Registry-dependent potential energy and lattice corrugation of twisted bilayer graphene from quantum Monte Carlo

Kittithat Krongchon[1], Tawfiqur Rakib,[2] Shivesh Pathak,[3] Elif Ertekin,[2,4] Harley T. Johnson[2,5] and Lucas K. Wagner[1]

[1]*Department of Physics, Institute for Condensed Matter Theory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[2]*Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[3]*Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*
[4]*Materials Research Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[5]*Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

An uncertainty in studying twisted bilayer graphene (TBG) is the minimum-energy geometry, which strongly affects the electronic structure. The minimum-energy geometry is determined by the potential-energy surface, which is dominated by van der Waals (vdW) interactions. In this paper, large-scale diffusion quantum Monte Carlo (QMC) simulations are performed to evaluate the energy of bilayer graphene at various interlayer distances for four stacking registries. An accurate registry-dependent potential is fit to the QMC data and is used to describe interlayer interactions in the geometry of near-magic-angle TBG. The band structure for the optimized geometry is evaluated using the accurate local-environment tight-binding model. We find that compared to QMC, DFT-based vdW interactions can result in errors in the corrugation magnitude by a factor of 2 or more near the magic angle. The error in corrugation then propagates to the flat bands in twisted bilayer graphene, where the error in corrugation can affect the bandwidth by about 30% and can change the nature and degeneracy of the flat bands.

I. INTRODUCTION

Twisted bilayer graphene (TBG) exhibits a multitude of correlated electronic phases and has emerged as a platform for studying correlated electron physics [1–9]. These correlation-driven phases are attributed to flattening of the band structure near the Fermi level due to the superlattice interaction in moiré patterns [10–15]. However, a complete model Hamiltonian for these systems remains unknown. An important piece of the puzzle of predicting flat bands is the van der Waals (vdW) interactions, which lead to symmetry breaking, corrugation, and other distortions in the layers. The lattice reconstruction in the moiré superlattices in turn significantly affects the electronic behavior of this system [16–22]. Therefore, the interplay between lattice corrugation and electronic structure provides strong impetus to accurately model vdW contributions in TBG.

Evaluating vdW interactions, however, is a challenging task for two reasons. First, this type of interaction results from long-range electron correlations, which means that local or semilocal exchange-correlation functionals from density functional theory (DFT) cannot describe them [23–25]. Although a large set of computational methods has been developed in the DFT regime to account for the long-range interactions [26–28], they are empirical in nature. Within various iterations of the so-called DFT-D scheme, which adds the dispersion corrections to the standard Kohn–Sham DFT energy, the estimated binding energy errors can differ by up to 250% [29] in bilayer graphene. Thus, the large uncertainty in this set of techniques warrants a more accurate first-principles approach. The second reason for the difficulty is that accurate treatment

of vdW interactions becomes computationally intractable for a large system, such as small-angle TBG, which consists of $\sim 10^4$ atoms. A systematic potential training approach is needed to model lattice and electronic degrees of freedom for vdW systems, while maintaining first-principles accuracy.

Diffusion quantum Monte Carlo (QMC) has been shown to closely reproduce experimental values for vdW materials due to explicit treatment of electron interactions [30–33]. Specifically, in AB-stacked bilayer graphene, the binding-energy curve from QMC is able to predict the out-of-plane phonon frequency and relaxed interlayer spacing in agreement with available experimental results [29], which shows that QMC is a promising technique for investigating bilayer graphene. The interlayer-energy curve from QMC data for this system is available for only a single registry: AB-stacked bilayer graphene. To fit a registry-dependent potential, multiple registries are needed. Therefore, more QMC data is needed to accurately parametrize the potential-energy surface of the entire moiré superlattice.

In this paper, we use large-scale QMC simulations to compute a reference quality Born–Oppenheimer ground-state energy for bilayer graphene as a function of the displacement for four stacking registries, which allows for an accurate assessment of the stacking-fault energy. To make this high-quality potential usable for large-scale molecular dynamics calculation, we fit the QMC results to an interatomic potential. We find that the QMC data is closer to the random-phase approximation (RPA) [34] and a previous atomic potential parametrization [35] than to DFT with dispersion corrections. The data is well fit by the Kolmogorov–Crespi (KC) potential model [36], whose refined parameters for the KC potential
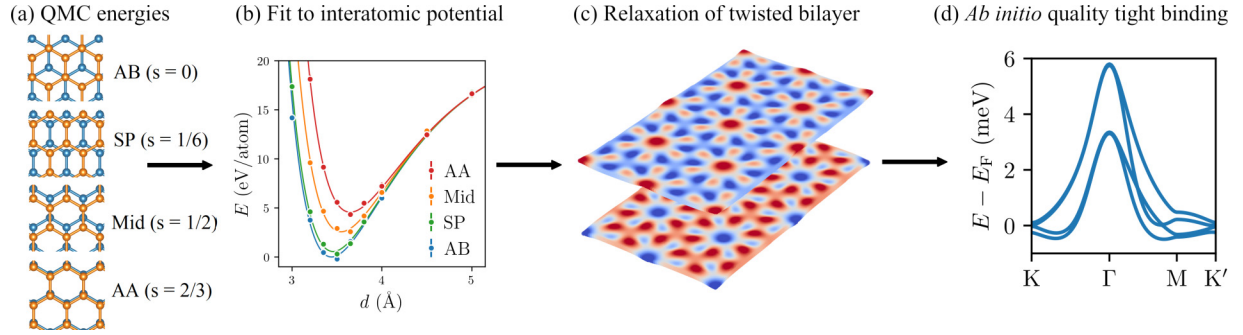
FIG. 1. Workflow for training the KC potential and calculating the band structure. (a) First, the QMC energy for bilayer graphene is calculated at various interlayer distances for four stacking registries. Stacking registries are labeled according to the rigid translation distance s of the top graphene layer along the armchair direction in units of $\sqrt{3}a$, where $a = 2.46$ Å is the in-plane lattice constant. (b) In the second step, the KC potential is fit to the QMC data. The error bars represent the single σ statistical uncertainties associated with the QMC results, which are explained in detail in the Appendix. (c) Next, using the fitted KC interlayer interaction, the TBG structure is relaxed. (d) Lastly, the band structure is computed on the minimum-energy structure.

are made available in the Supplementary Information in LAMMPS [37,38] format. We then assess the importance of an accurate vdW energy calculation by investigating the effect of corrugation computed by QMC, DFT-D2 [26,27], DFT-D3 [28], and Ref. [35] (labeled as "KC-Ouyang"). For each relaxed structure from different parametrizations of the KC potential, the band structure is evaluated using the accurate local-environment tight-binding (LETB) model [20]. The DFT-D2 and DFT-D3 vdW interactions result in large errors in the band structure due to the poor estimation of the structural corrugation. We show that the difference in corrugation due to QMC evaluation of the vdW interaction leads to significant changes in the electronic band structure.

II. METHOD

The procedure for parametrizing the interlayer interaction of TBG and evaluating the sensitivity of the minimum-energy structure and electronic structure is outlined as follows.

(1) First, the energy of rigid bilayer graphene as a function of the interlayer distance, is sampled using QMC for four stacking registries, defined in Fig. 1.

(2) The sampled QMC energy data are used to fit the KC potential.

(3) Using the fitted KC potential, the structure of TBG is optimized by minimizing the total energy.

(4) The electronic structure is calculated on the relaxed structure of TBG using the LETB model [20].

The steps are illustrated in Fig. 1 and are discussed in further details in the following sections.

A. QMC calculations

We consider four stacking registries of bilayer graphene, which are constructed by performing rigid translations of the top layer of AB-stacked bilayer graphene along the armchair direction by different distances. We label these registries according to the translation distances s as listed in Fig. 1. The sliding parameter s is defined such that starting from the AB structure ($s = 0$), the translation of $s = 1$ brings the structure back to AB. The nomenclature of these stacking orders is

described in Ref. [34], except for the "Mid" stacking type, which is additionally defined here to be $s = 1/2$. For each of the four registries, we perform QMC to sample the energy at 11 interlayer distances, ranging from $d = 3$ Å to 7 Å, for a total of 44 energy data points.

In the fixed-node QMC scheme as implemented in the QMCPACK software package [39], the ground-state wave function is projected out of the Slater–Jastrow trial wave function of the form

$$\Psi(\mathbf{R}) = \text{Det}[\phi_i^\uparrow(\mathbf{r}_j^\uparrow)]\text{Det}[\phi_i^\downarrow(\mathbf{r}_j^\downarrow)]\exp(J), \quad (1)$$

where $\mathbf{R} = \{\mathbf{r}_1, \ldots, \mathbf{r}_M\}$ is the collection of electron coordinates of the M-electron system, ϕ is the Kohn–Sham orbital, i and j denote electron indices, $\uparrow$ and $\downarrow$ indicate spins, and J is the Jastrow correlation factor as defined in Ref. [40].

The set of Kohn–Sham orbitals is produced by the Quantum ESPRESSO plane-wave DFT code [41–43], using 200 Ry kinetic energy cutoff and the **k**-point mesh of $20 \times 20 \times 1$. The core electrons are removed using Dirac–Fock pseudopotentials described by Refs. [44,45]. We verify that the orbitals generated by the Perdew–Burke–Ernzerhof approximation [46] and two vdW functionals, namely DFT-D2 [26] and DFT-D3 [28], result in the same QMC energy within error bars [47]. The QMC energy is twist averaged [48] over 4×4 **k**-point mesh. The finite-size errors are eliminated by extrapolating the twist-averaged QMC energies of 3×3, 4×4, 5×5, and 6×6 bilayer graphene supercells to the thermodynamic limit. The full analysis of the QMC errors is provided in the Appendix.

B. Potential parametrization

The stacking-dependent KC potential [36] is fit using least squares to the 44 calculated QMC data points to obtain a smooth curve that can describe the energy at any registry and interlayer distance. Since the KC potential does not provide a good fit across the entire range of interlayer distances, we weight the low-energy configurations higher than the large-distance configurations. In this paper, we assume that this is enough to estimate the interlayer interaction in the twisted bilayer case; essentially we are assuming that the interaction

energy only depends on the local stacking. We thus expect the potential to be most accurate for small twist angles, where the local alignment of the layers varies slowly. The detailed procedure to select the most suitable model for a low-energy structure is described in the Appendix.

The KC potential model fit to QMC data is labeled as KC-QMC. We perform a similar procedure for DFT-D2 [26] and DFT-D3 [28] but with many more samples of interlayer distances (0.01 Å apart) as these vdW correction schemes are not computationally expensive. The KC potentials fitted to DFT-D2 and DFT-D3 are labeled as KC-DFT-D2 and KC-DFT-D3 respectively in order to make distinctions between the computed training data and the fitted curves.

C. Stacking-fault energy calculation

The stacking-fault energy (SFE) is defined as the difference between the energy of a displaced configuration ($s \neq 0$) and the energy of the most stable configuration ($s = 0$), which is the AB-stacked bilayer graphene in this case. Given a KC potential, we find the minimum point of the energy curve for each registry to obtain the relaxed interlayer spacing and its corresponding minimum energy. We plot the SFE as a function of registry s along the armchair direction for our results from KC-QMC, KC-DFT-D2, and KC-DFT-D3 in Fig. 2(a). The KC parameters for KC-QMC (Table II), KC-DFT-D2, and KC-DFT-D3 are obtained by the fitting procedure as described in the Appendix, while the KC parameters for KC-Ouyang are taken from Ref. [35]. The data points for KC-QMC, KC-Ouyang, KC-DFT-D2, and KC-DFT-D3 are obtained from a quadratic fit within the range of 0.05 Å from the minimum of the potential-energy surface for each stacking registry. The data points for RPA are taken directly from Ref. [34].

We fit the SFE as a function of registry using the formula [34]

$$
\begin{aligned}
F(s) = \; & c_0 + c_1[2\cos(2\pi s) + \cos(4\pi s)] \\
& + c_2[2\cos(6\pi s) + 1] \\
& + c_3[2\cos(4\pi s) + \cos(8\pi s)] \\
& + \sqrt{3}c_1[-2\sin(2\pi s) + \sin(4\pi s)] \\
& - \sqrt{3}c_3[-2\sin(4\pi s) + \sin(8\pi s)],
\end{aligned}
\tag{2}
$$

where $F(s)$ is the SFE, s is the registry in the armchair direction in units of $\sqrt{3}a$, and a is the in-plane lattice constant. We also fit the interlayer distance $d_{\min}$ for a given registry using the same functional form as Eq. (2). The fitting constants c_0, c_1, c_2, c_3 are reported in Table VI and VII.

D. Structural relaxation

The relaxed geometries of TBG are obtained using the conjugate gradient method with a stopping tolerance of 10^{-11} eV as implemented in the LAMMPS molecular dynamics program. The intralayer interaction is given by the reactive empirical bond order potential [49], while the interlayer interaction is described by one of the four KC potentials that we have parametrized, namely KC-QMC, KC-Ouyang, KC-DFT-D2, and KC-DFT-D3. The initial structures are de-
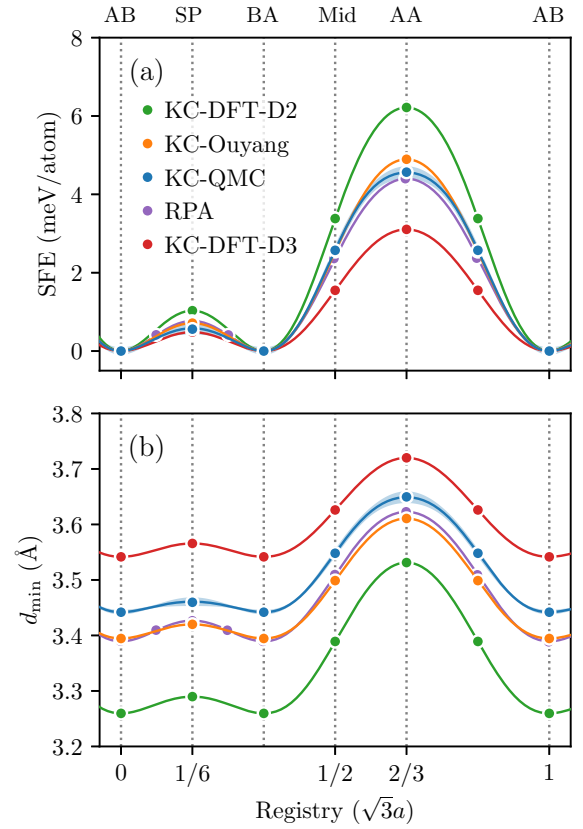

FIG. 2. (a) Stacking-fault energy given by different potential-energy surfaces. Data points of KC-QMC, KC-Ouyang, KC-DFT-D2, and KC-DFT-D3 are obtained from quadratic fits in the range 0.05 Å around the potential-energy surface minimum. The RPA data points are obtained from Ref. [34]. Lines are fits of Eq. (2) to the data points. QMC errors from the bootstrapping technique are represented by the light blue regions around the QMC fit. (b) Minimum interlayer spacings. Data points are obtained from the same procedure as the stacking-fault energy.

fined by two rigid sheets of graphene at 3.4 Å interlayer distance. The top layer is rotated at a twist angle θ with respect to the bottom layer. For this paper, we perform relaxation calculations for the commensurate twist angles of $\theta = 0.84°$, $0.93°$, $0.99°$, $1.05°$, $1.08°$, and $1.16°$. Despite the large number of atoms in a simulation cell, which is on the order of 10^4, the geometry optimizations remain computationally tractable due to the low cost of the classical potentials.

E. Local-environment tight-binding model

To determine the band structure of the relaxed geometry, the local-environment tight-binding model is employed due to its high accuracy within twisted bilayer graphene [20]. This model accounts for the detailed local environment of the nearby atoms and has the following form:

$$
H_{\text{LETB}} = \sum_{ij\sigma} t_{ij}^{\text{LETB}}(\mathbf{R}_i, \mathbf{R}_j, \{\mathbf{R}_{ij}\}) c_{i\sigma}^{\dagger} c_{j\sigma} + \text{H.c.},
\tag{3}
$$

where σ denotes the spin index, $\mathbf{R}_i$ and $\mathbf{R}_j$ represent the location of atoms i and j, $\{\mathbf{R}_{ij}\}$ is a set of atomic positions in the vicinity of atoms i and j. The functional form of the
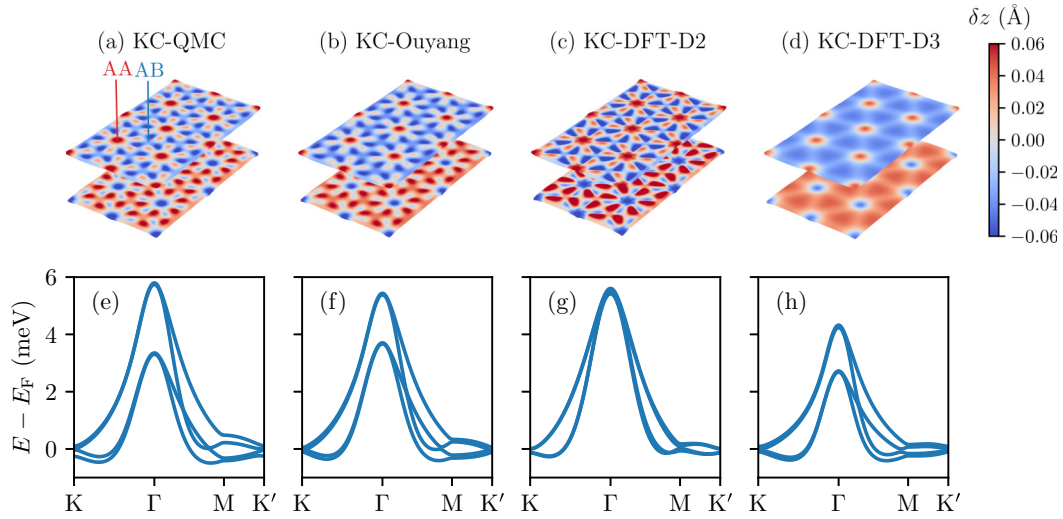
FIG. 3. Corrugation of twisted bilayer graphene at 0.99° from various sources and their respective flat bands. [(a)–(d)] Corrugation in angle twisted bilayer graphene geometry, relaxed using KC potentials with parameter sets from difference sources: (a) QMC, (b) Ref. [35], (c) DFT-D2, and (d) DFT-D3. Corrugation δz is defined as the deviation of the z coordinate of a carbon atom from the average of the smallest and largest z coordinates within the layer. [(e)–(h)] Flat bands near the Fermi level for four twisted bilayer structures obtained from different potentials as described in (a)–(d).

hopping parameters t_{ij}^{LETB} is classified into intralayer and interlayer contributions based on the z projection of the distance $\mathbf{R}_i - \mathbf{R}_j$. The detailed description of the functional form of t_{ij}^{LETB} is provided in Ref. [20].

III. RESULTS

The results of our calculations are discussed using QMC as the reference. This decision is based on the ability to reproduce the experimental values of relaxed interlayer spacing in the AB stacking and out-of-plane zone-center optical phonon frequency in bilayer graphene [29].

A. Stacking-fault energy

Figure 2(a) shows the SFE as a function of registry. We find a close agreement between KC-QMC, KC-Ouyang, and RPA. Let us consider the AA stacking, shown by the vertical dotted line at the registry of $s = 2/3$, where the differences are the most notable. The difference between KC-QMC and RPA is 0.2 meV in the stacking-fault energy. We consider the agreement between KC-QMC and RPA to be an indication that both might be accurate in this case. A similar agreement between QMC and RPA was found for water on boron nitride [50]. The difference between KC-QMC and KC-Ouyang is approximately 0.3 meV in the stacking-fault energy. While the agreement might appear to be fortuitous, a similar agreement between QMC and the so-called DFT-MBD [51], on which the KC-Ouyang potential was trained, was found in DNA-ellipticine molecules [31].

Relative to the consensus SFE of KC-QMC, RPA, and KC-Ouyang, KC-DFT-D2 overestimates SFE by 1.7 meV for the AA stacking compared to KC-QMC, while KC-DFT-D3 underestimates the SFE by 1.5 meV. This result agrees with previous comparisons to RPA only [34], so our results increase confidence in the RPA results.

Figure 2(b) shows the relaxed interlayer distance d_{min} as a function of registry. Similar to the case of SFE in Fig. 2(a), there is an agreement between KC-QMC, KC-Ouyang, and RPA. In this case, however, both RPA and KC-Ouyang underestimate the relaxed interlayer distance for the AA stacking by a small amount of 0.03 and 0.04 Å, respectively. Meanwhile, the two dispersion-corrected DFT methods show the opposite trend of SFE, as shown by the fact that KC-DFT-D2 has the smallest relaxed interlayer distances overall. KC-DFT-D2 underestimates the relaxed interlayer distance by 0.12 Å, while KC-DFT-D3 overestimates it by 0.07 Å. This result is consistent with the result of SFE shown in Fig. 2(a) in the sense that a smaller interaction leads to larger interlayer distance, which in turn results in smaller SFE.

B. Effects of the more accurate interlayer potential on the minimum-energy structure of TBG

The corrugation δz, defined as the deviation of the z coordinate of a carbon atom from the average of the smallest and largest z coordinates within the layer, of both layers for each interlayer potential is visualized in Figs. 3(a)–3(d). In the AA regions, the relaxed structures manifest an upward bulge in the top layer and a downward bulge in the bottom layer, which are denoted by the most prominent red and blue regions in the visualization. Around these AA peaks, the sixfold symmetric structure is observed. The formation of small-amplitude structure around the AA regions results in the threefold symmetry around the AB/BA regions, which can be identified by the centroid of three adjacent AA nodes as labeled in Fig. 3(a). This result is in qualitative agreement with previous studies of breathing mode structure [17,21,52]. The heights of these peaks depend on the potential being used to describe the interlayer interactions. In the case of KC-QMC [Fig. 3(a)], the maximum out-of-plane corrugation is $\delta z_{\text{KC-QMC}} = 0.075$ Å. The maximum corrugation of KC-Ouyang is $\delta z_{\text{KC-Ouyang}} =$

TABLE I. Flat bandwidths in meV of the structures from different potentials at various commensurate twist angles.

Potential	0.84°	0.93°	0.99°	1.05°	1.08°	1.16°
KC-QMC	45.680	18.627	6.294	16.098	26.350	49.134
KC-Ouyang	44.998	18.075	5.848	16.901	27.108	49.933
KC-DFT-D2	44.218	16.938	5.880	18.974	29.404	52.534
KC-DFT-D3	42.936	17.744	4.740	16.352	26.175	48.243

0.071 Å, or 5% smaller than $\delta z_{\text{KC-QMC}}$. This similarity is expected from the agreement in SFE (Fig. 2). For KC-DFT-D2, the chiral relaxations are also observed between layers, where the top layer has the opposite chirality to the bottom layer, in agreement with KC-QMC and KC-Ouyang. However, the difference in maximum out-of-plane corrugation from KC-QMC is more pronounced as $\delta z_{\text{KC-DFT-D2}}$ is 0.112 Å, which is 50% larger than $\delta z_{\text{KC-QMC}}$. A small difference from KC-QMC is also observed in around AA nodes, in which an inner sixfold symmetric structure is observed, while in the case of KC-QMC and KC-Ouyang, the AA regions have smooth round-shaped bulges. On the other hand, the structure from KC-DFT-D3 has much smaller overall corrugation, with the maximum of only $\delta z_{\text{KC-DFT-D3}} = 0.045$ Å. As a result, the chiral symmetry between both layers and sixfold symmetric structure around AA nodes are not observed in KC-DFT-D3.

C. Effects of the more accurate minimum-energy structure on the band structure of TBG

To investigate the effects of the more accurate minimum-energy structure on the band structure, the LETB model is employed [20]. In Figs. 3(e)–3(h), the LETB band structures are plotted for four twisted bilayer structures at 0.99°, relaxed using KC parameters from four different sources, namely KC-QMC, KC-Ouyang, KC-DFT-D2, and KC-DFT-D3. The band structures from KC-QMC, KC-Ouyang, and KC-DFT-D3 show similar features, in which the four bands exhibit two degenerate energy states at the gamma point, as reported in previous studies [20,53,54]. The band structure from a KC-DFT-D2 geometry shows a slightly different feature from the other three structures, where in this case, all the four bands seem to form a single degenerate energy state at the gamma point.

The flat bandwidths, defined as the difference between the maximum energy and the minimum energy of the flat bands, for different potentials and twist angles are reported in Table I. At the twist angle of 0.99°, the bandwidth for KC-QMC is 7% larger than KC-Ouyang and KC-DFT-D2, and 33% larger than KC-DFT-D3. In Fig. 4(a), the flat bandwidth is plotted as a function of the twist angle. Reducing the twist angle from 1.16° to 0.99°, all four potentials show a similar downward trend, where the inflection points occurs, which defines the first magic twist angle. The identification of the magic angle at 0.99° agrees with the previous result [20].

The electron and hole gaps are defined by the separation of the flat bands from the remote bands, and have been noted [6,55] to be particularly sensitive to the corrugation. For our
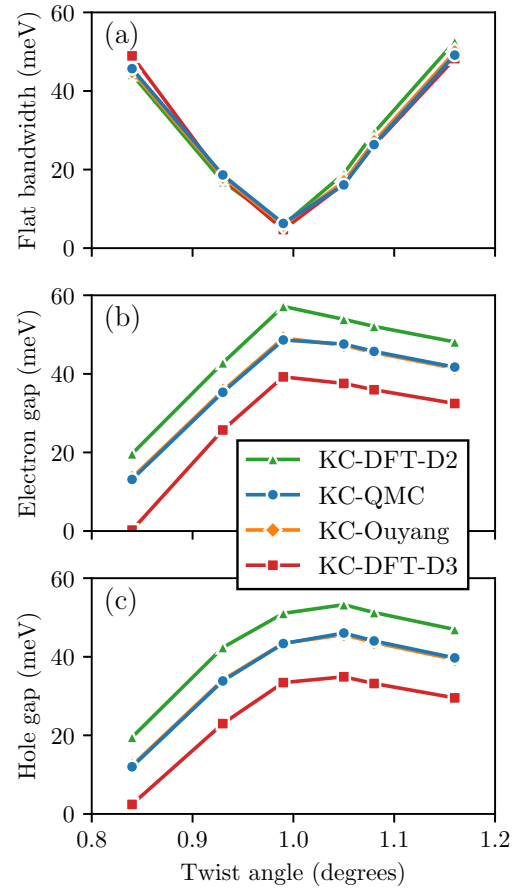

FIG. 4. (a) Flat bandwidth, (b) electron band gap, and (c) hole band gap as a function of twist angle. The four sets of band structure are computed using the same LETB model [20] but with geometries relaxed under different interlayer KC potentials: KC-QMC, KC-Ouyang, KC-DFT-D2, and KC-DFT-D3.

geometries, these gaps are presented in Figs. 4(b) and 4(c). The treatment of the vdW interaction can change the estimated hole and electron gaps by a significant amount. For example, near the magic angle of 0.99°, the gaps can vary by almost a factor of two. Corrugation tends to increase the electron and hole gaps, so the underestimation of corrugation in DFT-D3 results in gaps that are too small, while the reverse in DFT-D2 results in gaps that are too large.

IV. CONCLUSIONS

As has been noted previously in the literature, the careful treatment of the van der Waals interaction appears to be very important to obtain accurate corrugation in twisted bilayer graphene [19,34], and small changes in the corrugation can affect the electronic structure significantly [16,21]. In this paper, we provided a state-of-the-art benchmark of the van der Waals interaction in bilayer graphene and found that while a commonly used atomic potential [35] is fairly accurate, DFT-based methods may not improve the description, and in fact may lead to worse results. Models of the electronic structure, which depend on van Hove singularities [56] and other features of the flat bands should take this into consideration.

For the accuracy of the interlayer interaction in graphene, we found that the commonly used Kolmogorov–Crespi [36] model as parametrized in Ref. [35] was surprisingly accurate compared to QMC results, as well as the RPA estimation. On the other hand, commonly used corrections to DFT, the DFT-D2 and DFT-D3 functionals, resulted in large errors in the interlayer potential and stacking-fault energy, which leads to an over (DFT-D2) or under (DFT-D3) estimation of the degree of corrugation.

We found that the electronic band structure varied significantly depending on the treatment of the corrugation. While structures relaxed from different potentials result in the same prediction of the magic twist angle at $0.99°$, the error in the electronic structure due to using DFT geometries near the magic angle (here found to be roughly $0.99°$) results in rearrangements of bands up to around 50% (≈ 3 meV) in the two lower flat bands for the DFT-D2 structure. While the separation of the flat bands from the other bands as quantified by electron and hole gaps shows the same inflection point at $0.99°$, at the magic twist angle, the relaxed structures from DFT-D2 and DFT-D3 result in a similar error of roughly 10 meV, which is an error of about 50%.

This study has focused on freestanding bilayer graphene and has employed fixed-node QMC as a reference method due to its established reliability. We believe this technique is the most accurate available that can treat the required system size. Comparison between QMC and less accurate approximation techniques serves as a benchmark for testing and improving numerical schemes, and it would be interesting to see whether other high-accuracy first-principles techniques obtain similar results. We only considered freestanding bilayer graphene in this study, while most experiments are performed on a substrate and encapsulated with materials such as boron nitride. It would be a fruitful future direction to use this model to explore the effects of encapsulation on the structure of bilayer graphene.

Alongside this paper, we provide the interlayer energy of bilayer graphene from QMC and a potential fit to the QMC data suitable for use in LAMMPS, which is available online [57]. This data is suitable for use to perform atomic scale simulations of bilayer graphene, and has the level of accuracy comparable to the underlying quantum Monte Carlo calculations in the binding region.
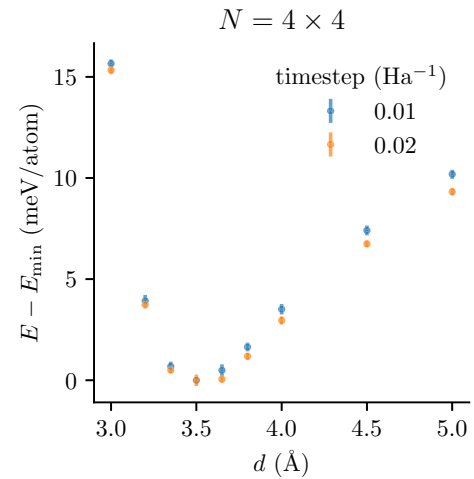
FIG. 5. QMC energy vs interlayer distance for AB-stacked bilayer graphene with the supercell size of $N = 4 \times 4$, where N is the number of primitive cells in a simulation cell. Energies are referenced to the minimum-energy point at 3.5 Å. The energies for the time step of 0.02 Ha^{-1} are roughly 0.01 meV away from the energies for the time step of 0.01 Ha^{-1}.

APPENDIX

1. Errors in diffusion quantum Monte Carlo data

Figure 5 shows that the results for the time step of 0.02 Ha^{-1} have roughly 0.01 to 0.02 meV errors from the energies for the time step of 0.01 Ha^{-1}. Since the discrepancy is small at small interlayer distances, we consider the time step of 0.02 Ha^{-1} a good approximation for describing the potential energy near equilibrium distances.

The diffusion quantum Monte Carlo (QMC) calculations are performed using a constant 4×4 twist grid on each of the
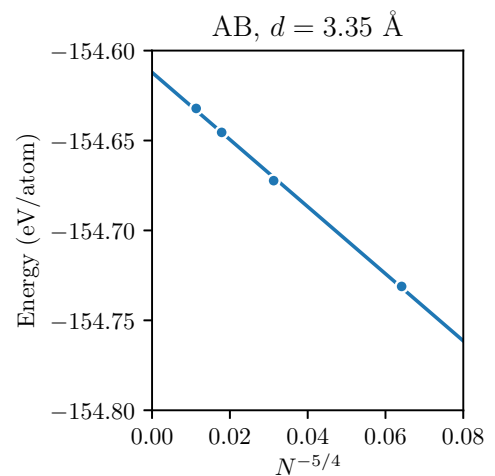
FIG. 6. QMC energy vs $N^{-5/4}$ for AB-stacked bilayer graphene at 3.35 Å interlayer distance, where N is the number of primitive cells in a simulation cell. The statistical error bars from QMC range from 0.2 to 0.8 meV/atom, which are much smaller than the y-axis scale and are therefore not displayed in the figure. The line is fitted through the energy data points using Eq. (A1) for the purpose of extrapolating the energy to the thermodynamic limit.

TABLE II. List of KC parameter values for the carbon-carbon interaction. The energy within the parentheses after KC-QMC indicate the value of $k_B T$, which is used to tune the fitting weights described by Eq. (A5). The model KC-QMC (4 meV) are used to describe interlayer interactions to obtain relaxed structures shown in Figs. 3(a)–3(d). The value of R_{cut} is fixed to 16 Å.

Potential	z_0 (Å)	C_0 (meV)	C_2 (meV)	C_4 (meV)	C (meV)	δ (Å)	λ (Å^{-1})	A (meV)
KC-QMC (2 meV)	3.44323	14.33525	12.56000	0.41916	7.40186	0.65357	3.25395	13.91535
KC-QMC (3 meV)	3.40756	16.29563	13.09544	0.01848	6.65729	0.68934	3.28558	13.74788
KC-QMC (4 meV)	3.37942	18.18467	13.39421	0.00356	6.07494	0.71935	3.29308	13.90678
KC-QMC (5 meV)	3.37042	19.26299	12.89983	0.00524	5.07182	0.75587	3.29479	13.84456
KC-QMC (6 meV)	3.35062	21.22680	13.17347	0.20199	4.18306	0.78726	3.29340	14.17078
KC-QMC (7 meV)	3.35104	22.97568	12.37219	0.31184	1.46736	0.86239	3.28298	14.03609
KC-QMC (8 meV)	3.33406	25.26089	14.07200	2.20093	0.01279	0.84472	3.27746	14.38600
KC-QMC (9 meV)	3.34121	24.58252	13.29192	1.28582	0.00060	0.86790	3.27689	14.15974
KC-QMC (10 meV)	3.30334	27.74635	15.07551	1.53471	0.00001	0.86547	3.27683	15.12378
KC-QMC (11 meV)	3.31271	26.85873	15.25734	3.13552	0.00087	0.82689	3.27726	14.83406
KC-QMC (12 meV)	3.31412	26.68834	15.26599	3.70204	0.00005	0.81525	3.27729	14.77050
KC-QMC (13 meV)	3.33134	25.18924	14.44755	4.21503	0.00008	0.80145	3.27754	14.29565
KC-QMC (14 meV)	3.33377	24.96005	14.28375	4.75430	0.00000	0.79129	3.27772	14.21445
KC-QMC (15 meV)	3.36987	22.14523	12.65076	4.42759	0.00341	0.78733	3.27777	13.31070
KC-QMC (16 meV)	3.37850	21.51294	12.19151	4.82753	0.00005	0.77777	3.27803	13.09407
KC-QMC (17 meV)	3.40957	19.41387	10.93804	4.61657	0.00002	0.77283	3.27814	12.38366
KC-QMC (18 meV)	3.37184	21.95300	12.29456	5.48004	0.00011	0.76864	3.27827	13.22834
KC-QMC (19 meV)	3.38030	21.33727	11.89459	5.50172	0.00005	0.76582	3.27834	13.02214
KC-QMC (20 meV)	3.38045	21.31321	11.82651	5.65686	0.00002	0.76330	3.27840	13.01085
KC-QMC (unweighted)	3.37006	21.78334	10.46939	8.86496	0.00001	0.72395	3.28315	13.09016
KC-Ouyang [35]	3.41608	20.02158	10.90551	4.27564	0.01001	0.84471	2.93606	14.31326

supercells $N = 3 \times 3$, 4×4, 5×5, and 6×6, where N is the number of unit cells in a simulation cell. The grid of twists is offset from the Γ point by 0.076923 in reciprocal lattice units in order to avoid the Dirac point, which would otherwise result in slower convergence. The largest cell size results in an equivalent 24×24 **k** mesh, which we confirmed is converged in DFT. The energy is extrapolated to the infinite system size by fitting the following linear equation [58]:

$$E(N) = E(\infty) + cN^{-5/4}, \tag{A1}$$

where $E(\infty)$ is a fitting parameter, which represents the energy extrapolated to the thermodynamic limit, and N is the number of primitive cells in a simulation cell. The error bars of the extrapolated energies are calculated using the bootstrapping technique. The linear model fits our energy data well, as shown in Fig. 6 for AB-stacked bilayer graphene with 3.35 Å interlayer distance. The time step is chosen to be 0.02 Ha^{-1}.

2. Fitting procedure

We sampled 44 energy data points using QMC and fit them to the KC potential [36]. The KC model is a registry-dependent potential designed to improve on the classical potential by taking into account the anisotropy of the π
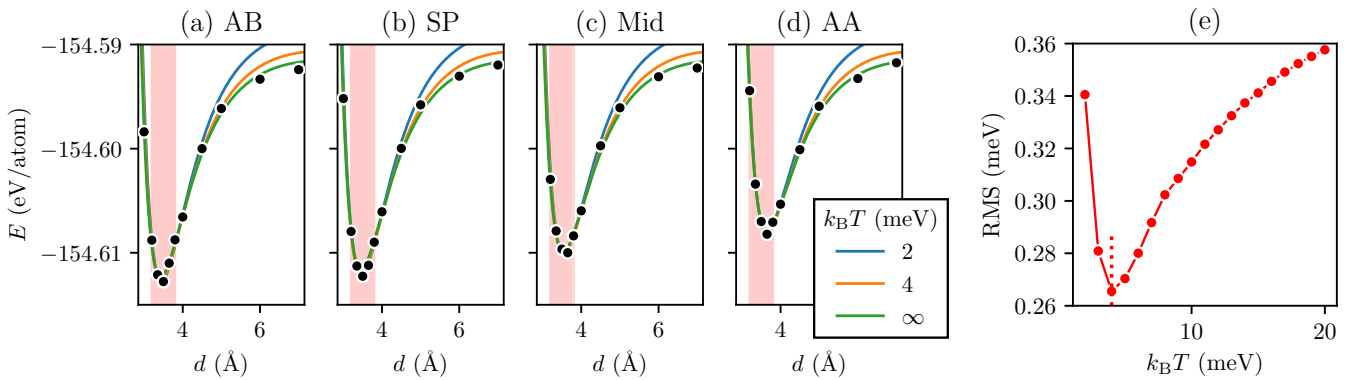

FIG. 7. Potential-energy surface of bilayer graphene at registry (a) AB, (b) SP, (c) Mid, and (d) AA. Data points (black) are calculated using QMC. Lines are fits using the KC model from Eq. (A2) with the fitting weights controlled by the parameter $k_B T$ as given by Eq. (A5). Red highlighted regions cover the range of interlayer distances near the equilibrium spacing, chosen to be $d \in [3.2, 3.8]$ Å. (e) RMS as a function of $k_B T$ for the data points within the range of interlayer distances $d \in [3.2, 3.8]$ Å. For the region near the equilibrium spacing, $k_B T = 4$ meV minimizes RMS (shown as the dotted vertical line). The KC model associated with this value of $k_B T$ is therefore chosen to represent the QMC data (labeled as KC-QMC) and is used in the subsequent relaxation calculations of the twisted bilayer graphene.

TABLE III. Binding energy in meV/atom at different registries. The chosen KC-QMC model corresponds to KC fitting with no weights. The error bars from KC-QMC are obtained from the bootstrap method.

Method	AB	SP	Mid	AA
KC-QMC	21.96(9)	21.17(8)	19.52(9)	17.41(9)
KC-Ouyang	24.82	24.11	22.28	19.92
KC-DFT-D2	23.99	22.93	20.65	17.82
KC-DFT-D3	19.52	19.02	17.95	16.36
RPA [59]	25			
QMC [29]	17.7(9)			11.5(9)

TABLE IV. The minimum interlayer spacing d_{min} in Å. The chosen KC-QMC model corresponds to $k_B T = 4$ meV. The error bars from KC-QMC are obtained from the bootstrap method.

Method	AB	SP	Mid	AA
KC-QMC	3.442(4)	3.460(4)	3.548(5)	3.650(6)
KC-Ouyang	3.395	3.420	3.499	3.611
KC-DFT-D2	3.260	3.290	3.389	3.531
KC-DFT-D3	3.542	3.566	3.626	3.720
RPA [59]	3.39			
Exp. [60]	3.48(10)			

overlap between layers. The full model is given by

$$E^{KC} = \frac{1}{2} \sum_i \sum_{j \neq i} \text{Tap}(r_{ij}) V_{ij},$$

$$V_{ij} = e^{-\lambda(r_{ij}-z_0)}[C + f(\rho_{ij}) + f(\rho_{ji})] - A\left(\frac{r_{ij}}{z_0}\right)^{-6},$$

$$\rho_{ij}^2 = r_{ij}^2 - (\mathbf{r}_{ij} \cdot \mathbf{n}_i)^2, \qquad \text{(A2)}$$

$$\rho_{ji}^2 = r_{ij}^2 - (\mathbf{r}_{ij} \cdot \mathbf{n}_j)^2,$$

$$f(\rho) = e^{-(\rho/\delta)^2} \sum_{n=0}^{2} C_{2n}\left(\frac{\rho}{\delta}\right)^{2n},$$

where $\mathbf{r}_{ij}$ is the distance vector pointing from atom i to atom j from different layers, $\mathbf{n}_k$ is the surface normal at atom k, and ρ_{ij} is the transverse distance from atom i to atom j. The taper function given by

$$\text{Tap}(x_{ij}) = 20x_{ij}^7 - 70x_{ij}^6 + 84x_{ij}^5 - 35x_{ij}^4 + 1, \qquad \text{(A3)}$$

$$x_{ij} = \frac{r_{ij}}{R_{cut}}, \qquad \text{(A4)}$$

provides a continuous long-range cutoff, and R_{cut} is fixed to 16 Å throughout all the calculations.

The weighted fit is determined by optimizing the function

$$\chi^2 = \sum_l^{N_d} w_l \left(E_l - E_l^{KC}\right)^2,$$

where E_l is the computed energy for the lth configuration of bilayer graphene, E_l^{KC} is the calculated energy from Eq. (A2), and N_d is the number of data points, which is 44 for the QMC method and 268 for DFT-D2 and DFT-D3. The fitting weight w_l depends on the energy and a tunable parameter $k_B T$ according to the formula

$$w_l = \exp\left(-\frac{E_l - E_{min}}{k_B T}\right). \qquad \text{(A5)}$$

Here, E_{min} is the smallest data point, which in our case is the equilibrium point of the AB stacking potential-energy surface, k_B is the Boltzmann constant, and T is the temperature, which serves as a parameter that controls the relative weight of the low- and high-energy data points.

The weights are introduced so that we can tune how much the fitting model favors the data points near equilibrium by adjusting the value of $k_B T$. A smaller value of $k_B T$ corresponds

to a model that highly favors the points near equilibrium, while giving up the goodness of fit at very small or large interlayer distances. Thus, the model is selected according to the goodness of fit in the region of interest, which can be quantified by the root-mean-square (RMS) error as described in the next section. The values of fitting parameters for different values of $k_B T$ are reported in Table II along with the KC parameters from Ref. [35] labeled as KC-Ouyang. In Figs. 7(a)–7(d), the computed QMC energies for four stacking types are displayed as black data points along with the fitted curves for three different sets of weights, i.e., $k_B T = 2$, 4, and ∞ meV (unweighted).

We find that while the KC potential fits well to the energy data from the DFT-D scheme because they both have the functional form of r^{-6}, the QMC potential energy is not exactly proportional to r^{-6}, which is why the KC potential struggles to describe the entire potential-energy surface as shown in Fig. 7. Modifications to the model might be needed to describe the potential-energy surface at any interlayer separation range.

3. Choosing the KC parameter set

In order to decide which KC parameter set to use, we investigate the RMS errors between the QMC data and the fitting within the region of interest. The RMS errors for the data within $d = 3.2$ to 3.8 Å as illustrated by the highlighted regions in Figs. 7(a)–7(d) are plotted as a function of $k_B T$ in Fig. 7(e). This range of interlayer distances is chosen based upon the fact that the relaxed interlayer distances of all the methods we explore fall within this range as shown in Fig. 2(b). Figure 7(e) suggests that the most preferable model is corresponding to $k_B T = 4$ meV due to having the smallest RMS. Therefore, all the subsequent calculations of relaxed structures in this work are performed using this set of KC parameters to describe the interlayer interactions. In the

TABLE V. Comparison of key factors that could lead to discrepancy between the QMC results of this study and Ref. [29].

Factor	Present paper	Reference [29]
Data points	44	5
Fitting potential	KC [36]	"16-12-8-4" [29]
d_{min} for AA	3.650(6) Å	3.495 Å [61]
d_{min} for AB	3.442(4) Å	3.384 Å [61]
Method to obtain d_{min}	QMC	vdW-DF [62–64]

TABLE VI. Fitting constants in meV for the stacking-fault energy as a function of registry as given by Eq. (2). The error bars from KC-QMC are obtained from the bootstrap method.

Method	c_0	c_1	c_2	c_3
KC-DFT-D2	2.46701	–0.71409	–0.13147	0.02322
KC-Ouyang	1.85743	–0.56041	–0.07521	0.01648
KC-QMC	1.815(38)	–0.546(9)	–0.098(19)	0.039(8)
RPA [34]	1.74278	–0.49974	–0.09179	0.01101
KC-DFT-D3	1.15472	–0.34776	–0.03980	0.00266

TABLE VII. Fitting constants in Å for the relaxed interlayer spacing as a function of registry as given by Eq. (2). The error bars from KC-QMC are obtained from the bootstrap method.

Method	c_0	c_1	c_2	c_3
KC-DFT-D2	3.35457	–0.03093	–0.00147	0.00073
KC-Ouyang	3.47106	–0.02458	–0.00149	0.00056
KC-QMC	3.5171(21)	–0.0245(5)	–0.0021(9)	0.0015(4)
RPA [34]	3.47805	–0.02632	–0.00363	0.00039
KC-DFT-D3	3.60477	–0.01990	–0.00122	0.00007

case of KC-QMC, the chosen parameters are the third row of Table II. The minimum interlayer spacing d_{min} for different stacking registries and potentials is reported in Table IV.

On the other hand, the binding energy (BE) is defined as the energy required to separate the two graphene sheets in equilibrium to the infinite interlayer distance. Therefore, the BE is a property at a large interlayer distance and is evaluated using KC parameters that are fitted with no weights because data points at large interlayer distances now deserve the same weights as data points near equilibrium. The BEs for different stacking registries and potentials are reported in Table III. Our BEs from QMC for AB-stacked and AA-stacked bilayer graphene are 21.96(9) meV and 17.41(9) meV, while the BEs for these two stacking configurations are reported to be 17.7(9) and 11.5(9) meV in the previous QMC study [29], which are slightly smaller than our BE results. The small discrepancy between the QMC BEs may arise from various factors, including differences in QMC data points, fitting functions, and equilibrium spacing d_{min} being used to estimate the BEs between the two studies (Table V). We also observed that precise optimization of the Jastrow factor is crucial for obtaining consistent results, with errors similar to the observed discrepancies when the Jastrow factor is not optimized extremely carefully.

[1] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. USA **108**, 12233 (2011).

[2] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature (London) **556**, 43 (2018).

[3] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras et al., Nature (London) **556**, 80 (2018).

[4] M. H. Naik and M. Jain, Phys. Rev. Lett. **121**, 266401 (2018).

[5] B. Padhi, C. Setty, and P. W. Phillips, Nano Lett. **18**, 6175 (2018).

[6] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael et al., Nat. Phys. **15**, 1174 (2019).

[7] C. Shen, Y. Chu, Q. Wu, N. Li, S. Wang, Y. Zhao, J. Tang, J. Liu, J. Tian, K. Watanabe et al., Nat. Phys. **16**, 520 (2020).

[8] D. Basov, A. Asenjo-Garcia, P. J. Schuck, X. Zhu, and A. Rubio, Nanophotonics **10**, 549 (2021).

[9] B. Padhi, R. Chitra, and P. W. Phillips, Phys. Rev. B **103**, 125146 (2021).

[10] C. R. Dean, L. Wang, P. Maher, C. Forsythe, F. Ghahari, Y. Gao, J. Katoch, M. Ishigami, P. Moon, M. Koshino et al., Nature (London) **497**, 598 (2013).

[11] K. Liu, L. Zhang, T. Cao, C. Jin, D. Qiu, Q. Zhou, A. Zettl, P. Yang, S. G. Louie, and F. Wang, Nat. Commun. **5**, 4966 (2014).

[12] P.-C. Yeh, W. Jin, N. Zaki, J. Kunstmann, D. Chenet, G. Arefe, J. T. Sadowski, J. I. Dadap, P. Sutter, J. Hone et al., Nano Lett. **16**, 953 (2016).

[13] S. Carr, D. Massatt, S. Fang, P. Cazeaux, M. Luskin, and E. Kaxiras, Phys. Rev. B **95**, 075420 (2017).

[14] R. Ribeiro-Palau, C. Zhang, K. Watanabe, T. Taniguchi, J. Hone, and C. R. Dean, Science **361**, 690 (2018).

[15] S. Lisi, X. Lu, T. Benschop, T. A. de Jong, P. Stepanov, J. R. Duran, F. Margot, I. Cucchi, E. Cappelli, A. Hunter et al., Nat. Phys. **17**, 189 (2021).

[16] K. Uchida, S. Furuya, J.-I. Iwata, and A. Oshiyama, Phys. Rev. B **90**, 155451 (2014).

[17] S. Dai, Y. Xiang, and D. J. Srolovitz, Nano Lett. **16**, 5923 (2016).

[18] N. N. T. Nam and M. Koshino, Phys. Rev. B **96**, 075311 (2017).

[19] H. Yoo, R. Engelke, S. Carr, S. Fang, K. Zhang, P. Cazeaux, S. H. Sung, R. Hovden, A. W. Tsen, T. Taniguchi et al., Nat. Mater. **18**, 448 (2019).

[20] S. Pathak, T. Rakib, R. Hou, A. Nevidomskyy, E. Ertekin, H. T. Johnson, and L. K. Wagner, Phys. Rev. B **105**, 115141 (2022).

[21] T. Rakib, P. Pochet, E. Ertekin, and H. T. Johnson, Commun. Phys. **5**, 242 (2022).

[22] X.-F. Zhou, Y.-W. Liu, C.-Y. Hao, C. Yan, Q. Zheng, Y.-N. Ren, Y.-X. Zhao, K. Watanabe, T. Taniguchi, and L. He, Phys. Rev. B **107**, 125410 (2023).

[23] M. J. Allen and D. J. Tozer, J. Chem. Phys. **117**, 11113 (2002).

[24] U. Zimmerli, M. Parrinello, and P. Koumoutsakos, J. Chem. Phys. **120**, 2693 (2004).

[25] S. Tsuzuki and H. P. Lüthi, J. Chem. Phys. **114**, 3949 (2001).

[26] S. Grimme, J. Comput. Chem. **27**, 1787 (2006).

[27] V. Barone, M. Casarin, D. Forrer, M. Pavone, M. Sambi, and A. Vittadini, J. Comput. Chem. **30**, 934 (2009).

[28] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[29] E. Mostaani, N. D. Drummond, and V. I. Fal'Ko, Phys. Rev. Lett. **115**, 115501 (2015).

[30] P. Ganesh, J. Kim, C. Park, M. Yoon, F. A. Reboredo, and P. R. Kent, J. Chem. Theory Comput. **10**, 5318 (2014).

[31] A. Benali, L. Shulenburger, N. A. Romero, J. Kim, and O. A. von Lilienfeld, J. Chem. Theory Comput. **10**, 3417 (2014).

[32] H. Shin, J. Kim, H. Lee, O. Heinonen, A. Benali, and Y. Kwon, J. Chem. Theory Comput. **13**, 5639 (2017).

[33] J. T. Krogel, S. F. Yuk, P. R. Kent, and V. R. Cooper, J. Phys. Chem. A **124**, 9867 (2020).

[34] S. Zhou, J. Han, S. Dai, J. Sun, and D. J. Srolovitz, Phys. Rev. B **92**, 155438 (2015).

[35] W. Ouyang, D. Mandelli, M. Urbakh, and O. Hod, Nano Lett. **18**, 6009 (2018).

[36] A. N. Kolmogorov and V. H. Crespi, Phys. Rev. B **71**, 235415 (2005).

[37] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen *et al.*, Comput. Phys. Commun. **271**, 108171 (2022).

[38] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[39] J. Kim, A. D. Baczewski, T. D. Beaudet, A. Benali, M. C. Bennett, M. A. Berrill, N. S. Blunt, E. J. L. Borda, M. Casula, D. M. Ceperley *et al.*, J. Phys.: Condens. Matter **30**, 195901 (2018).

[40] L. Mitáš and R. M. Martin, Phys. Rev. Lett. **72**, 2438 (1994).

[41] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).

[42] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni *et al.*, J. Phys.: Condens. Matter **29**, 465901 (2017).

[43] P. Giannozzi, O. Baseggio, P. Bonfà, D. Brunato, R. Car, I. Carnimeo, C. Cavazzoni, S. De Gironcoli, P. Delugas, F. Ferrari Ruffino *et al.*, J. Chem. Phys. **152**, 154105 (2020).

[44] J. Trail and R. Needs, J. Chem. Phys. **122**, 014112 (2005).

[45] J. Trail and R. Needs, J. Chem. Phys. **122**, 174109 (2005).

[46] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[47] K. Krongchon, B. Busemeyer, and L. K. Wagner, J. Chem. Phys. **146**, 124129 (2017).

[48] C. Lin, F. H. Zong, and D. M. Ceperley, Phys. Rev. E **64**, 016702 (2001).

[49] D. W. Brenner, O. A. Shenderova, J. A. Harrison, S. J. Stuart, B. Ni, and S. B. Sinnott, J. Phys.: Condens. Matter **14**, 783 (2002).

[50] Y. Wu, L. K. Wagner, and N. R. Aluru, J. Chem. Phys. **144**, 164118 (2016).

[51] A. Ambrosetti, A. M. Reilly, R. A. DiStasio, Jr., and A. Tkatchenko, J. Chem. Phys. **140**, 18A508 (2014).

[52] S. Dai, Y. Xiang, and D. J. Srolovitz, Phys. Rev. B **93**, 085410 (2016).

[53] S. Carr, S. Fang, P. Jarillo-Herrero, and E. Kaxiras, Phys. Rev. B **98**, 085144 (2018).

[54] Z. A. Goodwin, V. Vitale, X. Liang, A. A. Mostofi, and J. Lischner, Electron. Struct. **2**, 034001 (2020).

[55] P. Lucignano, D. Alfè, V. Cataudella, D. Ninno, and G. Cantele, Phys. Rev. B **99**, 195419 (2019).

[56] S. Wu, Z. Zhang, K. Watanabe, T. Taniguchi, and E. Y. Andrei, Nat. Mater. **20**, 488 (2021).

[57] https://github.com/qmc-hamm/qmc_graphene_stacking_fault

[58] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes, Phys. Rev. B **78**, 125106 (2008).

[59] T. Olsen and K. S. Thygesen, Phys. Rev. B **88**, 115131 (2013).

[60] I. Razado-Colambo, J. Avila, D. Vignaud, S. Godey, X. Wallart, D. Woodruff, and M. Asensio, Sci. Rep. **8**, 1 (2018).

[61] I. Brihuega, P. Mallet, H. González-Herrero, G. T. De Laissardière, M. M. Ugeda, L. Magaud, J. M. Gómez-Rodríguez, F. Ynduráin, and J.-Y. Veuillen, Phys. Rev. Lett. **109**, 196802 (2012).

[62] H. Rydberg, M. Dion, N. Jacobson, E. Schröder, P. Hyldgaard, S. I. Simak, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **91**, 126402 (2003).

[63] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **95**, 109902(E) (2005).

[64] G. Román-Pérez and J. M. Soler, Phys. Rev. Lett. **103**, 096102 (2009).