the Initial training set is augmented with the pairs $(\sigma^{(1)}(k), E_{mix}^{DFT}(\sigma^{(1)}(k)))$ and the new CE hamiltonian $\hat{H}_{CE}^{(2)}$ is fitted according to the procedure described in the Fitting Cluster Expansion model point. Then the new set of configurations $\sigma^{(2)}(k)$ ($k = 0-9$) that minimize the total energies of the CE hamiltonian $\hat{H}_{CE}^{(2)}$ with energies $E^{CE(2)}(\sigma^{(2)}(k))$ is found in the MC procedure. Further the DFT energies, $E_{mix}^{DFT}(\sigma^{(2)}(k))$, are computed for the alloy configurations $\sigma^{(2)}(k)$. When the energies obtained from the actual CE hamiltonian differ too much from the corresponding DFT energies, the procedure is iteratively repeated, and it is stopped for iteration i where $R^{(i)} := \sum_{k=1}^9 |E_{mix}^{DFT}(\sigma^{(i)}(k)) - E^{CE(i)}(\sigma^{(i)}(k))| \leq \varepsilon$. Evaluation of the reliable CE models for GL-N and GL-B typically requires few iterations, as it will be presented later on. In the context of this procedure leading to the improvement of the CE hamiltonian, one speaks very often about the investigation of the prediction power of the CE model, in the sense that CE model (and the hamiltonian) gets ability to predict the energies very close to the DFT ones for identical configurations.

Having the CE model and hamiltonian $\hat{H}_{CE}$ determined, one is in the position to perform extended Monte Carlo calculations for alloys. Here the GL-N and GL-B alloys for N and B concentrations in range 0.01-0.30 have been studied within the MC method on the 800 site hexagonal lattice with periodic boundary conditions. For each concentration, the dopants are initially distributed randomly over the lattice, and Metropolis algorithm is used to swap atoms in search for the configuration minimizing the total energy. When such configuration is achieved, it is then analyzed in the context of short range order (SRO) formation.

3.2 Reference DFT data for fitting CE models

Cluster Expansion methodology requires creation of the reference data set (Initial training set) and performing the most reliable calculations of the mixing energies on the DFT level. Such DFT calculations can be performed either on the rigid lattice of the alloy or taking into account the volume and the atomic positions relaxation. In this thesis, we focus mainly on the *chemical order* of the CBN alloys on graphene honeycomb lattice and try to shed light on the chemical mechanisms leading to the formation of such systems. Therefore, it is also a prerequisite to investigate the

chemical order in the GL-N and GL-B binary alloys on the first-principles levels. In the following, we describe the construction of the two CE models for GL-N and GL-B alloys based on the same Initial training set of 20 structures (see figure Fig. 3.1) treating them either as (i) *rigid* or (ii) fully *relaxed* systems. Altogether, we consider four CE models, hereafter indicated as: **GL-N-RIGID** - GL-N on rigid lattice; **GL-N-RELAX** - GL-N with fully relaxed geometry; and corresponding models for boron-doped graphene, **GL-B-RIGID** - GL-B on rigid lattice; and **GL-B-RELAX** - GL-B with fully relaxed geometries.

First, we make a survey of the DFT results for initial training set of structures, on rigid and fully relaxed lattice. This allows us to assess the degree of geometrical relaxations in the investigated alloyed systems. We perform DFT calculations employing the Quantum Espresso simulation package (Giannozzi *et al.*, 2009) taking into account electronic spin-polarization. The total energies (for both rigid and relaxed systems) and optimized geometries are obtained with the use of Perdew, Burke and Ernzerhof (PBE) exchange-correlation functional of the GGA type for description of the electron-electron interaction. To account for the interactions of the valence electrons with atomic cores, we employ fully relativistic Projector-Augmented Wave (PAW) pseudopotential type with non-linear core correction. In the calculations, we use the PAWs generated considering as valence electrons 2s and 2p for carbon and nitrogen, while 2p for boron (specific PAWs were downloaded from the library <http://www.quantum-espresso.org/pseudopotentials>, names of pseudopotentials' files are following: C.rel-pbe-n-kjpaw_psl.1.0.0.UPF, N.rel-pbe-n-kjpaw_psl.1.0.0.UPF, B.rel-pbe-n-kjpaw_psl.1.0.0.UPF). The calculations for the structures of the initial training set are done in the 4x4 lateral graphene supercell, with 12 Å distance between the layers (so-called vacuum layer) to avoid spurious interlayer interaction. The kinetic energy cut-off for wavefunction expansion into plane waves is chosen to 80 Ry, and 8 times larger for the electronic density. The Monkhorst-Pack scheme is used for the Brillouin zone sampling, with 4x4x1 of k-points grid. During the geometry relaxation, the convergence threshold was set to 10^{-5} [Ry] in energy and 10^{-4} [Ry/Bohr] in force.

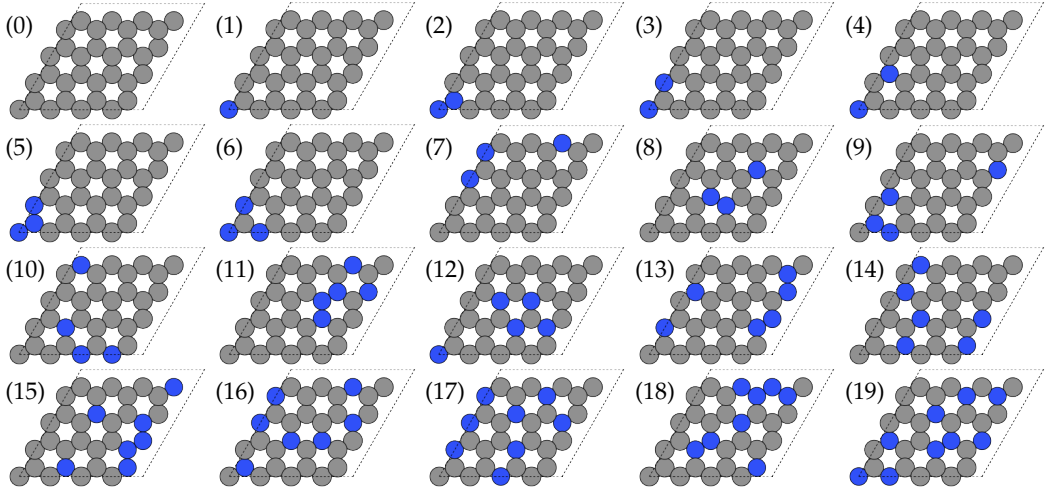


FIGURE 3.1: Clusters (with corresponding atomic configurations of superlattices) constituting initial training set (REFSET1) used for fitting the first iteration of Cluster Expansion (CE1) models. Gray dots are carbon atoms, while blue ones indicate dopants (N or B).

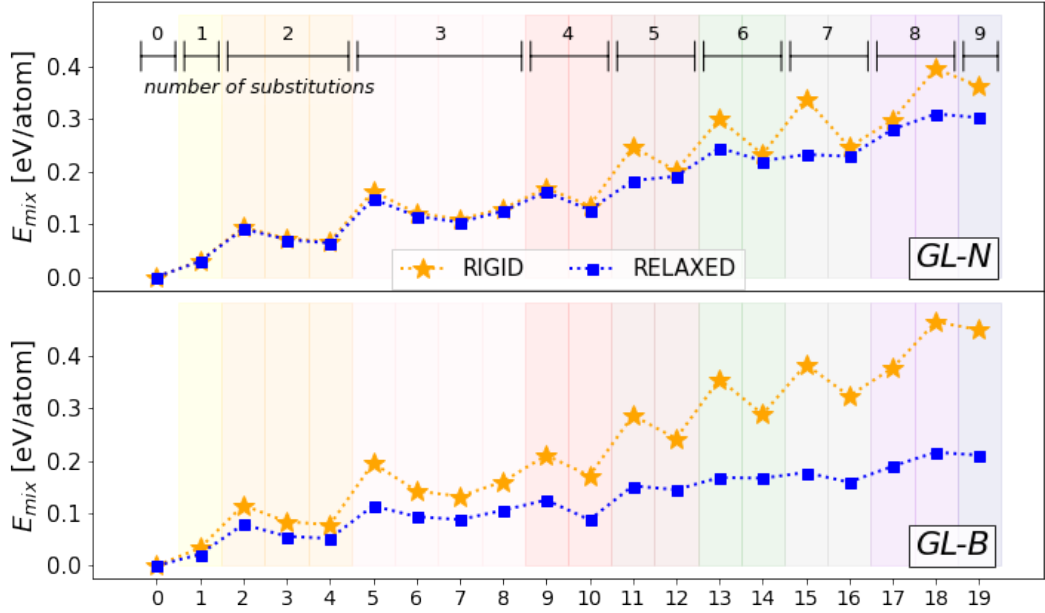


FIGURE 3.2: Mixing energies for nitrogen and boron alloyed graphene (GL-N and GL-B, respectively) for RIGID and RELAXED atomic positions as obtained from DFT calculations for atomic configurations indicated as 0-19 (depicted in Fig. 3.1). The number of dopants in the supercells for each configuration is shown in the upper part of the figure.

The results of performed calculations on initial training sets for GL-N and GL-B are summarized in Tab. 3.1. For each geometry (with index as in Fig. 3.1), we list the values of mixing energy per atom on the rigid lattice, the optimized lattice constant a_{lat}^{rlx} of the relaxed structure expressed as percentage of the pristine graphene lattice constant, the mixing energy for the system with relaxed atomic positions and volume. It turns out that some of the relaxed structures remain (as graphene) completely flat, whereas other exhibit some buckling. This fact is annotated in the column 'flat?' of the Tab. 3.1. The characteristic side view of such buckled structure is exemplified in Fig. 3.3, which should provide some idea about the possible degree of corrugation in the alloyed structures. For all studied N(B) concentrations, the mixing energies are positive, indicating that mixing C atoms with N(B) on the graphene lattice (either rigid or relaxed) is not energetically favorable, and rather non-equilibrium incorporation of N(B) into graphene lattice could lead to the desirable GL-N and GL-B alloys. The energetics of the studied GL-N and GL-B alloyed structures is nicely presented in Fig. 3.2. As can be seen, for GL-N the relaxation reduces the mixing energy rather moderately. The only larger reduction of the GL-N mixing energy (resulting from relaxation) occurs for structures, where two nitrogen atoms have been placed on the lattice as the nearest neighbors. It is obviously not energetically favorable, and the lattice undergoes relatively strong deformation, including the corrugation. The similar behavior is observed for the GL-B alloys, however, generally the relaxation effects are stronger here than in the GL-N alloys. We would like to point out relatively large increase of the GL-B lattice constants in comparison to the pristine graphene, particularly for the larger B concentrations. However, even in the case of GL-B alloy, the CE model based on rigid lattice could lead to correct qualitative (or even semi-quantitative) description of the alloy.

TABLE 3.1: Mixing alloy energies obtained from DFT calculations for GL-N and GL-B for rigid and relaxed atomic positions in clusters as depicted in Fig. 3.1. Number of substitutions n_{dop} , change of the lattice constant a_{lat}^{rlx} for the 4×4 superlattice, and the information whether the relaxed structure remains flat or gets corrugated are also listed.

		GL-N E_{mix} [eV/atom]				GL-B E_{mix} [eV/atom]			
index	n_{dop}	RIGID	RELAXED	a_{lat}^{rlx}	flat?	RIGID	RELAXED	a_{lat}^{rlx}	flat?
0 (GL)	0	0.00000	0.00000	2.468Å	✓	0.00000	0.00000	2.468Å	✓
1	1	0.03017	0.02926	-0.0%	✓	0.03368	0.02193	+0.7%	✓
2	2	0.09412	0.09124	-0.0%	✓	0.11526	0.07870	+1.4%	✓
3	2	0.07325	0.07074	-0.7%	✓	0.08374	0.05601	+0.7%	✓
4	2	0.06668	0.06437	-0.7%	✓	0.07791	0.05174	+0.7%	✓
5	3	0.16283	0.14801	-0.7%	✗	0.19538	0.11397	+0.7%	✗
6	3	0.12139	0.11543	-0.7%	✓	0.14230	0.09356	+1.4%	✗
7	3	0.10950	0.10386	-0.7%	✓	0.13199	0.08762	+1.8%	✓
8	3	0.12995	0.12508	-0.7%	✓	0.16001	0.10538	+1.8%	✓
9	4	0.16859	0.16148	-0.7%	✓	0.21027	0.12509	+1.8%	✓
10	4	0.13520	0.12734	-0.7%	✓	0.17134	0.08726	+1.8%	✓
11	5	0.24759	0.18295	-1.1%	✗	0.28601	0.15198	+1.4%	✗
12	5	0.20169	0.19138	-1.1%	✓	0.24214	0.14458	+1.8%	✗
13	6	0.30003	0.24544	-1.1%	✗	0.35486	0.16785	+2.8%	✓
14	6	0.23315	0.22063	-1.1%	✓	0.28843	0.16664	+3.9%	✓
15	7	0.33932	0.23295	-2.1%	✗	0.38312	0.17743	+2.8%	✓
16	7	0.24686	0.22931	-1.1%	✓	0.32311	0.15884	+4.6%	✓
17	8	0.29851	0.28069	-1.1%	✓	0.37641	0.18929	+3.9%	✗
18	8	0.39611	0.31012	-2.1%	✗	0.46416	0.21602	+2.8%	✗
19	9	0.36318	0.30303	-2.1%	✗	0.44912	0.21030	+4.6%	✗

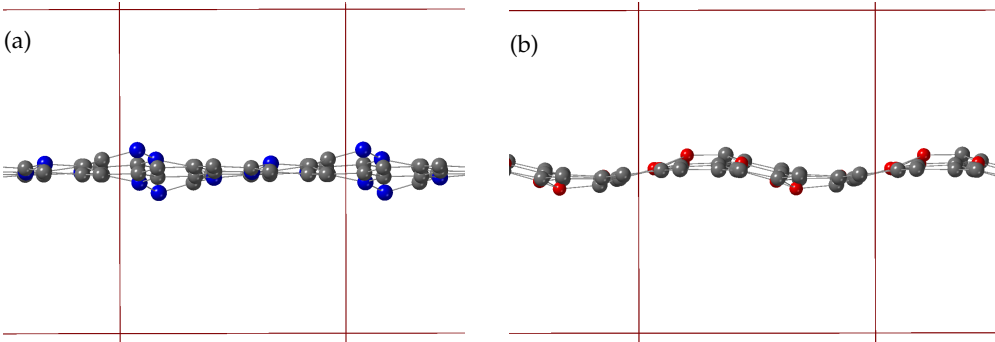


FIGURE 3.3: Side view of an exemplary corrugated relaxed structure from the REFSET1 (Fig. 3.1) for doping with (a) nitrogen (blue dots) (b) boron (red dots) after geometry and lattice constant optimization. Red lines show the boundaries of simulation boxes.

3.3 Results of CE models determination

Just having the data for initial training sets generated and presented in the previous section, we are now in the position to employ the procedure outlined in Sec. 3.1, leading to development of the CE models of suitable predictive power. We perform four iterations in the fine tuning of the CE model **GL-N-RIGID**, while three iterations are executed in remaining cases, **GL-B-RIGID**, **GL-N-RELAX**, and **GL-B-RELAX**. The characteristics of the CE models obtained at the stages of the iterative procedure are summarized in Tab. 3.2. To illustrate better the convergence of the CE i models, we plot in Fig. 3.4 histograms depicting the difference between energies of the lowest energy configurations predicted with the CE i hamiltonian and corresponding DFT energies, utilizing residual formula $R^{(i)}$. Systematically augmenting the number of configurations in the training set, as indicated by the REFSET i size variable in Tab. 3.2, one is able to develop better CE model (or hamiltonian). In performing this tuning of the CE hamiltonian, we rely on the experiences of the Berlin's collaborators.

TABLE 3.2: Details of the Cluster Expansion models determined during the iterative procedure as described in the text. CE i denotes the i^{th} iteration for the model for GL-N (GL-B) in the RIGID or RELAX framework, fitted to the mixing energies of REFSET i . $n_{clusters}$ is the number of clusters included in the CE model.

	GL-N				GL-B		
RIGID							
	CE1	CE2	CE3	CE4	CE1	CE2	CE3
$n_{clusters}$	13	13	11	13	13	13	13
max. cluster size	3	3	3	3	4	4	3
max. cluster radius	5.12	5.12	5.12	5.12	2.84	2.84	5.12
REFSET <i>i</i> size	20	29	40	43	20	28	41
RMSE [eV/atom]	0.00044	0.00073	0.00138	0.00149	0.00044	0.00131	0.00262
CV score [eV/atom]	0.00068	0.00115	0.00188	0.00215	0.00113	0.00196	0.00387
RELAX							
	CE1	CE2	CE3		CE1	CE2	CE3
$n_{clusters}$	13	13	13		12	13	13
max. cluster size	3	3	3		3	4	3
max. cluster radius	5.12	5.12	5.68		4.920	3.757	5.681
REFSET <i>i</i> size	20	32	42		20	33	41
RMSE [eV/atom]	0.00051	0.00538	0.00913		0.00082	0.00184	0.00406
CV score	0.00091	0.1306	0.03024		0.00157	0.00287	0.00563

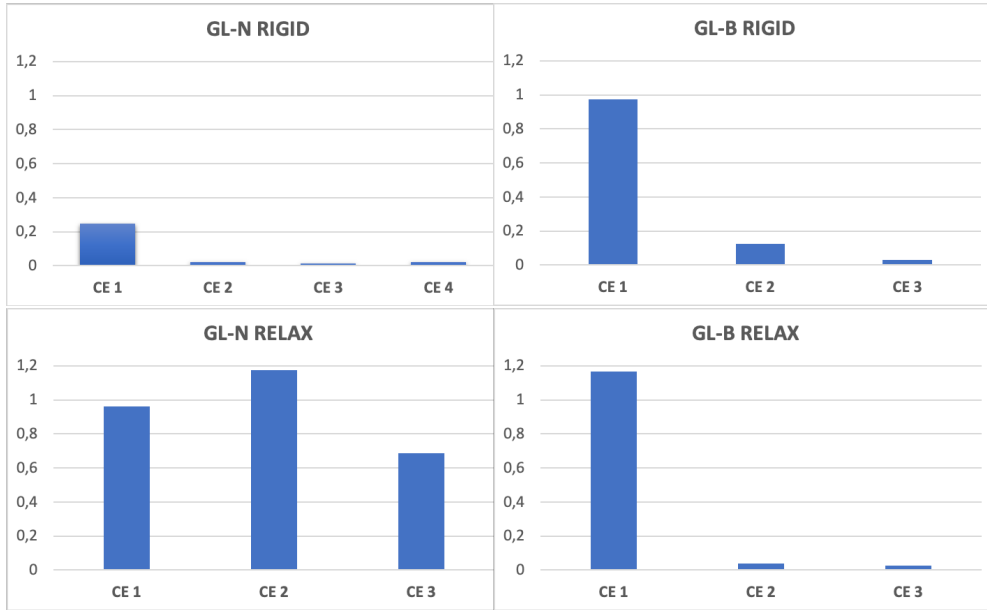


FIGURE 3.4: Histograms showing the value of $R^{(i)}$ for all iterations of CE_i model in all investigated classes of models.

Here, we would like to restrict the presentation of the details of the iterative procedure for the CE model to the case of the GL-N alloy on the rigid lattice **GL-N RIGID**, whereas the description of the iterative steps in the case of three remaining CE models is moved to Appendix A, for clarity.

CE model for graphene alloyed with nitrogen on rigid lattice (GL-N RIGID). Tab. 3.3 collects values of Effective Cluster Interactions determined for CE1-CE4 models, as well as the size and radius of respective clusters used for the construction of the CE hamiltonian. It is interesting to notice, that only zero-point, one-point, and some of two-point ECIs remain their values in transfer from one iteration of CE model to the next one ($i \rightarrow i + 1$), whereas the most of the *ECIs* change their values in some kind of statistical manner. Predictive power of subsequent iterations is shown in Fig. 3.5. Namely we compare CE_i models predictions for the REFSET i with the DFT calculations, and relate them to energies of those systems after Monte Carlo simulation (that should give configuration of atoms corresponding to the lowest energy). Energies of those minimum structures in CE_i model is then computed with DFT to estimate how well CE_i model predicts energy, and whether it finds truly the most stable configurations. Generally, one can say that even the first iteration model (CE1) delivers reasonable accuracy, as the structures obtained from Monte Carlo

simulations (reference set as initial structures) have lower energies than initial set, when calculated in DFT. Subsequent iterations 2-4 result in even better convergence. CE4 model is in very good agreement with DFT energies.

TABLE 3.3: Cluster Expansion models (CE1, CE2, CE3, CE4) for N-alloyed graphene on a rigid lattice obtained in the iterative procedure. For each iteration, clusters included in CE i model are characterized by: number of nitrogen atoms (called commonly points) in the cluster N_{pts} , its radius R , followed by value of Effective Cluster Interaction associated with this cluster.

	CE1 GL-N RIGID			CE2 GL-N RIGID			CE3 GL-N RIGID			CE4 GL-N RIGID		
	N_{pts}	$R[\text{\AA}]$	ECI	N_{pts}	$R[\text{\AA}]$	ECI	N_{pts}	$R[\text{\AA}]$	ECI	N_{pts}	$R[\text{\AA}]$	ECI
0	0	0.000	0.9600	0	0.000	0.9399	0	0.000	0.9766	0	0.000	0.9957
1	1	0.000	1.3357	1	0.000	1.2573	1	0.000	1.3704	1	0.000	1.4543
2	2	1.420	0.3223	2	1.420	0.2817	2	1.420	0.3025	2	1.420	0.3489
3	2	2.460	-0.0310	2	2.460	0.0644	2	2.460	0.0674	2	2.460	0.0495
4	2	2.840	-0.0495	2	2.840	-0.0683	2	2.840	-0.607	2	2.840	0.0527
5	2	3.757	0.0795	2	3.757	0.0146	2	3.757	0.0272	2	3.757	0.0070
6	2	5.681	-0.0258	2	4.260	0.0260	3	3.757	-0.0518	3	2.840	0.0642
7	3	2.460	-0.0342	2	4.920	-0.0741	3	3.757	-0.0207	3	3.757	-0.0568
8	3	4.260	-0.0085	3	2.460	-0.0327	3	4.920	-0.0238	3	4.260	-0.0579
9	3	4.920	-0.1117	3	3.757	-0.0315	3	4.920	-0.0172	3	4.260	0.0235
10	3	4.920	0.0458	3	4.920	-0.0734	3	5.120	-0.0231	3	4.920	0.0160
11	3	5.120	0.0179	3	4.920	-0.0052				3	4.920	-0.0108
12	3	5.120	0.0106	3	5.120	0.0415				3	5.120	-0.0254

As it has been already indicated in this thesis, one of the main research purposes is to find out whether the incorporation of N (B) atoms onto the graphene lattice can result in a random distribution of dopants or some kind of the Short Range Order forms in the alloy. Therefore, the SRO is definitely the property that we would like to obtain from the CE model with as high accuracy as possible. We use Warren-Cowley parameters as a measure of the Short Range Order in this thesis (described in Section 2.4.2), which have been defined as $\Gamma_{AB}^{(i)} = 1 - P_{AB}^{(i)} / (c_A c_B)$, where $P_{AB}^{(i)}$ is probability to find atoms of types A and B in the i^{th} coordination shell in the system with concentrations of A and B species equal to c_A and c_B , respectively. We perform (in addition to the RMSEs in Tab. 3.2) tests of the CE i model at each iteration level i in the context of predictions of the Warren-Cowley parameters. For each generation of the CE i model, we perform Monte Carlo calculations of the alloy configurations minimizing the total energy of the systems (supercells) with 800 lattice points and a given alloy concentration. The distribution of N dopants on the 800 supercell for

N concentrations equal to 2%, 10%, and 25% as obtained by the four generations of the CE models are depicted in Fig. 3.6. The corresponding Warren-Cowley SRO parameters for N-N pairs within coordination shells 1-4 are presented for the CE1 - CE4 models of GL-N in Fig. 3.7.

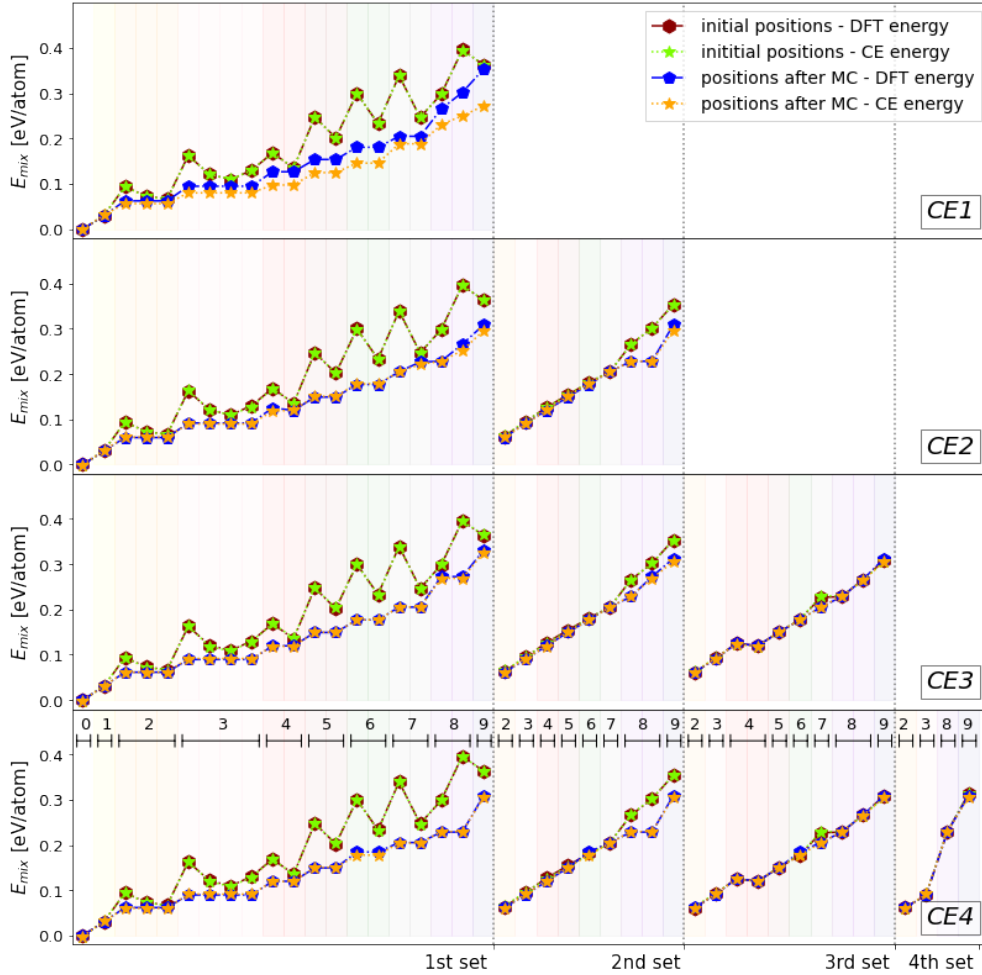


FIGURE 3.5: Predictive power of consecutive iterations CE1-CE4 during iterative procedure of Cluster Expansion model fitting for N-doped graphene in rigid lattice (**GL-N RIGID**). For iteration i energies before and after Monte Carlo simulations are plotted, calculated both with CE_i model and with DFT. Consecutive reference sets are shown on x-axis. The number of dopants in each configuration is shown in the upper part of the lowest panel for CE4.

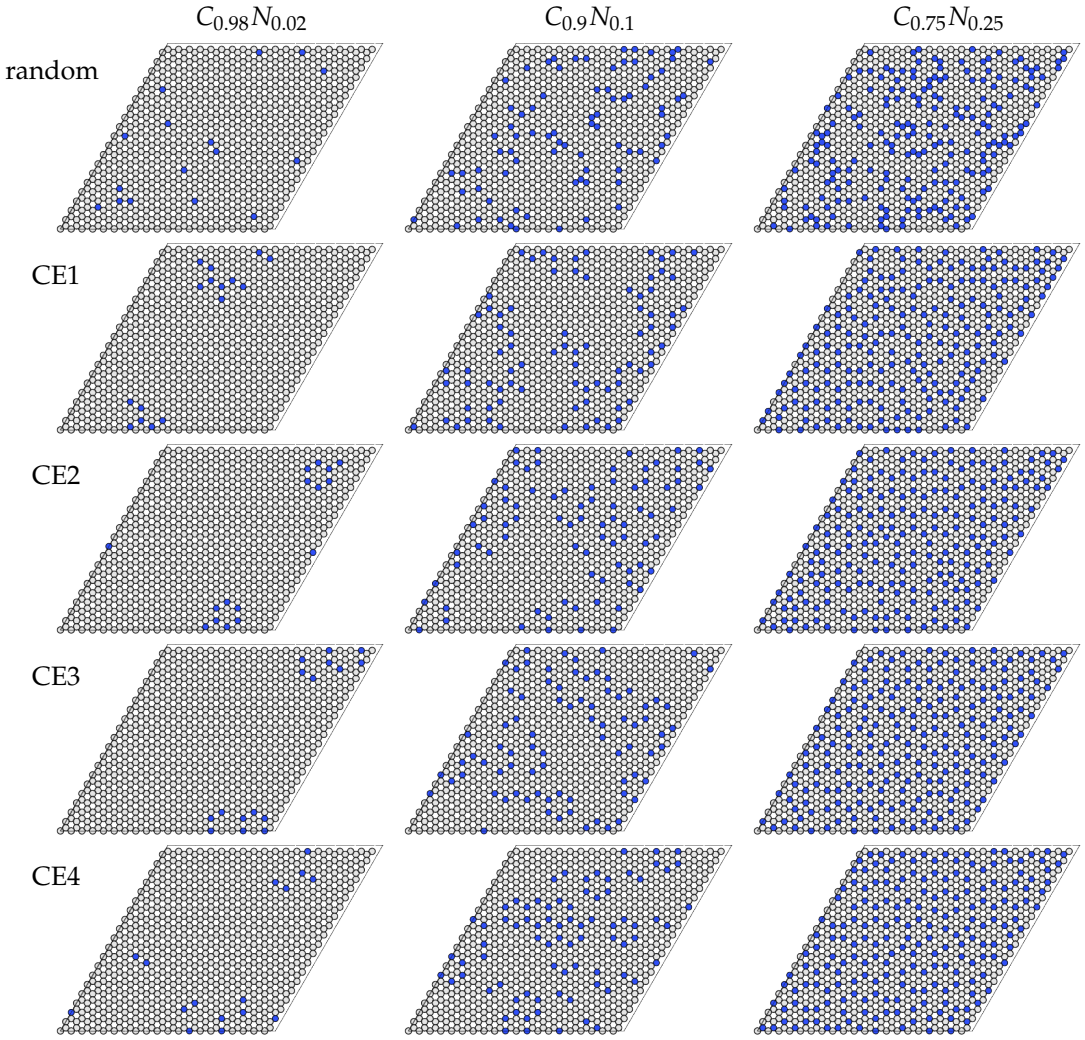


FIGURE 3.6: Evolution of nitrogen dopants (blue dots) distribution over the rigid carbon honeycomb lattice after performing Monte Carlo simulations (at $T = 10K$) for progressing generations of Cluster Expansion models (CE_i) presented for three exemplary concentrations of N, 2%, 10% and 25%. For each N concentration, initial distributions of nitrogen over the graphene lattice were assumed to be a random one.

It is apparent that the third coordination shell is the most favorable distance for N-N pairs on the 800 atom supercell. The W-C parameters Γ_{NN}^3 are negative markedly for all the concentrations. This means that probability to find two N atoms as 3^{rd} neighbors is much higher than it would eventuate from random distribution. Remaining W-C parameters Γ_{NN}^i (with $i \neq 3$) are positive, just indicating that the coordination shells 1, 2, and 4 are unfavorable for N-N pairs in the GL-N alloy. One can see this

also in Fig. 3.6, where configurations of atoms in GL-N alloy after Monte Carlo simulations in every CE model iteration are alike, and deviate significantly from random distribution.

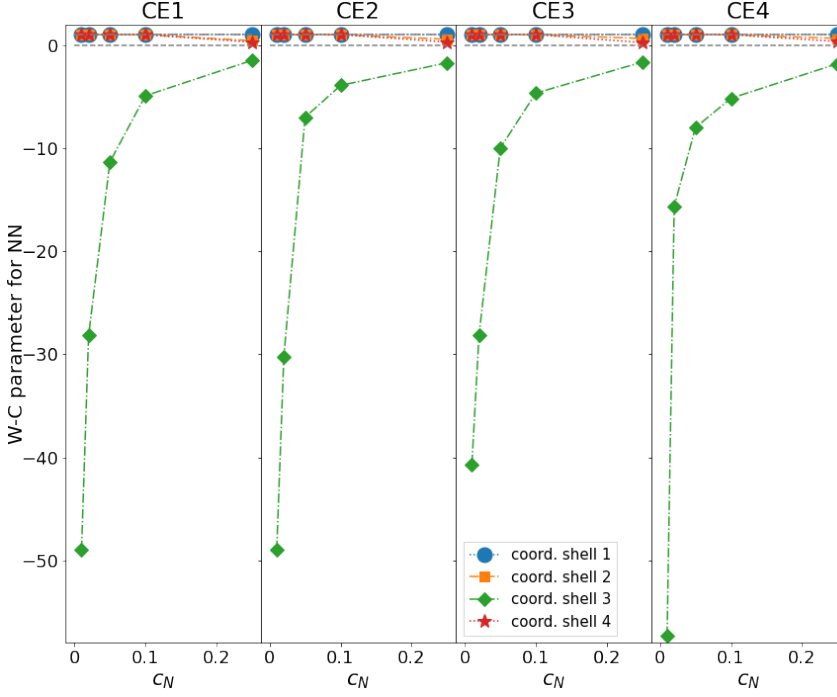


FIGURE 3.7: Warren-Cowley SRO parameters for N-N pairs extracted from minimum energy structures obtained in Monte Carlo Simulations with i^{th} generation of CE model. Parameters were calculated for 1st – 4th coordination shells.

Fig. 3.7 clearly demonstrates that the Warren-Cowley parameters do not exhibit stronger dependence on the CE model, and, in spite of the fact that the CE4 model is the best, any iteration of the CE model for GL-N alloy predicts correctly existence of the short range order (SRO) and reasonable magnitudes of the SRO parameter.

Having discussed in detail the procedure to generate the CE models for rigid GL-N lattice (**GL-N RIGID**), we would like now to conclude our efforts to generate the CE models for GL-B alloy on rigid lattice (**GL-B RIGID**) and GL-N and GL-B with full atomic relaxation, **GL-N RELAX**, and **GL-B RELAX**, respectively. The details concerning the *ECIs* and calculations of the Warren-Cowley parameters with the CE models on the 800 atom superlattices can be found in Appendix A.

CE model for graphene alloyed with boron on rigid lattice (GL-B RIGID).

First iteration CE1 has a very poor predicting power. CE2 gives better predictions for small systems, yet we observe clustering of B atoms in bigger supercell, for the highest investigated concentrations (25% of boron atoms). This problem does not occur in third iteration model CE3. It favors structures with B atoms situated on the third coordination shell, similarly to the result obtained for **GL-N RIGID** model. Thus, we suspect that this effect might derive from the size of the cell of the structures constituting reference sets, rather than from physical grounds. Final model in this class results in proper behavior of the first neighbors (B-B bonds are unfavorable), however predicting power of the model is worse than in nitrogen-alloyed class. We conclude that adding bigger clusters (*i.e.*, more than four substituted atoms, or with bigger radius), and further extension the reference set for the fitting procedure would be beneficial.

CE model for graphene alloyed with nitrogen on relaxed lattice (GL-N RELAX).

In this class, we did not obtain satisfactory predicting power in any of the iterations. First iteration model CE1 gives completely wrong predictions, favoring arrangements with domains of nitrogen within graphene lattice. Configurations resulting from Monte Carlo simulations with models CE1-CE3 do not present much lower DFT energy compared to initial distributions, in fact, many final configurations result in increased DFT energies. Also we do not get any instructive information from analysis of Monte Carlo results in bigger systems (800 atoms), every model gives different predictions on ordering, see Figs. A.6-A.7. We conclude then, that the reference data that consists of energies of structures after geometry and volume relaxation is harder to be reproduced in Cluster Expansion approach, where energy depends only on arrangement of atoms. It is also possible that iterative procedure presented in this thesis is not appropriate method, and CE model could be fitted accurately, but with broader set of reference structures, and/or a bigger Clusters Pool.

CE model for graphene alloyed with boron on relaxed lattice (GL-B RELAX).

Data obtained here suggests that (as with **GL-N RELAX**) fitting CE model using procedure presented above is a difficult target (if achievable at all). First fitted model CE1 gives also invalid predictions, however, subsequent iterations improve predicting power of the model. Nonetheless, third iteration model is still far from acceptable, thus, as in respective nitrogen case - alternative fitting strategy might be more successful.

In conclusion, fitting accurate Cluster Expansion models for N- or B-alloyed graphene class of structures seems to be better accessible for models on the rigid lattice than for relaxed systems. Cluster Expansion method proved great performance in many bulk systems, but the question arises, if energy landscape of doped, layered graphene systems affected considerably by lattice distortions for higher concentrations of substituted atoms are possible to be mapped into Cluster Expansion Ising-like hamiltonian dependent only on arrangement of species. Our experience shows that one of the reasons for the slower convergence of the CE models for relaxed lattices lies in the relatively small dimensions of the superlattices (32 atoms) used to generate DFT reference data. Our calculations show that in the GL-B alloy with higher B concentrations the lattice constant of the 32 atom supercell increases by up to 4.6% in comparison to pristine graphene. It would be, therefore, desirable to realize such B concentrations on the graphene supercell containing few times more atoms, which would increase the computational burden dramatically. Nevertheless, all the studied CE models of GL-N and GL-B lead to clear prediction of the SRO in these alloys. The DFT calculations performed for the CE models, can be also used to verify the parametrization of the VFF models that will be used later on for the description of CBN 2D systems.

In the last section of this chapter, we would like to use the CE model on superlattice with 800 lattice points to calculate the free energy of the GL-N and GL-B alloys, and make an attempt to construct the phase diagram for these alloys.

3.4 Results of the Cluster Expansion models for GL-N and GL-B on rigid lattice

We use Cluster Expansion models for N- and B-doped graphene on rigid lattice (**GL-N RIGID** and **GL-B RIGID**) with N (B) concentrations of 1% - 30%. We consider the 800 points honeycomb lattice and use the CE hamiltonian to find the atomic configurations with the minimal total energy. To do this, we employ the Monte Carlo Method with Metropolis algorithm choosing the temperature in the simulation typically to 10K and employing the random distribution of dopants as an initial

guess. However, we would like to stress that we performed Monte Carlo simulations with higher temperatures (even up to few thousands K), but Metropolis algorithm in each case was able to find atomic positions leading to the minimal total energy.

Averaged mixing energies of structures before and after MC simulations are presented for both GL-N and GL-B systems in Fig. 3.8. In both cases (GL-N and GL-B), the structures exhibiting short range order lead to the lower energies than in configurations with completely random distribution of dopants.

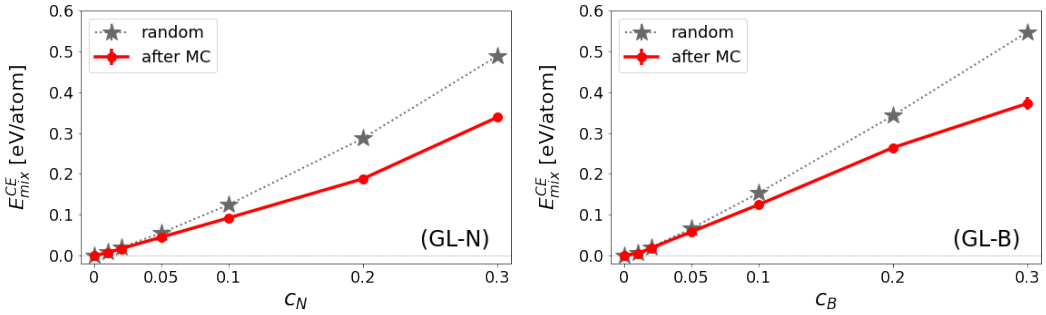


FIGURE 3.8: Average mixing energy for the GL-N and GL-B structures for two distributions of atoms: (i) random, and (ii) after Monte Carlo Metropolis simulations with Cluster Expansion models fitted to energies on rigid lattice.

For random arrangement of atoms, the entropy of mixing is given by the famous formula $S_{BW} = -k_B \sum_i c_i \ln(c_i)$ that for a binary alloy acquires the form $S(c) = -k_B(c \ln c + (1 - c) \ln(1 - c))$, taking maximum equal to $2k_B$ for $c = \frac{1}{2}$. This form of entropy, very often is named Bragg-Williams (BW) point approximation. In the case of the presence of the short range order, the entropy expression requires to be modified. The lowest order correction to the configurational entropy can be ascribed to the corrections originating from the pair-pair correlations, and the entropy in the presence of the short range order and in the so-called pair approximation can be written as $S_{pair} = k_B(2 \sum_i c_i \ln(c_i) - \frac{3}{2} \sum_{i,j} p_{ij} \ln(p_{ij}))$ (Eq. (2.48)), where the coordination number of $Z = 3$ is used in this expression. The averaged configurational entropies in the point and pair approximations, $S_{BW}(c_x)$ and $S_{pair}(c_x)$, as a function of dopants concentration c_x ($x = N, B$) for GL-N and GL-B alloys are plotted in Fig. 3.9. Now one is in position to determine the free energy for the alloy system with short range order as $F_{mix}^{CE} = E^{CE} - TS_{pair}$.

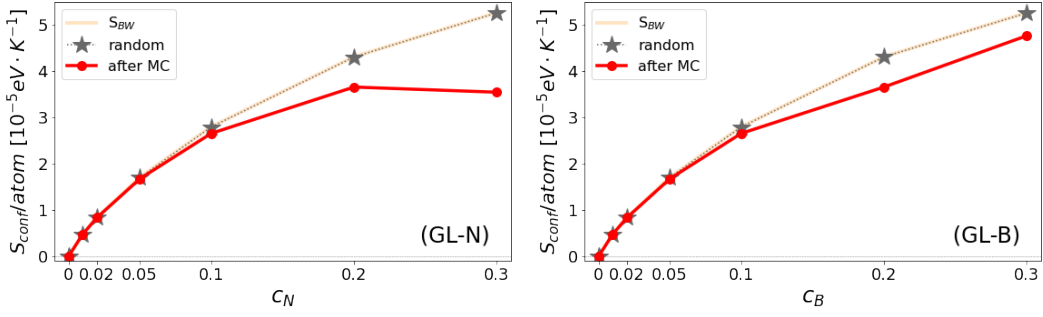


FIGURE 3.9: Average entropy S_{pair} for the GL-N and GL-B structures random and after Monte Carlo Metropolis simulations with Cluster Expansion models fitted to energies on rigid lattice. For comparison, entropy for randomly distributed systems as well as values of Bragg-Williams point approximation is also plotted.

These free energies as a functions of N(B) concentrations in the GL-N (GL-B) alloys, and for the temperature range between 0 and 9000 K are depicted in Fig. 3.10. It is clearly seen than the negative free energy of the alloys GL-N and GL-B can only appear for temperatures exceeding 3000K, so very close to the melting temperature. Therefore, these cluster expansion calculations clearly corroborate numerous experimental suggestions that incorporation of nitrogen and boron into graphene lattice can happen only through non-equilibrium processes. It could be of interest, to plot the free energy as a function of N (or B) concentrations assuming an hypothetical completely random alloy, *i.e.*, with the total energy obtained from the random distribution of atoms and point approximation for the entropy. Such free energy for GL-N and GL-B is plotted in Fig. 3.11 and resembles similarities to the free energy computed taking the short range order into account and presented in Fig. 3.10.

In the summary of this chapter, we would like to point out that cluster expansion methodology provides insights into physics of the GL-N and GL-B binary alloys, and will be used as some kind of guidelines for assertion of the VFF models.

We are aware that there are earlier papers employing the Cluster Expansion methodology to systems being alloys of h-BN and graphene, of the formula $(C_2)_x(BN)_{1-x}$. For example, in (S. Du *et al.*, 2020) the authors create quasi-ternary CE model for C-BN alloys, yet assuming BN pairs as one ‘point’. There are also other CE studies of C-BN alloys (Yuge, 2009). We will relate their findings with results for alloys obtained within VFF approach in Chapter 5 of this thesis.

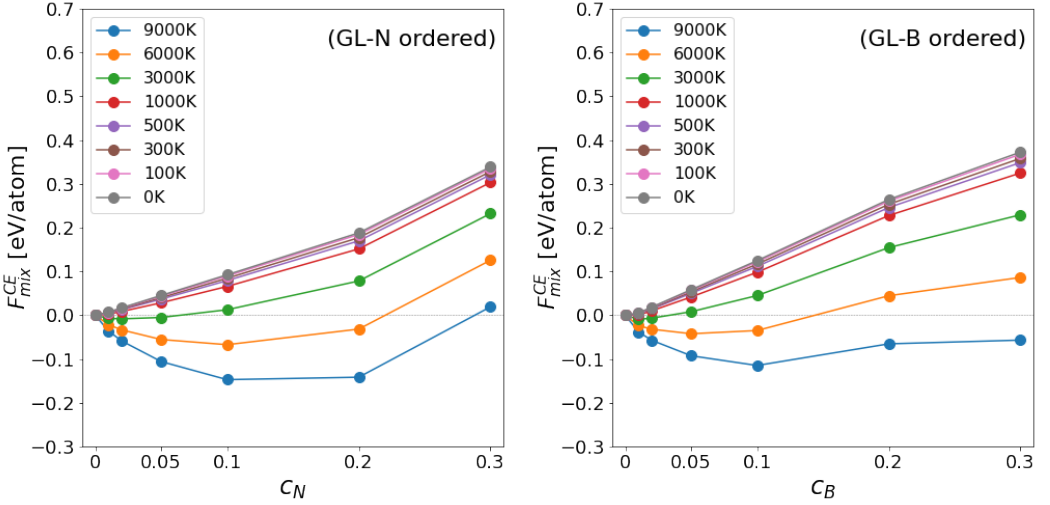


FIGURE 3.10: Dependence of average free energy F_{mix}^{CE} on temperature for GL-N (on the left) and GL-B (on the right) alloys exhibiting short range order as obtained from Monte Carlo Metropolis simulations with respective CE models on rigid lattice.

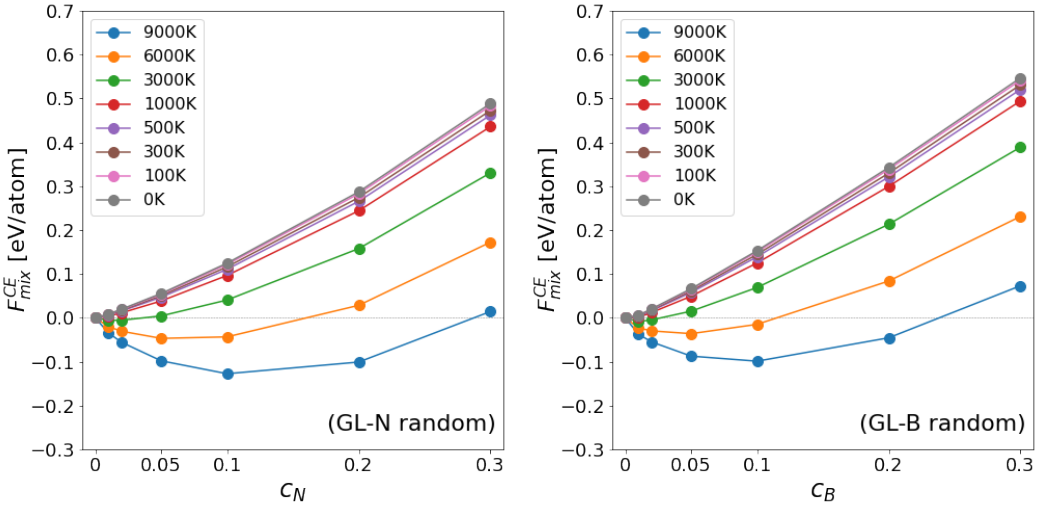


FIGURE 3.11: Dependence of average free energy F_{mix}^{CE} on temperature for random GL-N (on the left) and GL-B (on the right) structures calculated with respective CE models on rigid lattice.

Next chapters are devoted to calculations of the energy, stability and ordering in CBN binary and ternary alloys obtained using Valence Force Field methods. The huge advantage of VFFs is possibility to describe energy of the system not only

depending on atomic configuration but also as a function of all atomic positions. Moreover, strength of interaction between two atoms depends on their surroundings (bond-order potentials). This allows not only to investigate periodic structures but also ribbons, flakes, *etc.*, as well as to perform simulations of bond creating and breaking, *e.g.*, in chemical reactions (Molecular Dynamics simulations). Geometry optimization performed with VFFs allows to inquire into morphology of systems with defects, grain boundaries, effects of strain and other mechanical properties. These topics are out of reach of Cluster Expansion methodology.

Chapter 4

Assesment of VFF methods for simulations of CBN alloys

Choice of the proper Valence Force Field (VFF) and its parametrization is paramount to the success of all the subsequent calculations. Driven by this motivation, the sections below present results of tests of various empirical potentials described in Sec. 2.3 in the area of hexagonal layered CBN alloys. This is preliminary step before performing large-scale simulations, but I find it the most important. It gives us the knowledge about which properties of a given range of materials can be calculated resulting in appropriate values and which aren't properly described. All the simulations involving VFF mentioned in this and the next chapter were performed with use of LAMMPS package (Thompson *et al.*, 2022), that is widely-used tool for modelling in multiscale systems.

4.1 Tersoff potential: available and developed parametrizations

Below, the parametrizations available for Tersoff and extended Tersoff potentials for systems containing C, B, and N atoms are shortly described.

- **Ters-SET0:** Matsunaga (Matsunaga *et al.*, 2000) parametrization derived for cubic carbon boron nitrides. Values of the parameters are listed in Tab. 4.1. Parameters for carbon have their source in original Tersoff paper (Tersoff, 1989),

parameters for nitrogen, and nitrogen-carbon interactions in Kroll PhD dissertation (Kroll, 1996), while boron part as well as parameters for heteroatomic bonds were introduced in the abovementioned paper. Although intended for cubic systems, this parametrization has been widely used for simulations in hexagonal systems as reported in literature, *e.g.*, in recent paper (Jam *et al.*, 2018).

- **Ters-SET1:** It is parametrization equal to **Ters-SET0** with corrections for carbon as reported in (Lindsay *et al.*, 2010). It was developed for graphene and carbon nanotubes. In comparison to **Ters-SET0**, only two parameters are varied: $B_C = 430.0$ and $h_C = -0.930$. This modification of C-C interactions has been used by various authors, *e.g.*, (Senturk *et al.*, 2019).
- **Ters-SET2:** This set is our combination of parameters by Lindsay (Lindsay *et al.*, 2010) for C-C interactions, Kinaci (Kinaci *et al.*, 2012) parameters for B-N, C-N and C-B (developed for h-BN/graphene interfaces in lateral graphene/h-BN junctions), and Albe (Albe *et al.*, 1997) for B-B and N-N interactions (first introduced for description of the hexagonal boron nitride-based systems). Values of the parameters are listed in Tab. 4.3. We propose this mix of parametrizations, gathering the best (in our opinion) descriptions for each pair of atomic types. It was not refitted as a whole set, though. In the following part of this chapter, we show decent overall performance of **Ters-SET2**.
- **Ters-BNC:** parametrization implemented in the LAMMPS package in the file named Ters-BNC. It is different from **Ters-SET2** only in description of N-N and B-B interactions. Specifically $A_{NN} = 128.86866$, $B_{NN} = 138.77866$, $A_{BB} = 40.0520156$, $B_{BB} = 43.132016$. We include this set for completeness of the thesis.
- **extep:** extended Tersoff potential parametrization for BN systems (Los *et al.*, 2017). The parameters are gathered in Tab. 4.2. This extended form of Tersoff original potential is described in Sec. 2.3.1. This VFF was developed specifically to describe accurately a wide range of BN-based structures, including proper description of defects in hBN. Although quite recent, this method has been employed by other authors *e.g.*, (Ghaderzadeh *et al.*, 2021).
- **hybrid: Ters-SET2 + extep.** In LAMMPS it is possible to calculate different interactions (or groups of atoms) with distinct potentials. Thus, we decided to test such a hybrid approach, where B-B, B-N and N-N interactions are calculated with **extep** while C-C, C-N and C-B - with **Ters-SET2**.

TABLE 4.1: Tersoff interatomic potential parameters following (Matsunaga *et al.*, 2000) constituting **Ters-SET0** parametrization.

Elements	C	N	B
A(eV)	$1.3936 \cdot 10^3$	$1.1000 \cdot 10^4$	$2.7702 \cdot 10^2$
B(eV)	$3.467 \cdot 10^2$	$2.1945 \cdot 10^2$	$1.8349 \cdot 10^2$
$\lambda(A^{-1})$	3.4879	5.7708	1.9922
$\mu(A^{-1})$	2.2119	2.5115	1.5856
β	$1.5724 \cdot 10^{-7}$	$1.0562 \cdot 10^{-1}$	$1.6000 \cdot 10^{-6}$
n	0.72751	12.4498	3.9929
c	$3.8049 \cdot 10^4$	$7.9934 \cdot 10^4$	$\cdot 10^{-1}$
d	$4.3484 \cdot 10^0$	$1.3432 \cdot 10^2$	$1.5870 \cdot 10^{-3}$
h	-0.57058	-0.9973	0.5000
R(A^{-1})	1.8	2.0	1.8
S(A^{-1})	2.1	2.3	2.1
	C-N	C-B	N-B
χ_{i-j}	0.9685	1.0025	1.1593
ω_{i-j}	0.6381	1.000	1.000

The parameters (excluding χ and ω) are determined for single elements. In the heterogeneous bonds the interpolation formulas are used $A_{ij} = \sqrt{A_i A_j}$, $B_{ij} = \sqrt{B_i B_j}$, $R_{ij} = \sqrt{R_i R_j}$, $S_{ij} =$

$$\sqrt{S_i S_j}, \lambda_{ij} = \frac{\lambda_i + \lambda_j}{2}, \mu_{ij} = \frac{\mu_i + \mu_j}{2}.$$

c, d, h determine the bond-angle dependence of bonding interactions between neighboring atoms. χ and ω take into account strengthening or weakening of heteropolar bonds. For identical species involved in the bond, they are unitary.

TABLE 4.2: Extended Tersoff interatomic potential parameters for BN systems following (Los *et al.*, 2017) constituting **extep** parametrization.

Interactions	B-B	N-N	B-N
A(eV)	3376.335	2978.952	3330.0655
B(eV)	2768.736	2563.560	2595.686
$\lambda(A^{-1})$	2.686	2.829	2.95
$\mu(A^{-1})$	2.521	2.627	2.693
β	1	1	1
n	1.142	0.665	0.658
c	26617.3	23.5	306.587
d	141.2	3.75	10
h	-0.13	-0.4	-0.7218
R(A^{-1})	1.8	1.8	1.8
S(A^{-1})	2.2	2.2	2.2

χ and ω equal one in this parametrization, as it defines explicitly heteroatom interactions.

TABLE 4.3: Tersoff interatomic potential parameters constituting **Ters-SET2**.
 $\omega_{ij} = 1$ for all ij , and $B' = B \cdot \chi$

$V_{ij} = V_{ji}$	C-C	N-N	B-B	C-N	C-B	N-B
$A(eV)$	1393.6	2978.95279	1404.4752	1386.78	1386.78	1380
$B'(eV)$	430.	2563.56	1173.197	387.575152	339.06891	340
$\lambda(A^{-1})$	3.4879	2.82931	2.23726	3.5279	3.5279	3.568
$\mu(A^{-1})$	2.2119	2.627272	2.0775	2.2054	2.2054	2.199
$R(A^{-1})$	1.80	1.90	1.90	1.85	1.85	1.90
$S(A^{-1})$	2.10	2.10	2.10	2.05	2.05	2.00
$V_{ij} \neq V_{ji}$	C-x	N-N	B-B	other N-x	other B-x	
β	1.5724e-7	0.019251	1.6e-6	1.25724e-7	1.25724e-7	
n	0.72751	0.6184432	3.9929061	0.72751	0.72751	
c	38049.	17.7959	0.52629	25000.	25000.	
d	4.3484	5.9484	0.001587	4.3484	4.3484	
h	-0.93	0.	0.5	-0.89	-0.89	

4.2 ReaxFF: available parametrizations

Each ReaxFF parametrization is developed for the specific group of materials, and the transferability must be checked each time before performing large-scale simulations. The most of the available sets of parameters are reviewed in (Senftle *et al.*, 2016). Below, the parametrizations concerning the CBN systems are mentioned and shortly described, paying attention to the training sets that were used in each case, and results obtained by other authors for systems close to our range of interest.

1. **Reax-C2013**: parameters for carbon (Srinivasan *et al.*, 2015) (named $\text{ReaxFF}_{\text{C-2013}}$). It was designed for modeling carbon condensed phases. Its training set was based on previous set for hydrocarbons (Chenoweth *et al.*, 2008), and extended with data for graphite and diamond. The performance of $\text{ReaxFF}_{\text{C-2013}}$ was revised broadly in (Jensen *et al.*, 2015).
2. **Reax-PAI**: C\B\N parameters reported in (Pai *et al.*, 2016): $\text{ReaxFF}_{\text{Reax-PAI}}$. This force field was intended for studying liquid CBN hydrogen-storage materials. The authors targeted molecular systems that contained BN heterocycles with C, H, B and N atoms included. $\text{ReaxFF}_{\text{Reax-PAI}}$ was utilized in the simulations of hydrogenated hBN-graphene (Dethan *et al.*, 2020), but the validity of this result remains unknown.

3. **Reax-PAUP**: C\B\N parameters from (Paupitz *et al.*, 2014): ReaxFF_{PAUP}. This parametrization was fitted for the simulations of B/N containing fullerenes. ReaxFF_{PAUP} was utilized for MD simulations of B- and/or N- doped graphene oxide in (Bawari *et al.*, 2020). Also, hBN nanosheets doped with -OH groups were investigated using this parametrization in (B. B. Sharma *et al.*, 2020). The validity of the parametrization was evaluated there by estimating a characteristic fracture toughness of pristine hBN nanosheet. No evaluation of geometry results was reported.

Unfortunately, none of the sets presented above was trained specifically against CBN layered, hexagonal structures. We do not list here values of respective ReaxFF parameters, due to their vast count, whole set can be found in the cited papers.

4.3 COMB3: available parametrizations

Parameters of the COMB3 potential were developed for several elements, starting from C-O-H parametrization (Liang, Devine, *et al.*, 2012). Good transferability of this set was shown in (Fonseca, Liang, D. Zhang, Choudhary, and Sinnott, 2016), where the authors performed simulations of graphene oxide (GO) systems, resulting in quite good predictions of chemical and structural properties of GO, despite it hasn't been included to the training set. Martinez and coworkers (Martinez *et al.*, 2017) developed nitrogen parameters. This allows us for calculations of graphene alloyed with N, but up to date, there is unfortunately no parametrization for boron. Clearly, it limits availability of COMB3 potential for studies of CBN alloys in the present thesis. COMB3 is the most promising formulation of VFF that we have found in literature at the beginning of our research, as one needs only to generate one set of parameters for each species, that should be applicable to various geometries and modelled materials. Therefore, we decided to include COMB3 in our tests (henceforth referred to as **COMB**), hoping that in the near future parametrization for boron would be developed as well. We do not present values of C and N parameters here, due to their number, exceeding even ReaxFF list of parameters.

The remaining part of this chapter is devoted to the tests of all the VFFs with the parametrizations described above. We will refer to the each VFF with specific parametrization using short names, bolded throughout their presentation.

4.4 Tests of the empirical potentials

In order to make the assessment of the suitability of the empirical VFF methods considered in the thesis, we have performed a series of tests for:

- (1) dimer molecules consisting of C, B, and N atoms,
- (2) pristine 2D structures, *i.e.*, graphene (GL) and hexagonal boron nitride (hBN) monolayers,
- (3) GL and hBN with point defects.

During the tests, we compare results of all VFFs with results of *ab initio* simulations. Our reference DFT data was calculated with Quantum Espresso simulation package. The total energy and geometry was obtained with the use of Perdew, Burke and Ernzerhof (PBE) exchange-correlation functional of the GGA type, together with Projector-Augmented Wave (PAW) pseudopotential type (with non-linear core correction, during the full relativistic calculations with spin). For (1) we use $10\text{\AA} \times 10\text{\AA} \times 10\text{\AA}$ supercell containing 2 atoms. For the 2D structures in (2) are calculated using unit cell with $12\text{\AA}$ of vacuum in direction perpendicular to the layer in order to avoid spurious interlayer interaction. Point defects (3) simulations are carried out in 4×4 graphene/hBN supercell (32 atoms), with $12\text{\AA}$ distance between the layers. Kinetic energy cutoff for wavefunctions is set to (1 and 2) 100 Ry, and (3) 80 Ry. The Monkhorst-Pack scheme is used for the Brillouin zone sampling, with the number of k-points dependent on the system size: (1) $4 \times 4 \times 1$, (2) $8 \times 8 \times 1$, (3) $4 \times 4 \times 1$. During the geometry relaxation, the convergence threshold was set to 10^{-5} Ry in energy and 10^{-4} Ry/Bohr in force. In the following sections, we will refer to these DFT simulations as to **QE-DFT**.

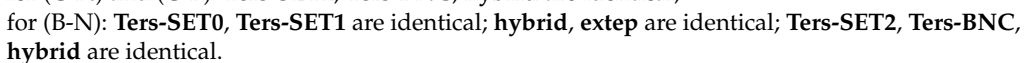
(1) Dimers

The simplest possible interaction is that occurring between two atoms. There are six possible pairs consisting of C, B, and N. For all of them, we calculate cohesive energy as a function of interatomic distance and depict it in Fig. 4.1. The optimum bondlengths and respective energies are summarized in the Tab. 4.4 for all VFFs, followed by the DFT results (**QE-DFT**) and the values reported in the literature.

TABLE 4.4: Equilibrium binding energy and bond length for dimer molecules built of C, B and N atoms, calculated using the methods and parametrizations tested in the present thesis - results of Tersoff, ReaxFF and COMB force fields. For comparison the values of E_{bond} and d_{bond} obtained from DFT calculations **QE-DFT** and by other authors are depicted.

	C-C		N-N		B-B		C-N		C-B		B-N	
	E_{bond} [eV]	d_{bond} [Å]	E_{bond} [eV]	d_{bond} [Å]	E_{bond} [eV]	d_{bond} [Å]	E_{bond} [eV]	d_{bond} [Å]	E_{bond} [eV]	d_{bond} [Å]	E_{bond} [eV]	d_{bond} [Å]
Ters-SET0	-5.164	1.447	-3.197	1.456	-3.084	1.574	-3.963	1.481	-4.460	1.505	-5.654	1.448
Ters-SET1	-9.304	1.278	-3.197	1.456	-3.084	1.574	-4.937	1.433	-6.333	1.376	-5.654	1.448
Ters-SET2	-9.304	1.278	-9.910	1.110	-3.080	1.590	-7.920	1.319	-5.544	1.420	-6.317	1.377
Ters-BNC	-9.304	1.278	-9.910	0.000	-3.080	0.000	-7.920	1.319	-5.544	1.420	-6.317	1.377
extep			-9.910	1.110	-3.080	1.590					-6.360	1.325
hybrid	-9.304	1.278	-9.910	1.110	-3.080	1.590	-7.920	1.319	-5.544	1.420	-6.360	1.325
COMB	-6.275	1.323	-9.798	1.099			-8.569	1.177				
Reax-C2013	-7.846	1.204										
Reax-PAI	-9.287	1.260	-8.123	1.096	-3.715	1.919	-12.118	1.088	-3.432	1.471	-8.197	1.096
Reax-PAUP	-11.498	1.153	-10.090	1.056	-3.218	1.724	-11.061	1.090	-6.872	1.348	-9.023	1.091
QE-DFT	-6.252	1.262	-10.350	1.110	-2.486	1.634	-7.997	1.178	-3.879	1.434	-4.840	1.276
literature	-6.25 (a)	1.312 (a)	-9.763 (a)	1.098 (a)	-3.0 (g)	1.590 (b),(c)	-7.721 (a)	1.117 (a)	-4.597 (d)			1.281 (b)
		1.242 (b)		1.116 (b)				1.172 (b)				
		1.240 (c)		1.108 (b)								
	-6.244 (d)											
(QMC)	1.242 (e)		-6.612 (e)	1.094 (e)	-9.902 (e)	1.590 (e)	-2.917 (e)					
(OF-DFT)	1.408 (f)			1.106 (f)		1.630 (f)		1.180 (f)				
(exp)	1.244 (f)			1.095 (f)		1.587 (f)		1.169 (f)				

Values from the following papers: (a)(Kaye *et al.*, 2021), (b) (Huber, 1979), (c) (Jemmis *et al.*, 2006), (d) Ref. Verhaegen *et al.*, 1964, (e) (Filippi *et al.*, 1996), (f) (Finzel, 2020).



We would like to assess this test with the following remarks: **COMB** gives the best results for C-C, N-N and C-N dimers. Unfortunately, we lack parametrization for boron. **Ters-BNC** parametrization gives completely wrong N-N and B-B interactions that are vital in the simulations in general CBN alloys. Therefore, we exclude this method from the further tests. **Ters-SET0** and **Ters-SET1** do not perform very well, especially for N-N and C-N. **Ters-SET2** and **hybrid** give quite similar predictions (parameters are partly identical). These two perform the best among the parametrizations covering all the C-B-N interactions. **Extep** gives good results in B-N, N-N and B-B dimers, and **hybrid** is equal to **extep** for B-N, N-N and B-B, while remaining C-C, C-B and C-N are equal to **Ters-SET2**. **Reax-PAI** and **Reax-PAUP** parametrizations are acceptable for N-N, while for the other pairs, none of the values of minimum energy and bond length or the shape of the curve are satisfactory. **Reax-C2013** performs best for C-C out of all ReaxFF parametrizations, but it is not the best result in the group of all considered potentials.

(2) Pristine graphene and hexagonal boron nitride

Graphene and hexagonal boron nitride monolayers are base structures for the present thesis, so using empirical potential describing them properly is of the greatest importance. Hence, we calculate energy of GL and hBN as a function of lattice constant using all VFFs to determine equilibrium lattice parameter as well as the cohesive energy E_{coh} in each of the methods. Tab. 4.5 compares VFF energies to **QE-DFT** and available literature data, moreover dependence of cohesive energy on the lattice constant is shown in Fig. 4.2.

Remarks: In this test, **COMB** performs very well for GL and so are (identical this case) **Ters-SET1**, **Ters-SET2**, and **hybrid**. **ReaxFF-PAI** and **ReaxFF-PAUP** (nearly overlapping here) are much worse again, **ReaxFF-C2013** is acceptable but note that it was fitted to different *ab initio* results, differing from **QE-DFT** values. In hBN, **extep**, and **hybrid** (identical here) give excellent results for the lattice constants greater than 2.2 Å, other Tersoff sets are in good agreement with **QE-DFT** as well. Cohesive energy curves sharply rise with decreasing distance for smaller values, but we are not likely to be in this range of lattice constants in our simulations.

TABLE 4.5: Equilibrium lattice constant and cohesive energy of graphene and monolayer hexagonal boron nitride, calculated using empirical methods tested in the present thesis, compared to **QE-DFT** results and the values reported in the literature.

	graphene		hBN	
	$E_{coh} [\frac{eV}{atom}]$	$a_{lat} [\text{\AA}] (d_{bond})$	$E_{coh} [\frac{eV}{atom}]$	$a_{lat} [\text{\AA}] (d_{bond})$
Ters-SET0	-7.393	2.530 (1.461)	-7.565	2.560 (1.478)
Ters-SET1	-7.978	2.492 (1.439)	-7.565	2.560 (1.478)
Ters-SET2	-7.978	2.492 (1.439)	-7.507	2.498 (1.442)
extep			-6.690	2.504 (1.446)
hybrid	-7.978	2.492 (1.439)	-6.690	2.504 (1.446)
COMB	-7.352	2.462 (1.421)		
Reax-C2013	-7.524	2.462 (1.422)		
Reax-PAI	-8.582	2.502 (1.444)	-7.129	2.554 (1.474)
Reax-PAUP	-8.589	2.501 (1.444)	-7.392	2.571 (1.485)
QE-DFT	-7.854	2.468 (1.425)	-6.994	2.515 (1.451)
literature (theory)	-7.464 (QMC) ^(a)			2.45 (DFT-HSE06) ^(b)
	-7.906 (PBE-PP) ^(a)			2.491 (DFT-LDA) ^(b)
	-7.972 (PAW) ^(a)			2.512 (DFT-PBE) ^(b)
		2.465(DFT-PAW) ^(c)	-7.072 ^(d)	2.501 (DFT-PAW) ^(d)
literature (exp)		2.459 (1.42) ^(e)		2.49 (1.44) ^(f)

(a) H. Shin *et al.*, 2014, (b) Hunt *et al.*, 2020, (c) Pozzo *et al.*, 2011, (d) Yin *et al.*, 2010,

(e) Cooper *et al.*, 2012, (f) C. Jin *et al.*, 2009.

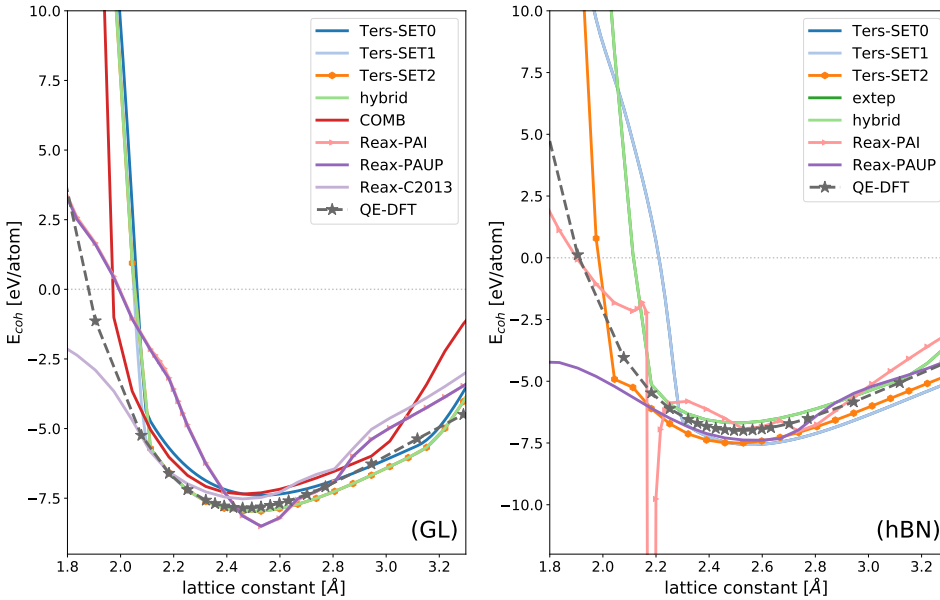


FIGURE 4.2: Energy of pristine (a) GL, and (b) hBN monolayer as a function of the lattice constant calculated using various empirical potentials and compared to the **QE-DFT**.

Note that for (GL): **Ters-SET1**, **Ters-SET2**, **hybrid** are identical;

for (hBN): **Ters-SET0**, **Ters-SET1** are identical; and **hybrid**, **extep** are identical.

Fig. 4.2 does not show the full course of the **Reax-PAI** energy. There appears some discontinuity, that results from the parameters' values. However, it is worth noticing that **Reax-PAI** was fitted to the small molecules and liquid phases of C-B-N. Evidently it is not suited for layered CBN structures.

Planar GL and hBN structures from calculations discussed so far are the result of assumed space groups: P6/mmm and P6₃, respectively. Thus, we perform also simulations in 4x4 supercell (32 atoms) with slightly distorted initial atomic positions, which induces the reduced symmetry. For a range of lattice constants (2.252-2.736 Å), we examine when the systems after geometry optimization start to corrugate (*e.g.*, as a method to release stress). Maximum lattice constant for which deviation from planar occurs is denoted a_{lat}^{corr} . The results of calculations with all VFFs are summarized in Tab. 4.6 and compared to the DFT result.

TABLE 4.6: Maximum values of the lattice constant in graphene and hBN structures for which the corrugation occurs after geometry optimization (simulations performed in the range of lattice constants 2.252-2.736 Å). For lattice constants larger than a_{lat}^{corr} , structures are planar.

	graphene	hBN
	a_{lat}^{corr} [Å]	a_{lat}^{corr} [Å]
Ters-SET0	2.460	2.546
Ters-SET1	2.321	2.546
Ters-SET2	2.321	2.460
extep		2.460
hybrid	2.321	2.460
COMB	2.321	
Reax-C2013	2.390	
Reax-PAUP	2.511	2.563
Reax-PAI	2.511	>2.736
QE-DFT	2.373	2.442

Remarks: We observe quite good agreement between **QE-DFT** and all Tersoff results, especially **Ters-SET2** and **hybrid**. **COMB** and **Reax-C2013** are also doing well in graphene case. Importantly, the remaining ReaxFF parametrizations favor (energetically) planar structures in the examined range of lattice constants. However,

a small initial displacement of atoms leads to corrugated final structure, with the energies higher by $\approx 3 - 5$ eV/atom, which indicates energetic instability of the system.

Reax-PAI and **Reax-PAUP** parametrizations show improper behavior in tests (1) and especially (2). It is hard to expect accuracy in the systems with in GL/hBN with defects or dopants. Thus, we exclude them from the further considerations. Despite promising idea to describe atomic interactions, ReaxFF potential demands careful fitting to serve as a useful tool, unfortunately it is out of scope of the present thesis.

(3) Point defects in graphene and hexagonal boron nitride

In the third test, we study energy and geometry of point defects in graphene: single boron substitution - denoted **GL: B_C**, single nitrogen substitution **GL: N_C**, single vacancy **GL: V_C** and BN pair substitution **GL: B_CN_C**. In hexagonal boron nitride we examine: single carbon on nitrogen sublattice **hBN: C_N**, single carbon on boron sublattice **hBN: C_B**, single nitrogen vacancy **hBN: V_N**, and single boron vacancy **hBN: V_B**. The geometry of these defects optimized with **QE-DFT** is shown in Fig. 4.3.

We analyze formation energy of defects E_f in GL or hBN as defined in Eq. (2.43) (description of components is given there)

$$E_f(\text{GL}_d) = E_{tot}(\text{GL}_d) - E_{tot}(\text{GL}) + n_C\mu_C + n_B\mu_B + n_N\mu_N$$

$$E_f(\text{hBN}_d) = E_{tot}(\text{hBN}_d) - E_{tot}(\text{hBN}) + n_C\mu_C + n_B\mu_B + n_N\mu_N.$$

Important part of above equations are chemical potentials μ_i . They are equal to the Gibbs free energy that isolated atoms i exchange with the heat reservoir, and they *should* represent experimental growth conditions. The ultimate values of the chemical potentials are denoted μ_i^0 (with $i = \text{C, B, N}$), representing the energies of the most stable phases of the elements. In our simulations, μ_C^0 is referenced to as energy per atom in graphene. Quantum Monte Carlo simulations (H. Shin *et al.*, 2014) show that differences in cohesive energies of graphite, graphene, and diamond are very small: 7.520, 7.464, and 7.503 eV/atom, respectively).

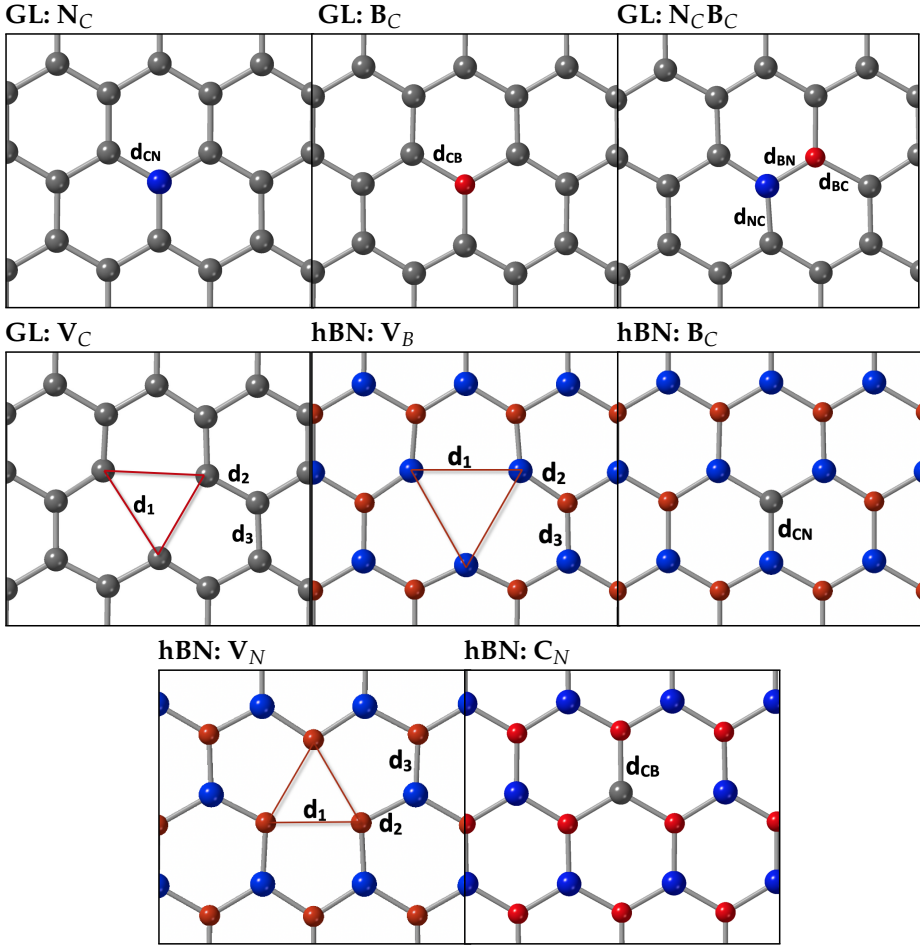


FIGURE 4.3: Geometry of point defects in graphene and hBN after QE-DFT geometry optimization. Gray dots denote carbon, blue - nitrogen and red - boron. We refer to distances marked on the figures in Tabs. 4.8-4.9.

For boron, we select several possible two-dimensional structures as reported in literature (D. Li *et al.*, 2020) and compare their cohesive energies. QE-DFT as well as all VFF methods result in the lowest energy for α -borophene, thus we set it as μ_B^0 . Finally, μ_N^0 is the energy per atom of nitrogen molecule. Nonetheless, it is quite arbitrary to choose the reference structures for μ_X^0 , see, *e.g.*, (Azevedo, 2006; Miranda *et al.*, 2021), where authors choose alternative structures for determination of μ_i^0 . Also in (McDougall *et al.*, 2017) chemical potential of carbon is computed from CH₄ molecule, in contrast to many other papers where graphene/graphite is taken as a reference.

The values of computed chemical potentials using all the investigated methods, compared to the **QE-DFT** are presented in Tab. 4.7.

TABLE 4.7: Values of maximal chemical potential for C, N, and B in the computational approaches investigated in this thesis, followed by the BN pair energy in hBN - $E_f(\text{BN})$, and hBN enthalpy of formation $\Delta H_f(\text{BN})$.

	$\mu_{\text{C}}^0 [\frac{\text{eV}}{\text{atom}}]$	$\mu_{\text{N}}^0 [\frac{\text{eV}}{\text{atom}}]$	$\mu_{\text{B}}^0 [\frac{\text{eV}}{\text{atom}}]$	$E_f(\text{BN}) [\frac{\text{eV}}{\text{pair}}]$	$\Delta H_f(\text{BN}) [\text{eV}]$
Ters-SET0	-7.393	-1.5986	-7.855	-15.129	-5.676
Ters-SET1	-7.978	-1.5986	-7.855	-15.129	-5.676
Ters-SET2	-7.978	-4.955	-7.413	-15.014	-2.645
extep		-4.955	-4.757	-13.380	-3.668
hybrid	-7.978	-4.955	-4.757	-13.380	-3.668
COMB	-7.352	-4.899			
Reax-C2013	-7.524				
Reax-PAI	-8.588	-4.062	-4.959	-14.258	-5.238
ReaxPAUP	-8.582	-5.045	-6.108	-14.785	-3.632
QE-DFT* (E_{coh})	-7.854	-5.171	-5.836	-13.989	-2.981
QE-DFT (E_{tot})	-250.943	-383.912	-155.507	-542.400	-2.981

* here we present cohesive energy $E_{\text{coh}} = E_{\text{tot}} - \sum E_{\text{atomic}}$ in addition to total energy resulting from DFT calculations, for better comparison to empirical potentials, that only give interaction strength between the atoms, and the energy of an isolated atom is always set to zero.

In the *real* growth conditions, the chemical potential μ_i of element i can differ from its maximum value μ_i^0 , and this difference is commonly indicated as $\Delta\mu_i$. Considering chemical compounds, like hexagonal boron nitride in our study, the thermodynamic equilibrium (S.-H. Wei *et al.*, 2000) is achieved when the differences in chemical potentials for two elements are equal to the compound's enthalpy of formation ΔH_f . For B and N pair, this rule reads

$$\Delta\mu_{\text{B}} + \Delta\mu_{\text{N}} = \Delta H_f(\text{BN}) = E_{\text{pair}}(\text{BN}) - \mu_{\text{B}}^0 - \mu_{\text{N}}^0, \quad (4.1)$$

where $\Delta H_f(\text{BN})$ is the enthalpy of formation of BN pair in hBN, and $E_{\text{pair}}(\text{BN})$ - the total energy of BN pair in hBN.

In considerations of doped hBN systems, we look at the two limiting cases:

- N-rich conditions (*i.e.*, when the system is in equilibrium with N₂ atmosphere), with $\mu_N = \mu_N^0$, which gives $\Delta\mu_N = 0$ and $\Delta\mu_B = E_{pair}(BN) - \mu_N^0$,
- N-poor (B-rich) conditions, when $\Delta\mu_B = 0$ and $\Delta\mu_N = E_{pair}(BN) - \mu_B^0$.

The knowledge of chemical potentials is necessary to relate the growth conditions of a material to the probability of defect appearance in the structure. Here, we attempt to describe only the relations between chemical potentials in equilibrium. However, typically the growth process does not occur in equilibrium conditions. The knowledge of chemical potentials is also important for comparison of the results obtained in this work with results of other authors and particularly with experimental data. For the studies of defects in graphene, we employ the simplified version of the formation energy Eq. (2.43) that we denote as E_f^0 (zero-formation energy), where we assume $\mu_i = \mu_i^0$

$$E_f^0(GL_d) = E_{tot}(GL_d) - E_{tot}(GL) + n_C\mu_C^0 + n_B\mu_B^0 + n_N\mu_N^0 \quad (4.2)$$

Tabs. 4.8-4.9 compare formation energies and significant interatomic distances around the defects as marked in Fig. 4.3 calculated with various empirical potentials. For comparison, **QE-DFT** results and the values reported in literature are also listed.

Remarks:

- **GL: B_C.** **Ters-SET2** and **hybrid** give the best agreement with **QE-DFT**, while in **Ters-SET0** and **Ters-SET1** boron atom sticks out of plane in contrast to DFT geometry.
- **GL: N_C.** **Ters-SET2** and (identical) **hybrid** give the best N-C bond length; **Ters-SET0**, **Ters-SET1** and **COMB** overestimate remarkably N-C bond (1.5 Å in contrast to 1.415 Å resulting from **QE-DFT**).
- **GL: B_CN_C.** **Ters-SET2** performs the best, while **Ters-SET0** and **Ters-SET1** again result in distorted structure. **Hybrid** returns slightly worse energy and bond lengths estimations.
- **GL: V_C.** In contrast to previous defects, **QE-DFT** optimum geometry is slightly corrugated, and the distances between three carbon atoms neighboring the vacancy (d_1 in Fig. 4.3 for GL: V_C) are not equal. Initially symmetric defects reconstruct into a form of 5- and 9-membered pair of rings, known as Jahn-Teller

TABLE 4.8: Cohesive energies per atom, zero-formation energies (following Eq. (4.2)), and the values of characteristic interatomic distances shown in Fig. 4.3 for various defects in graphene. The results obtained with tested here VFF models are compared to the DFT calculations in 4x4 (32 atoms) superlattices employing Quantum Espresso numerical package (**QE-DFT**). The results of DFT calculations available in literature are also presented.

GL: Bc					GL: Bc Nc				
	$E_{coh} [\frac{eV}{atom}]$	$E_f^0 [eV]$	d_{BC}	$E_{coh} [\frac{eV}{atom}]$	$E_f^0 [eV]$	$d_{BC} [\text{\AA}]$	$d_{NC} [\text{\AA}]$	$d_{BN} [\text{\AA}]$	
QE-DFT	-7.768	0.741	planar	-7.731	-0.771	planar	1.487	1.399	1.451
Ters-SET0	-7.336	2.296	B out of plane	-7.307	-2.568	B out of plane	1.506	1.487	1.450
Ters-SET1	-7.910	2.057	B out of plane	-7.883	-3.479	B and N shift	1.483	1.494	1.448
Ters-SET2	-7.830	4.171	planar	-7.874	-0.252	planar	1.515	1.431	1.432
hybrid	-7.830	4.171	planar	-7.944	-5.163	planar	1.519	1.418	1.379
(VASP:DFT-GGA)			planar ^(b)			1.49 ^(d)	1.40 ^(d)	1.44 ^(d)	
(SIESTA-GGA)			1.50 ^(c)			1.47 ^(c)	1.38 ^(c)	1.44 ^(c)	
GL: Nc					GL: Vc				
	$E_{coh} [\frac{eV}{atom}]$	$E_f^0 [eV]$	$d_{NC} [\text{\AA}]$	$E_{coh} [\frac{eV}{atom}]$	$E_f^0 [eV]$	$d_1 [\text{\AA}]$	$d_2 [\text{\AA}]$	$d_3 [\text{\AA}]$	
QE-DFT	-7.741	0.936	planar	-7.599	7.904	slightly corrugated	2.329, 2x2.594	1.394, 1.432, 1.420, 1.426, 1.397, 1.427	
Ters-SET0	-7.299	-2.777	planar	-7.171	6.886	planar	2.620	1.446	1.459
Ters-SET1	-7.900	-3.908	planar	-7.961	0.505	planar	3.061	1.356	1.415
Ters-SET2	-7.983	-3.188	planar	-7.961	0.505	planar	3.061	1.356	1.415
hybrid	-7.983	-3.188	planar	-7.961	0.505	planar	3.061	1.356	1.415
COMB	-7.039	7.563	planar	-7.092	8.046	planar	2.786	1.416	1.412
(VASP-GGA)			planar ^(b)						
(SIESTA-GGA)			1.398 ^(a) , 1.42 ^(h)		7.7 ^(e)		2.02 (f-T) ^(e)		
(review-DFT) ^(f)			1.42 ^(c)		7.6-7.9 (7.3-858)		1.8-2.37		
(OpenMX-GGA) ^(g)					7.98	corrugated	2.09, 2x2.58		
(exp)					7.0 ± 0.5 ^(h)		1.9 ^(f)		

(a) Rani *et al.*, 2013, (b) M. Wu *et al.*, 2010, (c) Al-Aqtash *et al.*, 2012, (d) Rani *et al.*, 2014, (e) Ma *et al.*, 2004, (f) Skowron *et al.*, 2015, (g) Rodrigo *et al.*, 2016, (h) Thrower *et al.*, 1978.

TABLE 4.9: Cohesive energies per atom, formation energies for N-rich and N-poor conditions, and values of characteristic interatomic distances shown in Fig. 4.3 for various defects in hBN. The results obtained with tested here VFF models are compared to the DFT calculations in 4x4 (32 atoms) superlattices employing Quantum Espresso numerical package (**QE-DFT**). The results of DFT calculations available in literature are also presented.

	hBN: C _N				hBN: V _N			
	$E_{coh}[\frac{eV}{atom}]$	$E_f[eV]$ (N-rich)	(N-poor)	$d_{CN}[\text{\AA}]$	$E_{coh}[\frac{eV}{atom}]$	$E_f[eV]$ (N-rich)	(N-poor)	
QE-DFT	-6.939	4.443	1.462	planar	1.509	7.736	4.756	planar
Ters-SET0	-7.497	7.951	2.275	corrugated	1.516	13.470	7.794	planar
Ters-SET1	-7.523	7.707	2.031	corrugated	1.493	13.470	7.794	planar
Ters-SET2	-7.388	6.834	4.189	planar	1.509	7.645	5.000	planar
hybrid	-6.667	3.751	0.083	v	1.509	8.573	4.905	planar
extep						8.573	4.905	planar
(SIESTA-GGA) ^(a)		4.8	1.88	1.52		7.68	5.17	2.06, 2x10
(bulk; CASTEP-LSDA) ^(b)		7.86*		1.49		9.85*		1.34
(PBE)		4.1 ^(d) , 4.3 ^(f)		1.509 ^(f)		7.7 ^(d) , 6.2 ^(f)	3.6 ^(f)	2.312, 2x2.311 ^(d)
(HSE)		4.3 ^(d) , 4.41 ^(f)		7.65 ^(d)		8.48 ^(d)		
(bulk; HSE)		4.3 ^(e)	1.4 ^(e)	1.50 ^(e)		8.4 ^(e)	6.0 ^(e)	2.28 ^(e)
	hBN: C _B				hBN: V _B			
	$E_{coh}[\frac{eV}{atom}]$	$E_f[eV]$ (N-rich)	(N-poor)	$d_{CN}[\text{\AA}]$	$E_{coh}[\frac{eV}{atom}]$	$E_f[eV]$ (N-rich)	(N-poor)	
QE-DFT	-6.905	1.893	4.873	planar	1.404	-6.687	10.665	slightly corrugated
Ters-SET0	-7.502	-4.138	1.538	corrugated	1.514	-7.437	3.677	corrug.
Ters-SET1	-7.526	-4.318	1.358	corrugated	1.496	-7.437	3.677	corrug.
Ters-SET2	-7.542	-3.194	-0.549	planar	1.429	-7.343	5.186	planar
hybrid	-6.709	-1.041	2.627	planar	1.420	-6.341	9.089	planar
extep						-6.341	12.756	planar
(SIESTA-GGA) ^(a)		1.41	3.84	1.41		7.05	10.08	2.64, 2.65, 2.7
(bulk; CASTEP-LSDA) ^(b)		5.59*		1.39		13.1*		1.41
(PBE)		1.8 ^(c) , 1.69 ^(d)		1.409 ^(d)		7.4 ^(d) , 7.5 ^(f)	10.2 ^(f)	2.783, 2.694, 2.695 ^(d)
(HSE)		1.95 ^(d) , 1.87 ^(f)				7.65 ^(d)		
(bulk; HSE)		1.5 ^(e)	4.8 ^(e)	1.40 ^(e)		8.4 ^(e)	>10 ^(e)	2.71, 2x2.257 ^(e)

(a) Azevedo *et al.*, 2007, (b) McDougall *et al.*, 2017, (c) Berseneva *et al.*, 2011, (d) B. Huang *et al.*, 2012,

(e) Weston *et al.*, 2018, (f) V. Wang *et al.*, 2012.

* the authors use zero-formation energy as defined in the text.

defect, and described in experimental papers (Robertson *et al.*, 2013). All Tersoff and COMB parametrizations return planar structures. However, **COMB** returns structure where C atoms with ‘dangling bonds’ move towards each other, which agrees with DFT. **Ters-SET0** is also acceptable, but **Ters-SET1** and **Ters-SET2** equal to **hybrid** here, shorten excessively C-C bonds in the neighborhood of the vacancy.

- **hBN: C_N.** **Hybrid** and **Ters-SET2** give excellent results of energy and geometry, but formation energies are different from **QE-DFT** and some results reported in literature. **Ters-SET0** and **Ters-SET1** - predict corrugated geometry (in contrast to **QE-DFT**) and worse values of cohesive energy.
- **hBN: C_B.** **Hybrid** and **Ters-SET2** result in quite good cohesive energy and geometry, but we observe formation energies to be negative in both cases. **Ters-SET0** and **Ters-SET1** again return corrugated structures (in contrast to **QE-DFT**), while values of cohesive energy are similar to **Ters-SET2**.
- **hBN: V_N.** All the methods result in planar structures of the vacancy surrounding. **Extep** and (identical) **hybrid** predict very well values of cohesive and formation energies. However, they overestimate noticeably the distance between boron atoms neighboring the vacancy (denoted d_1 in Tab. 4.9 and marked in Fig. 4.3). All the Tersoff parametrizations result in similar and correct cohesive energies, while d_1 is better captured with **Ters-SET0** and (identical) **Ters-SET1**, than in **Ters-SET2**.
- **hBN: V_B.** Values of cohesive energies are quite correct for all the discussed VFF methods. However, the geometry is not perfect, as Tersoff sets overestimate the distance between N atoms neighboring the vacancy, while **extep** and (identical) **hybrid** underestimate it. Equal in this case **Ters-SET0** and **Ters-SET1**, result in corrugated structures as **QE-DFT**. But neither **extep** nor **Ters-SET2** differentiate between the N-N bonds, which happens in the case of all reported DFT results.

4.5 Evaluation of VFF approaches in the view of CBN 2D alloys simulations

At this point, we evaluate and eliminate approaches that are definitely not very well suited for CBN 2D alloys. We discard **Reax-PAI** and **Reax-PAUP** parametrizations due to the problems with description of pristine graphene and hBN. **Ters-BNC** is

rejected due to non-optimized parameters for B-B and N-N bonds. **Extep**, **COMB** and **Reax-2013** do not provide complete parameters set for CBN systems. However, they are interesting candidates, if the missing parametrizations were provided, especially **COMB**. Among the other VFFs, the overall performance of **Ters-SET0** and **Ters-SET1** is not sufficient, or simply wrong in multiple cases discussed in this chapter. Thus, we have two candidates remaining: **Ters-SET2** and **hybrid** with the best overall performance. **Hybrid** gives better results in hBN systems (as extended Tersoff potential which is responsible for B-N interactions there, it was fitted specifically to hBN structures). However, it describes interactions of certain pairs of atoms by two different potentials. We are not sure, if such a mixture would properly picture general CBN systems. Eventually, we choose **Tersoff potential with Ters-SET2 parametrization** for further studies. Visual summary of assessment of VFF methods is presented in Tab. 4.10.

TABLE 4.10: General performance of all the investigated empirical potentials in the tests presented in Sec. 4.4. "inc." denotes incomplete parametrization lacking some of C-B-N interactions, "rej." means parametrization rejected during the first part of the tests. For each tested system and VFF method, we conclude its performance with: '++' that means *excellent*, '+' refers to *decent*, '+/-' marks the result partly correct and partly wrong, '-' is *inaccurate*, while '- -' is a *very bad result*. We sum pluses and minuses to get approximate general evaluation of all the VFFs.

Test no.	system tested	Ters SET0	Ters SET1	Ters SET2	Ters BNC	extep	hybrid	COMB	Reax C2013	Reax PAI	Reax PAUP
(1)	C-C	+	+/-	+/-	+/-		+/-	++	+/-	-	-
(1)	N-N	--	--	++	--	++	++	++		+	+
(1)	B-B	+	+	+	--	+	+			+/-	-
(1)	C-N	-	-	+	+		+	++		-	-
(1)	C-B	+	-	+/-	+/-		+/-			-	+
(1)	B-N	+/-	+/-	+/-	+/-	+/-	+/-			--	--
(2)	GL	+/-	+	+			+	++	+/-	-	-
(2)	hBN	-	-	+		++	++			--	--
(3)	GL: B _C	+	+/-	++			++				
(3)	GL: N _C	-	-	+			+	-			
(3)	GL: B _C N _C	-	-	++			+				
(3)	GL: V _C	+	-	-			-	+			
(3)	hBN: C _N	-	-	+			++				
(3)	hBN: C _B	-	-	+			++				
(3)	hBN: V _N	+	+	-		+	+				
(3)	hBN: V _B	+/-	+/-	+/-		+/-	+/-				
summary		6+ 8-	3+ 9-	13+ 2-	rej.	inc.	16+ 1-	inc.	inc.	rej.	rej.

(4) Performance of Tersoff potential in small CBN systems

As a final validation of the chosen potential, we compare **Ters-SET2** and **QE-DFT** energies (calculated identically as systems described in Sec. 3.2 and Sec. 4.4 for point defects) for a number of small GL and hBN structures (32 atoms in supercell) with 1-4 substituted atoms. They are listed and analyzed below:

- (4-i) 44 clusters in graphene with nitrogen substitutions, with energies calculated on the rigid lattice, and similar set with boron substitutions. This set is used in Chapter 3 as a Clusters Pool, and it is visualized in Fig. A.1. We see in Fig. 4.4 (upper panel) that **Ters-SET2** correctly favors N-doped GL structures with N atoms separated, compared to GL with N-N pair. However, value of this difference is overestimated ≈ 3 -4 times. Moreover, total cohesive energy decreases with increasing number of N atoms, in contrast to DFT results. This would impact our considerations of, *e.g.*, phase transitions, when we compare different concentrations of N in GL. At the same time, lower panel in Fig. 4.4 shows improper behavior of **Ters-SET2** in B-alloyed graphene, as it assigns lower energy to GL with B-B bond than for the systems with B atoms separated. It is probably due to overestimated Tersoff energy in B-B dimer interaction as shown in Fig. 4.1 (B-B). This peculiarity needs to be fixed, before performing large-scale simulations. Even so, the tendency to increase the total energy with increasing concentration of B in GL is consistent with DFT results. Finally, note that Tersoff does not distinguish energies of configurations where N-dopants are second or further neighbors (due to cutoff function present in its definition).
- (4-ii) 20 clusters in graphene with nitrogen substitutions, with energies calculated in relaxed lattice, and similar set with boron substitutions. This set is introduced in Chapter 3 as a Initial Training Set (**REFSET1**), and displayed in Fig. 3.1. Fig. 4.5 collates Tersoff and DFT energies. Relaxation reduces differences between 'clustered' and 'separated' configurations in both approaches, but trends are much the same as in (4-i). However, geometry optimization leads to *e.g.*, reverse relation of structures 11 - 12 in GL-N case. More complex effects play the role in DFT, that Tersoff is unable to capture.
- (4-iii) For tests of ternary systems, we choose 52 graphene structures doped with B and N simultaneously (with energies calculated in rigid lattice) and 28 hBN systems doped with C (with energies calculated both in rigid and relaxed lattice). Exemplary subset of this set (as it would be challenging to show all the data in legible

way) is shown in Fig. 4.6, and respective energies are plotted in Fig. 4.7. In ternary alloys, many predictions of **Ters-SET2** are correct, but issues raised in the case of binary systems, deteriorate its performance here as well.

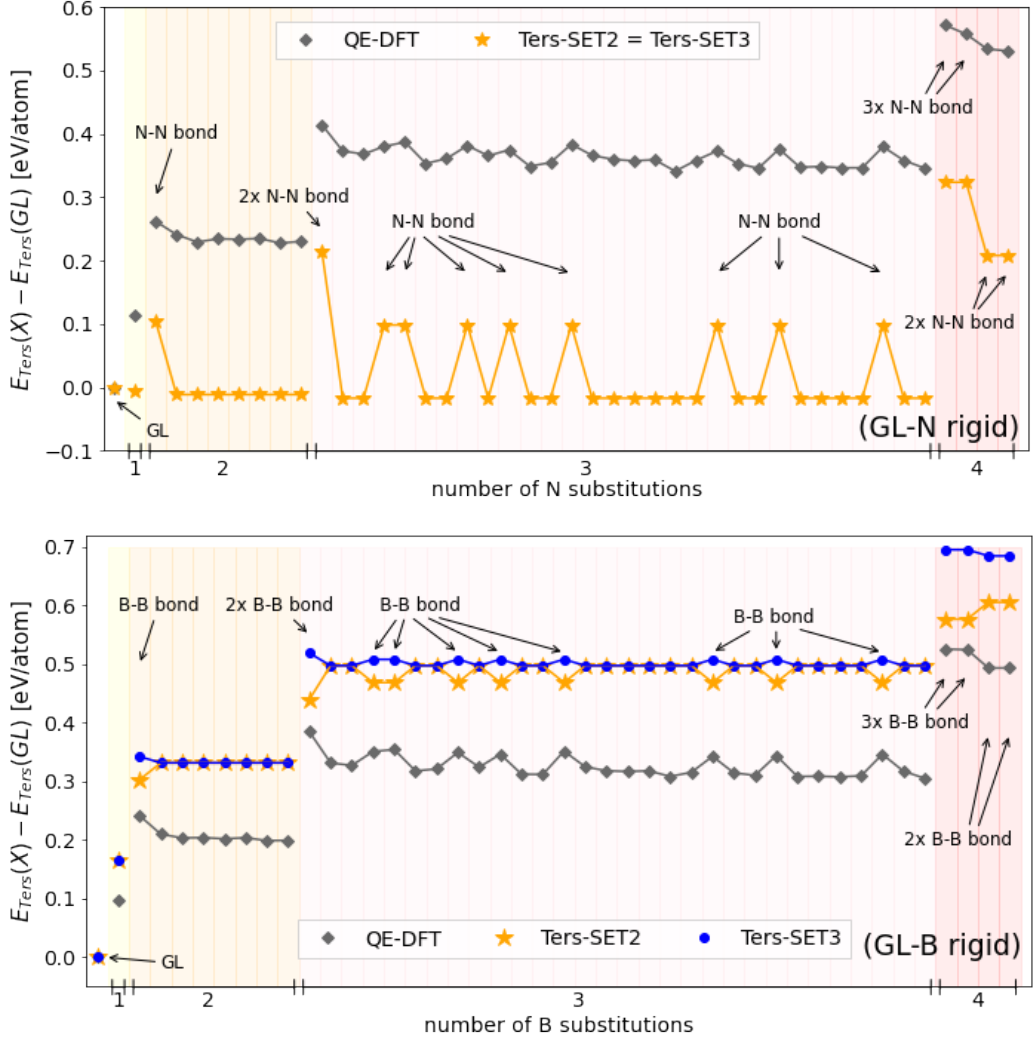


FIGURE 4.4: Comparison of **QE-DFT**, **Ters-SET2**, and **Ters-SET3** energies of GL-N (upper panel) and GL-B structures (lower panel) (calculated on the rigid lattice, relative to pristine GL energy), constituting test set (4-i), *i.e.*, Clusters Pool introduced in Sec. 3.1. Configurations are plotted in Fig. A.1. Number of substituted atoms for all configurations is marked at the bottom of the figures, moreover we point out structures with N or B atoms arranged as first neighbors. For GL-N structures **Ters-SET2** and **Ters-SET3** are identical.

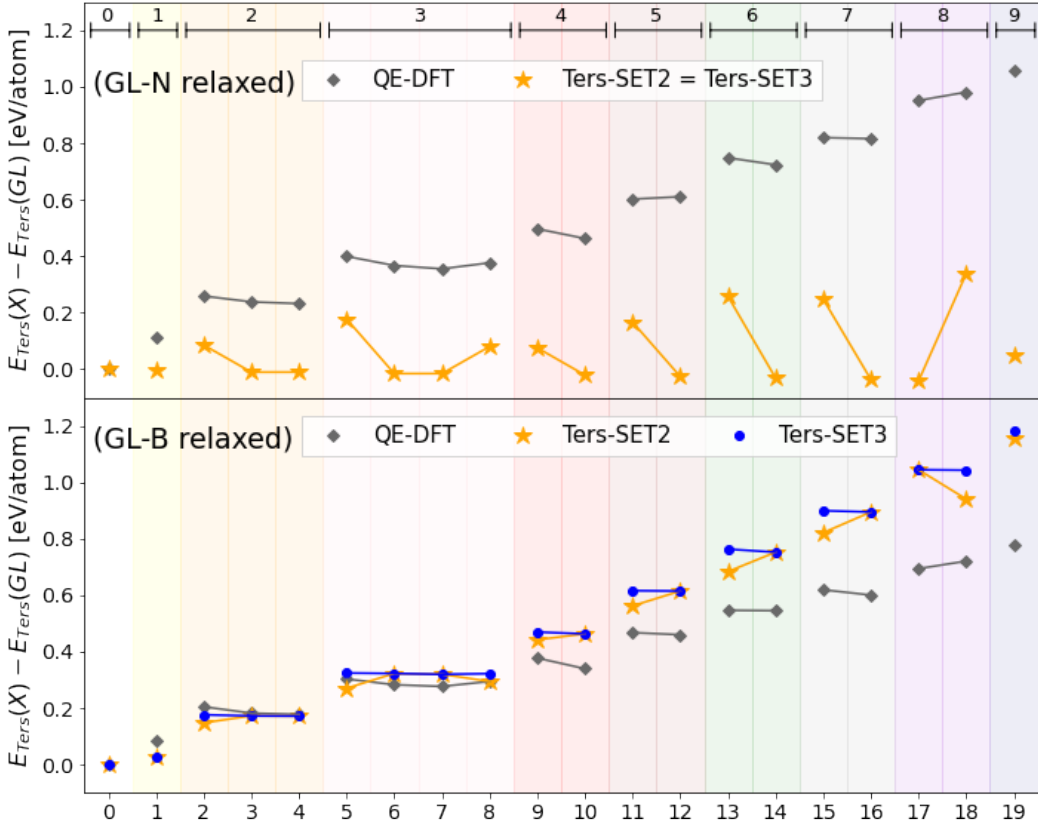


FIGURE 4.5: Comparison of **QE-DFT**, **Ters-SET2**, and **Ters-SET3** energies of GL-N (upper panel) and GL-B structures (lower panel) (calculated on the relaxed lattice, relative to pristine GL energy), constituting test set (4-ii) *i.e.*, Initial training set introduced in Sec. 3.1. Atomic configurations indicated as 0-19 are depicted in Fig. 3.1. Number of dopants in the supercells for each configuration is shown in the upper part of the figure. For GL-N structures **Ters-SET2** and **Ters-SET3** are identical.

Therefore, we have overtaken some efforts to improve **Ters-SET2** parametrization. First, we refit description of B-B interaction, to obtain appropriate relation of B-B and B-C energies. After several attempts, we find that varying single parameter $B_{BB} = 1173.197 \rightarrow 1149.00$ results in qualitatively correct description of B-B pairs within graphene. Fig. 4.8 shows comparison of calculations for B-B molecule as obtained with **QE-DFT**, **Ters-SET2** and **Ters-SET3** approaches.

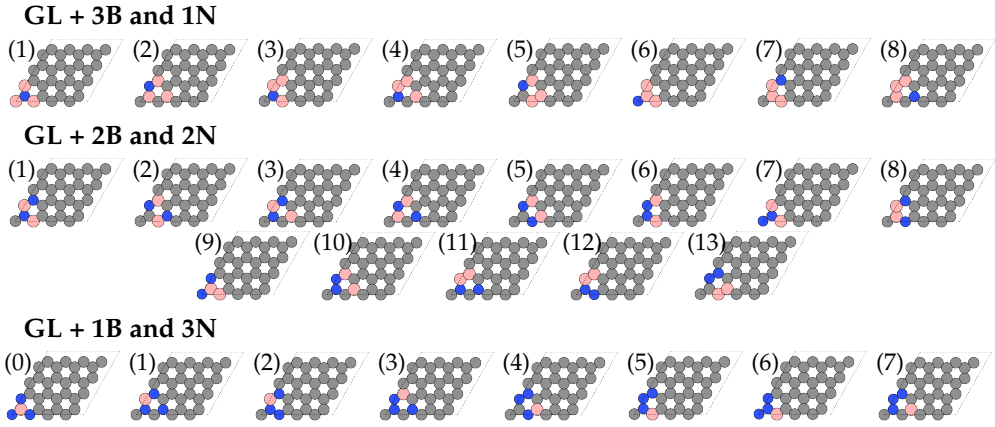


FIGURE 4.6: Subset of set of ternary systems tested in (4-iii), consisting of graphene with four introduced B and N substitutions. Configurations are categorized by number of B and N and structures with the same chemical formula are ordered from lowest to highest DFT energy on a relaxed lattice.

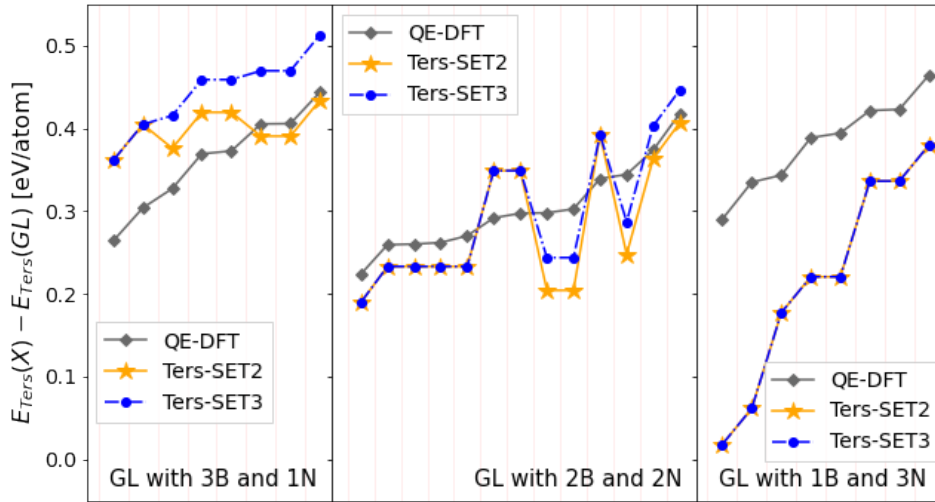
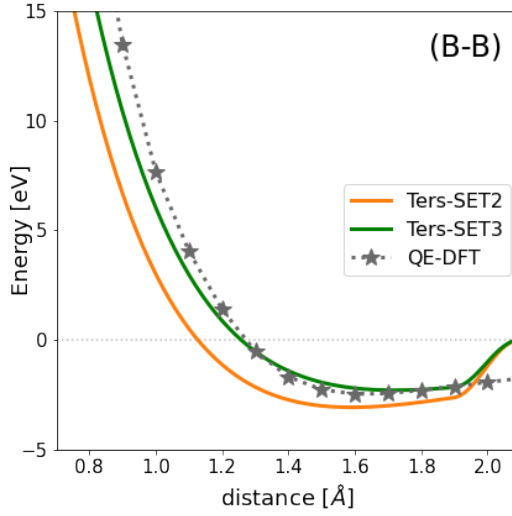


FIGURE 4.7: Comparison of QE-DFT, Ters-SET2 and Ters-SET3 energies of GL-BN structures (calculated on relaxed lattice, relative to pristine GL energy) constituting subset of test set (4-iii), as depicted in Fig. 4.6, where 4 substitutions are introduced into graphene lattice. The number of N and B dopants in the supercells for various configurations is marked in the lower part of the figure. The structures with the same chemical formula are sorted from the lowest to highest DFT energies.

FIGURE 4.8: Comparison of **QE-DFT**, **Ters-SET2**, and **Ters-SET3** energies as a function of interatomic distance for B-B molecule.



Minimum of BB dimer binding energy obtained with **Ters-SET3** amounts -2.3 eV (*versus* -3.08 eV for **Ters-SET2**) is for distance equal 1.72 Å (*versus* 1.59 Å). However, the shape is much closer to the desired DFT results. It is important to mention that change of B_{BB} parameter does not influence any other result from first part of the tests. It only modifies the energy when two boron atoms appear in graphene lattice as first neighbors.

What is more, we try to refit parameters concerning overestimated strength of C-N pair, but we have failed in completing this task. In conclusion, for the given concentration x of N atoms in graphene (the structures $GL_{1-x}N_x$), predictions of the lowest energy structures should be reliable. However, we cannot depend on Tersoff potential in C-N systems in the subject of calculating mixing energy, and free energy, that is aimed in Section 5.1.2. Henceforth, we refer to the parametrization with altered B_{BB} parameter as **Ters-SET3**.

Fig. 4.4 shows comparison of **Ters-SET2** and improved **Ters-SET3** in application to rigid N-doped and B-doped graphene systems. The difference is visible only for the B-graphene, where correct tendency to increase the energy when B-B bond is present in the system is achieved with **Ters-SET3**. One can notice that energy differences between B-B forming a bond and structures with separated dopants is

smaller in **Ters-SET3**, when compared to **QE-DFT**. This would influence quantitative description of properties in these structures.

Fig. 4.5 shows analogous comparison in structures with relaxed geometry. It confirms previous conclusions, that **Ters-SET3** improves performance of Tersoff potential in GL-B systems. Finally, we show exemplary result for tests in ternary structures. Fig. 4.7 compares Tersoff and DFT results in graphene doped with four foreign atoms, that might be realized in three ways with B and N dopants. Here again, we notice that **Ters-SET3** improves quality of the Tersoff energies, but due to the problems with description of C-N interaction, there are also wrong predictions here.

In the view of all the results presented in this chapter we decide to use Tersoff potential with **Ters-SET3** parametrization. We are aware of drawbacks and areas that Tersoff results do not return correct values.

4.6 Conclusions: empirical potentials for simulations in solid state physics

Let's summarize shortly the main features of Valence Force Fields for materials' simulations.

- VFFs portray the energy as a function of atomic positions (and charge in some cases), and in the family of bond-order potentials, interaction between two atoms is influenced also by their surroundings. In many formulas, one can recognize known physical interactions (*e.g.*, attractive, repulsive parts, van der Waals, Coulomb interactions), but it might be arbitrarily chosen function as well, that in authors' opinion would give the best accuracy.
- Empirical potentials allow to perform simulations of very large structures, with any periodicity and shape, also with defects, to optimize geometry, run Molecular Dynamics for time evolution of the system, and much more. Despite increasing computational power available in computer science, simulations of systems corresponding to experimentally obtained structures are still hardly reachable for *ab initio* methods.
- Parameters in VFFs are fitted to a chosen set of structures and their particular properties, and (unfortunately) always some effects would be skipped. The more

complex formula of the force field, the more computational effort is needed to perform simulations, but the results are typically more accurate.

- Parametrizing a given VFF is a complex task. It can take several months and typically whole teams are involved to generate well-fitted set of parameters. With increasing demands towards VFF potentials and resulting their complication and number of parameters introduced in VFF formula, difficulty is also increasing. There are, however, new procedures being developed, that employ machine learning techniques to find the best set of parameters for a given reference data (Sengul *et al.*, 2021). Taking into account the importance of the VFF methods for the material science, the new machine learning based fitting procedure definitely opens new perspectives for very accurate studies of materials properties and definitely should be followed in the next future.
- Recently machine learning branch has offered also new approach to determine parameters 'on the fly', fitting them for any given configuration repeatedly (Kolodzeiski *et al.*, 2021). This might be a great chance to develop new generation of Valence Force Fields, exhibiting accuracy comparable to *ab initio* methods.

In the following chapter, we employ Tersoff potential with parametrization **Ters-SET3** and perform simulations in large supercells.

Chapter 5

Tersoff potential results on ordering in large scale CBN systems

This chapter presents the results of large-scale simulations of CBN hexagonal monolayer systems performed in the Valence Force Field (VFF) framework. The VFF "engine" here is Tersoff potential with **Ters-SET3** parametrization as presented in Chapter 4. This approach allows in principle for atomistic calculations for very extended systems. Section 5.1.1 introduces the results of Metropolis Monte Carlo simulations of CBN alloys on rigid lattice, where one focuses on the chemical order effects, and determines distribution of atoms over the lattice points (configuration) that leads to the minimal total energy of alloy. Analysis of the simulations includes presentation of the lowest energy configurations, discussion of stability, ordering, and configurational entropy. From this, we draw conclusions about the applicability of Tersoff potential to calculations of free energy and phase transitions from ordered to random CBN alloys. In Sec. 5.2 we compare Tersoff and Cluster Expansion results (the latter presented in Sec. 3.4) and ability of both approaches to simulate monolayer CBN systems. Finally we present Density of States calculated with Tight Binding method (Sec. 5.4), and analyze electronic properties of doped CBN alloys.

The last part of this chapter is devoted to the Tersoff potential simulations that include geometry optimization. Optimized graphene and hBN flakes are presented in Sec. 5.5. This is followed by results of Monte Carlo simulations in B-doped graphene flakes.

5.1 Monte Carlo simulations of 2D CBN systems

We divide the study of lateral C-B-N alloys into four *classes* of topics: (i) graphene doped with nitrogen (**GL-N**), (ii) graphene doped with boron (**GL-B**), (iii) graphene doped with boron and nitrogen simultaneously (**GL-BN**), and (iv) hexagonal boron nitride doped with carbon (**hBN-C**). The parameters of Monte Carlo simulations performed in all aforementioned classes of systems are summarized in Tab. 5.1.

TABLE 5.1: Parameters characterizing the Monte Carlo simulations for CBN systems on rigid hexagonal lattice. c_x denotes range of concentrations of substituted atoms x , *cell* describes supercell size used in the simulations, *cell dimensions* refers to spacial dimensions of the supercell, MCS is number of Monte Carlo Steps, and N_{stat} - number of instances for averaging.

	GL-N	GL-B	GL-BN	hBN-C
c_x	$c_N \in [0.01, 0.50]$	$c_B \in [0.01, 0.50]$	$c_N = c_B \in [0.01, 0.50]$	$c_C \in [0.01, 0.50]$
cell	15x30 (1800 atoms) supercell, with periodic boundary conditions			
cell dimensions	64.75x74.77 Å			64.89x 74.93 Å
MCS	1 000 000			
N_{stat}	10			

For any class and concentration of substituting atoms, we prepare N_{stat} initial random distributions. This allows for gathering sufficient statistics and suitable averaging results of the investigated systems' properties. The Monte Carlo algorithm is performed for all the samples, *i.e.*, MCS swaps of a random pair of atoms (of different species) are attempted, with all the swaps that lower total energy of the system being accepted. For the swaps causing increase of the energy, acceptance rate equals $p_i = \exp(\frac{-\Delta E_{Ters}}{k_B T})$, where ΔE_{Ters} denotes difference in total Tersoff energy of the system before and after the swap, T is temperature (chosen here to be 10K), see discussion of role of temperature in Monte Carlo simulations, and k_B - Boltzmann constant. This procedure searches configuration space for atomic arrangements resulting in the lowest energy. We have tested that the system size is sufficient to neglect effects of periodic boundary conditions, mainly due to the short-range interactions included in Tersoff potential. We have also performed numerous MC test runs with T values: 150, 300, 500, 800, 1500 and even 3000 K, just to find out that in all cases the energetically most favorable configuration is independent on the T chosen employing Metropolis algorithm. Typically the simulation have continued

for few millions of MC steps, however, generally the convergence is reached in 10^6 MC steps.

5.1.1 Results of Monte Carlo simulations for C-B-N systems on rigid hexagonal lattice.

Initial (random) and equilibrium distribution of atoms (after Monte Carlo) are depicted in Figs. 5.1-5.4. It is easy to notice that in every class, final atomic distributions deviate considerably from the random ones. Specifically, in (i) N-alloyed graphene: N atoms avoid creating N-N bonds within the lattice, even for the highest concentrations; (ii) B-alloyed graphene is a twin case, here B atoms avoid each other for all the concentrations; (iii) in BN-alloyed graphene we observe very strong drive of B and N atoms to form hBN domains within graphene lattice; (iv) in C-alloyed hexagonal boron nitride, similarly to (iii), we observe formation of C areas surrounded by hBN ones. Note, that here we do not start from completely random arrangement of species, but in ordered hBN lattice, we substitute certain number (equal) of B and N in place of C. To confirm these visual observations, we calculate mean energies of the structures before and after Monte Carlo simulations (exhibiting short range ordering). The average is calculated over instances of each concentration, but also, in every instance, we gather energies after certain number of MC steps when its value fluctuates around some constant value (*i.e.*, after *thermalization*). The values of average energy (per atom) for random and MC-optimized structures are plotted in Fig. 5.5. Note, that the more substituted atoms, the more system can 'gain' energetically *via* proper arrangement. Additionally, we plot in Fig. 5.5 (ii)-(iv) blue dashed line, which shows weighted average of systems that each class is built of: graphene and nitrogen (N_2) in (i), graphene and α -borophene (B_α) in (ii), while graphene and hexagonal boron nitride for (iii) and (iv). From this, we can estimate that mixing energy in classes (ii)-(iv) is greater than zero for all concentrations, as

$$\begin{aligned}
 E_{mix}(GL_{1-x}B_x) &= E(GL_{1-x}B_x) - (1-x)E(GL) - xE(B_\alpha) \\
 E_{mix}(GL_{1-x}N_x) &= E(GL_{1-x}N_x) - (1-x)E(GL) - xE(N_2) \\
 E_{mix}(GL_{1-x}(BN)_x) &= E(GL_{1-x}(BN)_x) - (1-x)E(GL) - xE(hBN)
 \end{aligned} \tag{5.1}$$

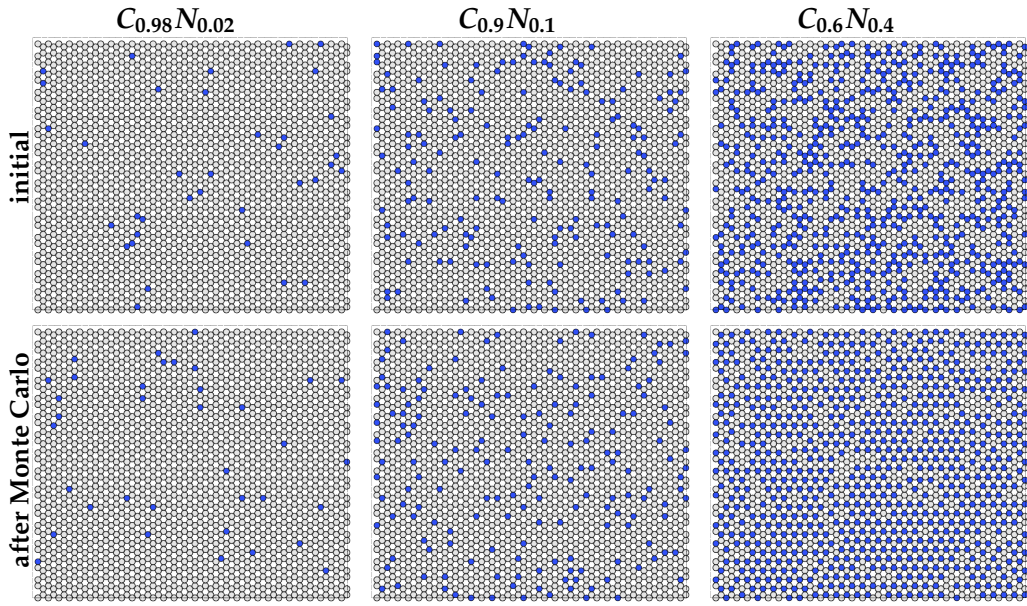


FIGURE 5.1: Evolution of nitrogen dopants (blue dots) distribution over the rigid honeycomb lattice after performing Monte Carlo simulations for exemplary concentrations of N being equal to 2%, 10%, and 40%. Initial distributions of N atoms were assumed to be a random ones.

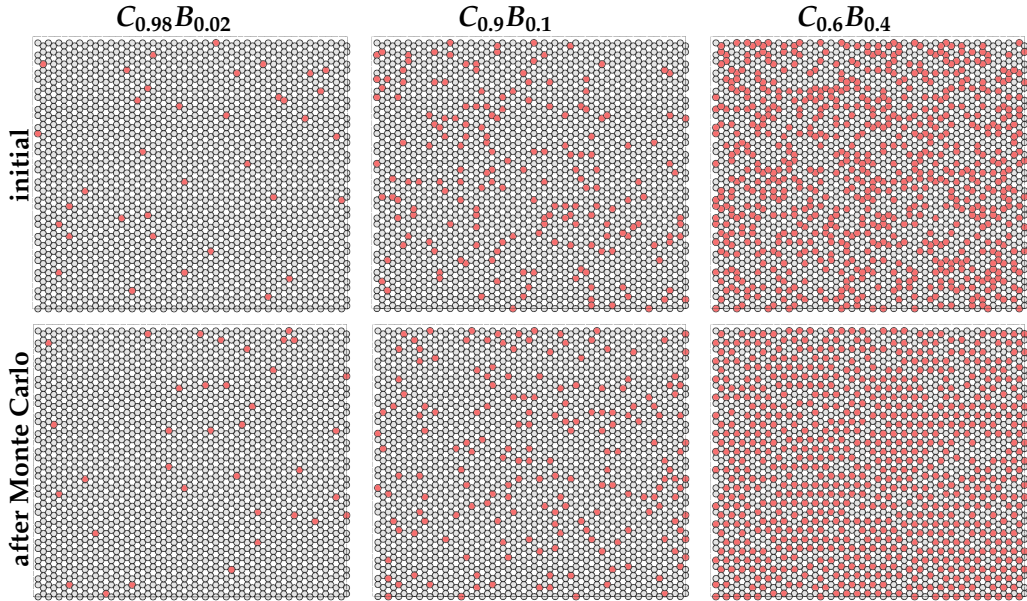


FIGURE 5.2: Evolution of boron dopants (red dots) distribution over the rigid honeycomb lattice after performing Monte Carlo simulations for exemplary concentrations of B being equal to 2%, 10%, and 40%. Initial distributions of B atoms were assumed to be a random ones.

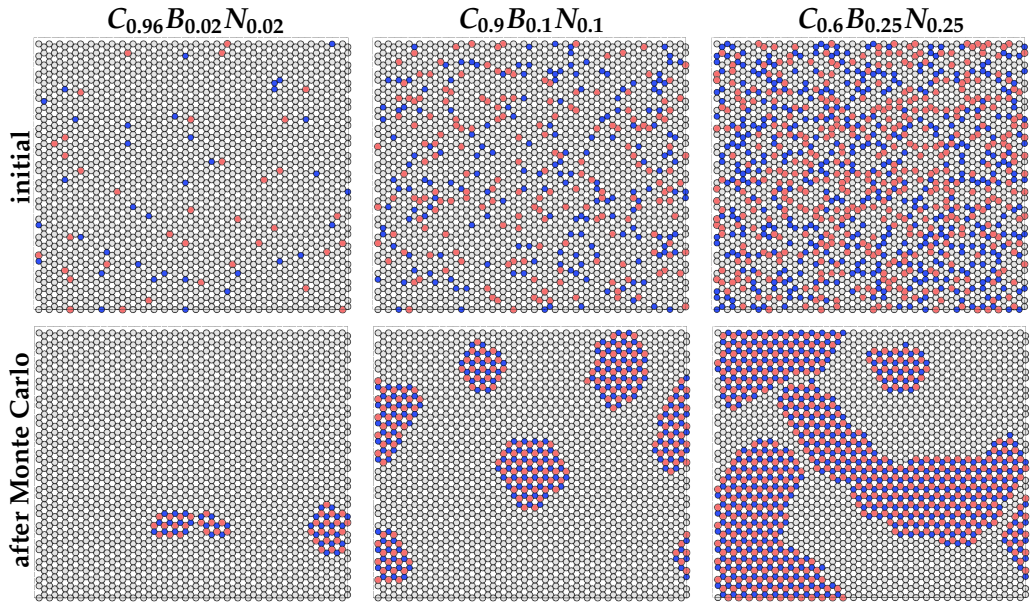


FIGURE 5.3: Evolution of boron (red dots) and nitrogen (blue dots) dopants distribution over the rigid honeycomb lattice after performing Monte Carlo simulations for exemplary concentrations of B and N ($c_B = c_N$), 2%, 10%, and 25%. Initial distributions of B and N atoms were assumed to be a random ones.

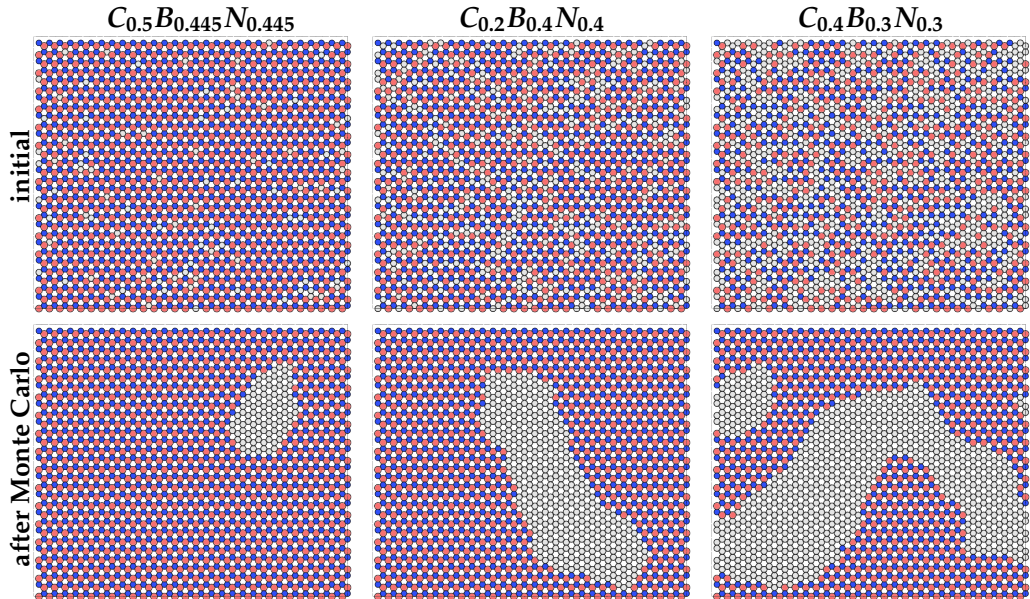


FIGURE 5.4: Evolution of carbon (gray dots) dopants distribution over the rigid hexagonal boron nitride honeycomb lattice (red dots are boron, blue - nitrogen atoms) after performing Monte Carlo simulations for exemplary concentrations of C, 5%, 20%, and 40%. Initial distribution of C over the hBN lattice was assumed to be a random one with equal number of substitutions in each of the sublattices.

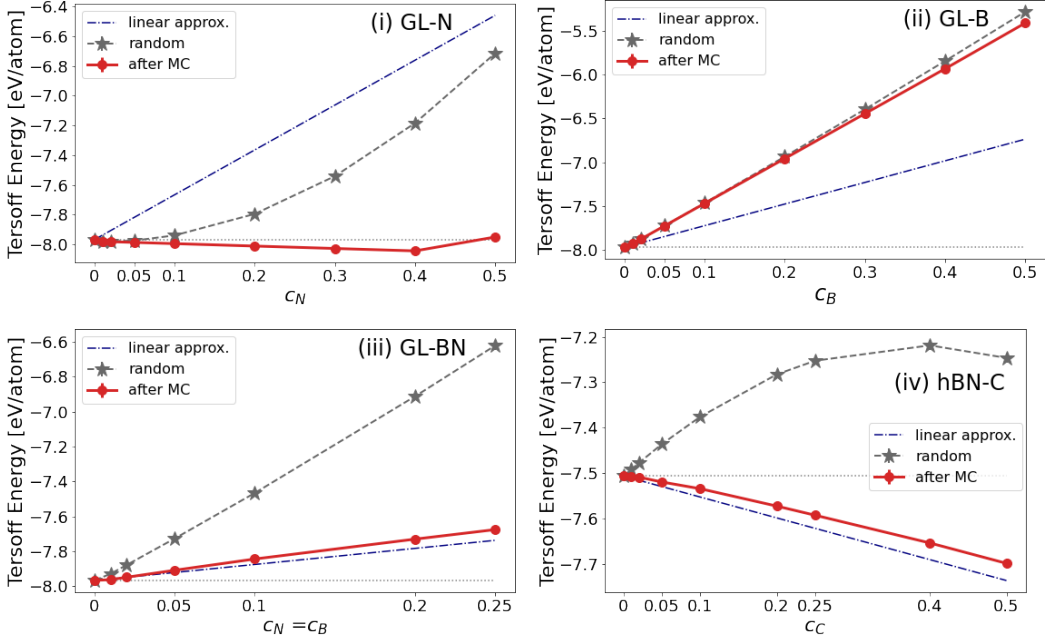
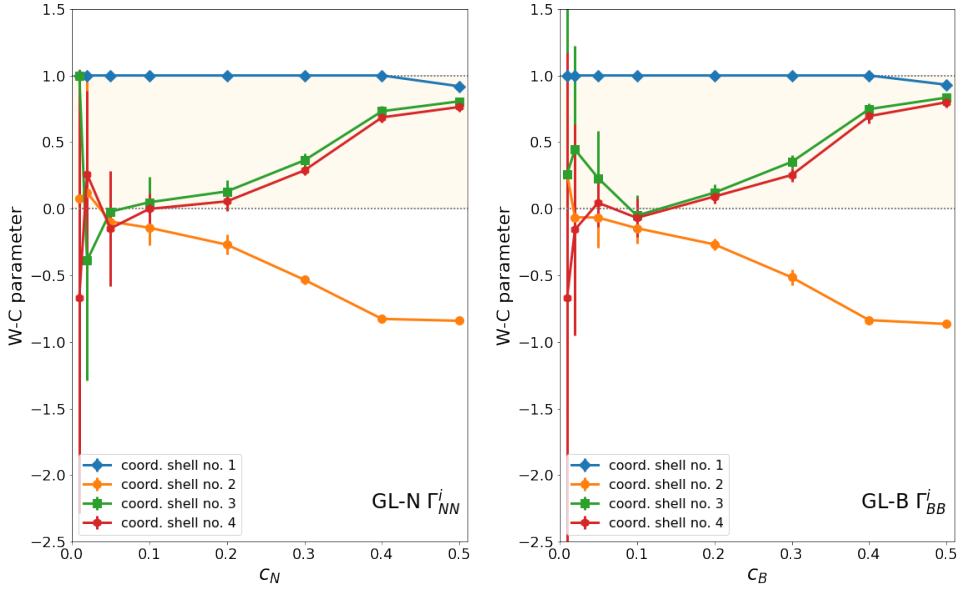


FIGURE 5.5: Average energy for all investigated classes of CBN alloys of various concentrations: for random initial distributions, and distributions averaged during MC run, as well as over N_{stat} instances. Horizontal lines mark energies of un-doped system, while *linear approx.* means weighed energy of (i) graphene and nitrogen molecule, (ii) graphene and borophene, (iii) and (iv) of graphene and hBN, according to respective concentrations of species.

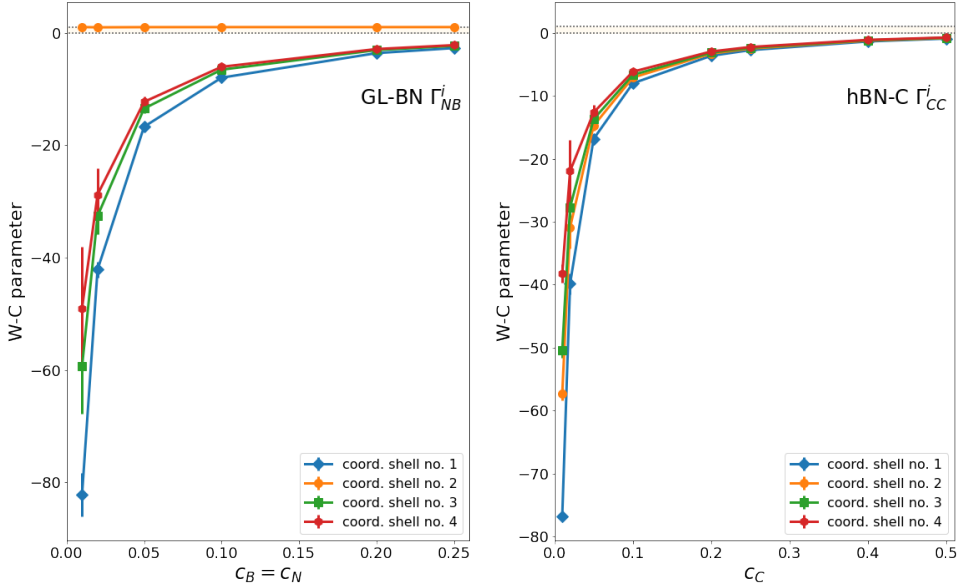
For GL-N mixing energy is negative, due to the problems with C-N parametrization as discussed in Sec. 4.5.

Finally, we calculate Warren-Cowley (W-C) Short-Range Order parameters for initial and final distributions. The Warren-Cowley approach is described in Sec. 2.4.2. Let's remind here only the formula $\Gamma_{IJ}^{(i)} = 1 - P_{IJ}^{(i)} / (c_I c_J)$, with $P_{IJ}^{(i)}$ referring to probability of finding I and J type of atoms on i^{th} coordination shell in a given system, and c_I, c_J - concentrations of I, J species respectively. The average values of characteristic Warren-Cowley parameters for each class are depicted in Fig. 5.6. Note that $\Gamma_{IJ}^{(i)}$ is zero for ideal random distribution of I and J atoms. Positive values correspond to the situation when I and J atoms avoid each other, with maximum value equal to 1, when there are no I-J bonds in the system. W-C parameter smaller than zero indicates that pairing of I and J types (on i^{th} coordination shell) is preferable.



(i) Warren-Cowley parameters for N-N pairs in GL-N for N concentration c_N in the range 0.01-0.5.

(ii) Warren-Cowley parameters for B-B pairs in GL-B for B concentration c_B in the range 0.01-0.5.



(iii) Warren-Cowley parameters for B-N pairs in GL-BN for equal N and B concentrations $c_N = c_B$ in the range 0.01-0.25.

(iv) Warren-Cowley parameters for C-C pairs in hBN-C for C concentration c_C in the range 0.01-0.5.

FIGURE 5.6: Characteristic Warren-Cowley $\Gamma_{IJ}^{(i)}$ parameters for the 1st – 4th coordination shells of atoms I and J in four classes of investigated CBN alloys. The values of $\Gamma_{IJ}^{(i)}$ are averaged over N_{stat} Monte Carlo runs for the studied cases.

Analysis of Warren-Cowley parameters for CBN alloys confirms our previous (visual) impression:

- (i) $\Gamma_{NN}^{(1)} > 0$ for all the concentrations. Moreover, only for $c_N = 0.5$ value of $\Gamma_{NN}^{(1)}$ is noticeably smaller than 1. This indicates very strong antipreference of N atoms to form the (nearest) bonds within graphene lattice. One can observe a related tendency of $\Gamma_{NN}^{(2)}$ to decrease with increasing N concentration, due to the 'lack of space', so N atoms tend to be situated as the second nearest neighbors.
- (ii) $\Gamma_{BB}^{(1)} > 0$ for all the concentrations. Situation is similar to the one described above.
- (iii) $\Gamma_{BN}^{(1)} < 0$ for all the concentrations, B and N atoms prefer to occupy sites next to each other. Large absolute value of $\Gamma^{(1)}$ (especially for small concentrations of B and N) indicate substantial deviation from random distribution, towards segregation into hBN and graphene domains. $\Gamma_{BN}^{(2)}$ is very close to 1, so we don't find B and N atoms as second neighbors. Here also $\Gamma_{CC}^{(1)} \ll 0$ (not plotted here) indicating presence of much more C-C bonds than it would result from random distributions.
- (iv) Situation is similar to GL-BN class. We observe $\Gamma_{CC}^{(i)} < 0, i = 1, 2, 3, 4$ which proves the presence of domains of pristine graphene within the lattice. We find out that $\Gamma_{BN}^{(1)} < 0$ (not presented here), similarly to the GL-BN class for all concentrations of carbon substitutions investigated.

We have calculated Warren-Cowley parameters for all pairs of atomic types, but we choose to present the data as above, since it shows clearly the preferences in the CBN alloys. Moreover, in binary alloys, when we know the value of $\Gamma_{AB}^{(i)}$, the remaining values (for a given coordination shell) are mutually dependent due to the symmetry and sum rules (described in Sec. 2.4.2). In our opinion, ordering phenomena in CBN alloys investigated with Tersoff potential on a rigid lattice result from basic preference (or antipreference) of species to create the bonds within hexagonal structure.

Equipped with the knowledge about energy and optimum configurations of CBN alloys (on rigid lattice), let's focus now on the configurational entropy and free energy of mixing for alloys configurations corresponding to the lowest total energy.

5.1.2 Entropy and Free Energy of Rigid CBN systems

One of the goals of the present thesis is to predict phase transitions in CBN alloys, and determination of the configurational entropy is necessary part of this task. Therefore, remembering that all investigated CBN systems exhibit short range order, we take it into account during entropy calculations. We utilize formula of configuration entropy in pair approximation (specified in Sec 2.4.4), applying value of coordination number in hexagonal lattice $Z=3$

$$S_{pair} = k_B \left(2 \sum_{i=C,B,N} c_i \ln(c_i) - \frac{3}{2} \sum_{i,j \in C,B,N} p_{ij} \ln(p_{ij}) \right).$$

Values (averaged over N_{stat} instances) of S_{pair} for configurations resulting from Monte Carlo simulations in classes (i)-(iv) are depicted in Fig. 5.7 and compared to the values for initial arrangements of species. Entropy resulting from Bragg-Williams point approximation $S_{BW} = -k_B \sum_i c_i \ln(c_i)$ is also plotted for comparison, and we observe its very good agreement with S_{pair} of randomly doped graphene systems (i)-(iii). It is not the case of (iv) hBN-C initial configurations, as they are generated via carbon substitution into hBN randomly, equally on both sublattices, but areas of ordered hBN are preserved anyway. As we know from theoretical considerations, value of S_{conf} is negative for ordered binary systems, it equals $-\frac{1}{2}k_B \ln(2)$, e.g., for pristine hBN or any other binary system. This feature of the formula Eq. (2.48) is clearly visible in Fig. 5.7. In GL-N and GL-B (Fig. 5.7 (i) and (ii) respectively) we observe rise of entropy value with increasing concentration. Then, around $c_i = 0.25$ ($i = N, B$), entropy starts decreasing, and for the highest concentration $c_i = 0.5$ S_{pair} is negative. It is important to notice that pair approximation was meant as correction to the point approximation for alloys deviating from random (Mohri, 2007). Moreover, all ternary structures resulting from MC simulations exhibit negative S_{pair} . Clearly, this formula is not suitable for systems exhibiting long-range ordering, e.g., when large areas of hBN are present, or when we are close to ordered binary alloys (as in GL-N and GL-B for concentrations of substituting atoms equal 50%). In ternary CBN alloys and for GL with high content of B/N atoms, we would need to apply another procedure to calculate entropy, to be able to account for ordering present in the systems correctly. Possibly, in those classes of systems, we need to apply integral formula of entropy defined in Eq. (2.45). However, that demands performing extensive Monte Carlo simulations, with large

number of temperatures, which is out of scope of this thesis.

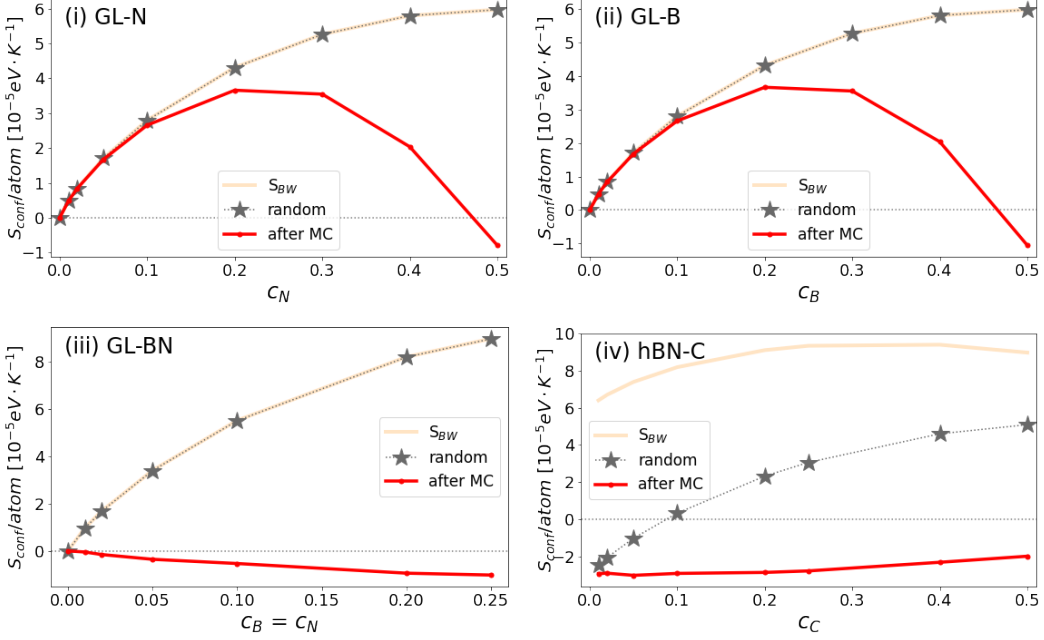


FIGURE 5.7: Average configurational entropy S_{pair} in pair approx. for the (i) GL-N, (ii) GL-B, (iii) and (iv) hBN-C structures before and after Monte Carlo Metropolis simulations with Tersoff potential. Values of Bragg-Williams entropy S_{BW} (point approx.) is also plotted for comparison.

That being so, we have decided to determine free energy of mixing $F_{mix} = E_{mix} - TS_{conf}$ only for binary structures investigated here, *i.e.*, GL-N and GL-B. We combine then Tersoff mixing energy from Eq. (5.1) and entropy as defined in Eq. (2.48) and plot the free energy of mixing as a function of dopants concentration and for a range of temperatures in Fig. 5.8. It was discussed in Sec. 4.5 and in the previous part of this chapter, that Tersoff parametrization gives wrong prediction in GL-N systems, when we compare structures with different nitrogen concentrations. Clearly, this obstacle does not allow for obtaining valid picture of the free energy, this would demand reparametrizing Tersoff potential. In GL-B (ii), we observe that free energy of mixing is positive for temperatures up to 6000K, whereas, for larger values, there is an area of negative values, that would suggest possibility of intermixing C and B. Regardless of that, it is hardly possible to obtain crystal 2D structure in such a temperature, thus we conclude that it is not possible to dissolve boron in graphene lattice under equilibrium conditions.

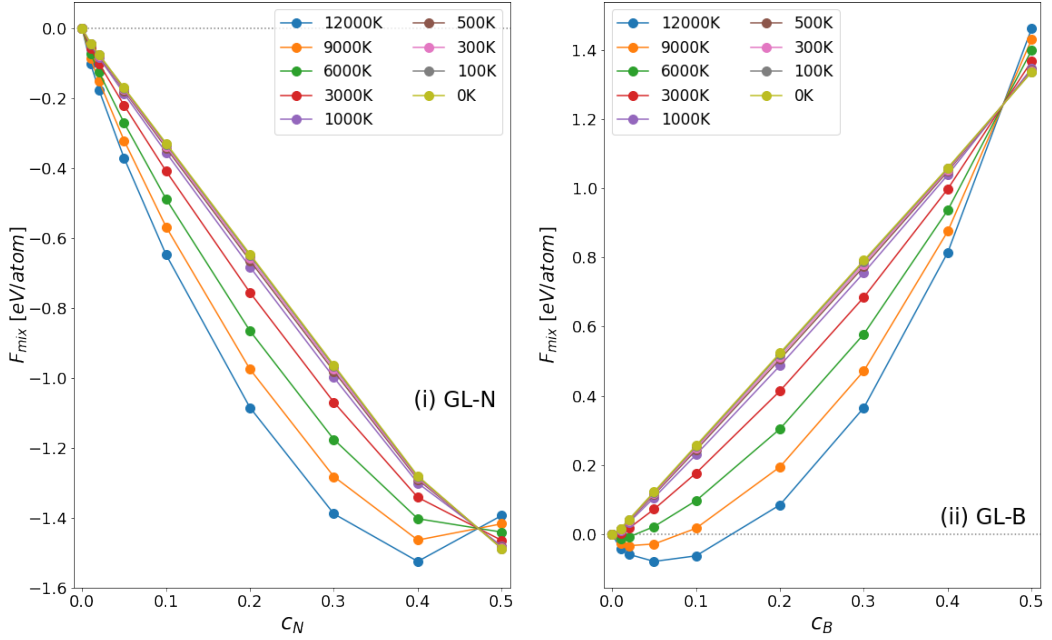


FIGURE 5.8: Average free energy F_{mix} for the lowest energy structures obtained in Monte Carlo Metropolis simulations for (i) GL-N, and (ii) GL-B, plotted for the range of temperatures.

5.2 Comparison of Tersoff and Cluster Expansion results

Several similarities and differences of both CE and Tersoff approaches can be noted, for calculations on rigid hexagonal GL-N and GL-B systems.

- Both approaches show correct behavior in the first coordination shell. N-N (B-B) bonds are energetically unfavorable. Warren-Cowley parameters for the first coordination shell are similar, they confirm anti-preference of N atoms to create bonds within GL lattice. The same trends are observed for B atoms.
- CE method describes better total energy of formation as a function of concentration. Tersoff represents acceptably differences in energy within the same stoichiometry (*e.g.*, various configurations of $C_{0.9}B_{0.1}$) but is not quantitatively correct in relating energies of systems with various concentrations of species, especially in the case of C-N systems. As a result, mixing energy of N-alloyed graphene is described properly in CE approach, but Tersoff gives wrong result owing to the

overestimation of C-N interaction. In B-alloyed graphene, values of mixing energies are approximately two times larger in Tersoff when compared to Cluster Expansion results.

- Shape of the entropy in pair approximation is quite similar, due to the fact that S_{conf} takes into account only atomic pairs creating bonds in the lattice. Both Tersoff and Cluster Expansion models give the same answer here - N-N and B-B bonds are unfavorable, so configurations after Monte Carlo simulations result in similar number of bonds of each type.
- Cluster expansion models show preference of N-N and B-B atoms to occupy third coordination shell. This might be artifact of the fitting procedure and size of the supercell in reference set of structures, and such behavior is not observed in results obtained within VFF.
- Free energy for B-alloyed graphene is positive for both methods, in the temperatures as high as 3000K. Both approaches show that boron is insoluble in graphene under thermodynamic equilibrium. The same result holds for nitrogen solubility in CE model.

5.3 Note on the 2D simulations with relaxed geometry.

In the course of our research, we have performed also geometry optimization, using Conjugate Gradient (CG) algorithm (Hestenes *et al.*, 1952). We were trying to extend our Monte Carlo scheme *via* adding CG steps inbetween of Monte Carlo swaps of atomic types. We have tested various frequencies of CG optimization insertion in range 1-20000 MC steps, and alternating number of CG steps from 1 to 50 for each CG instance. However, it turns out that the lowest energy is obtained for all classes of CBN alloys, when no geometry optimization is included in MC algorithm, but only at the end, after whole MC simulation of rigid lattice. It is quite easy to interpret though. If we optimize geometry 'too early' before system finds its best configuration, we would fall into some local minimum, when reverse swaps would be very unlikely. As a reminder, in graphene (doped) lattice C-N bond is shorter compared to C-C, while C-B is much longer. Similar situation occurs in hBN doped with C. C-N and C-B optimum bondlength deviates from B-N length in pristine system, see analysis of point defects in GL and hBN in Tabs. 4.8-4.9. Therefore, results of ordering in all classes of investigated systems would be exactly the same in the systems on rigid and relaxed lattice. According to energy values, we have

seen that Tersoff potential has many drawbacks in describing specific values of energy in CBN systems. There is a chance that applying another simulation scheme would allow for including geometry optimization to the search of optimum configurations. However, that would demand other tools, and developing completely new approach.

Employing VFF methodology for modeling of 2D graphene based CBN structures could turn out to be particularly useful in the case of nonperiodic or even amorphous systems such as nanoribbons or nanoflakes with extended defects. In such systems, in addition to the chemical order effects, the local relaxation of system's geometry might be of particular importance. In the last section of this chapter, we signalize the prospects of the Tersoff type VFF approach turning to exemplary graphene and boron nitride flakes. In such a case, finding equilibrium bondlengths between the atoms is much more important, as the system has more degrees of freedom.

However, now we would like to focus in the next section on the relation between morphology of CBN alloys and their electronic structure.

5.4 Density of states in CBN alloys

The performed studies of the large variety of CBN binary and ternary alloys clearly demonstrate the existence of the pronounced short range ordering in the most energetically preferable configurations. Among the studied systems there is semi-metallic graphene as well as hexagonal boron nitride - large gap insulator. Therefore, it is extremely interesting to figure out to which extent the band gap could be tuned through synthesizing a suitable composition of these three elements. In the present section, we investigate the effects of specific arrangements of atoms on the resulting electronic structure of CBN alloys. For systems including thousands of atoms, the natural way of presenting the electronic properties is Density of States (DOS) and projected Density of States (PDOS). To determine DOS and PDOS for CBN alloys, we employ simple Empirical Tight Binding (ETB) approach. As it is well known, the low energy (*i.e.*, in the neighborhood of the Fermi energy) dispersion relations in graphene are very well described considering only π -type orbitals

(p_z) per atom and only first neighbor hopping characterized by single hopping parameter t , as was described in Sec. 2.4.5. To perform the calculations, we use values of parameters obtained through fitting the TB electronic structure of hBN to the predictions of the density functional theory (Giraud, 2012), and specifically, the on-site energies equal -2.55 eV, 0.0, and 2.46 eV, for nitrogen, carbon, and boron, respectively, while $t = -2.16$ eV. This choice of parameters leads to the energy gap of hBN equal to 5.01 eV. For comparison, we tested also another set of TB parameters (DiBenedetto, 2017) to calculate the Density of States of several structures and observed only insubstantial changes of the results. We determine 1800 energies (ϵ_α) of the TB hamiltonian for the simulation supercell used in Monte Carlo calculation for alloys through the direct diagonalization. Then the DOS(E) is obtained by summing over 1800 gaussians, each characterized by the energy (ϵ_α) and gaussian's broadening $\sigma = 0.04 - 0.08$ eV.

For all classes of CBN structures, we calculate and plot the Total Density of States (TDOS) as well as DOS projected on the lattice sites occupied by atoms C, B, and N (PDOS). This allows us to analyze the contributions of various atomic species to the total DOS. We also determine Fermi energy E_F in all cases, which we take as the energy of the Highest Occupied Molecular Orbital (HOMO), ϵ_{HOMO} . We determine also the energy of the Lowest Unoccupied Molecular Orbital (LUMO), ϵ_{LUMO} , and the band gap $E_{GAP} = \epsilon_{LUMO} - \epsilon_{HOMO}$.

Figs. 5.9-5.11 show TDOS and PDOS for exemplary CBN alloy structures, that were obtained from TB hamiltonian for atomic distributions over the hexagonal lattice sites corresponding to configurations with the lowest total energy, as found out in Monte Carlo simulations presented in Sec. 5.1.1.

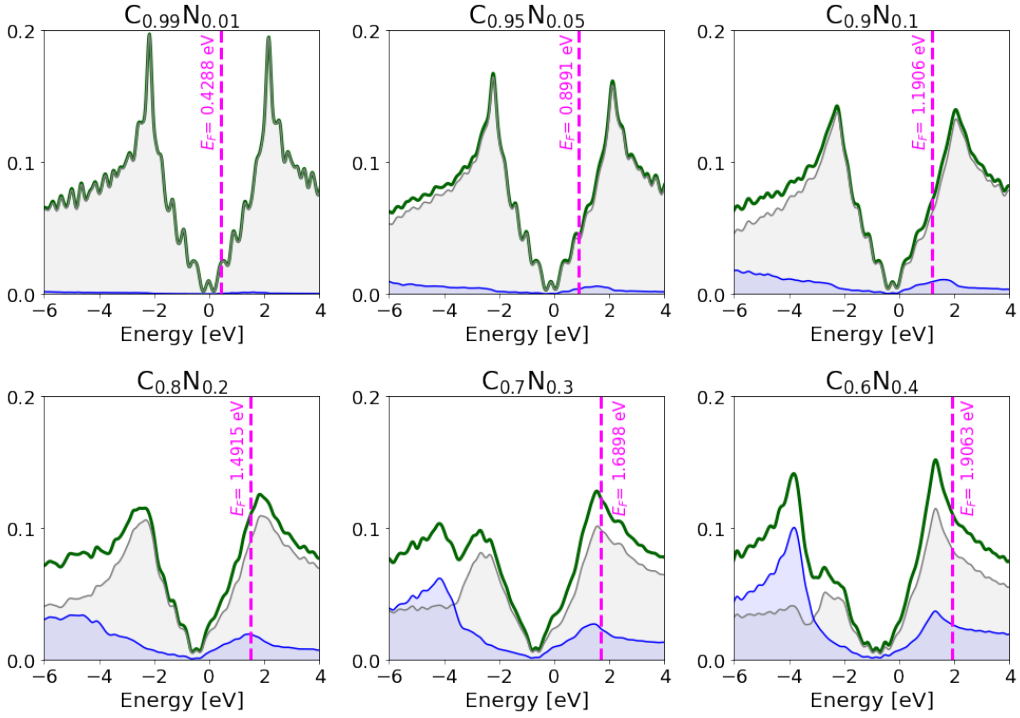


FIGURE 5.9: Total and Projected Density of States (DOS) for GL-N alloys with C and N concentrations indicated. Fermi energy is also marked in each case. Green line denotes Total DOS, gray - DOS projected on carbon atoms, and blue - DOS projected on nitrogen atoms. E_F shows the value of Fermi energy.

We start the presentation of the alloys' electronic structures with results for graphene-nitrogen binary alloy in Fig. 5.9. The results for six configurations with nitrogen concentrations ranging from 1% to 40% are depicted. We can easily notice that DOS for the smallest N concentration is very close to pristine graphene one. With increasing number of N atoms in the system, we observe the evolution of DOS and its gradual distortion, caused by the presence of nitrogen states, mostly in the valence band. Moreover, Fermi energy moves towards higher energies, which apparently demonstrates n-doping effect of introducing nitrogen into graphene lattice.

Analogous situation for B-doping in graphene is visualized in Fig. 5.10, where p-type character of systems is obvious. In this case, the more boron atoms are substituted in place of carbon, the more Fermi energy shifts towards lower energies, deeper into the valence band. At higher concentrations, boron levels modify graphene-like DOS, and B states modify the conduction band rather stronger than

the valence one.

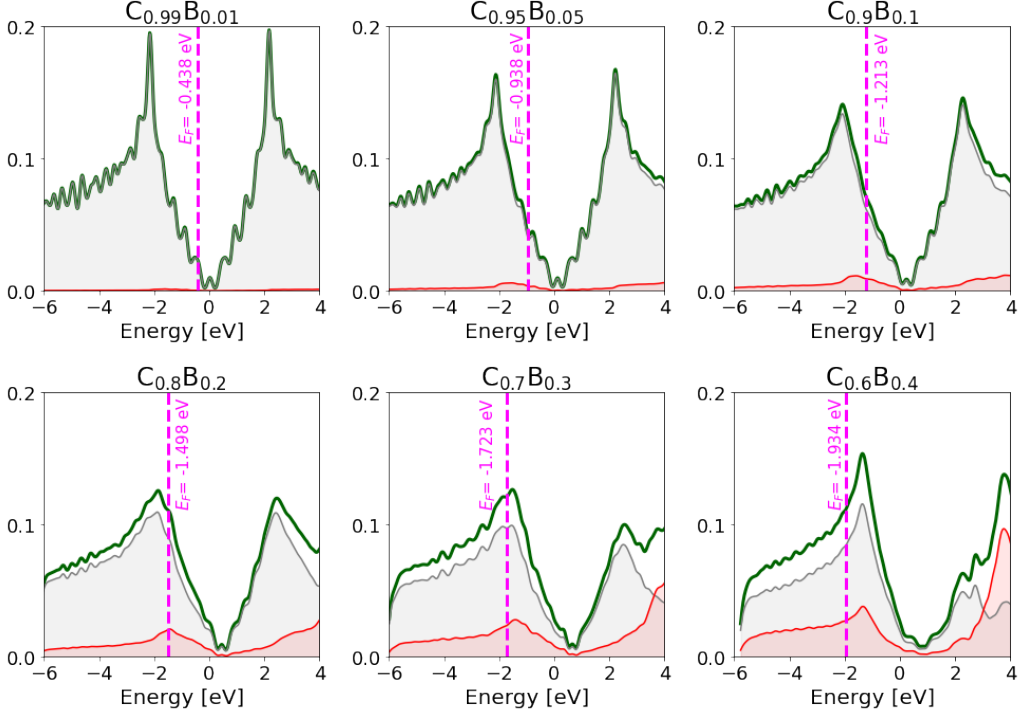


FIGURE 5.10: Total and Projected Density of States (DOS) for GL-B binary alloys with C and B concentrations indicated. Position of the Fermi energy is also marked in each case. Green line denotes Total DOS, gray - DOS projected on carbon atoms, and red line - DOS projected on boron atoms. E_F shows the value of Fermi energy.

Now we turn to the discussion of the electronic structure of ternary CBN alloys. Fig. 5.11 presents DOS and PDOS for ternary $C_x(BN)_{1-x}$ alloys in the whole range of considered concentrations x . We decided to show the two classes of the ternary alloys: GL-BN and hBN-C together, to provide a general overview of the gradual transition from graphene density of states with the close to zero bandgap, into the hBN-like DOS with the band gap approaching 5 eV for the pristine hBN. It is seen that nitrogen states contribute mostly to the valence band, while boron states - to the conduction band. Note that all these ternary alloys are completely compensated, there are always as many occupied states as empty ones. The Fermi energy is close to zero in the systems with prevalent number of C atoms. When B and N atoms (or BN pairs) start to constitute majority of atoms, the energy gap gets clearly pronounced and E_F , taken as ϵ_{HOMO} , shifts towards the top of the valence band.

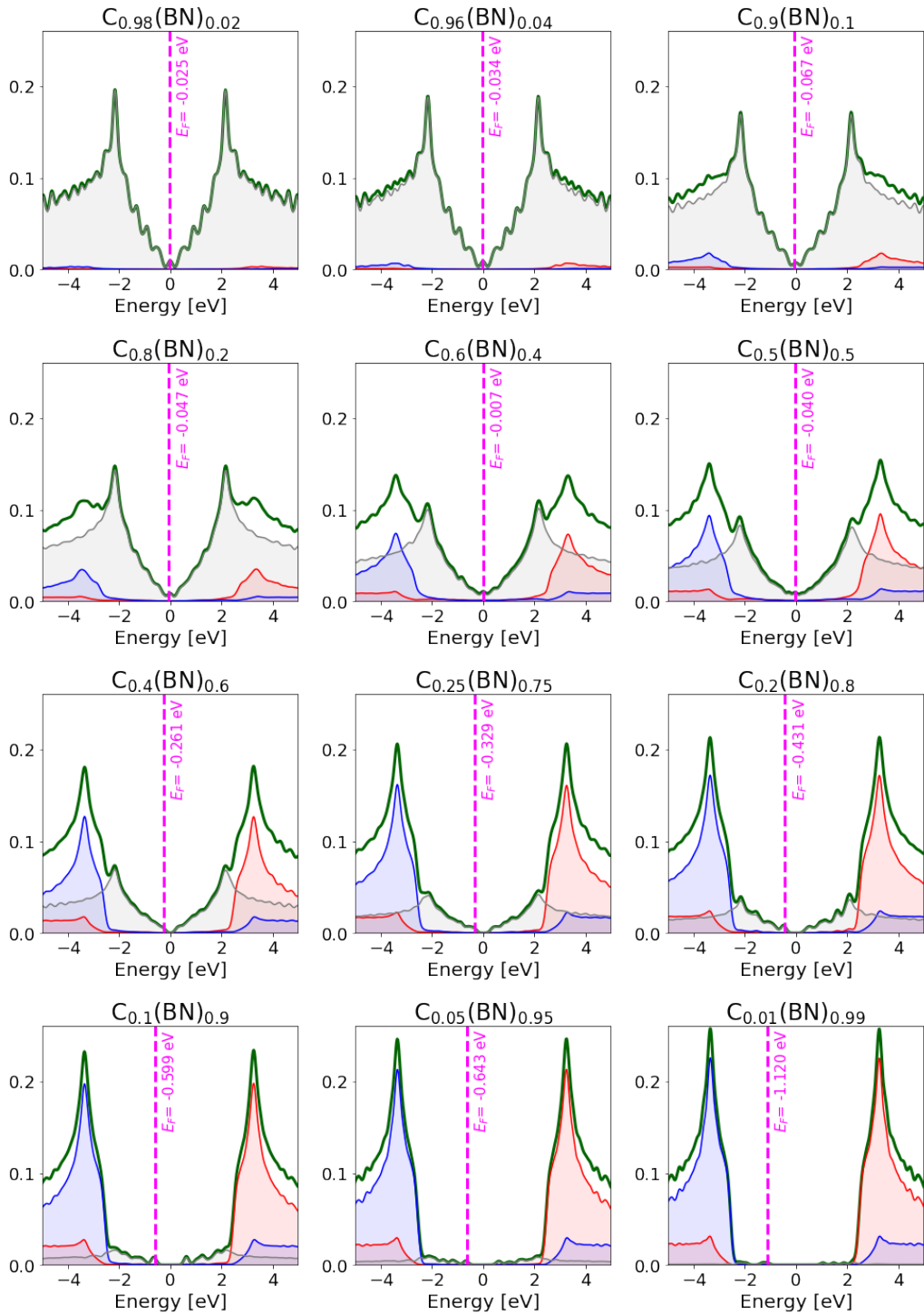
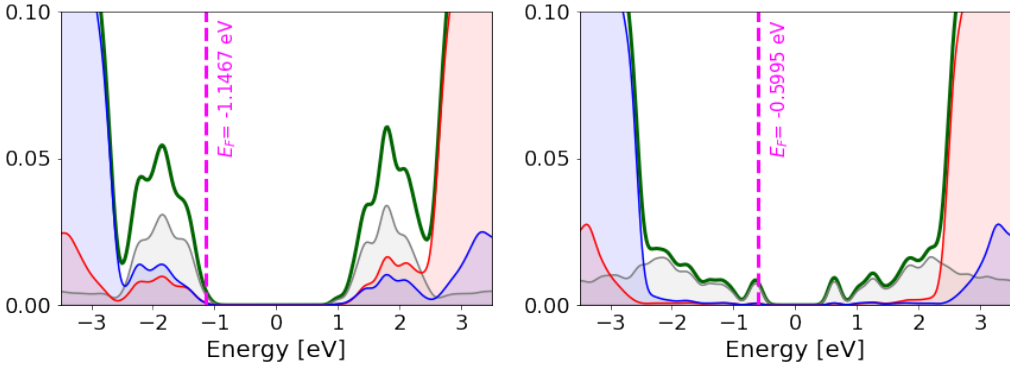


FIGURE 5.11: Total and Projected Density of States (DOS) for $C_x(BN)_{1-x}$ alloys exhibiting short range ordering for $x \in [0.01, 0.98]$ (arranged from the highest to lowest C concentration). Fermi energy E_F is also marked in each case. Green line denotes Total DOS, gray - DOS projected on carbon atoms, blue - DOS projected on nitrogen atoms, and red - DOS projected on boron atoms.

It turns out that the details of the electronic structure are related strongly to the distribution of atoms over the lattice sites, and particularly the effect of SRO forming is extremely relevant. Since all Monte Carlo searches of the atomic configuration leading to the lowest total energy have been started employing the random distribution of atoms, one can easily compare the details of the electronic structure for these two distribution of atoms in an alloy.

As an illustrative example, we present comparison of DOS and PDOS in the case of hBN with 10% of carbon atoms (substituted equally on both sublattices) in Fig. 5.12, where the electronic structure for (A) random distribution, and (B) distribution exhibiting the SRO, are depicted. One can notice significant difference of PDOS connected with carbon atoms and with the total DOS as the result. For random distribution of C, contributions of carbon states in the area of hBN band gap are clearly visible. They can be identified as the states of isolated C_B and C_N impurities in hBN lattice, which act like donors and acceptors, respectively. The character of the electronic structure changes in the system with SRO. Carbon states are more uniformly distributed (C-C nearest neighbor get preferable), graphene like DOS starts to build, and the band-gap gets considerably smaller.



(A) Random distribution of C substitutions (B) Arrangement exhibiting segregation over hBN lattice. into graphene and hBN domains.

FIGURE 5.12: Density of States in the area of hBN band gap for $C_{0.1}(BN)_{0.9}$ alloy with various arrangements of atoms, showing the influence of short range order on DOS shape. Fermi energy E_F is also marked for both configurations. Green line denotes Total DOS, gray - DOS projected on carbon atoms, blue - DOS projected on nitrogen atoms, and red - DOS projected on boron atoms.

Finally, we show the evolution of the band gap with the varying ratio of carbon to

BN pairs in the ternary $C_x(BN)_{1-x}$ systems as presented in Fig. 5.13. In the left hand side of this figure, we consider band gaps of hBN-C system that is treated as hBN doped with carbon atoms. In pristine hBN, boron and nitrogen atoms occupy opposite sublattices, each N has 3 B nearest neighbors and vice versa. This placement of B and N atoms over sublattices remains, even if 50% of them gets substituted by C atoms. Substitution of C atoms is random as an initial guess but changes to the one exhibiting the SRO when the configuration with the minimal total energy is found in the MC algorithm. The energy gaps for these two configurations are depicted for x up to 0.5. As one can see the configuration with SRO, minimizing the total energy, leads to the considerably smaller energy gaps, and we have discussed origin of this issue above.

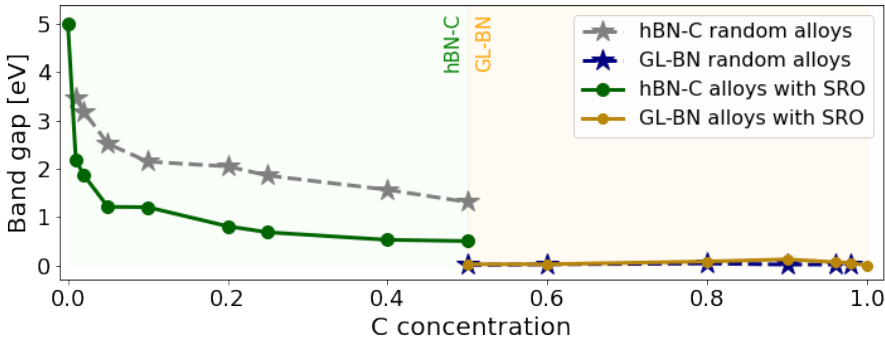


FIGURE 5.13: Value of the band gap in ternary $C_x(BN)_{1-x}$ alloys exhibiting short range order for the whole range of carbon concentrations x . Respective gaps for hBN-C quasi-random, and GL-BN random alloys are also presented.

We conclude that, with increasing number of carbon atoms in hBN, the band gap decreases very fast, from 5 eV for $x = 0$ to 0.5 eV for $x = 0.5$. This would suggest possible way for tuning the band gap in $C_x(BN)_{1-x}$ alloys. Much the same non-linear absolute decline of the band gap was previously reported in several experimental and theoretical studies (B.-Y. Wang *et al.*, 2016; Xie *et al.*, 2016). However, the band gap bowing observed in the bulk hBN (Uddin, Doan, *et al.*, 2014) is much smaller than the value reported here. In other studies (D’Souza *et al.*, 2015), the decrease of the energy gap in alloy with concentration is very abrupt in full analogy to our findings. We would like to point out that provided one could realize $C_x(BN)_{1-x}$ alloy without mixing B and N atoms on the same sublattice, the trend in the band gap between $x > 0.5$ and $x = 1$ (pristine graphene) should be maintained, and the band gap should decrease more or less smoothly from 0.5 eV to 0 eV.

However, the alloys $C_x(BN)_{1-x}$ with $x > 0$ of the type GL-BN have been realized in such a way that all three atoms have been distributed over the two sublattices. Also in this type of CBN alloys, the SRO forms, but N and B atoms are still distributed over two sublattices, their energy states are not separated, and in result the band gaps are very tiny, as illustrated in the right hand part of Fig. 5.13. In this type of GL-BN alloys, the gaps for complete random distribution of atoms are not different than in the case of SRO.

To conclude this section, we would like to point out that the studies of the electronic structure of CBN alloys shed light on physical mechanisms that could facilitate the band gap tuning. However, the resulting electronic structure can depend crucially on the process of alloy synthesis and these aspects require definitely further investigation.

5.5 Graphene and hBN flakes

We would like to finish this thesis with some proposal for further research of CBN nanostructures. Such systems should be particularly suitable for employment of the VFF and TB methodology we have used here. As exemplary system, let us take triangular and hexagonal nanoflakes, since they are reported in the literature (P.-C. Lin *et al.*, 2017; J. Zhang, Tan, *et al.*, 2021). We perform geometry optimization of the flakes (using Conjugate Gradient method). The optimized structures are shown in Fig. 5.14. They aren't planar, which is shown on the projections perpendicular to the flake plane. We can observe in all cases considerable shortening of the bonds on the edges. It is worth noting that triangular hBN flake has different number of boron and nitrogen atoms. However, when B and N positions are swapped, Tersoff potential gives exactly the same energy and geometry due to the chosen (symmetric) parameters of B-N interactions.

5.5.1 Boron doped graphene flakes

We consider also few exemplary doped flakes. Fig. 5.15 shows arrangements of B atoms in the hexagonal flake containing 1734 atoms. One can easily observe that the tendency of boron atoms is to occupy specific area of the flake - third the most outer row. When those positions are already occupied, B atoms situate themselves

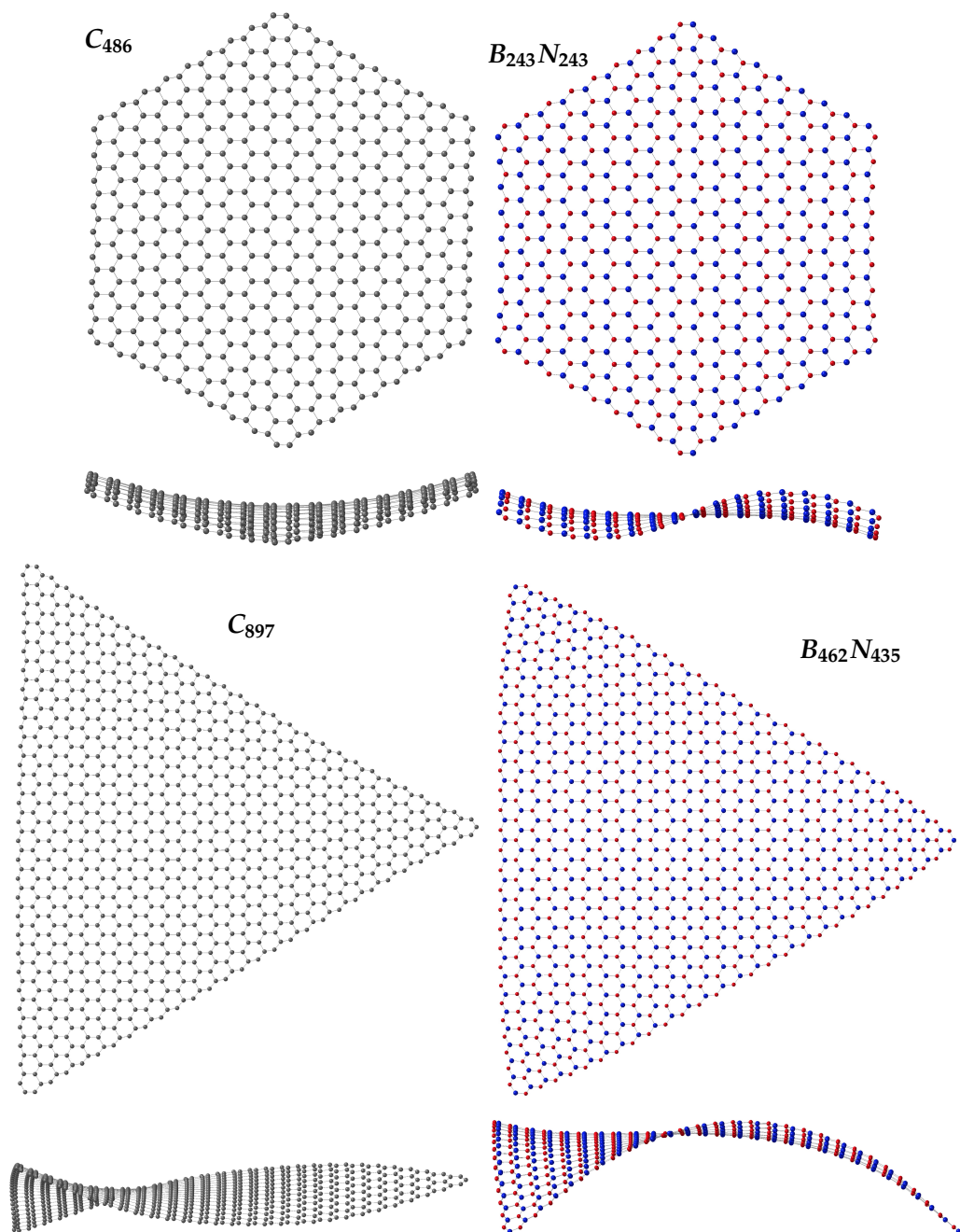


FIGURE 5.14: Top and side views of exemplary graphene and hBN flakes, where geometry was optimized with the use of Tersoff potential. Grey dots denote carbon, blue - nitrogen and pink - boron.

as close as possible to this row, but closer to the inside of the flake, however, avoiding creation of B-B bonds. This results in structures exhibiting specific short-range ordering.

The CBN nanostructures that are finding more and more applications in the nanotechnology are definitely worth of being investigated within VFF methodology. However, we find out that one should carry out a series of methodological developments for reliable modeling of these systems such as better parametrization of valence force field, include charge transfer effects within VFF, work out better procedures of geometry optimisations, develop tight-binding hamiltonians allowing for hopping integrals dependent on the distances between the atoms.

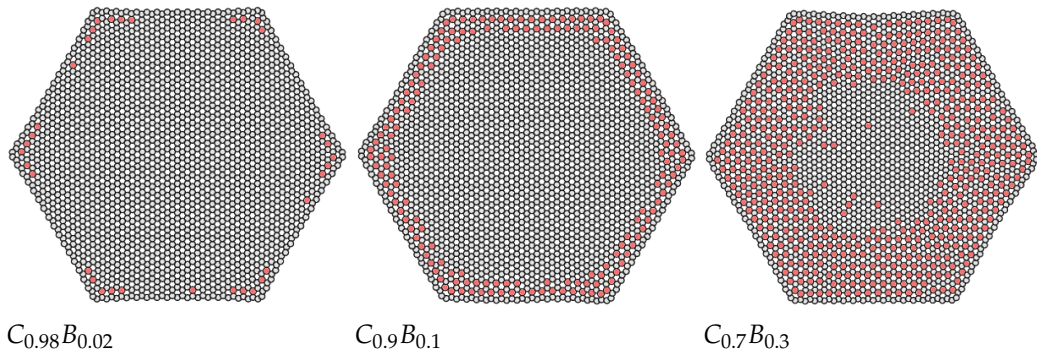


FIGURE 5.15: Boron (red dots) dopants distribution over the relaxed carbon hexagonal flake (1734 atoms) after performing Monte Carlo simulations for exemplary concentrations of B, $c_B = c_N = 2\%$, 10% , and 30% . Initial distribution of B atoms was assumed to be a random one.

Chapter 6

Summary and Outlook

Within this thesis, we employ the multi-scale approach to provide deep understanding of energetic stability of the technologically important class of CBN hexagonal two-dimensional non-periodic alloys. We focus on the chemical order of boron, nitrogen, and carbon atoms on honeycomb graphene lattice and determine the equilibrium distribution of atoms in binary C-N and C-B, and ternary C-B-N alloys. We determine how the distribution of atoms on the lattice determines electronic properties of these alloys. To perform these studies, we have surveyed and verified well-known methods of condensed matter physics: empirical atomistic energy calculations within valence force field (VFF) schemes, *ab initio* methods, Cluster Expansion (CE) approach, Monte Carlo (MC) optimization technique, determination of electronic structure. The complexity of the physico-chemistry of the CBN 2D alloys practically restricts possibility of their reliable research to empirical in nature atomistic valence force field approaches. Thorough careful tests of several VFFs, *i.e.*, Tersoff and extended Tersoff potentials, ReaxFF, and COMB3, we have decided to base our studies on suitably modified parametrization of Tersoff potential, and named it **Ters-SET3**. These test-studies clearly demonstrate that the present stage of parametrization of all types of VFFs is not satisfactory for the problem of 2D CBN alloys, and requires further development. In particular, in Tersoff potential we improve: (i) behavior regarding the strength of B-B bond in graphene structure, and (ii) to some extent, the originally strongly overestimated strength of interaction between carbon and nitrogen atoms.

The situation in the field of VFFs for 2D CBN alloys strongly suggests that reliable studies of these materials should be augmented by methods allowing for some kind of cross-checking. The method of choice for alloys is definitely Cluster Expansion

(CE) approach that is based on *ab initio* calculations for relatively small number of prototypical systems that are considered as constituent elements of alloy. Here, we develop CE models for binary B- or N- alloyed graphene systems, following an iterative scheme for rigid and relaxed alloy systems. Similarly to other authors (see e.g., (Laks *et al.*, 1992)), we find that the construction of the good CE model with lattice relaxation taken into account is more difficult. Nevertheless, rigid and relaxed CE models find the existence of the short range order in GL-B and GL-N alloys, and distribution of atoms in these binary alloys is far from being random. Since we focus mostly on the issue of chemical order in alloys, we use the total energy and atomic positions determined in the rigid CE models to find corrected by the presence of the SRO configurational entropy, and further the free energy. All these studies clearly show that even at temperatures around 3000K the binary GL-B and GL-N alloys are not thermodynamically stable. Our results corroborate findings of other authors, which strongly suggest that incorporation of B and N into graphene lattice is rather governed by non equilibrium processes.

We have used VFF based on optimised Tersoff potential to perform large scale modeling of binary (GL-B and GL-N) and ternary $(C_{2x}(BN)_{1-2x})$ alloys employing large supercell with thousands of atoms, and therefore, avoiding artificial long-range order that is very often spuriously built-in in methodology based on superlattices containing several or dozens of atoms.

For all investigated concentrations of C, B, N species, the most energetically favorable arrangements of atoms stray from random, which is confirmed by calculations of short range order Warren-Cowley parameters. The description of SRO within VFF agrees very well with CE models for binary alloys. Uniformly distributed dopants in graphene and hBN are not found in equilibrium conditions, which is confirmed in calculations of entropy and free energy.

Employing the low energy Tight Binding model for calculation of the electronic structure (total and projected density of states), we shed light on physical mechanisms that determine the dependence of the alloy's electronic structure on the distribution of atoms over the lattice. This reveals relation between distribution of atoms and resulting electronic structure, should facilitate the interpretation of the experimental data, and also should provide some hints for design of suitable functional electronic structure through synthesis of alloys with required composition.

In the view of our results, general conclusions can be driven:

- Bond order potentials are valuable tools for investigation of ordering, geometry and phase diagrams of *realistic* systems, as they allow for performing simulations in supercells of basically any size. Their construction often allows for direct physical interpretation of interactions present in the systems.
- Choice of appropriate empirical potential and its careful parametrization is crucial for the quality of simulations.
- VFFs investigated in this thesis, despite many advantages, demand certain improvements to capture more accurate features of CBN lateral alloys.
- Cluster Expansion approach can reproduce *ab initio* results for small superlattices with fairly high precision, however, does not provide, as VFF does, the possibility of direct check for systems with thousands of atoms and with spurious long-range order effects eliminated.
- Short range order is detected in all investigated classes of CBN lateral alloys, and the temperature of phase transition from ordered to disordered systems would be higher than the melting temperature.
- Band gap engineering is possible in CBN alloys, especially *via* doping hBN with carbon.

Six years is a lot of time in terms of scientific development nowadays, and numerous papers concerning CBN graphene-like systems (both experimental and theoretical) have been published, since I started my research. New methods and algorithms for determination interatomic interactions are being developed right now, especially using machine learning (ML). Several reports on ML application in the area of CBN graphene-like systems are available (Y. Du *et al.*, 2021; M. Dieb *et al.*, 2018; Qian *et al.*, 2021; Thiemann *et al.*, 2020), showing improved accuracy as compared to VFFs mentioned in this thesis. This novel approach for designing models for energy calculations shows surprising efficiency and potential for progress. Thus, we would like to point out two directions for future extension of research presented here. First, development of better parametrization of VFFs describing CBN alloy materials, and second - investigating the most up-to-date methods of description of condensed phases via machine learning approach. We believe also that hybrid approaches involving VFF and CE could contribute to the reliable modeling of nano-alloys.

As my personal summary, there are several life lessons I have learned throughout this research. Humbleness. Thoroughness. Perseverance. Honesty. I could appreciate the value of cooperation and mutual support. If the results presented here contribute in a small part to understanding physical phenomena in the field of 2D

materials and the methods utilized to describe them, I would consider my work meaningful.

Appendix A

Details of determination of the Cluster Expansion models

A.1 Clusters Pool

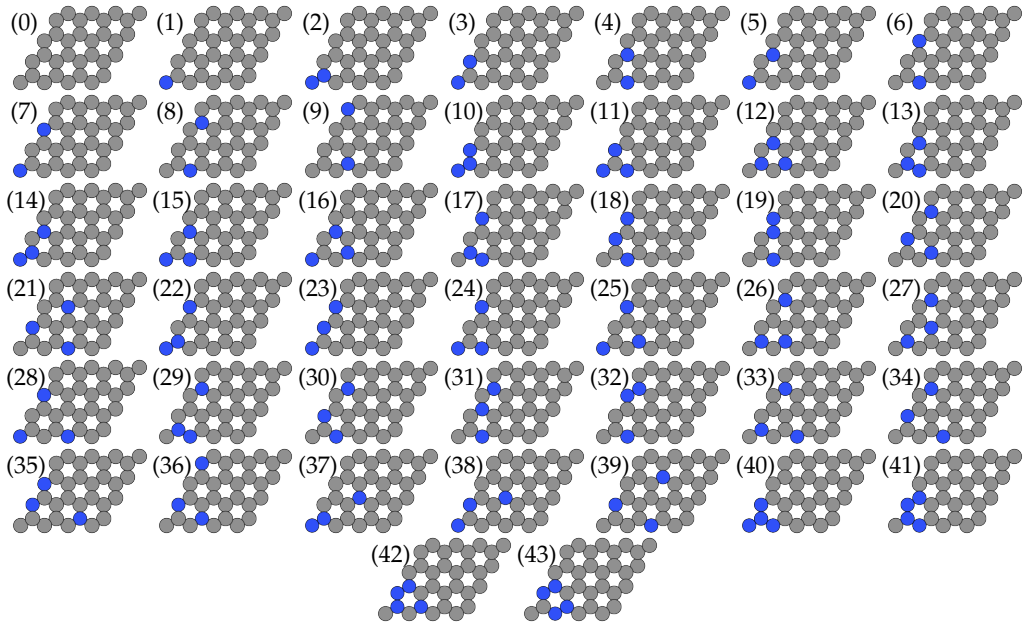


FIGURE A.1: Clusters Pool used as a source of clusters for fitting all the instances of Cluster Expansion models considered in this thesis.

A.2 Evolution of CE models determination

The details of recursive method for determination Cluster Expansion model for graphene alloyed with nitrogen on rigid lattice (**GL-N RIGID**) are given in Sec. 3.3. Remaining variants: graphene alloyed with boron in rigid lattice (**CE2 GL-B RIGID**), graphene alloyed with nitrogen in relaxed lattice (**CE2 GL-N RELAX**) and graphene alloyed with boron in relaxed lattice (**CE2 GL-B RELAX**) are presented below. RMSE and CV scores for all iterations are compared in the main text in Tab. 3.2. Sec. 3.3 contains also analysis and conclusions about the fitted models. For each class of CE models, we present in sequence:

- values of Effective Cluster Interactions determined for CE_i model, as well as the size and radius of respective clusters in Tab. A.1, Tab. A.2 and Tab. A.3,
- energies of subsequent reference sets before and after Monte Carlo (MC) simulations computed with CE_i models are compared to respective DFT data, showing accuracy of successive iterations in Fig. A.2, Fig. A.5, Fig. A.8,
- exemplary initial (random) positions and positions after MC simulations in structures containing 800 atoms for exemplary concentrations of nitrogen (boron) are set side by side in Fig. A.3, Fig. A.6, Fig. A.9. This allows for visual analysis of ordering of substituted atoms in each case,
- values of Warren-Cowley short range order parameter for NN (BB) pairs distributed within the lattice, calculated and depicted in Fig. A.4, Fig. A.7, Fig. A.10.

GL-B (RIGID)

TABLE A.1: Cluster Expansion models (CE1, CE2, CE3) for graphene doped with boron on a rigid lattice obtained in the iterative procedure. For each iteration, clusters included in CE_i model are characterized with a number of points in cluster N_{pts} and its radius R , followed by value of Effective Cluster Interaction parameter associated with this cluster.

	CE1 GL-B RIGID			CE2 GL-B RIGID			CE3 GL-B RIGID		
	N_{pts}	$R[\text{\AA}]$	ECI	N_{pts}	$R[\text{\AA}]$	ECI	N_{pts}	$R[\text{\AA}]$	ECI
0	0	0.000	1.1203	0	0.000	0.8072	0	0.000	0.9156
1	1	0.000	1.4714	1	0.000	0.3774	1	0.000	0.7791
2	2	1.420	0.1775	2	1.420	0.0182	2	1.420	-0.0112
3	2	2.460	0.3755	2	2.460	-0.4368	2	2.460	-0.0616
4	2	2.840	-0.1931	2	2.840	-0.1851	2	2.840	0.0562
5	2	3.757	-0.1085	2	3.757	-0.2515	2	3.757	-0.2422
6	3	4.260	-0.2786	2	4.920	-0.1185	2	4.920	-0.2525
7	3	4.920	0.2076	3	3.757	-0.2451	3	2.460	-0.0913
8	3	4.920	0.0359	3	3.757	-0.0998	3	4.260	-0.0915
9	3	5.120	-0.1105	3	4.920	-0.1077	3	4.260	0.0795
10	3	5.120	0.1507	3	4.920	-0.0534	3	4.920	-0.2495
11	4	2.840	0.0375	3	5.120	0.0484	3	4.920	-0.0521
12	4	2.840	0.0670	4	2.840	0.0825	3	5.120	0.0383

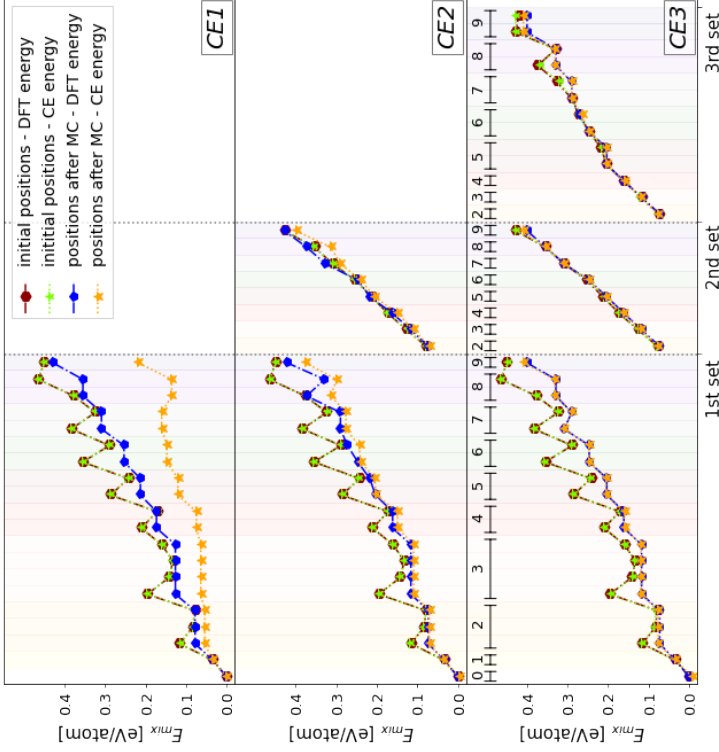


FIGURE A.2: Predictive power of consecutive iterations CE1-CE3 during iterative procedure of Cluster Expansion model fitting for B-doped graphene in rigid lattice (**GL-B RIGID**). For iteration i , energies before and after Monte Carlo simulations are plotted, calculated both with CE_i model and with DFT. x-axis shows consecutive reference sets. The number of dopants in each configuration is shown in the lower part of the figure.

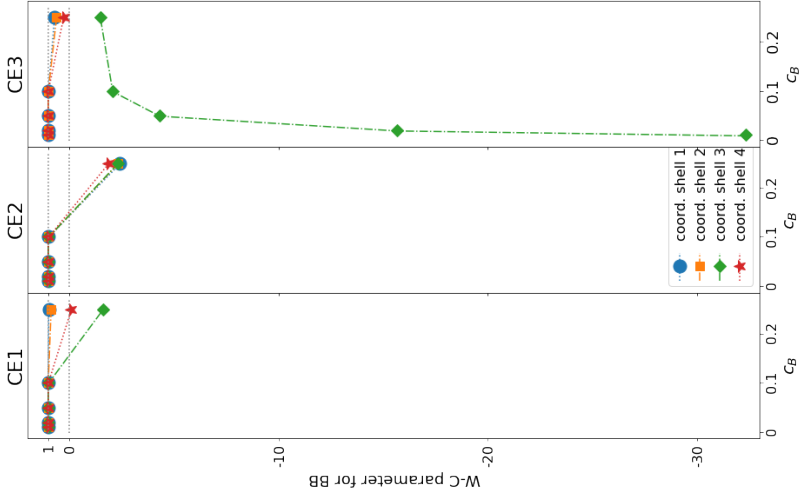


FIGURE A.4: Warren-Cowley SRO parameters for B-B pairs extracted from minimum energy structures obtained in Monte Carlo Simulations with i^{th} generation of CE model on rigid lattice (denoted MCI). Parameters were calculated for $1^{st} - 4^{th}$ coordination shells.

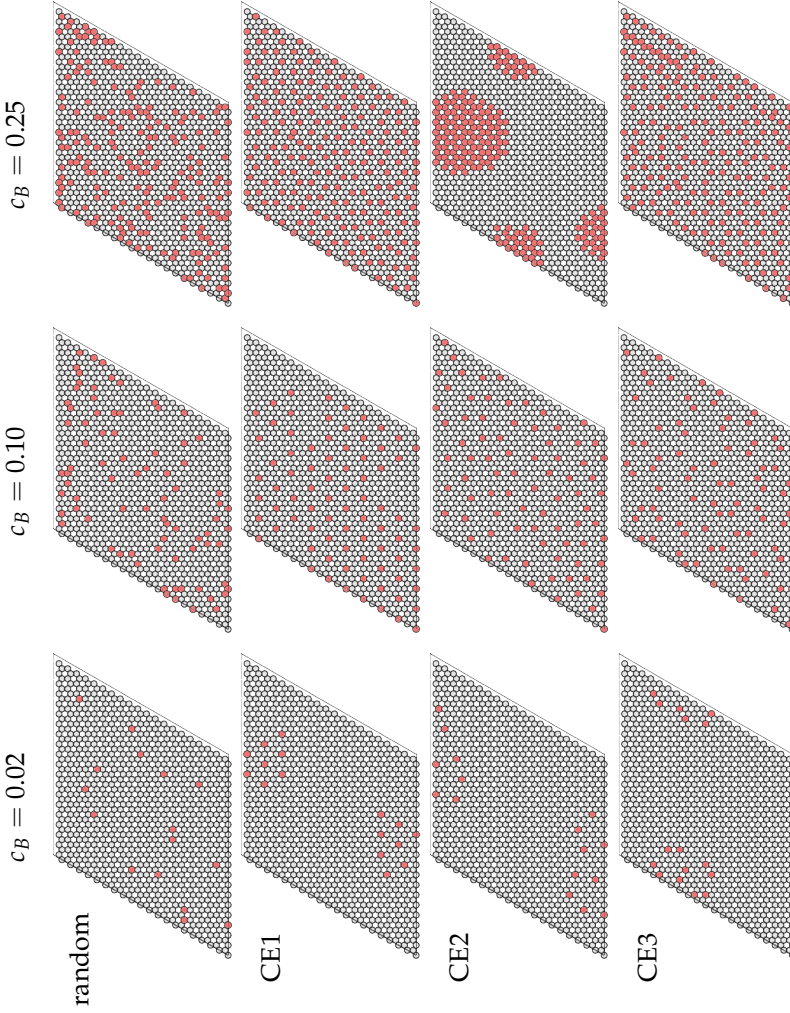


FIGURE A.3: Evolution of boron dopants (red dots) distribution over the rigid carbon honeycomb lattice after performing Monte Carlo simulations (at $T = 10K$) for progressing generations of Cluster Expansion models (CE*i*) presented for three exemplary concentrations of B, 2%, 10%, and 25%. For each B concentration, initial distributions over the graphene lattice were assumed to be a random one.

GL-N (RELAX)

TABLE A.2: Cluster Expansion models (CE1, CE2, CE3) for graphene doped with nitrogen on a relaxed lattice obtained in the iterative procedure. For each iteration, clusters included in CE i model are characterized with a number of points in cluster N_{pts} and its radius R , followed by value of Effective Cluster Interaction parameter associated with this cluster.

	CE1 GL-N RELAX			CE2 GL-N RELAX			CE3 GL-N RELAX		
	N_{pts}	$R[\text{\AA}]$	ECI	N_{pts}	$R[\text{\AA}]$	ECI	N_{pts}	$R[\text{\AA}]$	ECI
0	0	0.000	0.3033	0	0.000	0.6922	0	0.000	0.8298
1	1	0.000	-0.583	1	0.000	1.6922	1	0.000	1.4253
2	2	1.420	-0.2754	2	1.420	0.1003	2	1.420	0.2063
3	2	2.460	-0.5848	2	2.460	0.1451	2	2.460	0.0050
4	2	2.840	-0.0250	2	2.840	0.2346	2	2.840	-0.0348
5	2	3.757	-0.8354	2	3.757	0.0748	2	3.757	-0.0315
6	2	5.120	0.1297	3	3.757	0.0745	2	4.260	0.8050
7	2	5.681	0.0191	3	4.920	0.1266	3	4.260	0.4095
8	3	2.460	-0.1379	3	4.920	-0.1599	3	4.920	0.2769
9	3	2.460	-0.0274	3	4.920	0.1219	3	4.920	0.0505
10	3	3.757	-0.5413	3	5.120	0.3261	3	4.920	-0.3503
11	3	4.920	-0.1699	3	5.120	-0.1435	3	5.120	0.0373
12	3	5.120	0.1887	4	3.757	0.1014	3	5.681	-0.0692

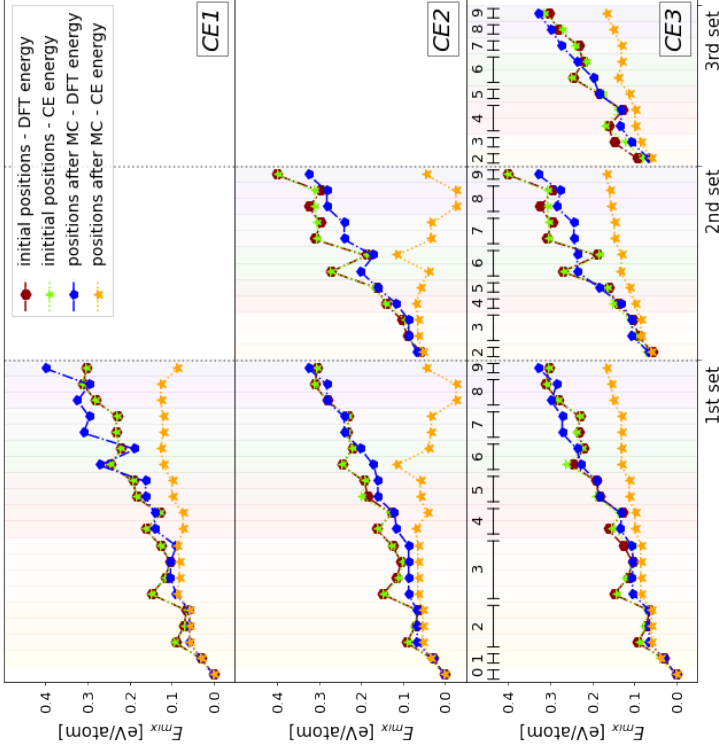


FIGURE A.5: Predictive power of consecutive iterations CE1-CE3 during iterative procedure of Cluster Expansion model fitting for N-doped graphene in rigid lattice (GL-N RELAX). For iteration i , energies before and after Monte Carlo simulations are plotted, calculated both with CE i model and with DFT. x-axis shows consecutive reference sets. The number of dopants in each configuration is shown in the lower part of the figure.

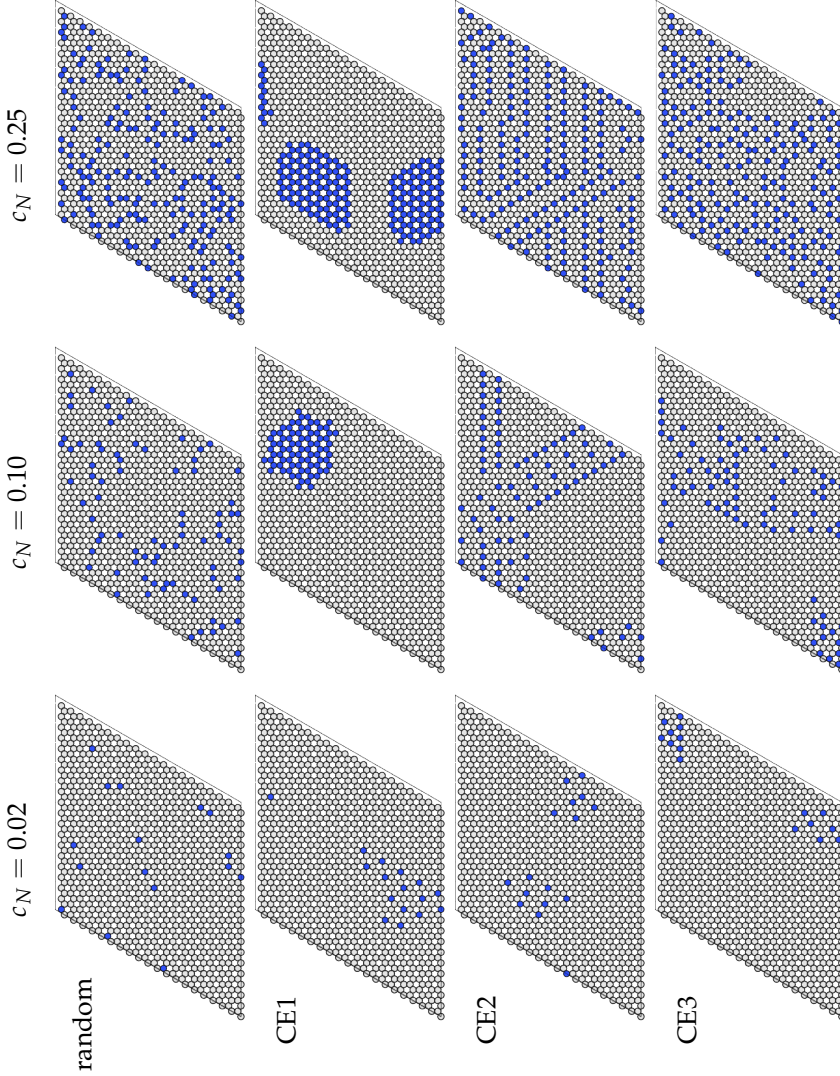


FIGURE A.6: Evolution of nitrogen dopants (blue dots) distribution over the relaxed carbon honeycomb lattice after performing Monte Carlo simulations (at $T = 10\text{K}$) for progressing generations of Cluster Expansion models (CE i) presented for three exemplary concentrations of N, 2%, 10%, and 25%. For each B concentration, initial distributions were assumed to be a random one.

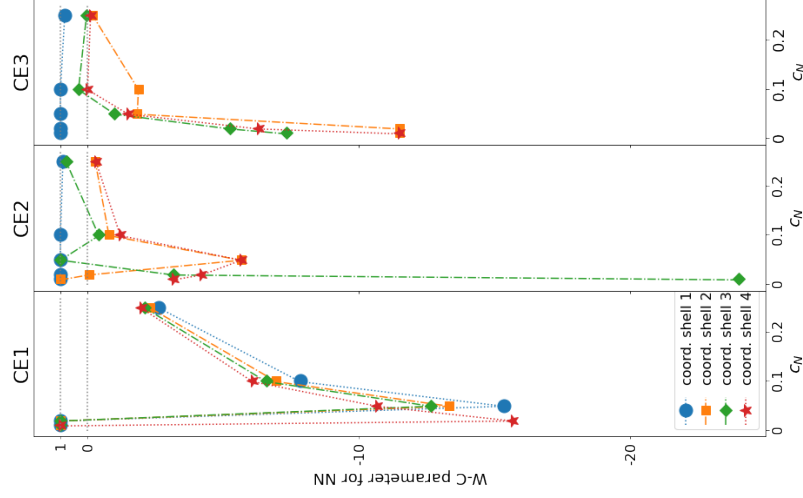


FIGURE A.7: Warren-Cowley SRO parameters for N-N pairs extracted from minimum energy structures obtained in Monte Carlo Simulations with i^{th} generation of CE model on relaxed lattice. Parameters were calculated for $1^{\text{st}} - 4^{\text{th}}$ coordination shells.

GL-B (RELAX)

TABLE A.3: Cluster Expansion models (CE1, CE2, CE3) for graphene doped with boron on a relaxed lattice obtained in the iterative procedure. For each iteration, clusters included in CE_i model are characterized with a number of points in cluster N_{pts} and its radius R , followed by value of Effective Cluster Interaction parameter associated with this cluster.

CE1 GL-B RELAX				CE2 GL-B RELAX				CE3 GL-B RELAX			
N_{pts}	$R[\text{\AA}]$	ECI		N_{pts}	$R[\text{\AA}]$	ECI		N_{pts}	$R[\text{\AA}]$	ECI	
0	0.000	-0.2875		0	0.000	0.2980		0	0.000	0.4682	
1	1.000	-2.1714		1	0.000	0.1370		1	0.000	0.5871	
2	1.420	-0.9958		2	1.420	-0.1341		2	1.420	-0.0986	
3	2.460	-0.7555		2	2.460	0.2505		2	2.460	0.0835	
4	2.840	-0.1934		2	2.840	-0.0158		2	2.840	0.1042	
5	3.757	-0.4925		2	3.757	-0.1442		2	3.757	0.1434	
6	4.260	-0.0941		2	4.260	0.0517		3	2.460	-0.1031	
7	4.920	-0.5831		3	2.460	0.1997		3	4.920	-0.0768	
8	2.460	-0.3969		3	4.920	0.0570		3	4.920	0.1148	
9	3.757	-0.1393		3	5.120	0.1299		3	5.120	0.1059	
10	3.4920	-0.4831		4	3.757	0.1334		3	5.120	0.2389	
11	3.4920	-0.2111		4	3.757	0.1790		3	5.120	-0.1335	
12				4	3.757	-0.0964		3	5.681	-0.0319	

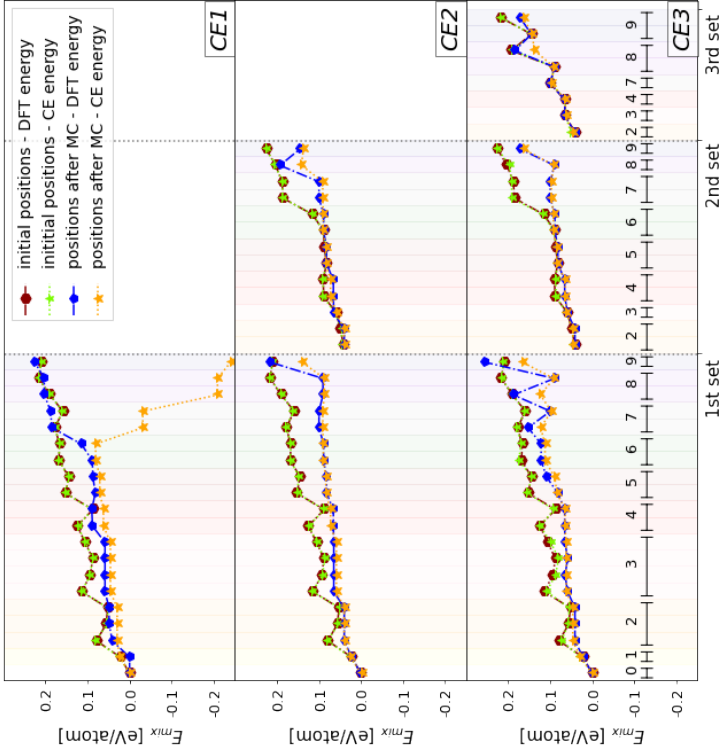


FIGURE A.8: Predictive power of consecutive iterations CE1-CE3 during iterative procedure of Cluster Expansion model fitting for B-doped graphene in rigid lattice (GL-B RELAX). For iteration i , energies before and after Monte Carlo simulations are plotted, calculated both with CE i model and with DFT. x-axis shows consecutive reference sets. The number of dopants in each configuration is shown in the lower part of the figure.

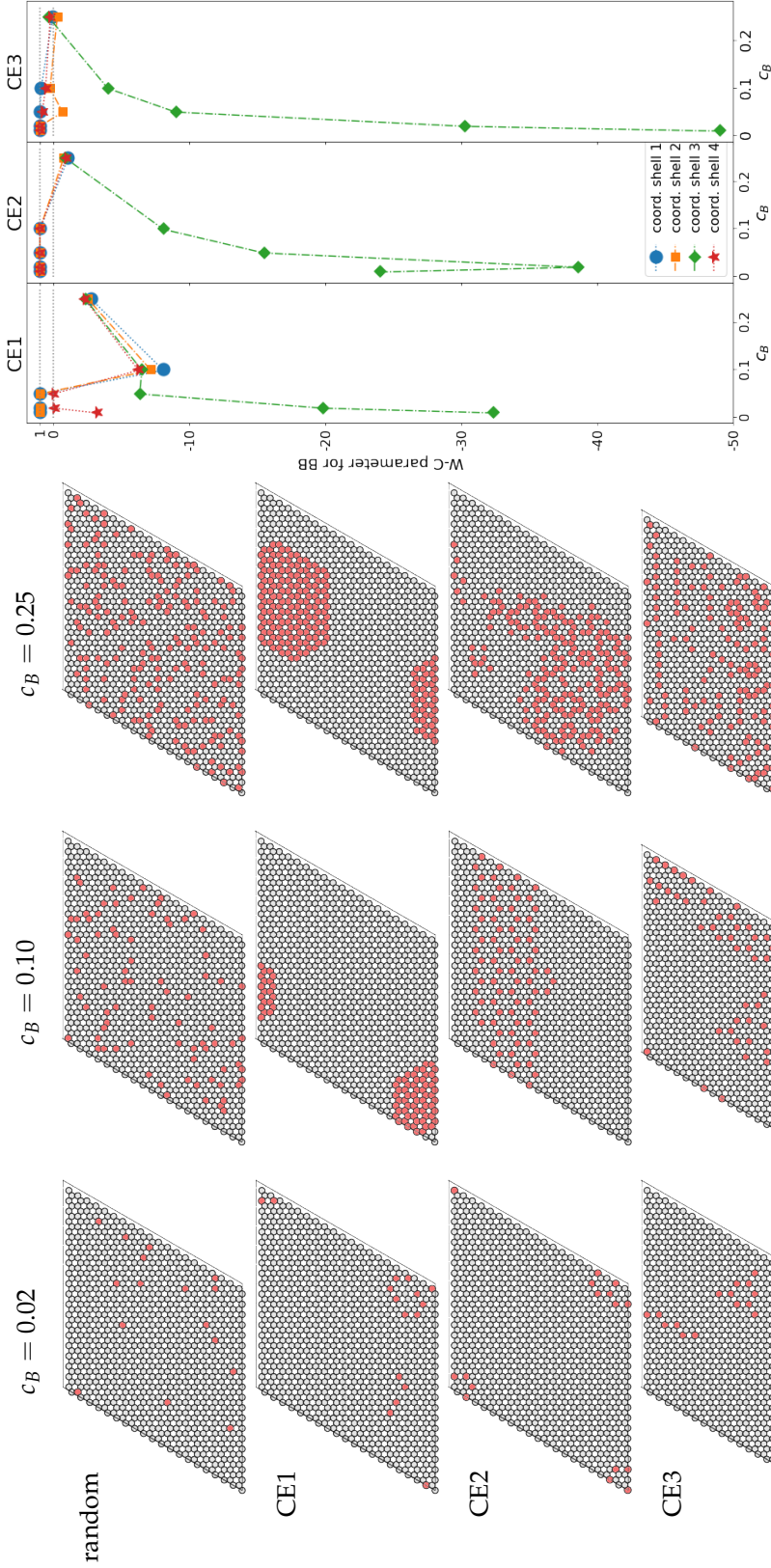


FIGURE A.9: Evolution of boron dopants (red dots) distribution over the relaxed carbon honeycomb lattice after performing Monte Carlo simulations (at $T = 10\text{K}$) for progressing generations of Cluster Expansion models (CE*i*) presented for three exemplary concentrations of B, 2%, 10%, and 25%. For each B concentration, initial distributions over the graphene lattice were assumed to be a random one.

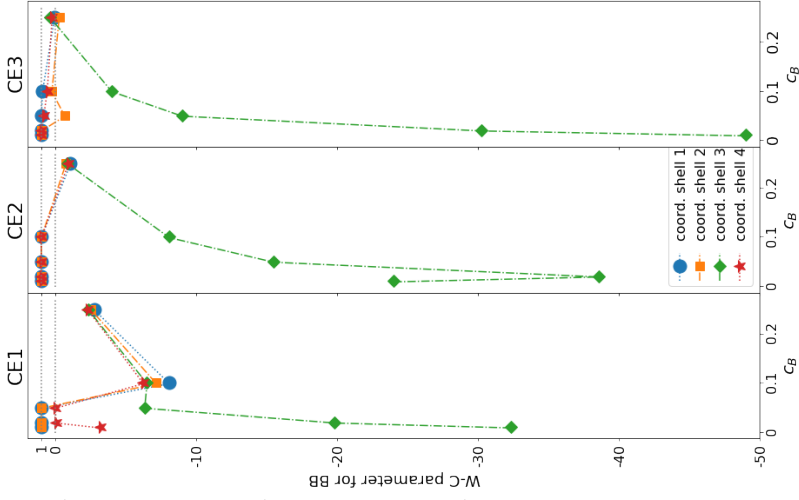


FIGURE A.10: Warren-Cowley SRO parameters for B-B pairs extracted from minimum energy structures obtained in Monte Carlo simulations with i^{th} generation of CE model on relaxed lattice (denoted MCI). Parameters were calculated for $1^{\text{st}} - 4^{\text{th}}$ coordination shells.

Bibliography

- Abdullah, N. R. *et al.* (2020) "Effect of BN dimers on the stability, electronic, and thermal properties of monolayer graphene", *Results in Physics* 18, p. 103282. DOI: <https://doi.org/10.1016/j.rinp.2020.103282>.
- Abergel, D. *et al.* (2010) "Properties of graphene: a theoretical perspective", *Advances in Physics* 59 (4), p. 261. DOI: 10.1080/00018732.2010.487978.
- Agnoli, S. and M. Favaro (2016) "Doping graphene with boron: a review of synthesis methods, physicochemical characterization, and emerging applications", *J. Mater. Chem. A* 4 (14), p. 5002. DOI: 10.1039/C5TA10599D.
- Albe, K. (1998) "Computersimulationen zu Struktur und Wachstum von Bornitrid". PhD thesis. Technische Universität Dresden.
- Albe, K., W. Möller, and K.-H. Heinig (1997) "Computer simulation and boron nitride", *Radiation Effects and Defects in Solids* 141 (1-4), p. 85. DOI: 10.1080/10420159708211560.
- Ångqvist, M. and P. Erhart (Sept. 2017) "Understanding Chemical Ordering in Intermetallic Clathrates from Atomic Scale Simulations", *Chemistry of Materials* 29 (17), p. 7554.
- Al-Aqtash, N., K. Tarawneh, T. Tawalbeh, and I. Vasiliev (Aug. 2012) "Ab Initio Study of Interaction between Boron and Nitrogen Dopants in Graphene", *Journal of Applied Physics* 112, p. 034304. DOI: 10.1063/1.4742063.
- Azevedo, S., J. Kaschny, C. M. C. de Castilho, and F. B. Mota (2007) "A theoretical investigation of defects in a boron nitride monolayer", *Nanotechnology* 18 49, p. 495707. DOI: 10.1088/0957-4484/23/48/489501.
- Azevedo, S. (2006) "Energetic stability of B-C-N monolayer", *Physics Letters A* 351 (1), p. 109. DOI: 10.1016/j.physleta.2005.10.053.
- Barrios-Vargas, J. E. *et al.* (2017) "Electrical and Thermal Transport in Coplanar Polycrystalline Graphene-hBN Heterostructures", *Nano Letters* 17 (3), p. 1660. DOI: 10.1021/acs.nanolett.6b04936.
- Bawari, S. *et al.* (2020) "Structural evolution of BCN systems from graphene oxide towards electrocatalytically active atomic layers", *Mater. Chem. Front.* 4 (8), p. 2330. DOI: 10.1039/D0QM00220H.
- Becke, A. D. (2014) "Perspective: Fifty years of density-functional theory in chemical physics", *The Journal of Chemical Physics* 140 (18), 18A301. DOI: 10.1063/1.4869598.
- Beniwal, S. *et al.* (2017) "Graphene-like Boron-Carbon-Nitrogen Monolayers", *ACS Nano* 11 (3), p. 2486. DOI: 10.1021/acs.nano.6b08136.

- Bepete, G. *et al.* (2013) "Incorporation of small BN domains in graphene during CVD using methane, boric acid and nitrogen gas", *Nanoscale* 5 (14), p. 6552. DOI: 10.1039/C3NR01699D.
- Berseneva, N., A. V. Krashennnikov, and R. M. Nieminen (July 2011) "Mechanisms of Post-synthesis Doping of Boron Nitride Nanostructures with Carbon from First-Principles Simulations", *Phys. Rev. Lett.* 107 (3), p. 035501. DOI: 10.1103/PhysRevLett.107.035501.
- Bhat, S. *et al.* (Mar. 2016) "Electrochemical Li Storage Properties of Carbon-Rich B–C–N Ceramics", *Journal of Carbon Research C* 2, p. 9. DOI: 10.3390/c2020009.
- Binder, K. and D. W. Heermann (2010) *Monte Carlo Simulation in Statistical Physics: An Introduction; 5th ed.* Graduate texts in physics. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-642-03163-2.
- Blöchl, P. E. (Dec. 1994) "Projector augmented-wave method", *Phys. Rev. B* 50 (24), p. 17953. DOI: 10.1103/PhysRevB.50.17953.
- Blum, V. *et al.* (2009) "Ab initio molecular simulations with numeric atom-centered orbitals", *Computer Physics Communications* 180 (11), p. 2175. DOI: <https://doi.org/10.1016/j.cpc.2009.06.022>.
- Botello-Mendez, A., A. Lherbier, and J.-C. Charlier (Dec. 2013) "Modeling electronic properties and quantum transport in doped and defective graphene", *Solid State Communications* 175, p. 90. DOI: 10.1016/j.ssc.2013.08.029.
- Bragg, W. L. and E. J. Williams (1934) "The effect of thermal agitation on atomic arrangement in alloys", *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 145 (855), p. 699. DOI: 10.1098/rspa.1934.0132.
- Bundaleska, N. *et al.* (2018) "Large-scale synthesis of free-standing N-doped graphene using microwave plasma", *Scientific Reports* 8 (1), p. 12595.
- Canadell, E., M.-L. Doublet, and C. Iung (2012) *Orbital Approach to the Electronic Structure of Solids*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199534937.001.0001.
- Chandra, V. *et al.* (2012) "Highly selective CO₂ capture on N-doped carbon produced by chemical activation of polypyrrole functionalized graphene sheets", *Chem. Commun.* 48 (5), p. 735. DOI: 10.1039/C1CC15599G.
- Chen, S. *et al.* (2018) "Carbon doping of hexagonal boron nitride porous materials toward CO₂ capture", *J. Mater. Chem. A* 6 (4), p. 1832. DOI: 10.1039/C7TA08515J.
- Chenoweth, K., A. C. T. van Duin, and W. A. Goddard (2008) "ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation", *The Journal of Physical Chemistry A* 112 (5), p. 1040. DOI: 10.1021/jp709896w.
- Cooper, D. R. *et al.* (Apr. 2012) "Experimental Review of Graphene", *Nature materials* 6, p. 501686. DOI: 10.5402/2012/501686.
- Cowley, J. M. (Mar. 1950) "An Approximate Theory of Order in Alloys", *Phys. Rev.* 77 (5), p. 669. DOI: 10.1103/PhysRev.77.669.

- D'Souza, R. and S. Mukherjee (2015) "Electronic structure, phase stability and resistivity of hybrid hexagonal Cx(BN)1-x two-dimensional nanomaterial: A first-principles study", *Physica E: Low-dimensional Systems and Nanostructures* 69, p. 138. DOI: <https://doi.org/10.1016/j.physe.2015.01.026>.
- Dethan, J. F. N. and V. Swamy (2020) "Tensile properties of hydrogenated hybrid graphene – hexagonal boron nitride nanosheets: a reactive force field study", *Molecular Simulation* 46 (15), p. 1220. DOI: 10.1080/08927022.2020.1810854.
- DiBenedetto, A. C. (2017) "The electronic properties of hexagonal boron nitride and graphene nanoribbons". MA thesis. Ball State University, Muncie, Indiana, USA.
- Du, S. *et al.* (Feb. 2020) "Giant enhancement of solid solubility in monolayer alloys by selective orbital coupling", *Phys. Rev. B* 101 (5), p. 054201. DOI: 10.1103/PhysRevB.101.054201.
- Du, Y., P. Ying, and J. Zhang (2021) "Prediction and optimization of the thermal transport in hybrid carbon-boron nitride honeycombs using machine learning", *Carbon* 184, p. 492. DOI: <https://doi.org/10.1016/j.carbon.2021.08.035>.
- Duan, J., S. Chen, M. Jaroniec, and S. Z. Qiao (2015) "Heteroatom-Doped Graphene-Based Materials for Energy-Relevant Electrocatalytic Processes", *ACS Catalysis* 5 (9), p. 5207. DOI: 10.1021/acscatal.5b00991.
- Ducastelle, F. (1991) *Order and Phase Stability in Alloys (Cohesion and Structure)*. North Holland.
- Duin, A. C. T. van, S. Dasgupta, F. Lorant, and W. A. Goddard (2001) "ReaxFF: A Reactive Force Field for Hydrocarbons", *The Journal of Physical Chemistry A* 105 (41), p. 9396. DOI: 10.1021/jp004368u.
- Elias, C. *et al.* (2019) "Direct band-gap crossover in epitaxial monolayer boron nitride", *Nature Communications* 10 (1), p. 2639. DOI: 10.1038/s41467-019-10610-5.
- Elstner, M. *et al.* (Sept. 1998) "Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties", *Phys. Rev. B* 58 (11), p. 7260. DOI: 10.1103/PhysRevB.58.7260.
- Feng, L. *et al.* (2020) "Boron-, sulfur-, and phosphorus-doped graphene for environmental applications", *Science of The Total Environment* 698, p. 134239. DOI: <https://doi.org/10.1016/j.scitotenv.2019.134239>.
- Filippi, C. and C. J. Umrigar (1996) "Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules", *The Journal of Chemical Physics* 105 (1), p. 213. DOI: 10.1063/1.471865.
- Finzel, K. (2020) "Equilibrium Bond Lengths from Orbital-Free Density Functional Theory", *Molecules* 25 (8). DOI: 10.3390/molecules25081771.
- Fonseca, A. F., T. Liang, D. Zhang, K. Choudhary, S. R. Phillpot, *et al.* (2017) "Graphene – Titanium Interfaces from Molecular Dynamics Simulations", *ACS Applied Materials & Interfaces* 9 (38), p. 33288. DOI: 10.1021/acsami.7b09469.

- Fonseca, A. F., T. Liang, D. Zhang, K. Choudhary, and S. B. Sinnott (2016) "Probing the accuracy of reactive and non-reactive force fields to describe physical and chemical properties of graphene-oxide", *Computational Materials Science* 114, p. 236. DOI: <https://doi.org/10.1016/j.commatsci.2015.12.030>.
- Ghaderzadeh, S. *et al.* (2021) "Atomistic Simulations of Defect Production in Monolayer and Bulk Hexagonal Boron Nitride under Low- and High-Fluence Ion Irradiation", *Nanomaterials* 11 (5). DOI: 10.3390/nano11051214.
- Giannozzi, P. *et al.* (Sept. 2009) "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials", *Journal of Physics: Condensed Matter* 21 (39), p. 395502. DOI: 10.1088/0953-8984/21/39/395502.
- Giraud, P. (2012) "Study of the Electronic Structure of hexagonal Boron Nitride on Metals Substrates". MA thesis. Universite des Sciences et Technologies, San Sebastian, Spain.
- Gong, Y. *et al.* (2014) "Direct chemical conversion of graphene to boron- and nitrogen- and carbon-containing atomic layers", *Nature Communications* 5 (1), p. 3193. DOI: 10.1038/ncomms4193.
- Gopalakrishnan, K., K. Moses, A. Govindaraj, and C. Rao (2013) "Supercapacitors based on nitrogen-doped reduced graphene oxide and borocarbonitrides", *Solid State Communications* 175-176 (Complete), p. 43. DOI: 10.1016/j.ssc.2013.02.005.
- Goringe, C. M., D. R. Bowler, and E. R. Hernández (1997) "Tight-Binding Modelling of Materials", *Reports on Progress in Physics* 60, p. 1447. DOI: 10.1088/0034-4885/60/12/001.
- Guilhon, I., M. Marques, L. K. Teles, and F. Bechstedt (Jan. 2017) "Optical absorbance and band-gap engineering of $(\text{BN})_{1-x}(\text{C}_2)_x$ two-dimensional alloys: Phase separation and composition fluctuation effects", *Phys. Rev. B* 95 (3), p. 035407. DOI: 10.1103/PhysRevB.95.035407.
- Hamann, D. R., M. Schlüter, and C. Chiang (Nov. 1979) "Norm-Conserving Pseudopotentials", *Phys. Rev. Lett.* 43 (20), p. 1494. DOI: 10.1103/PhysRevLett.43.1494.
- Han, G. H. *et al.* (2013) "Continuous Growth of Hexagonal Graphene and Boron Nitride In-Plane Heterostructures by Atmospheric Pressure Chemical Vapor Deposition", *ACS Nano* 7 (11), p. 10129. DOI: 10.1021/nn404331f.
- Harrison, W. A. (1989) *Electronic Structure and the Properties of Solids: The physics of the Chemical Bond*. New York: Dover Publications.
- Hastings, W. K. (Apr. 1970) "Monte Carlo sampling methods using Markov chains and their applications", *Biometrika* 57 (1), p. 97. DOI: 10.1093/biomet/57.1.97.
- Herder, L. M., J. M. Bray, and W. F. Schneider (2015) "Comparison of cluster expansion fitting algorithms for interactions at surfaces", *Surface Science* 640, p. 104. DOI: <https://doi.org/10.1016/j.susc.2015.02.017>.
- Hestenes, M. R. and E. Stiefel (1952) "Methods of conjugate gradients for solving linear systems", *J Res NIST* 49 (6), p. 409. DOI: 10.6028/jres.049.044.
- Hohenberg, P. and W. Kohn (Nov. 1964) "Inhomogeneous Electron Gas", *Phys. Rev.* 136 (3B), B864. DOI: 10.1103/PhysRev.136.B864.

- Huang, B. and H. Lee (Dec. 2012) "Defect and impurity properties of hexagonal boron nitride: A first-principles calculation", *Phys. Rev. B* 86 (24), p. 245406. DOI: 10.1103/PhysRevB.86.245406.
- Huang, C. *et al.* (July 2015) "Carbon-doped BN nanosheets for metal-free photoredox catalysis", *Nature communications* 6, p. 7698. DOI: 10.1038/ncomms8698.
- Huber, K. (1979) *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules*. New York: Springer US. DOI: 10.1007/978-1-4757-0961-2.
- Hunt, R. J., B. Monserrat, V. Zolyomi, and N. D. Drummond (May 2020) "Diffusion quantum Monte Carlo and GW study of the electronic properties of monolayer and bulk hexagonal boron nitride", *Phys. Rev. B* 101 (20), p. 205115. DOI: 10.1103/PhysRevB.101.205115.
- Jam, A. N., R. Abadi, M. Izadifar, and T. Rabczuk (2018) "Molecular dynamics study on the mechanical properties of carbon doped single - layer polycrystalline boron - nitride nanosheets", *Computational Materials Science* 153, p. 16. DOI: <https://doi.org/10.1016/j.commatsci.2018.06.011>.
- Jemmis, E. D., B. Pathak, R. B. King, and H. F. Schaefer III (2006) "Bond length and bond multiplicity: σ -bond prevents short π -bonds", *Chem. Commun.* (20), p. 2164. DOI: 10.1039/B602116F.
- Jensen, B. D., K. E. Wise, and G. M. Odegard (2015) "Simulation of the Elastic and Ultimate Tensile Properties of Diamond, Graphene, Carbon Nanotubes, and Amorphous Carbon Using a Revised ReaxFF Parametrization", *The Journal of Physical Chemistry A* 119 (37), p. 9710. DOI: 10.1021/acs.jpca.5b05889.
- Ji, Y. *et al.* (2016) "Heptazine-based graphitic carbon nitride as an effective hydrogen purification membrane", *RSC Adv.* 6 (57), p. 52377. DOI: 10.1039/C6RA06425F.
- Jin, C., F. Lin, K. Suenaga, and S. Iijima (May 2009) "Fabrication of a Freestanding Boron Nitride Single Layer and Its Defect Assignments", *Phys. Rev. Lett.* 102 (19), p. 195505. DOI: 10.1103/PhysRevLett.102.195505.
- Jin, J. *et al.* (Apr. 2014) "Catalyst-Free Synthesis of Crumpled Boron and Nitrogen Co-Doped Graphite Layers with Tunable Bond Structure for Oxygen Reduction Reaction", *ACS Nano* 8 (4), p. 3313.
- Kaur, M., M. Kaur, and V. K. Sharma (2018) "Nitrogen-doped graphene and graphene quantum dots: A review on synthesis and applications in energy, sensors and environment", *Advances in Colloid and Interface Science* 259, p. 44. DOI: <https://doi.org/10.1016/j.cis.2018.07.001>.
- Kaye and Laby (2021) www.kayelaby.npl.co.uk. URL: https://web.archive.org/web/20190502165701/http://www.kayelaby.npl.co.uk/chemistry/3_2/3_2.html (visited on 03/26/2021).
- Kinaci, A., J. B. Haskins, C. Sevik, and T. Ça ğ in (Sept. 2012) "Thermal conductivity of BN-C nanostructures", *Phys. Rev. B* 86 (11), p. 115410. DOI: 10.1103/PhysRevB.86.115410.
- Kohn, W. and L. J. Sham (Nov. 1965) "Self-Consistent Equations Including Exchange and Correlation Effects", *Phys. Rev.* 140 (4A), A1133. DOI: 10.1103/PhysRev.140.A1133.

- Kolodzeiski, E. and S. Amirjalayer (2021) "On-the-Fly Training of Atomistic Potentials for Flexible and Mechanically Interlocked Molecules", *Journal of Chemical Theory and Computation* 17 (11), p. 7010. DOI: 10.1021/acs.jctc.1c00497.
- Koski, K. J. and Y. Cui (2013) "The New Skinny in Two-Dimensional Nanomaterials", *ACS Nano* 7 (5), p. 3739. DOI: 10.1021/nm4022422.
- Krivanek, O. L. *et al.* (2010) "Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy", *Nature* 464 (7288), p. 571. DOI: 10.1038/nature08879.
- Kroll, P. M. (1996). PhD thesis. Technische Hochschule Darmstadt.
- Kumar, N. *et al.* (2013) "Borocarbonitrides, $B_xC_yN_z$ ", *J. Mater. Chem. A* 1 (19), p. 5806. DOI: 10.1039/C3TA01345F.
- Kumar, R. *et al.* (2021) "Heteroatom doping of 2D graphene materials for electromagnetic interference shielding: a review of recent progress", *Critical Reviews in Solid State and Materials Sciences* 0 (0), p. 1. DOI: 10.1080/10408436.2021.1965954.
- Laks, D. B., L. G. Ferreira, S. Froyen, and A. Zunger (Nov. 1992) "Efficient cluster expansion for substitutional systems", *Phys. Rev. B* 46 (19), p. 12587. DOI: 10.1103/PhysRevB.46.12587.
- Leong, Z. and T. L. Tan (2019) "Robust cluster expansion of multicomponent systems using structured sparsity", *Phys. Rev. B* 100 (13), p. 134108. DOI: 10.1103/PhysRevB.100.134108.
- Li, D. *et al.* (2020) "2D Boron Sheets: Structure, Growth, and Electronic and Thermal Transport Properties", *Advanced Functional Materials* 30 (8), p. 1904349. DOI: <https://doi.org/10.1002/adfm.201904349>.
- Li, Q. *et al.* (2021) "Localized Wannier function based tight-binding models for two - dimensional allotropes of bismuth", *New Journal of Physics* 23 (6), p. 063042. DOI: 10.1088/1367-2630/ac04c9.
- Li, X.-H. and M. Antonietti (2013) "Polycondensation of Boron- and Nitrogen-Codoped Holey Graphene Monoliths from Molecules: Carbocatalysts for Selective Oxidation", *Angewandte Chemie International Edition* 52 (17), p. 4572. DOI: <https://doi.org/10.1002/anie.201209320>.
- Liang, T., B. Devine, S. R. Phillpot, and S. B. Sinnott (2012) "Variable Charge Reactive Potential for Hydrocarbons to Simulate Organic-Copper Interactions", *The Journal of Physical Chemistry A* 116 (30), p. 7976. DOI: 10.1021/jp212083t.
- Liang, T., T.-R. Shan, *et al.* (2013) "Classical atomistic simulations of surfaces and heterogeneous interfaces with the charge-optimized many body (COMB) potentials", *Materials Science and Engineering: R: Reports* 74 (9), p. 255. DOI: <https://doi.org/10.1016/j.mser.2013.07.001>.
- Lima, J. S. *et al.* (2021) "Mechanical and electronic properties of boron nitride nanosheets with graphene domains under strain", *RSC Adv.* 11 (56), p. 35127. DOI: 10.1039/D1RA05831B.

- Lin, P.-C. *et al.* (2017) "Nano-sized graphene flakes: insights from experimental synthesis and first principles calculations", *Phys. Chem. Chem. Phys.* 19 (9), p. 6338. DOI: 10.1039/C6CP08354D.
- Lin, T.-W. *et al.* (2012) "Converting Graphene Oxide Monolayers into Boron Carbonitride Nanosheets by Substitutional Doping", *Small* 8 (9), p. 1384. DOI: <https://doi.org/10.1002/sml1.201101927>.
- Lindsay, L. and D. A. Broido (2010) "Optimized Tersoff and Brenner empirical potential parameters for lattice dynamics and phonon thermal transport in carbon nanotubes and graphene", *Phys. Rev. B* 81 (20), p. 205441. DOI: 10.1103/PhysRevB.81.205441.
- Liu, Y. *et al.* (2016) "Elemental superdoping of graphene and carbon nanotubes", *Nature Communications* 7 (1), p. 10921. DOI: 10.1038/ncomms10921.
- Liu, Z. *et al.* (2013) "Carbon Doping of Hexagonal Boron Nitride by Using CO Molecules", *The Journal of Physical Chemistry C* 117 (18), p. 9332. DOI: 10.1021/jp402297n.
- Los, J. H. *et al.* (Nov. 2017) "Extended Tersoff potential for boron nitride: Energetics and elastic properties of pristine and defective *h*-BN", *Phys. Rev. B* 96 (18), p. 184108. DOI: 10.1103/PhysRevB.96.184108.
- Lu, J. *et al.* (Oct. 2013) "Order-disorder transition in a two-dimensional boron-carbon-nitride alloy", *Nature communications* 4, p. 2681. DOI: 10.1038/ncomms3681.
- Lv, R. *et al.* (Aug. 2012) "Nitrogen-doped graphene: Beyond single substitution and enhanced molecular sensing", *Scientific reports* 2, p. 586. DOI: 10.1038/srep00586.
- Lynch, R. W. and H. G. Drickamer (1966) "Effect of High Pressure on the Lattice Parameters of Diamond, Graphite, and Hexagonal Boron Nitride", *The Journal of Chemical Physics* 44 (1), p. 181. DOI: 10.1063/1.1726442.
- M. Dieb, T., Z. Hou, and K. Tsuda (2018) "Structure prediction of boron-doped graphene by machine learning", *The Journal of Chemical Physics* 148 (24), p. 241716. DOI: 10.1063/1.5018065.
- Ma, Y., P. O. Lehtinen, A. S. Foster, and R. M. Nieminen (2004) "Magnetic properties of vacancies in graphene and single-walled carbon nanotubes", *New Journal of Physics* 6, p. 68. DOI: 10.1088/1367-2630/6/1/068.
- Mardirossian, N. and M. Head-Gordon (2017) "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals", *Molecular Physics* 115 (19), p. 2315. DOI: 10.1080/00268976.2017.1333644.
- Martinez, J., T. Liang, S. B. Sinnott, and S. R. Phillpot (2017) "A third-generation charge optimized many body (COMB3) potential for nitrogen-containing organic molecules", *Computational Materials Science* 139, p. 153. DOI: <https://doi.org/10.1016/j.commatsci.2017.07.019>.
- Marzari, N. *et al.* (Oct. 2012) "Maximally localized Wannier functions: Theory and applications", *Rev. Mod. Phys.* 84 (4), p. 1419. DOI: 10.1103/RevModPhys.84.1419.

- Matsunaga, K., C. Fisher, and H. Matsubara (2000) "Tersoff Potential Parameters for Simulating Cubic Boron Carbonitrides", *Japanese Journal of Applied Physics* 39 (Part 2, No. 1A/B), p. L48. DOI: 10.1143/jjap.39.L48.
- Mayer, J. E. and E. Montroll (1941) "Molecular Distribution", *The Journal of Chemical Physics* 9 (1), p. 2. DOI: 10.1063/1.1750822.
- McDougall, N. L. *et al.* (Oct. 2017) "Influence of point defects on the near edge structure of hexagonal boron nitride", *Phys. Rev. B* 96 (14), p. 144106. DOI: 10.1103/PhysRevB.96.144106.
- Mendelson, N. *et al.* (Mar. 2021) "Identifying carbon as the source of visible single-photon emission from hexagonal boron nitride", *Nature Materials* 20, p. 1. DOI: 10.1038/s41563-020-00850-y.
- Metropolis, N. *et al.* (1953) "Equation of State Calculations by Fast Computing Machines", *The Journal of Chemical Physics* 21 (6), p. 1087. DOI: 10.1063/1.1699114.
- Meyer, J. C. *et al.* (2009) "Selective Sputtering and Atomic Resolution Imaging of Atomically Thin Boron Nitride Membranes", *Nano Letters* 9 (7), p. 2683. DOI: 10.1021/nl9011497.
- Miranda, Y., R. de Paiva, and S. Azevedo (2021) "Electronic properties and stability of 4–8 BxCyNz monolayers", *Solid State Communications* 326, p. 114174. DOI: <https://doi.org/10.1016/j.ssc.2020.114174>.
- Mohri, T. (2007) "Statistical Thermodynamics and Model Calculations". *Alloy Physics*. John Wiley and Sons, Ltd. Chap. 10, p. 525. DOI: <https://doi.org/10.1002/9783527614196.ch10>.
- Mortier, W. J., S. K. Ghosh, and S. Shankar (1986) "Electronegativity-equalization method for the calculation of atomic charges in molecules", *Journal of the American Chemical Society* 108 (15), p. 4315. DOI: 10.1021/ja00275a013.
- Mueller, T. and G. Ceder (July 2009) "Bayesian approach to cluster expansions", *Phys. Rev. B* 80 (2), p. 024103. DOI: 10.1103/PhysRevB.80.024103.
- Nagashima, A. *et al.* (Nov. 1995) "Electronic Structure of Monolayer Hexagonal Boron Nitride Physisorbed on Metal Surfaces", *Phys. Rev. Lett.* 75 (21), p. 3918. DOI: 10.1103/PhysRevLett.75.3918.
- Nakhaee, M., S. A. Ketabi, and F. M. Peeters (2020) "Tight-Binding Studio: A technical software package to find the parameters of tight-binding Hamiltonian", *Computer Physics Communications* 254, p. 107379. DOI: <https://doi.org/10.1016/j.cpc.2020.107379>.
- Nelson, L. J., G. L. W. Hart, F. Zhou, and V. Ozoliņš (Jan. 2013) "Compressive sensing as a paradigm for building physics models", *Phys. Rev. B* 87 (3), p. 035125. DOI: 10.1103/PhysRevB.87.035125.
- Norris, J. R. (1997) "Continuous-time Markov chains I". *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, p. 60. DOI: 10.1017/CB09780511810633.004.
- Novoselov, K. S. *et al.* (2004) "Electric Field Effect in Atomically Thin Carbon Films", *Science* 306 (5696), p. 666. DOI: 10.1126/science.1102896.

- Ozoliņš, V., C. Wolverton, and A. Zunger (Mar. 1998) "Cu-Au, Ag-Au, Cu-Ag, and Ni-Au intermetallics: First-principles study of temperature-composition phase diagrams and structures", *Phys. Rev. B* 57 (11), p. 6427. DOI: 10.1103/PhysRevB.57.6427.
- Pai, S. J., B. C. Yeo, and S. S. Han (2016) "Development of the ReaxFFCBN reactive force field for the improved design of liquid CBN hydrogen storage materials", *Phys. Chem. Chem. Phys.* 18 (3), p. 1818. DOI: 10.1039/C5CP05486A.
- Panchakarla, L. S. *et al.* (2009) "Synthesis, Structure, and Properties of Boron- and Nitrogen-Doped Graphene", *Advanced Materials* 21 (46), p. 4726. DOI: 10.1002/adma.200901285.
- Papaconstantopoulos, D. and M. Mehl (Mar. 2003) "The Slater-Koster tight-binding method: A computationally efficient and accurate approach", *Journal of Physics: Condensed Matter* 15, R413. DOI: 10.1088/0953-8984/15/10/201.
- Paupitz, R., C. E. Junkermeier, A. C. T. van Duin, and P. S. Branicio (2014) "Fullerenes generated from porous structures", *Phys. Chem. Chem. Phys.* 16 (46), p. 25515. DOI: 10.1039/C4CP03529A.
- Pedersen, T. G. and J. G. Pedersen (Apr. 2013) "Self-consistent tight-binding model of B and N doping in graphene", *Phys. Rev. B* 87 (15), p. 155433. DOI: 10.1103/PhysRevB.87.155433.
- Perdew, J. P., K. Burke, and M. Ernzerhof (Oct. 1996) "Generalized Gradient Approximation Made Simple", *Phys. Rev. Lett.* 77 (18), p. 3865. DOI: 10.1103/PhysRevLett.77.3865.
- Porezag, D. *et al.* (1995) "Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon", *Phys. Rev. B* 51 (19), p. 12947. DOI: 10.1103/PhysRevB.51.12947.
- Pozzo, M. *et al.* (Mar. 2011) "Thermal Expansion of Supported and Freestanding Graphene: Lattice Constant versus Interatomic Distance", *Phys. Rev. Lett.* 106 (13), p. 135501. DOI: 10.1103/PhysRevLett.106.135501.
- Qian, C., B. McLean, D. Hedman, and F. Ding (2021) "A comprehensive assessment of empirical potentials for carbon materials", *APL Materials* 9 (6), p. 061102. DOI: 10.1063/5.0052870.
- Rani, P. and V. K. Jindal (2013) "Designing band gap of graphene by B and N dopant atoms", *RSC Adv.* 3 (3), p. 802. DOI: 10.1039/C2RA22664B.
- (2014) "Stability and electronic properties of isomers of B/N co-doped graphene", *Applied Nanoscience* 4 (8), p. 989. DOI: 10.1007/s13204-013-0280-3.
- Rao, C. N. R. and K. Gopalakrishnan (2017) "Borocarbonitrides, BxCyNz: Synthesis, Characterization, and Properties with Potential Applications", *ACS Applied Materials & Interfaces* 9 (23), p. 19478. DOI: 10.1021/acsami.6b08401.
- Rao, C., K. Gopalakrishnan, and A. Govindaraj (2014) "Synthesis, properties and applications of graphene doped with boron, nitrogen and other elements", *Nano Today* 9 (3), p. 324. DOI: <https://doi.org/10.1016/j.nantod.2014.04.010>.
- Rappe, A. K. and W. A. Goddard (1991) "Charge equilibration for molecular dynamics simulations", *The Journal of Physical Chemistry* 95 (8), p. 3358. DOI: 10.1021/j100161a070.

- Rigamonti, S. *et al.* (n.d.) "CELL: a python package for cluster expansion with a focus on complex alloys". In preparation.
- Robertson, A. W. *et al.* (2013) "Structural Reconstruction of the Graphene Monovacancy", *ACS Nano* 7 (5), p. 4495. DOI: 10.1021/nn401113r.
- Rodrigo, L., P. Pou, and R. Pérez (2016) "Graphene monovacancies: Electronic and mechanical properties from large scale ab initio simulations", *Carbon* 103, p. 200. DOI: <https://doi.org/10.1016/j.carbon.2016.02.064>.
- Russo, M. F. and A. C. van Duin (2011) "Atomistic-scale simulations of chemical reactions: Bridging from quantum chemistry to engineering", *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 269 (14). Computer Simulations of Radiation Effects in Solids, p. 1549. DOI: <https://doi.org/10.1016/j.nimb.2010.12.053>.
- Ryou, J., J. Park, and S. Hong (2017) "Investigations of Vacancy Structures Related to Their Growth in h-BN Sheet", *Nanoscale Research Letters* 12 (1), p. 445.
- Sanchez, J., F. Ducastelle, and D. Gratias (1984) "Generalized cluster description of multi-component systems", *Physica A: Statistical Mechanics and its Applications* 128 (1), p. 334. DOI: [https://doi.org/10.1016/0378-4371\(84\)90096-7](https://doi.org/10.1016/0378-4371(84)90096-7).
- Schiro, T. *et al.* (2016) "Atomistic Interrogation of B-N Co-dopant Structures and Their Electronic Effects in Graphene", *ACS Nano* 10 (7), p. 6574. DOI: 10.1021/acsnano.6b01318.
- Seifert, G., D. Porezag, and T. Frauenheim (1996) "Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme", *International Journal of Quantum Chemistry* 58 (2), p. 185. DOI: [https://doi.org/10.1002/\(SICI\)1097-461X\(1996\)58:2<185::AID-QUA7>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-461X(1996)58:2<185::AID-QUA7>3.0.CO;2-U).
- Senftle, T. P. *et al.* (2016) "The ReaxFF reactive force-field: development, applications and future directions", *npj Computational Materials* 2 (1), p. 15011. DOI: 10.1038/npjcompumats.2015.11.
- Sengul, M. Y. *et al.* (2021) "INDEEDopt: a deep learning-based ReaxFF parameterization framework", *npj Computational Materials* 7 (1), p. 68.
- Senturk, A., A. Oktem, and A. Konukman (Jan. 2019) "An investigation on the thermomechanical properties of boron-doped g-C₃N₄", *Applied Physics A* 125. DOI: 10.1007/s00339-018-2355-1.
- Sharma, B. B. and A. Parashar (2020) "Mechanical and fracture behaviour of hydroxyl functionalized h-BN nanosheets", *Journal of Materials Science* 55 (8), p. 3228. DOI: 10.1007/s10853-019-04163-7.
- Sharma, D. K. *et al.* (2017) "Defect concentration in nitrogen-doped graphene grown on Cu substrate: A thickness effect", *Physica B: Condensed Matter* 513, p. 62. DOI: <https://doi.org/10.1016/j.physb.2017.03.004>.
- Shi, Z., A. Kutana, and B. I. Yakobson (2015) "How Much N-Doping Can Graphene Sustain?", *The Journal of Physical Chemistry Letters* 6 (1), p. 106. DOI: 10.1021/jz502093c.

- Shin, H. *et al.* (2014) "Cohesion energetics of carbon allotropes: Quantum Monte Carlo study", *The Journal of Chemical Physics* 140 (11), p. 114702. DOI: 10.1063/1.4867544.
- Shin, Y. K. *et al.* (2012) "Variable charge many-body interatomic potentials", *MRS Bulletin* 37 (5), p. 504. DOI: 10.1557/mrs.2012.95.
- Skowron, S. T., I. V. Lebedeva, A. M. Popov, and E. Bichoutskaia (2015) "Energetics of atomic scale structure changes in graphene", *Chem. Soc. Rev.* 44 (10), p. 3143. DOI: 10.1039/C4CS00499J.
- Slater, J. C. and G. F. Koster (1954) "Simplified LCAO Method for the Periodic Potential Problem", *Phys. Rev.* 94 (6), p. 1498. DOI: 10.1103/PhysRev.94.1498.
- Sobieraj, D. *et al.* (2020) "Chemical short-range order in derivative Cr-Ta-Ti-V-W high entropy alloys from the first-principles thermodynamic study", *Phys. Chem. Chem. Phys.* 22 (41), p. 23929. DOI: 10.1039/D0CP03764H.
- Song, L. *et al.* (Sept. 2012) "Binary and ternary atomic layers built from carbon, boron, and nitrogen", English. *Advanced Materials*, p. 4878. DOI: 10.1002/adma.201201792.
- Spiegelman, F. *et al.* (2020) "Density-functional tight-binding: basic concepts and applications to molecules and clusters", *Advances in Physics: X* 5 (1), p. 1710252. DOI: 10.1080/23746149.2019.1710252.
- Srinivasan, S. G., A. C. T. van Duin, and P. Ganesh (2015) "Development of a ReaxFF Potential for Carbon Condensed Phases and Its Application to the Thermal Fragmentation of a Large Fullerene", *The Journal of Physical Chemistry A* 119 (4), p. 571. DOI: 10.1021/jp510274e.
- Stern, H. A. *et al.* (1999) "Fluctuating Charge, Polarizable Dipole, and Combined Models: Parameterization from ab Initio Quantum Chemistry", *The Journal of Physical Chemistry B* 103 (22), p. 4730. DOI: 10.1021/jp984498r.
- Streitz, F. H. and J. W. Mintmire (Oct. 1994) "Electrostatic potentials for metal-oxide surfaces and interfaces", *Phys. Rev. B* 50 (16), p. 11996. DOI: 10.1103/PhysRevB.50.11996.
- Sutter, P., J. Lahiri, P. Albrecht, and E. Sutter (2011) "Chemical Vapor Deposition and Etching of High-Quality Monolayer Hexagonal Boron Nitride Films", *ACS Nano* 5 (9), p. 7303. DOI: 10.1021/nn202141k.
- Tang, Y.-B. *et al.* (Mar. 2012) "Tunable Band Gaps and p-Type Transport Properties of Boron-Doped Graphenes by Controllable Ion Doping Using Reactive Microwave Plasma", *ACS Nano* 6 (3), p. 1970.
- Terrones, H., R. Lv, M. Terrones, and M. Dresselhaus (June 2012) "The role of defects and doping in 2D graphene sheets and 1D nanoribbons", *Reports on progress in physics. Physical Society (Great Britain)* 75, p. 062501. DOI: 10.1088/0034-4885/75/6/062501.
- Tersoff, J. (Nov. 1988) "Empirical interatomic potential for silicon with improved elastic properties", *Phys. Rev. B* 38 (14), p. 9902. DOI: 10.1103/PhysRevB.38.9902.
- (Mar. 1989) "Modeling solid-state chemistry: Interatomic potentials for multicomponent systems", *Phys. Rev. B* 39 (8), p. 5566. DOI: 10.1103/PhysRevB.39.5566.

- Thiemann, F. L., P. Rowe, E. A. Müller, and A. Michaelides (Oct. 2020) "Machine Learning Potential for Hexagonal Boron Nitride Applied to Thermally and Mechanically Induced Rippling", *The Journal of Physical Chemistry C* 124 (40), p. 22278.
- Thompson, A. P. *et al.* (2022) "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales", *Computer Physics Communications* 271, p. 108171. DOI: <https://doi.org/10.1016/j.cpc.2021.108171>.
- Thrower, P. A. and R. M. Mayer (1978) "Point defects and self-diffusion in graphite", *physica status solidi (a)* 47 (1), p. 11. DOI: <https://doi.org/10.1002/pssa.2210470102>.
- Toufar, H. *et al.* (1996) "Dependence of the Hardness of Atoms in Molecules on the Local Environment: An ab Initio Study", *The Journal of Physical Chemistry* 100 (38), p. 15383. DOI: 10.1021/jp960693z.
- Troppenz, M., S. Rigamonti, and C. Draxl (2017) "Predicting Ground-State Configurations and Electronic Properties of the Thermoelectric Clathrates $\text{Ba}_8\text{Al}_x\text{Si}_{46-x}$ and $\text{Sr}_8\text{Al}_x\text{Si}_{46-x}$ ", *Chemistry of Materials* 29 (6), p. 2414. DOI: 10.1021/acs.chemmater.6b05027.
- Uddin, M. R., S. Majety, *et al.* (2014) "Layer-structured hexagonal (BN)C semiconductor alloys with tunable optical and electrical properties", *Journal of Applied Physics* 115 (9), p. 093509. DOI: 10.1063/1.4867641.
- Uddin, M. R., T. Doan, *et al.* (Aug. 2014) "Electrical transport properties of (BN)-rich hexagonal (BN)C semiconductor alloys", *AIP Advances* 4, p. 087141. DOI: 10.1063/1.4894451.
- Usachov, D. Y. *et al.* (2016) "Large-Scale Sublattice Asymmetry in Pure and Boron-Doped Graphene", *Nano Letters* 16 (7), p. 4535. DOI: 10.1021/acs.nanolett.6b01795.
- van de Walle, A. (2009) "Multicomponent multisublattice alloys, nonconfigurational entropy and other additions to the Alloy Theoretic Automated Toolkit", *Calphad* 33 (2). Tools for Computational Thermodynamics, p. 266. DOI: <https://doi.org/10.1016/j.calphad.2008.12.005>.
- Vanderbilt, D. (Apr. 1990) "Soft self-consistent pseudopotentials in a generalized eigenvalue formalism", *Phys. Rev. B* 41 (11), p. 7892. DOI: 10.1103/PhysRevB.41.7892.
- Verhaegen, G., F. E. Stafford, and J. Drowart (1964) "Mass Spectrometric Study of the Systems Boron—Carbon and Boron—Carbon—Silicon", *The Journal of Chemical Physics* 40 (6), p. 1622. DOI: 10.1063/1.1725370.
- Vesel, A., R. Zaplotnik, G. Primc, and M. Mozeti (2020) "A Review of Strategies for the Synthesis of N-Doped Graphene-Like Materials", *Nanomaterials* 10.
- Wakabayashi, K. (2012) "Electronic Properties of Graphene Nanoribbons". *Graphene Nanoelectronics: Metrology, Synthesis, Properties and Applications*. Ed. by H. Raza. Berlin, Heidelberg: Springer Berlin Heidelberg, p. 277. DOI: 10.1007/978-3-642-22984-8_9.
- Walle, A. van de and G. Ceder (Jan. 2002) "The effect of lattice vibrations on substitutional alloy thermodynamics", *Rev. Mod. Phys.* 74 (1), p. 11. DOI: 10.1103/RevModPhys.74.11.
- Wang, S. *et al.* (2018) "Band gap-Tunable Porous Borocarbonitride Nanosheets for High Energy-Density Supercapacitors", *ACS Applied Materials & Interfaces* 10 (23), p. 19588. DOI: 10.1021/acsami.8b02317.

- Wang, V., N. Ma, H. Mizuseki, and Y. Kawazoe (2012) "First-principles study of intrinsic defect properties in hexagonal BN bilayer and monolayer", *Solid State Communications* 152 (9), p. 816. DOI: <https://doi.org/10.1016/j.ssc.2012.01.034>.
- Wang, B.-Y. *et al.* (2016) "Nonlinear bandgap opening behavior of BN co-doped graphene", *Carbon* 107, p. 857. DOI: <https://doi.org/10.1016/j.carbon.2016.06.091>.
- Wei, S.-H., S. B. Zhang, and A. Zunger (2000) "First-principles calculation of band offsets, optical bowings, and defects in CdS, CdSe, CdTe, and their alloys", *Journal of Applied Physics* 87 (3), p. 1304. DOI: 10.1063/1.372014.
- Wei, X., M.-S. Wang, Y. Bando, and D. Golberg (2011) "Electron-Beam-Induced Substitutional Carbon Doping of Boron Nitride Nanosheets, Nanoribbons, and Nanotubes", *ACS Nano* 5 (4), p. 2916. DOI: 10.1021/nn103548r.
- Weston, L. *et al.* (June 2018) "Native point defects and impurities in hexagonal boron nitride", *Phys. Rev. B* 97 (21), p. 214104. DOI: 10.1103/PhysRevB.97.214104.
- Wolf, D., P. Keblinski, S. R. Phillpot, and J. Eggebrecht (1999) "Exact method for the simulation of Coulombic systems by spherically truncated, pairwise r⁻¹ summation", *The Journal of Chemical Physics* 110 (17), p. 8254. DOI: 10.1063/1.478738.
- Wolverton, C. and D. de Fontaine (Apr. 1994) "Cluster expansions of alloy energetics in ternary intermetallics", *Phys. Rev. B* 49 (13), p. 8627. DOI: 10.1103/PhysRevB.49.8627.
- Wu, M., C. Cao, and J. Z. Jiang (Nov. 2010) "Light non-metallic atom (B, N, O and F)-doped graphene: a first-principles study", *Nanotechnology* 21 (50), p. 505202. DOI: 10.1088/0957-4484/21/50/505202.
- Wu, Q. *et al.* (2016) "Cluster expansion method and its application in computational materials science", *Computational Materials Science* 125, p. 243. DOI: <https://doi.org/10.1016/j.commatsci.2016.08.034>.
- Wu, Z.-S. *et al.* (2011) "Doped Graphene Sheets As Anode Materials with Superhigh Rate and Large Capacity for Lithium Ion Batteries", *ACS Nano* 5 (7), p. 5463. DOI: 10.1021/nn2006249.
- Xie, W., T. Yanase, T. Nagahama, and T. Shimada (Jan. 2016) "Carbon-Doped Hexagonal Boron Nitride: Analysis as π -Conjugate Molecules Embedded in Two Dimensional Insulator", *C* 2, p. 2. DOI: 10.3390/c2010002.
- Yang, G., L. Li, W. B. Lee, and M. C. Ng (2018) "Structure of graphene and its disorders: a review", *Science and Technology of Advanced Materials* 19 (1), p. 613. DOI: 10.1080/14686996.2018.1494493.
- Yap, Y. K. (Jan. 2009) *B-C-N Nanotubes and Related Nanostructures*. Berlin: Springer.
- Yin, L.-C., H.-M. Cheng, and R. Saito (Apr. 2010) "Triangle defect states of hexagonal boron nitride atomic layer: Density functional theory calculations", *Phys. Rev. B* 81 (15), p. 153407. DOI: 10.1103/PhysRevB.81.153407.
- Yu, J., S. B. Sinnott, and S. R. Phillpot (Feb. 2007) "Charge optimized many-body potential for the Si/SiO₂ system", *Phys. Rev. B* 75 (8), p. 085311. DOI: 10.1103/PhysRevB.75.085311.

- Yuge, K. (Apr. 2009) "Phase stability of boron carbon nitride in a heterographene structure: A first-principles study", *Phys. Rev. B* 79 (14), p. 144109. DOI: 10.1103/PhysRevB.79.144109.
- (Aug. 2011) "Concentration effects on segregation behavior of Pt-Rh nanoparticles", *Phys. Rev. B* 84 (8), p. 085451. DOI: 10.1103/PhysRevB.84.085451.
- Zabet-Khosousi, A. *et al.* (2014) "Segregation of Sublattice Domains in Nitrogen-Doped Graphene", *Journal of the American Chemical Society* 136 (4), p. 1391. DOI: 10.1021/ja408463g.
- Zhang, D. *et al.* (Nov. 2019) "Dynamics of graphene/Al interfaces using COMB3 potentials", *Phys. Rev. Materials* 3 (11), p. 114002. DOI: 10.1103/PhysRevMaterials.3.114002.
- Zhang, J., B. Tan, *et al.* (2021) "Atomically Thin Hexagonal Boron Nitride and Its Heterostructures", *Advanced Materials* 33 (6), p. 2000769. DOI: <https://doi.org/10.1002/adma.202000769>.
- Zhang, J., C. Zhao, *et al.* (2016) "Tunable electronic properties of graphene through controlling bonding configurations of doped nitrogen atoms", *Scientific Reports* 6 (1), p. 28330.
- Zhang, J., R. Sun, *et al.* (2020) "Point defects in two-dimensional hexagonal boron nitride: A perspective", *Journal of Applied Physics* 128 (10), p. 100902. DOI: 10.1063/5.0021093.
- Zhang, J., W. Xie, *et al.* (2016) "Structural and Electronic Properties of Interfaces in Graphene and Hexagonal Boron Nitride Lateral Heterostructures", *Chemistry of Materials* 28 (14), p. 5022. DOI: 10.1021/acs.chemmater.6b01764.
- Zhao, L. *et al.* (2013) "Local Atomic and Electronic Structure of Boron Chemical Doping in Monolayer Graphene", *Nano Letters* 13 (10), p. 4659. DOI: 10.1021/nl401781d.
- Zhao, S. and J. Xue (Mar. 2013) "Mechanical properties of hybrid graphene and hexagonal boron nitride sheets as revealed by molecular dynamic simulations", *Journal of Physics D: Applied Physics* 46 (13), p. 135303. DOI: 10.1088/0022-3727/46/13/135303.
- Zhao, W. *et al.* (2012) "Production of Nitrogen-Doped Graphene by Low-Energy Nitrogen Implantation", *The Journal of Physical Chemistry C* 116 (8), p. 5062. DOI: 10.1021/jp209927m.
- Zheng, Y. *et al.* (2013) "Two-Step Boron and Nitrogen Doping in Graphene for Enhanced Synergistic Catalysis", *Angewandte Chemie International Edition* 52 (11), p. 3110. DOI: <https://doi.org/10.1002/anie.201209548>.
- Zhou, F., T. Maxisch, and G. Ceder (2006) "Configurational Electronic Entropy and the Phase Diagram of Mixed-Valence Oxides: The Case of Li_xFePO_4 ", *Phys. Rev. Lett.* 97 (15), p. 155704. DOI: 10.1103/PhysRevLett.97.155704.