Modeling mechanical relaxation in incommensurate trilayer van der Waals heterostructures

Ziyan Zhu [1], Paul Cazeaux,[2] Mitchell Luskin [3], and Efthimios Kaxiras[1,4]

[1]*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
[2]*Department of Mathematics, University of Kansas, Lawrence, Kansas 66045, USA*
[3]*School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455, USA*
[4]*John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*

The incommensurate stacking of multilayered two-dimensional materials is a challenging problem from a theoretical perspective and an intriguing avenue for manipulating their physical properties. Here we present a multiscale model to obtain the mechanical relaxation pattern of twisted trilayer van der Waals (vdW) heterostructures with two independent twist angles, a generally incommensurate system without a supercell description. We adopt the configuration space as a natural description of such incommensurate layered materials, based on the local environment of atomic positions, bypassing the need for commensurate approximations. To obtain the relaxation pattern, we perform energy minimization with respect to the relaxation displacement vectors. We use a continuum model in combination with the generalized stacking fault energy to describe the interlayer coupling, obtained from first-principles calculations based on density functional theory. We show that the relaxation patterns of twisted trilayer graphene and WSe_2 are "moiré of moiré," as a result of the incommensurate coupling two bilayer moiré patterns. We also show that, in contrast to the symmetry-preserving in-plane relaxation in twisted bilayers, trilayer relaxation can break the two fold rotational symmetry about the xy plane when the two twist angles are equal.

I. INTRODUCTION

Assemblies of multilayers of two-dimensional (2D) materials, referred to as van der Waals (vdW) heterostructures, have become a favorite platform for exploring strongly correlated states, because they offer a large parameter space and great tunability. Superconductivity and strongly-correlated states have been discovered in these heterostructures with one small twist angle (e.g., twisted bilayer graphene, twisted double-bilayer graphene) [1–8] and systems with a small misalignment (e.g., hBN + graphene) [9,10]. Due to the twist angle or the small lattice misalignment, the system forms moiré patterns and the periodicity becomes much larger than the original unit cell size. To understand the electronic and mechanical properties of the multilayered structures, we view them as a series of conventional crystals with a weak perturbative interaction between layers [11].

Crystal relaxation in vdW heterostructures has been studied in the continuum limit by energy minimization [12–18]. Relaxation in incommensurate stacked 2D materials has been known to form domain walls separating large commensurate areas [19]. For example, in twisted bilayer graphene, relaxation enlarges the AB/BA stacking regions (the equilibrium stacking) and forms a thin domain line in between adjacent AB/BA stacking regions, while the AA stacking regions (high energy stacking order) shrink to localized spots. The domain formation in twisted bilayer systems has been observed experimentally [20–25]. The relaxation pattern has a dependence on the twist angle: The magnitude of relaxation displacement

vectors decreases as the twist angle increases. At a twist angle $\theta > \theta^*$ ($\theta^* \sim 1°$ in twisted bilayer graphene), the relaxation effects are weak. As the twist angle becomes smaller, the domain wall thickness stays roughly constant and only the AB/BA stacking regions get enlarged.

A natural extension is to add another layer, which introduces a new degree of freedom—a second twist angle—to tune the system. We emphasize that the system we discuss here is different than multilayered systems such as double-bilayer graphene, in which there is only one independent parameter or one twist angle despite the presence of four layers. The experimental technique to realize twisted trilayer graphene is readily available and unconventional correlated states have been observed [26]. There have been preliminary studies of electronic states of twisted trilayer graphene using $k \cdot p$ perturbation theory [26–28]. In these models, mechanical relaxation has not been taken into consideration. However, atomic relaxation will modify the band structure by opening up a single-particle gap in twisted bilayer graphene [14,29,30] and thus is an important effect in obtaining realistic description of electronic behaviors.

The extension from twisted bilayers to twisted trilayers is challenging for the following reasons. (1) The system size becomes much larger. In twisted bilayer systems the relevant length scale is on the order of $1/\theta$. In twisted trilayers, assuming the two twist angles are identical, the length scale becomes $1/\theta^2$, where θ is the relevant twist angle, typically a small quantity for systems of interest. (2) Because of the independent second twist angle, a supercell description is no
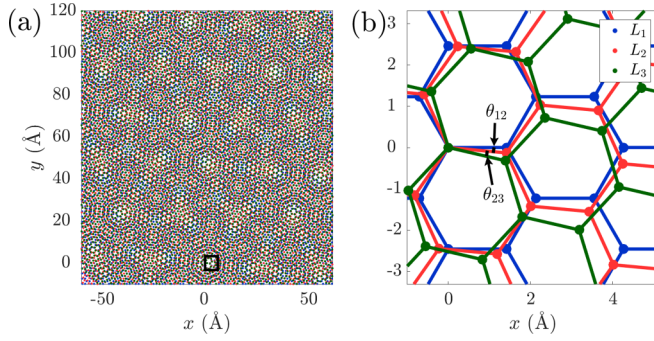
FIG. 1. (a) An example of a twisted trilayer honeycomb lattice in real space with $\theta_{12} = 5.3°$ and $\theta_{23} = 7.7°$. (b) Magnified view at the black box marked in (a). The twist angle between L_1 and L_2, θ_{12}, and the twist angle between L_2 and L_3, θ_{23}, are marked by black arrows.

longer valid, except for specific cases. The trilayer system is truly incommensurate, while there always exists a moiré supercell in the continuum limit or the small angle limit of twisted bilayers.

To address the challenge of the large system size and the lack of periodicity, we introduce a multiscale method based on the concept of configuration space to minimize the energy directly in the large system limit without having to use a supercell approximation. Specifically, we perform energy minimization in the continuum approximation over a collection of local atomic environments with respect to the other layers, which we call the configuration space [31–34]. This approach has recently been applied to describing the crystal relaxation in stacked bilayer materials [17,34].

The rest of the paper is organized as follows: In Sec. II, we introduce the configuration space to describe the trilayer system and formulate the energy minimization problem in configuration space using a continuum approximation. In Sec. III, we present the results for twisted trilayer graphene and WSe$_2$, as representative heterostructures of a 2D semimetal and a 2D semiconductor. Finally, we summarize the results and discuss potential applications in Sec. IV.

II. CONTINUUM APPROXIMATION AND ENERGY MINIMIZATION IN CONFIGURATION SPACE

A. Setup and notation

We first define the Bravais lattice basis vectors of a monolayer, a_1 and a_2, respectively, in a 2×2 matrix form as the columns of the following matrix:

$$A = a_0 \begin{pmatrix} \sqrt{3}/2 & \sqrt{3}/2 \\ -1/2 & 1/2 \end{pmatrix} = (a_1 \quad a_2), \qquad (1)$$

where a_0 is the lattice constant. We denote the unit cell of layer i by $\Gamma^{(i)}$. We consider the twisted trilayer systems with two independent twist angles; an example of such a system is given in Fig. 1(a). Suppose the second layer L_2 is fixed at the origin. The first layer (L_1) is twisted counterclockwise by θ_{12} with respect to L_2, and L_3 is twisted clockwise by θ_{23} with respect to L_2 [see Fig. 1(b)]. The primitive Bravais lattice basis vectors in the matrix form of L_i is defined as A_i. With $A_2 = A$,
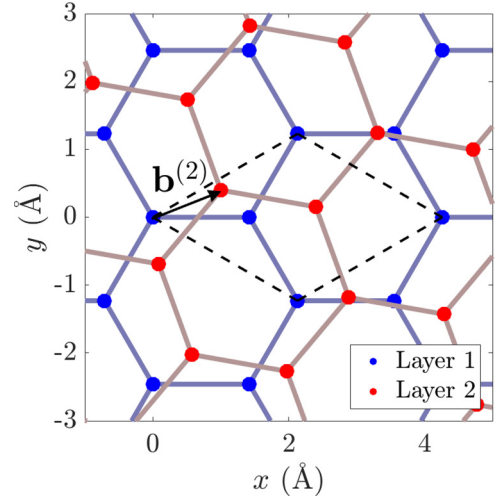

FIG. 2. The definition of the local shift or the disregistry $b^{(j)}$. Here, we take $j = 2$ as an example, defining $b^{(2)}$ between L_1 and L_2.

A_1 and A_3 are given as follows,

$$A_1 = \begin{pmatrix} \cos\theta_{12} & -\sin\theta_{12} \\ \sin\theta_{12} & \cos\theta_{12} \end{pmatrix} A = (a_1^{(1)} \quad a_2^{(1)}),$$

$$A_3 = \begin{pmatrix} \cos\theta_{23} & \sin\theta_{23} \\ -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} A = (a_1^{(3)} \quad a_2^{(3)}). \qquad (2)$$

The primitive reciprocal lattice vectors are given by the columns of $2\pi A_i^{-T}$.

B. Relaxation in configuration space

The issue with applying the real space continuum model to obtain the relaxation pattern in trilayer systems is the lack of periodicity. While bilayer systems always have a moiré supercell [17,34], trilayer systems are generally incommensurate and thus lack a periodic supercell even in the continuum limit. Therefore, a more general description beyond the supercell approximation is required. Here, we introduce the configuration space to describe the local environment of every position in the continuum and parametrize the trilayer system in configuration space [31–33] based on the formalism introduced by Cazeaux *et al.* [34]. We also reformulate the energy minimization problem in configuration space.

The local configuration or the environment for an arbitrary position r in layer L_i, which may be a lattice position on layer L_i, can be uniquely determined by three relative shift vectors or disregistries $b^{(j)}$ describing the relative position from this position in L_i with respect to the lattice of layer L_j (see Fig. 2 for an example), which is defined as the lattice position with respect to r modulo the unit cell $\Gamma^{(j)}$. More explicitly,

$$b^{(j)}(r) = R^{(j)} - r \in \Gamma^{(j)}, \qquad (3)$$

where $R^{(j)}$ is any lattice site in layer L_j. The shift vectors $b^{(j)}$ take values in the periodic unit cell $\Gamma^{(j)}$ of layer L_j by the translational invariance of the lattice. Figure 3 illustrates this mapping between real space and the configuration space of the three relative shift vectors for r sampled at lattice points of the middle layer L_2 of a triangular lattice trilayer for a representative atomic structure. The collection of these shift
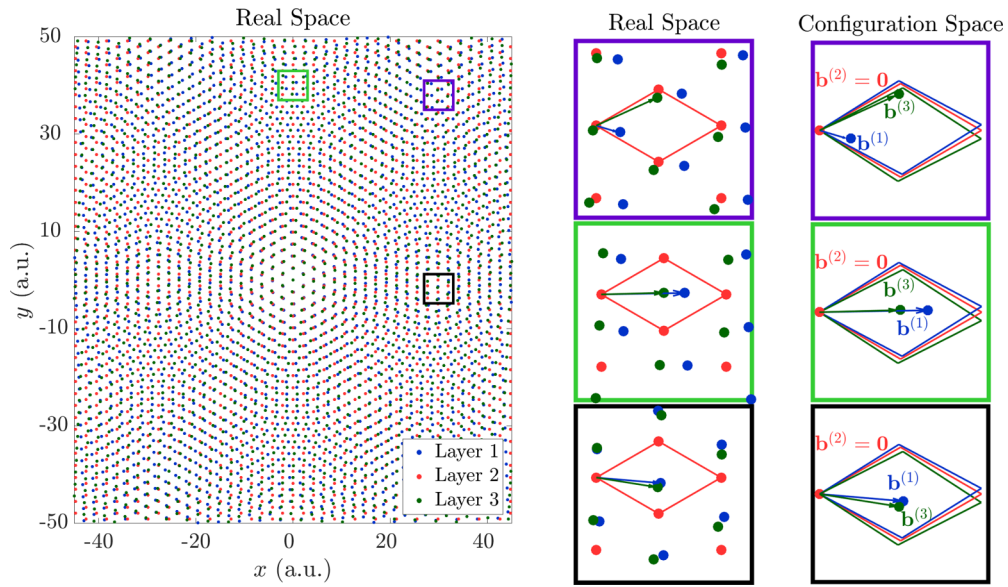
FIG. 3. Mapping between real and configuration space. Left: Real space atomic positions of a trilayer triangular lattice. The first layer (blue) is twisted by $2°$ with respect to the second layer (red), and the third layer (green) is twisted by $3°$ with respect to the second. Middle: magnified view of the three boxed area on the left. Red parallelograms are the unit cell of the layer 2. Arrows are pointing from the atom of interest (the red atom on the left corner) to the neighboring atoms on the other two layers. Right: the corresponding configuration space to the atom of interest. Blue, red, and green parallelograms are the unit cells of L_i for $i = 1, 2, 3$ respectively.

vectors forms a six-dimensional configuration space:

$$\Omega = \{\omega \mid \omega = (\boldsymbol{b}^{(1)}, \boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)})\}. \quad (4)$$

The shifts $\boldsymbol{b}^{(j)}(\boldsymbol{r})$ vary quickly on the scale of the lattice size and thus do not correspond to the intuitive notion of disregistry which varies on the scale of the moiré pattern. Instead, we construct functions $\boldsymbol{b}^{i \to j}(\omega)$ measuring smoothly the disregistry between layer L_i and L_j as a function of the local configuration ω, by using bilinear interpolation between lattice points [17,34]

$$\boldsymbol{b}^{i \to j}(\omega) = \boldsymbol{b}^{(j)} - A_j A_i^{-1} \boldsymbol{b}^{(i)} \in \Gamma^{(j)}. \quad (5)$$

This expression is periodic as a function of $\boldsymbol{b}^{(i)}$ and $\boldsymbol{b}^{(j)}$, respectively, in the unit cells of L_i and L_j, and the disregistry $\boldsymbol{b}^{i \to j}(\omega(\boldsymbol{r}))$ varies slowly on the atomic scale.

Figure 4 demonstrates this construction. We first choose a reference point $\boldsymbol{r}$; here, we use a lattice site of L_1 as an example. At the reference point, the configuration is $\omega = (\boldsymbol{b}^{(1)}, \boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)})$. At each one of the four lattice sites in layer L_i (respectively, L_j) denoted by $\boldsymbol{R}_{mn}^{(i)}$ (respectively, $\boldsymbol{R}_{mn}^{(j)}$) neighboring the reference point $\boldsymbol{r}$, the shift vector $\boldsymbol{b}^{(j)}(\boldsymbol{R}_{mn}^{(i)})$ coincides with the local disregistry of layer L_j with respect to layer L_i:

$$\boldsymbol{b}_{mn}^{i \to j} = \boldsymbol{R}_{mn}^{(j)} - \boldsymbol{R}_{mn}^{(i)} \equiv \boldsymbol{b}^{(j)}(\boldsymbol{R}_{mn}^{(i)}), \quad (6)$$

where we use the definition in Eq. (3). Because the lattices are almost aligned, these four disregistries only differ slightly from each other. We can then define $\boldsymbol{b}^{i \to j}(\omega)$ by bilinear interpolation of the four relative shifts $\boldsymbol{b}_{mn}^{2 \to 3}$, yielding Eq. (5) [34]. Note that the interpolated disregistry given in Eq. (5) is consistent with Eq. (6) at any lattice site of L_i.

In practice, we need to only consider reference points that coincide with lattice points in layer L_i, $\boldsymbol{r} = \boldsymbol{R}^{(i)}$, in order to

parametrize atomic displacements. This leads us to introduce the reduced, four-dimensional atomic configuration spaces:

$$\Omega_i = \{\omega \mid \omega = (\boldsymbol{b}^{(1)}, \boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)}) \mid \boldsymbol{b}^{(i)} = \boldsymbol{0}\}. \quad (7)$$

The relaxation displacement $\boldsymbol{u}^{(i)}(\omega)$ at each lattice point is a function of its local environment of shifts $\omega \in \Omega_i$. The above construction allows us to define an interpolated displacement $\widehat{\boldsymbol{u}}^{(j)}(\omega)$ of another layer L_j from the viewpoint of a lattice site in L_i by

$$\widehat{\boldsymbol{u}}^{(j)}(\omega) = \boldsymbol{u}^{(j)}(\boldsymbol{b}^{j \to 1}(\omega), \boldsymbol{b}^{j \to 2}(\omega), \boldsymbol{b}^{j \to 3}(\omega)), \quad (8)$$
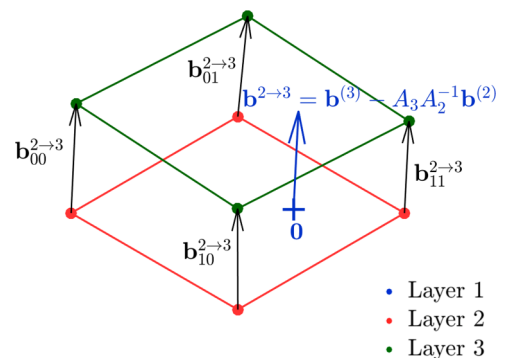

FIG. 4. A demonstration of the bilinear interpolation of the local shift vector $\boldsymbol{b}^{2 \to 3}$. This corresponds to the same real space position in the green (the second) box in Fig. 3 but with a different origin. Here, the origin is chosen to be a lattice site of L_1 with $\boldsymbol{b}^{(1)} = \boldsymbol{0}$, which is marked by a blue cross in the figure, while in Fig. 3, the origin is the red atom on the left corner. $\boldsymbol{b}_{mn}^{2 \to 3}$ are the relative shifts between L_2 and L_3 at the four nearest lattice sites to the origin.

where we simply use the definition given in Eq. (5). For example, the interpolated displacement $\widehat{\boldsymbol{u}}^{(2)}(\omega)$ for $\omega \in \Omega_1$ is

$$\widehat{\boldsymbol{u}}^{(2)}(\omega) = \boldsymbol{u}^{(2)}(\boldsymbol{b}^{2\to1}(\omega), \boldsymbol{b}^{2\to3}(\omega)) \qquad (9)$$

$$= \boldsymbol{u}^{(2)}\big(-A_1 A_2^{-1}\boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)} - A_3 A_2^{-1}\boldsymbol{b}^{(2)}\big),$$

where we omit $\boldsymbol{b}^{j\to j}(\omega) = 0$ from the argument of $\boldsymbol{u}^{(j)}$. We can then describe the relaxation-modulated local shift $\boldsymbol{B}^{i\to j}(\omega)$ of L_i with respect to L_j by

$$\boldsymbol{B}^{i\to j}(\omega) = \boldsymbol{b}^{i\to j}(\omega) + \widehat{\boldsymbol{u}}^{(j)}(\omega) - \boldsymbol{u}^{(i)}(\omega) \text{ for } \omega \in \Omega_i, \quad (10)$$

where $\widehat{\boldsymbol{u}}^{(j)}(\omega)$ is the interpolated displacement of L_j.

The relaxed site energy has two contributions, interlayer and intralayer site energies, as a function of the relaxation displacement vectors, $\boldsymbol{u} = (\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)})$. For a given layer L_i,

$$\Phi_{i,\omega}(\boldsymbol{u}) = \Phi_{i,\omega}^{\text{intra}}(\boldsymbol{u}) + \Phi_{i,\omega}^{\text{inter}}(\boldsymbol{u}). \qquad (11)$$

By the uniform sampling of the continuum of local configurations due to the ergodicity of incommensurate lattices [31,33], we can derive the total energy by integrating the site energies over all configurations [34]

$$E^{\text{tot}}(\boldsymbol{u}) = E^{\text{intra}}(\boldsymbol{u}) + E^{\text{inter}}(\boldsymbol{u}) = \sum_{i=1}^{3} \int_{\Omega_i} \Phi_{i,\omega}(\boldsymbol{u}) \, d\omega$$

$$= \sum_{i=1}^{3} \int_{\Omega_i} \Phi_{i,\omega}^{\text{intra}}(\boldsymbol{u}) \, d\omega + \sum_{i=1}^{3} \int_{\Omega_i} \Phi_{i,\omega}^{\text{inter}}(\boldsymbol{u}) \, d\omega. \qquad (12)$$

We then minimize $E^{\text{tot}}(\boldsymbol{u})$ to obtain the relaxation vector in configuration space $\boldsymbol{u}$. We will provide the form of E^{intra} and E^{inter} in Secs. II C and II D, respectively.

C. Intralayer energy

We proceed to describe the intralayer energy contribution in configuration space. From linear elasticity theory, the intralayer energy due to lattice relaxation depends on the real space gradient of displacement vectors,

$$E^{\text{intra}}(\boldsymbol{u}) = \sum_{i=1}^{3} \int_{\Omega_i} \Phi_{i,\omega}^{\text{intra}}(\boldsymbol{u}) d\omega, \qquad (13)$$

where $\omega = (\boldsymbol{b}^{(j)}, \boldsymbol{b}^{(k)}) \in \Omega_i$, and

$$\Phi_{i,\omega}^{\text{intra}}(\boldsymbol{u}) = \frac{1}{2} \nabla_{\boldsymbol{r}}\widehat{\boldsymbol{u}}^{(i)}(\omega) : C : \nabla_{\boldsymbol{r}}\widehat{\boldsymbol{u}}^{(i)}(\omega). \qquad (14)$$

In Eq. (14), C is the linear elasticity tensor. For in-plane deformation, C is a rank 4 tensor with its components defined as follows:

$$C_{11ij} = \begin{pmatrix} K+G & 0 \\ 0 & G \end{pmatrix} \quad C_{12ij} = \begin{pmatrix} 0 & K-G \\ G & 0 \end{pmatrix} \qquad (15)$$

$$C_{21ij} = \begin{pmatrix} 0 & G \\ K-G & 0 \end{pmatrix} \quad C_{22ij} = \begin{pmatrix} G & 0 \\ 0 & K+G \end{pmatrix},$$

where G and K are the shear and bulk modulus of a monolayer (anisotropic models can be used for anisotropic 2D materials such as black phosphorous). These values can be obtained from first-principles total-energy calculations based on density functional theory (DFT) by isotropically straining and

compressing the monolayer and performing a quadratic fit of the total energy as a function of the applied strain or shear. In order to incorporate out-of-plane relaxation, the elastic tensor would be a $3 \times 3 \times 3 \times 3$ tensor [13], which is beyond the scope of this work. We can compute the real space gradient of the interpolated displacement $\widehat{\boldsymbol{u}}^{(i)}$ at an arbitrary configuration $\omega \in \Omega_i$ by the chain rule applied to Eqs. (3), (5), and (8):

$$\nabla_{\boldsymbol{r}}\widehat{\boldsymbol{u}}^{(i)}(\omega) = \sum_{i\neq j} \nabla_{\boldsymbol{b}^{(j)}}\boldsymbol{u}^{(i)} \cdot \big(A_j A_i^{-1} - I\big). \qquad (16)$$

Defining the following matrix,

$$M_{ji} = A_j A_i^{-1} - I, \qquad (17)$$

we then have the Fourier representation of $\widehat{\boldsymbol{u}}^{(i)}$:

$$\nabla_{\boldsymbol{r}}\widehat{\boldsymbol{u}}^{(i)}(\boldsymbol{b}^{(j)}, \boldsymbol{b}^{(k)}) = i \sum_{\boldsymbol{G}^{(j)}} \sum_{\boldsymbol{G}^{(k)}} e^{i\boldsymbol{G}^{(j)}\cdot\boldsymbol{b}^{(j)}} e^{i\boldsymbol{G}^{(k)}\cdot\boldsymbol{b}^{(k)}}$$

$$\times \tilde{\boldsymbol{u}}^{(i)}(\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)}) \otimes \big[M_{ji}^T \boldsymbol{G}^{(j)} + M_{ki}^T \boldsymbol{G}^{(k)}\big], \qquad (18)$$

where $\boldsymbol{G}^{(i)}$ are reciprocal lattice vectors for L_i and the Fourier coefficients $\tilde{\boldsymbol{u}}^{(i)}$ are defined according to

$$\boldsymbol{u}^{(i)}(\boldsymbol{b}^{(j)}, \boldsymbol{b}^{(k)}) = \sum_{\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)}} \tilde{\boldsymbol{u}}^{(i)}(\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)})$$

$$\times e^{i(\boldsymbol{G}^{(j)}\cdot\boldsymbol{b}^{(j)} + \boldsymbol{G}^{(k)}\cdot\boldsymbol{b}^{(k)})}, \qquad (19)$$

for $j, k \neq i$ and $j < k$. The contribution of the mode $(\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)})$ to the variational elastic energy is

$$\mathcal{E}_i(\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)})$$

$$= \frac{1}{2}\big[\tilde{\boldsymbol{u}}^{(i)}(\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)}) \otimes \big(M_{ji}^T \boldsymbol{G}^{(j)} + M_{jk}^T \boldsymbol{G}^{(k)}\big)\big] : C :$$

$$\times \big[\tilde{\boldsymbol{u}}^{(i)}(\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)}) \otimes \big(M_{ji}^T \boldsymbol{G}^{(j)} + M_{jk}^T \boldsymbol{G}^{(k)}\big)\big]. \qquad (20)$$

In terms of the $\mathcal{E}_i$'s, the total intralayer energy of the trilayer system is

$$E^{\text{intra}}(\boldsymbol{u}) = \sum_{i=1}^{3} \sum_{\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)}} \mathcal{E}_i(\boldsymbol{G}^{(j)}, \boldsymbol{G}^{(k)}). \qquad (21)$$

Each layer in the trilayer system has a four-dimensional configuration space Ω_i, but the in-plane interactions only span a two-dimensional submanifold. The Euler-Lagrange PDE for the intralayer strain energy is thus not properly elliptic, but it is nonetheless nonsingular since $\mathcal{E}_1(\boldsymbol{G}^{(2)}, \boldsymbol{G}^{(3)}) = 0$ if and only if $\tilde{\boldsymbol{u}}_1(\boldsymbol{G}^{(2)}, \boldsymbol{G}^{(3)}) = 0$ or $\boldsymbol{G}^{(2)} = \boldsymbol{G}^{(3)} = 0$, and similarly for $\mathcal{E}_2(\boldsymbol{G}^{(1)}, \boldsymbol{G}^{(3)})$ and $\mathcal{E}_3(\boldsymbol{G}^{(1)}, \boldsymbol{G}^{(2)})$. To see this, we take $i = 1$ in Eq. (21) and $\mathcal{E}_1(\boldsymbol{G}^{(2)}, \boldsymbol{G}^{(3)}) \propto |M_{21}^T \boldsymbol{G}^{(2)} + M_{31}^T \boldsymbol{G}^{(3)}|^2$. We note that

$$M_{21}^T \boldsymbol{G}^{(2)} + M_{31}^T \boldsymbol{G}^{(3)} = \boldsymbol{G}'^{(1)} - \boldsymbol{G}^{(2)} - \boldsymbol{G}^{(3)}, \qquad (22)$$

where $\boldsymbol{G}'^{(1)} = A_1^{-T} A_2^T \boldsymbol{G}^{(2)} + A_1^{-T} A_3^T \boldsymbol{G}^{(3)}$ is in the reciprocal lattice of L_1, and $\boldsymbol{G}'^{(1)} - \boldsymbol{G}^{(2)} - \boldsymbol{G}^{(3)} = 0$ if and only if $\boldsymbol{G}'^{(1)} = \boldsymbol{G}^{(2)} = \boldsymbol{G}^{(3)} = 0$ by the incommensurability of the trilayer [34]. Observe that $|\boldsymbol{G}'^{(1)} - \boldsymbol{G}^{(2)} - \boldsymbol{G}^{(3)}|^2$ and hence $\mathcal{E}_1(\boldsymbol{G}^{(2)}, \boldsymbol{G}^{(3)})$ can be small even though $\boldsymbol{G}^{(2)}$ and $\boldsymbol{G}^{(3)}$ are large, which is contrary to the ellipticity condition that $\mathcal{E}_1(\boldsymbol{G}^{(2)}, \boldsymbol{G}^{(3)})$ is a uniformly positive definite quadratic form

TABLE I. Equilibrium lattice constant of the primitive unit cell a_0 (in Å), unit cell area $\mathcal{A}$ (in Å^2), elastic constants K and G, and GSFE Fourier components for graphene, parallel, and antiparallel WSe$_2$. Units of elastic constants and GSFE components are in meV per unit cell.

System	a_0 (Å)	$\mathcal{A}$ (Å^2)	K	G	c_0	c_1	c_2	c_3	c_4	c_5
Graphene	2.47	5.28	69518	47352	6.832	4.064	−0.374	−0.095	0.000	0.000
Parallel WSe$_2$	3.28	9.32	47123	31153	27.85	14.08	−2.442	−0.773	0.000	0.000
Antiparallel WSe$_2$	3.28	9.32	47123	31153	35.03	12.17	−1.899	−0.1501	3.865	0.542

in $(\boldsymbol{G}^{(2)}, \boldsymbol{G}^{(3)})$ for all Fourier coefficients of the relaxation vectors $\tilde{\boldsymbol{u}}^{(1)}(\boldsymbol{G}^{(2)}, \boldsymbol{G}^{(3)})$. Physically, nonellipticity implies the instability due to the resonance of long wavelength modes in an infinite system. However, note that a finite physical system does not support such an instability if the elastic energy Eq. (20) is simply positive. In contrast, the PDEs for the intralayer energy of bilayer vdW heterostructures are elliptic. For bilayers, the variational elastic energy of L_i takes the following form:

$$\mathcal{E}_i(\boldsymbol{G}^{(j)}) = \frac{1}{2}\left(\tilde{\boldsymbol{u}}^{(i)}(\boldsymbol{G}^{(j)}) \otimes M_{ji}^T \boldsymbol{G}^{(j)}\right) : C :$$
$$\times \left(\tilde{\boldsymbol{u}}^{(i)}(\boldsymbol{G}^{(j)}) \otimes M_{ji}^T \boldsymbol{G}^{(j)}\right). \quad (23)$$

Similarly, taking $i = 1$, we have $\mathcal{E}_1(\boldsymbol{G}^{(2)}) \propto |M_{21}^T \boldsymbol{G}^{(2)}|^2 = |\boldsymbol{G}'^{(1)} - \boldsymbol{G}^{(2)}|^2$, where $\boldsymbol{G}'^{(1)} = A_1^{-T} A_2^T \boldsymbol{G}^{(2)}$. In this case, $|\boldsymbol{G}'^{(1)} - \boldsymbol{G}^{(2)}|$ is bounded from zero and takes its minimum values when $\boldsymbol{G}'^{(1)} - \boldsymbol{G}^{(2)}$ are one of the columns of $2\pi(A_1^{-T} - A_2^{-T})$, which can be understood as the primitive vectors of the bilayer moiré supercell reciprocal lattice.

D. Interlayer energy

We now describe the interlayer energy due to the misfit from the twist of the relaxation displacement. The origin of the interlayer energy is the slight misalignment between the two layers. This interaction is represented by the generalized stacking fault energy (GSFE). The GSFE is obtained with DFT by calculating the difference in total energy between configurations that involve relative rigid shifts between the two layers of a bilayer, with the zero corresponding to the reference ground-state-energy configuration. The GSFE can be expressed as a Fourier sum of the first few terms [35], and its functional form is given in Appendix A. The computational details of the GSFE are provided in Appendix B and the Fourier coefficients for graphene and WSe$_2$ are provided in Table I. The GSFE for these two materials are shown in Fig. 5. We note that due to a lower symmetry, for WSe$_2$, there are two possible stackings of a bilayer, referred to as parallel and antiparallel, which denote the relative orientation of the two layers in the reference configuration. In graphene and parallel WSe$_2$, the origin (AA stacking) is when the two layers are exactly aligned, and AB/BA stacking is when one sublattice is aligned with the other. In antiparallel WSe$_2$, we refer to the origin (BB stacking) as the stacking configuration when the chalcogen atoms (Se) are aligned, the AA stacking order as when the metal atoms (W) are aligned, and the AB stacking order is when the metal atom is aligned with the chalcogen atom. The stacking configuration is consistent with the notation in Phillips and Hellberg [36].

Unrelaxed localized misfit site energies $\Phi_{i\pm}^{\text{misfit}}: \Gamma^{i\pm1} \mapsto \mathbb{R}$ are defined by

$$\Phi_{i+}^{\text{misfit}}(\boldsymbol{b}^{i\to i+1}(\omega)) = \Phi_{(i+1)-}^{\text{misfit}}(\boldsymbol{b}^{i+1\to i}(\omega)). \quad (24)$$

These site energies thus satisfy the following symmetry relation for $\boldsymbol{s} \in A_i^{-1}\Gamma_i = A_{i+1}^{-1}\Gamma_{i+1} \equiv [0, 1)^2$:

$$\Phi_{(i+1)-}^{\text{misfit}}(A_i \boldsymbol{s}) = \Phi_{i+}^{\text{misfit}}(-A_{i+1}\boldsymbol{s}), \quad (25)$$

and we see that the misfit energies $\Phi_{(i+1)-}^{\text{misfit}}$ and $\Phi_{i+}^{\text{misfit}}$ can be computed from the same misfit energy between L_i and L_{i+1}. As mentioned above, the simulations in this paper approximate the misfit energy $\Phi_{i\pm}^{\text{misfit}}$ by the GSFE, $V_{i\pm}^{\text{GSFE}}$, defined in Eq. (A2) in Appendix A.

When the layers are relaxed, the argument of V^{GSFE} is modified to be the relaxation-modulated local shift $\boldsymbol{B}^{i\to j}$ defined in Eq. (10). The interlayer contribution of each layer can be given by

$$\Phi_{i,\omega}^{\text{inter}}(\boldsymbol{u}) = \begin{cases} \frac{1}{2}V_{1+}^{\text{GSFE}}(\boldsymbol{B}^{1\to2}(\omega)), & \text{for } \omega \in \Omega_1, \\ \frac{1}{2}\left[V_{2+}^{\text{GSFE}}(\boldsymbol{B}^{2\to3}(\omega)) + V_{2-}^{\text{GSFE}}(\boldsymbol{B}^{2\to1}(\omega))\right], & \text{for } \omega \in \Omega_2, \\ \frac{1}{2}V_{3-}^{\text{GSFE}}(\boldsymbol{B}^{3\to2}(\omega)), & \text{for } \omega \in \Omega_3. \end{cases} \quad (26)$$

The total interlayer energy can therefore be obtained by integrating over all local configurations ω:

$$E^{\text{inter}} = \frac{1}{2}\int_{\Omega_1} V_{1+}^{\text{GSFE}}(\boldsymbol{B}^{1\to2}(\omega))\,d\omega$$
$$+ \frac{1}{2}\int_{\Omega_2}\left[V_{2-}^{\text{GSFE}}(\boldsymbol{B}^{2\to1}(\omega)) + V_{2+}^{\text{GSFE}}(\boldsymbol{B}^{2\to3}(\omega))\right]d\omega$$
$$+ \frac{1}{2}\int_{\Omega_3} V_{3-}^{\text{GSFE}}(\boldsymbol{B}^{3\to2}(\omega))\,d\omega. \quad (27)$$

Note that for heterotrilayer materials, V_{2-}^{GSFE} and V_{2+}^{GSFE} have different functional forms.

E. Relation between real space and configuration space

Finally, we describe the mapping between real space and configuration space. Each adjacent pair of layers i and j forms a moiré supercell, Γ^{ij}, given by the columns of the matrix $(I - A_i A_j^{-1})^{-1}A_i = (A_i^{-1} - A_j^{-1})^{-1}$. At the real space position at a
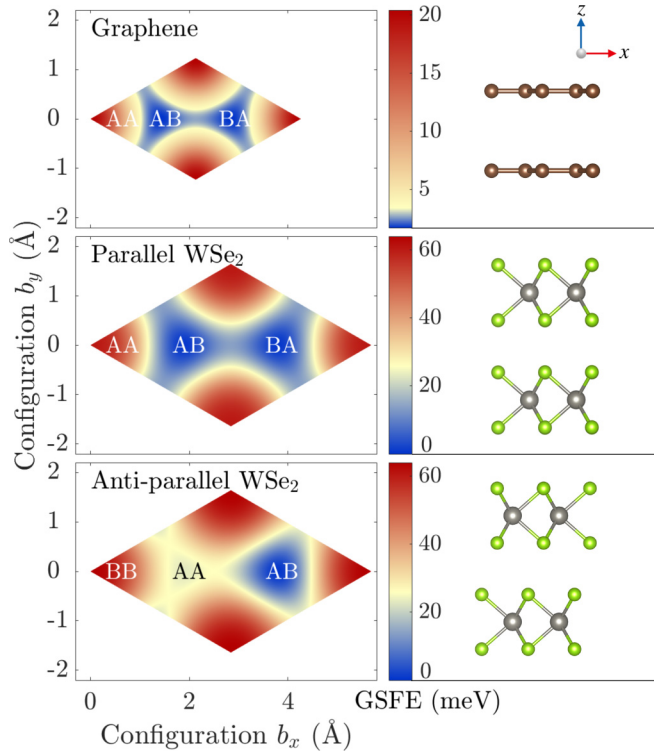
FIG. 5. Left: the GSFE for bilayer graphene, parallel WSe$_2$, and antiparallel WSe$_2$. Right: side views of the AA stacking orientation for each bilayer corresponding to the three cases on the left.

lattice site in L_i, $\boldsymbol{R}_{mn}^{(i)}$, we can calculate the local configuration relative to L_j by calculating the local shift vector [see Eq. (6)]:

$$\boldsymbol{b}^{(j)}\big(\boldsymbol{R}_{mn}^{(i)}\big) = \boldsymbol{R}_{mn}^{(j)} - \boldsymbol{R}_{mn}^{(i)} = (A_j - A_i)\binom{m}{n}$$

$$= \big(A_j A_i^{-1} - I\big)\boldsymbol{R}_{mn}^{(i)} \in \Gamma^{(j)}. \quad (28)$$

Note that this construction agrees with evaluating the interpolated disregistry $\boldsymbol{b}^{(j)}$ given by Eq. (3) at a lattice site $\boldsymbol{R}_{mn}^{(i)}$ as the reference point. This result can be generalized to any real space position $\boldsymbol{r}$:

$$\boldsymbol{b}^{i \to j}(\omega(\boldsymbol{r})) = \big(A_j A_i^{-1} - I\big)\boldsymbol{r} \in \Gamma^{(j)}, \quad (29)$$

where we notice that $\boldsymbol{b}^{i \to j}(\boldsymbol{r})$ is periodic in the bilayer moiré supercell Γ^{ij} in the continuum approximation, meaning it sends a position in the bilayer supercell to the unit cell of L_j. This generalization is equivalent to evaluating the interpolated disregistry given in Eq. (5) at a configuration ω.

The relation between the displacement defined in real space, $\boldsymbol{U}^{(i)}(\boldsymbol{r})$, and in configuration space, $\boldsymbol{u}^{(i)}(\boldsymbol{r})$, can then be given by evaluating $\widehat{\boldsymbol{u}}^{(i)}(\omega)$ at $\boldsymbol{b}^{i \to j}(\boldsymbol{r})$ and $\boldsymbol{b}^{i \to k}(\boldsymbol{r})$ with Eq. (29) above at a position $\boldsymbol{r}$ to obtain

$$\boldsymbol{U}^{(i)}(\boldsymbol{r}) = \boldsymbol{u}^{(i)}(\boldsymbol{b}^{i \to j}(\omega(\boldsymbol{r})), \boldsymbol{b}^{i \to k}(\omega(\boldsymbol{r}))), \quad (30)$$

where $j, k \neq i$ and $j < k$. We see that $\boldsymbol{U}^{(i)}(\boldsymbol{r})$ is not generally periodic since $\boldsymbol{b}^{(j)}(\boldsymbol{r})$ and $\boldsymbol{b}^{(k)}(\boldsymbol{r})$ are each periodic on the incommensurate moiré supercells Γ^{ij} and Γ^{ik}. In comparison, for incommensurate bilayers $\boldsymbol{U}^{(i)}(\boldsymbol{r}) = \widehat{\boldsymbol{u}}^{(i)}(\omega)$ is periodic on Γ^{ij}.

III. RESULTS

A. General twist angles, $\theta_{12} \neq \theta_{23}$

We perform energy minimization in configuration space with the discretized grid $N \times N \times N \times N$ using the JULIA OPTIM package [37]. We present our results in Fig. 6 for $\theta_{12} = 0.5°, \theta_{23} = 0.2°$ by taking two-dimensional line cuts of the configuration space, using $N = 36$. We show the results for twisted trilayer graphene in Figs. 6(a)–6(d) and for trilayer antiparallel WSe$_2$ in Figs. 6(e)–6(h). For the L_1, L_2 bilayer pair, we take the average over the configurations of L_3 (left), and similarly for the L_2, L_3 bilayer pair, we take the average over the configurations of L_1 (right). For the relaxation displacement vectors, we show the change in modulated local shift vector compared to the local shift vector without relaxation between the bilayer pair, i.e., $(\boldsymbol{B}^{i \to 2} - \boldsymbol{b}^{i \to 2})$ for $i = 1, 3$, where $\boldsymbol{b}^{i \to 2}$ and $\boldsymbol{B}^{i \to 2}$ are the local shift vector and the modulated local shift vector, respectively, given in Eqs. (5) and (10), respectively. For example, for the relative relaxation
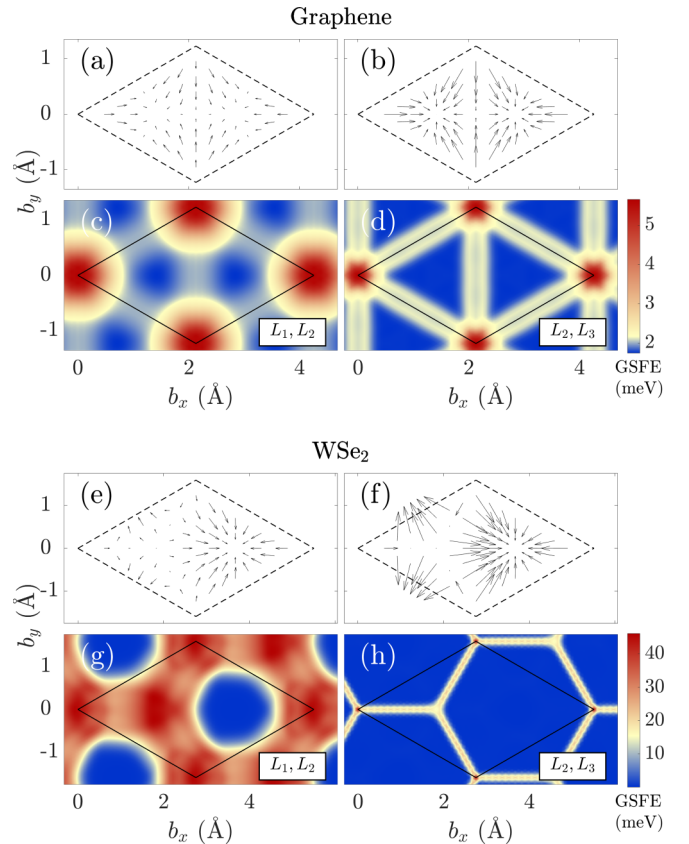

FIG. 6. Two-dimensional line cuts of the relaxation pattern and relaxed GSFE in configuration space for (a)–(d) trilayer graphene and (e)–(h) trilayer antiparallel WSe$_2$ at $\theta_{12} = 0.5°, \theta_{23} = 0.2°$. (a) and (b) show the change in the local shift vectors, $(\boldsymbol{B}^{i \to 2} - \boldsymbol{b}^{i \to 2})$. (a) shows $(\boldsymbol{B}^{1 \to 2} - \boldsymbol{b}^{1 \to 2})$, averaged over the L_3 configuration, and (b) shows $(\boldsymbol{B}^{3 \to 2} - \boldsymbol{b}^{3 \to 2})$, averaged over the L_1 configuration. (c) and (d) show the relaxed GSFE. (c) shows the sum of the GSFE between the L_1, L_2 bilayer pair, averaged over the L_3 configuration, and (d) shows the sum of the GSFE between the L_2, L_3 bilayer pair, averaged over the L_1 configuration. (e)–(h) are the corresponding plots for antiparallel WSe$_2$.
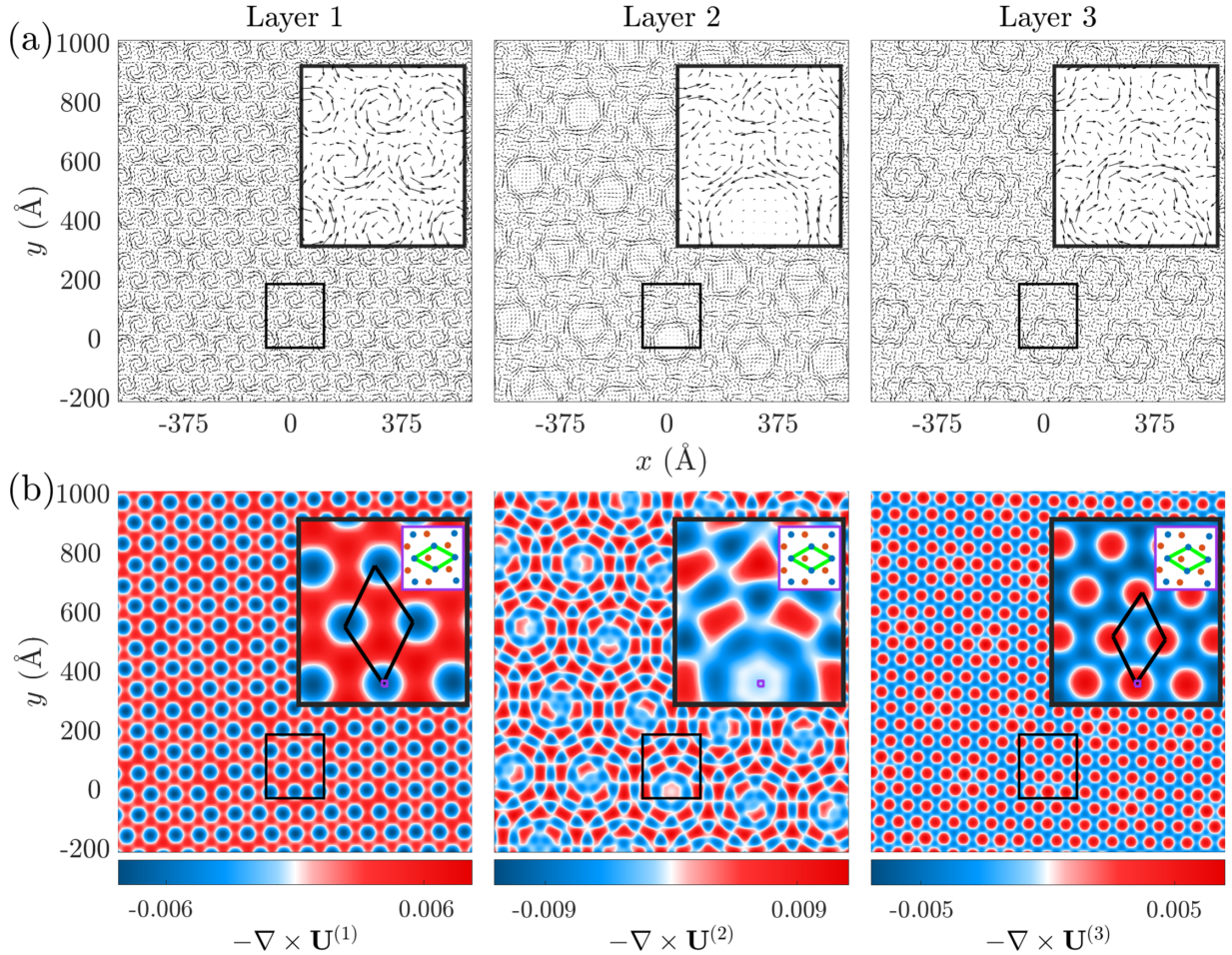
FIG. 7. Real space relaxation pattern for twisted trilayer graphene with $\theta_{12} = 1.73°$, $\theta_{23} = 2.24°$, using discretization $N = 81$. Insets on the upper right corner zoom in on the black boxed areas near the origin. (a) The relaxation displacement $\boldsymbol{U}^{(i)}(\boldsymbol{r})$. (b) Color plot shows the curl of the displacement vectors $-\nabla_{\boldsymbol{r}} \times \boldsymbol{U}^{(i)}(\boldsymbol{r})$ indicating the rotation direction: blue is clockwise rotation and red is counterclockwise rotation, and white means no rotation. Black parallelograms on L_1 and L_3 show the approximate supercell vectors that correspond to θ_{12} and θ_{23}, respectively, calculated using Eq. (32). The small insets on the upper right corner zoom in on the small purple boxed areas in the black insets. In the purple insets, blue and red dots show A and B sublattices, respectively, and the green parallelogram shows the monolayer cell.

between the L_1 and L_2 pair, we plot the following quantity:

$$
\begin{aligned}
&\boldsymbol{B}^{1\to2}(\boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)}) - \boldsymbol{b}^{1\to2}(\boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)}) \\
&= \boldsymbol{u}^{(2)}\big(A_1 A_2^{-1} \boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)} - A_3 A_2^{-1} \boldsymbol{b}^{(2)}\big) - \boldsymbol{u}^{(1)}(\boldsymbol{b}^{(2)}, \boldsymbol{b}^{(3)}),
\end{aligned}
$$

$$(31)$$

as a function of $\boldsymbol{b}^{(2)}$, averaged over the $\boldsymbol{b}^{(3)}$. Physically, this means displaying the relative relaxation vectors from L_2 to L_1, averaged over all the L_3 shifts. Similarly, for L_2 and L_3, we show $(\boldsymbol{B}^{3\to2} - \boldsymbol{b}^{3\to2})$ as a function of $\boldsymbol{b}^{(2)}$, averaged over all the L_1 configurations. The relaxation between L_2 and L_3 for both graphene and WSe$_2$ is stronger because of the smaller corresponding twist angle θ_{23}. The relaxation displacement vectors point towards the AB/BA stacking in graphene and AB stacking in antiparallel WSe$_2$, which are the equilibrium stacking orders. For both graphene and WSe$_2$, the relaxed GSFE exhibits domain wall formation between the adjacent bilayer pairs with different thicknesses. A smaller corresponding twist angle gives rise to stronger relaxation and thinner domain walls, similar to twisted bilayer graphene.

The shape of the domains are different depending on the lattice symmetry: Graphene has triangular domains while antiparallel WSe$_2$ has hexagonal domains due to its broken inversion symmetry.

Unlike in the case of twisted bilayer heterostructures, there does not exist in trilayers a linear mapping at the continuous level between configuration space and a periodic moiré supercell in real space. From the energy minimization in configuration space, we obtain a set of displacement vectors, $\boldsymbol{u}^{(i)}(\boldsymbol{b}^{(j)}, \boldsymbol{b}^{(k)})$, for each layer. The real space displacement $\boldsymbol{U}^{(i)}(\boldsymbol{R}^{(i)})$ at lattice points $\boldsymbol{R}^{(i)}$ in L_i is then given by evaluating $\boldsymbol{U}^{(i)}(\boldsymbol{R}^{(i)}) = \boldsymbol{u}^{(i)}(\boldsymbol{b}^{(j)}(\boldsymbol{R}^{(i)}), \boldsymbol{b}^{(k)}(\boldsymbol{R}^{(i)}))$. In contrast, the configuration space for the bilayer case is two-dimensional and the displacement is more simply given by $\boldsymbol{U}^{(i)}(\boldsymbol{R}^{(i)}) = \boldsymbol{u}^{(i)}(\boldsymbol{b}^{(j)}(\boldsymbol{R}^{(i)}))$. Figure 7 shows the real space relaxation displacement vectors $\boldsymbol{U}^{(i)}(\boldsymbol{r})$ and the curl of $-\boldsymbol{U}^{(i)}(\boldsymbol{r})$, $-(\nabla_{\boldsymbol{r}} \times \boldsymbol{U}^{(i)}(\boldsymbol{r}))$, which measures the rotation component of the modulation displacement vector field, for twisted trilayer graphene [38]. Physically, this quantity corresponds to the change of the local twist angle, which can be measured using the SQUID-on-tip technique [39]. In Fig. 7(a), we approximate the
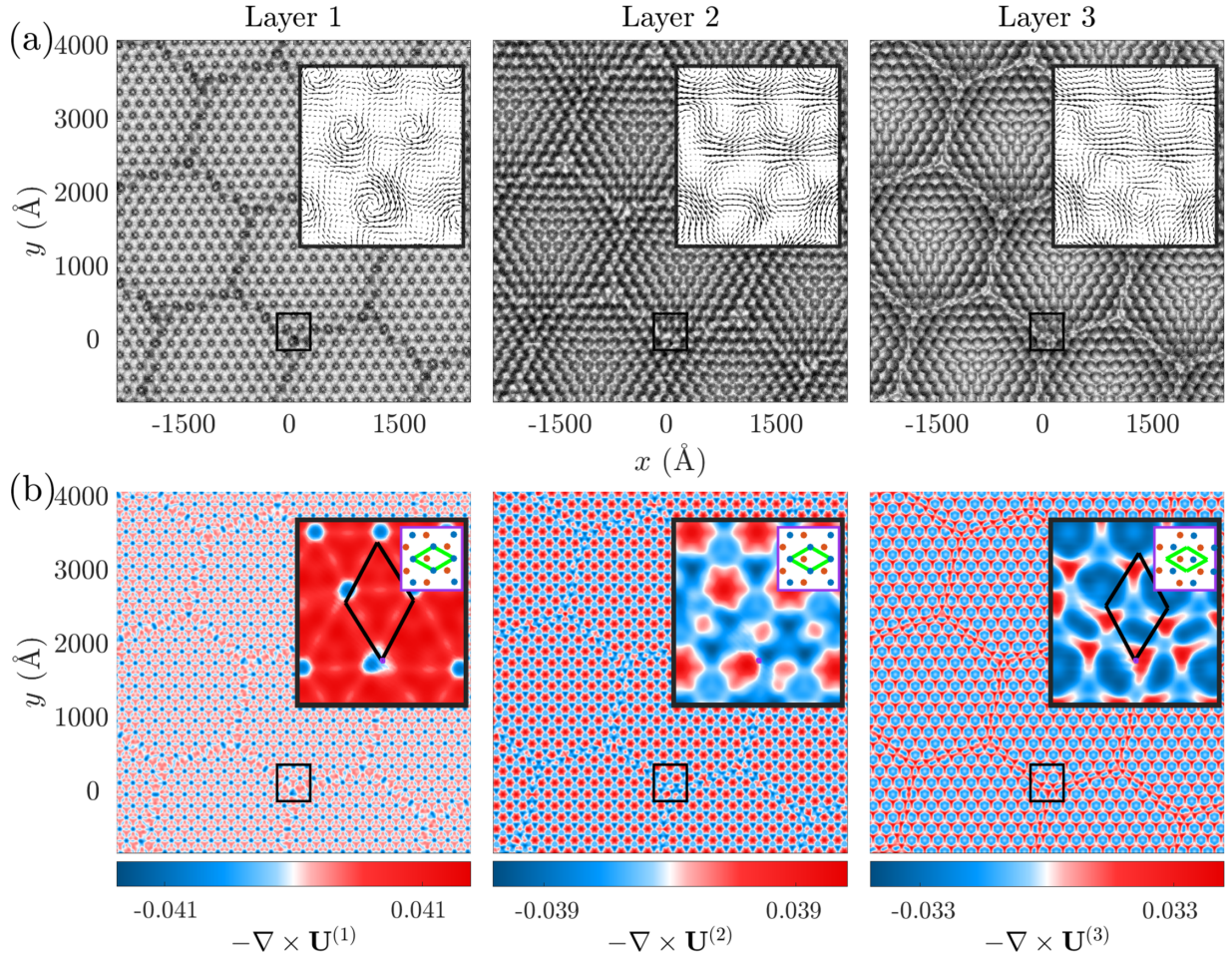
FIG. 8. Same as Fig. 7, for trilayer WSe$_2$, with parallel stack between L_1 and L_2 at $\theta_{12} = 1.0°$ and antiparallel stack between L_2 and L_3 at $\theta_{23} = 1.1°$, using discretization $N = 108$.

supercell lattice constant that corresponds to the twist angle $\theta_{i,i+1}$ according to

$$a_{sc}(\theta_{i,i+1}) = \frac{a_0}{2 \sin(\theta_{i,i+1}/2)}, \tag{32}$$

where a_0 is the lattice constant of the graphene/WSe$_2$ unit cell given in Table I. The length scale $a_{sc}(\theta_{i,i+1})$ can be understood as the periodicity of a twisted bilayer with the twist angle $\theta_{i,i+1}$ in the continuum limit. Figure 7(a) shows the relaxation displacement vectors. Near the origin, the relaxation tends to follow the local rotation to increase the local twist angle to reduce the size of the AA spots, which agrees with the bilayer results [38].

To better understand the relaxation pattern, Fig. 7(b) shows $-(\nabla_r \times U^{(i)}(r))$. On the bilayer moiré scales (see insets on the upper right corners), the relaxation pattern has two length scales: In L_1, the length scale of the relaxation is dominated by the continuum supercell lattice constant of θ_{12}, $a_{sc}(\theta_{12})$; in L_3, the relaxation has a length scale $a_{sc}(\theta_{23})$ instead. In the middle layer (L_2), the relaxation pattern is a result of the interference between the two bilayer moiré lengths, as a result of the coupling to both L_1 and L_3. Zooming out, we observe structures at a larger length scale than both of the bilayer moiré lengths, which is the "moiré of moiré," due to the interference between the two bilayer moiré cells.

Note that these larger structures are not identical to each other [most visible in Fig. 7(b), layer 2] due to the incommensurability of the system. The emergence of the larger length scales can also be explained by the nonellipticity of the system (see Sec. II C). The reciprocal space variational elastic energy components given in Eq. (20), $\mathcal{E}_i(G^{(j)}, G^{(k)})$, have multipliers $|G'^{(i)} - G^{(j)} - G^{(k)}|^2$ which can take an arbitrarily small value, which means the real space relaxation pattern can have arbitrarily large structures. The quasiperiodic large structures we observe in Fig. 7 manifests the nonellipticity of the Euler-Lagrangian PDE, and since these structures are aperiodic, zooming out further will lead to the appearance of even longer quasiperiodic structures. This shows our model's capability of capturing the incommensurability of the trilayer relaxation pattern.

To show the generality of our model, we can examine the results for other materials. For example, Fig. 8 shows the relaxation displacement vectors and the corresponding curl in trilayer WSe$_2$ heterostructures, with parallel stacking between L_1 and L_2 and antiparallel stacking between L_2 and L_3. Similarly to the graphene case, we observe the coupling between two length scales that corresponds to θ_{12} and θ_{23}. On the bilayer moiré length scales, the relaxation forms triangular domains between L_1 and L_2 and hexagonal domains between L_2 and L_3. On the moiré of moiré length scale, the overall
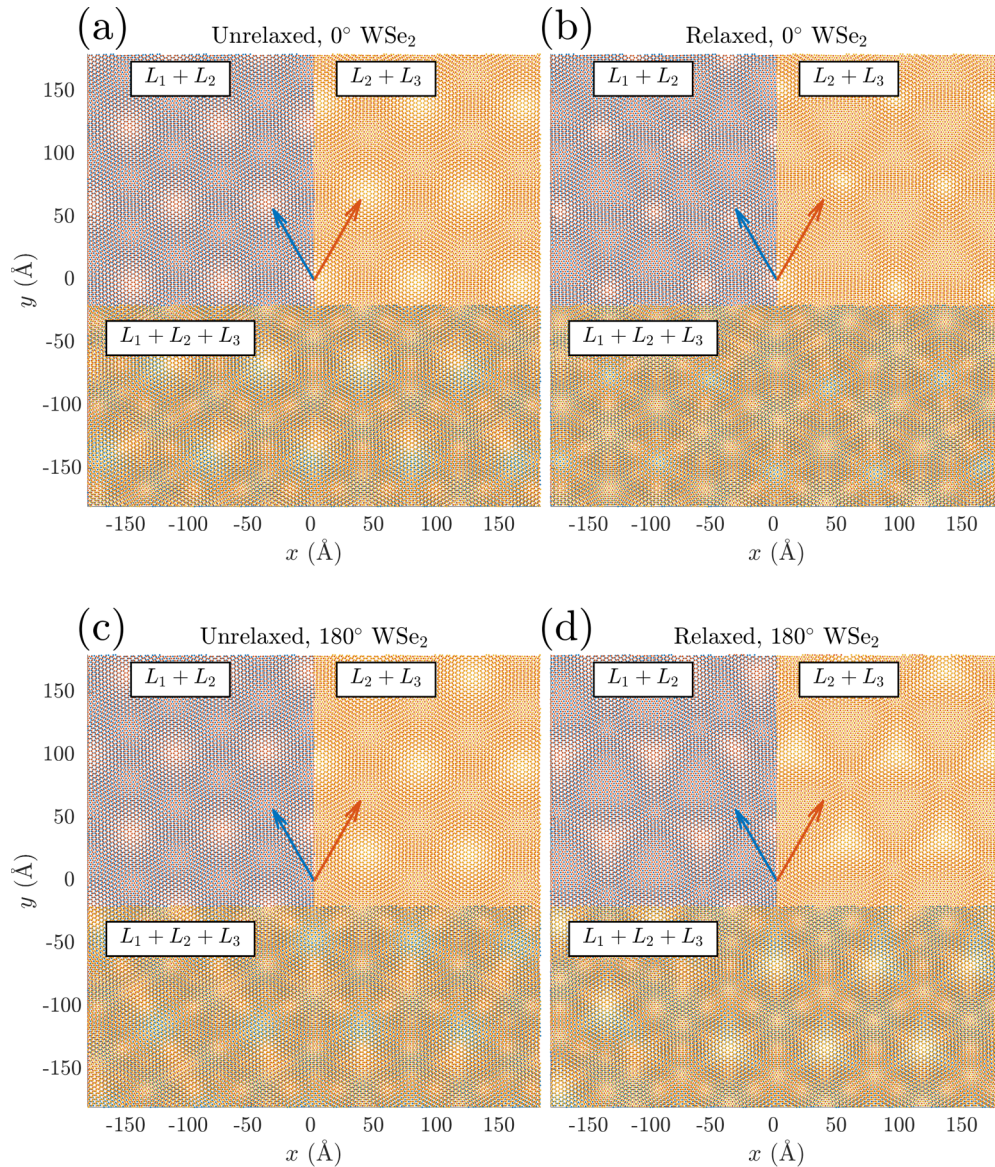
FIG. 9. Comparison between unrelaxed and relaxed atomic positions in twisted trilayer WSe_2 with (a),(b) parallel and (c),(d) antiparallel stackings at $\theta_{12} = 2.6°$ and $\theta_{23} = 2.3°$. In all panels, top left: atomic positions of L_1 and L_2 only (removing L_3); top right: atomic positions of L_2 and L_3 only (removing L_1); bottom: atomic positions of the three layers. Red and blue arrows are guides to eyes, showing the bilayer moiré lengths that correspond to $\theta_{12} = 2.6°$ and $\theta_{23} = 2.3°$.

domains are also hexagonal, since the overall inversion symmetry of the system is broken by antiparallel WSe_2. The shape can also be understood as the interference pattern between the triangular domains and hexagonal domains.

We note that when $\theta_{12} = -\theta_{23}$, L_1 and L_3 are aligned with a twisted middle layer. The system can also be modeled as a bilayer system, the interlayer energy of L_2 doubles and the strength of relaxation in L_2 doubles. We used our trilayer model to calculate the relaxation pattern for such systems and obtain a similar result as the relaxation pattern shown in Carr *et al.* [40].

We examine the real space atomic positions after relaxation. Figure 9 compares the unrelaxed and relaxed atomic positions for parallel and antiparallel WSe_2 at $\theta_{12} = 3°$, $\theta_{23} = 2.3°$. Note that the unrelaxed atomic positions are different in the parallel and antiparallel stackings [Fig. 9(a), 9(c)]. In

both cases, the relaxation forms domain between L_1 and L_2, as well as between L_2 and L_3, but at different length scales, to enlarge equilibrium stacking orders. Figure 9(b) shows that in parallel WSe_2, AB/BA stacking regions are enlarged and the AA stacking regions shrink. Figure 9(d) shows the domain formation in antiparallel WSe_2, in which AB stacking regions are enlarged.

B. Equal twist angles, $\theta_{12} = \theta_{23}$

In the special case where $\theta_{12} = \theta_{23}$, we found that there exist two types of solutions: symmetric and asymmetric. An example is shown in Fig. 10. The symmetric solution [Figs. 10(a), 10(b)] preserves the twofold rotational symmetry along the z axis (out-of-plane direction) of the unrelaxed trilayers, whereas the asymmetric solution [Figs. 10(c),
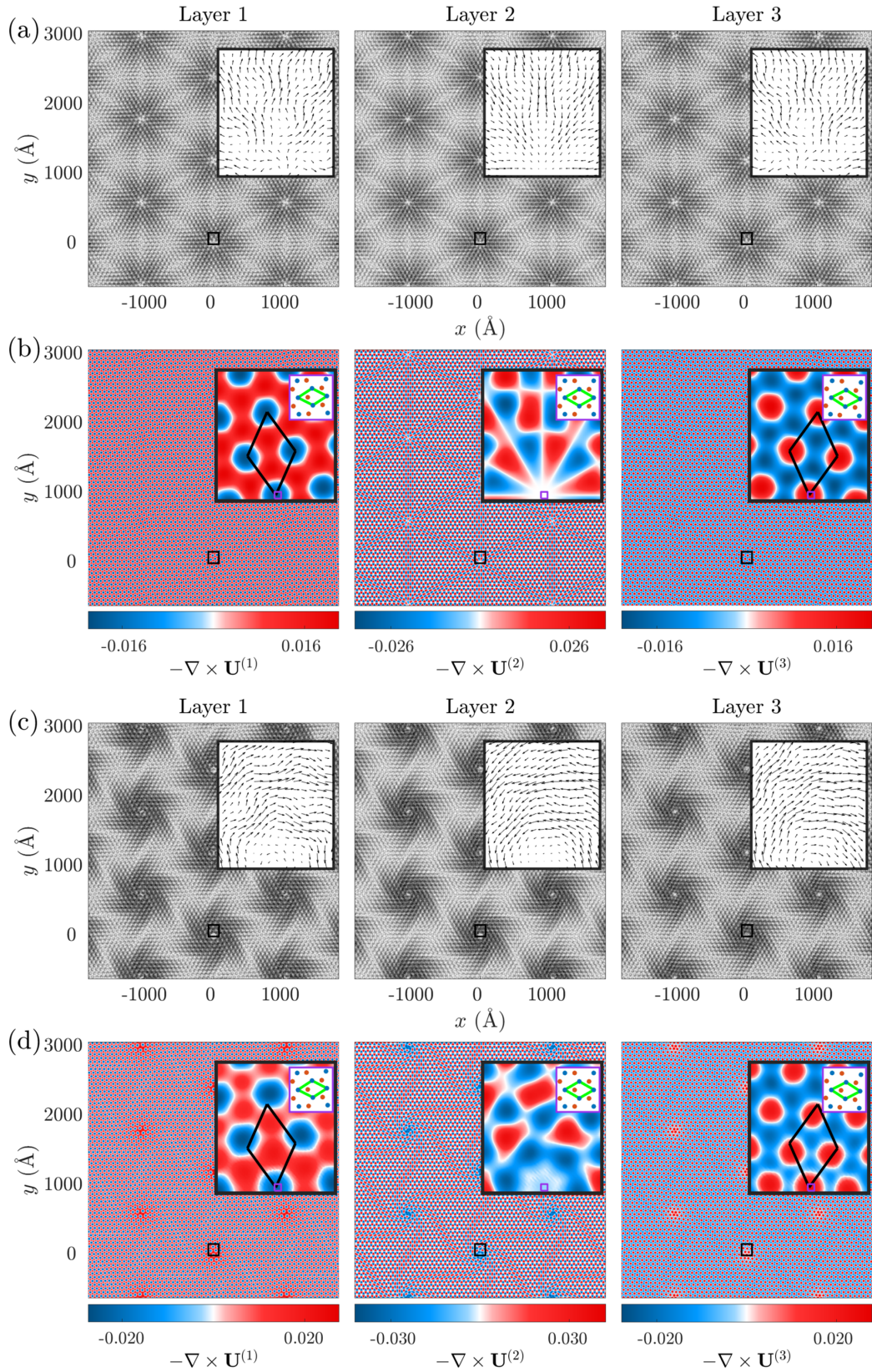
FIG. 10. Relaxation pattern for parallel WSe$_2$ at $\theta_{12} = \theta_{23} = 3°$, same as Fig. 7, showing the symmetric solution in (a),(b) and the asymmetric solution in (c),(d).
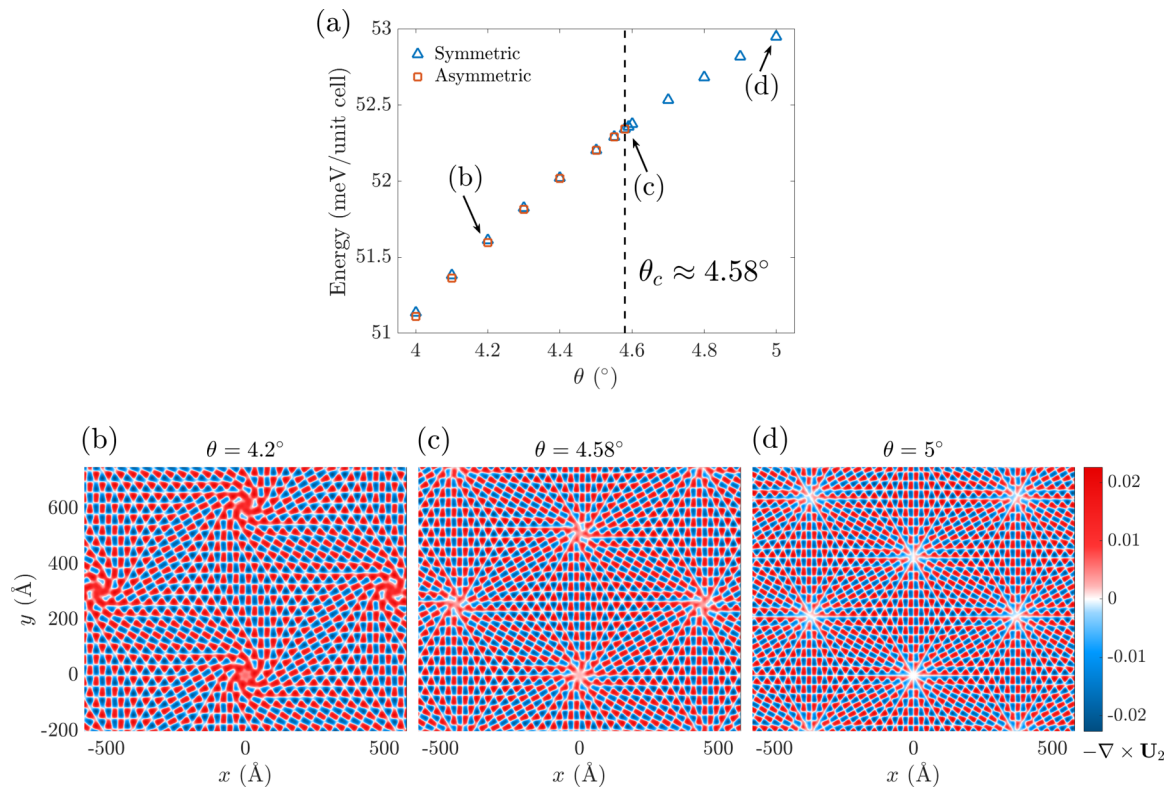
FIG. 11. (a) Relaxed total energy as a function of the twist angle shows the bifurcation critical point is around $\theta = 4.58°$. (b)–(d) The curl of relaxation displacement vectors in L_2 at the global minimum, (b) below, (c) at, and (d) above the bifurcation point.

10(d)] exhibits a pinwheel-shaped relaxation pattern, which remarkably breaks this symmetry. Near the origin (local AAA stacking) of L_2, the relaxation has zero magnitude for the symmetric solution, whereas the relaxation is finite for the asymmetric solution. The asymmetric solution is found to have a lower energy in this case: for example, at $\theta_{12} = \theta_{23} = 3°$ for parallel WSe$_2$, the total energies per monolayer unit cell after relaxation are 39.34 meV and 39.30 meV for the symmetric and asymmetric solutions, respectively. Note that there exists two asymmetric solutions with equal energy, in which the pinwheel rotates in the opposite directions.

Physically, the asymmetric solution redistributes the local AA stacking density between the neighboring bilayer pairs to reduce the total AA areas and consequently the total interlayer energy, at the cost of additional elastic energy due to the pinwheel relaxation pattern. This can be seen from the insets in Fig. 10. In the insets of Fig. 10(b) for L_1 and L_3, the black parallelograms (bilayer moiré cells) agrees with the length scale of the relaxation pattern. In L_2, relaxation is canceled out at the AAA spots (white color near the origin) to preserve the twofold rotational symmetry. In Fig. 10(d), however, the insets show that the relaxation near the origin in L_1 and L_3 now has different length scales than the bilayer moiré length corresponding to 3°. We also observe that the relaxation in L_2 obtains a preferred net rotation, which breaks the symmetry between the top and bottom layers. These results suggest that relaxation can spontaneously break the symmetry of the system, unlike in twisted bilayers in which the in-plane relaxation always preserves the symmetry.

A bifurcation point occurs at a critical twist angle, $\theta_c \approx 4.55°$. Above θ_c, the only solution is the symmetric one; below θ_c, the symmetric solution becomes an unstable saddle point and the two counter-rotating asymmetric solutions become the global minima. The existence of a critical twist angle can be understood from the competition between reducing the misfit energy and increasing elastic energy. As the twist angle increases, the moiré of moiré length scale becomes smaller, and therefore elastic energy cost for the pinwheel pattern increases. Eventually, the reduction in the misfit energy is not enough to compensate for the elastic energy cost of the pinwheel pattern and the system no longer supports the asymmetric solution.

Figure 11(a) shows the total energy after relaxation as a function of twist angle in twisted trilayer parallel WSe$_2$ for the two types of solution, confirming the existence of a critical bifurcation point at $\theta_c \approx 4.58°$. Above this θ_c, the symmetric solution is the only solution. Below θ_c, the asymmetric solution is the global minimum and the symmetric solution becomes an unstable saddle point and has slightly higher energy. Figures 11(b)–11(d) shows the three types of local minima below, at, and above the critical twist angle. In particular, Fig. 11(c) shows that at θ_c, the asymmetric solution has very weak pinwheellike rotation and it almost becomes a symmetric one. While we use twisted parallel WSe$_2$ as an example, the bifurcation also occurs in graphene or any trilayers with inversion symmetry between AB and BA stacking orders.

IV. CONCLUSION

We have presented a general formalism to obtain relaxation patterns in twisted trilayer systems, which cannot be achieved with the traditional continuum model in real space due to the aperiodicity of the system. We give results for twisted trilayer graphene and WSe$_2$. We introduce configuration space as a natural description of the twisted trilayer system. We show that the relaxation pattern of trilayer systems is the "moiré of moiré," as a result of the incommensurate coupling between two bilayer moiré lengths that correspond to the two twist angles. The relaxation pattern manifests the nonellipticity of the energy minimization problem in the sense that there are infinitely many large length scales emerging in the system, and our model is capable of capturing these large length scale modes. Similar to the bilayer relaxation, in trilayers, the relaxation forms domains at the bilayer moiré lengths to maximize the equilibrium stacking orders and minimize the high energy stacking order (AA stack). In contrast to the in-plane relaxation pattern in twisted bilayer graphene where the symmetry of the system is preserved, we found that the twisted trilayers can have symmetry-breaking relaxation patterns. When $\theta_{12} = \theta_{23}$, the trilayer relaxation has two types of solutions: one symmetric solution that preserves the twofold rotational symmetry along the z axis and the other asymmetric solution that breaks this rotational symmetry. This symmetry breaking might be experimentally observable effect, for example, through second harmonic generation or scanning transmission electron microscopy.

Although we have only presented results for homotrilayers in this work, our approach can be also applied to heterotrilayer materials. For materials with similar lattice constants (such as graphene and hBN), we can simply apply a small strain to the monolayer by assuming all three layers have the same lattice constant. In this way, we can calculate the GSFE by performing rigid shifts in a unit cell using DFT total energies as the interlayer coupling, which effectively takes care of the small strain. For materials with very different lattice constants, for example, graphene on top of transition metal dichalcogenides, we need to modify the definition of A_i matrices and be careful about mapping from configuration space to real space. The configuration space approach can also be generalized to four or more layers, and each additional layer adds two dimensions to configuration space. However, the extension would be very computationally expensive—the computational cost scales as $N^{2(d-1)}$, where d is the dimension and N is the discretization. Therefore, more sophisticated minimization methods should be introduced for computational efficiency. Moreover, we would expect a symmetry-breaking solution to be present for heterostructures with an odd number of layers.

The mechanical relaxation effect is an important correction to be taken into account for in modeling the electronic structure of twisted trilayers. The relaxation pattern in twisted trilayer may be used as a new platform to engineer strain in vdW heterostructures. The second twist angle is an additional degree of freedom to manipulate the electronic states and explore unconventional strongly correlated states [26].

APPENDIX A: GENERALIZED STACKING FAULT ENERGY

The GSFE is the total energy at different stacking configurations without rotation between a bilayer. Letting (b_x, b_y) be the relative displacement or the local stacking order between two adjacent layers, we define the following vector $\boldsymbol{v} = (v, w) \in [0, 2\pi]^2$:

$$\begin{pmatrix} v \\ w \end{pmatrix} = \frac{2\pi}{\alpha} \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{pmatrix} b_x \\ b_y \end{pmatrix}. \tag{A1}$$

In terms of $\boldsymbol{v}$, the GSFE of a hexagonal lattice can be expressed as a Fourier sum as [17,35,41],

$$\begin{aligned} V_{j\pm}^{\text{GSFE}} &= c_0 + c_1(\cos v + \cos w + \cos(v + w)) \\ &\quad + c_2(\cos(v + 2w) + \cos(v - w) + \cos(2v + w)) \\ &\quad + c_3(\cos(2v) + \cos(2w) + \cos(2v + 2w)) \\ &\quad + c_4(\sin v + \sin w - \sin(v + w)) \\ &\quad + c_5(\sin(2v + 2w) - \sin(2v) - \sin(2w)), \quad \text{(A2)} \end{aligned}$$

where we use the subscript $j\pm$ to denote the GSFE between the L_j and $L_{j\pm1}$. We note that we are only including the first few terms in the Fourier series because of the lattice symmetry and higher order terms are three or more orders of magnitude smaller than the leading order terms. For other lattice geometries, it may be necessary to include higher order terms in the Fourier series.

APPENDIX B: DENSITY FUNCTIONAL THEORY CALCULATIONS

The coefficients for the interlayer and intralayer energies are obtained from DFT calculations with the Vienna Ab initio Simulation Package (VASP) [42,43]. The van der Waals force is implemeneted through the vdW-DFT method using the SCAN+rVV10 functional [44–47]. The values are presented in Table I. The elasticity constants and interlayer coupling coefficients of graphene and MoS$_2$ are also given in S. Carr *et al.* [17].

To obtain the GSFE, we performed a rigid shift between a bilayer using a 9×9 grid in the unit cell. We extract the ground state energy from DFT, by allowing relaxation in the out-of-plane direction but fixing the in-plane direction. Figure 5 shows the GSFE of graphene and WSe$_2$. In graphene, the AB/BA stacking (when the A sublattice of one layer is aligned with the B sublattice of the second layer) is the equilibrium stacking configuration and the AA stacking (the

origin) has the highest energy. In TMDCs, MX_2, with M being a transition metal atom (e.g., Mo, W) and X being a chalcogen atom (S, Se, Te), the bilayer can have two orientations when performing the rigid shifts because the two sublattices have different atomic species. In parallel WSe_2, the two bilayers are aligned. The GSFE energy landscape is similar to that of graphene. In antiparallel WSe_2, however, the inversion sym-

metry is broken. In this case, the equilibrium stacking order is when the metal atoms and the chalcogen atoms are aligned (AB stacking). The GSFE is a more accurate description of interlayer coupling compared to empirical interlayer coupling potentials when they exist, such as for graphene and hBN, and it enables the description of interlayer coupling in TMDCs for which empirical interlayer coupling potentials do not exist.

[1] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Unconventional superconductivity in magic-angle graphene superlattices, Nature (London) **556**, 43 (2018).

[2] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Correlated insulator behavior at half-filling in magic-angle graphene superlattices, Nature (London) **556**, 80 (2018).

[3] M. Yankowitz, J. Jung, E. Laksono, N. Leconte, B. L. Chittari, K. Watanabe, T. Taniguchi, S. Adam, D. Graf, and C. R. Dean, Dynamic band-structure tuning of graphene moiré superlattices with pressure, Nature (London) **557**, 404 (2018).

[4] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, Spin-polarized Correlated Insulator and Superconductor in Twisted Double Bilayer Graphene, arXiv:1903.08130.

[5] Y. Cao, D. Rodan-Legrain, O. Rubies-Bigorda, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Tunable correlated states and spin-polarized phases in twisted bilayer-bilayer graphene, Nature (2020), doi:10.1038/s41586-020-2260-6.

[6] G. W. Burg, J. Zhu, T. Taniguchi, K. Watanabe, A. H. MacDonald, and E. Tutuc, Correlated Insulating States in Twisted Double Bilayer Graphene, Phys. Rev. Lett. **123**, 197702 (2019).

[7] C. Shen, Y. Chu, Q. Wu, N. Li, S. Wang, Y. Zhao, J. Tang, J. Liu, J. Tian, K. Watanabe, T. Taniguchi, R. Yang, Z. Y. Meng, D. Shi, O. V. Yazyev, and G. Zhang, Correlated states in twisted double bilayer graphene, Nat. Phys. **16**, 520 (2020).

[8] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, and A. Bachtold, Superconductors, orbital magnets, and correlated states in magic-angle bilayer graphene, Nature **574**, 653 (2019).

[9] G. Chen, L. Jiang, S. Wu, B. Lyu, H. Li, B. L. Chittari, K. Watanabe, T. Taniguchi, Z. Shi, J. Jung, Y. Zhang, and F. Wang, Evidence of a gate-tunable Mott insulator in a trilayer graphene moiré superlattice, Nat. Phys. **15**, 237 (2019).

[10] S. Moriyama, Y. Morita, K. Komatsu, K. Endo, T. Iwasaki, S. Nakaharai, Y. Noguchi, Y. Wakayama, E. Watanabe, D. Tsuya, K. Watanabe, and T. Taniguchi, Observation of superconductivity in bilayer graphene/hexagonal boron nitride superlattices, arXiv:1901.09356.

[11] G. A. Tritsaris, S. N. Shirodkar, E. Kaxiras, P. Cazeaux, M. Luskin, P. Plecháč, and E. Cancès, Perturbation theory for weakly coupled two-dimensional layers, J. Mater. Res. **31**, 959 (2016).

[12] P. San-Jose, A. Gutiérrez-Rubio, M. Sturla, and F. Guinea, Spontaneous strains and gap in graphene on boron nitride, Phys. Rev. B **90**, 075428 (2014).

[13] S. Dai, Y. Xiang, and D. J. Srolovitz, Twisted Bilayer Graphene: Moiré with a Twist, Nano Lett. **16**, 5923 (2016).

[14] N. N. T. Nam and M. Koshino, Lattice relaxation and energy band modulation in twisted bilayer graphene, Phys. Rev. B **96**, 075311 (2017).

[15] M. I. Español, D. Golovaty, and J. P. Wilber, Discrete-to-continuum modeling of weakly interacting incommensurate chains, Phys. Rev. E **96**, 033003 (2017).

[16] M. I. Español, D. Golovaty, and J. P. Wilber, Discrete-to-continuum modeling of weakly interacting incommensurate two-dimensional lattices, Proc. R. Soc. London A **474**, 20170612 (2018).

[17] S. Carr, D. Massatt, S. B. Torrisi, P. Cazeaux, M. Luskin, and E. Kaxiras, Relaxation and domain formation in incommensurate two-dimensional heterostructures, Phys. Rev. B **98**, 224102 (2018).

[18] N. Leconte, S. Javvaji, J. An, and J. Jung, Relaxation Effects in Twisted Bilayer Graphene: a Multi-Scale Approach, arXiv:1910.12805.

[19] S. N. Coppersmith, D. S. Fisher, B. I. Halperin, P. A. Lee, and W. F. Brinkman, Dislocations and the commensurate-incommensurate transition in two dimensions, Phys. Rev. B **25**, 349 (1982).

[20] J. S. Alden, A. W. Tsen, P. Y. Huang, R. Hovden, L. Brown, J. Park, D. A. Muller, and P. L. McEuen, Strain solitons and topological defects in bilayer graphene, Proc. Natl. Acad. Sci. USA **110**, 11256 (2013).

[21] C. R. Woods, L. Britnell, A. Eckmann, R. S. Ma, J. C. Lu, H. M. Guo, X. Lin, G. L. Yu, Y. Cao, R. V. Gorbachev, A. V. Kretinin, J. Park, L. A. Ponomarenko, M. I. Katsnelson, Yu. N. Gornostyrev, K. Watanabe, T. Taniguchi, C. Casiraghi, H. J. Gao, A. K. Geim, and K. S. Novoselov, Commensurate-incommensurate transition in graphene on hexagonal boron nitride, Nat. Phys. **10**, 451 (2014).

[22] M. M. van Wijk, A. Schuring, M. I. Katsnelson, and A. Fasolino, Relaxation of moiré patterns for slightly misaligned identical lattices: graphene on graphite, 2D Materials **2**, 034010 (2015).

[23] M. Yankowitz, K. Watanabe, T. Taniguchi, P. San-Jose, and B. J. Leroy, Pressure-induced commensurate stacking of graphene on boron nitride, Nat. Commun. **7**, 13168 (2016).

[24] N. Yeon Kim, H. Y. Jeong, J. H. Kim, G. Kim, H. S. Shin, and Z. Lee, Evidence of local commensurate state with lattice match of graphene on hexagonal boron nitride, ACS Nano **11**, 7084 (2017).

[25] H. Yoo, R. Engelke, S. Carr, S. Fang, K. Zhang, P. Cazeaux, S. H. Sung, R. Hovden, A. W. Tsen, T. Taniguchi, K. Watanabe, G.-C. Yi, M. Kim, M. Luskin, E. B. Tadmor, E. Kaxiras, and P. Kim, Atomic and electronic reconstruction at the van der Waals interface in twisted bilayer graphene, Nat. Mater. **18**, 448 (2019).

[26] K.-T. Tsai, X. Zhang, Z. Zhu, Y. Luo, S. Carr, M. Luskin, E. Kaxiras, and K. Wang, Correlated superconducting and insulating states in twisted trilayer graphene moire of moire superlattices, arXiv:1912.03375.

[27] C. Mora, N. Regnault, and B. A. Bernevig, Flatbands and Perfect Metal in Trilayer Moiré Graphene, Phys. Rev. Lett. **123**, 026402 (2019).

[28] Z. Ma, S. Li, Y.-W. Zheng, M.-M. Xiao, H. Jiang, J.-H. Gao, and X. C. Xie, Topological flat bands in twisted trilayer graphene, arXiv:1905.00622.

[29] S. Carr, S. Fang, Z. Zhu, and E. Kaxiras, An exact continuum model for low-energy electronic states of twisted bilayer graphene, Phys. Rev. Research **1**, 013001 (2019).

[30] F. Guinea and N. R. Walet, Continuum models for twisted bilayer graphene: Effect of lattice deformation and hopping parameters, Phys. Rev. B **99**, 205134 (2019).

[31] D. Massatt, M. Luskin, and C. Ortner, Electronic density of states for incommensurate layers, Multiscale Model. Simul. **15**, 476 (2017).

[32] P. Cazeaux, M. Luskin, and E. B. Tadmor, Analysis of rippling in incommensurate one-dimensional coupled chains, Multiscale Model. Simul. **15**, 56 (2017).

[33] E. Cancès, P. Cazeaux, and M. Luskin, Generalized Kubo formulas for the transport properties of incommensurate 2D atomic heterostructures, J. Math. Phys. **58**, 063502 (2017).

[34] P. Cazeaux, M. Luskin, and D. Massatt, Energy Minimization of Two Dimensional Incommensurate Heterostructures, Arch. Ration. Mech. Anal. **235**, 1289 (2020).

[35] E. Kaxiras and M. S. Duesbery, Free Energies of Generalized Stacking Faults in Si and Implications for the Brittle-Ductile Transition, Phys. Rev. Lett. **70**, 3752 (1993).

[36] M. Phillips and C. S. Hellberg, Commensurate structures in twisted transition metal dichalcogenide heterobilayers, arXiv:1909.02495.

[37] P. K. Mogensen and A. N. Riseth, Optim: A mathematical optimization package for Julia, J.Open Source Softw. **3**, 615 (2018).

[38] K. Zhang and E. B. Tadmor, Structural and electron diffraction scaling of twisted graphene bilayers, J. Mech. Phys. Solids **112**, 225 (2018).

[39] A. Uri, S. Grover, Y. Cao, J. A. Crosse, K. Bagani, D. Rodan-Legrain, Y. Myasoedov, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero, and E. Zeldov, Mapping the twist angle and unconventional Landau levels in magic angle graphene, Nature **581**, 47 (2020).

[40] S. Carr, C. Li, Z. Zhu, E. Kaxiras, S. Sachdev, and A. Kruchkov, Coexistence of ultraheavy and ultrarelativistic Dirac quasiparticles in sandwiched trilayer graphene, Nano Lett. **20**, 3030 (2020).

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

[42] G. Kresse and J. Hafner, Ab initio molecular dynamics for liquid metals, Phys. Rev. B **47**, 558 (1993).

[43] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[44] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Van der Waals Density Functional for General Geometries, Phys. Rev. Lett. **92**, 246401 (2004).

[45] O. A. Vydrov and T. V. Voorhis, Nonlocal van der Waals density functional: The simpler the better, J. Chem. Phys. **133**, 244103 (2010).

[46] J. Klimeš, D. R. Bowler, and A. Michaelides, Van der Waals density functionals applied to solids, Phys. Rev. B **83**, 195131 (2011).

[47] H. Peng, Z.-H. Yang, J. P. Perdew, and J. Sun, Versatile Van Der Waals Density Functional Based on a Meta-Generalized Gradient Approximation, Phys. Rev. X **6**, 041005 (2016).