Aalto University
School of Electrical
Engineering

Master's Programme in Chemical, Biochemical and Materials Engineering

Moire ferroelectricity in twisted hBN with machine learning potentials

Henrik Stenbrink

Master's Thesis
2025

Aalto University
School of Electrical Engineering

Author Henrik Stenbrink

Title Moire ferroelectricity in twisted hBN with machine learning potentials

Degree programme Chemical, Biochemical and Materials Engineering

Major Chemistry

Supervisor Prof. Miguel Caro

Advisors Prof. Jose Lado, Patricia Hernández-León

Date 14 August 2025 **Number of pages** 53 **Language** English

Abstract

This thesis is a study on the properties of twisted bilayer hexagonal boron nitride (hBN) by computational methods. The methods consists of both density functional theory (DFT) and machine learning (ML) potentials. We focus on developing an ML potential to study larger hBN structures that are computationally too expensive to be studied with DFT. We concentrate on relaxation, atomic displacement during relaxation, phonon dispersion relations as well as local density of states (LDOS). First, a functional ML potential was trained for this system, comparing its results with DFT calculations. We also computed phonon dispersion relations, which were used as a secondary performance test. When the performance of the ML potential was satisfactory, we relaxed twisted hBN systems with varying twist angles and calculated both the in-plane and out-of-plane atomic displacements during relaxation. The average and atom-wise interlayer distance was also calculated. The phonon dispersion relations were also used to gain valuable information about the twisted systems. We studied them to gain information about the stability of the system as well as finding evidence for phase transitions. Lastly, we computed LDOS of three twisted systems and an AA stacked structure for comparison. The results indicate that there is some driving force which gave rise to the final atomic positions in the relaxed structure. While the results did not prove conclusively the presence of polarization or ferroelectricity, they support this hypothesis.

Keywords Machine learning, computational chemistry, boron nitride, layered structure

Aalto-yliopisto
Sähkötekniikan
korkeakoulu

Tekijä Henrik Stenbrink

Työn nimi Moire ferroelektrisyys kierteisessä hBN:ssä käyttäen koneoppimispotentiaaleja

Koulutusohjelma Kemian-, Bio- ja Materiaalitekniikan Maisteriohjelma

Pääaine Kemia

Työn valvoja Prof. Miguel Caro

Työn ohjaajat Prof. Jose Lado, M.Sc. Patricia Hernández-León

Päivämäärä 14.8.2025 **Sivumäärä** 53 **Kieli** englanti

Tiivistelmä

Tässä lopputyössä tutkitaan kierteisen kaksikerroksisen boorinitridin ominaisuuksia laskennallisin menetelmin. Käytetyt menetelmät ovat DFT ja koneoppiminen. Me keskitymme kehittämään koneoppimispotentiaalin jotta voimme tutkia suurempia rakenteita, jotka ovat laskennallisesti liian kalliita DFT:lle. Keskeiset tutkittavat asiat ovat relaksaatio, atomien siirtymä relaksaation aikana, fononidispersio relaatiot ja paikallinen DOS. Ensin kehitimme koneoppimispotentiaalin ja vertasimme sen tuloksia, vertailuarvoina olivat DFT:llä lasketut arvot. Laskimme myös fononidispersio relaatiota ja käytimme näitä toissijena testinä koneoppimispotentiaalin suoriutumiselle. Kun potentiaalinen suoriutuminen johti tarpeaksi tarkkoihin tuloksiin, aloitiminne kierteisten rakenteiden varsinaisen tutkimisen. Relaksoimme eri kiertokulmien rakenteita, laskimme näiden rakenteiden atomien sekä tason sisäiset siirtymät että tasosta pois, relaksaation aikana. Lisäksi laskimme tasojen keskimääräisen ja atomikohtaisen välimatkan. Näiden lisäksi vertasimme 3 eri kierteisen rakenteen fononidispersio relaatioita keskenään sekä AA-pinotun ja kierteisen rakenteen eroa LDOS:ssa. Nämä tulokset eivät yksiselitteisesti todista polarisaatiota tai ferroelektrisyyttä, mutta tukevat käsitystä että polarisaatio olisi mahdollinen selitys.

Avainsanat Laskennallinen kemia, DFT, koneoppiminen, boorinitridi, kerrosrakenne

Aalto-universitetet
Högskolan för
elektroteknik

Författare Henrik Stenbrink

Titel Moiré-ferroelektrisitet i tvinnad hBN med hjälp av maskininlärningspotentialer

Utbildningsprogram Kemi-, biokemi- och materialteknik

Huvudämne Kemia

Övervakare Prof. Miguel Caro

Handledare Prof. Jose Lado, M.Sc. Patricia Hernández-León

Datum 14.8.2025 **Sidantal** 53 **Språk** engelska

Sammandrag

I denna avhandling studerar vi egenskaperna av tvinnade bilagrade hexagonala bornitrider (hBN). Detta utförs med hjälp av beräkneliga metoder, både täthetsfunktionalteori (DFT) och maskininlärningspotential. Vi fokuserar oss på att utveckla en maskininlärningspotential för att kunna studera större system som inte är tillgängliga eller kräver för mycket resurser med DFT. Vi fokuserar oss på relaxation, atomernas förflytnning under relaxation, fonondispersionsrelationer och lokal tillståndstäthet (LDOS). Först tränade vi en fungerande maskininlärningspotential for våra system. Prestationen av vår maskininlärningspotential jämfördes med DFT beräknade resultat. Som sekundär prestations test, beräkande vi också fonondispersionsrelationer. När prestationen av maskininlärningspotentialen var tillräckligt bra, relaxerade vi tvinnade hBN system med olika vinklar. För dessa tvinnade system, beräknade vi förskjutning, både i plan och ut ur plan, av atomer under relaxation, distansen mellan lager, både i genomsnitt och per atom. Efter detta studerade vi fonondispersionsrelationer av tvinnade hBN system med 3 olika vinklar och vi jämförde också LDOS av AA-staplad hBN och tvinnad hBN. Resultaten tyder på att det finns en kraft som påverkar atomernas positioner. Detta bevisar inte entydgt att polarisation eller ferroelektrisitet uppstår, men antyder att det kan vara orsaken.

Nyckelord Maskinlärning, beräknelig kemi, bornitrid, skiktad struktur

Preface

I want to thank Professor Jose Lado, Professor Miguel Caro and Patricia Hernández-León for their guidance and assistance.

I also want to thank my partner Outi Ylätalo for keeping me sane and alive.

Otaniemi, 15 August 2025

Henrik Stenbrink

Contents

Symbols and abbreviations

Symbols

$\mathbf{v}$	velocity
δ	Dirac delta function
E	Energy
ψ	Wavefunction
ρ	density of states
r	distance
σ	Gaussian smoothing function
Y	spherical harmonics
g	orthonormal radial function
c	weighting factor
z	position in z-direction
d	distance
n	number of objects
Å	Ångström

Operators

$\frac{\mathrm{d}\omega}{\mathrm{dk}}$	derivative of the frequency ω with respect to the wave number k
$\sum_i$	sum over index i

Abbreviations

DFT	density functional theory
ML	machine learning
DOS	density of states
LDOS	local density of states
hBN	hexagonal boron nitride
GP	Gaussian process
GPR	Gaussian process regression
GAP	Gaussian approximation potential
SOAP	smooth overlap of atomic positions
QUIP	quantum mechanics and interatomic potentials
VASP	Vienna ab initio simulations package
LAMMPS	large-scale atomic/molecular massively parallel simulator
TMD	transition-metal dichalcogenide
IP	in-plane displacement
OOP	out-of-plane displacement
vdW	van der Waals
RMSE	root mean square error
MAE	mean average error
TBG	twisted bilayer graphene
PES	potential energy surface
RAM	read-access memory

1 Introduction

In this thesis, we computationally researched twisted bilayer hexagonal boron nitride (hBN). Certain materials, from a newly discovered subclass called moiré materials, has recently been proven to exhibit properties that they do not show in stacked form [2]. The first of them was graphene, which was shown to exhibit superconductivity in 2018 [6]. This achievement piqued scientists' interest in doing the same with other materials. Some of the researched materials are WS_2 [43], WSe_2 [25],α-MoO_3 [13, 52], and $NbSe_2$ [34]. Additionally, twisted hBN has been shown to exhibit ferroelectricity [51], which is absent in its stacked version. While this property is not new, its impact in today's society makes it valuable and new ways of engineering it are extremely interesting from a scientific standpoint.

Ferroelectric materials, having the option of switching their polarization, can be used in effectively any device with memory, from smartphones to computers, from cars to home electronics. In particular, ferroelectric RAM offers some benefits in comparison to more traditional flash memories, most notably better reading and writing properties in general. On the other hand, its limitations include higher production costs and storage properties. This would most likely limit its use to more specialized applications. Boron nitride has been used for a long time but its structure has been only recently engineered in a way that presents ferroelectric properties. Ferroelectricity in hBN in particular is also interesting due to its general properties as a material. Therefore, it is reasonable to study further the material in question to find out what the controlling parameters are for this property.

This study focuses on understanding the stability and emergent mechanical properties of moiré patterns of hBN. This pattern in turn will lead to changes in its properties, potentially enabling ferroelectricity instead of behaving as a non-ferroelectric insulator. Such transition state is hypothesized to appear in the phonon dispersion relations as an anomaly. In order for a material to have wider-scale real-life applications, it is important to study larger structures. Therefore, the aim for this study was to create a reliable and cost-effective way to relax large structures to first and foremost find if moiré patterns are energetically stable. Once this was accomplished, the resulting structures were studied.

When it comes to simulating and computing large structures with high accuracy, machine learning (ML) is the prime candidate nowadays. While its accuracy highly depends on its implementation and training process, it is a more cost-effective alternative to relying purely on ab initio methods such as density functional theory (DFT). Because of this, ML-based atomistic modeling was deemed to be the best approach for this study and an ML potential was developed. This was done by constructing a training set specifically for this purpose. The training set contains hundreds of structures, whose energies and forces were computed using DFT by means of the Vienna ab initio simulation package (VASP) [28–30]. On the other hand, the ML potential was generated with the quantum mechanics and interatomic potentials (QUIP) software [9]. This ML potential was used to study different stacked and twisted bilayer hBN structures. The main focuses were the relaxation of structures, the atomic displacements during relaxation as well as the computation of the energies and forces associated with the displacements. Relaxation was done with the TurboGAP software [12], which uses the Gaussian approximation potential (GAP) [4, 5], while the atomic displacements and distances were computed using Python scripts developed independently for this research. The ML potential was also used in conjunction with the large-scale atomic/molecular massively parallel simulator (LAMMPS)[42] to study the phonon dispersion relations. On top of these, we used phonopy [45, 47] to compute the local density of states (LDOS), as well as to generate supercells with the displacements needed for the phonon dispersion relations. The LDOS was computed for selected structures to support possible findings.

This thesis is divided in 5 sections, including this introduction. Sect. 2 presents the background information, mainly focusing on boron nitride structures and the ML framework applied in this thesis. Sect. 3 describes the computations in detail, including the generation of an accurate ML potential, relaxation, phonon dispersion relations and LDOS. Sect. 4 includes all the results and discussion, and Sect. 5 includes a summary, conclusions of the results and suggestions for further study.

2 Literature review

In this section we will go over the background information needed to understand the reason for this study. Information includes general information about boron nitride, different structures and some of the interesting properties that are of key interest. Later in this section, machine learning is discussed as well.

2.1 Boron nitride and crystal structure

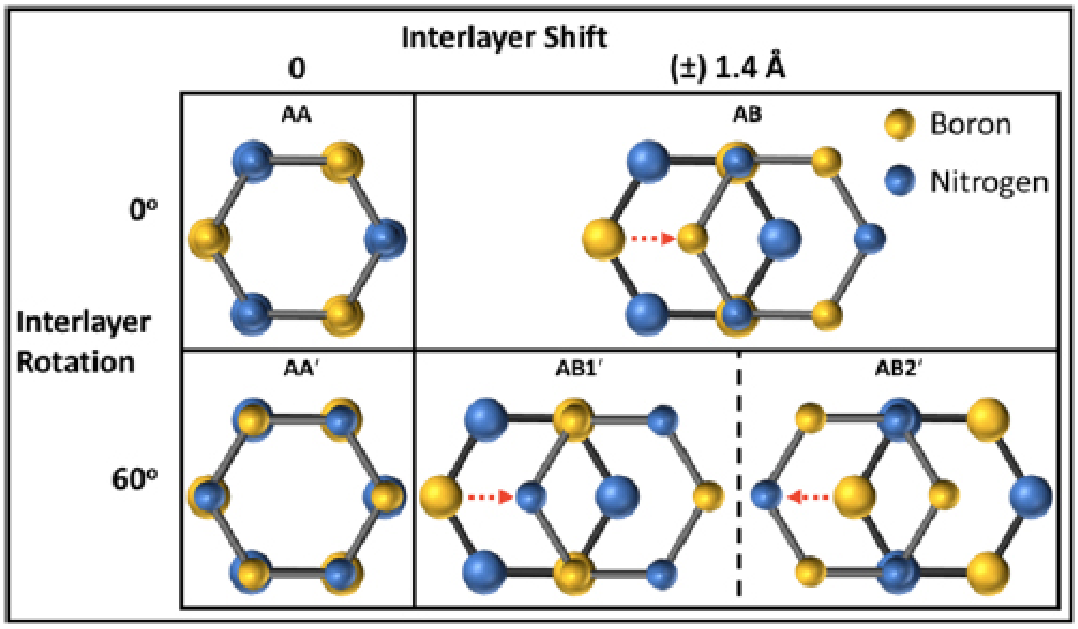


Figure 1: Stacking orders of hBN. Reprinted from [15].

Boron nitride comes in multiple forms, the most common ones are the hexagonal and cubic forms. The hexagonal form is one of the most studied forms, due to its structural similarity to graphene and its chemical and thermal stability. Hexagonal boron nitride is also a relatively common material, having a wide area of use. It is generally considered an insulator, but can alternatively be considered a wide bandgap semiconductor, having a bandgap between 5 and 6 eV. It also has an excellent thermal conductivity, noted as the best thermal conductivity, recorded between 1700 and 2000 W/mK [1]. The allotropes of BN are generally also inert to acids and corrosion [1]. The cubic form of boron nitride, however, has a band gap of 5.9 – 6.9 eV [8]. It has been shown to be a good thermal conductor as well. The crystalline forms of BN, cubic and hexagonal, are isoelectronic to carbon and structurally similar to diamond and graphite respectively.

Even when a material has a specific elemental composition, including the ratio of elements, there can be differences. A material can have different stacking orders. In the case of bilayered hBN, there are 5 distinct ones. The stacking orders are AA, AA', AB, AB1' and AB2'. In both AA and AA' stacking orders, all atoms are directly above each other, the difference being that in AA same species are below and above each other, while in AA' the order is alternating. In AB, AB1' and AB2' stacking orders, there is a shift along the x-axis, resulting in some of the atoms being directly on top of each other as well as having an atom directly above or below the midpoint of the hexagonal structure [15]. These stacking orders are visually represented in Fig. 1. The interlayer distances vary depending on the stacking order and the elements in question. For bilayer hBN, the interatomic distances between layers vary from 3.57 Å to 3.75 Å [37].

The aforementioned applies to 2 layers. One could go further, adding layers adds variation. For example, for graphene (and other materials with similar structure, including hBN) with 3 layers there is the possibility of having what is called ABC stacking, in which the third layer is shifted in relation to the second layer similarly to when the second layer is shifted in relation to the first layer [49]. This is presented in figure 2.

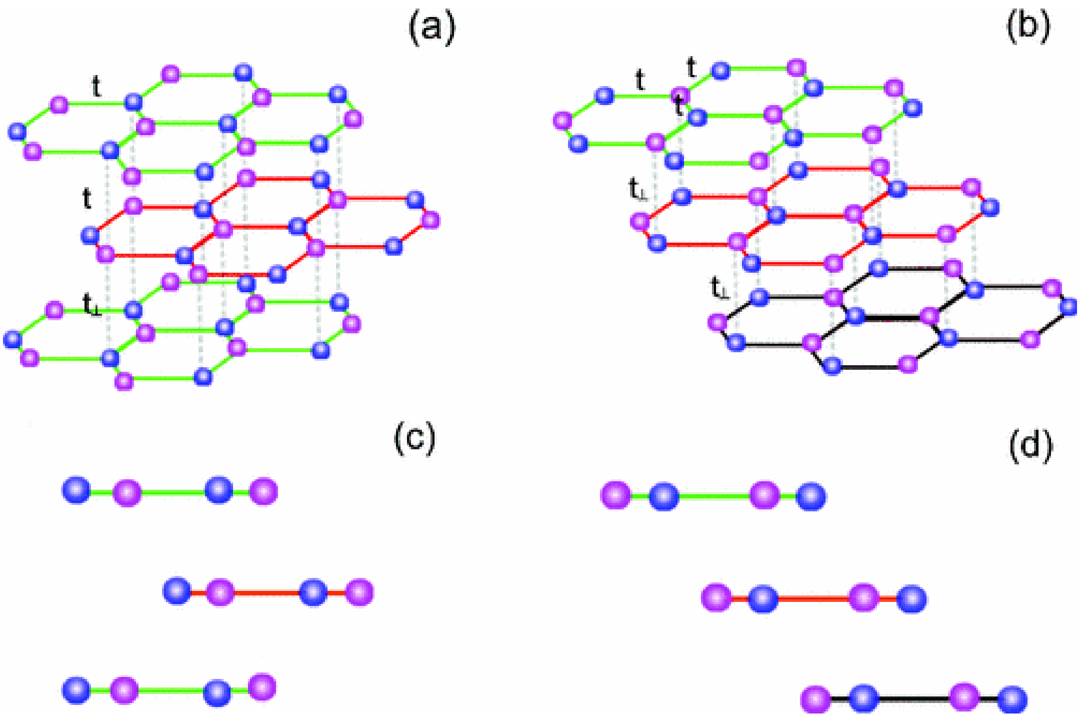


Figure 2: ABA and ABC stacking orders. A and B are viewed from a side and above point of view, while C and D are sideviews. Reprinted from [49].

The aforementioned stacking orders apply to hexagonal layered structures in general. The possible stacking orders increase with the number of layers, however all the stacking orders have at least some atoms directly above each other.

A subclass of materials are nanomaterials. There are different types of nanomaterials, with a common feature that at least one dimension is in the nano scale, sometimes required to be under 100 nm. However, this is not a strict requirement. Some examples of nanomaterials are quantum dots (0D), chains (1D) and surfaces (2D). The key feature of these nanomaterials is that one or more dimensions of the material is either smaller or in the same length scale as the wavelength of an electron. Here we will focus on 2D structures and these will be discussed next.

2.2 2D materials

As the synthesis methods improved, so did the number of possible materials. Nanosheets and similar surface type materials are examples of them. 2D materials, as well as any nanomaterial, have gained interest due to their interesting properties that are unlike their bulk counterparts. The reason for the difference in properties is due to the restrictions the dimensionality of the materials has. Nanomaterials also have an unusually high surface-to-volume ratio. A single layered material would strictly speaking only have surface. Surface atoms are less restricted than bulk atoms and have fewer neighbouring atoms than their bulk counterparts. This leads to effectively behaving in a different way. On a general level, surface atoms have more of what are called dangling bonds and higher entropy due to higher degrees of freedom. There are multiple 2D materials that have been studied, the most researched one being graphene. Sharing many similarities with graphene, hBN has also been extensively studied, among with other materials. Other well researched materials include, but are not limited to, $NbSe_2$, MoS_2, $MoSe_2$, and WS_2 [14, 22].

2.2.1 Interlayer forces and van der Waals interactions

Some layered structures, including but not limited to van der Waals (vdW) heterostructures, do not have strong bonding, i.e., chemical bonds, between the layers. They are instead kept together with weaker forces called vdW interactions or forces. These forces can roughly be divided into 3 groups: Keesom, Debye and London dispersion forces. The Keesom forces are forces between permanent dipoles, Debye forces are between permanent and induced dipoles and the London dispersion forces are between induced dipoles. The London dispersion forces are caused by the variance of electron clouds around the nuclei, therefore they are actually present in every single material. The London dispersion forces are the most relevant type of interlayer force for a vdW heterostructure and are also the reason for the interlayer spacing. For bilayer graphene, the interlayer distance is approximately 3.4 Å [27], which is its vdW equilibrium distance. The distance varies from material to material, i.e. it is material and element dependent.

One can also engineer layers that have a twist angle between the layers that results in a structure that is remarkably different. These are simply called twisted structures or in this case, twisted bilayer hBN. For the twisted structures to be distinct from the regular stacked structures, the angle can not be 60 degrees for hexagonal systems and 90 degrees for orthogonal. These angles will result in the regular stacked structures. As altering the crystal structure while keeping the elemental composition the same will change the properties of a material, it stands to reason that twisting of a layer will have a similar effect on the material. This is the case for the energetics and forces. Twisting can have other effects as well. It can result in a moiré pattern, and the material is then called a moiré material. These materials will be discussed next.

2.3 Moiré materials

The moiré pattern is a superstructure, forming a distinct interference pattern on a larger scale. In the case of twisted hBN, combining suitable angles and system sizes, a hexagonal pattern on a large scale can be seen. The reason why these moiré materials are interesting is that it has been shown that it can result in exotic properties such as superconductivity. [31] The moiré superlattice length can be mathematically expressed as follows:

$$\lambda = \frac{a}{sin(\frac{\theta}{2})}, \tag{1}$$

where a is the lattice length and θ is the twist angle. This effectively means that, with a large twist angle, the moiré superlattice length is smaller, which leads to less atoms needed to have a repeating pattern. Simultaneously, it means that, with a small twist angle, the moiré length becomes large, i.e., in order to have a repeating pattern, one needs more atoms. Since the moiré pattern is produced by having a lattice mismatch, often small, between layers, an obvious choice for a moiré material is therefore a material that is normally layered, such as graphite. Multiple materials have been used in research, such as twisted bilayer graphene, hBN, and In_2Se_3 [16, 50]. There are some limitations to the material properties to create a moiré material. First and foremost, the most interesting phenomena often appears at small twisting angles. A requirement is a layered structure, so that it is possible to adjust one layer at a time without reconstruction. Strong interactions (i.e., chemical bonding) between layers would most likely lead to reconstruction, meaning the loss of moiré pattern. However, there needs to be some interaction (i.e., weak interactions, vdW forces) between the layers, to keep it intact [26]. While twisting a layer with respect to others is generally interesting, some angles have been more interesting than others. The twist angle of 1.1° is one such angle for graphene, dubbed as the magic angle. It has been shown to result in merging of 2 Van Hove singularities, resulting in a flat energy band. Flat bands have been shown to host correlated electron states [2]. Van Hove singularities are critical points, minimum, maximum or saddle points, of a electronic band structure [32]. While they have generally been discussed in the context of electronic band structures, similarly they could be discussed in phonon dispersion relations.

As the crystal structure of a material has an impact on its properties, as mentioned in Sect 1, the band gap of a material can be tuned by altering the crystal structure. Even the stability of a material may change as the structure is altered. As was mentioned above, hBN is a non-ferroelectric insulator. However, in a twisted structure, ferroelectricity emerges, which has been shown in studies [51]. Next, ferroelectricity and its mechanisms will be discussed.

2.4 Ferroelectricity and structural properties

Ferroelectricity is an important property of many materials. Ferroelectric materials are defined as having a spontaneous polarization that is reversible with an external electric field. Even when the external field is removed, the polarization remains.

This effectively means that the polarization can be switched as needed, which in turn leads to being able to use ferroelectrics as memory for any device. It is worth noting that ferroelectricity can also emerge in smaller parts of a material. These are called ferroelectric domains. In these, the polarization is aligned so that there appears a negative and positive net polarization within the domain. Each domain, in turn, is separated by what is called a domain wall, in which the polarization switches abruptly [41].

There are multiple mechanisms of how ferroelectricity can be achieved. It can be displacive ferroelectricity [20], in which the symmetry is broken by the displacement of some atom, usually a central atom. If a material is centrosymmetric, the net polarization would cancel itself out, leading to no polarization and therefore no ferroelectricity either. Disturbing this balance would therefore lead to a net polarization. It is worth noting that all ferroelectric materials are also piezo- and pyroelectic but not all piezo- and pyroelectric materials are ferroelectric materials [3]. Another possibility is order-disorder ferroelectricity [20], where a material has permanent dipoles that are randomly oriented, again resulting in a net zero polarization. A disturbance, or an ordering of these dipoles, would in turn lead to polarization, which in turn allows for the ferroelectricity. A combination of the previously mentioned mechanisms, called hybrid mechanism, is a third possibility. There are also other mechanisms, including geometric, improper and lone-pair ferroelectricity [20]. The geometric, improper and lone-pair ferroelectricities will not be discussed here.

Depending on the mechanism, there are different controlling factors. In the case of order-disorder, it can be controlled with temperature. High temperatures lead to faster vibrations or movement in general, which in turn leads to randomness of orientations, due to statistical reasons. Statistical reasons in this context means that the distribution of the orientations can be assumed to be equally divided at any given time. This is because, as mobility increases, the time an ion spends in each orientation decreases to the point of it only being instantaneous. This effectively results in no polarization, meaning no ferroelectricity. If the temperature is lowered below a threshold temperature, the situation changes. Low mobility leads to ordering of dipoles, leading to ferroelectricity.

As was previously mentioned, hBN is an insulator, becoming ferroelectric in a twisted structure. Phase transitions between different ferroelectric phases can therefore emerge with the twist angle. Phase transitions can be detected by studying the phonon dispersion relations. The next section will go into more detail about phonon dispersion relations.

2.5 Phonon dispersion relations

Another way to study a material is computing phonon dispersion relations. Phonon dispersion relations are the equivalent of electronic band structures but for phonons, in which phonon frequencies are plotted against q-vector, which is equivalent to the k-space used in electronic band structures. To fully understand the importance of phonon dispersion relations, one must understand the intricacies of phonons in general. While the curves in electronic band structures are often divided into valence band and conduction band, a similar division is used in phonon dispersion relations. In the case of phonons, we divide them into acoustic and optical phonons. The acoustic phonons have lower frequency (i.e., energy) than the optical phonons. Acoustic phonons are describing vibrations that go through the material, much like sound moves through objects, while optical phonons describe vibrations that are relative to one another. Each of these can be further divided into transverse and longitudinal acoustic/optical phonons. In a transverse acoustic phonon, the atomic displacements are perpendicular to the motion of the wave itself. This can be visualized as how the waves in the ocean move forward, while the water can be seen as moving up and down. In a longitudinal acoustic phonon, however, the atomic displacement is along the wave. This can be visualized as some atoms in a plane getting closer to each other, while others are further away [53].

The different groupings of phonons have another reason as well, they are usually associated with different properties. Acoustic phonons are generally associated with thermal properties, such as thermal conductivity, while optical phonons are associated with dielectric properties and phase transitions. This is reasonable, since phase transitions, for example, require (more) energy and the optical phonons are higher in energy. Because optical phonons are associated with phase transitions, they are the ones to study closely, when a material is known to transition from one state to another, e.g., insulating state to ferroelectric state. This is one of the reasons to study phonon dispersion relations.

Another reason to study them is to check for stability. If a phonon dispersion relation of a structure shows no imaginary frequencies (often referred to as negative frequencies due to computational software output using negative values), it can be a sign of dynamic instability. Vice versa, if no imaginary frequencies are found, the structure should be stable and not in transition (i.e., unstable) [46]. This case, as the energetics case previously, have a bit of a tolerance. Due to the nature of the methodology, some imaginary frequencies can be present, especially at or near the gamma point. Computational methods by nature are not 100 % exact, but rather there is some approximation included. Therefore, small inaccuracies occur which can lead to imaginary frequencies.

Understanding the lattice dynamics is crucial in understanding multiple phenomena. The phonon dispersion relations can be used for that and other reasons as well, such as the study of thermal properties, but it has more exotic applications [46]. One such is studying a material for phonon softening. Phonon softening is essentially the lowering of phonon frequencies, meaning a lower energy. If a system requires less energy to create vibrations, i.e., disturb the system, it is a possible sign of an incoming phase transition. While phase transitions are universally well known and common, they are still interesting. They can be divided into 2 groups, first-order or instantaneous and second-order phase transitions. First-order transitions are the more relatable ones from real life situations, such as melting or boiling. These happen in a sense instantaneously, at specific temperatures and pressures. Second-order phase transitions however do not suddenly occur, but rather they are slowly approached. Such is often the case of ferroelectric phase transition. First-order phase transitions involve latent heat while second-order phase transitions do not [3].

When discussing phonon dispersion relations, it is worth mentioning the phonon flat bands. A phonon flat band appears when the group velocity is near-zero. It is mathematically expressed as follows,

$$v_g = \frac{dw}{dk} \tag{2}$$

Where v stands for velocity, subscript g stand for group, w and k being the frequency and wavevector, respectively. Simply put, the group velocity is therefore the change in frequency in relation to the change in wavevector. When the group velocity is near-zero, it can be considered to be a localized vibration instead of the usual vibrations that propagates through the system. These can exist for both the acoustic and optical modes. The phonon flat bands can affect the system in multiple ways, for example the thermal properties of a material, meaning that heat is not transported effectively, possibly leading to the material in question being an insulator. Flat bands can also affect the electron-phonon coupling. If a material has both electrons and phonon flat bands, it can enhance the coupling, leading to polaron formation. Polarons are electron and its surrounding lattice distortion combined, effectively meaning an electron coupled with a phonon. This interaction leads to increased effective mass, which can lead to entrapment of the electron, leading to a strong electron-phonon coupling [17].

The elecron-phonon coupling determines how the vibrations of a system affect its electronic properties and vice versa. There is a coupling between these two particles, as the deformation of the lattice(s) can and will lead to electron-phonon collisions. These happen in both polar and non-polar semiconductors and are called deformation potential scattering. In non-polar semiconductors, a local compression or decompression increases or decreases the band gap, leading to increase or decrease of the energy of states of the conduction band, which in turn leads to the electron being affected by a potential that is relative to the amplitude of the phonon or lattice vibration. The same happens in polar materials as well, but with polar materials there are more interactions, mainly electrostatic, from the polarization. Depending on the phonon type, the electrostatic forces can lead to the piezoelectric effect with acoustic phonons and, in the case of optical phonons, the electrostatics can lead to Fröhlich interaction [38].

A way of further studying the system is studying the DOS of the system.

2.6 DOS and phonon DOS

Studying the DOS of a system can be a useful way to obtain more information of a material. It describes how many available states of each energy are available in the system. Mathematically, it is defined as follows:

$$g(E) = \Sigma_n \delta(E - E_n), \tag{3}$$

where δ is the Dirac delta function, E is the energy in question and E_n are the eigenvalues of the system's Hamiltonian. The DOS can also be computed for phonons, seeing as they can be viewed as packets of energy as well. For phonons, the DOS is defined as follows [44]:

$$g(w) = \frac{1}{N} \sum_{\lambda} \delta(w - w_A) \tag{4}$$

where N is the number of unit cells and $\lambda = (\nu, \mathbf{g})$ where ν is the band index, $\mathbf{g}$ is the q-point and lastly w_λ is the phonon frequency. Now, the DOS can also be projected into cartesian coordinates as follows [44]:

$$g^j(w, \hat{\mathbf{n}}) = \frac{1}{N} \sum_{\lambda} \delta(w - w_A) |\hat{\mathbf{n}} \cdot \mathbf{e}_\lambda^j|^2, \tag{5}$$

where j is the atom index, and $\hat{\mathbf{n}}$ is the unit projection vector. Dirac delta function is a function that has an infinite value when the input, in this case $w - w_A$, is 0. Otherwise it has a value of 0.

While projecting the DOS onto Cartesian coordinates can be interesting, it is often more useful and interesting to see the local projection, resulting in the PDOS of a single atom, also called LDOS. This is done by simply summing the PDOS's of each direction, resulting in the LDOS as follows [44]:

$$g^j(w) = \sum_{\hat{\mathbf{n}}=\{x,y,z\}} g^j(w, \hat{\mathbf{n}}). \tag{6}$$

As it can be worthwhile to find differences between regions or domains within a structure, the LDOS can be used for this as well. This could be done taking into account the LDOSs of the atoms in a specific domain, assuming the borders are well defined.

One will often observe some sort of grouping in a binary system with the phonon LDOS, for the simple reason of elements being different. There can also be further grouping, due to the atomic neighbourhoods being different, depending on the system. If we consider the two cases of having a infinite single layer of hBN vs. a tri-layer infinite hBN, the difference is simple to understand. All of the borons and nitrogens in a single infinite layer have identical surroundings, no atoms in the z-direction (below or above) and always having alternating atoms in the in-plane (hexagonal) structure. Now, in the case of a tri-layer system, all the borons or nitrogens would still have the same in-plane atomic neighbourhoods of alternating borons and nitrogens in the hexagon, but now there is a difference between the layers. The top layer would have 2 layers of atoms below, same with bottom layer having atoms above, but the middle layer in turn would have a layer of atoms both above and below. One could argue that all 3 layers have the same number of atoms "close enough", however that would not take into account the direction nor the distance between the atoms. It is therefore logical that there would be a difference between the LDOS of the layers. Even in bilayer systems, there should be differences between the layers. In this case, both layers will only have one layer of atoms either above or below, the in-plane atomic environments are identical, alternating atoms and the hexagonal structure. The distance between the atoms is also the same, because the distance between the top layer and bottom layer is of course the same as the distance from bottom layer to top layer. However, the directionality is not the same. This might not be important in all cases, but in a binary layer it does make a difference. Seeing as how phonons can be thought of as vibrations, they are affected by forces and forces have directions, there should be a difference between the vibrations. The layers could either be attracted (or repelled) up or down, depending on the case. Now, this is in an idealized system, where there are no distortions in the hexagons and with equal distances between layers, but this is not always the case.

Further groupings could happen, depending on the system. Multiple materials have domains within them, for example in magnetic behaviour. This is also the case for other properties, such as ferroelectricity, there are materials that have ferroelectric domains. This is one of the reasons why it is worthwhile to not only look at the DOS of a system but also the LDOS of a system. With LDOS computations, we can at the very least understand if a system has domains in which the vibrations are different. This could in turn be used to further study the differences between these domains. In the case of ferroelectricity, in which there is polarization, it stands to reason that there are differences in vibrations, due to the differences in forces caused by the polarization. Studying the phonon dispersion relations and the LDOS could lead to the discovery of ferroelectric domains, be it multiple smaller ones or a larger one. Finding ferroelectric domains in turn might be useful information in further research into a material for practical applications.

Another atomistic property is the electron-phonon coupling (EPC). Simply put, the vibrations of the lattices of a material have an effect on the electrons and vice versa. If, for instance, the lattice vibrates in a way that the distance between layers is decreased, it will influence the behaviour of the electrons and therefore the electric conductivity. The opposite way of thinking of it is having a charge difference between lattices, causing an electrostatic force that would attract the lattices closer (or vice versa). This would then result in the vibrations of the lattice changing compared to the same system without any polarization. How much these coupling affects the system is harder the determine, due to it depending on multiple factors such as, but not limited to, size, elements in the material and distances between atoms. What is clear, however, is that as the charge difference increases, so will the effect. The same is true for the distance, except that they have an inverse relation. The electron density of an element should increase the effect as well. Technically, the coupling can go both ways, EPC can both increase and decrease the ferroelectricity of a material.

Due to the increasing size of the moiré material with decreasing twist angle, one needs to have effective computational methods. An excellent method for big systems is machine learning. In the following section, we shall go over what machine learning is and how it works.

2.7 Machine learning-based atomistic modeling

There are multiple reasons to justify the application of computational methods to material research. It is a safer way to study hazardous materials, reducing potentiallydangerous human mistakes while lowering costs and time frames. Moreover, atomistic simulations enhance experimental searches of new materials, giving useful insights on structures and properties, while decreasing the amount of experiments needed. This is especially useful for engineered materials that are non-trivial to synthesize, such as vdW heterostructures.

Among the available methods, ML approaches are significantly more cost-effective, by an order of magnitude faster in systems of same size [10]. This can vary depending on various factors, such as system size and complexity. This effectively means that ML approaches provide access to larger systems compared to other computationally efficient methods, such as DFT. Density functional theory is considered one of the fastest ab initio computational methods. Generally, in computational methods, the more accurate results we want, the more the computational cost grows. However, this is not strictly true when it concerns ML-based models. While ML approaches may introduce some error, it can be minimized with a proper training set and model parametrization.

In simple words, ML aims at finding patterns and correlations between input and output data. This is achieved with the help of empirical data, called training set, which trains the model to predict the outputs of a given input. In the context of atomistic modeling, these input-output pairs are used to map a potential energy surface (PES). The key concepts mentioned above will be discussed next in this section.

The learning task happens through mathematical modeling, leading to the prediction of new data points. Learning in an ML context means finding a suitable function to which the data points fit. This is done with the use of a loss function, which calculates the difference between the (approximated) fitted value and the exact value at each point [21]. The goal is therefore to minimize the loss function at each data point, without overfitting (i.e., avoiding a perfect fit of the training data). As a result, the model is able to interpolate between data points, effectively giving the user access to an output for every data point within the configuration space. A simple way to understand this is how humans learn from experience. If we know from past experience that it tends to be hot in both June and August (i.e., training data), we will expect that it will also be hot in July (i.e., interpolation). This is effectively what is meant by ML prediction. Ideally, an ML model relies on interpolation exclusively. However, that depends on the range and diversity of the training data, leading to extrapolation in practice. Continuing with the previous example, we could apply that same past experience to predict trends outside summer months (i.e., extrapolation), but our uncertainty increases considerably. Extrapolation is usually not a desired behavior for these models, as we discuss in sect. 2.7.3. These characteristics are key to understand the cost-effectiveness of ML methods in general compared to ab initio methods. Due to the abstract mathematical nature of ML, there are many design choices that enable its application to diverse fields, using empirical data as input that ranges from language usage to atomic configurations.

The classification of ML models is intricate, although useful to understand their properties and applications. One way to divide them is considering the notion of supervised and unsupervised approaches [19]. Their main difference is that supervised ML models rely on labeled data as opposed to unsupervised ones. Label in this context means that the output is known for specific data points, i.e., already computed or otherwise confirmed. Supervised and unsupervised models have their own strengths. A supervised model learns from the known connection between input and output, to make predictions within the configuration space of the data. In this study, atomic positions correspond to the input data and the outputs are energies and forces, as we will explain in Sect. 2.7.1. That information is used to map out an approximated version of the PES, which allows the prediction of energies and forces of structures that are not part of the training set [19]. That said, a supervised model should be used extremely carefully for structures too dissimilar to those in the training data, since extrapolated values are unreliable. This is especially true for compositions with chemical species not included in the training set. An unsupervised model on the other hand tries to find connections between data points, learning underlying patterns. Those connections can then be used to group data points, e.g., to find which structures are closer to each other based on a similarity measure [19]. An example of this would be to group carbon-based molecules based on functional groups, e.g., grouping into hydrocarbons, aldehydes, alcohols, etc.

Other possible classification considers model parametrization. Parametric and nonparametric approaches differ on their assumptions of the model parameters. The former fixes the set of parameters required to estimate the output, independently of the amount of input data. While reducing the problem by assuming a parametric form for the function, it can also cause issues without prior knowledge of the functional form. If the initial assumption is poor, it will result in poor results [19]. In contrast, nonparametric ML models avoid any explicit assumptions about the underlying structure of the input data. While this approach requires a larger amount of training data, it is able to adapt to wider range of functional forms, resulting in more accurate results [21].

This research aims at accurate prediction and the functional form was not known by the author of this manuscript. For these reasons, the ML model of choice was a supervised nonparametric technique known as the Gaussian approximation potential (GAP), based on Gaussian process regression (GPR). In the rest of this section, we will introduce the specific methodology followed in this study. First, we will discuss GPR and its general framework, Gaussian processes (GP), before focusing on their application to atomistic modeling. We will continue with details about the datasets involved and their mathematical representation in Sect. 2.7.2 and 2.7.3 .

2.7.1 Gaussian approximation potential (GAP)

The starting point to understand our methodology is Gaussian processes (GP). A GP corresponds to a stochastic, i.e., probabilistic, mathematical model that generalizes the Gaussian probability distribution to describe functions [36]. Such model is not restricted by dimensionality or number of variables. In a GP, we assume that the predictive distribution of the outputs is a multivariate Gaussian.

As we want to estimate the values of a function (the PES) between known data points, we need to interpolate. The goal is to approximate the underlying relation between the variables and the ouput, which may benefit from Gaussian process regression (GPR). For simple correlations, e.g., how the price of a product affects its sales, it is easy to predict without the use of regression. In the case of atomistic modeling, the correlations tend to be more complex and therefore more advanced techniques such as regression are necessary. With regression, we are able to map the approximate PES and use it to predict the energy values of structures that were not involved in the generation of that map. One of the benefits of using GPR-based models is that they can both predict and give an error assessment of their performance [23].

The Gaussian approximation potential (GAP) is a GPR-based framework adapted to atomic systems [4, 5], where the fitted function is the PES of the system. In GAP, the most relevant assumption of PES fitting is that the total energy of a system is the sum of the local energy contributions. These local contributions depend on the descriptors, i.e., mathematical representations used to describe the chemical and/or physical information of the system. Mathematically, the total energy E is formulated as follows:

$$E = \sum_{d}^{Descriptors} \sum_{i=1}^{N_d} \varepsilon_d(\mathbf{q_i}) , \tag{7}$$

where d denotes the type of descriptor, N_d is the number of descriptors $\mathbf{q}$ per type, and ε_d is the atomic contribution to the total energy. Different descriptors are used for different energy terms: 2-body, 3-body and n-body descriptors for 2-body, 3-body and n-body energy terms [24].

Each of the local energies is expressed as an indepentent GP as follows:

$$\varepsilon_d(\mathbf{q}) = \sum_{m=1}^{M_d} c_m k_d(\mathbf{q}, \mathbf{q_m}) , \tag{8}$$

where c_m stands for the fitting coefficient, and M_d is the number of representative points of the descriptor type d. These representative points can include all the data points or a subset of them, known as sparse set. Additionally, k_d represents a kernel function, which describes how similar an atomic environment is compared to another, with $k_d = 1$ meaning identical and $k_d = 0$ completely dissimilar. Following a training step, the model tries to adjust the fitting coefficients from eq. (8) to approximate the structural energies and forces of previously DFT-computed structures, part of the training set. Since the total energy is assumed to be the sum of the local energy contributions, it follows that the forces depend on the derivative of the local energy with respect to the Cartesian coordinates [24], as shown in the following equation:

$$\frac{\partial \varepsilon_d}{\partial r_{i\alpha}} = \sum_{m=1}^{M_d} c_m \nabla_{\mathbf{q}} k(\mathbf{q}, \mathbf{q_m}) \frac{\partial \mathbf{q}}{\partial r_{i\alpha}} , \tag{9}$$

where $r_{i\alpha}$ denotes one of the Cartesian coordinates and ∇ is the differential operator. [24].

Up to now, we have introduced the basis of GAP, a kernel-based ML interatomic potential. This framework provides a recipe based on GPR to compute the PES of atomic systems. In order to effectively use this framework for predictions, we need a way to express the system we wish to study mathematically via some representation. That is the role of atomic descriptors, which will be discussed in the following subsection.

2.7.2 Atomic descriptors

A key component of ML is the so-called descriptors. The descriptors are mathematical representations of a system, i.e., atomic environments in our case. Formally, they encode an independent variable, the input, into a vector that the output depends on [11]. Given an atomic environment, a descriptor is a function mapping any atomic positions in its neighborhood to such vector. This description is not unique, meaning we can define several functions for the same system. For instance, atomic representations range from interatomic distances and bond angles (2-body and 3-body descriptors respectively) to relationships involving several atoms (many-body descriptors). When choosing a descriptor for our ML model, we should keep in mind the output of interest and its symmetries. For example, the local energy and forces of a structure are invariant under translations, rotations and permutations of identical atoms. Any descriptor used in their computation has to preserve these symmetries.

While 2-body and 3-body descriptors are fundamental descriptions of any atomic system, they are not enough to describe energies and forces in general, especially in larger structures. This is why we rely on many-body descriptors too, such as the smooth overlap of atomic positions (SOAP) descriptors [5, 7]. The starting point of the SOAP formalism is representing the atomic density of the neighborhood of a certain atom i. Such neighborhood is defined as the sphere S of cutoff radius r_{cut} centered on atom i, where each atom included in it contributes with a Gaussian function centered at its nucleus. This is mathematically done as shown in the following equation:

$$\rho^{(i)}(\mathbf{r}) = \sum_{j \in S_i(r_{cut})} \rho_j(\mathbf{r}) \,, \tag{10}$$

where $\rho(\mathbf{r})$ is the atomic density as a function of distance from the nucleus. A thing to note is that introducing a cutoff radius implies that we are localizing the contributions of the atom in question. While the exact calculation of the density would require setting r_{cut} to infinity, the atomic contributions diminish as the distance grows. This motivates the cutoff approximation, whose accuracy will depend on the choice of cutoff and the relevance of long-range contributions in each particular system. Also, computing the exact atomic density is not feasible due to the increase in computational cost, especially for large systems. This locality assumption is one of the fundamental sources of error in the GAP-SOAP formalism [24].

We can define a kernel function to compare two atomic environments as follows:

$$k^{SOAP}(i,j) \propto \int d\hat{R} \left| \int d\mathbf{r} \rho_i^*(\mathbf{r}) \rho_j(\hat{R}\mathbf{r}) \right|^2 . \tag{11}$$

Here, k^{SOAP} gives the similarity between atomic environments i and j, taking values between zero and one, with zero meaning completely different environments and one completely identical. Note that all possible rotations are taken into account through the integration. Additionally, the exponent could have any positive integer value n, but it has to be at least $n = 2$ to retain angular information. This is because one does not have relative positions without a reference point (i.e., with only 1 atom).

The SOAP kernel in eq. (11) could already be computed, but it is computationally expensive. To make the computations more efficient (i.e., cheaper), the atomic densities are expanded into a radial basis and spherical harmonics, which allows the construction of a rotationally invariant descriptor in vector form. The density is then expressed as follows:

$$\rho(\mathbf{r}) = \sum_{j \in S(r_{cut})} \sum_{nlm} c^j_{nlm} g_n(r) Y_{lm}(\theta, \phi). \tag{12}$$

In eq. (12), c_{nlm} is the expansion coefficient, $\{g_n(r)\}$ is the radial basis and $Y_{lm}(\theta, \phi)$ are the spherical harmonics. Note the lack of indexing for atomic density, this was done for clarity and compactness of the equation, making it easier to understand. After some manipulation of the equation and normalization, the resulting SOAP descriptor can be expressed as:

$$\mathbf{q}^{SOAP}(i) = \frac{\mathbf{p}(i)}{\sqrt{\mathbf{p}(i) \cdot \mathbf{p}(i)}}, \tag{13}$$

where $\mathbf{p}(i)$ is a vector whose components are known as the power spectrum of the atomic density:

$$p_{nn'l}(i) = \sum_m c^i_{nlm} \left(c^i_{n'lm}\right)^* . \tag{14}$$

Calculating the dot product of the descriptors of atoms i and j gives us the SOAP kernel:

$$k^{SOAP}(i, j) = (\mathbf{q}^{SOAP}(i) \cdot \mathbf{q}^{SOAP}(j))^{\zeta}, \tag{15}$$

where ζ is effectively a sensitivity parameter, allowing for finer distinction between similar but not identical atomic neighborhoods when using a value greater than one. It is worth noting that, although chemical species are not explicitly stated here, we included multispecies treatment in our calculations, i.e., we did not limit our GAP model to a specific element.

In this study, we use the state-of-the-art SOAP descriptors known as soap_turbo descriptors [7]. This representation is an optimized version of the earlier established SOAP descriptor. It shares similarities and differences when compared to the SOAP descriptor. The main difference is how the atomic density is expressed. In the original SOAP descriptor, the atomic densities were represented in the following way:

$$\rho_j(\mathbf{r}) = exp\left(-\frac{|\mathbf{r} - \mathbf{r_j}|^2}{2\sigma^2}\right), \tag{16}$$

where $\mathbf{r_j}$ is the atomic position of the central atom and σ is a parameter controlling how quickly the atomic density decays [7]. The soap_turbo descriptor on the other hand has the following density:

$$\rho_j(\mathbf{r}) = exp\left(-\frac{1}{2}\frac{(r-r_j)^2}{\sigma_r^2}\right) exp\left(-\frac{1}{2}\frac{r_{\perp,j}^2}{\sigma_\perp^2}\right) . \tag{17}$$

From eq. (16) and (17), it is clear that they share both similarities and differences. As far as similarities are concerned, they both describe a local environment defined by a radius, clearly visible by the first term of the equations. However, the soap_turbo descriptor has the atomic densities divided. The difference in their atomic densities is due to the approximation of the density with a separable form:

$$\rho_j(\mathbf{r}) \approx \rho_{r,j}(r)\rho_{\perp,j}(\theta, \phi) . \tag{18}$$

That is, soap_turbo assumes that the total atomic density can be divided into radial and angular parts. This mathematical trick allows more efficient computations, up to 20 times faster, while keeping their accuracy [7].

While the SOAP descriptors implicitly carry the information about distances and angles, it is still useful to have them in separate descriptors. This is because when having multiple contributions at the same time, it might become hard to differentiate which parameter affects them and how much. Including the fundamental descriptors, such as angle and distance, allows the ML model to fit the data more accurately. While the computational cost increases as a result, such increase is usually insignificant and justifiable.

Something ML models in general have in common is their reliance on datasets. Depending on the situation, two or three types of datasets are required, namely training set, validation set and test set. These are discussed in the upcoming section.

2.7.3 Datasets

The main datasets in ML modeling are the training set, the test set, and the validation set. The training set consists of experimental data, synthetic data (i.e., computer generated) or a mixture of both. This dataset is arguably the most important one as it controls how well the ML potential performs. That is, an ML potential will be as good as its training data. Theoretically, such potential will have minimal error when fed with an infinite amount of information. This is not possible in practice, otherwise using ML would be redundant. Note that an ideal ML potential still carries the error of its training data, i.e., it will have zero error only when the training data has no uncertainty. As a result, we need to decide how to balance the size of our training set versus the accuracy of our results.

Perhaps one of the most important aspects of a training set is its diversity. The wider the scope of the ML potential, the richer the training data needed. For example, if one wishes to study structures of a single chemical element, there is no reason to include atomic data involving other species to properly fit its PES. However, in the presence of more elements, the training set must include information on them. The same logic applies to any prior knowledge about the modeled systems. As the training set grows, along with different types of structures, so does the computational cost. Therefore, one should always optimize the amount and diversity of input data to create a computationally cost-efficient training set. Note that increasing the presence of a specific class of data (e.g., having large amounts of cubic BN or hexagonal BN structures) can also affect the generated ML potential, introducing a bias towards that class. Generally, one should avoid bias, but there are advantages in overrepresenting data for certain applications. Overall, this discussion boils down to using a training set that covers enough of the relevant configuration space so the ML potential interpolates between those regions. By definition, GPR models perform well within the interpolation range, but poorly outside it. Extrapolation can lead to faulty predictions and, consequently, false conclusions. One should therefore always check the results carefully.

The second type of dataset, the validation set, is useful for checking the performance of the ML potential. The validation set is not strictly speaking needed, but it is nevertheless a very useful tool, which can make the process of generating a working ML potential both smoother and faster. The validation set is utilized as a first checkpoint to see roughly how well the ML potential performs with similar structures to those on the training set. The real use of a validation set is to optimize the hyperparameters of the ML model, such as regularization, to avoid under- or overfitting the data. A simple way of looking at overfitting is thinking about the input data as points on a xy-graph and the ML potential as a function. Overfitting means that one would create a function that will pass exactly through each point. In that case, the results for the training set would be excellent, but what would happen if you were to check how close a new data point is approximated by that function? Most likely, the ML potential would poorly fit any new result. In this context, that would mean that while the ML potential might give excellent results for the training set, it has now lost all predictive power. On the other hand, underfitting is the opposite issue: the function is not required to go near any of the training data points, resulting in the allowance of massive errors, which in turns means that one does not get any meaningful results. Both situations are avoided via proper regularization, which is why it is key when generating an ML potential.

After the parameters have been optimized, the test set is used. The test set is the final test for the potential and should comprise structures not included in the training set, as diverse as possible within the problem. The reason is that one wants to evaluate how the ML potential performs without having the exact data in the training. If the ML potential was not over- or underfitted, the results should be reasonable, but generally there will always be some error.

Depending on the field of study and the type of problem one is trying to solve, the method to generate those datasets can vary. In this study, we used DFT calculations as reference data. That is, we trained our ML potential with DFT data and measured its performance by comparing its predictions with DFT results. This approach is the standard procedure in ML-based models for computational chemistry and material sciences. In particular, our datasets are a collection of BN structures defined by their atomic species and positions, along with the DFT-computed energies and forces acting on the structure.

While all of the previous datasets have their importance, not all of them need to be treated separately. In an iterative process, which is essentially building up the training set until the results are good enough, one can use a previous test set as a validation set. This is easier to understand with an example. Let us assume we have a training and a test set, consisting of 100 and 10 structures respectively, and that the training set is not enough for the ML model to be able to predict with the needed precision. We could then add the information from this test set to the validation set and expand the training set with another 100 structures. At this point, we would create a new test set and check the performance of our model again. This process could be repeated as many times as necessary. This is the basic principle of iterative training, applied in this research.

3 Computational methods

As this study was computational in nature, the main tools used were DFT and ML for atomistic modeling. DFT was used to generate the training data, while ML was used for the rest. The reason for this division was that, as mentioned in Sect. 2.3, the smaller the twist angle, the larger the system. While the system sizes were in the low hundreds, DFT is a good choice. However, as the system sizes in this study were up to roughly 25000, the ML approach was more approriate.

We started with the preparation of the datasets from sect. 2.7.3, using Python-based scripts and the atomic simulation environment (ASE) package [18]. We generated a total of 900 hBN structures: 500 stacked bilayers, 300 sliding bilayers, and 100 monolayer structures. The stacked structures include AA and AA' stacking orders in a 1:1 ratio. Both sliding and stacked structures consist of 8x8 atoms per species and layer, giving a total of 256 atoms per structure. All stacked and sliding bilayers had an interlayer distance of 3.305 Å. The sliding was done by starting from a AA-stacked structure, moving the top layer with 100 steps, each step being 0.01 Å in x-, y- or xy-direction. Finite point DFT calculations were performed on all the structures with VASP [28–30]. The parameters of those computations are shown in table 1. The most notable ones are the following: EDIFF, which sets the difference in energy between steps to determine convergence of the calculation; KSPACING, which establishes the distance between points in the k-point grid and ISIF, which controls what is calculated and the degrees of freedom (in this case forces are calculated, atomic positions are free to change while cell volume and shape are locked). However, the atomic positions were locked by setting NSW to zero.

Afterwards, we trained an ML potential with the QUIP package [9]. In particular, we relied on the GAP implementation [4] and the soap_turbo library [7], setting their parameters as shown in table 2. We used seven atomic descriptors in total: one 2-body descriptor per possible dimer (B-B, N-N, B-N) to describe the interatomic distances, two 3-body descriptors to account for the bonding angles, and one soap_turbo descriptor per chemical species. While the many-body descriptors carry implicitly the information regarding distances and angles, including the 2- and 3-body representations

Table 1: VASP settings for the computations of all datasets.

Setting	Value
ENCUT	600 eV
EDIFF	10^{-8} eV
EDIFFG	$(-1) \cdot 10^{-3}$ eV/Å
ISIF	2
KSPACING	0.2 Å^{-1}
NSW	0

Table 2: GAP and soap_turbo parameters for ML potential generation in this thesis.

Setting	2-body	3-body	SOAP turbo
l_{max}	-	-	8
α_{max}	-	-	{8,8}
rcut_hard (Å)	6.5	6.5	6.5
rcut_soft (Å)	-	-	5.5
ζ	-	-	4
δ (eV)	0.5	0.01	0.1
sparse_method	uniform	uniform	cur_points
n_sparse	40	200	-
n_rattled_8x8	-	-	500
n_sliding-x	-	-	100
n_sliding-y	-	-	100
n_sliding-diag	-	-	100
n_monolayer	-	-	100

improves the overall accuracy of the potential.

Once we obtained an ML potential, we did some preliminary tests of its performance using the TurboGAP software [12] and a validation sample. This dataset contains a small collection of twisted structures with an interlayer distance of 4.35 Å. This gave us a rough understanding of the accuracy of our ML potential to adjust the regularization and other hyperparameters. The process of training plus testing was repeated until the energies and forces were close enough to the DFT-computed values. This technique is known as cross-validation [11]. When the ML potential was deemed accurate enough, it was properly tested on a richer test set, following the same approach. The test set consists of 50 bilayered hBN, consisting of AA-stacked bilayer hBN, amount of each atom is 4x4 or 8x8 per element per layer.

The next step of this research was to further study the effects of twisting one layer on the simulated material. As the twisted hBN structures were relaxed with the ML potential, atomic displacements and interlayer distances were calculated. The former were computed as the difference between the initial and final atomic positions relative to the center of mass (CM). First, we calculated the relative coordinates of each atom as follows:

$$(x, y, z)_{n,rel} = (x, y, z)_n - (x, y, z)_{\mathrm{CM}}\,, \tag{19}$$

where $n = 0, 1$ denotes the starting and ending coordinates. This was performed per atom and axis separately, for both initial and final coordinates. The displacement follows from the Pythagorean theorem:

$$d = \sqrt{(x_{1,rel} - x_{0,rel})^2 + (y_{1,rel} - y_{0,rel})^2 + (z_{1,rel} - z_{0,rel})^2}\,. \tag{20}$$

We also considered two variations of this quantity: in-plane displacement (xy-plane) and out-of-plane displacement (z-axis). On the other hand, interlayer distances were defined as the difference between the average z-component of the atomic positions in the top and bottom layers, respectively.

The next analysis of this thesis work was the phonon dispersion relations. For this purpose, we relied on the Python package Phonopy [45, 47]. First, the original structures were expanded into supercells, increasing their size to roughly 30 Å in x- and y-directions. While not so important in the study of 2D-structures, we doubled their length in the z-axis as a precaution. In conjunction with the creation of supercells, we generated hundreds of displaced structures. The displacements were performed for each atom in all directions randomly, considering the existing symmetries. The resulting structures were then used to compute force constants, by means of the finite differences method in VASP and running our ML potential with LAMMPS. The most relevant phonopy parameters are presented in table 3. On top of these, the chosen high symmetry path was $\Gamma - \mathbf{A} - \mathbf{K} - \mathbf{H} - \mathbf{M} - \mathbf{L} - \Gamma$. This path was chosen for completeness sake, since a smaller one would have sufficed. Regarding the force

Table 3: Phonopy parameters for supercell and displacement generation. The varying dimensionality was chosen depending on the original size of the structure, aiming for 30 Å in x- and y-directions.

Setting	Value
displacement (Å)	0.01
primitive axis	auto
dimension	varying
mesh	20 20 20
band-points	101

constant computations, we used symfc [39].

Another aspect of the computational study was the computation of LDOS, which gave us insights on the differences in the available energy states for each atom. These were computed as presented in eq.6 and phonopy's default values were used.

4 Results

Following the methodology described in sect. 3, we obtained several results regarding the performance of the trained ML potential for hBN bilayer structures, as well as atomic displacements and phonon dispersion relations. In this section, we present those results and discuss some effects of twisting layers on hBN structures, such as the emergence and evolution of moiré patterns.

The primary performance check of the ML potential was cross-validation of energy and forces. Figure 3 shows the difference on energy per atom between DFT and ML calculations. Note that we will refer to the dataset used in Fig.3 as validation set from now on. The root-mean square error (RMSE) for GAP energies is 0.415 meV/atom. Moreover, the error on average for GAP forces is 0.504 eV/Å. Their component-wise errors are 0.274 eV/Å, 0.348 eV/Å and 0.888 eV/Å for x-, y- and z-direction, respectively. All these errors were deemed small enough to guarantee an accurate potential, especially for energies when compared to other GAP-based hBN potentials [33]. While the error in the computed forces is considerably larger in the z-direction, it is reasonable since it is the most relevant direction to the vdW forces in hBN bilayers. It is also the direction with higher variance of atomic displacements, captured by the fitting of the corresponding force component.

A secondary performance check consisted in analyzing the phonon dispersion relation of different GAP-generated structures, to reveal possible dynamic instabilities in those systems. Since these relations did not show significant imaginary frequencies, i.e., the structures were properly relaxed, we concluded that the performance of our ML potential is sufficiently accurate and reliable to proceed with its application to twisted hBN bilayers.

We continued with the study of the properties and any emerging phenomena in the material, through both the phonon dispersion relations and the atomic displacements. In addition to these, we supported the analysis with LDOSs. Note that the dataset used in these calculations, from here on called position dataset, is a separate dataset from the validation one and, therefore, the structure numbers in both Tables 4 and 5 and Figs. 4, 5 and 6 are not related to those in Fig. 3.

Concerning the energy characterization of the structures, AA-stacked structures have a GAP energy per atom of -8.786 eV/atom, while twisted structures have -8.799 eV/atom. These results, included in Table 4, suggest that twisted structures are not only stable but even more than AA-stacked structures, which is known to be metastable [40].

Furthermore, the displacements of atoms from their initial to their relaxed positions was measured as described in Sect. 3, and included in Table 4. We noticed that the in-plane (IP) displacements are fairly small, especially for larger twist angles, but they increase up to an order of magnitude as the twist angle decreases. This seems to start around $2-3°$ twist angle. Regarding the out-of-plane (OOP) displacements, they present some variation with larger twist angles, as opposed to smaller angles where they fluctuate around 0.275 ± 0.001 Å. These behaviors are shown in Fig. 4 and 5, respectively.

On the other hand, there is a clear increase in the average interlayer distance, from

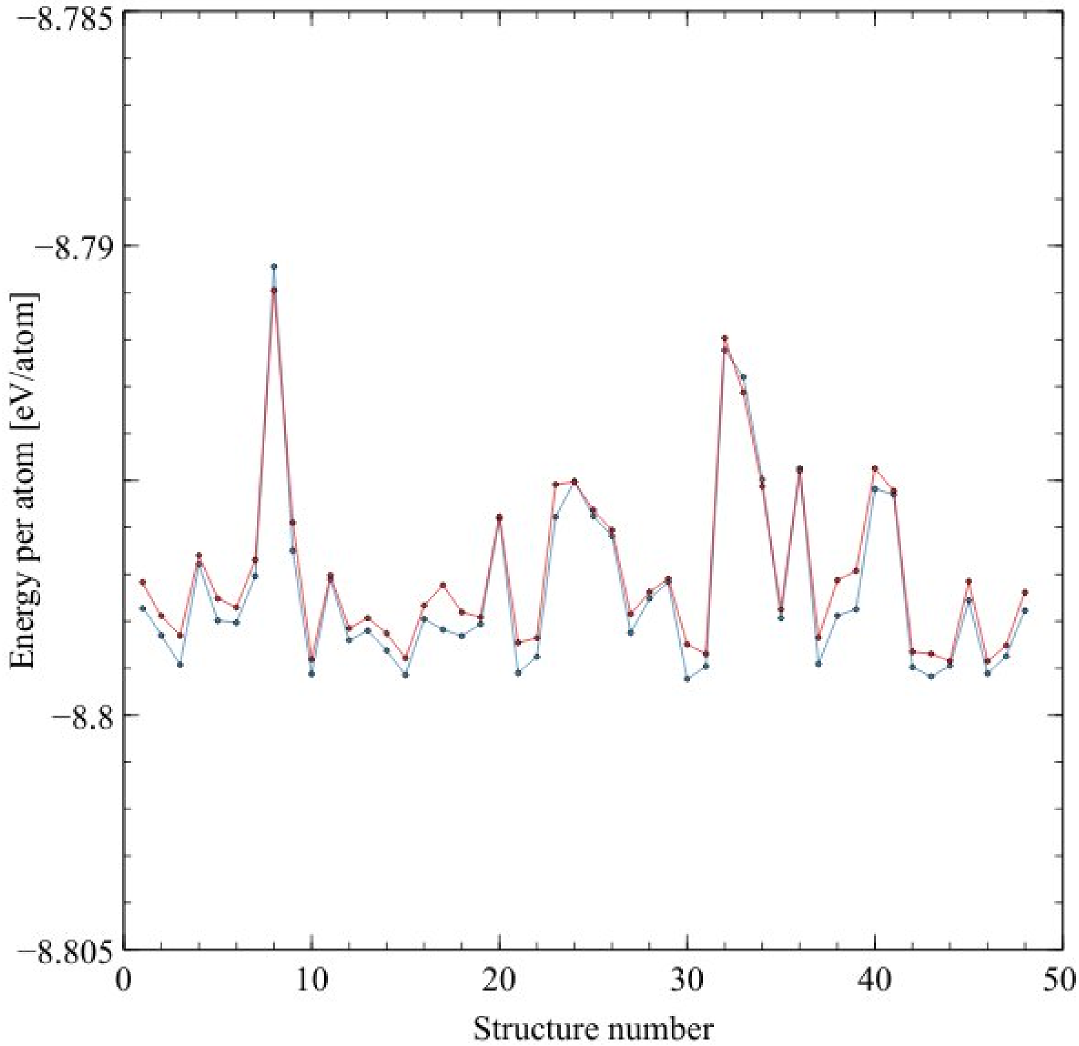


Figure 3: Comparison of energies per atom computed with DFT (blue) versus GAP (red), for 49 hBN structures. Note that the first 24 structures are 4x4 sized bilayers while the last 25 structures are 8x8.

3.66 Å for the AA-stacked structure to a maximum of 3.84 Å for the 21.79° twist angle structure. As shown in table 5, the interlayer distance becomes practically constant for smaller twists. These results are also shown in Fig. 6. This trend is dissimilar to what has been recently reported, as their interlayer distance is seemingly constant at higher angles while it starts to decrease with lower twist angles. The similarity is that in each study, the interlayer distance overall is decreases $\tilde{0}.05$ Å. In our study, the interlayer distance is also $\tilde{0}.4$ Å larger than the previous reported results. Without having access to more information on how they conducted their study, it is difficult, if even possible, to compare the results of our and their study. [33].

Additionally, the interlayer displacement computed per atom is not homogeneous, but rather it varies depending on the location. As can be seen from the subplots in fig. 7, the displacements are clearly smaller in the middle of the structure compared to

Table 4: Out-of-plane (OOP) and in-plane (IP) average displacements ($\bar{d}$) as well as the energies obtained during GAP-computed structure relaxation for selected twist angles. The 0° twist corresponds to an AA-stacked structure, included here as energy reference. The angles shown here were chosen evenly, with 5 steps in between, including the highest and lowest ones in the dataset. Note that the structure number relates to the twist angle: the higher the former, the lower the later and N stands for structure number. Position dataset was used in this table.

N	Atoms	Twist angle (°)	OOP $\bar{d}$ (Å)	IP $\bar{d}$ (Å)	E_{atom} (eV/atom)
N/A	256	0	0	0	-8.786
1	28	21.79	0.255	0.002	-8.799
2	76	13.17	0.272	0.002	-8.799
3	148	9.43	0.279	0.000	-8.799
6	508	6.01	0.284	0.001	-8.799
11	1588	3.15	0.271	0.002	-8.799
16	3268	2.13	0.275	0.006	-8.799
21	5548	1.61	0.276	0.009	-8.799
26	8428	1.30	0.275	0.012	-8.799
31	11908	1.08	0.277	0.013	-8.799
36	15988	0.93	0.274	0.014	-8.799
41	20668	0.82	0.276	0.015	-8.799
43	22708	0.76	0.283	0.018	-8.799
44	23764	0.74	0.275	0.015	-8.799
45	24844	0.73	0.277	0.015	-8.799

the outer parts. We also noticed distinct areas in the bottom right and top left corners, where the displacement is larger than in the surrounding area. These results indicate a difference in the forces acting between the layers in specific regions of the material. A possible explanation for this is the presence of polarization across the material. If there are opposite polarizations between the top and bottom layers within a certain domain, the layers would be drawn closer together due to attractive forces. Conversely, identical polarization between layers would result in repulsive forces and areas with lower interlayer distance. It is worth noting that the OOP-displacement could be manipulated by altering the starting interlayer distance. If the system relaxes to the same interlayer distance, indepedendant of starting positions, one could effectively increase or decrease the OOP-displacement. In our case, we had a reasonable 4.35 Å distance.

As per figures in Fig. 6, the interlayer distance varies from 3.78 Å to 3.84 Å. Comparing these results with those of the bilayer AA-stacked structure, it would seem that the average interlayer distances in twisted structures are a bit higher. While the difference int interlayer distance between AA stacked and twisted structures is significant, roughly an increase of 5 %, it is not too worrisome.

Up to now, we have focused our discussion purely on displacements within the material. Analyzing phonon-related properties can add a valuable extra layer

Table 5: Average interlayer distance ($\bar{d}_{interlayer}$) in twisted hBN structures. The twist angles were chosen with 5 steps in between, resulting in effectively randomizing the twist angle. Note that N stands for the structure number. The position dataset was used in this table.

N	Atoms	Twist angle (°)	$\bar{d}_{interlayer}$(Å)
N/A	256	0	3.66
1	28	21.79	3.84
2	76	13.17	3.81
3	148	9.43	3.79
6	508	6.01	3.78
11	1588	3.15	3.81
16	3268	2.13	3.80
21	5548	1.61	3.79
26	8428	1.30	3.80
31	11908	1.08	3.80
36	15988	0.93	3.80
41	20668	0.82	3.80
43	22708	0.76	3.78
44	23764	0.74	3.80
45	24844	0.73	3.80

of information. On one hand, phonon dispersion relations can indicate whether a structure is dynamically stable or unstable based on the absence or presence of imaginary frequencies. Figures 8, 9 and 10 show the dispersion relations of bilayer hBN for twist angles 21.79 °, 13.17 ° and 9.43 ° respectively. The most striking differences between fig. 8 and 9 are the number of curves and the increase of the highest frequency from 41 THz to 44 THz. The former is a consequence of the size: each atom in a system contributes three curves, resulting in a higher number of dispersion curves for larger systems. As was shown in eq. 1 in Sect. 2.3, smaller twist angles require larger structures, explaining this difference.

However, the latter difference, i.e., the increase of frequency for the specific phonons, is not as evident. One plausible explanation is the existence of a phase transition at a certain twist angle, with the structure in fig. 8 undergoing it while finished in fig. 9. Phonon softening, as mentioned in Sect. 2, is a phenomenon where phonon frequencies are lowered and could indicate an upcoming transition. On top of that, certain phase transitions are associated with transverse optical phonons [48], which correspond to one of the modes for high frequency phonons. For these reasons, it is possible that the structure is going through a phase transition starting at a twist angle bigger than 13.17°. While these results do not pinpoint the exact type of phase transition, we would expect a transition from insulating to ferroelectric state, since it has been experimentally shown that it should happen, although with small twist angles [51]. Further research in this direction is needed to prove this hypothesis. The phonon dispersion relation for bilayer hBN with a twist angle of 9.43°, presented in

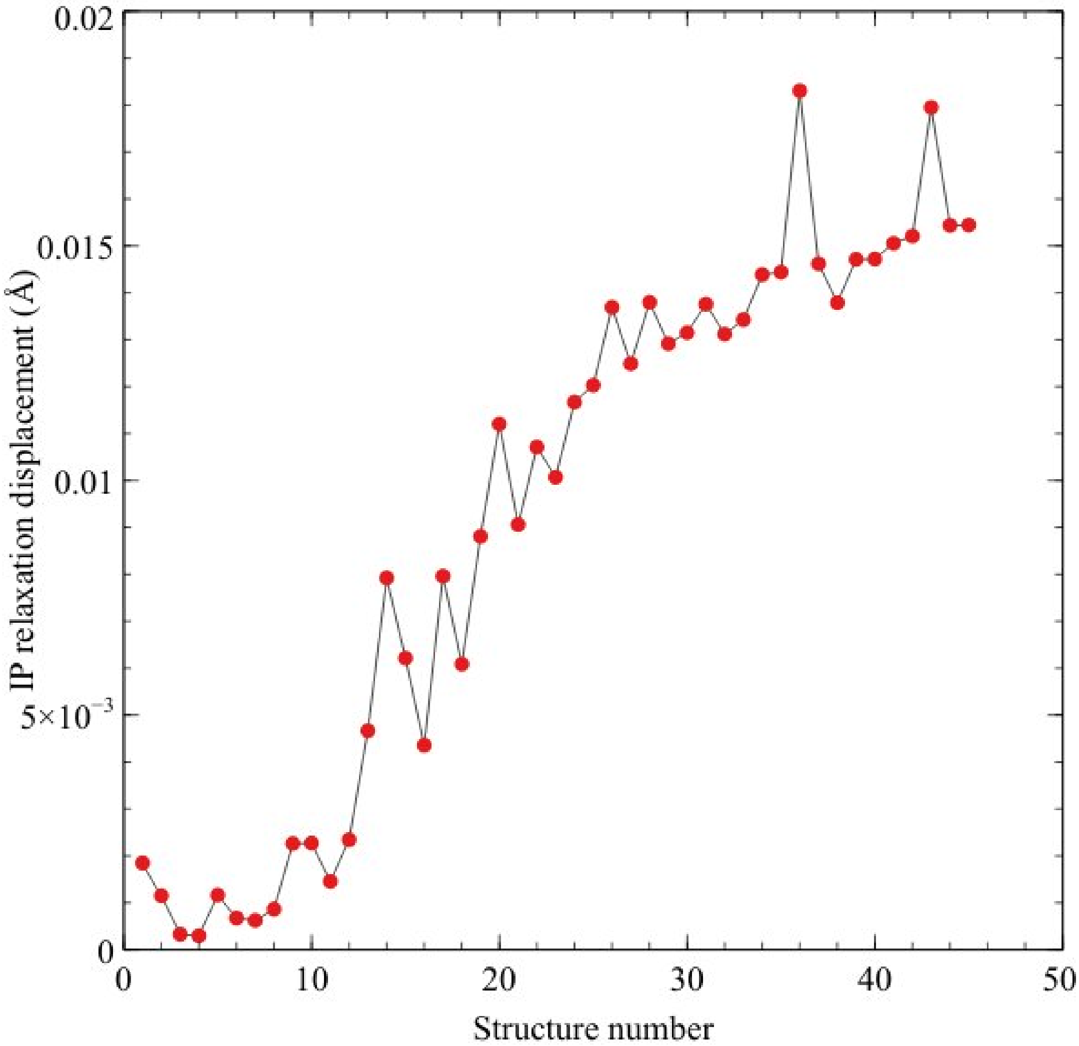


Figure 4: The average IP displacement computed as the difference between z-position of each atom before and after relaxation. Twist angle decreases as structure number increases, starting from 21.79 ° and ending in 0.73 °

fig. 10, also supports the above explanation. It has the higher frequency phonons, which agrees with the phase transition being over and the phonons returning to their original energy level. However, this is not the only explanation. The structure could also have become mechanically more rigid, which could also explain the increase in the phonon frequencies.

As far as the phonon frequencies are concerned, both fig. 9 and 10 share similarities and seem to be the same, although they are somewhat hard to decipher. It is worth noting that all these phonon dispersion relations exhibit flat phonon bands, starting from the Γ, K and M points. However, seeing as all the phonon dispersion relations also have some imaginary frequencies, it does cast some doubt on the reliability of the results. As ML techniques in general are approximations and the magnitude of the imaginary frequencies, $0.1 - 0.6$ THz, is small, this is not a major concern. Post-hoc

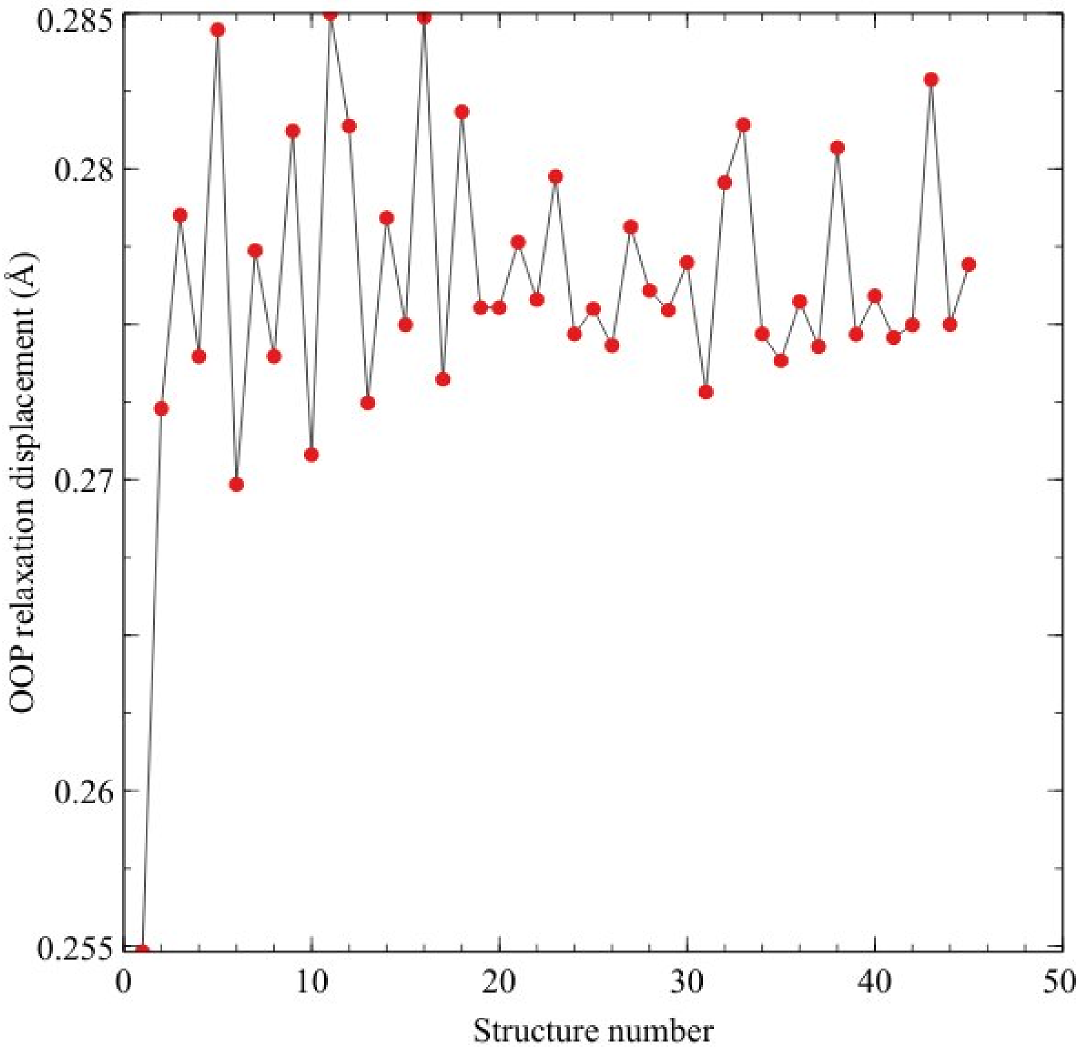


Figure 5: The average OOP displacement during relaxation with ML potential. These results indicate a decrease in interlayer distance after relaxation.

methods have been suggested to remove the imaginary frequencies [35]. In this thesis, post-hoc methods were not incorporated and the imaginary frequencies are present.

Finally, the phonon dispersion relation of a monolayer hBN was computed to have a better understanding of the effect of the vdW forces between layers. The monolayers themselves only contribute to six curves in total, starting from 0, 23.7 and 26.7 THz. This is presented in Fig. 11.

On the other hand, we performed an LDOS analysis to obtain more information about the material. We found differences between a regularly stacked, in this case AA-stacked, and a twisted hBN. In AA-stacked hBN, there are four different groupings of atoms present, as shown in Fig. 12: two nitrogens and two borons.

Twisting layers reduces the symmetry of the system; therefore, they cannot be grouped up in a similar fashion as was the case in fig. 12. As a consequence, there are 28 different curves, one per atom. However, it is hard to differentiate them due to the

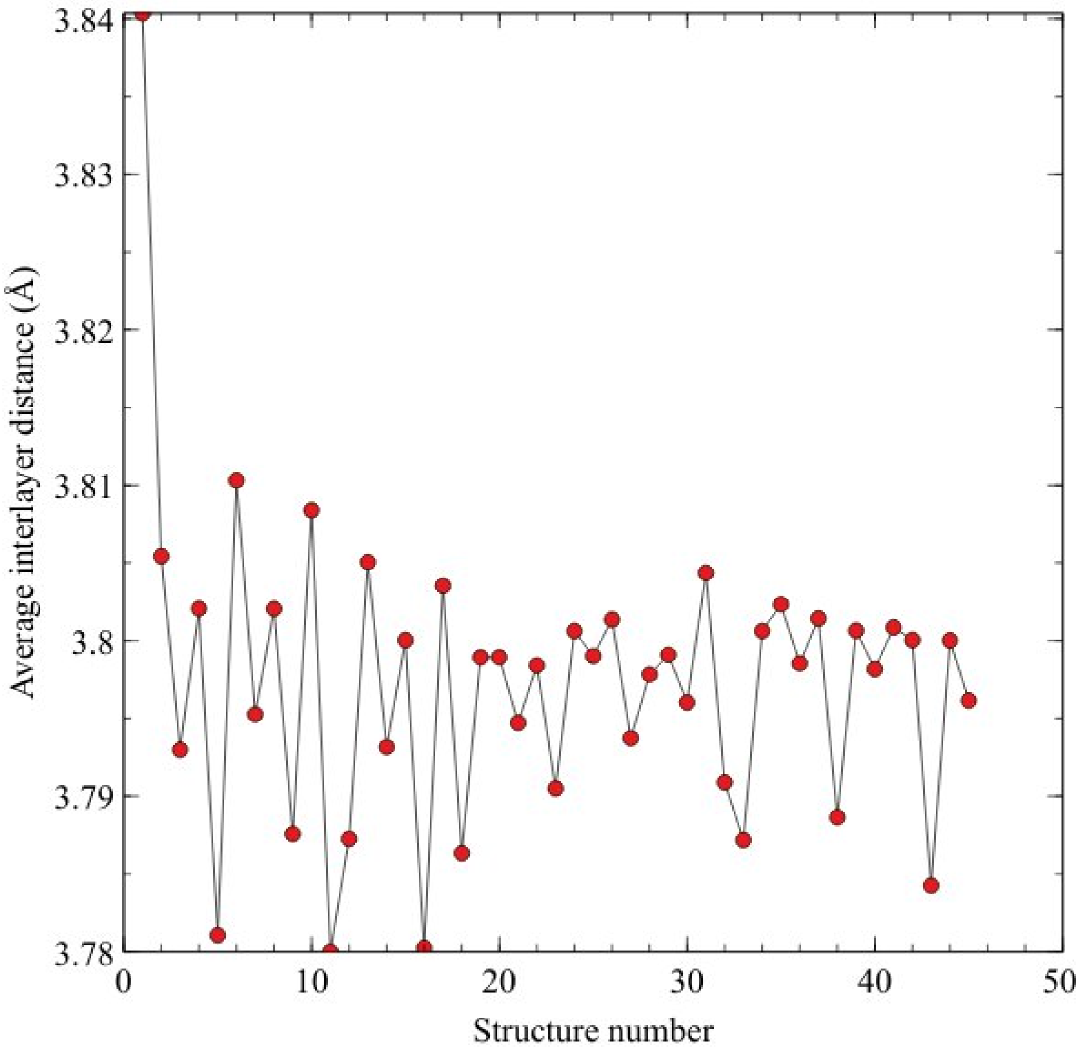
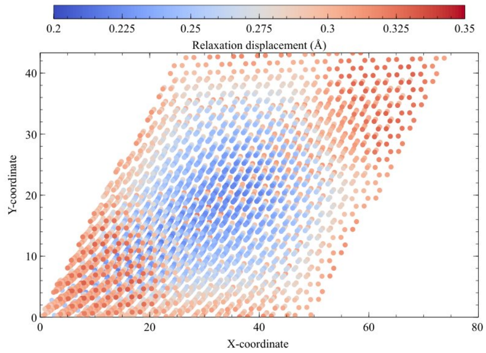
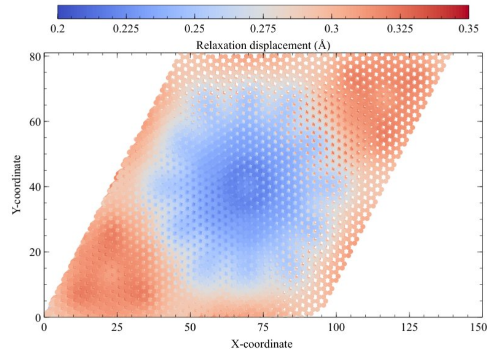


Figure 6: Average interlayer distance after GAP-driven relaxation of twisted hBN bilayers with varying twist angles.

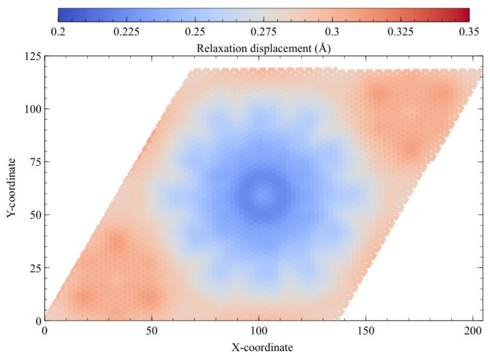
energy levels being close to each other. This is shown in fig. 13. The LDOS analysis of the twisted system suggests that its energy states are more split in comparison to the stacked structure, clearly indicated by the amount of local maxima in Fig. 13. Another point of interest is the high spike at near-zero frequency. There is a drastic difference at that energy level between the stacked and twisted systems. This supports a large difference between the systems. The high spike could partly be explained by the system having energy states too close to each other to be resolved by our ML potential, being grouped together as a result.



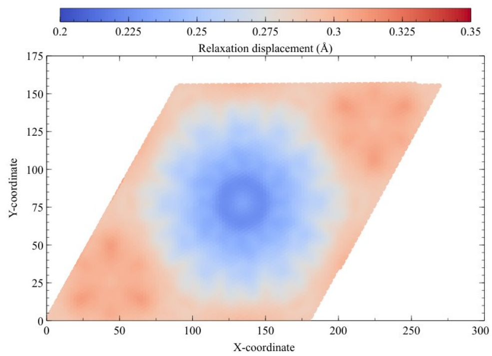
(a) Angle 3.15 °



(b) Angle: 1.61 °.



(c) Angle 1.08 °



(d) Angle 0.82 °

Figure 7: Displacement heatmaps during relaxation for twisted bilayer hBN with different angles. Red color signifies a large atom-wise interlayer displacement, while blue signifies a smaller displacement.

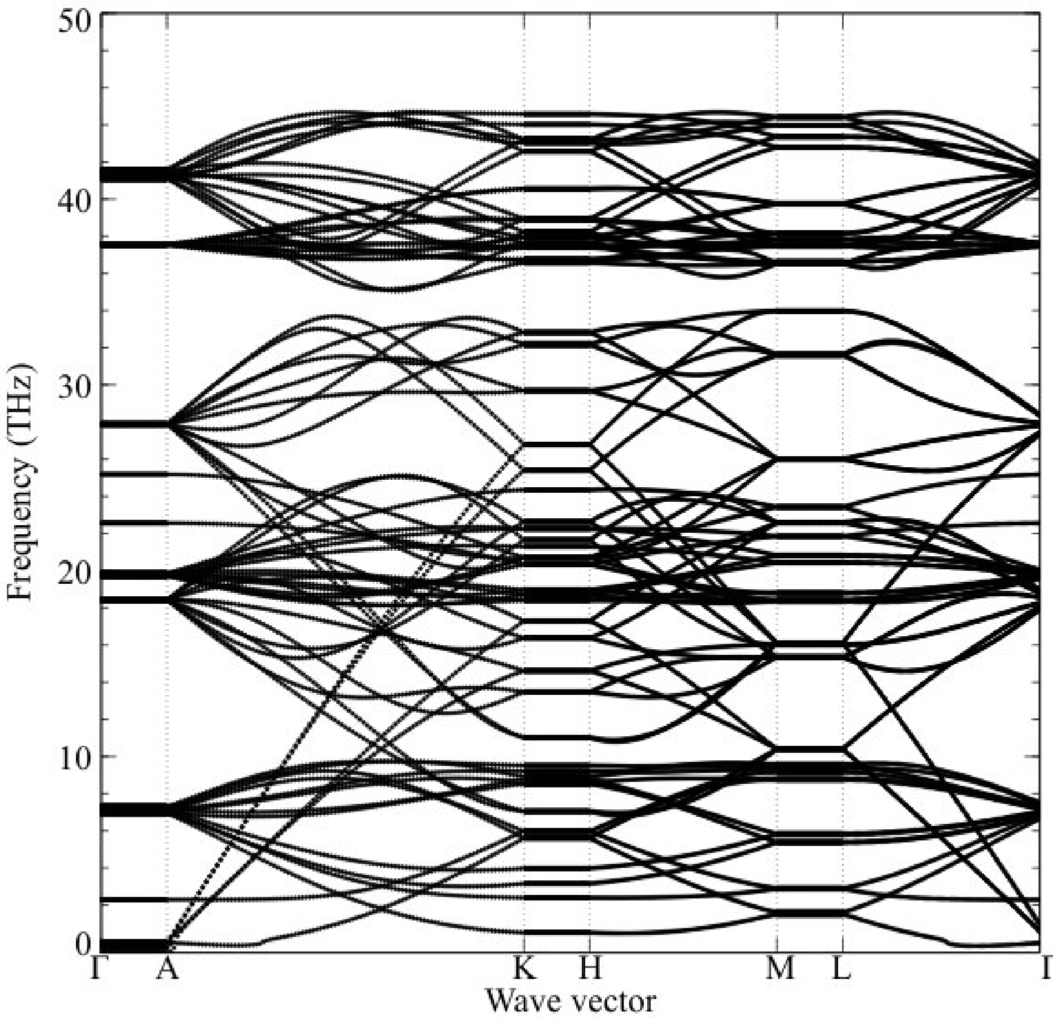


Figure 8: Phonon dispersion relations of a hBN bilayer with 21.79° twist angle, consisting of 28 atoms.

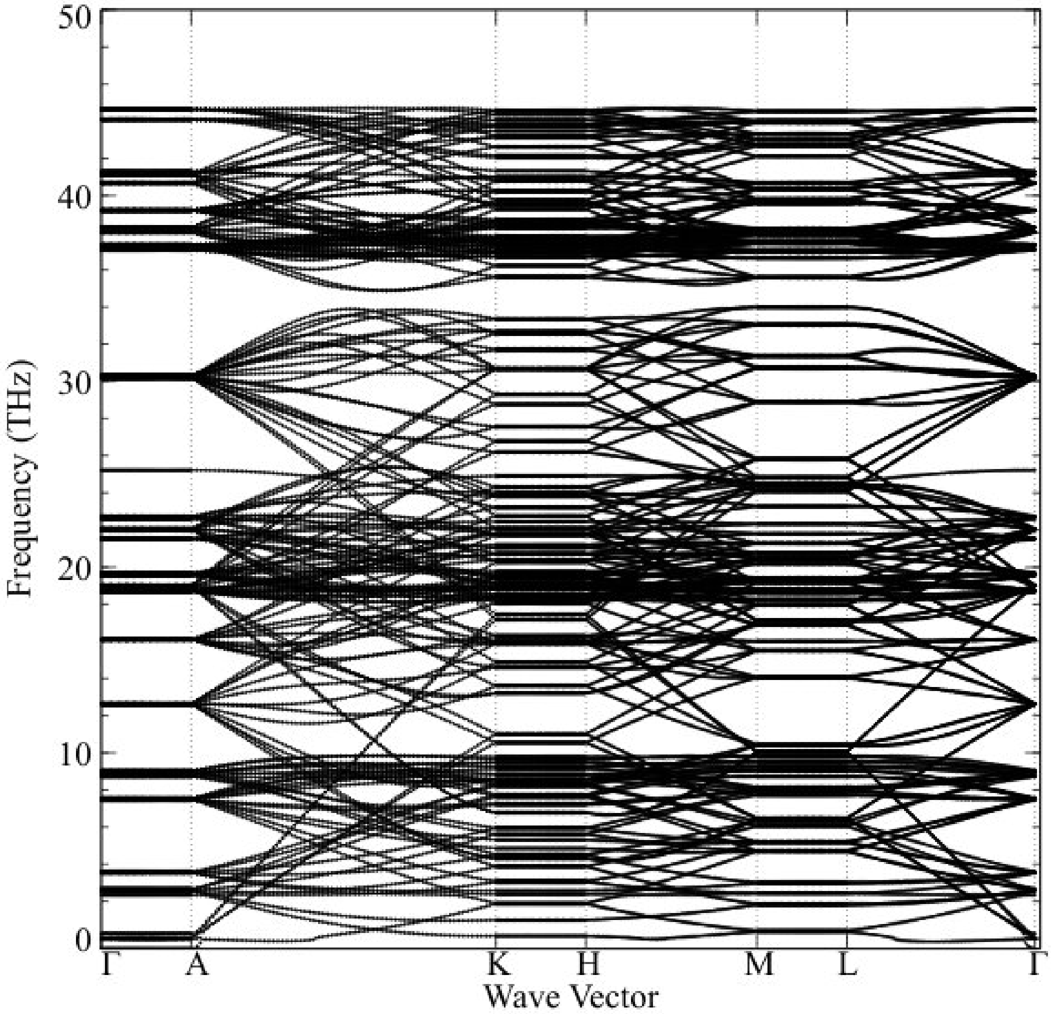


Figure 9: Phonon dispersion relations of a hBN bilayer with 13.17° twist angle, consisting of 76 atoms.

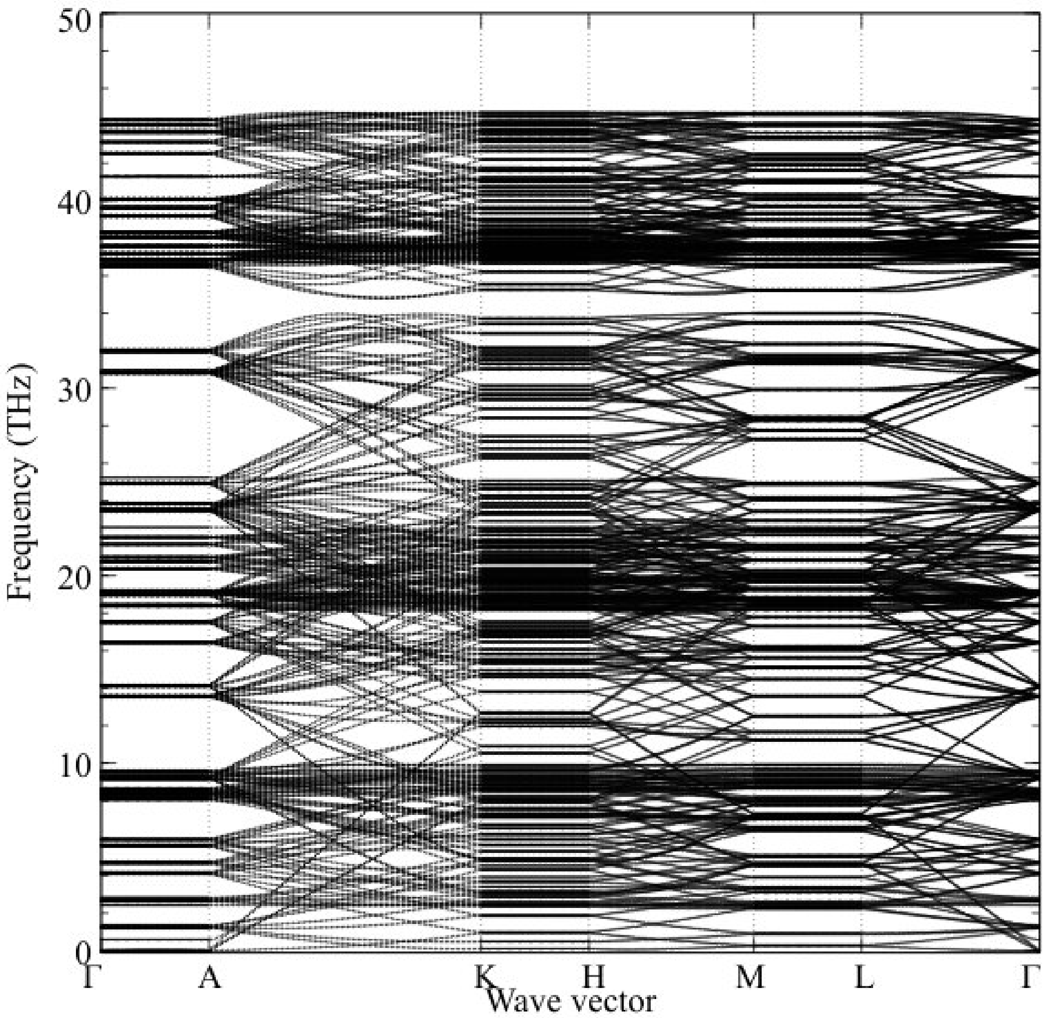


Figure 10: Phonon dispersion relations of a hBN bilayer with 9.43° twist angle, consisting of 148 atoms.

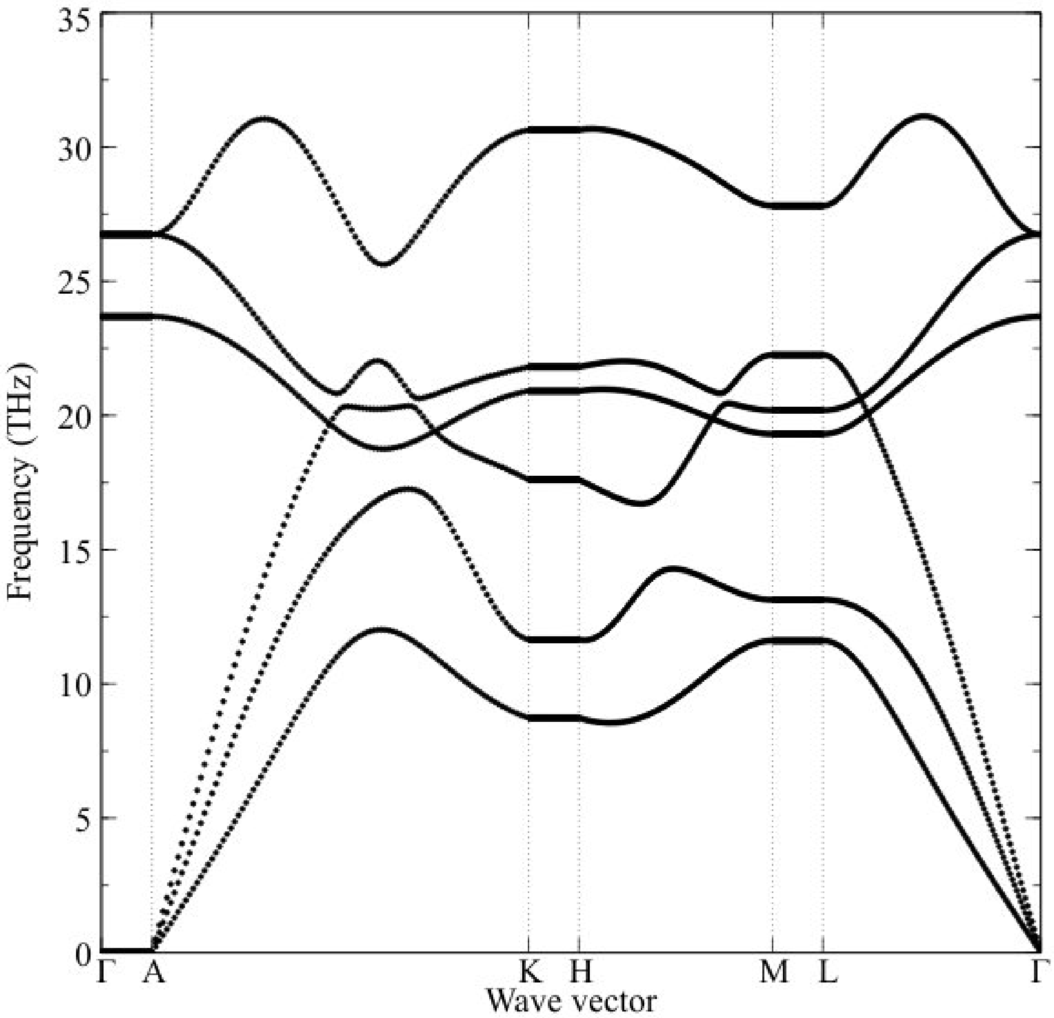


Figure 11: Phonon dispersion relations of monolayer hBN.

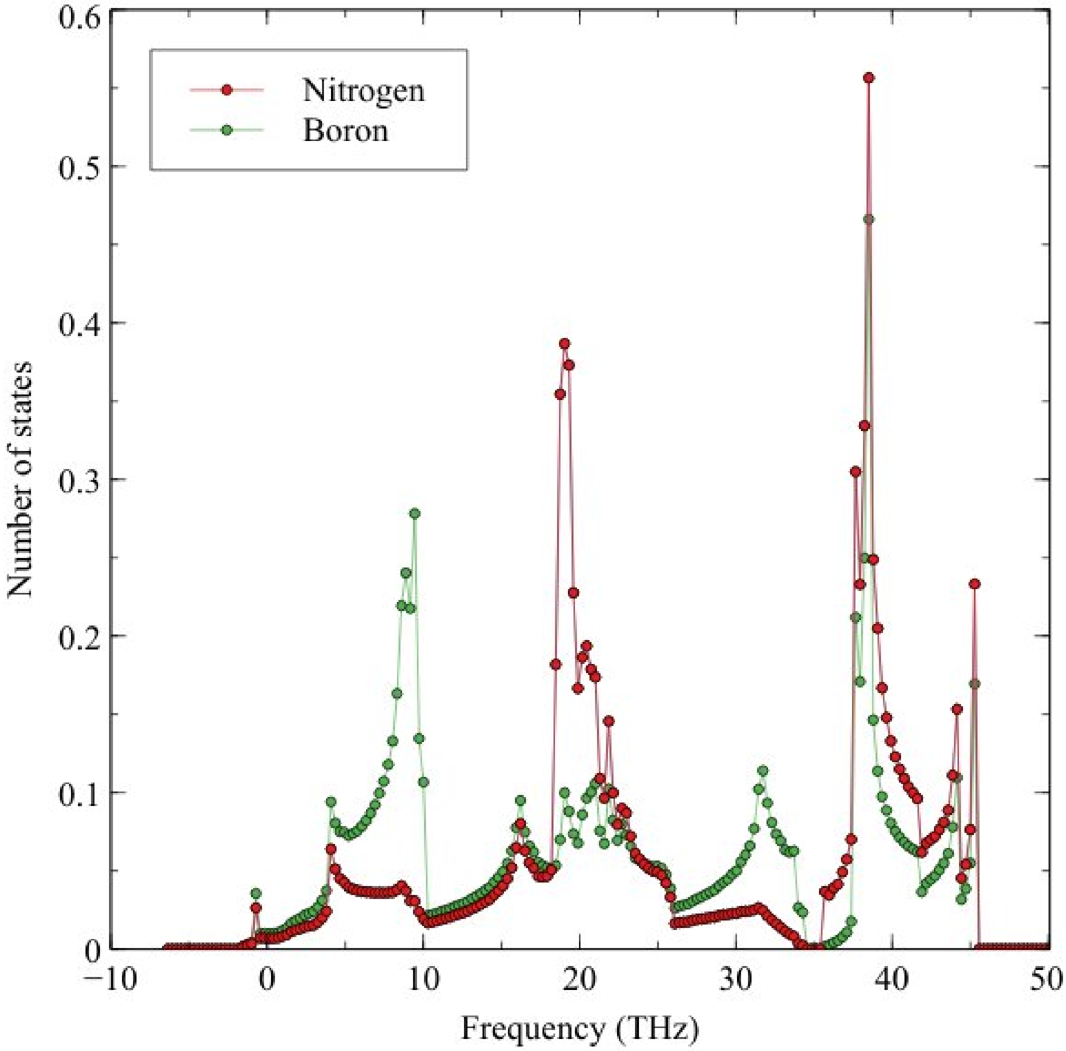


Figure 12: LDOS of an 8x8 bilayer AA-stacked hBN. Green curves correspond to boron atoms, while red curves to nitrogen ones. Note that there are two of each, from top and bottom layers, but stacked on top of each other.

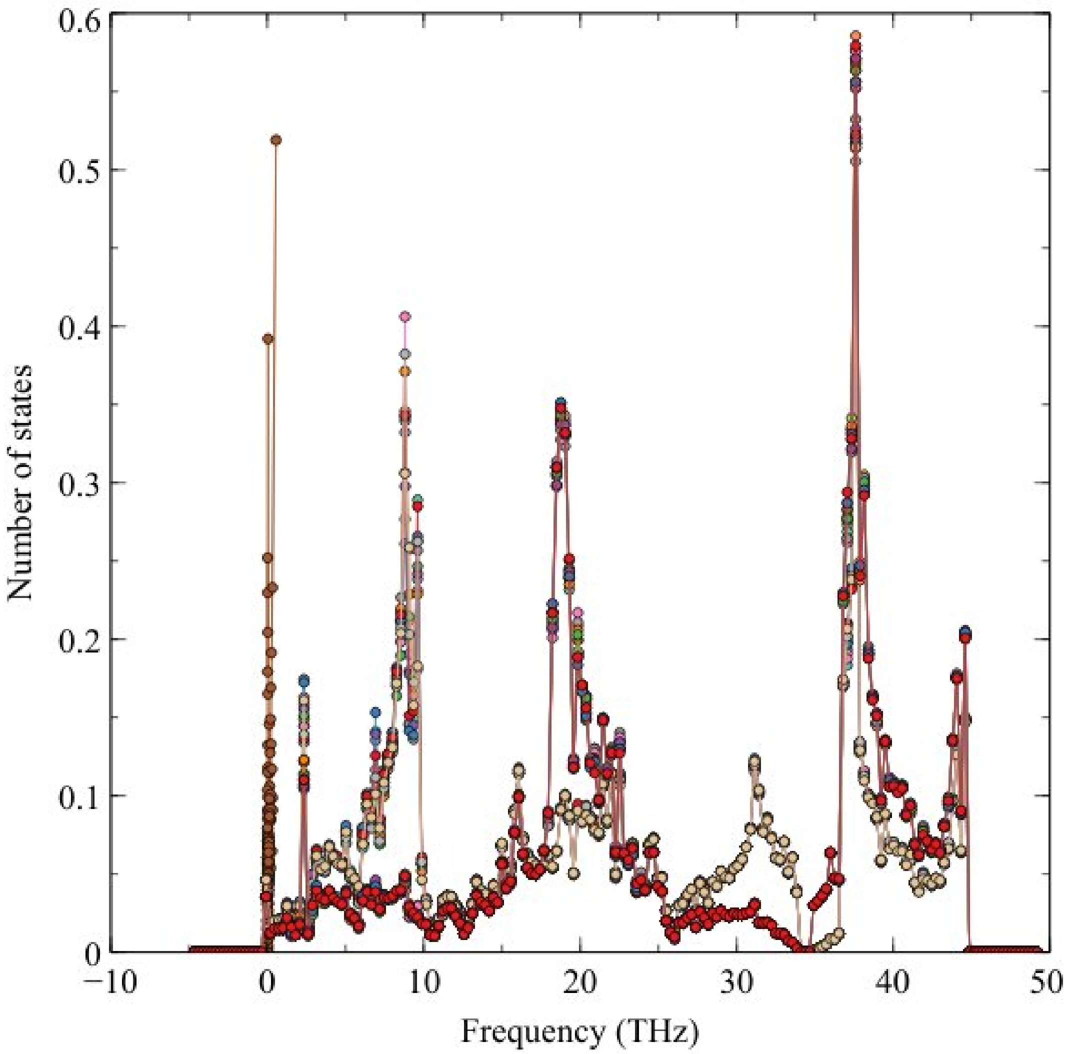


Figure 13: LDOS of bilayer hBN with a 21.79° twist angle. Note that not all curves are visible due to stacking and limitations in distinct colors.

5 Summary/Conclusions

In this study, we used an ML approach to study the atomic displacements, interlayer distances, phonon dispersion relations and the energetics of bilayer hBN structures with varying twist angles. We also compared the LDOS of AA-stacked and twisted hBN. The ML potential was tested and we concluded that its performance is accurate enough. There is some margin of error in comparison to DFT calculations. As an extra performance test, the phonon dispersion relations were computed. The ML-generated phonon dispersion relations do not show significant imaginary frequencies, which supports the reliability of the ML potential. There is substantial amount of linearity in the phonon dispersion relations. Most are related to the acoustic branch, but some are in the optical branch too.

As far as the displacements are concerned, there is some difference between the interlayer distance of AA-stacked and the twisted structures. The domain formation of smaller and larger interlayer distances for twisted structures tells us that there is some force in play. A possible explanation is the presence of polarization or polarized domains, which is a prerequisite of ferroelectric domains.

A previous study showed the emergence of ferroelectricity for a single low twist angle, but it failed to answer at which angle the phase transition from non-ferroelectric into ferroelectric phase starts [51]. While the results are not conclusive, they do indicate that as the twist angle decreases, the changes in the structure become stronger. This could indicate that the polarization is stronger as the twist angle is lowered. Due to the prerequisite of moiré pattern, not every angle could be computed. Therefore, we could not find the exact angle at which the phase transition starts, but this study does show that there is a specific angle, namely around $2-3°$, where some change happens. This can be seen in the increase of IP-displacement. The OOP-displacement seem to fluctuate around 0.27 Å. This fluctuation seems to diminish as the twist angle is lowered.

The last part of this study was the LDOS comparison of twisted vs. stacked structure, which resulted in the anticipated outcomes. There is a significant difference between their results. The Stacked structure, due to its high degree of symmetry, had fewer available energy states, while the twisted structure had clearly more division.

The findings of this study do merit a broader study. One approach would be to improve the ML potential to increase the accuracy of it, either disproving or substantiating the findings of this study. The same ML potential could also be used to find other signs of phase transition. If the findings of this research are further supported, we could apply it to extended hBN systems with more layers. This could be done by either adding layers one by one, changing which layer is twisted and by what angle, as well as having bilayer on top of bilayer with different twist angles to see the effects of increasing the dimensionality of the effect. It could be worth looking closer into selected twist angles showing irregularities, e.g., 2.13°. A closer look into the highest twist angle is of high interest, due to the presence of high frequency phonons. In addition to the aforementioned suggestions, the LDOSs and phonon dispersion relations could be studied with higher precision, to be able to differentiate between curves more reliably and effectively and find if there are more differences than those seen here. The ML potential itself could be improved, training it more and including more diverse data in the training.

References

[1] Sajjad Ali, Pir Muhammad Ismail, Muhammad Humayun, and Mohamed Bououdina. Hexagonal boron nitride: From fundamentals to applications. *Desalination*, 599:118442, April 2025. ISSN 0011-9164. doi: 10.1016/j.desal.2024.118442. URL http://dx.doi.org/10.1016/j.desal.2024.118442.

[2] Eva Y. Andrei, Dmitri K. Efetov, Pablo Jarillo-Herrero, Allan H. MacDonald, Kin Fai Mak, T. Senthil, Emanuel Tutuc, Ali Yazdani, and Andrea F. Young. The marvels of moiré materials. *Nature Reviews Materials*, 6(3):201–206, March 2021. ISSN 2058-8437. doi: 10.1038/s41578-021-00284-1. URL http://dx.doi.org/10.1038/s41578-021-00284-1.

[3] Ashim Kumar Bain and Prem Chand. *Ferroelectrics: Principles and Applications*. Wiley, February 2017. ISBN 9783527805310. doi: 10.1002/9783527805310. URL http://dx.doi.org/10.1002/9783527805310.

[4] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.*, 104:136403, Apr 2010. doi: 10.1103/PhysRevLett.104.136403. URL https://link.aps.org/doi/10.1103/PhysRevLett.104.136403.

[5] Albert P. Bartók, Risi Kondor, and Gábor Csányi. On representing chemical environments. *Phys. Rev. B*, 87:184115, May 2013. doi: 10.1103/PhysRevB.87.184115. URL https://link.aps.org/doi/10.1103/PhysRevB.87.184115.

[6] Yuan Cao, Valla Fatemi, Shiang Fang, Kenji Watanabe, Takashi Taniguchi, Efthimios Kaxiras, and Pablo Jarillo-Herrero. Unconventional superconductivity in magic-angle graphene superlattices. *Nature*, 556(7699):43–50, March 2018. ISSN 1476-4687. doi: 10.1038/nature26160. URL http://dx.doi.org/10.1038/nature26160.

[7] Miguel A. Caro. Optimizing many-body atomic descriptors for enhanced computational performance of machine learning based interatomic potentials. *Physical Review B*, 100(2), July 2019. ISSN 2469-9969. doi: 10.1103/physrevb.100.024112. URL http://dx.doi.org/10.1103/PhysRevB.100.024112.

[8] R.M. Chrenko. Ultraviolet and infrared spectra of cubic boron nitride. *Solid State Communications*, 14(6):511–515, March 1974. ISSN 0038-1098. doi: 10.1016/0038-1098(74)90978-8. URL http://dx.doi.org/10.1016/0038-1098(74)90978-8.

[9] Gábor Csányi, Steven Winfield, J R Kermode, A De Vita, Alessio Comisso, Noam Bernstein, and Michael C Payne. Expressive programming for computational physics in fortran 95+. *IoP Comput. Phys. Newsletter*, page Spring 2007, 2007. URL https://github.com/libAtoms/QUIP.

[10] Beatriz G. del Rio, Brandon Phan, and Rampi Ramprasad. A deep learning framework to emulate density functional theory. *npj Computational Materials*, 9(1), August 2023. ISSN 2057-3960. doi: 10.1038/s41524-023-01115-3. URL http://dx.doi.org/10.1038/s41524-023-01115-3.

[11] Volker L. Deringer, Albert P. Bartók, Noam Bernstein, David M. Wilkins, Michele Ceriotti, and Gábor Csányi. Gaussian process regression for materials and molecules. *Chemical Reviews*, 121(16):10073–10141, August 2021. ISSN 1520-6890. doi: 10.1021/acs.chemrev.1c00022. URL http://dx.doi.org/10.1021/acs.chemrev.1c00022.

[12] Miguel A. Caro et al. Turbogap. https://github.com/mcaroba/turbogap. Accessed: 2025-07-25.

[13] Shuting Fu, Yiliang Guo, Tongbiao Wang, Tianbao Yu, and Qinghua Liao. Spontaneous emission controlled by the twisted -moo3 bilayers. *Optics amp; Laser Technology*, 192:113412, December 2025. ISSN 0030-3992. doi: 10.1016/j.optlastec.2025.113412. URL http://dx.doi.org/10.1016/j.optlastec.2025.113412.

[14] Yohanes S. Gani, Hadar Steinberg, and Enrico Rossi. Superconductivity in twisted graphene nbse2 heterostructures. *Physical Review B*, 99(23), June 2019. ISSN 2469-9969. doi: 10.1103/physrevb.99.235404. URL http://dx.doi.org/10.1103/PhysRevB.99.235404.

[15] S Matt Gilbert, Thang Pham, Mehmet Dogan, Sehoon Oh, Brian Shevitski, Gabe Schumm, Stanley Liu, Peter Ercius, Shaul Aloni, Marvin L Cohen, and Alex Zettl. Alternative stacking sequences in hexagonal boron nitride. *2D Materials*, 6(2):021006, March 2019. ISSN 2053-1583. doi: 10.1088/2053-1583/ab0e24. URL http://dx.doi.org/10.1088/2053-1583/ab0e24.

[16] Feng He, Yongjian Zhou, Zefang Ye, Sang-Hyeok Cho, Jihoon Jeong, Xianghai Meng, and Yaguo Wang. Moiré patterns in 2d materials: A review. *ACS Nano*, 15 (4):5944–5958, March 2021. ISSN 1936-086X. doi: 10.1021/acsnano.0c10435. URL http://dx.doi.org/10.1021/acsnano.0c10435.

[17] Friedrich Hensel and William W. Jr. Warren. 2.3.3.3 self-trapping (polaron) effects, 1999. URL https://app.knovel.com/hotlink/khtml/id:kt00UQGCR3/fluid-metals-liquid-vapor/self-trapping-polaron.

[18] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen,

Michael Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, June 2017. ISSN 1361-648X. doi: 10.1088/1361-648x/aa680e. URL http://dx.doi.org/10.1088/1361-648X/aa680e.

[19] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor. *An Introduction to Statistical Learning: with Applications in Python*. Springer International Publishing, 2023. ISBN 9783031387470. doi: 10.1007/978-3-031-38747-0. URL http://dx.doi.org/10.1007/978-3-031-38747-0.

[20] Jinhyuk Jang and Si-Young Choi. Reduced dimensional ferroelectric domains and their characterization techniques. *Microstructures*, 4(2), 2024. doi: 10.20517/microstructures.2023.67. URL http://dx.doi.org/10.20517/microstructures.2023.67.

[21] Alexander Jung. *Machine Learning: The Basics*. Springer Nature Singapore, 2022. ISBN 9789811681936. doi: 10.1007/978-981-16-8193-6. URL http://dx.doi.org/10.1007/978-981-16-8193-6.

[22] Jun Kang, Jingbo Li, Shu-Shen Li, Jian-Bai Xia, and Lin-Wang Wang. Electronic structural moiré pattern effects on mos2/mose2 2d heterostructures. *Nano Letters*, 13(11):5485–5490, October 2013. ISSN 1530-6992. doi: 10.1021/nl4030648. URL http://dx.doi.org/10.1021/nl4030648.

[23] Pathan Mohsin Khan, Gopala Krishna Jillella, and Kunal Roy. *Recent advancements in QSAR and machine learning approaches for risk assessment of organic chemicals*, page 167–185. Elsevier, 2024. ISBN 9780443153396. doi: 10.1016/b978-0-443-15339-6.00035-7. URL http://dx.doi.org/10.1016/B978-0-443-15339-6.00035-7.

[24] Sascha Klawohn, James P. Darby, James R. Kermode, Gábor Csányi, Miguel A. Caro, and Albert P. Bartók. Gaussian approximation potentials: Theory, software implementation and application examples. *The Journal of Chemical Physics*, 159(17), November 2023. ISSN 1089-7690. doi: 10.1063/5.0160898. URL http://dx.doi.org/10.1063/5.0160898.

[25] Patrick Knüppel, Jiacheng Zhu, Yiyu Xia, Zhengchao Xia, Zhongdong Han, Yihang Zeng, Kenji Watanabe, Takashi Taniguchi, Jie Shan, and Kin Fai Mak. Correlated states controlled by a tunable van hove singularity in moiré wse2 bilayers. *Nature Communications*, 16(1), February 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-57235-5. URL http://dx.doi.org/10.1038/s41467-025-57235-5.

[26] Kurt W. Kolasinski. 1.1.8 advanced topic: Stacked two-dimensional materials and moiré superlattices, 2020. URL https://app.knovel.com/hotlink/khtml/id:kt0137SIL1/surface-science-foundations/advanced-topic-stacked.

[27] Zbigniew Kozioł, Grzegorz Gawlik, and Jacek Jagielski. Van der waals interlayer potential of graphitic structures: From lennard–jones to kolmogorov–crespy and lebedeva models. *Chinese Physics B*, 28(9):096101, September 2019. ISSN 1674-1056. doi: 10.1088/1674-1056/ab38a5. URL http://dx.doi.org/10.1088/1674-1056/ab38a5.

[28] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, July 1996. ISSN 0927-0256. doi: 10.1016/0927-0256(96)00008-0. URL http://dx.doi.org/10.1016/0927-0256(96)00008-0.

[29] G. Kresse and J. Furthmüller. Efficient iterative schemes forab initiototal-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16): 11169–11186, October 1996. ISSN 1095-3795. doi: 10.1103/physrevb.54.11169. URL http://dx.doi.org/10.1103/PhysRevB.54.11169.

[30] G. Kresse and J. Hafner. Ab initiomolecular dynamics for liquid metals. *Physical Review B*, 47(1):558–561, January 1993. ISSN 1095-3795. doi: 10.1103/physrevb.47.558. URL http://dx.doi.org/10.1103/PhysRevB.47.558.

[31] Chun Ning Lau, Marc W. Bockrath, Kin Fai Mak, and Fan Zhang. Reproducibility in the fabrication and physics of moiré materials. *Nature*, 602(7895):41–50, February 2022. ISSN 1476-4687. doi: 10.1038/s41586-021-04173-z. URL http://dx.doi.org/10.1038/s41586-021-04173-z.

[32] Guohong Li, A. Luican, J. M. B. Lopes dos Santos, A. H. Castro Neto, A. Reina, J. Kong, and E. Y. Andrei. Observation of van hove singularities in twisted graphene layers. *Nature Physics*, 6(2):109–113, November 2009. ISSN 1745-2481. doi: 10.1038/nphys1463. URL http://dx.doi.org/10.1038/nphys1463.

[33] Wilson Nieto Luna, Robin Smeyers, Cem Sevik, Lucian Covaci, and Milorad V. Milošević. Machine-learning interatomic potential for twisted hexagonal boron nitride: Accurate structural relaxation and emergent polarization, 2025. URL https://arxiv.org/abs/2503.11797.

[34] James G. McHugh, Vladimir V. Enaldiev, and Vladimir I. Fal'ko. Moiré superstructures in marginally twisted nbse2 bilayers. *Physical Review B*, 108 (22), December 2023. ISSN 2469-9969. doi: 10.1103/physrevb.108.224111. URL http://dx.doi.org/10.1103/PhysRevB.108.224111.

[35] Ioanna Pallikara, Prakriti Kayastha, Jonathan M Skelton, and Lucy D Whalley. The physical significance of imaginary phonon modes in crystals. *Electronic Structure*, 4(3):033002, July 2022. ISSN 2516-1075. doi: 10.1088/2516-1075/ac78b3. URL http://dx.doi.org/10.1088/2516-1075/ac78b3.

[36] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, November 2005. ISBN 9780262256834.

doi: 10.7551/mitpress/3206.001.0001. URL http://dx.doi.org/10.7551/mitpress/3206.001.0001.

[37] R. M. Ribeiro and N. M. R. Peres. Stability of boron nitride bilayers: Ground-state energies, interlayer distances, and tight-binding description. *Physical Review B*, 83(23), June 2011. ISSN 1550-235X. doi: 10.1103/physrevb.83.235312. URL http://dx.doi.org/10.1103/PhysRevB.83.235312.

[38] Emmanuel Rosencher and Borge Vinter. 6 electronic properties of semiconductors, 2002. URL https://app.knovel.com/hotlink/khtml/id:kt006APK91/optoelectronics/boltzmanns-equation.

[39] Atsuto Seko and Atsushi Togo. Projector-based efficient estimation of force constants. *Physical Review B*, 110(21), December 2024. ISSN 2469-9969. doi: 10.1103/physrevb.110.214302. URL http://dx.doi.org/10.1103/PhysRevB.110.214302.

[40] Cong Su, Eli Janzen, Mingze He, Chi Li, Alex Zettl, Joshua D. Caldwell, James H. Edgar, and Igor Aharonovich. Fundamentals and emerging optical applications of hexagonal boron nitride: a tutorial. *Advances in Optics and Photonics*, 16(2):229, May 2024. ISSN 1943-8206. doi: 10.1364/aop.502922. URL http://dx.doi.org/10.1364/AOP.502922.

[41] Jie Sun, Yiming Li, Di Hu, Bowen Shen, Boyang Zhang, Zilong Wang, Haiyue Tang, and Anquan Jiang. Roadmap for ferroelectric domain wall memory. *Microstructures*, 4(1), January 2024. ISSN 2770-2995. doi: 10.20517/microstructures.2023.52. URL http://dx.doi.org/10.20517/microstructures.2023.52.

[42] Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. in 't Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, Ray Shan, Mark J. Stevens, Julien Tranchida, Christian Trott, and Steven J. Plimpton. Lammps - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, 271:108171, February 2022. ISSN 0010-4655. doi: 10.1016/j.cpc.2021.108171. URL http://dx.doi.org/10.1016/j.cpc.2021.108171.

[43] Zhiwen Tian, Dawei He, Mohan Huang, Wenwen Wu, Yinglin Zhang, Xiaojing Liu, Fangying Ren, Jiarong Wang, Guili Li, Kun Zhao, Yongsheng Wang, and Xiaoxian Zhang. Regulating transient optical responses in twisted bilayer ws2. *Nanotechnology*, 36(18):185202, March 2025. ISSN 1361-6528. doi: 10.1088/1361-6528/adc311. URL http://dx.doi.org/10.1088/1361-6528/adc311.

[44] Atsushi Togo. phonopy-dos. https://phonopy.github.io/phonopy/index.html, 2009. Accessed: 2025-07-25.

[45] Atsushi Togo. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.*, 92(1):012001, 2023. doi: 10.7566/JPSJ.92.012001.

[46] Atsushi Togo and Isao Tanaka. First principles phonon calculations in materials science. *Scripta Materialia*, 108:1–5, November 2015. ISSN 1359-6462. doi: 10.1016/j.scriptamat.2015.07.021. URL http://dx.doi.org/10.1016/j.scriptamat.2015.07.021.

[47] Atsushi Togo, Laurent Chaput, Terumasa Tadano, and Isao Tanaka. Implementation strategies in phonopy and phono3py. *J. Phys. Condens. Matter*, 35(35): 353001, 2023. doi: 10.1088/1361-648X/acd831.

[48] Urszula D. Wdowik, Krzysztof Parlinski, Stéphane Rols, and Tapan Chatterji. Soft-phonon mediated structural phase transition in gete. *Physical Review B*, 89 (22), June 2014. ISSN 1550-235X. doi: 10.1103/physrevb.89.224306. URL http://dx.doi.org/10.1103/PhysRevB.89.224306.

[49] Dong-Hui Xu, Jie Yuan, Zi-Jian Yao, Yi Zhou, Jin-Hua Gao, and Fu-Chun Zhang. Stacking order, interaction, and weak surface magnetism in layered graphene sheets. *Physical Review B*, 86(20), November 2012. ISSN 1550-235X. doi: 10.1103/physrevb.86.201404. URL http://dx.doi.org/10.1103/PhysRevB.86.201404.

[50] Qiaoling Xu, Yuzheng Guo, and Lede Xian. Moiré flat bands in twisted 2d hexagonal vdw materials. *2D Materials*, 9(1):014005, December 2021. ISSN 2053-1583. doi: 10.1088/2053-1583/ac3a98. URL http://dx.doi.org/10.1088/2053-1583/ac3a98.

[51] Kenji Yasuda, Xirui Wang, Kenji Watanabe, Takashi Taniguchi, and Pablo Jarillo-Herrero. Stacking-engineered ferroelectricity in bilayer boron nitride. *Science*, 372(6549):1458–1462, June 2021. ISSN 1095-9203. doi: 10.1126/science.abd3230. URL http://dx.doi.org/10.1126/science.abd3230.

[52] Lei Zhou, Xiang Ni, Zerui Wang, Enrico M. Renzi, Junbo Xu, Zhou Zhou, Yu Yin, Yanzhen Yin, Renkang Song, Zhichen Zhao, Ke Yu, Di Huang, Zhanshan Wang, Xinbin Cheng, Andrea Alù, and Tao Jiang. Engineering shear polaritons in 2d twisted heterostructures. *Nature Communications*, 16 (1), March 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-58197-4. URL http://dx.doi.org/10.1038/s41467-025-58197-4.

[53] J.M. Ziman. *Principles of the Theory of Solids*, page 27–76. Cambridge University Press, July 1972. ISBN 9781139644075. doi: 10.1017/cbo9781139644075.004. URL http://dx.doi.org/10.1017/CBO9781139644075.004.