

NOVEL APPROACH FOR ELECTRONIC STRUCTURE
OF NON-PERIODIC MATTER

A Dissertation

Presented to the Physics Department
of Cornell University

by

Brandon Li

May 2024

ACKNOWLEDGEMENTS

I would like to extend my deep gratitude towards my advisor Prof. Tomás Arias who has been a wonderful teacher, mentor, and collaborator throughout my undergraduate years. I feel I owe much of my intellectual development as a physicist and researcher to him and he has inspired me to be passionate in everything to do with physics. Working with Prof. Arias has shown me what it means for someone to care about his students as much as his research, and I think that is very important.

I would also like to thank Drake Niedzielski who has provided much support and guidance on this project and who I have had many interesting conversations with. It is always a great feeling to work with someone who is as equally as motivated and passionate.

Finally, I would like to thank my close friends and especially my partner for providing me with ideas and motivating me to complete this work. Without these people the endeavor of physics would be a much less fun.

TABLE OF CONTENTS

Acknowledgements	2
Table of Contents	3
List of Figures	4
1 Introduction	1
2 Perturbation Theory for Effective Atoms and Linear Response	3
2.1 Density-Functional Perturbation Theory	4
2.1.1 First Order Energy Response	4
2.1.2 Effective atoms	5
2.1.3 First Order Wavefunction Response	6
2.1.4 $\psi\psi$ derivative	7
2.1.5 $\psi\tau$ derivative	8
2.1.6 Notes on Implementation	9
2.1.7 Convergence Tests for VPT and EAT	10
2.2 Incommensurate Perturbation Theory	11
2.2.1 $\psi\tau$ Derivative	13
2.2.2 $\psi\psi$ Derivative	14
2.2.3 Summary of Expressions	19
2.3 Minimization Algorithms for Incommensurate Perturbations	19
2.3.1 Application to dielectric response	20
3 Wannier Function Based Theory of Incommensurate Systems	21
3.1 Introduction	21
3.2 Ab Initio Tight Binding Models	22
3.3 Variational Constraints in DFT	23
3.3.1 Arbitrary constraint	25
3.3.2 Flake Boundary Conditions	26
3.4 Preliminary Results for Twisted Bilayer Graphene	27
4 Conclusion	30
5 Appendix	31
5.1 Summary of Algebraic DFT Notation [10]	31
5.2 Derivative of the Q operator	33
5.3 Metallic Fillings	34
5.4 Momentum Resolved Density of States	35
Bibliography	36

LIST OF FIGURES

2.1	Density response for a Gaussian charge perturbation near graphene sheet.	10
2.2	Finite difference convergence test of VPT.	11
2.3	Finite difference convergence test of EAT.	12
3.1	Graphical depiction of MINT Flake system. (Light green) Substrate unit cell atoms, (Dark purple) Flake atoms, (Arrows) Flake atoms connected by boundary conditions	28
3.2	Momentum resolved density of states for graphene bilayer flake system, <u>no boundary conditions applied.</u> k-path goes through BZ of substrate. Intensity is plotted on a log scale. Red vertical line shows the expected location of mirrored Dirac cone.	29
3.3	Momentum resolved density of states for graphene bilayer flake system, <u>with periodic boundary conditions.</u> k-path goes through BZ of substrate. Intensity is plotted on a log scale. Red vertical line shows the expected location of mirrored Dirac cone.	29

CHAPTER 1

INTRODUCTION

Studying condensed matter systems is a difficult endeavor because of the complexity associated with the sheer number of atoms in a small piece of matter. The simplest way to approximate these systems is by assuming the electrons are non-interacting, and feel a potential that is derived purely from the atomic nuclei. A very useful result holds if the potential is periodic, for example in the case of a crystal lattice. Bloch's theorem states that in the presence of a periodic potential, the electronic eigenstates of a system take the form $u(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$ where $u(\mathbf{r})$ has the same periodicity as the potential and \mathbf{k} is an arbitrary wavevector. Besides giving us a convenient and simple way to index all the eigenstates, this decomposition also allows for a computationally feasible way to represent them. Specifically, we can choose a finite Fourier basis set for the periodic function $u(\mathbf{r})$.

This formulation is still not sufficient to describe real-world materials because it neglects the correlations and interactions between electrons. This is where the framework of density-functional theory (DFT) attains its greatest success. DFT is based on the Nobel Prize-winning work of Kohn and Hohenberg that showed the ground state energy of a many-electron system depends only on the total density of electrons [7]. Under this framework, the ground state energy is computed by minimizing a set of effective one-body Schrödinger equations for each electron that are coupled to each other by a functional of the total electron density. This functional incorporates all the non-trivial correlations as well as the effect of exchanging identical fermions, and thus bears the name of the exchange-correlation functional.

By combining DFT with Bloch's theorem we obtain a powerful way of studying

the ground state properties of periodic materials. For example, calculations of band structures and lattice constants are usually quite accurate. With all the success it has seen, we are left wondering how applicable these techniques are to other kinds of systems, in particular those without exact periodicity. There are many interesting systems that fit into this category and equally as many reasons to study them. For example, high-entropy alloys (HEAs) are materials that have a crystal structure but are non-periodic because each lattice site is occupied by a random type of atom [4]. In many instances these materials exhibit desirable physical characteristics. One application is as electrocatalysts where the existence of many distinct local configurations of atoms is beneficial because some of those configurations act as extremely efficient binding sites. We develop perturbation theory techniques to study responses to non-periodic perturbations as well as optimize high-entropy alloys by treating each atom as an effective combination of different species.

A different kind of non-periodic system is a layered material. These are made of two or more 2D lattices that are stacked on top of each other. If the lattice constants of each layer are not commensurate, the system loses its overall periodicity. There exist ab-initio techniques to study systems with incompatible lattices, for example mismatched interface theory (MINT) [5]. We use MINT to develop a way of computing the electronic structure of a twisted bilayer graphene system. This material consists of two graphene sheets stacked on top of each other with a 30-degree twist angle between them. There are interesting effects such as mirrored Dirac cones that make this system a popular subject of research.

CHAPTER 2

PERTURBATION THEORY FOR EFFECTIVE ATOMS AND LINEAR RESPONSE

The study of high-entropy alloys begins with an old technique from solid state physics. The virtual crystal approximation (VCA) was used in early electronic structure calculations to approximate the properties of periodic systems with different species of atoms in a unit cell. To reduce the computational load, VCA pretends that each atom is the average of all the atoms in the sense that the atomic potential is set to the average potential of each of the atoms. Following this, symmetry can be used to reduce the total number of atoms that need to be considered in the calculation. With the steep increase in computing power, VCA's original necessity has been eliminated. The reason we mention it is that the fundamental idea behind it still holds power. It is the assumption that the behavior of a bunch of different atoms is similar to the behavior of their average.

The new idea of effective atom theory (EAT) is motivated by the fact that a given group of atoms in a section of a HEA could have every configuration allowed by the makeup of the HEA. If we want to find a configuration that maximizes some quantity, such as the binding energy in the case of an electrocatalyst, then we would have to search through a combinatorially large number of such configurations. Instead, the solution is actually to enlarge the search space by making each atom a weighted linear combination of the different species in the HEA. This way, we can use any continuous optimization algorithm to search through the space to find the optimal energy solution and discretize the solution to find the configuration that achieves the maximum possible energy. This requires computing the gradient of energy in continuous parameter space, which is a problem for perturbation theory.

2.1 Density-Functional Perturbation Theory

Applying perturbation to density-functional theory involves taking a variational approach, which we refer to by variational perturbation theory (VPT). DFT itself is can be formulated variationally because we can think of the ground state solution as one that extremizes the energy and hence the first derivative of energy is zero when evaluated at the solution. This allows us to obtain the first order change in energy easily. Computing the first order change in wavefunctions is more complicated and requires computation of second derivatives of the total energy. We will discuss both in the following sections.

2.1.1 First Order Energy Response

The result that allows us to compute the first order response of ground state energy is the Hellman-Feynman theorem. Let ψ denote the wavefunction of the system, and define τ as a set of parameters that determine the quantity we wish to perturb and $E(\psi, \tau)$ the total energy which depends on both variables. Then,

$$\frac{d}{d\tau} E(\psi_*(\tau), \tau) = \left. \frac{\partial E}{\partial \psi} \right|_{\psi_*} \frac{\partial \psi_*(\tau)}{\partial \tau} + \frac{\partial E}{\partial \tau} \quad (2.1)$$

where $\psi_*(\tau)$ represents the energy minimizing wavefunction at a given parameter τ . Because ψ_* sits at a global minimum of E , then the term $\left. \frac{\partial E}{\partial \psi} \right|_{\psi_*}$ must necessarily be zero and we find

$$\frac{d}{d\tau} E(\psi_*(\tau), \tau) = \frac{\partial E}{\partial \tau}. \quad (2.2)$$

This result states that the first order change in energy can be computed solely by taking the derivative with respect to the perturbation itself without any reference to the wavefunction dependent parts.

2.1.2 Effective atoms

The first order energy response can now be applied to effective atoms to calculate the energy gradient used in minimization. Effective atom theory is a potential works by parameterizing atom species with continuous variables. In algebraic DFT notation (see appendix), the ion dependent part of the total energy, E_{ion} , is equal to

$$(Jn)^\dagger V_{loc} + \sum_{k,n,s} w_k \text{tr}(C_k^\dagger V_{snk} M_s V_{snk}^\dagger C_k F) \quad (2.3)$$

where k runs over k-points and V_{loc} , the local part of the pseudopotential, is equal to $V_{loc} = \sum_{n,s} (V_{loc})_{sn}$. The second term in the expression involving VMV^\dagger represents the nonlocal part of the pseudopotential. Here s, n run over the species and atom numbers, respectively. A mixed atom has a potential that is a linear combination of different atomic species, weighted by parameters θ . We write the components of the mixed atom's potential as

$$((V_{loc})_{mixed})_n = \sum_s \theta_{sn} (V_{loc})_{sn}, \quad (2.4)$$

$$((V_{nl})_{mixed})_n = \sum_s \theta_{sn} V_{snk} M_s V_{snk}^\dagger. \quad (2.5)$$

Optimizing energy over the weights θ_{sn} requires finding the gradient of E_{ion} with respect to the weights. Applying the Hellman-Feynman theorem tells us

$$\frac{dE_{ion}}{d\theta_{sn}} = \frac{\partial E_{ion}}{\partial \theta_{sn}} = (Jn)^\dagger (V_{loc})_{sn} + \sum_k w_k \text{tr}(C_k^\dagger V_{snk} M_s V_{snk}^\dagger C_k F). \quad (2.6)$$

There is also an additional term involving the change in energy due to the total number of electrons. Going through the math will yield the extra term

$$\frac{dE_{ion}}{d\theta_{sn}} = \frac{\partial E_{ion}}{\partial \theta_{sn}} + \mu \frac{dN}{d\theta_{sn}} \quad (2.7)$$

where μ is the chemical potential of the system and N is the total number of electrons.

2.1.3 First Order Wavefunction Response

Knowing the response of the ground state wavefunctions to external perturbations gives us all the information about the linear response of the system. For example, we can take the perturbed wavefunctions and use them to compute the first order density perturbation which in turn determines the dielectric response of the system. Computing the response wavefunctions within the variational method necessarily involves taking the second derivatives of energy. Let us express the wavefunction $\boldsymbol{\psi}$ in terms of its basis elements ψ_i and system parameters $\boldsymbol{\tau}$ as the vector τ_i . Starting with the fact that $\frac{\partial E}{\partial \boldsymbol{\psi}} = 0$ at $\boldsymbol{\psi} = \boldsymbol{\psi}(\boldsymbol{\tau})$, we expand the energy as a second order Taylor series:

$$\frac{\partial E}{\partial \psi_j}(\boldsymbol{\psi} + d\boldsymbol{\psi}, \boldsymbol{\tau} + d\boldsymbol{\tau}) = \frac{\partial E(\boldsymbol{\psi}, \boldsymbol{\tau})}{\partial \psi_j} + \frac{\partial E(\boldsymbol{\psi}, \boldsymbol{\tau})}{\partial \psi_i \psi_j} d\psi_i + \frac{\partial E(\boldsymbol{\psi}, \boldsymbol{\tau})}{\partial \tau_\alpha \psi_j} d\tau_\alpha + O(d\boldsymbol{\psi}d\boldsymbol{\tau}). \quad (2.8)$$

We also know that $\frac{\partial E}{\partial \psi_j}(\boldsymbol{\psi}, \boldsymbol{\tau}) = \frac{\partial E}{\partial \psi_j}(\boldsymbol{\psi} + d\boldsymbol{\psi}, \boldsymbol{\tau} + d\boldsymbol{\tau}) = 0$ because the ground state wavefunctions must minimize the energy. After dropping the second order terms, we obtain the equation

$$\frac{\partial E}{\partial \psi_i \psi_j} d\psi_i + \frac{\partial E}{\partial \tau_\alpha \psi_j} d\tau_\alpha = 0. \quad (2.9)$$

Note that this equation is linear in $d\boldsymbol{\psi}$ and $d\boldsymbol{\tau}$, and solving it for $d\boldsymbol{\psi}$ gives first order change of the ground state wavefunction w.r.t. any kind of perturbation parameterized by $d\boldsymbol{\tau}$. Since a typical calculation can involve upwards of 10^6 basis elements for $\boldsymbol{\psi}$, it is not feasible to compute the matrix $\frac{\partial E}{\partial \psi_i \psi_j}$ and invert it directly. Rather, we must resort to iterative techniques for solving large linear systems (eg. conjugate gradients). To get a sense for how this works in practice, let us symbolically write $\frac{\partial E}{\partial \psi_j}$ as ∇E and the second derivative $\frac{\partial E}{\partial \psi_i \psi_j}$ as $d_\psi \nabla E$. Similarly, we write $\frac{\partial E}{\partial \tau_\alpha \psi_j}$ as $d_\tau \nabla E$. Thinking about these linear operators as functions applied

on vectors the equation looks like

$$d_\psi \nabla E(d\boldsymbol{\psi}) = d_\tau \nabla E(d\boldsymbol{\tau}). \quad (2.10)$$

If we have then the algebraic form for ∇E , it possible to analytically take it's derivative w.r.t. $\boldsymbol{\psi}$ and $\boldsymbol{\tau}$. Once we have this, applying $d_\psi \nabla E$ to a wavefunction $d\boldsymbol{\psi}$ is the same as plugging $d\boldsymbol{\psi}$ into the mathematical expression.

2.1.4 $\psi\psi$ derivative

From now on, instead of the wave functions, we will be working with their unconstrained expansion coefficients Y_k related to the normalized coefficients C_k by $C_k = Y_k U_k^{-1/2}$ and $U_k = Y_k^\dagger O Y_k$. Computing the second derivative is most easily done by starting from the first derivative, or the gradient. In algebraic notation (see appendix), the derivative of energy is written as

$$\nabla E = \sum_k w_k [(\mathbb{1} - O C_k C_k^\dagger) H_k C_k F U_k^{-1/2} + O C_k Q_k([\tilde{H}_k, F])]. \quad (2.11)$$

We will now make the assumption that the fillings are scalar, which is true for semiconducting systems. A discussion of how to generalize these results for non-uniform fillings is found in the appendix. With this assumption, the term involving the Q operator drops out. Combining this with the fact that $\nabla E = 0$ at the energy minimum and $U_k = \mathbb{1}$ gives us a simple expression for the derivative of ∇E with respect dY :

$$d_Y \nabla E = \sum_k w_k [dH_k C_k F - H_k dC_k F - O C_k d\tilde{H}_k F - O dC_k \tilde{H}_k F]. \quad (2.12)$$

Here the symbol d denotes the first order change in a particular quantity with respect to dY . The quantities dH_k , $d\tilde{H}_k$ and dC_k depend linearly on dY . Their

forms are easily derived from looking at the definitions. The perturbation to the subspace Hamiltonian is

$$d\tilde{H}_k = dC_k^\dagger H_k C_k + C_k^\dagger dH_k C_k + C_k^\dagger H_k dC_k. \quad (2.13)$$

The change of the orthonormalized wavefunctions are related to the unconstrained coefficients by

$$dC_k = dY_k + Y_k dU_k^{-1/2}, \quad (2.14)$$

$$dU_k^{-1} = -\frac{1}{2}(dY^\dagger OY + Y^\dagger OdY). \quad (2.15)$$

Finally, the perturbation to the self-consistent Hamiltonian is

$$dH_k = I^\dagger (\text{Diag } dV_{\text{sc}}) I \quad (2.16)$$

where

$$dV_{\text{sc}} = -4\pi J^\dagger O L^{-1} O J dn + J^\dagger O J \text{Diag}[E'_{xc}(n)] dn + \text{Diag}[E'_{xc}(n)] J^\dagger O J dn + \text{Diag}[E''_{xc}(n) \cdot dn] J^\dagger O J n, \quad (2.17)$$

$$dn = \text{diag}(IdC_k F C_k^\dagger I^\dagger + IC_k F dC_k^\dagger I^\dagger) = 2 \text{Re } \text{diag}(IdC_k F C_k I^\dagger), \quad (2.18)$$

$$(a \cdot b)_i = a_i b_i. \quad (2.19)$$

These expressions give us a way of applying the operator $d_\psi \nabla E$, to a perturbation wavefunction dY to obtain the left side of equation 2.9. Having completed this, we move on to the other second derivative of interest.

2.1.5 $\psi\tau$ derivative

The process of computing $d_\tau \nabla E$ is much simpler as only the Hamiltonian H_k is directly affected by perturbations. We find that

$$d_\tau \nabla E = dH_k C_k F - O C_k C_k^\dagger dH_k C_k F, \quad (2.20)$$

where

$$dH_k = I^\dagger(\text{Diag } dV_{\text{sc}})I. \quad (2.21)$$

The derivative of the potential dV_{sc} depends on the exact type of perturbation. In the case of perturbations to atomic positions, we must calculate the derivative of the structure factor $S_f(G) = \sum_I e^{-iG \cdot X_I}$ that contributes a portion of the total potential $V_{\text{sc}} = J^\dagger \left(\frac{-4\pi Z}{G^2} \right) S_f(G)$. Taking the first differential gives

$$dS_f(G) = -iG_j e^{-iG \cdot X_{\alpha j}}, \quad (2.22)$$

$$dV_{\text{sc}} = J^\dagger \left(\frac{-4\pi Z}{G^2} \right) dS_f(G). \quad (2.23)$$

There are many other kinds of perturbations we can consider. For example, a perturbation to V_{ext} gives $dV_{\text{sc}} = dV_{\text{ext}}$. Finally we can also have charge density and electric field perturbations.

2.1.6 Notes on Implementation

The results of variational perturbation theory have been realized in the JDFTx software package, which is an open source density-functional theory code designed for large scale calculations [10]. The implementation of VPT consists of a set of commands to calculate the response to specific types of perturbations. These include perturbations to external potential, external charge, electric field, and ionic positions. Supported features include ultrasoft pseudopotentials, symmetries, and the majority of exchange-correlation functionals. Support for non-uniform fillings has yet to be implemented. In terms of performance, VPT is computationally efficient. The speed of a VPT calculation typically matches or exceeds an electronic structure calculation for the same system. Furthermore, the advantage of VPT

over the finite difference method for approximating a derivative is that it is more accurate and does not require the user to manually adjust parameters such as step size. Finally, as an example of theory being put into use, a calculation of the density response of a graphene sheet to a charge perturbation is shown below.

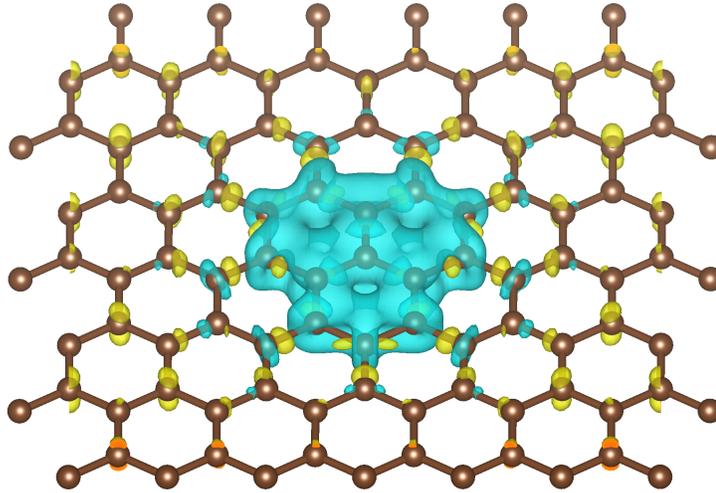


Figure 2.1: **Density response for a Gaussian charge perturbation near graphene sheet.**

2.1.7 Convergence Tests for VPT and EAT

To test the validity of density-functional perturbation theory and the correctness of the derivations, we may choose to compare the derivatives of certain quantities compared using the analytical methods against a finite difference approximation of the derivative. In the first test, we started with a Silicon lattice and applied a periodic perturbation to the external potential. We then computed the expected response density dn using both VPT and finite difference and compared the two with different step sizes Δ . The relative difference in the Euclidean norm between the results of the two methods are shown in (2.2).

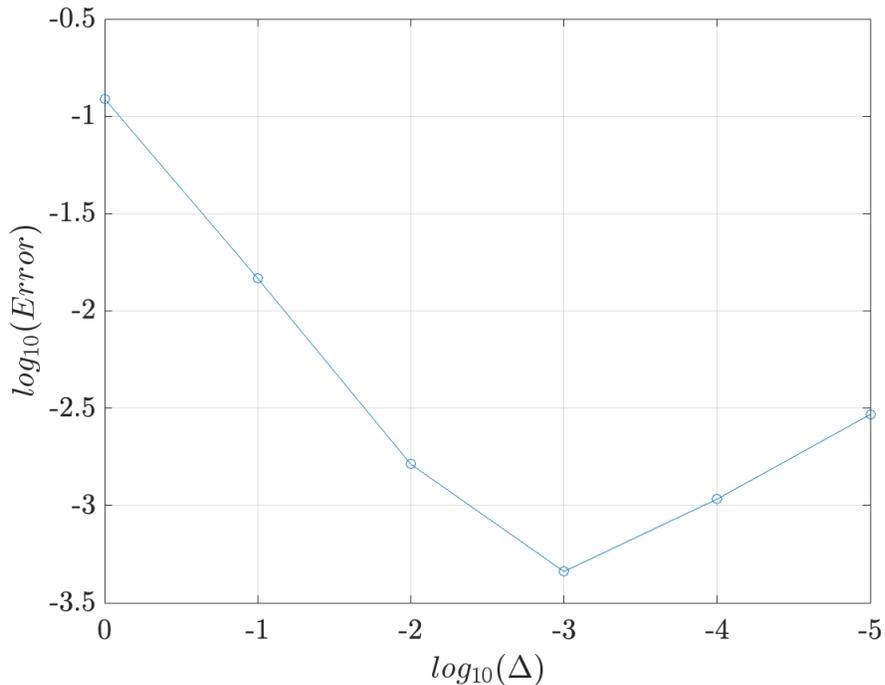


Figure 2.2: **Finite difference convergence test of VPT.**

Following this, we tested effective atom theory by having the same Silicon lattice and replacing the two atoms by a mixture of different species. We then computed the gradient of the total energy analytically and compared it to the central finite difference approximation computed by shifting the weights slightly (2.3). Both figures show errors that decrease with decreasing step size, indicating that the responses computed by VPT and EAT are correct and represent the true derivatives.

2.2 Incommensurate Perturbation Theory

With VPT working for periodic perturbations, let us turn our attention to the case when the perturbation has a non-periodic phase attached to it. Consider a

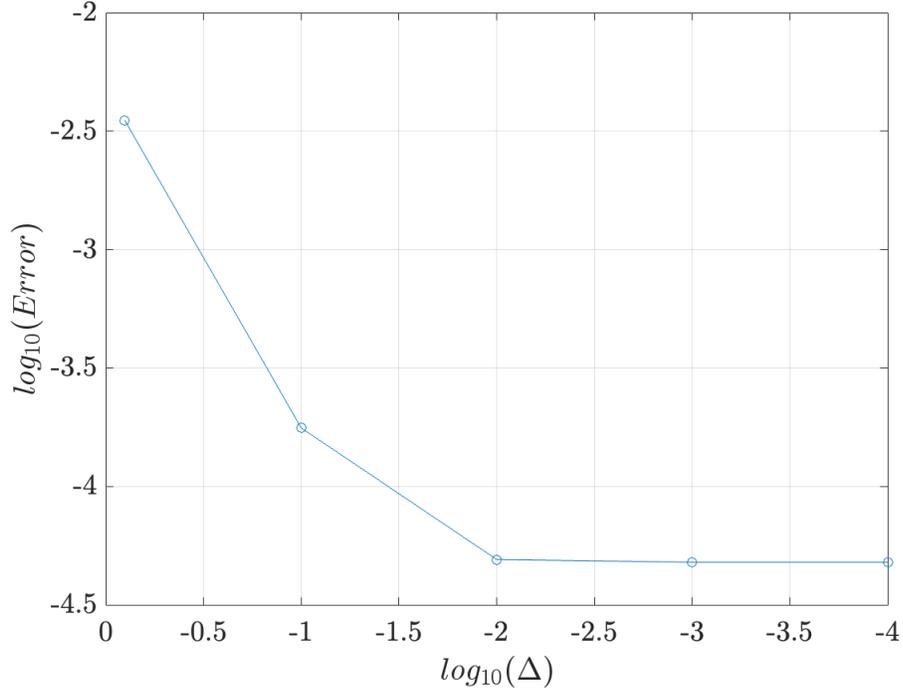


Figure 2.3: **Finite difference convergence test of EAT.**

perturbation of the form

$$dV_{\text{ext}} = du(r)e^{iqr}. \quad (2.24)$$

In traditional perturbation theory the change in wavefunctions is given by a sum over states $n'k'$:

$$d|\psi_{nk}\rangle = \sum_{n'k'} \frac{|\psi_{n'k'}\rangle \langle \psi_{n'k'} | du(r)e^{iqr} | \psi_{nk}\rangle}{E_{nk} - E_{n'k'}}. \quad (2.25)$$

We notice that the numerator most of the time is zero unless the phase accumulated by $|\psi_{nk}\rangle$ after applying an extra e^{iqr} matches the phase of $\langle \psi_{n'k'} |$ up to a reciprocal lattice vector. That is,

$$\langle \psi_{n'k'} | du(r)e^{iqr} | \psi_{nk}\rangle \neq 0 \iff k' = k + q \pmod{BZ}. \quad (2.26)$$

This suggests that the perturbation to a wavefunction with wavevector \mathbf{k} will be at wavevector $\mathbf{k} + \mathbf{q}$. Given this knowledge, we can begin to formulate a way

of representing these kinds of perturbations in the algebraic DFT notation using variational perturbation theory. We will start with the fact that a system with an $n \times n \times n$ k-point mesh is equivalent to a supercell with $n \times n \times n$ unit cells at the Γ point. The ground state wavefunctions C in the supercell can be grouped according to blocks according to their k-point, and so can the basis elements. This means the matrix C can be written as a block-diagonal matrix of the form

$$C = \begin{bmatrix} C_{11} & & \\ & C_{22} & \\ & & \ddots \end{bmatrix}, \quad (2.27)$$

where C_{ii} is a block corresponding to the wavefunctions C_i in the small calculation for k-point i . We note that the multiplication rule for two block matrices A, B with blocks A_{ij} and B_{ij} respectively is the same form as regular matrix multiplication:

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}. \quad (2.28)$$

Now, let A and C be block diagonal matrices and let B be an arbitrary matrix with nonzero blocks. We have

$$(ABC)_{ij} = A_i B_{ij} C_j, \quad (AC)_k = A_k C_k. \quad (2.29)$$

These results allow us to pretend like we are working in a large supercell approximation of the true crystal and systematically handle the effects of perturbations with an attached phase.

2.2.1 $\psi\tau$ Derivative

Using this block matrix framework, we can derive the second derivatives of energy. Here it is easier to start with the mixed $\psi\tau$ derivative. We begin by looking at

the perturbation stated in the previous section. Because we wish to maintain the hermitianness of the Hamiltonian, we will require the perturbation to be real, so the previous perturbation acquires a counterpart in the form of its complex conjugate to make $du(r)e^{iqr} + du^*(r)e^{-iqr}$. Now, define the map $T_+(\mathbf{k})$ that takes a wavevector \mathbf{k} to the unique vector in the unit reciprocal lattice cell that has the same Bloch character as $\mathbf{k} + \mathbf{q}$, and similarly $T_-(\mathbf{k})$ to map to the equivalence class of $\mathbf{k} - \mathbf{q}$ such that T_+ and T_- are inverses of each other. Applying this to the $\psi\tau$ derivative, if there is a perturbation to the potential at wavevector \mathbf{q} then dH will have non-zero components $dH_{T_+(k)k}$ and $dH_{T_-(k)k}$. Starting from the form of the second derivative in the case of a single k-point,

$$d_\tau \nabla E = dHCF - OCC^\dagger dHCF, \quad (2.30)$$

we find that the matrix $d_\tau \nabla E$ has two sets of non-zero blocks, whose values are given by

$$(d_\tau \nabla E)_{T_+(k)k} = dH_{T_+(k)k} C_k F - OC_{T_+(k)k} C_{T_+(k)k}^\dagger dH_{T_+(k)k} C_k F, \quad (2.31)$$

$$(d_\tau \nabla E)_{T_-(k)k} = dH_{T_-(k)k} C_k F - OC_{T_-(k)k} C_{T_-(k)k}^\dagger dH_{T_-(k)k} C_k F. \quad (2.32)$$

In the next section we will derive what the the expression for $dH_{T_\pm(k)k}$, applicable for both kinds of derivatives.

2.2.2 $\psi\psi$ Derivative

The second derivative with respect to the wavefunctions has more moving parts. Starting from the expression for the derivative in the case of the periodic potential

$$d_Y \nabla E = dHCF - HdCF - OCd\tilde{H}F - OdC\tilde{H}F. \quad (2.33)$$

To have the two sides of the linear system 2.9 have the same block structure, we require perturbations to C_k at both $k - q$ and $k + q$. Hence, $C_{T_+(k)k}$ and $C_{T_-(k)k}$ must be non-zero and we have

$$(d_Y \nabla E)_{T_{\pm}k} = dH_{T_{\pm}(k)k}C_k F - H_{T_{\pm}(k)k}dC_{T_{\pm}(k)k}F - OC_{T_{\pm}(k)k}d\tilde{H}_{T_{\pm}(k)k}F - OdC_{T_{\pm}(k)k}\tilde{H}_k F. \quad (2.34)$$

Now, when we try to expand the wavefunction dC in block matrix form, something interesting happens. We have

$$dC_{T_{\pm}(k)k} = dY_{T_{\pm}(k)k} + Y_{T_{\pm}(k)k}dU_{T_{\pm}(k)k}^{-1/2}. \quad (2.35)$$

These expressions seem to present an issue. The problem is that the perturbation to the wavefunction at k should theoretically not depend on the wavefunctions at $k \pm q$ based on the results of perturbation theory, and yet they appear interlinked in the formulae. This raises the question of what exactly is happening here, and the answer has to do with orthogonality.

Essentially, if C remains normalized under a perturbation, then $(C + dC)^\dagger(C + dC) = \mathbb{1} \implies dC^\dagger C + C^\dagger dC = 0$ and in terms of block matrices the equation reads $(dC_{kT_{\pm}(k)})^\dagger C_k + C_{T_{\pm}(k)}^\dagger dC_{T_{\pm}(k)k} = 0$, and this is where the interdependence on the $k + q$ wavefunctions comes from. To see if there is a way to reframe this problem, let us start with the idea that the perturbed wavefunctions dC solve the equation $\frac{\partial^2 E}{\partial \psi^2} dC = \frac{\partial^2 E}{\partial \psi \tau}$. If we want to solve this equation in the space of unnormalized wavefunctions, then we may project the equation into the orthonormal subspace. The projection operator is $P = \mathbb{1} - OCC^\dagger$ which gives $P \frac{\partial^2 E}{\partial \psi^2} P dY = \frac{\partial^2 E}{\partial \psi \tau}$ when applied to the linear system. We will show that $P \frac{\partial^2 E}{\partial \psi^2} = \frac{\partial^2 E}{\partial \psi^2}$ on the orthonormal subspace where $dC^\dagger C = 0$. This can be seen by expanding

$$P \frac{\partial^2 E}{\partial \psi^2} dC = (\mathbb{1} - OCC^\dagger)(-OdC\tilde{H}F - OCd\tilde{H}F + dHCF + HdCF) \quad (2.36)$$

$$\begin{aligned}
&= -OdC\tilde{H}F + dHCF + HdCF + OCC^\dagger OdC\tilde{H}F \\
&\quad -OCC^\dagger dHCF - OCC^\dagger HdCF.
\end{aligned} \tag{2.37}$$

We now use $d\tilde{H} = dC^\dagger HC + C^\dagger dHC + C^\dagger HdC$ and the fact that C contains eigenvectors of the Hamiltonian meaning $HC = C\epsilon$ where ϵ is diagonal, hence $dC^\dagger HC = dC^\dagger C\epsilon = 0$. This means

$$P \frac{\partial^2 E}{\partial \psi^2} dC = -OdC\tilde{H}F + dHCF + HdCF - OCC^\dagger dHCF = \frac{\partial^2 E}{\partial \psi^2} dC. \tag{2.38}$$

It is also the case that $dC = PdY \implies dC^\dagger C = 0$. So, we see solving

$$P \frac{\partial^2 E}{\partial \psi^2} PdY = \frac{\partial^2 E}{\partial \psi^2} PdY \tag{2.39}$$

is the same as solving

$$\frac{\partial^2 E}{\partial \psi^2} (PdY) = \frac{\partial^2 E}{\partial \psi^2} PdY. \tag{2.40}$$

In our notation we now have

$$dC_{T_\pm(k)k} = (PdY)_{T_\pm(k)k} = dY_{T_\pm(k)k} - OC_{T_\pm(k)} C_{T_\pm(k)}^\dagger dY_{T_\pm(k)k} \tag{2.41}$$

$$\begin{aligned}
d_Y \nabla E &= -OdC_{T_\pm(k)k} \tilde{H}_k F - OC_{T_\pm(k)} C_{T_\pm(k)}^\dagger dH_{T_\pm(k)k} C_k F \\
&\quad + dH_{T_\pm(k)k} C_k F + H_{T_\pm(k)} dC_{T_\pm(k)k} F.
\end{aligned} \tag{2.42}$$

This means we still need access to the wavefunctions $C_{T_\pm(k)}$, which can be obtained via an inexpensive band structure calculation, but their purpose is now clear: to project out the components of the perturbation in the direction of existing eigenvectors.

Moving on to the Hamiltonian term in the second derivative, we note that the electron density $n(r)$ can be written as $\psi(r)^* \psi(r)$ where $\psi(r) = \sum C_{k\alpha} e^{i(k+G_\alpha)r} = e^{ikr} \sum C_{k\alpha} e^{i(G_\alpha)r} \equiv e^{ikr} u_k(r)$ and u_k is the corresponding periodic part of the wavefunction. We are assuming $d\psi_k$ has two components at $T_+(k)$ and $T_-(k)$. We will

label the periodic parts u_k and u'_k , and thus the total perturbing wavefunction is $e^{iT_+(k)r} du_k(r) + e^{iT_-(k)r} du'_k(r)$. The differential of electron density can be expressed as

$$dn = \sum d\psi(r)^* \psi(r) + \psi(r)^* d\psi(r) \quad (2.43)$$

$$= \overline{(e^{iT(\mathbf{k})\cdot\mathbf{r}} du_k(r) + e^{iT^{-1}(\mathbf{k})\cdot\mathbf{r}} du'_k(r))} u_k(r) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (2.44)$$

$$+ \overline{u_k(r) e^{i\mathbf{k}\cdot\mathbf{r}}} (e^{iT(\mathbf{k})\cdot\mathbf{r}} du_k(r) + e^{iT^{-1}(\mathbf{k})\cdot\mathbf{r}} du'_k(r))$$

$$= \sum_{\mathbf{k},n} \left[e^{i(\mathbf{k}-T^{-1}(\mathbf{k}))\cdot\mathbf{r}} du_k'^*(r) u_k(r) + e^{i(T(\mathbf{k})-\mathbf{k})\cdot\mathbf{r}} u_k^*(r) du_k(r) \quad (2.45)$$

$$+ e^{i(\mathbf{k}-T(\mathbf{k}))\cdot\mathbf{r}} du_k^*(r) u_k(r) + e^{i(T^{-1}(\mathbf{k})-\mathbf{k})\cdot\mathbf{r}} u_k^*(r) du_k'(r) \right].$$

We notice that the following quantities are all periodic:

$$\frac{e^{i(\mathbf{k}-T^{-1}(\mathbf{k}))\cdot\mathbf{r}}}{e^{i\mathbf{q}\cdot\mathbf{r}}}, \quad \frac{e^{i(T(\mathbf{k})-\mathbf{k})\cdot\mathbf{r}}}{e^{i\mathbf{q}\cdot\mathbf{r}}}, \quad \frac{e^{i(\mathbf{k}-T(\mathbf{k}))\cdot\mathbf{r}}}{e^{-i\mathbf{q}\cdot\mathbf{r}}}, \quad \frac{e^{i(T^{-1}(\mathbf{k})-\mathbf{k})\cdot\mathbf{r}}}{e^{-i\mathbf{q}\cdot\mathbf{r}}}. \quad (2.46)$$

This means we can rewrite the expression as

$$= \sum_{\mathbf{k},n} \left[\frac{e^{i(\mathbf{k}-T^{-1}(\mathbf{k}))\cdot\mathbf{r}}}{e^{i\mathbf{q}\cdot\mathbf{r}}} du_k'^*(r) u_k(r) + \frac{e^{i(T(\mathbf{k})-\mathbf{k})\cdot\mathbf{r}}}{e^{i\mathbf{q}\cdot\mathbf{r}}} u_k^*(r) du_k(r) \right] e^{i\mathbf{q}\cdot\mathbf{r}} \quad (2.47)$$

$$+ \left[\frac{e^{i(\mathbf{k}-T(\mathbf{k}))\cdot\mathbf{r}}}{e^{-i\mathbf{q}\cdot\mathbf{r}}} du_k^*(r) u_k(r) + \frac{e^{i(T^{-1}(\mathbf{k})-\mathbf{k})\cdot\mathbf{r}}}{e^{-i\mathbf{q}\cdot\mathbf{r}}} u_k^*(r) du_k'(r) \right] e^{-i\mathbf{q}\cdot\mathbf{r}}$$

In the operator notation, this becomes

$$(dn)_q = \sum_k \text{diag}[M_{\mathbf{k}-T^{-1}(\mathbf{k})-\mathbf{q}} (IdC_{T_-(k)k})^\dagger FIC_k \quad (2.48)$$

$$+ M_{-\mathbf{k}+T(\mathbf{k})-\mathbf{q}} (IC_k)^\dagger FIdC_{T_+(k)k}]$$

$$(dn)_{-q} = \sum_k \text{diag}[M_{\mathbf{k}-T(\mathbf{k})+\mathbf{q}} (IdC_{T_+(k)k})^\dagger FIC_k \quad (2.49)$$

$$+ M_{-\mathbf{k}+T^{-1}(\mathbf{k})+\mathbf{q}} (IC_k)^\dagger FIdC_{T_-(k)k}]$$

We now define the M operator to act on real space vectors by multiplying by a complex exponential:

$$M_\alpha \psi = e^{i\alpha\cdot\mathbf{r}} \cdot \psi(r). \quad (2.50)$$

We see that the Hamiltonian resulting from this perturbation will have two types of non-zero components, $dH_{T_-(k)k}$ and $dH_{T_+(k)k}$. This is because in the plane-wave basis, we have

$$dV_{sc} \propto [-4\pi IL^{-1}Jdn + E'_{xc}(n) \cdot dn + E'_{xc}(n) \cdot dn + E''_{xc}(n) \cdot dn \cdot n] \quad (2.51)$$

Clearly the perturbing potential dV_{sc} has Bloch characters \mathbf{q} and $-\mathbf{q}$. Let $dV_{sc}(r) = v(r)e^{i\mathbf{q}r} + v'(r)e^{-i\mathbf{q}r}$, so that $(dV_{sc})_q$ has periodic part $v(r)$ and $(dV_{sc})_{-q}$ has periodic part $v'(r)$. We write

$$\begin{aligned} (dV_{sc})_{\pm q} &= -4\pi J^\dagger OL_{\pm q}^{-1} OJdn_{\pm q} + J^\dagger OJ \text{Diag}[E'_{xc}(n)]dn_{\pm q} \\ &+ \text{Diag}[E'_{xc}(n)]J^\dagger OJdn_{\pm q} + \text{Diag}[E''_{xc}(n) \cdot dn_{\pm q}]J^\dagger OJn. \end{aligned} \quad (2.52)$$

When we pointwise multiply by a wavefunction $\psi(r) = w(r)e^{i\mathbf{k}\cdot r}$ with Bloch character \mathbf{k} , we get

$$v(r)w(r)e^{i(\mathbf{k}+\mathbf{q})\cdot r} + v'(r)w(r)e^{i(\mathbf{k}-\mathbf{q})\cdot r}. \quad (2.53)$$

and again, we see that the following are periodic:

$$\frac{e^{i(\mathbf{k}+\mathbf{q})\cdot r}}{e^{i\mathbf{T}(\mathbf{k})r}}, \quad \frac{e^{i(\mathbf{k}-\mathbf{q})\cdot r}}{e^{iT^{-1}(\mathbf{k})r}}. \quad (2.54)$$

Therefore, rewriting the expression gives

$$\text{Diag}(dV_{sc})IC = v(r)w(r)\frac{e^{i(\mathbf{k}+\mathbf{q})\cdot r}}{e^{i\mathbf{T}(\mathbf{k})r}}e^{i\mathbf{T}(\mathbf{k})r} + v'(r)w(r)\frac{e^{i(\mathbf{k}-\mathbf{q})\cdot r}}{e^{iT^{-1}(\mathbf{k})r}}e^{iT^{-1}(\mathbf{k})r}. \quad (2.55)$$

Applying the I^\dagger operator takes us back to Fourier space, and we see that

$$(I^\dagger \text{Diag}(dV_{sc})IC)_{T_+(k)} = I^\dagger M_{\mathbf{k}-\mathbf{T}(\mathbf{k})+\mathbf{q}} \text{Diag}((dV_{sc})_q)IC_k, \quad (2.56)$$

$$(I^\dagger \text{Diag}(dV_{sc})IC)_{T_-(k)} = I^\dagger M_{\mathbf{k}-\mathbf{T}^{-1}(\mathbf{k})-\mathbf{q}} \text{Diag}((dV_{sc})_{-q})IC_k. \quad (2.57)$$

Equivalently, we have

$$dH_{T_+(k)k} = I^\dagger M_{\mathbf{k}-\mathbf{T}(\mathbf{k})+\mathbf{q}} \text{Diag}((dV_{sc})_q)I, \quad (2.58)$$

$$dH_{T_-(k)k} = I^\dagger M_{\mathbf{k}-\mathbf{T}^{-1}(\mathbf{k})-\mathbf{q}} \text{Diag}((dV_{sc})_{-q})I. \quad (2.59)$$

This provides a way to calculate the response to incommensurate perturbations.

2.2.3 Summary of Expressions

Instead of labelling the non-zero blocks of dC by their k-point $T_{\pm}(k)$, we will imagine that the expansion coefficients actually have the phase $k \pm q$ attached which may differ from $T_{\pm}(k)$ by a reciprocal lattice vector. This substitution eliminates the extra phases and gives the following expressions for the $\frac{\partial^2 E}{\partial \psi \psi}$:

$$(d_Y \nabla E)_{k \pm q, k} = -O dC_{k \pm q, k} \tilde{H}_k F - O C_{k \pm q} C_{k \pm q}^\dagger dH_{k \pm q, k} C_k F + dH_{k \pm q, k} C_k F + H_{k \pm q} dC_{k \pm q, k} F, \quad (2.60)$$

$$(dn)_{\pm q} = \sum_k \text{diag} [(IdC_{k \mp q, k})^\dagger F I C_k + (I C_k)^\dagger F I dC_{k \pm q, k}], \quad (2.61)$$

$$(dV_{sc})_{\pm q} = -4\pi J^\dagger O L_{\pm q}^{-1} O J dn_{\pm q} + J^\dagger O J \text{Diag}[E'_{xc}(n)] dn_{\pm q} \quad (2.62)$$

$$+ \text{Diag}[E'_{xc}(n)] J^\dagger O J dn_{\pm q} + \text{Diag}[E''_{xc}(n) \cdot dn_{\pm q}] J^\dagger O J n,$$

$$dH_{k \pm q, k} = I^\dagger \text{Diag}((dV_{sc})_{\pm q}) I. \quad (2.63)$$

For the derivative $\frac{\partial^2 E}{\partial \tau \psi}$ we find

$$(d_\tau \nabla E)_{k \pm q, k} = dH_{k \pm q, k} C_k F - O C_{k \pm q, k} C_{k \pm q, k}^\dagger dH_{k \pm q, k} C_k F. \quad (2.64)$$

2.3 Minimization Algorithms for Incommensurate Perturbations

It was observed during testing that numerical inaccuracies resulting from the projection operator tends to result in a linear system corresponding to (2.9) that does not converge extremely well under conjugate gradients. We found that the MINRES algorithm is a more robust way to compute incommensurate perturbations, and one extra benefit of MINRES is that it can solve indefinite linear systems

whereas conjugate gradients only works with positive definite systems. Both algorithms are included in the JDFTx implementation of incommensurate perturbation theory.

2.3.1 Application to dielectric response

This method can be used to calculate responses to non-periodic potentials such as point charge perturbations when we wish to calculate the dielectric response of a material. The idea is that any perturbation $V(r)$ can be written as a Fourier sum

$$V(r) = \int \tilde{V}(q)e^{iqr} \propto \int_{BZ} \sum_G \tilde{V}(G+q)e^{i(G+q)r} \quad (2.65)$$

$$= \int_{BZ} \left[\sum_G \tilde{V}(G+q)e^{iGr} \right] e^{iqr}. \quad (2.66)$$

Hence $V(r)$ can be decomposed into a combination of periodic perturbations with an extra phase factor e^{iqr} , which can be handled by VPT. Let us take a brief look at the computational complexity of the new technique compared to a simple supercell calculation. In general, the computational time of a DFT calculation scales as N^3 where N is the number of atoms in the unit cell. If we go to an $n \times n \times n$ supercell, we have $N = n^3$ times more atoms and an N^3 times cost, whereas if we do n^3 calculations with a q mesh, we have simply $N \cdot N$ cost that we need to worry about, so we have a factor of N speed up.

CHAPTER 3
WANNIER FUNCTION BASED THEORY OF
INCOMMENSURATE SYSTEMS

3.1 Introduction

In the past decade or so, there has been an explosion in research studying the properties of 2D layered materials, which are created by taking single-atom thick sheets from a larger crystal and stacking two or more of these sheets on top of one another. Recently, Moiré materials, a specific class of layered materials, have garnered an especially large amount of attention. These are characterized by two layers that are rotated relative to one another so that their unit cells no longer align. If the angle of rotation is small, then the lattices will form interesting large-scale patterns caused by the relative (mis)alignment of the atoms. The long-range interactions arising from these Moiré patterns can lead to a host of interesting and unique effects, including superconductivity, ferromagnetism, and topological conducting channels [6]. In twisted bilayer graphene (BLG), people have found additional phenomena such as interaction-induced insulating states and anomalous quantum hall states [1].

The first models of such systems were built by coming up with a continuum Hamiltonian description of the layers and their interactions while ignoring the short-scale atomic structure [2]. The work of Bistritzer and MacDonald showed using this model that there exist certain “magic angles” at which the velocity at the Dirac point vanishes leading to a large density of states and counterflow conductivity. Subsequent works focused on developing better models of twisted BLG: One such study incorporated *ab initio* calculations into a tight binding model

to study the behavior of bands near the fermi level for certain angles [9].

There does exist a limitation to all the studies discussed so far, however, which is that none of them include self-consistency. If we wish to have a short wavelength, self-consistent model for mismatched systems incorporating *ab initio* calculations, then we can apply mismatched interface theory (MINT), which was recently developed here at Cornell. MINT works by superposing a small flake of one material on an infinite lattice of the other and examining the convergence of physical quantities as the flake size grows to infinity [5]. There are a few drawbacks to this technique that hamper its usefulness in certain situations. For example, one of the quantities we are interested in is the momentum resolved density of states (DOS) which is essentially an unfolding of the band structure. If we compute either the band structure directly or momentum resolved DOS using Wannier functions, we see a large number of completely flat bands. This is an effect of flake quantization that arises because different sides of the flake have no way of communicating with each other, and is referred to as the flat band problem. An additional barrier is that computational complexity increases rapidly with an increasing number of atoms. This makes large flake calculations expensive to perform.

3.2 Ab Initio Tight Binding Models

Creating an *ab initio* tight binding model involves performing a density-functional theory calculation of the flake material on an infinite substrate. From this, maximally-localized Wannier functions are computed from the output of the electronic structure calculation. The matrix elements of the Wannier Hamiltonian then be used as the tight-binding Hamiltonian. The issues with this approach are

the same as the ones mentioned previously. The limits of computational complexity means we can only feasibly perform small flake calculations (~ 20 flake atoms). Our group is in the process of developing a technique for overcoming this issue that involves performing multiple flake calculations for different parts of the flake material and stitching the resulting Hamiltonians together to effectively model a very large flake. This technique, which we refer to as MINT Quilt, gives us a way to efficiently create *ab initio* tight binding Hamiltonians, however the issue of flat bands still remains, even in large stitched flake systems. An approach to solving the flat band problem is by imposing custom boundary conditions on the flake edges to try and emulate an infinite system. We will discuss both how to impose general constraints and the specific flake boundary conditions in the following sections.

3.3 Variational Constraints in DFT

When thinking about how to impose constraints in the wavefunctions during minimization, it is helpful to look at a constraint that already exists. Consider the orthonormality requirement for the wavefunctions C . This constraint takes the form $C^\dagger C = \mathbb{1}$ and it can be enforced introducing a Lagrange multiplier to the total energy which we denote E_{DFT} . The Lagrangian is written as

$$\mathcal{L} = E_{\text{DFT}} + \text{Re tr}[\Lambda^\dagger(C^\dagger C - \mathbb{1})] \quad (3.1)$$

where Λ is an $N \times N$ hermitian matrix with each component being a separate Lagrange multiplier. Now the minimum of E_{DFT} over the set of orthonormal states is the stationary point of the Lagrangian \mathcal{L} . The gradient of the constraint

function is

$$d(\text{Re tr}[\Lambda^\dagger(C^\dagger C - \mathbb{1})]) = \text{Re tr}[\Lambda^\dagger(dC^\dagger C + C^\dagger dC)] \quad (3.2)$$

$$= \text{Re tr}[dC^\dagger C \Lambda^\dagger + (C \Lambda^\dagger)^\dagger dC] = 2 \text{Re tr}[dC^\dagger C \Lambda^\dagger]. \quad (3.3)$$

Hence, the gradient of the Lagrangian is

$$\nabla_C \mathcal{L} = HC + C \Lambda^\dagger. \quad (3.4)$$

The gradient must vanish when the wavefunctions minimize energy. Let us denote the minimizing wavefunctions by C_* . Setting the gradient to zero and multiplying by C_*^\dagger gives

$$C_*^\dagger HC_* + C_*^\dagger C_* \Lambda^\dagger = 0 \implies \Lambda^\dagger = -C_*^\dagger HC_*. \quad (3.5)$$

During, we do not know what C_* is, however if we are close to the minimum, we can write $C = C_* + \Delta C$ where and $|\Delta C|/|C_*| \ll 1$. Therefore, we know what Λ is approximately:

$$\Lambda^\dagger = -C^\dagger HC + \mathcal{O}(|\Delta C|). \quad (3.6)$$

Thus,

$$\nabla_C \mathcal{L} \approx HC - CC^\dagger HC \equiv HC - C\tilde{H}. \quad (3.7)$$

At the energy minimum, we can choose to diagonalize \tilde{H} by a unitary rotation of C , in which case case, H_{sub} turns into a diagonal matrix ϵ and the equation

$$\nabla_C \mathcal{L} = 0 \implies HC = C\epsilon \quad (3.8)$$

is just an eigenvector equation for the columns of C , and this shows that the Lagrange multipliers turn into the energy eigenvalues ϵ at the minimum. This corresponds to the equation $H|\psi_i\rangle = \epsilon_i|\psi_i\rangle$ in conventional notation.

3.3.1 Arbitrary constraint

We can apply this same method to enforce a constraint of the type

$$U^\dagger C = 0. \quad (3.9)$$

As before, we introduce a Lagrange multiplier that goes along with the original Lagrange Multiplier.

$$\mathcal{L} = E_{\text{DFT}} + \text{Re tr}[\Lambda^\dagger(C^\dagger C - \mathbb{1})] + 2 \text{Re tr}[C^\dagger U \Lambda_U]. \quad (3.10)$$

Taking the gradient with respect to wavefunctions gives

$$\nabla_C \mathcal{L} = HC + C\Lambda^\dagger + U\Lambda_U \quad (3.11)$$

We can solve for the Lagrange multipliers in a similar way to the previous part.

Take C_* to be the ground state wavefunctions. Then

$$HC_* + C_*\Lambda^\dagger + U\Lambda_U = 0. \quad (3.12)$$

Multiplying by C_*^\dagger still fixes $\Lambda^\dagger = -C_*^\dagger HC_*$ as we had before. Now, we can instead multiply by U^\dagger . This eliminates the middle term, leaving

$$(U^\dagger U)\Lambda_U = -U^\dagger HC_* \implies \Lambda_U = -(U^\dagger U)^{-1}U^\dagger HC_*. \quad (3.13)$$

Hence,

$$\nabla_C \mathcal{L} = HC - CH_{\text{sub}} - U(U^\dagger U)^{-1}U^\dagger HC_*. \quad (3.14)$$

This modification to the energy gradient allows us to perform energy minimization under an arbitrary constraint. If U is orthonormal and $C \approx C_*$ then $(U^\dagger U)^{-1} = \mathbb{1}$ and the expression simplifies to

$$\nabla_C \mathcal{L} = HC - C\tilde{H} - UU^\dagger HC. \quad (3.15)$$

During minimization we can follow the energy gradient while keeping the wavefunctions on both constraint manifolds all the time. For the orthonormality constraint this is enforced by orthonormalizing the wavefunctions at each step. For the arbitrary constraint, we can do something similar and project out the columns of U :

$$C \rightarrow (\mathbb{1} - U(U^\dagger U)^{-1}U^\dagger)C. \quad (3.16)$$

Alternatively, if we are solving a tight binding model then we know the entire matrix of H and we can rewrite the stationary condition as

$$C_*\epsilon = (\mathbb{1} - UU^\dagger)HC_*. \quad (3.17)$$

assuming $\tilde{H} = \epsilon$ has been diagonalized. We also use the condition that C_* satisfies the constraint to rewrite the equation as

$$C_*\epsilon = (\mathbb{1} - UU^\dagger)H(\mathbb{1} - UU^\dagger)C_*. \quad (3.18)$$

This explicitly turns it into an eigenvector problem for the hermitian matrix $(\mathbb{1} - UU^\dagger)H(\mathbb{1} - UU^\dagger)$. We can solve it for the ground state and discard all the eigenvectors that do not satisfy the constraints.

3.3.2 Flake Boundary Conditions

Using the previous result allows us to develop and implement boundary conditions on the flake. Assume for simplicity that the system contains a one-dimensional flake with a single Wannier function on the leftmost and rightmost atom which we call $|W_L\rangle$ and $|W_R\rangle$ respectively. We furthermore assume that the states of the full incommensurate system can be described by a Bloch wavevector k . It then follows that the projection of an eigenstate $|\psi\rangle$ on the rightmost edge should be equal to

the projection on the left edge, plus an additional Bloch phase. If W is the width of the flake the wavefunction should acquire the phase e^{ikW} hence the condition is

$$\langle W_R|\psi\rangle - e^{ikW}\langle W_L|\psi\rangle = 0. \quad (3.19)$$

This corresponds to the constraint matrix

$$U = |W_R\rangle - e^{-ikw}|W_L\rangle. \quad (3.20)$$

We now generalize this boundary condition to a two dimensional rectangular flake. We imagine copies of the rectangular flake tiling all of space where the left and right boundaries are identified with each other as well as the top and bottom boundaries. We then have a set of atom pairs that are supposed to be connected to each other. Let us label the corresponding Wannier functions by the pairs $(|W_i^A\rangle, |W_i^B\rangle)$. We then have N separate constraints

$$U_i = |W_i^B\rangle - e^{-ik(x_i^B - x_i^A)}|W_i^A\rangle \quad (3.21)$$

and the matrix U consists of each constraint stacked together:

$$U = \begin{bmatrix} U_1 & \dots & U_n \end{bmatrix}. \quad (3.22)$$

This boundary condition is intended to connect the flake edges in a natural way and reduce the quantization effects.

3.4 Preliminary Results for Twisted Bilayer Graphene

We applied the method to bilayer graphene using a tight binding Hamiltonian constructed with the MINT Quilt method. The flake system under consideration had a 6×6 unit cell graphene flake with 98 atoms, some extra atoms lying on

the boundary of the flake (see 3.1). We computed a momentum resolved density of states plot through reciprocal space that passed through the most important points (3.2). A discussion of the momentum resolved can be found in the appendix. Now the first thing we notice from the plot is the presence of many flat bands. We can see the band structure of the tight binding model shows a sharp band coming from the substrate layer as well as a fuzzy flake band that emerges from the many flat bands.

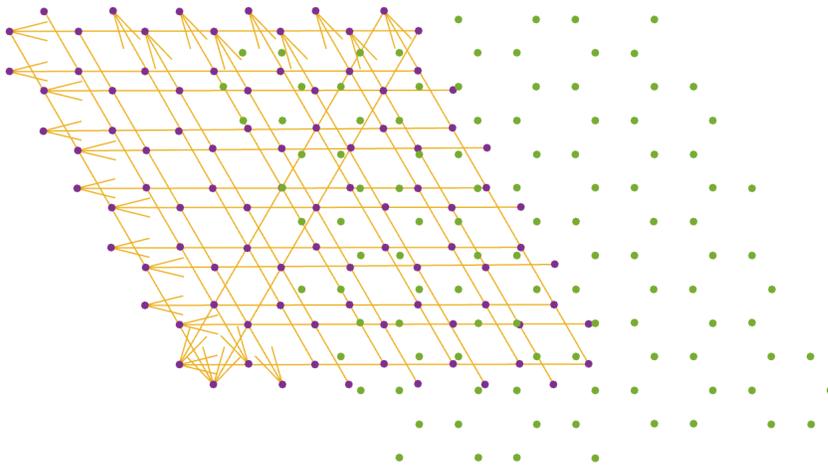


Figure 3.1: **Graphical depiction of MINT Flake system.** (Light green) Substrate unit cell atoms, (Dark purple) Flake atoms, (Arrows) Flake atoms connected by boundary conditions

We now turn on the boundary conditions and find that the flat bands attain significant curvature (3.3). In addition the band from the flake is much better resolved. This indicates that the periodic boundary conditions do help to create a more sharply defined band structure. One thing we do not see is the mirrored Dirac cone observed in experiment [11]. This shows that work is still needed to have this method produce physical results.

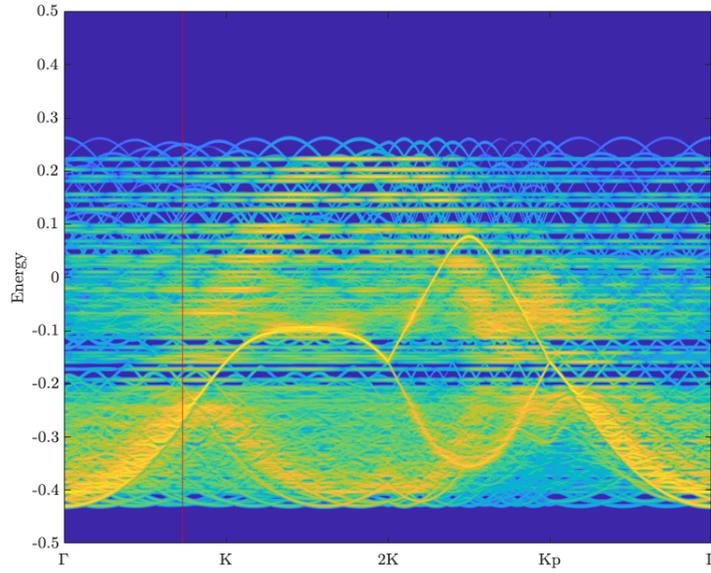


Figure 3.2: **Momentum resolved density of states for graphene bilayer flake system, no boundary conditions applied.** k-path goes through BZ of substrate. Intensity is plotted on a log scale. Red vertical line shows the expected location of mirrored Dirac cone.

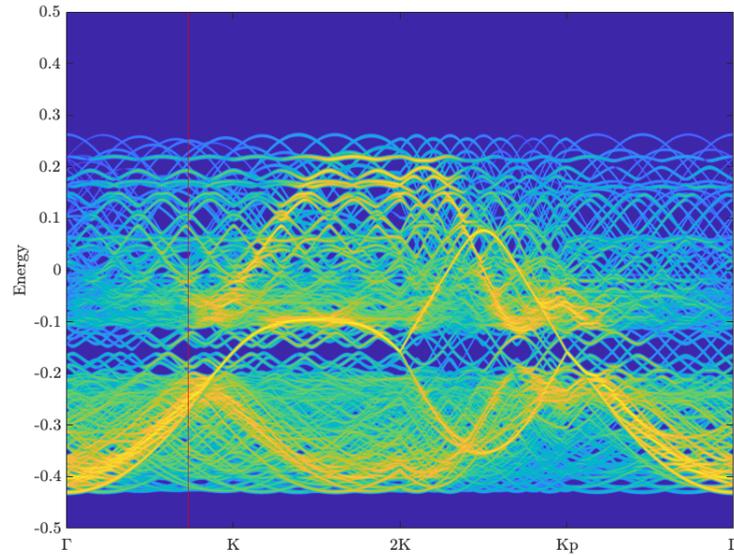


Figure 3.3: **Momentum resolved density of states for graphene bilayer flake system, with periodic boundary conditions.** k-path goes through BZ of substrate. Intensity is plotted on a log scale. Red vertical line shows the expected location of mirrored Dirac cone.

CHAPTER 4

CONCLUSION

We have shown that it is possible to adapt existing electronic structure techniques to calculate properties of non-periodic systems. Starting with the algebraic formulation of DFT, we derived perturbation theory techniques to find the responses of periodic systems to both periodic and non-periodic perturbations. We also outlined a method to optimize high-entropy alloys by replacing atoms with effective atoms and derived perturbation theory expressions to calculate the gradient of the free energy with respect to mixing parameters. We demonstrated the validity of this technique in a couple of simple test systems. Finally, we described a technique to generate large *ab initio* tight binding Hamiltonians from DFT calculations for incommensurate systems. We combined this with specialized flake boundary conditions to produce a better resolved flake band structure.

There is still work to be done in both variational perturbation theory and improving the MINT flake boundary conditions. The most important feature yet to be supported in VPT is solving for perturbations in metallic systems. Here the fillings are depend on the energy eigenvalues making the gradient more difficult to compute. Possible approaches for handling variables fillings are discussed in the appendix. On the other side, the flake boundary conditions are also not perfect. The band structure plot (3.3) shows what appears to be avoided crossings in the flake band structure, which ideally should not be present. Improved boundary conditions should lead to more symmetry between the flake and substrate bands.

5.1 Summary of Algebraic DFT Notation [10]

In algebraic DFT notation, all quantities are stored in complex matrices. States are encoded with a basis set $|m\rangle, 1 \leq m \leq M$ where M is the total number of basis elements. Let us now consider an energy minimization problem in which we minimize the N lowest energy eigenstates. These states $|\psi_n\rangle, 1 \leq n \leq N$ can be thought of as column vectors with length M . The combined state of the system is stored in an $M \times N$ matrix defined by

$$C_{mn} = \langle m|\psi_n\rangle. \quad (5.1)$$

We can alternatively think of this as the result of packing all the column vectors ψ_i into a single matrix:

$$C = \begin{bmatrix} \uparrow & & \uparrow \\ \psi_1 & \dots & \psi_n \\ \downarrow & & \downarrow \end{bmatrix}. \quad (5.2)$$

We call C a column bundle. The ground state is given by minimizing the density-functional energy $E_{\text{DFT}}(C)$ over the wavefunctions which includes contributions from external potentials, exchange-correlation and the Coulomb interaction. Assuming C is normalized, the energy is given by $E_{\text{DFT}} = \text{tr}(C^\dagger H C F)$. Here F is the fillings matrix which contains the number of electrons per state. The gradient of this energy is given by

$$\nabla_C E_{\text{DFT}} = H C F. \quad (5.3)$$

This gradient satisfies the requirement that the inner product of the gradient and a direction gives the change of the energy in that direction:

$$dE_{\text{DFT}} = 2 \text{Re tr}[dC^\dagger(\nabla_C E_{\text{DFT}})] \equiv \langle dC^\dagger, \nabla_C E_{\text{DFT}} \rangle \quad (5.4)$$

Here 2Re tr is the analog of the dot product, meaning

$$\text{Re tr}(A^\dagger B) = \sum_{ij} [\text{Re}(A_{ij}) \text{Re}(B_{ij}) + \text{Im}(A_{ij}) \text{Im}(B_{ij})]. \quad (5.5)$$

If we furthermore impose the condition that the gradient has lie in the tangent space of the set of orthonormal wavefunctions, then it must be projected:

$$\nabla_C E_{\text{DFT}} = (\mathbb{1} - OCC^\dagger)HCF. \quad (5.6)$$

Here O is the overlap operator, proportional to the identity in the plane wave basis. Now, if fillings are not constant, we have to add an extra term to account for subspace rotations. Also if the wavefunctions are not normalized, we have to add a normalization term as well. Unnormalized wavefunctions are represented by the letter Y and are related to the normalized C coefficients by

$$C = YU^{-1/2} \quad U = Y^\dagger OY. \quad (5.7)$$

In this case the full expression for the gradient with respect to the unnormalized coefficients with k -points included is

$$\nabla_{Y_k} E = (\mathbb{1} - OC_k C_k^\dagger)H_k C_k F U_k^{-1/2} + OC_k Q_k([\tilde{H}_k, F]), \quad (5.8)$$

where $\tilde{H} \equiv C^\dagger H C$ is called the subspace Hamiltonian. The full Hamiltonian H itself is composed of energy terms coming from the electron-nuclear interaction, electron-electron, and exchange correlation terms. There is also an external potential term V_{ext} . Explicitly, we write it as

$$H_k = -\frac{1}{2}L_k + I^\dagger(\text{Diag } V_{\text{sc}})I, \quad (5.9)$$

$$V_{\text{sc}} = \underbrace{J^\dagger O J V_{\text{nuc}}}_{\text{nuclear pot.}} + \underbrace{J^\dagger O \hat{\phi}}_{\text{electronic pot.}} + \underbrace{J^\dagger O J E_{xc}(n) + \text{Diag}[E'_{xc}(n)] J^\dagger O J n}_{\text{exchange-correlation term}} + V_{\text{ext}}. \quad (5.10)$$

We also have expressions for the Coulomb potential $\hat{\phi}$ and the density n :

$$\hat{\phi} = -4\pi L^{-1} O J n, \quad (5.11)$$

$$n = \sum_k w_k \text{diag}(I C_k F C_k^\dagger I^\dagger). \quad (5.12)$$

5.2 Derivative of the Q operator

The Q operator arises when the fillings are fixed but not constant. The derivative dQ comes into play when we need to take the second derivative of energy. First we give a formula for $Q_U(A)$. Let V be the matrix containing the eigenvectors of U and μ be the corresponding ordered list of eigenvalues. Define the matrix M to have the components

$$M_{nm} = \frac{1}{\sqrt{\mu_n} + \sqrt{\mu_m}}. \quad (5.13)$$

Then,

$$Q_U(A) = V[(V^\dagger A V) \cdot M] V^\dagger, \quad (5.14)$$

so

$$d(Q_U(A)) = dV[(V^\dagger A V) \cdot M] V^\dagger + dV[(V^\dagger A V) \cdot M] dV^\dagger \quad (5.15)$$

$$+ V[(dV^\dagger A V + V^\dagger dA V + V^\dagger A dV) \cdot M] V^\dagger + V[(V^\dagger A V) \cdot dM] V^\dagger. \quad (5.16)$$

Now, we must determine the expressions for the quantities dV and dM which are related to the first order change in eigenvectors and eigenvalues respectively. We see from [8] that

$$dV = V[(V^\dagger dU V) \cdot N] \quad N_{nm} = \begin{cases} \frac{1}{\mu_m - \mu_n} & n \neq m \\ 0 & n = m \end{cases}. \quad (5.17)$$

and also by taking a derivative,

$$dM_{nm} = -\frac{1}{2} \frac{d\mu_n/\sqrt{\mu_n} + d\mu_m/\sqrt{\mu_m}}{(\sqrt{\mu_n} + \sqrt{\mu_m})^2}, \quad (5.18)$$

$$d\mu_n = (V^\dagger dUV)_{nn}. \quad (5.19)$$

There is a special case in which $U = \mathbb{1}$. Here, $V = \mathbb{1}$ and all the eigenvalues are 1 so $Q_U(A) = \frac{1}{2}A$. Also, when calculating $d(Q_U(A))$ we see that N_{nm} is no longer well defined due to the denominators being zero. This is the same issue that arises in regular perturbation theory due to degeneracies. The solution in this case is just to see that we can set V to also be the eigenvectors of dU in which case $dV = 0$. Now, it is true that $dV = 0$ no matter what basis we're working in, and thus we can choose $V = \mathbb{1}$ again. Finally, this means

$$d(Q_U(A)) = V[(V^\dagger AV) \cdot dM]V^\dagger + \frac{1}{2}dA. \quad (5.20)$$

We can further simplify this expression to the following form:

$$-\frac{1}{8}(dUA + AdU) + \frac{1}{2}dA = \frac{1}{2}dA - \frac{1}{8}\{A, dU\}. \quad (5.21)$$

5.3 Metallic Fillings

Handling metallic fillings is most easily done by introducing a new set of variables that parameterize the fillings. The JDFTx software uses a matrix called the auxiliary Hamiltonian whose eigenvalues map onto the fillings via the smearing function [3]. Full support of metallic systems would involve taking second derivatives with respect to both the wavefunctions and the auxiliary Hamiltonian.

5.4 Momentum Resolved Density of States

The momentum resolved density of states is a function that tells us about the distribution of quantum states in the space of energy and momentum. This construct is useful both because it closely matches the raw data produced by experimental techniques such as ARPES, and it is defined even when the wavefunctions do not obey Bloch's theorem. It is defined as

$$n(\mathbf{q}, \epsilon) = \sum_i \left| \int d\mathbf{r} \psi_i(\mathbf{r}) e^{-i\mathbf{q}\mathbf{r}} \delta(\epsilon - \epsilon_i) \right|^2 \quad (5.22)$$

where i labels the eigenstates $|\psi_i\rangle$ with energies ϵ_i .

BIBLIOGRAPHY

- [1] Eva Y Andrei and Allan H MacDonald. Graphene bilayers with a twist. *Nat. Mater.*, 19(12):1265–1275, 2020.
- [2] Rafi Bistritzer and Allan H MacDonald. Moiré bands in twisted double-layer graphene. *Proc. Natl. Acad. Sci. U. S. A.*, 108(30):12233–12237, 2011.
- [3] Christoph Freysoldt, Sixten Boeck, and Jörg Neugebauer. Direct minimization technique for metals in density functional theory. *Phys. Rev. B*, 79:241103, Jun 2009.
- [4] Easo P. George, Dierk Raabe, and Robert O. Ritchie. High-entropy alloys. *Nature Reviews Materials*, 4(8):515–534, 2019.
- [5] Eli Gerber, Yuan Yao, Tomas A. Arias, and Eun-Ah Kim. Ab initio mismatched interface theory of graphene on α - ruct_3 : Doping and magnetism. *Phys. Rev. Lett.*, 124:106804, Mar 2020.
- [6] 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, 2021. PMID: 33769797.
- [7] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.
- [8] Sohrab Ismail-Beigi and T.A. Arias. New algebraic formulation of density functional calculation. *Computer Physics Communications*, 128(1–2):1–45, June 2000.
- [9] E. Suárez Morell, J. D. Correa, P. Vargas, M. Pacheco, and Z. Barticevic. Flat bands in slightly twisted bilayer graphene: Tight-binding calculations. *Phys. Rev. B*, 82:121407, Sep 2010.
- [10] Ravishankar Sundararaman, Kendra Letchworth-Weaver, Kathleen A. Schwarz, Deniz Gunceler, Yalcin Ozhabes, and T.A. Arias. Jdftx: Software for joint density-functional theory. *SoftwareX*, 6:278–284, 2017.
- [11] Wei Yao, Eryin Wang, Changhua Bao, Yiou Zhang, Kenan Zhang, Kejie Bao, Chun Kai Chan, Chaoyu Chen, Jose Avila, Maria C. Asensio, Junyi Zhu,

and Shuyun Zhou. Quasicrystalline 30° twisted bilayer graphene as an incommensurate superlattice with strong interlayer coupling. *Proceedings of the National Academy of Sciences*, 115(27):6928–6933, June 2018.