Local atomic stacking and symmetry in twisted graphene trilayers

Contents

1 Influence of Bragg Disk Variability on Virtual Apertures

Defining virtual apertures to target overlap regions of interest from the averaged diffraction patterns may introduce artifacts in heterogeneous samples wherein the variability of the Bragg disk positions leads to the inclusion of pixels from outside the interference region, so we verify that the chosen virtual aperture region does not include errant contributions from the diffuse background or other interference regions. Biasing from this effect was avoided by imaging samples with relatively low sample variability due to local tilting and/or strain as shown below and through using relatively small virtual apertures. We note that all samples used were very homogeneous, seen in the variation in disk position of the order of 0.01 nm^{-1} within the provided representative acquisition region in Supplemental Figure 1 (we note that there is some uncertainty associated with fitting disk locations associated with thin moiré structures such as these). Bleed-in of these unintended regions manifested in regions along the edges of a given disk's virtual dark field with considerably less intensity (or an unexpected low or high-frequency intensity variation from bleed-in of another overlap region), allowing us to further screen our virtual aperture masks. A down-sizing of the overlap regions of roughly 25% was therefore adequate to account for this variation in disk position of $\approx 10\%$ of the Bragg disk diameter. We note that the viability of this down-sizing approach requires a decent number of pixels in the overlap regions, placing restrictions on the choice of camera length and convergence angle as discussed in Supplemental section 4.

2 Choice of Acquisition Parameters

Typically the convergence semi-angle is chosen to minimize the probe size through achieving the right balance between abberations and the diffraction limit in approaches that do not make use of the probe tails. For the interferometric approach we use, the optimal convergence semi-angle is somewhat larger and depends on a few additional sample-dependent considerations, such as the moiré twist angle and the substrate choice and orientation. We note that a relatively large probe size is needed to obtain sufficiently many pixels in the disk overlap regions of interest, particularly for samples with larger twist angles or when the
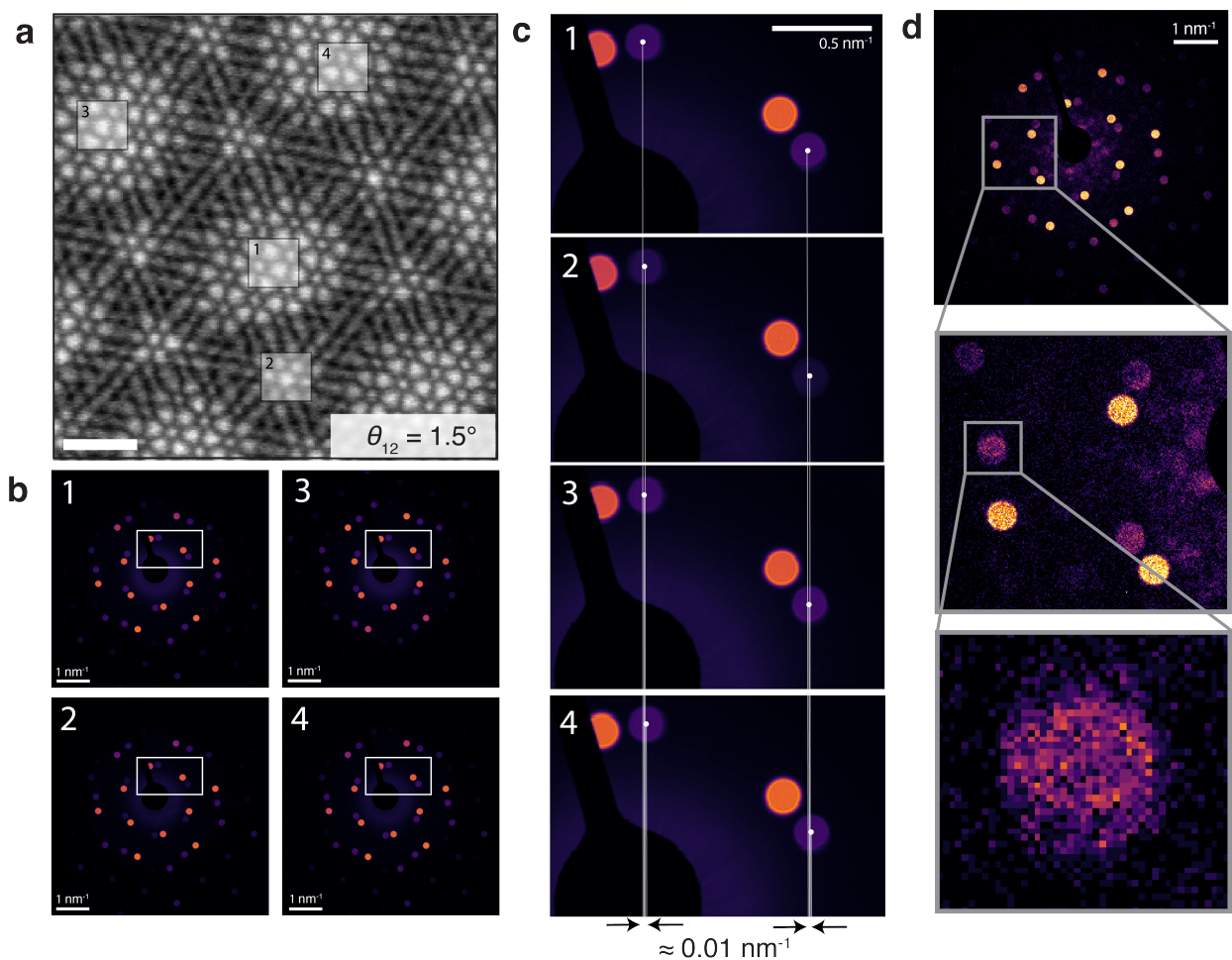
Supplemental Figure 1: Bragg Disk Variability (a) Virtual dark field images corresponding to the overlap of all three layers from the $\theta_{13}, \theta_{12} = 0.22^\circ, 1.5^\circ$ sample shown in Main Figure 1. Scale bar is 25 nm. **(b)** Representative convergent beam electron diffraction (CBED) patterns within the numbered regions depicted in **(a)**. **(c)** Enlarged regions from the CBED patterns illustrating degree of Bragg disk variability. **(d)** Representative CBED pattern and enlarged regions illustrating the signal to noise in the diffraction pattern obtained for a real space pixel.

camera factor is larger as a result of the small accelerating voltage. However increasing the probe size past its conventional optimal also increases the signal to noise associated with each pixel. Large probe sizes may also lead to overlap with undesired Bragg disks from the substrate or an inability to resolve separate overlap regions in multi-layered samples. As seen in figure 3a of the main text, a small bleed-in of the smaller moire pattern was hard to complete exclude. We give a few considerations and equations for choosing an appropriate set of imaging conditions below, but note that the optimal set-up will be dependent on choice of sample, substrate, and detector.

A set of two Bragg disks each with radius r at a distance D away from 000 reflection that are offset by θ_{ij} from one another have the area of overlap given below. In practice the probe radius should be multiplied by a factor representing the proportion of theoretically accessible pixels that can reasonably be extracted using our approach (we downsize each region by 25%).

$$A(\theta_{ij}) = 2r^2 cos^{-1}\left(\frac{D\sqrt{2 - 2\cos\theta_{ij}}}{2r}\right) - \sqrt{-D^2(-1 + \cos\theta_{ij})(-D^2 + 2r^2 + D^2\cos\theta_{ij})} \quad (1)$$

When we have a trilayer system (taking the convention that the largest angle is $\theta_{13} = \theta_{12} + \theta_{23}$), we have areas of overlap associated with layers 123, 23, and 12 of $A(\theta_{13})$, $A(\theta_{23}) - A(\theta_{13})$ and $A(\theta_{12}) - A(\theta_{13})$ respectively. To obtain any pixels at all in the 123 layer overlap region, this requires that $\sqrt{2}r/D > \sqrt{1 - \cos\theta_{13}}$. We found that a convergence angle of 1.71 mrad and camera length of 800 mm was sufficient for imaging features on the moiré length-scale for the range of samples presented in this work. At these imaging conditions the Bragg disk radius r is 0.18 nm^{-1} and the distance D between the central beam and a first order Bragg reflection is 2.09 nm^{-1}. Therefore the condition $\sqrt{2}r/D >> \sqrt{1 - \cos\theta_{13}}$ is satisfied for angles up to even a few degrees and we are left with sufficiently many pixels to obtained the virtual dark fields as described. We note that some of these limitations may be overcome by focusing on only higher order reflections instead of only the first and second peaks used in this work or from allowing the overlap regions to vary over the sample instead of using a

fixed regions identified from the averaged diffraction pattern (which would permit use of a larger fraction of accessible pixels).
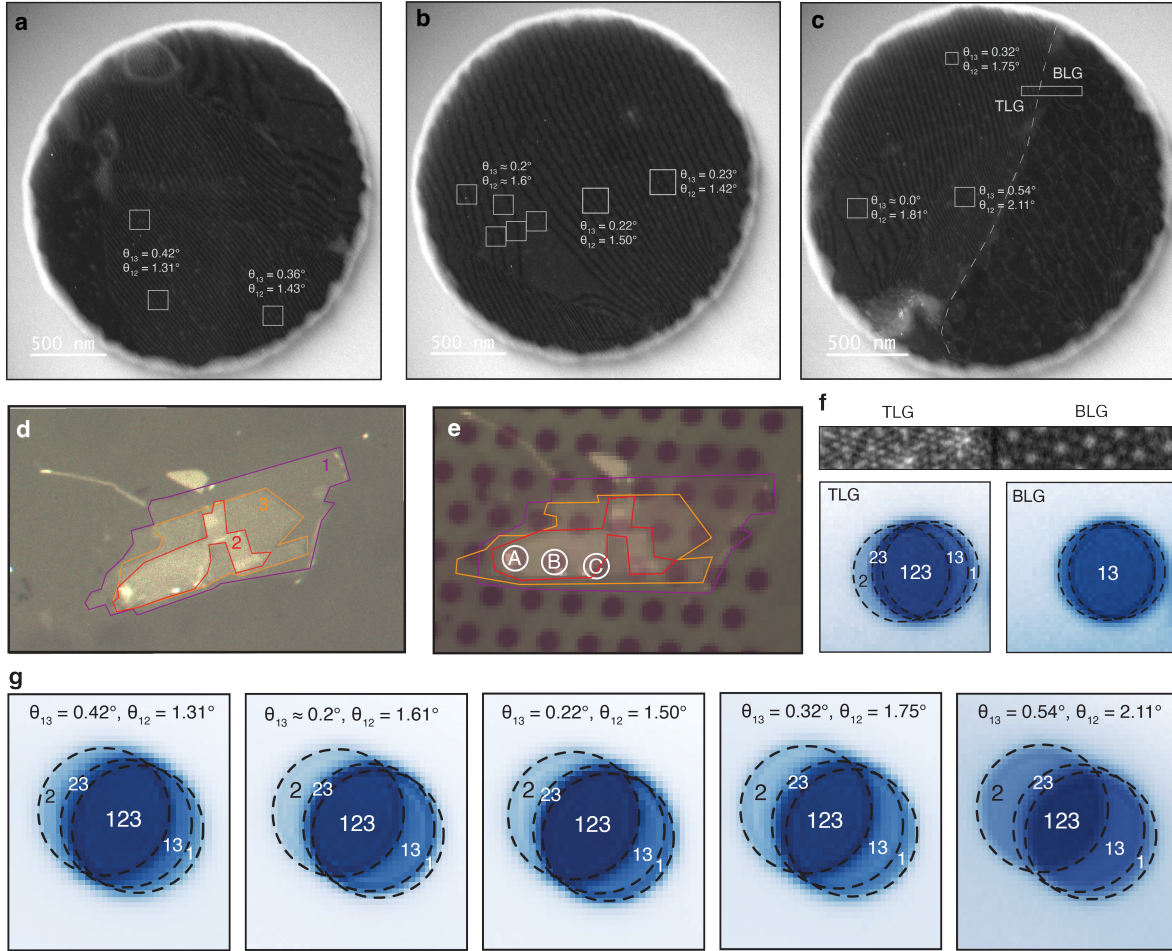
3 Attribution of Bragg Disks to Layers

We are unable to determine if a measured trilayer sample is in the $\theta_{13} < \theta_{23}$ or $\theta_{13} > \theta_{23}$ regime from the 4D-STEM data-set alone, as the approach described is insensitive to out-of-plane orientation. We therefore leverage additional information to attribute each of the measured Bragg disks to a given graphene layer. This is accomplished using the known order in which each graphene layer was picked up, optical micrographs of the individual graphene layers prior to, and after assembly into the heterostructure, and conventional dark-field electron micrographs of the samples assembled on TEM grids. Taken together we are able to attribute regions of the dark field images (Supplemental Figure 2A-C) to either twisted trilayer graphene or a region of twisted bilayer graphene in which only layers 1 and 3 are present. We then attribute Bragg disks to graphene layers by collecting a 4D-STEM data over the TLG/BLG interface as seen in Supplemental Figure 2F, which show that the smallest twist angle in the sample is between layers 1 and 3. Comparison of the graphene $2\overline{11}0$ Bragg reflections in other selected portions of the sample (Supplemental Fig. 2G) further suggests that θ_{13} only gradually increases from left to right in the sample and that $\theta_{13} << \theta_{23}$ for all of the investigated regions of this sample. We note that we are unable to distinguish between layers 1 and 3 in this approach and assume that $\theta_{12} > \theta_{23}$ throughout. As we do not consider any substrate effects in this work, this convention has no bearing on our analysis.

4 Predicted Intensities for Bilayer Interfaces

The intensity of two overlapping Bragg disks can be modeled by the following, where I and $\mathbf{g}$ are the intensity and reciprocal space vector for a given Bragg disk, A and B are coefficients, and $\mathbf{u}$ is the inter-layer displacement vector [1]. We note that this expression assumes a center of inversion symmetry within the sample [2].

$$I(\mathbf{g}) = A(\mathbf{g})\cos^2(\pi\mathbf{g}\cdot\mathbf{u}) + B(\mathbf{g})$$

Supplemental Figure 2: Attribution of Bragg disks to layers for $\theta_{13} << \theta_{23}$ regime TTLG sample (**a-c,f**) Dark field electron micrographs obtained from the trilayer structure, with optical microscope images shown in (**d-e**). BLG graphene region corresponds to the presence of only layers 1 and 3, seen by comparison to (**d-e**). 4D-STEM data-sets used in this work were obtained in the the boxes overlaid in (**a-c**). The 4D-STEM scan obtained over the TLG/BLG boundary and corresponding $10\bar{1}0$ Bragg disks illustrate that $\theta_{13} \ll \theta_{23}$. (**g**) Selected $2\bar{1}\bar{1}0$ Bragg disks for 4D-STEM analysis, illustrating that $\theta_{13} \ll \theta_{23}$ holds true for all data-sets shown here. Bragg disk intensities are shown on a logarithmic scale throughout.

In general the fitting parameters A and B depend on the Bragg reflection in question and thus reciprocal space vector. However whenever the underlying atomic lattice has local C_3 symmetry, the fitting parameters will be the same for all Bragg reflections of the same order (such that the intensity depends only on the magnitude of $\mathbf{g}$). In previous work this dependence of the overlap intensity on the orientation of $\mathbf{g}$ allowed us to distinguish between differently oriented soliton features. However independently fitting the overlap intensity along each of these reciprocal lattice directions is not strictly necessary if we aim to only distinguish soliton from pockets of high symmetry stacking but not determine their orientation. In this work, we therefore instead only investigated the average intensities in the first and second order Bragg disks. This simpler approach distinguishes between AA, AB, and SP-type stacking orders but averages away the information necessary to distinguish between different SP orientations. In practice we averaged the intensity associated with all six first order Bragg reflections to minimize measurement noise.

$$I_1 = (I_{0\bar{1}10} + I_{\bar{1}010} + I_{\bar{1}100} + I_{01\bar{1}0} + I_{10\bar{1}0} + I_{1\bar{1}00})/6$$

$$I_2 = (I_{2\bar{1}\bar{1}0} + I_{\bar{2}110} + I_{1\bar{2}10} + I_{\bar{1}2\bar{1}0} + I_{\bar{1}\bar{1}20} + I_{11\bar{2}0})/6$$

When we write the displacement vector in the lattice vector basis as $\mathbf{u} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2$ where $\mathbf{a}_1 = a_0(1,0)$, $\mathbf{a}_2 = a_0(1/2, \sqrt{3}/2)$ and a_0 is the lattice constant for graphene, we can compute the averaged intensity of all first and second order Bragg reflections as follows:

$$I_1 = A_1(\cos^2(\pi c_1) + \cos^2(\pi c_2) + \cos^2(\pi(c_1 - c_2))) + B_1$$

$$I_2 = A_2(\cos^2(\pi(c_1 + c_2)) + \cos^2(\pi(2c_1 - c_2)) + \cos^2(\pi(2c_2 - c_1))) + B_2$$

These resulting expressions highlight that the average intensities, once normalized to remove dependence on the constants (such that $A_i = 4/9$ and $B_i = -1/3$ for I_i in the range of $[0,1]$), can be used to distinguish between all relevant high-symmetry stacking orders in graphene trilayers without the need for any fitting parameters. Therefore only two intensity expressions and thus a bi-variate colormap are sufficient to resolve the two unknowns c_1, c_2.

However, avoiding rigorously fitting the diffraction intensities in favor of a simple normalization, while convenient, leaves the stacking order attributions sensitive to noise. This is addressed in further detail Supplemental section 10. Accompanying multi-slice simulations (Supplemental Figure 4) support the validity of this approach for the trilayer structures shown in this work.

We note that materials with in-plane projections that break inversion symmetry about $\mathbf{u} = 0$ (such as the ABt trilayers since ABC $\neq$ ABA) require the inclusion of additional odd terms [2]. We are however able to overcome this by instead shifting the origin of the displacement field $\mathbf{u}$ to coincide with the ABC-type stacking $\mathbf{u} = a_1/3 + a_2/3 = \mathbf{a}_0/2(1, 1/\sqrt{3})$, which is a center of inversion symmetry for the in-plane projection of $\mathbf{u}$. The resultant expression for the ABt trilayer intensities is then as follows for the normalized intensities I_i. We also note that now $A_1 < 0$ as the ABC-type stacking order at the origin is a associated with a local minima in scattering intensity instead of a local maxima as was the case for the AAA-stacking, as is seen in the accompanying multi-slice simulations. Practically this is accomplished using the following expressions for the normalized ABt intensities.

$$
\begin{aligned}
I_1 &= \frac{4}{3} - \frac{4}{9}(\cos^2(\pi(c_1 - 1/3)) + \cos^2(\pi(c_2 - 1/3)) + \cos^2(\pi(c_1 - c_2))) \\
I_2 &= \frac{4}{9}(\cos^2(\pi(c_1 + c_2 - 2/3)) + \cos^2(\pi(2c_1 - c_2 - 1/3)) + \cos^2(\pi(2c_2 - c_1 - 1/3))) - \frac{1}{3}
\end{aligned}
$$

It is also important to note as that the derived expressions for intensity modulation depend only on the in-plane projection of the displacement vectors, so we can investigate samples that differ only in their out-of-plane orientation (such as tAB and AtB) with an identical framework.

5 Predicted Intensities for Trilayers

We can derive the expected intensity within a region of three overlapping graphene Bragg disks from the weak phase object approximation in a similar manner to previous work [1, 2]. We first assume that we have a centrosymmetric trilayer structure in which each layer's
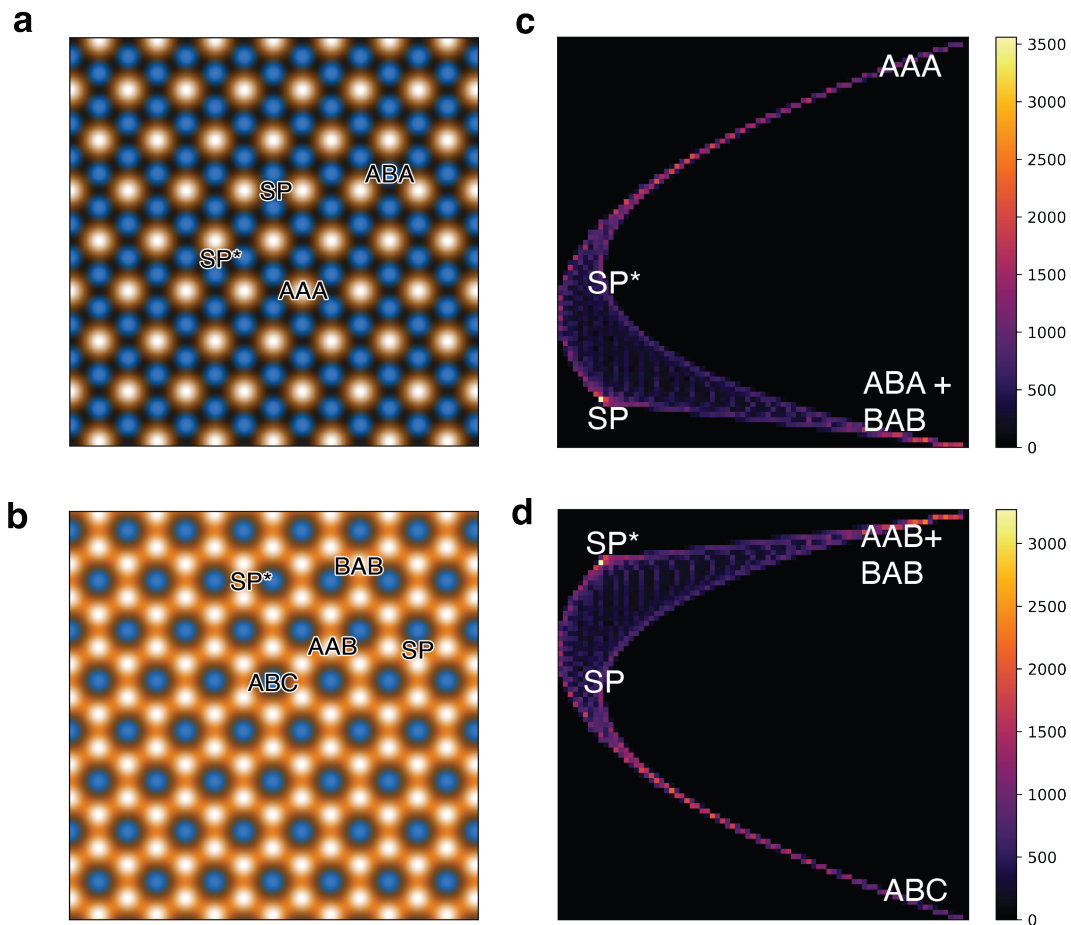
Supplemental Figure 3: Stacking order maps for rigid AtA and tAB. Schematics illustrating the stacking order map expected for a rigid AtA structure **(a)** and tAB structure **(b)**, obtained via the expressions provided in Supplemental section 7. (**c, d**) Histograms illustrating the relative prevalence of each stacking configuration, binned according to average first and second order Bragg disks intensities (using 80x80 bins) wherein color indicates counts.

projected electrostatic potential is $V(\mathbf{r})$. We define the displacement vectors $\mathbf{u}_{12}$, $\mathbf{u}_{23}$, and $\mathbf{u}_{13}$ associated with layers 1&2, layers 2&3, and layers 1&3 respectively. The total projected potential is then given by the sum $V(\mathbf{r} - \mathbf{u}_{12}) + V(\mathbf{r}) + V(\mathbf{r} + \mathbf{u}_{23})$. Assuming that diffraction occurs within a single plane and using the weak phase object approximation, the ψ associated with the outgoing electron beam can therefore be expressed as the following, in terms of the unscattered electron beam ψ_0 and the relativistic interaction parameter σ

$$\psi(\mathbf{r}) \approx (1 + i\sigma V(\mathbf{r} - \mathbf{u}_{12}) + i\sigma V(\mathbf{r}) + i\sigma V(\mathbf{r} + \mathbf{u}_{23}))\psi_0(\mathbf{r}) \tag{2}$$

The Fourier space intensity $I(\mathbf{k})$ is then given by the following, in which $\circledast$ denotes convolution.

$$I(\mathbf{k}) = |\psi(\mathbf{k})|^2 = |(\delta(\mathbf{k}) + i\sigma \sum_{\mathbf{g}} \delta(\mathbf{k} - \mathbf{g})V(\mathbf{g})(1 + e^{2\pi i \mathbf{g}\cdot\mathbf{u}_{12}} + e^{-2\pi i \mathbf{g}\cdot\mathbf{u}_{23}})) \circledast \psi_0(\mathbf{r})|^2 \tag{3}$$

Assuming that the convergence semi-angle is chosen to ensure that Bragg disks within a single layer do not overlap, we therefore have the following relationship between the projected displacement vectors and the modulation in intensity I_j for a set of selected Bragg disks at positions $\mathbf{g}_j$. We note that we defined $A(\mathbf{g}) = 4\sigma^2|V(\mathbf{g})|^2$, $B(\mathbf{g}) = -3\sigma^2|V(\mathbf{g})|^2$ and used that $e^{2ix} + e^{-2ix} = 2cos(2x) = 4cos^2(x) - 2$.

$$I(\mathbf{g}_j) = A(\mathbf{g}_j)(cos^2(\pi\mathbf{g}_j \cdot \mathbf{u}_{12}) + cos^2(\pi\mathbf{g}_j \cdot \mathbf{u}_{13}) + cos^2(\pi\mathbf{g}_j \cdot \mathbf{u}_{23})) + B(\mathbf{g}_j) \tag{4}$$

6   Comparison to Multi-slice Simulations and Finite Probe Effects

In order to validate the derived intensity expressions used to form the bi-variate colormap with which the diffraction patterns are analyzed, we preformed a series of multi-slice simulations. Multislice simulations were carried out using ABTEM [3] with an acceleration energy of 80 keV, convergence semi-angle cutoff of 4.0 mrad, and rolloff of 0.5, and a potential sampling of 0.2 Å. We made use of the infinite projection scheme for all plots shown. The intensities in the first and second order Bragg disks in the resultant theoretical diffraction

patterns were integrated to obtain the results shown in Supplemental Figure 4 using the same procedure outlined in Supplemental section 3.

The obtained dependence of Bragg disk intensity on stacking order from multi-slice (circles) agree well with the analytical expressions presented (solid lines) and do not lead to any qualitative differences in the colored plots shown. We do note a small (20%) discrepancy between the multi-slice intensities and those obtained from the expression used within the soliton regions between AAB and BAB regions within the ABt sample. Analyzing the data with respect to the derived expression will therefore impart a small overestimation of these soliton regions, however we note that this intensity discrepancy is less than noise-driven normalization bias discussed in Supplemental section 10 and standard deviation of averaged line-cuts shown (Fig. 2F, 2L) and therefore likely does not feature predominantly in analysis.
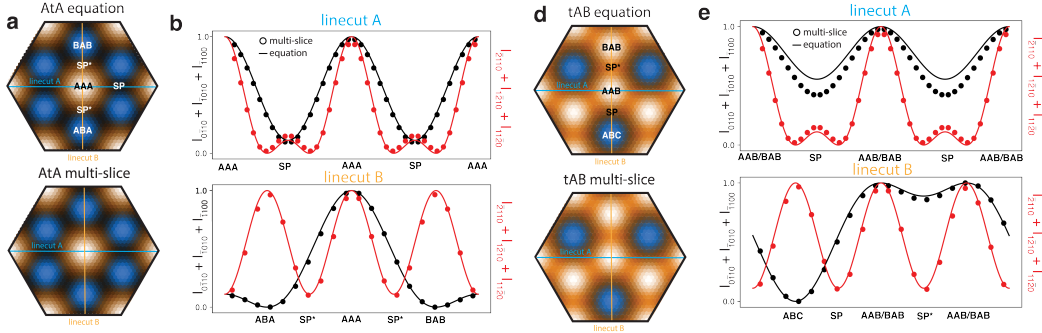
Supplemental Figure 4: Comparison between Simulated, Expected, and Experimental Intensities. (a,d) Comparison of the intensity distribution for AtA and tAB bilayers obtained via multi-slice simulation and from the expressions provided in Supplemental section 7 with **(b,e)** corresponding line-cuts.

We also note a small systematic reduction in the second order Bragg disk intensity associated with the AAA stacking regions when comparing the experimental data to the predicted intensities, which is robust to noise. We ascribe this to an averaging effect associated with the finite spatial resolution. This effect arises because the AAA regions display an appreciable contraction to form domains on the order of a nanometer across or less, below the 1.25 nm spatial resolution imposed by the finite probe width. The regions labeled AAA within the experimental stacking order maps and line-cuts therefore include a measurable bleed-in from neighboring low intensity regions, decreasing the intensity of the peak in the second order

11

Bragg disk intensities. This effect is observed predominantly in the second order intensity owing to the fact that the second order intensity peak is considerably sharper than the first order intensity peak associated with this region. This decrease in intensity persists following normalization due to the ABA regions, which are expected to have comparable second order intensity and instead retain sizes larger than the spatial resolution following reconstruction. We note however that while this prevents us from drawing quantitative claims regarding the atomic reconstruction of the smaller moiré pattern, it does not effect the measurements of the larger scale moiré reconstruction shown in main text Figure 4. Further we are still able to measure the rough trends in atomic reconstruction within the smaller moiré qualitatively as these intensity signatures can still provide evidence of a relative contraction of the AAA domains compared to the ABA domains.

7   Effects of Normalization Bias

In avoiding the use of fitting parameters, the simpler interferomic approach employed in this work is more susceptible to noise. In particular, variation in the peak and background scattering intensity over a sample (driven by factors such as sample tilting and impurities such as carbon build-up and hydrocarbon residues) will impart a spatial variation in the coefficients A and C presented in Supplemental section 7. This limitation is not specific to this work, as previous fitting of these interference fringes assumed that the fitting parameters were uniform over the field of view. However such an approach is still preferable to obtaining constants based on the maximum and minimum intensities within a data set. This bias was mitigated by smoothing the virtual dark field images with a Gaussian filter ($\sigma = 0.5$ nm) prior to normalization, however it remains important to determine if the observed stacking distributions are robust with respect to this noise-driven normalization bias.

Supplemental Figure 5 illustrates line profiles associated with the observed stacking order modulation within AtA (Supplemental Figure 5A, 5C) and tAB (Supplemental Figure 5B, 5D) samples, a portion of which (Supplemental Figure 5A-B) was already shown in Figure 2 and is repeated here for convenience in comparison. When the normalization constants
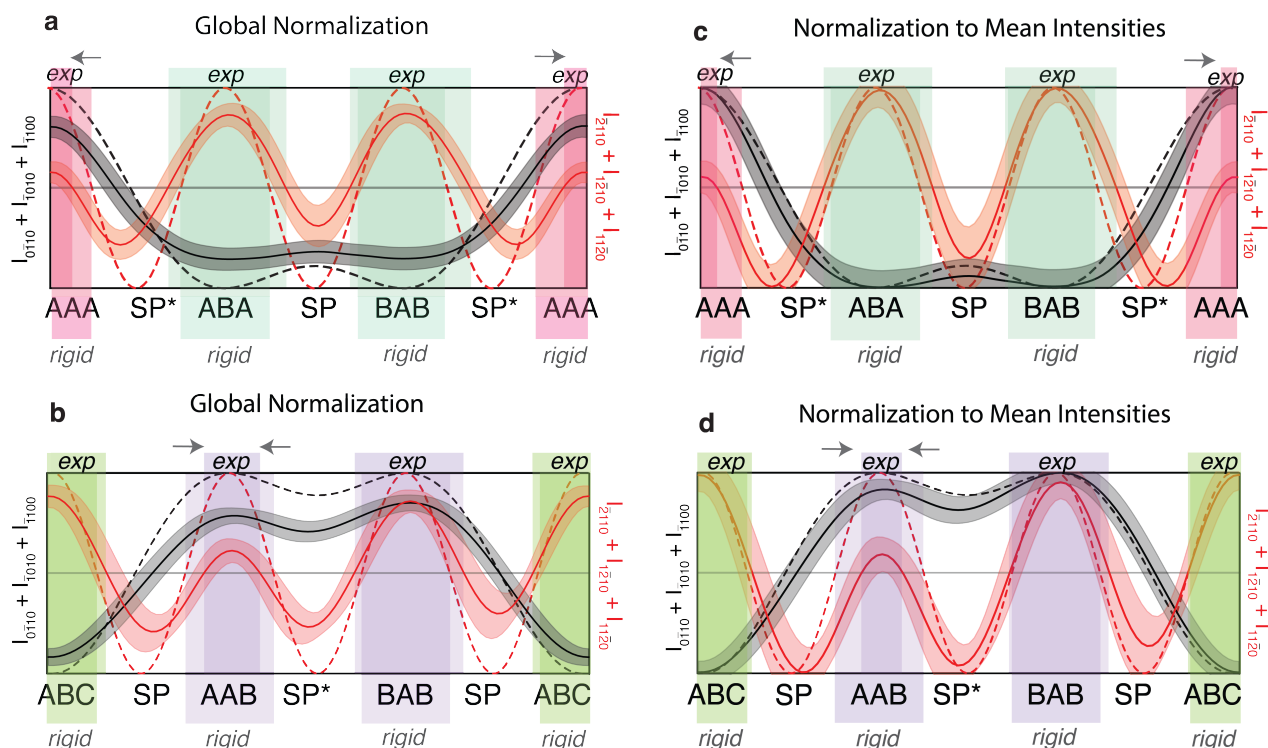
Supplemental Figure 5: Effects of Normalization Bias on Observed Reconstruction (**a,c**) Line-cut illustrating the progression of stacking order within an AtA sample as shown in Figure 2 of the Main text obtained from normalizing the entire data-set together (**a**) and from a locally optimal normalization (**c**). (**b,d**) Analogous line-cuts for the tAB sample. Rigid intensity profiles (dotted lines) are computed using the expressions in Supplemental section 7. Domain sizes were obtained from the full width at half max of $I_{2\bar{1}10} + I_{\bar{2}110} + I_{1\bar{2}10}$ with $I_{0\bar{1}10} + I_{\bar{1}010} + I_{\bar{1}100}$ only serving to distinguish between sites as in Fig 2. We note for comparison that geometric portioning in Fig 3 and Supplemental Figs 6,9,10 uses a slightly different convention described in Supplemental section 11. These two conventions yield comparable trends. Shaded regions represent the standard deviation throughout.

are chosen based on the maximum and minimum values in the entire smoothed data-set (Supplemental Figures 5A-B), we observe a contraction of the AAA domains, as noted in the main text. The same general trend is seen when normalizing to the averaged line profile (Supplemental Figure 5C-D), and the extent of contraction observed is roughly 13 percent larger in the AAA domains (the two methods yield domain sizes of 48% and 35% of the rigid domain size respectively). We note that the presented standard deviation predicts a reduction in domain size to 27-72% of the rigid lattice, and the error margin from the standard error is insignificant as over 200 line-cuts were used. For the tAB sample (Supplemental Figure 5B, 5D), wherein the AAB domains shrink, the two methods yield AAB domain sizes of 63% and 46% of the rigid domain size respectively and the standard deviation in line-cuts predicts a AAB domain size of 36-94% of the rigid domain (again with a negligible margin from the standard error).

We therefore note that the approach used is sufficient to draw qualitative claims regarding the approximate extent of reconstruction in these samples, but a systematic normalization bias motivates more a quantitative investigation in ongoing future work. All results presented in the main text were obtained from samples with an extent of noise and sample homogeneity comparable that in Supplemental Figure 5 and claims regarding reconstruction are avoided or discussed in the context of these caveats. We would like to emphasize that the reconstruction trends that we most seek to emphasize for this work pertain to the larger-scale moire pattern (and specifically the persistence of AtB regions), for which such a high uncertainty is not present. Further, the bias discussed does not obscure the trends drawn.

8 Supplemental Atomic Stacking Maps

Additional atomic stacking maps for AtA and tAB regions and corresponding stacking statistics are provided in Supplemental Figure 6. We can see from these atomic stacking maps that AtA samples (Supplemental Figure 6A-B) in the $1.81°$ and $1.0°$ regime are structurally alike, similar of twisted bilayers within this twist angle range [1]. The atomic stacking of tAB samples (Supplemental Figure 6C-D) in the 1.5 degree and 0.1 degree regime show a pattern

of reconstruction to decrease the proportion of AAB stacking with little preference between ABC and BAB, as noted in the main text, and is more dramatic at smaller twist angles. Additional atomic stacking maps for TTLG regions and corresponding stacking statistics (obtained via the procedure outlined in Supplemental section 11) are provided in Figures S7 and S8. Supplemental Figure 9 displays the portions of each stacking order seen within the larger AtA, AtB, and tAB domains. We see that the larger scale moire pattern has little effect on the extent of reconstruction within AtA domains, which appear very similar to the stacking distribution seen within the $\theta_{13} \approx 0$ regions (shown as dotted lines in Supplemental Figure 9). Maps of stacking portion against both variable twist angles in Supplemental Figure 10 shows that the relative portion of AtA/tAA, AtB/tAB, and SP stacking within the larger scale moiré pattern depends primarily on the smaller twist angle, motivating the analysis and visualization in Figure 3A of the main text.

9   Continuum Relaxation Model

To calculate the relaxation-redistributed local twist angle, we employ a continuum relaxation model in local configuration space [4]. In twisted trilayer graphene with two independent twist angles, there does not exist a largest length scale and the system is incommensurate [4]. Therefore, instead of formulating the problem in real space, we adopt configuration space, which describes the local environment of every position in layer L_ℓ and bypasses a periodic approximation [5]. Every position in real space $\boldsymbol{r}$ in L_i can be uniquely parameterized by three shift vectors $\boldsymbol{b}^{i \to j}$ for $j = 1, 2, 3$ that describes the relative position between a given point $\boldsymbol{r}$ with respect to all three layers. Note that $\boldsymbol{b}^{i \to j} = \boldsymbol{0}$ if $i = j$ since the separation between a position with itself is 0, which leads to a four-dimensional configuration space.

For a given real space position $\boldsymbol{r}$, the following linear transformation relates $\boldsymbol{r}$ and $\boldsymbol{b}^{i \to j}$ in layer i with respect to layer j, and following linear transformation maps the relaxation from the local configuration to the real space positions $\boldsymbol{r}$:

$$\boldsymbol{b}^{i \to j}(\boldsymbol{r}) = (E_j^{-1} E_i - \mathbb{1})\boldsymbol{r}, \tag{5}$$

15
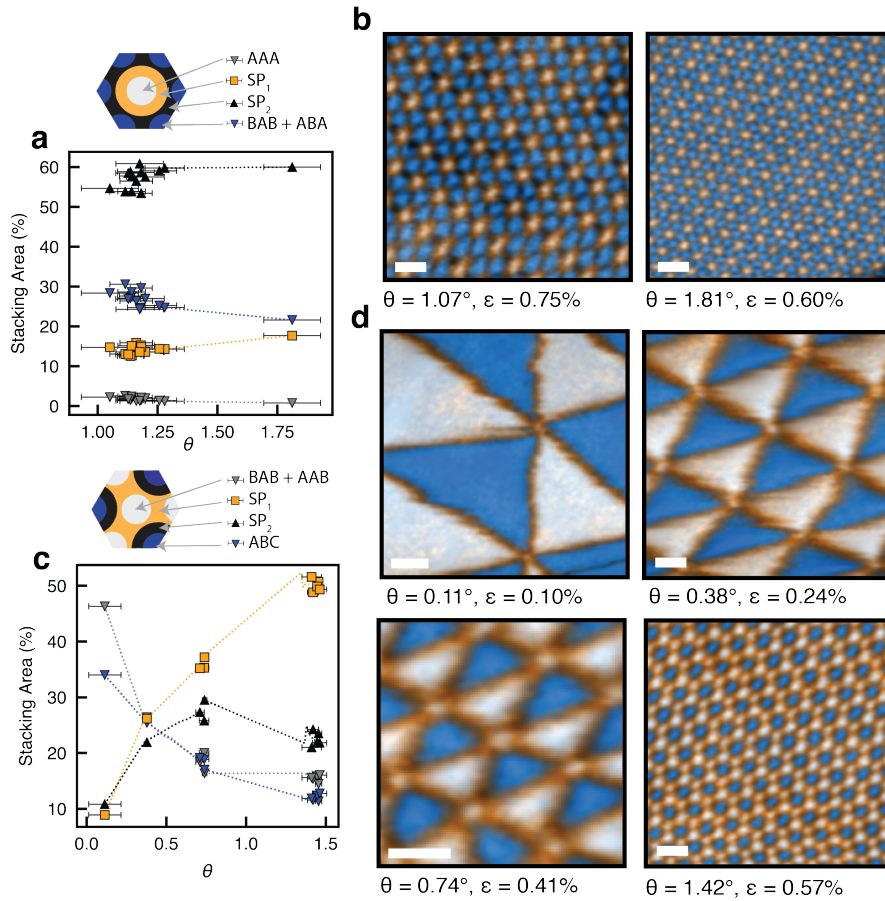
Supplemental Figure 6: Supplemental Atomic Stacking in AtA and tAB trilayers.
(**a**) Plot illustrating how the proportion of AAA stacking (grey), ABA + BAB stacking (blue), and lower symmetry stackings (black and yellow) change with twist angle in the AtA regions measured. Hexagonal insets illustrates how stacking areas are partitioned within the moiré unit cell. (**b**) Corresponding maps of local atomic stacking in representative AtA regions. (**c**) Plot illustrating how the proportion of BAB + AAB stacking (grey), ABC stacking (blue), and lower symmetry stackings (black and yellow) change with twist angle in the tAB regions measured. Hexagonal insets illustrates how stacking areas are partitioned within the moiré unit cell. (**d**) Corresponding maps of local atomic stacking in representative tAB regions. All scale bars are 25 nm. Scale bars represent standard deviations and dotted lines are piece-wise linear fits to guide the eye. Error bars denote standard deviations collected from 3-253 points for each data point.
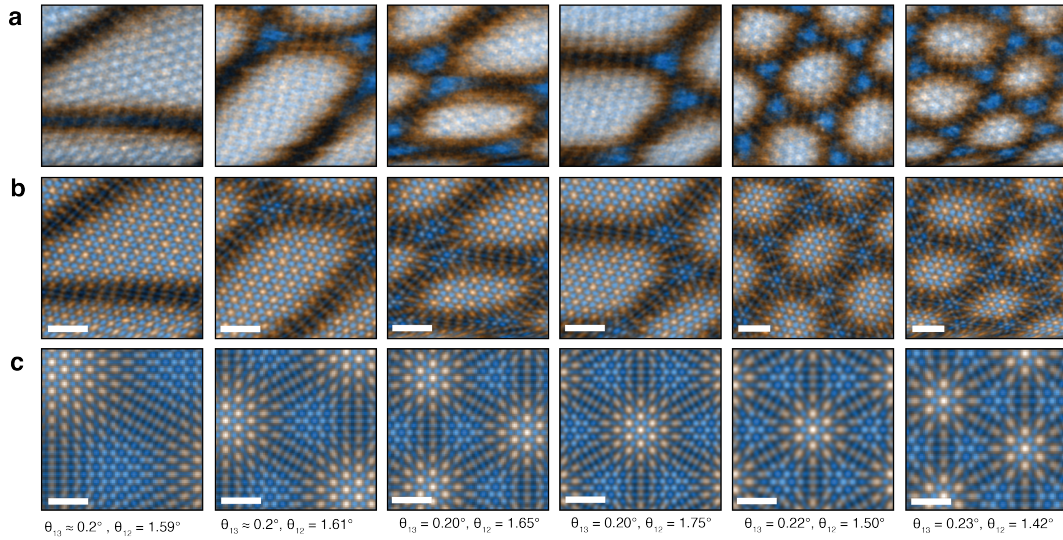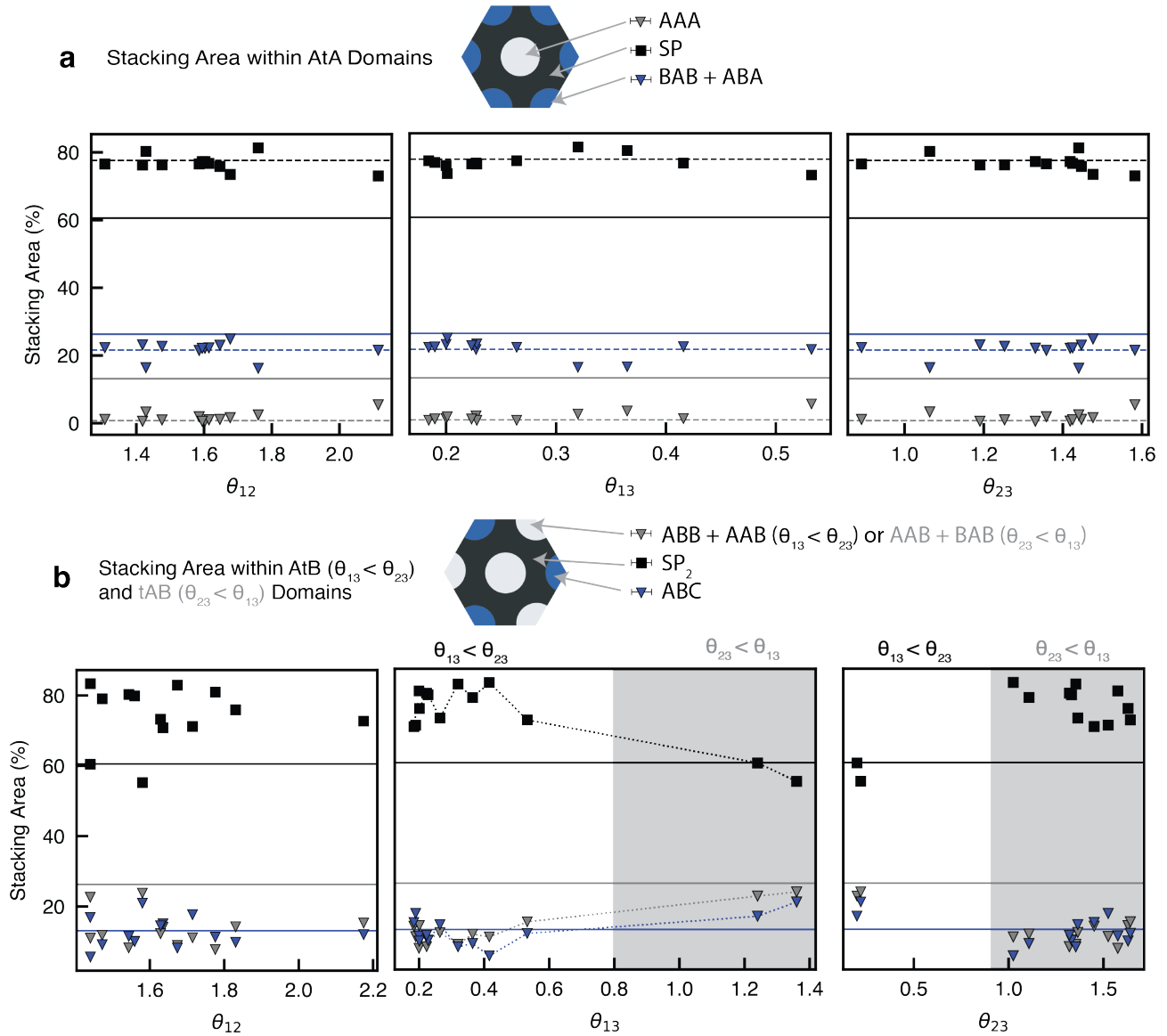
Supplemental Figure 7: Supplemental Atomic Stacking in TTLG trilayers (1/2).
(**a**) Maps of local atomic stacking from the larger moiré pattern only, corresponding to the local in-plane offset between between layers 1 and 3. (**b**) Local atomic stacking obtained from considering all three graphene layers. (**c**) Stacking order maps expected of rigid structures with the same twist angles, obtained from the equations in Supplemental section 8. All scale bars are 50 nm. Data-sets in panels 4-5 are shown in the main text.
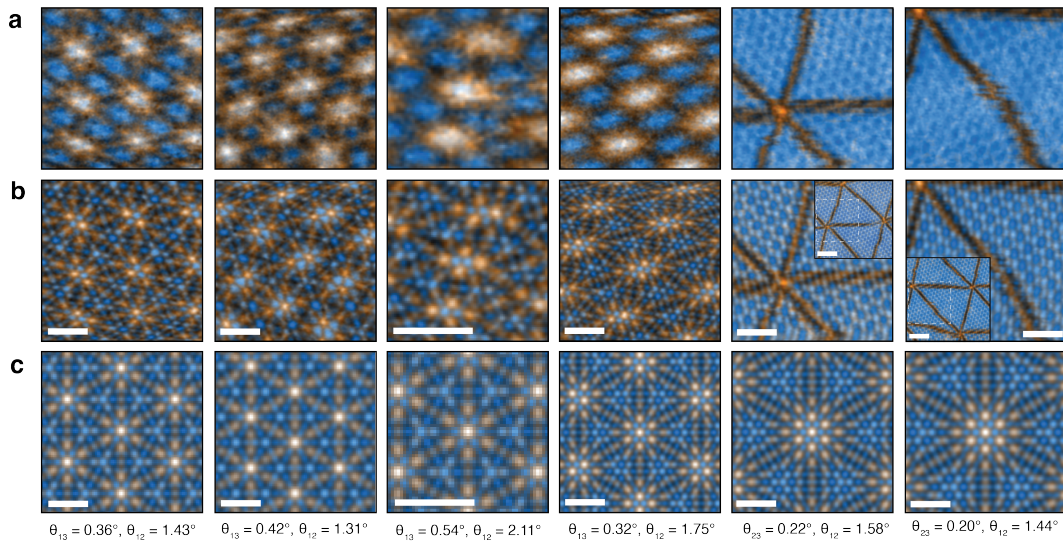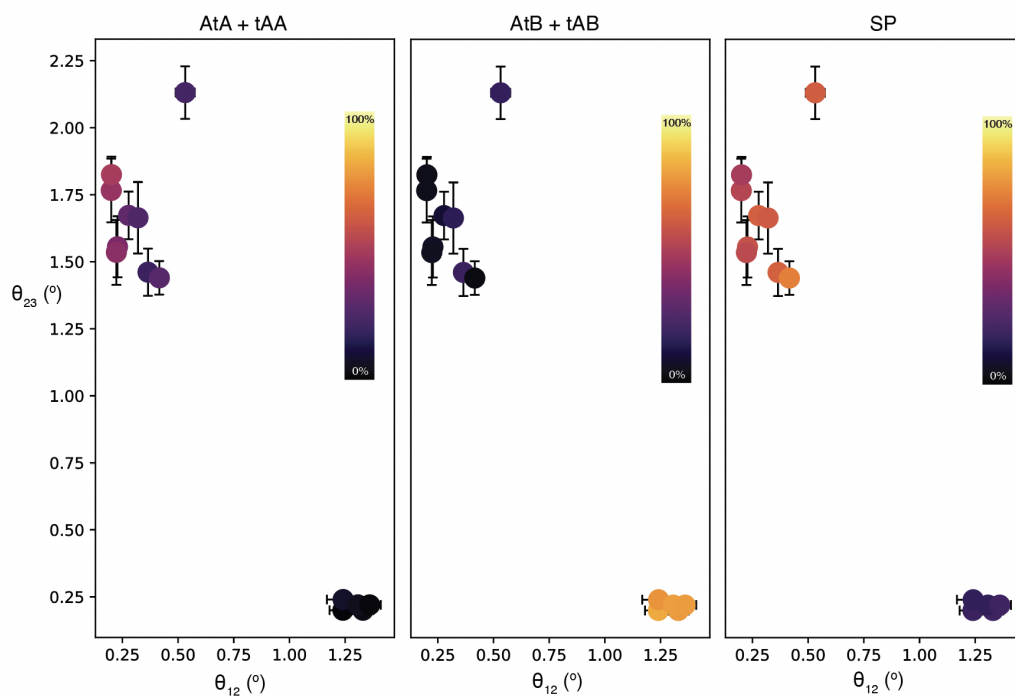

Supplemental Figure 8: Supplemental Atomic Stacking in TTLG trilayers (2/2).
(**a**) Maps of local atomic stacking from the larger moiré pattern only, corresponding to the local in-plane offset between between layers 1 and 3 for panels 1-4 and between layers 2 and 3 for panels 5-6. (**b**) Local atomic stacking obtained from considering all three graphene layers. (**c**) Stacking order maps expected of rigid structures with the same twist angles, obtained from equations presented in Supplemental section 8. All scale bars are 50 nm. Data-sets in panels 4-5 are shown in the main text.

Supplemental Figure 9: Atomic Stacking Trends in TTLG (**a**) Plot illustrating how the proportion of AAA (grey), ABA+BAB (blue), and SP stacking (black) depend on the three measured twist angles θ_{12}, θ_{13}, and θ_{23} change within AtA domains. Dashed lines represent the stacking portions measured for a pure AtA sample with $\theta_{13} \approx 0$ as shown in Supplemental Figure 6. (**b**) Plot illustrating how the proportion of ABB + AAB (grey), ABC (blue), and SP stacking (black) depend on the three measured twist angles change within AtB domains, and AAB + BAB (grey), ABC (blue), and SP stacking within tAB domains. Dotted lines are piece-wise linear fits to guide the eye. Solid lines are the stacking portions expected of rigid structures.

Supplemental Figure 10: Dependence of stacking order on both variable twist angles Plot illustrating how the proportion of AtA/tAA stacking, tAB/AtB stacking, and lower symmetry stackings change with each independent twist angle, illustrating that the overall AtA/tAA and tAB/AtB domain sizes are predominantly affected by the smallest twist angle. Error bar widths are standard deviations from 3-112 measurements for each data point.

where E_i and E_j are the unit cell vectors of layers i and j respectively, rotated by θ_{ij}. In the trilayer system, there is no simple linear transformation between real and configuration space. The relation between the displacement field defined in real space, $\boldsymbol{U}^{(i)}(\boldsymbol{r})$, and in configuration space, $\boldsymbol{u}^{(i)}(\boldsymbol{b})$, can be found by evaluating $\boldsymbol{u}^{(j)}(\boldsymbol{b})$ at the corresponding $\boldsymbol{b}^{i \to j}(\boldsymbol{r})$ and $\boldsymbol{b}^{i \to k}(\boldsymbol{r})$ with Eq. (5) to obtain

$$\boldsymbol{U}^{(i)}(\boldsymbol{r}) = \boldsymbol{u}^{(i)}(\boldsymbol{b}^{i \to j}(\boldsymbol{r}), \boldsymbol{b}^{i \to k}(\boldsymbol{r})), \tag{6}$$

where $j, k \neq i$ and $j < k$.

The relaxed energy has two contributions, intralayer and interlayer energies:

$$E^{\text{tot}}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)}) = E^{\text{intra}}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)}) + E^{\text{inter}}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)}), \tag{7}$$

where $\boldsymbol{u}^{(\ell)}$ is the relaxation displacement vector in layer ℓ. To obtain the relaxation pattern, we minimize the total energy with respect to the relaxation displacement vector.

We model the intralayer coupling based on linear elasticity theory:

$$\begin{aligned} E_{\text{intra}}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)}) = \sum_{\ell=1}^{3} \int \frac{1}{2}\Big[& G(\partial_x u_x^{(\ell)} + \partial_y u_y^{(\ell)})^2 \\ & + K((\partial_x u_x^{(\ell)} - \partial_y u_y^{(\ell)})^2 + (\partial_x u_y^{(\ell)} + \partial_y u_x^{(\ell)})^2)\Big] d\boldsymbol{b}, \end{aligned} \tag{8}$$

where G and K are shear and bulk moduli of monolayer graphene, which we take to be $G = 47352 \, \text{meV/unit cell}$, $K = 69518 \, \text{meV/unit cell}$ [4, 6]. Note that the gradient in Eq. (8) is with respect to the real space position $\boldsymbol{r}$.

The interlayer energy accounts for the energy cost of the layer misfit, which is described by generalized stacking fault energy (GSFE) [7, 8], obtained using first principles Density Functional Theory (DFT) with the Vienna Ab initio Simulation Package (VASP) [9–11]. GSFE is the ground state energy as a function of the local stacking with respect to the lowest energy stacking between a bilayer. For bilayer graphene, GSFE is maximized at the

AA stacking and minimized at the AB stacking. Letting $\boldsymbol{b} = (b_x, b_y)$ be the relative stacking between two layers, we define the following vector $\boldsymbol{v} = (v, w) \in [0, 2\pi]^2$:

$$
\begin{pmatrix} v \\ w \end{pmatrix} = \frac{2\pi}{a_0} \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{pmatrix} b_x \\ b_y \end{pmatrix},
\tag{9}
$$

where $a_0 = 2.4595\,\text{Å}$ is the graphene lattice constant. We parameterize the GSFE as follows,

$$
\begin{aligned}
V_{j\pm}^{\text{GSFE}} = c_0 &+ c_1(\cos v + \cos w + \cos(v + w)) \\
&+ c_2(\cos(v + 2w) + \cos(v - w) + \cos(2v + w)) \\
&+ c_3(\cos(2v) + \cos(2w) + \cos(2v + 2w)),
\end{aligned}
\tag{10}
$$

where we take $c_0 = 6.832\,\text{meV/cell}$, $c_1 = 4.064\,\text{meV/cell}$, $c_2 = -0.374\,\text{meV/cell}$, $c_3 = -0.0095\,\text{meV/cell}$ [4, 6]. The van der Waals force is implemented through the vdW-DFT method using the SCAN+rVV10 functional [12]. Note that we amplify the GSFE by a factor of 10. Physically, amplifying the GSFE enhances the strength of relaxation. It has been shown that the energy difference between AA and AB stackings can vary by a factor of 4 depending on the van der Waals functionals used [8]. In terms of $V_{\ell\pm}^{\text{GSFE}}$, the total interlayer energy can be expressed as follows:

$$
\begin{aligned}
E^{\text{inter}} = \frac{1}{2} \int V_{1+}^{\text{GSFE}}(\boldsymbol{B}^{1\to2}) \, \mathrm{d}\boldsymbol{b} + \frac{1}{2} \int \left[V_{2-}^{\text{GSFE}}(\boldsymbol{B}^{2\to1}) + V_{2+}^{\text{GSFE}}(\boldsymbol{B}^{2\to3}) \right] \mathrm{d}\boldsymbol{b} \\
+ \frac{1}{2} \int V_{3-}^{\text{GSFE}}(\boldsymbol{B}^{3\to2}) \, \mathrm{d}\boldsymbol{b},
\end{aligned}
$$

where $\boldsymbol{B}^{i\to j} = \boldsymbol{b}^{i\to j} + \boldsymbol{u}^{(j)} - \boldsymbol{u}^{(i)}$ is the relaxation modified local shift vector. Note that we neglect the interlayer coupling between layers 1 and 3. The total energy is obtained by summing over uniformly sampled configuration space. In this work, we discretize the four-dimensional configuration space by $54 \times 54 \times 54 \times 54$.

Finally, we calculate the local twist angle between layers i and j as redistributed by relaxation

as follows,

$$\theta_{ij,\text{local}}(\boldsymbol{r}) = \left| \theta_{ij} + \sin^{-1}\left(\frac{\boldsymbol{\nabla} \times \boldsymbol{U}^{(j)}(\boldsymbol{r})}{2} \right) - \sin^{-1}\left(\frac{\boldsymbol{\nabla} \times \boldsymbol{U}^{(i)}(\boldsymbol{r})}{2} \right) \right|. \tag{11}$$

As regions of local AtA, tAA, and or AAt stacking appear as domains where the local twist angle $|\theta_{ij,\text{local}}(\boldsymbol{r})| \approx 0$, we were able to use a threshold of 0.25 degrees to obtain the stacking statistics shown in Figure 4 of the main text. These stacking order trends and the local twist angle maps (Supplemental Figure 10) illustrate a pattern of atomic relaxation consistent with what we observed in the experimental data, with large AtA regions apparent when $\theta_{13} << \theta_{23}$ and large tAB regions apparent when $\theta_{13} >> \theta_{23}$.
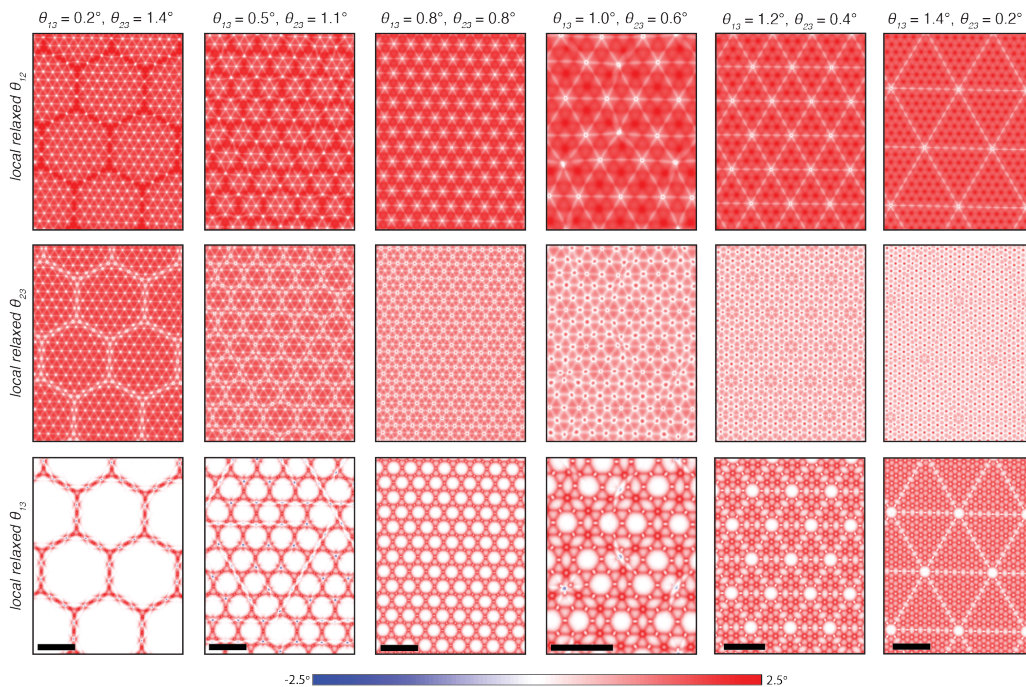
Supplemental Figure 11: Theoretical Predictions of Local Twist Angles Predictions for $\theta_{ij,\text{local}}(\boldsymbol{r})$ obtained from the procedure described in Supplemental section 12 for θ_{12} (top), θ_{23} (middle), and θ_{23} (bottom) for a series of twist angles. Twist angles are defined such that θ_{12} and θ_{23} are of opposite sign, and θ_{12} is 1.6 degrees in all twist angle maps shown.

10   Comparison to Scanning Probe Measurements

The nature of atomic reconstruction we measure is markedly different than that proposed in previous work on samples with comparable experimental conditions, wherein it was suggested that slightly misaligned MATTLG samples (θ_{13} of $\approx 0.25°$) relax to almost exclusively AtA regions, with the AtB and SP regions stretched into thin domain boundaries and/or topological defects that contribute insignificantly to the STM measurements [13] while we measure appreciable AtB regions down to $\theta_{13} = 0.20°$. We also see domains of either AtA and AtB (when $\theta_{13} \ll \theta_{23}$) or tAB and tAA (when $\theta_{23} \ll \theta_{13}$) containing distinct distributions of twist angles associated with the small-moiré pattern. We note that these domains appear similar to the clusters of stacking features, termed plaquettes (P) and twistons (T), presented previously as consisting of AAA domains [13]. One possible scenario that would yield qualitative agreement between our results, STM, and theory is that the plaquettes and twistons represent domains of local AtB (P) and AtA (T) or tAB (P) and tAA (T) symmetry. This would coincide with the bright spots in STM measurements (large local density of states, DOS) not all corresponding to AAA-type stacking, instead reflecting a richer distribution of stacking orientations. It is possible that the sample presented in [13] is structurally similar to the samples we present in Fig. 3 panels 2 and 3. In this case, the DOS maxima measured by STM would reflect both AAA and AAB symmetry (in the AtA and AtB regions, respectively) as flatbands are expected and observed in AAB regions as well [14]. In AtB regions, a lower local DOS of adjacent ABB regions would be expected from an STM measurement since the STM tip would be probing the Bernal bilayer side, as reported previously [14]. This effect may then result in the visualization, by STM, of a triangular lattice of the most prominent AAB symmetry. We also note that such logic would hold true if the sample presented in [13] was instead structurally similar to the samples we present in Fig. 3 panels 4, as we would still expect DOS maxima reflecting both AAA and (predominantly) AAB symmetry. Although we note this prior work included a careful consideration to exclude local DOS effects to attribute micrographs to sample topology. It is plausible given the presence of thick hBN on only one side that there is a large difference in height between these regions,

yielding a similar effect. However, we note that the discussion presented here is speculative and discrepancies between our two measurements may be influenced by subtle differences in annealing and encapsulation or hereto unknown nuances in our experimental conditions. We primarily hope to emphasize that interferomic transmission electron microscopy may serve as a helpful tool to complement scanning probe measurements and other techniques that measure sample morphology via indirect signatures such as height variation and are difficult to analyze without prior information regarding sample morphology.

References

1. Kazmierczak, N. P. *et al.* Strain fields in twisted bilayer graphene. *Nature materials* **20,** 956–963 (2021).

2. Van Winkle, M. *et al.* Quantitative Imaging of Intrinsic and Extrinsic Strain in Transition Metal Dichalcogenide Moir\'e Bilayers. *arXiv preprint arXiv:2212.07006* (2022).

3. Madsen, J. & Susi, T. The abTEM code: transmission electron microscopy from first principles. *Open Research Europe* **1,** 13015 (24 2021).

4. Zhu, Z., Carr, S., Massatt, D., Luskin, M. & Kaxiras, E. Twisted trilayer graphene: A precisely tunable platform for correlated electrons. *Physical review letters* **125,** 116404 (2020).

5. Cazeaux, P., Luskin, M. & Massatt, D. Energy Minimization of Two Dimensional Incommensurate Heterostructures. *Archive for Rational Mechanics and Analysis* **235,** 1289–1325. arXiv: `1806.10395 [physics.comp-ph]` (Aug. 2019).

6. Carr, S. *et al.* Relaxation and domain formation in incommensurate two-dimensional heterostructures. *Phys. Rev. B* **98,** 224102. `https://link.aps.org/doi/10.1103/PhysRevB.98.224102` (22 2018).

7. Kaxiras, E. & Duesbery, M. S. Free energies of generalized stacking faults in Si and implications for the brittle-ductile transition. *Phys. Rev. Lett.* **70,** 3752–3755. `https://link.aps.org/doi/10.1103/PhysRevLett.70.3752` (24 June 1993).

8. Zhou, S., Han, J., Dai, S., Sun, J. & Srolovitz, D. J. Van der Waals bilayer energetics: Generalized stacking-fault energy of graphene, boron nitride, and graphene/boron nitride bilayers. *Phys. Rev. B* **92,** 155438 (15 2015).

9. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47,** 558–561 (1 1993).

10. Kresse, G. & Furthmüller, J. Efficiency of ab initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* **6,** 15–50. ISSN: 0927-0256 (1996).

11. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54,** 11169–11186 (16 1996).

12. Peng, H., Yang, Z.-H., Perdew, J. P. & Sun, J. Versatile van der Waals Density Functional Based on a Meta-Generalized Gradient Approximation. *Phys. Rev. X* **6,** 041005 (4 2016).

13. Turkel, S. *et al.* Orderly disorder in magic-angle twisted trilayer graphene. *Science* **376,** 193–199 (2022).

14. Tong, L.-H. *et al.* Spectroscopic visualization of flat bands in magic-angle twisted monolayer-bilayer graphene: coexistence of localization and delocalization. *Physical Review Letters* **128,** 126401 (2022).