Electronic structure theory of weakly interacting bilayers

Shiang Fang¹ and Efthimios Kaxiras¹,²

¹Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA
²John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

(Received 18 April 2016; published 27 June 2016)

We derive electronic structure models for weakly interacting bilayers such as graphene–graphene and graphene–hexagonal boron nitride, based on density functional theory calculations followed by Wannier transformation of electronic states. These transferable interlayer coupling models can be applied to investigate the physics of bilayers with arbitrary translations and twists. The functional form, in addition to the dependence on the distance, includes the angular dependence that results from higher angular momentum components in the Wannier $p_z$ orbitals. We demonstrate the capabilities of the method by applying it to a rotated graphene bilayer, which produces the analytically predicted renormalization of the Fermi velocity, Van Hove singularities in the density of states, and moiré pattern of the electronic localization at small twist angles. We further extend the theory to obtain the effective couplings by integrating out neighboring layers. This approach is instrumental for the design of van der Walls heterostructures with desirable electronic features and transport properties and for the derivation of low-energy theories for graphene stacks, including proximity effects from other layers.

DOI: 10.1103/PhysRevB.93.235153

I. INTRODUCTION

Two-dimensional layered materials are becoming the focus of experimental and theoretical investigations aiming to realize the potential applications of these atomically thin structures [1–3]. The library of these layered materials is still expanding with properties that range from metals and semimetals to semiconductors and insulators. Graphene, a semimetallic atomically thin sheet of carbon atoms in the honeycomb lattice [4,5], is an important member of the layered materials family. Flakes of graphene can be obtained by the exfoliation method from a graphite crystal [6] or synthesized by methods such as chemical vapor deposition [7]. In addition to its outstanding electronic and mechanical properties, graphene is also an interesting platform to investigate the quasirelativistic strongly interacting many-body physics near the charge-neutrality point (CNP) when screening between charges is reduced [8]. This electron-hole plasma, known as the Dirac fluid, behaves differently from the conventional Fermi liquid, when the Fermi level is far from the CNP. The inclusion of graphene in the van der Waals (vdW) heterostructures, that is, stacks of various layered materials, can serve several purposes, for instance as an active layer, a spacer, or an electrode. The different layered materials exhibit a wide variety of physical properties such as topological phases [9], superconductivity [10], magnetism [11], and charge density waves [12]. Different ways of stacking, manipulating these materials, and intercalating with foreign atoms in the vdW heterostructure open even wider possibilities for interesting physics phenomena and for novel nanodevice applications [2,13]. Other control knobs include electrical gating, an external magnetic field, and various contacts and twist angles which affect strongly the Brillouin zone (BZ) alignment and the coupling between layers [14].

To investigate the electronic properties of these fascinating systems, large-scale density functional theory (DFT) [15,16], tight-binding models [17–20], and low-energy $k \cdot p$ expansions [21–24] have been employed. The parameter-free DFT approach is computationally demanding, while the computationally efficient tight-binding methods and low-energy $k \cdot p$ expansions are hampered by the absence of a universal form of the interlayer couplings. The interlayer couplings employed by empirical methods are often parametrized as functions of only the interatomic distances or more elaborate forms that depend on the local bonding environment [20,25] with the values of parameters obtained by fitting the band structure of selected crystal configurations. A set of such interlayer hopping terms for graphene has been determined from $ab$ initio calculations, but they are extracted only from a restricted subset of all possible bilayer orientations [26,27]. The dependence of interlayer hopping on both the distance between pairs of atoms and the relative orientation of bonds, as exemplified by the $\gamma_3$ and $\gamma_4$ terms in the Slonczewski-Weiss-McClure model [5], have not yet been addressed properly. An accurate and transferable theory of interlayer coupling would not only provide an efficient way of evaluating electronic properties, but would also shed light on transport properties across layers [28] and on the derivation of effective low-energy theories for arbitrary graphene stacking sequences.

We provide here a comprehensive and quantitative understanding of interlayer coupling in two prototypical bilayers, graphene–graphene (G-G) and graphene–hexagonal boron nitride (G-hBN). This type of $ab$ initio modeling based on the Wannier transformation is applied to derive a transferable potential applicable to bilayer configurations with arbitrary translations and twists between the two layers. In contrast to similar analysis applied to transition-metal dichalcogenides, in which interlayer coupling was shown to have a simple, orientation-independent scaling form that depends only on the distance between pairs of atoms [29], the G-G and G-hBN interlayer couplings include a dependence on both pair distances and relative orientations. The ensuing angular dependence is related to the crystal field distortions of the atomic $p_z$ Wannier orbitals. We derive the form of effective coupling terms that is suitable in the general vdW heterostructure with arbitrary twists and translations when some layers are integrated out. This scheme is relevant for obtaining the proximity effects on graphene due to other layered materials in the vdW heterostructure. These interlayer coupling models provide...
efficient ways for obtaining electronic properties and effective low-energy theories, as well as for estimating proximity effects in vdW heterostructures, especially when graphene layers are included in the stacks.

The paper is structured as follows: In Sec. II, we introduce the numerical methods employed for the DFT calculations and the Wannier transformation. In Sec. III, we investigate the Hamiltonians for weakly interacting bilayers, using G-G and G-hBN interfaces as the prototypical examples. In Sec. IV we discuss the physics of twisted bilayer graphene and in Sec. V we derive the effective theory for proximity effects by integrating out the neighboring layers. Our concluding Sec. VI gives a summary of the main points, makes comparisons with similar approaches in the literature, and contains some remarks on possible extensions and future applications.

II. NUMERICAL METHODS

The approach we adopt here is to derive the \textit{ab initio} tight-binding Hamiltonian based on the Wannier transformation of DFT calculations. Within DFT, we obtain the Bloch wave functions and energies using VASP \cite{vasp1,vasp2} with pseudopotentials of the projector augmented wave (PAW) type, the exchange-correlation functional of Perdew, Burke, and Ernzerhof (PBE) \cite{pbe}, a plane-wave energy cutoff 500 eV, and a \(17 \times 17 \times 1\) reciprocal space grid. A 20 Å distance is used to eliminate the coupling between periodic images of the layers in the direction perpendicular to the atomic planes. The diagonal Kohn-Sham Hamiltonian in the Bloch basis from the layers in the direction perpendicular to the atomic planes. The used to eliminate the coupling between periodic images of the exchange-correlation functional of Perdew, Burke, and pseudopotentials of the projector augmented wave (PAW) type, and the coupling dependence is characterized by the projected distance \(\theta\).

When two or more monolayers are brought into contact, the shape of the Wannier function has implications for the interlayer coupling. These couplings are described by the matrix elements: \(\langle \psi_i | H | \psi_j \rangle\) with \(\psi_i\) the Wannier orbital of the first (second) layer and \(H\) the total Hamiltonian. The angular momentum mixing as shown in Fig. 1(b) for the Wannier orbital in graphene translates into the angular dependence of such interlayer couplings, in addition to the usual dependence on the distance of the pair. Without loss of generality, we assume the projected vector \(r\) from \(\psi_i\) to \(\psi_j\) on the plane is along the positive \(x\) axis, and \(\theta_i\) (\(\theta_j\)) is the angle relative to \(r\) needed to determine the orientation of the crystal of the layer to which the Wannier orbital \(\psi_i\) belongs. The interlayer coupling can then be written as the function \(t(r, \theta_i, \theta_j)\). If the underlying crystal and the embedded Wannier orbital has \(N\)-fold rotation symmetry, then \(\theta\) is only defined up to modulo \(2\pi/N\). The above interlayer coupling

III. HAMILTONIAN FOR WEAKLY INTERACTING BILAYERS

Before modeling the interlayer coupling, we first reconstruct the \textit{ab initio} tight-binding Hamiltonian \(H_0\) for a graphene monolayer \cite{graphene_monolayer}. The unit cell for monolayer graphene is spanned by \(a_1 = (\sqrt{3} \hat{x} - \hat{y})a/2\) and \(a_2 = (\sqrt{3} \hat{x} + \hat{y})a/2\) with the lattice constant \(a = 2.46\ \text{Å}\). Two basis atoms are situated at \(\delta_A = 0\) and \(\delta_B = (a_1 + a_2)/3\). We extract the intralayer couplings up to the eighth nearest neighbors, which shows good agreements with DFT results. The numerical parameters for \(t_i\), the intralayer hopping parameter to the \(i\)th nearest neighbor, are listed in the left block \(\mathcal{H}_0\) of Table I.

The shape of the localized basis, also known as the Wannier orbital, provides intuition for the chemical bonding, hybridization, and the symmetry of the crystal. For monolayer graphene, the constructed Wannier orbital has a dominant \(p_z\) character but the azimuthal symmetry is broken by the crystal field distortion from the neighboring atoms. Locally, at the position of the carbon atom, the threefold rotation symmetry is restored. Thus, the angular momentum is defined up to modulo 3, which means that there is hybridization within each sector of angular momentum states. In Fig. 1(b), we decompose the Wannier function for graphene into \(m = 0\) (dominant \(p_z\)), \(m = \pm 3\), and \(m = \pm 6\) angular momentum components. This decomposition shows the range and the strength of each component, and the characteristic radius gets larger for larger angular momentum components.

<table>
<thead>
<tr>
<th>(H_0)</th>
<th>(H')</th>
<th>(V_0(r))</th>
<th>(V_5(r))</th>
<th>(V_6(r))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t_1)</td>
<td>-2.8922</td>
<td>(t_5)</td>
<td>0.0524</td>
<td>(\lambda_1)</td>
</tr>
<tr>
<td>(t_2)</td>
<td>0.2425</td>
<td>(t_6)</td>
<td>-0.0209</td>
<td>(\xi_1)</td>
</tr>
<tr>
<td>(t_3)</td>
<td>-0.2656</td>
<td>(t_7)</td>
<td>-0.0148</td>
<td>(x_1)</td>
</tr>
<tr>
<td>(t_4)</td>
<td>0.0235</td>
<td>(t_8)</td>
<td>-0.0211</td>
<td>(\kappa_1)</td>
</tr>
</tbody>
</table>

(a) (b)

FIG. 1. (a) The generic graphene bilayer configuration with arbitrary translations and twists. The constituent monolayer crystal is described by the primitive vectors \(a_1\) and \(a_2\). For an interlayer pair, the coupling dependence is characterized by the projected distance \(r\) and the angles \(\theta_{12}\) and \(\theta_{12}'\) between \(r\) and the nearest-neighbor bonds. (b) Decomposition of the Wannier function for monolayer graphene into the constant \((m = 0)\), \(\cos(3\theta)\) \((m = \pm 3)\), and \(\cos(6\theta)\) \((m = \pm 6)\) components.
can be simplified to

$$t(r, \theta_1, \theta_2) = \sum_{m_1, m_2 = -\infty}^{\infty} f_{m_1, m_2}(r) e^{im_1 \theta_1 + im_2 \theta_2}$$

with integers $m_i$. For real $t$, $f_{m_1, m_2}(r) = f_{m_1, m_2}^*(r)$. This decomposition can be viewed as the multichannel interlayer hopping process.

We next apply this general analysis to bilayers of graphene (G-G) and of graphene/hexagonal boron nitride (G-hBN). We consider two specific stackings, AA and AB, defined by the relative position of the basis $A$ or $B$ atom of the top layer to that of the basis $A$ atom of the bottom layer. The bilayers are assumed to be flat with the same constant interlayer bond distance $r$.

The interlayer hoppings are extracted from the interlayer bonds while keeping the underlying crystal orientation untouched. The interlayer hoppings are invariant under the sixfold rotation symmetry of the underlying crystal, these interlayer hoppings are invariant under $\theta \to 3\theta \to 6\theta$. For real $t$, these correspond to nonzero values for the terms $f_{0,0}, f_{\pm1,0}, f_{\pm2,0}$, and $f_{\pm2,\pm1}$, with $N_1 = N_2 = 3$. We use the following fitting functions for $V_i(r)$ with $\bar{r} = r/a$:

$$V_0(r) = \lambda_0 e^{-\xi_0 r^2} \cos(\kappa_0 \bar{r})$$
$$V_3(r) = \lambda_3 \bar{r}^2 e^{-\xi_3 r^2} \cos(\kappa_3 \bar{r})$$
$$V_6(r) = \lambda_6 \bar{r} e^{-\xi_6 r^2} \sin(\kappa_6 \bar{r})$$

These Fourier-projected components are plotted in Figs. 2(c) and 2(d) for the AA and AB stackings, respectively.

For the constant term, the AA/AB hoppings are very similar to each other, and we define $V_0(r)$ to be the average of the two. Projection into $\cos(3\theta)$ is significant for the AB type but vanishes identically for the AA type, and the curve in the AB case is defined as $2V_3(r)$. The two stackings have similar behavior for the much smaller $\cos(6\theta)$ term, and $2V_6(r)$ is the average of the two. The values of the fitting parameters are given in Table I from the analysis of the translated AA/AB structures.

A proper analysis of the inherent symmetry of the two-layer system provides the justification for the form of the interlayer hoppings and enables us to generalize the model to arbitrary configuration. Specifically, the sixfold symmetry of the crystal field allows mixing between $p_z$ and $m = \pm 3N$ orbitals with $N$ an integer. Due to the crystal symmetry, these components acquire a phase $(-1)^N$ when $A$ and $B$ basis atoms are interchanged, which are related by a $yz$ mirror operation. With the Wannier orbital viewed as composite objects of mixed angular momentum, the hopping between two such objects is determined by the superposition of the individual hopping channels between each component, within the two-center approximation [35]. The dominant channel is between the two $m = 0$ components which gives the constant part in the interlayer hopping. There is a $\cos(3\theta)$ term from the coupling between $m = 0$ and $m = \pm 3$ channels. Due to the symmetry, the two terms add up constructively (destructively) for the AB (AA) type. This can also be seen from the additional minus sign in exchanging $A$ and $B$ basis atoms. There are two types of contribution to the $\cos(6\theta)$ term, and they can be generated from the coupling between $m = 0$ and $m = \pm 6$, or between...
TABLE II. hBN intralayer and G-hBN interlayer tight-binding parameters. $\epsilon_g = 2.2021$ eV, $\epsilon_N = -1.9124$ eV with the convention $a = 2.46$ Å for $\tilde{r}$ in the fitting expression Eq. (3).

<table>
<thead>
<tr>
<th>$t_1$</th>
<th>$t_5$</th>
<th>$t_3$</th>
<th>$t_2$</th>
<th>$t_6$</th>
<th>$t_1$</th>
<th>$t_5$</th>
<th>$t_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-2.6490$</td>
<td>$0.3444 (0.0301)$</td>
<td>$0.3905$</td>
<td>$-2.0594 (0.2276)$</td>
<td>$-0.0374 (-0.0240)$</td>
<td>$1.5426$</td>
<td>$3.0827$</td>
<td>$3.7998$</td>
</tr>
<tr>
<td>$-0.2163$</td>
<td>$-0.0053$</td>
<td>$0.6085$</td>
<td>$0.0588$</td>
<td>$-0.0651$</td>
<td>$2.1909$</td>
<td>$0.5142$</td>
<td>$0.5264$</td>
</tr>
<tr>
<td>$0.0502$</td>
<td>$-0.0133$</td>
<td>$1.8229$</td>
<td>$0.3905$</td>
<td>$-0.0651$</td>
<td>$2.5177$</td>
<td>$0.3502$</td>
<td>$3.0464$</td>
</tr>
</tbody>
</table>

$m = \pm 3$ components of the two atoms. By symmetry, the first (second) part of the contribution is even (odd) in $AA/AB$. We can model the contribution of the first type from the average $\cos(6\theta)$ term of $AA/AB$ in Figs. 2(c) and 2(d). The channel between two $m = \pm 3$ can in general produce complicated angular dependence, but it is only a small correction (a few meV) and hence we ignore it in our model.

Following similar steps, we derive the tight-binding Hamiltonian for interlayer coupling in the case of a bilayer G-hBN. In the G-hBN interlayer coupling compared to the G-G coupling the symmetry is lower since the atoms are not identical, and this affects the amplitude for each angular momentum channel. From similar analysis for the shifted $AA/AB$ couplings in a single monolayer, which is quite different, showing the importance of having an accurate description of interlayer coupling.

IV. TWISTED BILAYER GRAPHENE PHYSICS

The coupling between layers is weak for $\theta \approx 30^\circ$, and gets stronger when the twist approaches angles near $0^\circ$ or $60^\circ$ [17]. When two layers are twisted, the bands are formed from the hybridization of monolayer bands [19] as in the schematic diagram of Fig. 4(a), inset. The characteristic kinetic energy scale is defined by $h\nu_F \Delta K$, with $\nu_F = 8.22 \times 10^3$ m/s the Fermi velocity of the monolayer graphene, and $\Delta K = |K - K'| = \frac{h}{\nu_F} \sin(\theta/2)$ the distance between displaced $K$ points of two layers ($K = \frac{2\pi}{a} \tilde{x}, K' = \frac{2\pi}{a} (\cos(\theta) \tilde{y} - \sin(\theta) \tilde{x})$). Under the hybridization at large twist angles, the bilayer bands retain the linear Dirac dispersion, but with a different slope around the Dirac point compared to the folded bands of the monolayer graphene as in Figs. 3(c) and 3(b). The Fermi velocity is renormalized by the interlayer coupling as can be seen from the effective low-energy theory around the Dirac point. A pair of states around the Dirac point with near-zero energy are coupled to three pairs of states of energy $\pm h\nu_F \Delta K$. The effective theory from a Schrieffer-Wolff transformation for the near-zero-energy doublet states has corrections to linear order in $K$ which renormalize the velocity [21]. We constructed a series of twisted supercell structures with decreasing angles from $30^\circ$, and the renormalized Fermi velocity calculated from the bands along $\Gamma-K$ indeed follows the theoretical prediction $\nu_F / \nu_F = 1 - C/ \sin^2(\theta/2)$ with $C = 1.953 \times 10^{-4}$ [21].

Another feature for the band structure of the twisted superstructure is the Van Hove singularities (VHS) in the density of states (DOS) near the Fermi level [19], which often leads to electronic instabilities such as superconductivity [37] and magnetism [38] in the many-body system. In Fig. 4(c), the DOS of a $(M,N) = (6,5)$ supercell with $\theta \approx 6.01^\circ$ is

![FIG. 3. Comparison between the tight-binding Hamiltonian (red lines) and ab initio DFT (blue circles) band structure calculations along $\Gamma-M-\Gamma$ in an AB-stacking (a) bilayer, (b) bulk; (c) $(M,N) = (6,5)$ twisted superstructure $(\theta^{(6,5)} \approx 6.01^\circ)$ and comparison to the folded monolayer band structure (green crosses).](image-url)
angles. They are driven by the different local mechanical energy for AA and AB stackings [43]. Though the prediction of mechanical deformations is beyond the scope of the current work, we comment that more general forms of interlayer couplings can be modeled by incorporating variable height or strain by modifying the initial AA/AB sliding bilayers, which will allow proper description of the effects of structural deformations.

V. EFFECTIVE THEORY FROM PROXIMITY EFFECTS

As a final comment on how our model can be applied, we discuss how to construct effective theories with limited degrees of freedom instead of having to solve the Hamiltonian in the full Hilbert space of the combined layers. In graphene bilayers, the low-energy Hamiltonian has the form of non-Abelian gauge theory [44]. To define and formulate the problem, we consider a vdW bilayer heterostructure with layer 1 as the main component where the low-energy degrees of freedom at the Fermi level $E_F$ resides. Layer 2 is brought close to layer 1 to introduce the desired proximity effects. In general, the presence of layer 2 will affect the electronic properties of layer 1 in two ways: (1) by introducing a direct additional potential generated by the neighboring atoms; (2) by introducing virtual interlayer hopping processes through hybridization to the states of the neighboring layer. The general vdW heterostructure Hamiltonian takes the form

$$
\mathcal{H}_{\text{vdw}} = \left[ \hat{H}_1 + \Delta \hat{H}_{12} \right] \hat{T}^\dagger \left[ \hat{H}_2 + \Delta \hat{H}_{21} \right].
$$

$\Delta \hat{H}_{12}$ and $\Delta \hat{H}_{21}$ are the direct corrections of the first type, and $\hat{T}$ is the interlayer coupling in the heterostructure. Integrating out the second layer gives a perturbation term for the first layer,

$$
\Delta V_{12}(E) = \hat{T} \frac{1}{E - (\hat{H}_2 + \Delta \hat{H}_{21})} \hat{T}^\dagger,
$$

where, since $\hat{T}$ is already small, and $\Delta \hat{H}_{21}$ is a higher order correction that disrupts the lattice translation symmetry in $\hat{H}_2$, this term can be ignored. The effective potential takes the following form in the spatial representation, evaluated at $E = E_F$:

$$
\Delta V_{12}(r_2, r_1, E_F) = \sum_{s_2, s_1} t_{r_2, s_2} \frac{1}{E_F - \hat{H}_2} \langle s_1 | t_{s_1, r_1}^\dagger |angle
$$

$$
= \sum_{s_2, s_1} t_{r_2, s_2} \frac{1}{\Omega_k} \int_{\text{BZ}} d^2k \langle \psi_{k, s_1} \rvert E_F - \hat{H}_2(k) \rvert \psi_{k, s_2} \rangle
$$

with $\Omega_k$ the BZ area, $r_j (s_j)$ the localized orbitals of layer 1 (2), which include both the position vector $r_j (s_j)$ and the orbital index $\delta_j$, and $t_{r_j, s_1}$ the interlayer coupling from $j$ to $i$ orbitals between the layers. In the usual perturbation framework, this expression describes hopping across the layers from $r_1$ to $s_1$, allowing for all paths $s_1 \rightarrow s_2$ within layer 2, and hopping back to layer 1, from $s_2$ to $r_2$. When applied to the G-hBN bilayer, the sublattice symmetry breaking mass terms of $\Delta V_{12}$ from hBN to carbon sites have opposite sign from the direct term $\Delta H_{12}$: the carbon site above a BN layer experiences the same sublattice potential as the BN layer itself in the

FIG. 4. (a) The renormalized Fermi velocity, $v_F / v_F$, as a function of the twist angle $\theta$ and VHS in the DOS (inset). (b) The band structure along K-M-K' in an $(M, N) = (6, 5)$ supercell ($\theta \approx 6.01'$) compared to the folded monolayer bands (black dots). (c) The DOS for a $(M, N) = (6, 5)$ supercell (green) compared with the DOS of the monolayer (black). The energy extrema in the bands in (b) correspond to the singular points of DOS. (d) Similar calculations for the $(M, N) = (31, 30)$ supercell ($\theta \approx 1.08'$). (e) The states around the Fermi level are localized at AA sites forming a moiré pattern.

In the discussion so far we have assumed the layers to be flat and focused only on their electronic properties. Structural relaxations such as rippling or more drastic commensurate-incommensurate transitions with domain-line formation, as in the G-hBN bilayer [42], could be relevant for small twist
direct contribution $\Delta \hat{H}_{12}$, while $\Delta V_{12}$ is opposite due to level repulsion in the framework of perturbation theory. The use of this effective potential will enable application to very large systems without loss of accuracy.

VI. CONCLUSION

In summary, we derived the ab initio G-G and G-hBN interlayer couplings based on the maximally localized Wannier function transformation of DFT calculations. We show that these interlayer couplings have both pair-distance and angular-orientation dependence. In contrast, the conventional way of modeling such couplings by fitting band structure calculations leads to ambiguities in the functional form and its dependence on important structural variables [17–20,25,45]. The success of the latter, simpler approach is due to the small number of parameters needed in effective low-energy theories near the Dirac energy [21–24], implying that only one set of dominant Fourier components for interlayer coupling is relevant; this enabled us to generalize the interlayer-coupling model to arbitrary stacking orientations with that involve any possible relative translation or rotation of the layers. Our model can also be generalized to incorporate local variations of in-plane strain and interlayer distance by varying the reference configurations in a systematic way. We expect our model to be relevant in investigating the derivation of low-energy theories appropriate for layer stackings [21–24,44], the band gap introduced by the presence of hBN [46], optical absorption [47], vertical transport across layers [28], phenomena such as the quantum Hall effects, and Hofstadter’s butterfly from the competition between magnetic field and supercell length scales [48–50]. The systematic Wannier approach also allows for further generalization to other two-dimensional layered materials.

ACKNOWLEDGMENTS

We thank Dennis Huang, Daniel Massatt, Stephen Carr, Min-Feng Tu, Sharmila N. Shirodkar, Georgios A. Titisarist, Paul Cazeaux, Eric Cances, Mitchell Luskin, Bertrand Halperin, and Philip Kim for useful discussions. This work was supported by the STC Center for Integrated Quantum Materials, NSF Grant No. DMR-1231319, and ARO MURI Award No. W911NF-14-0247. We used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by NSF Grant No. ACI-1053575.

[36] The \((M,N)\) supercell has twist angle \(\cos(\theta_{(M,N)}) = \frac{N^2 + 4NM + M^2}{2(N^2 + NM + M^2)}\). The supercell of the bottom layer has cell vectors \(a_1^{(M,N)} = Na_1 + Ma_2\) and \(a_2^{(M,N)} = -Ma_1 + (M + N)a_2\) while the top layer has cell vectors \(a_1' = Ma_1' + Na_2'\) and \(a_2' = -Na_1' + (M + N)a_2'\) with \(a_i'\), the counterclockwise rotation of \(a_i\) by \(\theta_{(M,N)}\).