Editors' Suggestion   Featured in Physics

Theory of phonon spectroscopy with the quantum twisting microscope

Jiewen Xiao,[1] Erez Berg,[1] Leonid I. Glazman ,[2] Francisco Guinea ,[3,4] Shahal Ilani ,[1] and Felix von Oppen [5]

[1]*Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel*
[2]*Department of Physics and Yale Quantum Institute, Yale University, New Haven, Connecticut 06520, USA*
[3]*Imdea Nanoscience, Faraday 9, 28015 Madrid, Spain*
[4]*Donostia International Physics Center, Paseo Manuel de Lardizábal 4, 20018 San Sebastian, Spain*
[5]*Dahlem Center for Complex Quantum Systems and Fachbereich Physik, Freie Universität Berlin, 14195 Berlin, Germany*

We develop a theory of probing phonon modes of van der Waals materials using the quantum twisting microscope. While elastic tunneling dominates the tunneling current at small twist angles, the momentum mismatch between the K points of tip and sample at large twist angles can only be bridged by inelastic scattering. This allows for probing phonon dispersions along certain lines in reciprocal space by measuring the tunneling current as a function of twist angle and bias voltage. We illustrate this modality of the quantum twisting microscope by developing a systematic theory for graphene-graphene junctions. We show that beyond phonon dispersions, the tunneling current also encodes the strength of electron-phonon couplings. Extracting the coupling strengths for individual phonon modes requires careful consideration of various inelastic tunneling processes. These processes are associated with the intralayer and interlayer electron-phonon couplings and appear at different orders in a perturbative calculation of the tunneling current. We find that the dominant process depends on the particular phonon mode under consideration. Our results inform the quest to understand the origin of superconductivity in twisted bilayer graphene and provide a case study for quantum-twisting-microscope investigations of collective modes.

I. INTRODUCTION

Beyond conventional transport experiments, much information on van der Waals materials derives from local probes such as scanning tunneling microscopy (STM) as well as scanning single-electron-transistor and scanning SQUID measurements. A particularly well-studied system is twisted bilayer graphene with its multitude of electronic phases when the twist angle is close to the magic angle [1–3]. Scanning tunneling microscopy has been instrumental in elucidating the effects of correlations on their flat-band dispersion [4–7], of superconductivity [8], and of the correlated insulating states [9]. Compressibility measurements using a scanning single-electron transistor revealed a cascade of flavor-polarized phases [10,11] as well as excess entropy associated with the formation of local moments [12]. It was also instrumental in uncovering anomalous Chern-insulator phases [13]. Scanning-SQUID measurements revealed the formation of Chern mosaics [14].

The quantum twisting microscope (QTM) [15] is a powerful new instrument complementing previously existing local probes. Rather than measuring the local tunneling current at the atomic scale as in scanning tunneling microscopy, it relies on coherent tunneling across a twistable finite-area junction formed at the interface between van der Waals systems placed on a scanning tip with a flat pyramidal top and on a substrate (see Fig. 1). Due to the finite contact area, tunneling conserves crystal momentum modulo reciprocal lattice vectors of the tip and sample layers. Except at special twist angles, umklapp processes involving larger reciprocal lattice vectors

will typically be suppressed. The twist imposes a relative rotation of the dispersions of tip and sample in momentum space. Moreover, the bias voltage introduces a relative shift of the dispersions in energy. Measuring the tunneling current as a function of bias voltage and twist angle will then provide direct signatures of momentum-resolved dispersions. This has been used to explore the electronic dispersion of graphene layers using graphene-graphene junctions as well as the flat-band dispersions of twisted bilayer graphene using junctions of graphene and twisted bilayer graphene [15]. Theoretical work has explored the use of the QTM to probe two-dimensional superconductors [16], spin liquids [17], as well as spin-ordered states close to metal-insulator transitions [18].

Beyond electronic dispersions, the QTM is also exquisitely suited to access momentum-resolved dispersions of the collective excitations of van der Waals systems as shown by a very recent experiment [19]. Whenever the electronic dispersions of tip and sample with their relative twist do not intersect, the momentum mismatch can be bridged by emission of a collective-excitation quantum. Here, we illustrate this modality of the QTM by developing a comprehensive theory of phonon spectroscopy. Our considerations focus on graphene-graphene junctions, but the approach readily generalizes to other junctions and other collective excitations. Along the way, we include analytical results for elastic tunneling between tip and sample for context and comparison.

Due to the semimetallic nature of graphene, its Fermi circles are typically small for relevant gate-induced densities. Thus, inelastic tunneling processes enabled by phonon emission dominate except at the smallest twist angles (i.e., for

$\theta \gtrsim 5°$ in current experiments [19]), and the phonon wave vector is approximately equal to the distance between Dirac points of tip and sample. This implies that the relevant phonon wave vector can be systematically varied by changing the twist angle, allowing for direct measurements of the phonon dispersions.

In addition to the phonon dispersion, the measurements also encode the strength of electron-phonon coupling as a function of wave vector. The phonon spectrum and electron-phonon coupling are important inputs in developing a theory of superconductivity [20–29] and the linear temperature dependence of the resistivity [23,30–37] in magic-angle twisted bilayer graphene. In both cases, there has been substantial debate whether the phenomenon has a conventional origin in the electron-phonon coupling or an underlying exotic mechanism due to electron-electron correlations. Measurements of graphene-graphene junctions in a QTM may be highly relevant in this context as van der Waals coupling tends to lock the interlayer distance between tip and sample layers of a QTM to the interlayer distance of a twisted bilayer. At the same time, QTM measurements cannot access inelastic phonon processes at very small twist angles, where they will be difficult to differentiate from the elastic-tunneling background. Gleaning information on phonon dispersions and electron-phonon couplings in magic-angle twisted bilayer graphene will thus require extrapolation from larger twist angles.

This paper is organized as follows. In Sec. II, we begin by summarizing the results of the detailed calculations in the subsequent sections. This summary section makes our results accessible without the need to read the more technical sections in detail. Section III collects background material and fixes notation. We first introduce the electronic properties of graphene layers (Sec. III A), review tunneling between twisted layers in Secs. III B and III C, and discuss the electrostatics of QTM junctions Sec. III D. Section IV focuses on the elastic tunneling current, giving analytical results for the threshold behaviors of the differential conductance. Section V discusses the phonon modes and the electron-phonon coupling, including both intralayer and interlayer couplings. The theory of phonon spectroscopy is finally addressed in Sec. VI. Following the general expressions for the inelastic-tunneling current in Sec. VI A, we discuss the contributions of the interlayer and intralayer electron-phonon couplings in Secs. VI B and VI C, respectively. We conclude in Sec. VII.

II. OVERVIEW OF RESULTS

A. Tunneling

We illustrate phonon spectroscopy by considering tunneling between two graphene layers. Twisting the tip and sample layers with respect to each other leads to a relative rotation of their Brillouin zones (Fig. 2). In particular, this induces a relative displacement of their K points by equal-length vectors $\mathbf{q}_j$, where $j = 0, 1, 2$ enumerates the three equivalent K points within the Brillouin zone [Fig. 2(b)]. Current flow between tip and sample is dominated by elastic tunneling as long as the Fermi circles of tip and sample intersect [15] [Fig. 2(c)]. This is a consequence of the fact that tunneling between tip and sample conserves crystal momentum modulo reciprocal
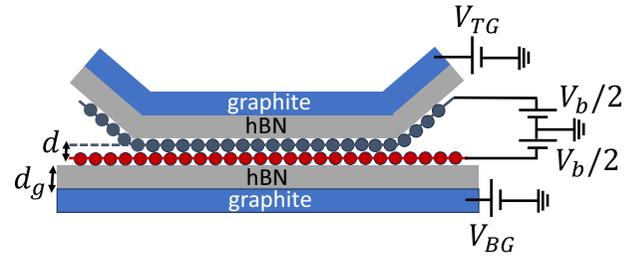
FIG. 1. Schematic setup of QTM with scanning tip (top) and fixed sample (bottom). The electron densities of the van der Waals layers (here, monolayer graphene) on tip and substrate are controlled by independent top (TG) and bottom (BG) gates (here, graphite). The gate electrodes are separated from the van der Waals layers by a gate dielectric (here, hBN). For elastic-tunneling experiments at small twist angles, tip and sample are separated by an additional barrier layer (e.g., WSe$_2$, not shown). No additional barrier is needed when probing inelastic tunneling currents at larger twist angles.

lattice vectors of the graphene layers. At larger twist angles, the momentum mismatch between the electrons in tip and sample requires inelastic processes involving phonon emission, with the phonon wave vector approximately equal to one of the $\mathbf{q}_j$ [Fig. 2(d)]. (Here, we assume that temperature is sufficiently low that phonon absorption can be neglected.) Bias voltages, at which eV_b equals the energy $\hbar\omega_{r,\mathbf{q}_j}$ of a phonon mode r, are associated with threshold features in the current-voltage characteristic. Tracking these inelastic-tunneling features as a function of twist angle θ (and hence
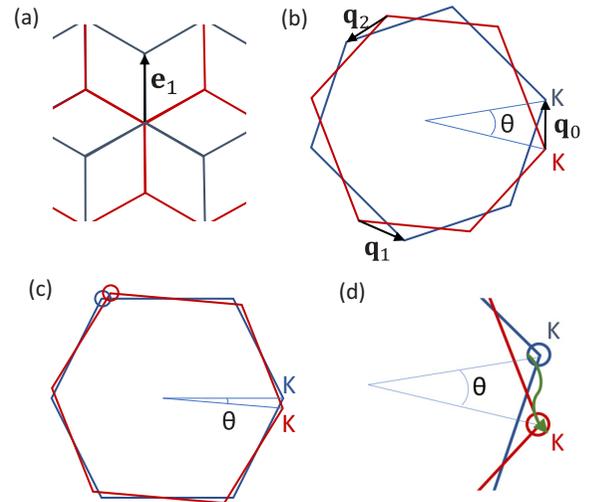
FIG. 2. (a) Bernal-stacked configuration (two layers shifted relative to each other by the bond vector $\mathbf{e}_1$) used as reference configuration for twisted layers. Top layer (tip, blue); bottom layer (sample, red). (b) Definition of vectors $\mathbf{q}_j$ connecting the K points of the Brillouin zones of tip and sample. (c) Overlapping Fermi circles of tip and sample at small twist angle θ. In this limit, current is dominated by elastic tunneling between tip and sample. (d) Tunneling at twist angles with mismatched Fermi circles occurs by phonon emission (green arrow connecting states on Fermi circles of tip and sample due to momentum conservation). The phonon wave vector $\mathbf{Q}$ is approximately equal to one of the $\mathbf{q}_j$.
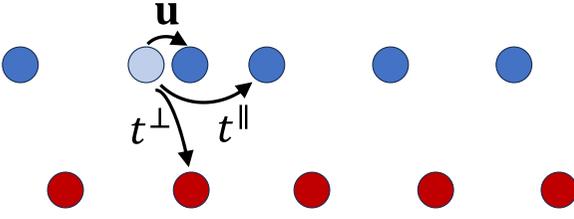
FIG. 3. Electron-phonon coupling in twisted graphene layers. Sketch of tip (blue) and sample (red) layers with phonon-induced atomic displacement **u** (illustrated for the in-plane displacement of one atom in the tip). The displacements **u** modify the bond lengths and consequently the hopping amplitudes $t^{\parallel}$ within the layer (intralayer electron-phonon coupling $\mathcal{H}_{\text{intra}}$) and between layers $t^{\perp}$ (interlayer electron-phonon coupling $\mathcal{H}_{\text{inter}}$).

$\mathbf{q}_j$) allows for mapping out the phonon dispersion along certain lines in momentum space. (Umklapp processes lead to additional sharp elastic scattering peaks at larger, commensurate twist angles due to overlap of Fermi surfaces in higher Brillouin zones [15].)

Electron-phonon coupling emerges from several mechanisms [38]. Modifications of the hybridization of carbon orbitals associated with phonon-induced changes in the bond lengths contribute as illustrated in Fig. 3. The intralayer coupling $\mathcal{H}_{\text{intra}}$ originates in changes of the hopping amplitudes within the layer, corresponding to electron-phonon coupling within the individual graphene layers. In twisted bilayers, phonons also affect the amplitude of interlayer tunneling, giving rise to the interlayer electron-phonon coupling $\mathcal{H}_{\text{inter}}$. In addition, longitudinal acoustic phonons lead to a local expansion or contraction of the lattice, which shifts the chemical potential [38,39], referred to as deformation potential. We find that the coupling mechanism dominating the tunneling current differs between phonon modes, so that it is important to account for the various couplings to understand phonon signatures in QTM measurements.

Twisted graphene layers are described by the Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_T + \mathcal{H}_{\text{intra}} + \mathcal{H}_{\text{inter}}. \qquad (1)$$

Here, $\mathcal{H}_0$ describes the two uncoupled graphene layers, including their phonon modes, and $\mathcal{H}_T$ accounts for the (purely electronic) interlayer tunneling. Retaining terms to first order in the interlayer tunneling, we can then expand the $\mathcal{T}$ matrix for electron scattering between tip and sample layers as

$$\mathcal{T} = \mathcal{H}_T + \mathcal{H}_{\text{inter}} + \mathcal{H}_T \mathcal{G}_0 \mathcal{H}_{\text{intra}} + \mathcal{H}_{\text{intra}} \mathcal{G}_0 \mathcal{H}_T + \cdots. \quad (2)$$

The first term $\mathcal{H}_T$ on the right-hand side gives rise to the elastic-tunneling processes at small twist angles. Inelastic tunneling involving the emission of a phonon due to the interlayer electron-phonon coupling is described by the second term. Both of these processes can be described in the lowest order in a Fermi-golden-rule calculation of the tunneling current. The remaining two terms describe higher-order inelastic processes involving both electron tunneling and intralayer electron-phonon coupling, with the Green function $\mathcal{G}_0 = [E - \mathcal{H}_0]^{-1}$ of the uncoupled layers accounting for the energy denominators of the virtual intermediate states. We find that the inelastic-tunneling current can be dominated by one

electron-phonon coupling or the other, despite their different orders in perturbation theory.

B. Electrostatics and characteristic voltages

Due to the small quantum capacitance of the graphene layers, a bias voltage applied between tip and sample will predominantly modify the chemical potentials. This is accompanied by a smaller relative shift $e\phi$ in energy of the Dirac points of tip and sample due to the electrostatic potential difference ϕ. The ratio of the shifts in electrostatic and chemical potentials is of order $q_{\text{TF}}d$, where q_{TF} is the Thomas-Fermi screening wave vector of graphene and d the distance between tip and sample layers [see Eq. (63) below and the discussion around it for more details]. In our analytical calculations, we focus on the limit in which $q_{\text{TF}}d \ll 1$ (small quantum capacitance) and assume overall charge neutrality. Then, tip and sample have opposite chemical potentials $\pm\mu$, so that the bias voltage (i.e., the difference in electrochemical potential) is $eV_b = 2\mu + e\phi$ with $e\phi \ll 2\mu$.

At small bias voltages, the electrostatic shift ϕ can be neglected and the Dirac points of tip and sample are aligned in energy, but offset in momentum by $\mathbf{q}_j$. The offset Dirac cones intersect at energies that are larger in magnitude than $\hbar v_D q_0/2$, where v_D is the Dirac velocity. As the chemical potentials of tip and sample are equal to $\pm eV_b/2$, this leads to a characteristic voltage of

$$eV_b^* = \hbar v_D q_0 \qquad (3)$$

for elastic scattering [Figs. 4(a) and 4(b)], with current flow due to elastic tunneling only possible for bias voltages $V_b > V_b^*$.

At larger bias voltages, the electrostatic potential leads to an appreciable relative shift of the Dirac points in energy. When this shift $e\phi$ becomes of order $\hbar v_D q_j$, there is approximate nesting of the Dirac dispersions of tip and sample [Fig. 4(c)]. Nesting defines a second characteristic voltage V_b^{**} through the condition

$$e\phi(V_b^{**}) = \hbar v_D q_0. \qquad (4)$$

Note that there is a sharp drop in the elastic-tunneling current as the bias increases past this characteristic voltage.

The dispersions depicted in Figs. 4(a) and 4(b) neglect interlayer tunneling, which opens gaps at the crossing points of the Dirac dispersions of the two layers. The magnitude of the resulting gaps is given by the interlayer tunneling strength w. Our perturbative approach requires that these gaps be small compared to eV_b^*. This is satisfied for twist angles

$$\theta > \frac{w}{v_D|\mathbf{K}|}. \qquad (5)$$

(Here, **K** denotes the location of the K point as measured from the Γ point.) This is equivalent to the condition that the twist angle be larger than the magic angle of twisted bilayer graphene.

For twist angles satisfying Eq. (5), inelastic tunneling processes are relevant provided that the phonon energy $\hbar\omega_{r,\mathbf{q}_j}$ is smaller than eV_b^*. In this case, the tunneling current at bias voltages $V_b < V_b^*$ is entirely due to inelastic processes, allowing for measurements of the phonon dispersions and the
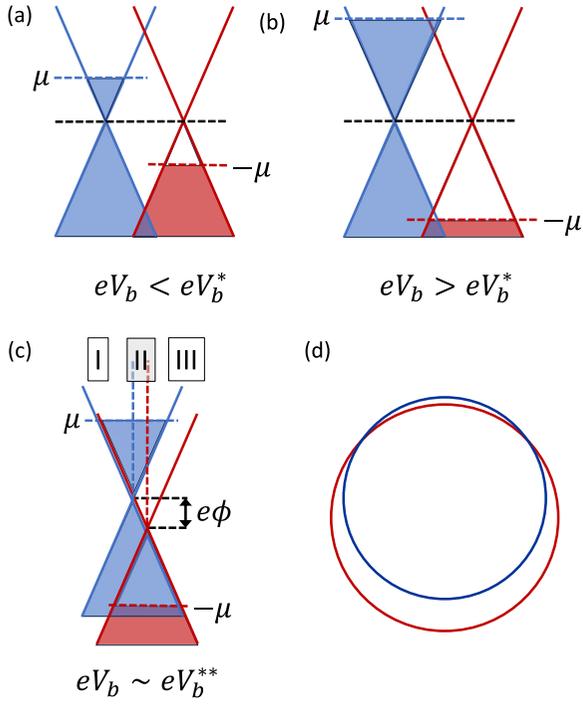
FIG. 4. Characteristic voltages for elastic QTM tunneling at small twist angles, assuming overall charge neutrality of tip and sample as well as small quantum capacitance. (a), (b) Tunneling onsets at eV_b^* [Eq. (67)], beyond which occupied states in the Dirac cone of tip overlap with empty states in the Dirac cone of the sample. Illustration of Dirac cones of tip and sample separated in momentum by q_0 and their fillings for a voltage (a) below and (b) above V_b^*. (c) Dirac cones of tip and sample for bias voltages of order V_b^{**} exhibiting nesting due to the relative electrostatic shift $e\phi$ of the Dirac points. I, II, III refer to momentum regions to the left of the blue dashed line (I), between blue and red dashed lines (II), and to the right of the red dashed line (III). (d) Constant-energy cut through the shifted Dirac-cone dispersions of tip and sample for bias voltages close to the nesting condition illustrated in (c).

electron-phonon coupling. For optical phonons with frequencies ~200 meV, this gives a minimal twist angle of $1°$– $2°$. For acoustic phonons, the condition is less stringent due to their smaller energy. There exists an energy window as long as the Dirac velocity is large compared to the mode velocity of the acoustic phonons, which is always the case away from the magic angle.

C. Elastic tunneling

Coherent tunneling across an extended contact (area $\Omega = L_x L_y$) is central to the QTM [15]. To bring out the importance of coherence, we consider a reference problem of Ω/λ_F^2 parallel incoherent local tunneling contacts. The Fermi wavelength $\lambda_F = 2\pi/\kappa_F$ defines the characteristic size of the local contacts, so that the coarse-grained local tunneling of a contact at **R** can be approximated as $w\lambda_F^2\delta(\mathbf{r} - \mathbf{R})$. Here, w denotes the underlying tunneling amplitude of the extended contact with units of energy. This gives an incoherent tunneling current of order

$$(2\pi e/\hbar)(w\lambda_F^2)^2 v(\mu)[N_f v(\mu)eV_b] \qquad (6)$$

per tunnel contact, where $v(\mu) = \mu/(2\pi\hbar^2 v_D^2)$ is the density of states per flavor at the chemical potential. This expression for the current is composed of the Fermi-golden-rule rate for tunneling of an incident electron and the number of incident electrons within the voltage window accounting for the spin and valley degeneracy $N_f = 4$. With the linear graphene dispersion $E = \hbar v_D \kappa$ and the relation $\mu = \hbar v_D \kappa_F = eV_b/2$ between voltage and chemical potential (for sufficiently small voltages, so that the electrostatic potential ϕ can be neglected), this gives a differential-conductance scale of order

$$G_{\text{incoh}} = \frac{e^2}{h}\frac{N_f w^2 \Omega}{\hbar^2 v_D^2} \qquad (7)$$

for an array of Ω/λ_F^2 parallel contacts.

We find that G_{incoh} provides a convenient scale for expressing our results for the tunneling current in the QTM. Elastic tunneling sets in at voltages $V_b > V_b^*$. The differential conductance in the vicinity of V_b^* is given by (see Sec. IV A)

$$\frac{dI}{dV_b} = 3\sqrt{2}G_{\text{incoh}}\sqrt{\frac{V_b^*}{V_b - V_b^*}}\theta(V_b - V_b^*), \qquad (8)$$

exhibiting a square-root divergence in the bias voltage. As $eV_b^* = \hbar v_D q_0$, measuring the threshold voltage as a function of twist angle (and hence $\mathbf{q}_0$) can be viewed as direct spectroscopy of the Dirac dispersion of the graphene layers.

Similarly, the current near the threshold voltage V_b^{**} for nesting is (see Sec. IV B)

$$I = \frac{3}{2}G_{\text{incoh}}\frac{(V_b^{**})^2}{\sqrt{V_b^*|\phi(V_b^{**}) - \phi(V_b)|}}\theta(V_b^{**} - V_b). \qquad (9)$$

This implies a large negative differential conductance at voltage V_b^{**}. The divergence at V_b^{**} is a consequence of the assumptions of strictly linear dispersion and momentum-conserving tunneling. The tunneling current originates from momenta, where the Dirac cones of tip and sample touch, so that the contributions to the current diverge concurrently at all energies within the bias window [Figs. 4(c) and 4(d)]. The divergence of the current at V_b^{**} will thus be cut off by nonlinear corrections to the dispersion relation as well as spatial inhomogeneities, which lift strict momentum conservation.

D. Inelastic tunneling

When the Fermi circles of tip and sample no longer intersect at larger twist angles, momentum-conserving tunneling at low bias voltages is enabled by emission (or absorption) of phonons. Neglecting the small electronic momenta relative to the K points, the relevant phonons have wave vectors equal to the wave vectors $\mathbf{q}_j$ connecting the K points of the tip and the sample [Fig. 2(b)]. Consequently, the inelastic tunneling channels for the various phonon modes r open beyond the threshold voltages $eV_b = \hbar\omega_{r,\mathbf{q}_j}$, where $\omega_{r,\mathbf{q}_j}$ denotes the phonon frequency of mode r at wave vector $\mathbf{q}_j$. This appears in the tunneling current as steps in the differential conductance and consequently as peaks in d^2I/dV_b^2.

For out-of-plane (flexural) phonons, the electron-phonon coupling is dominated by the interlayer electron-phonon coupling $\mathcal{H}_{\text{inter}}$. While $\mathcal{H}_{\text{inter}}$ depends linearly on the phonon displacements, the intralayer coupling $\mathcal{H}_{\text{intra}}$ to out-of-plane

phonons has a weaker quadratic dependence on the mode displacements. The dominant coupling due to $\mathcal{H}_{\text{inter}}$ emerges directly from the change in the interlayer tunneling amplitude due to the out-of-plane component of the atomic displacements and is thus of order $\ell_{\text{ZPM}}(\partial w/\partial d)$. Here, d is the interlayer distance and[1]

$$\ell_{\text{ZPM}} = \left(\hbar/M\omega_{r,\mathbf{q}_0}\right)^{1/2} \tag{10}$$

(with M the mass per unit cell of a single layer) the zero-point-motion amplitude of the relevant out-of-plane phonon mode. We then find [Eq. (104)]

$$\frac{d^2 I}{dV_b^2} \sim G_{\text{incoh}}(\kappa_F a)^2 \left(\ell_{\text{ZPM}} \frac{\partial \ln w}{\partial d}\right)^2 \delta\left(V_b - \hbar\omega_{r,\mathbf{q}_0}/e\right), \tag{11}$$

where the factor $(\kappa_F a)^2$ accounts for the size of the graphene Fermi circles. (a is the carbon-carbon bond length of graphene.)

There are two out-of-plane phonon modes r, which we denote by ZA and ZO, respectively. The mode ZA is an optical mode in the out-of-plane direction, i.e., the atomic displacements are antisymmetric between the two layers. At the same time, it is acoustic in nature within the plane, i.e., the long-wavelength displacements of the two sublattices are symmetric within each layer. Even when mechanical coupling between the tip and sample layers is negligible, the mode frequencies saturate to a constant at small wave vectors due to mechanical coupling between the graphene layers and their substrates. At larger wave vectors, the mode frequencies are quadratic in the phonon wave vector akin to the flexural modes of free-standing graphene membranes. The ZO mode is also optical in nature within the plane and thus characterized by larger mode frequencies for all wave vectors.

The interlayer electron-phonon coupling to in-plane phonons has a different origin. A relative displacement $\mathbf{u}$ of the two layers does not modify the magnitude of the interlayer tunneling, but changes its phase by $\exp(i\mathbf{K} \cdot \mathbf{u})$ [40]. If $\mathbf{u}$ originates from a phonon displacement, the phase becomes time dependent, akin to a time-dependent vector potential. Thus, $\mathcal{H}_{\text{inter}}$ can be viewed as being due to a synthetic electric field [41–43]. Importantly, the coupling is maximal for mode displacements $\mathbf{u}$, which are parallel to the vector $\mathbf{K}$ of the K point. At small twist angles, the phonon wave vector $\mathbf{q}_0$ is approximately perpendicular to $\mathbf{K}$, so that the coupling is predominantly to transverse phonon modes. This is illustrated in Fig. 5. Unlike the intralayer electron-phonon coupling, this coupling does not go to zero in the long-wavelength limit. Transverse phonons, both acoustic (TA) and optical (TO), will thus involve an interlayer electron-phonon coupling of order $w\mathbf{K} \cdot \mathbf{u} \sim w(\ell_{\text{ZPM}}/a)\cos\theta$ from expanding the phase factor to linear order. Combining this with the factor accounting for the size of the graphene Fermi circles gives

$$\left.\frac{d^2 I}{dV_b^2}\right|_{\text{inter,T}} \sim G_{\text{incoh}}(\kappa_F \ell_{\text{ZPM}})^2 \cos^2\theta \, \delta\left(V_b - \hbar\omega_{r,\mathbf{q}_0}/e\right) \tag{12}$$
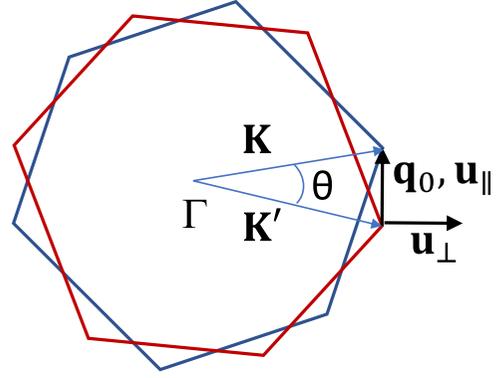
FIG. 5. Illustration of the interlayer electron-phonon coupling. At small twist angles θ, the phonon wave vector $\mathbf{q}_0$ is almost perpendicular to the vectors $\mathbf{K}$ and $\mathbf{K}'$ of the K points of tip and substrate, respectively. As a result, the phonon displacements are approximately parallel to $\mathbf{K}$ and $\mathbf{K}'$ for transverse phonons ($\mathbf{u}_\perp$) and perpendicular for longitudinal phonons ($\mathbf{u}_\parallel$).

for the contribution of the interlayer electron-phonon coupling for transverse acoustic and optical phonons. The expression for longitudinal acoustic (LA) and optical (LO) phonons is similar, differing only in its twist-angle dependence,

$$\left.\frac{d^2 I}{dV_b^2}\right|_{\text{inter,L}} \sim G_{\text{incoh}}(\kappa_F \ell_{\text{ZPM}})^2 \sin^2\theta \, \delta\left(V_b - \hbar\omega_{r,\mathbf{q}_0}/e\right). \tag{13}$$

The full result for both longitudinal and transverse phonons is given in Eq. (104).

The intralayer electron-phonon coupling $\mathcal{H}_{\text{intra}}$ gives contributions of the same order for acoustic phonons, albeit with a different twist-angle dependence. We can estimate the relevant electron-phonon coupling by noting that $\mathcal{H}_{\text{intra}}$ originates in the dependence of $t^\parallel$ on the bond length. Thus, the intralayer electron-phonon coupling is of order $\frac{\partial t^\parallel}{\partial a}\ell_{\text{ZPM}}(q_0 a)$ for acoustic phonons. Here, the last factor accounts for the fact that the relative displacements of neighboring atoms is suppressed in the long-wavelength limit. Combining this with the fact that the tunneling Hamiltonian is of order w and the energy denominators are of order $\hbar v_D q_0$, we find that the contribution of $\mathcal{H}_{\text{intra}}$ to the $\mathcal{T}$ matrix is of order

$$\frac{w(\partial t^\parallel/\partial a)\ell_{\text{ZPM}}(q_0 a)}{\hbar v_D q_0}. \tag{14}$$

With the estimates $\partial t^\parallel/\partial a \sim t^\parallel/a$ and $\hbar v_D/a \sim t^\parallel$, this becomes of order $w(\ell_{\text{ZPM}}/a)$, which is indeed of the same order as the contribution of $\mathcal{H}_{\text{inter}}$. At the same time, $\mathcal{H}_{\text{intra}}$ has only weak twist-angle dependence and is of the same order for longitudinal and transverse phonons due to the triad of bond vectors. This yields

$$\left.\frac{d^2 I}{dV_b^2}\right|_{\text{intra}} \sim G_{\text{incoh}}(\kappa_F \ell_{\text{ZPM}})^2 \delta\left(V_b - \hbar\omega_{r,\mathbf{q}_0}/e\right) \tag{15}$$

for acoustic phonons.

For optical phonons, the intralayer electron-phonon coupling $\mathcal{H}_{\text{intra}}$ does not involve the suppression factor $q_0 a$, so

---

[1]Notice that the actual real-space amplitude of the zero-point fluctuations for a particular phonon mode is smaller than ℓ_{ZPM} by a factor $1/\sqrt{N}$.

TABLE I. Summary of results for the step in dI/dV at the inelastic threshold $eV_b = \hbar\omega_{\mathbf{q}_0}$ for various phonon modes. Expressions given in the table give the parametric dependencies when multiplied by $G_{\text{incoh}}\theta(eV_b - \hbar\omega_{\mathbf{q}_0})$. The contribution of $\mathcal{H}_{\text{inter}}$ is obtained in first-order perturbation theory, while the contribution of $\mathcal{H}_{\text{intra}}$ follows from a second-order calculation treating both $\mathcal{H}_{\text{intra}}$ and $\mathcal{H}_T$ perturbatively [see Eq. (2)]. The equation numbers given in the table point to the full expressions in the text. We include the results for electron-phonon coupling originating from modifications of hopping amplitudes by phonons. For LA phonons, there is an additional contribution due to the deformation potential (strength D), which can be obtained from the contribution of $\mathcal{H}_{\text{intra}}$ by the replacement $\partial t^{\parallel}/\partial \ln a \leftrightarrow D$ (see Sec. V B 3). We note that the amplitude ℓ_{ZPM} of the zero-point motion contains an implicit dependence on the frequency of the phonon mode under consideration [cf. Eq. (10)].

dI/dV	Acoustic		Optical		Out of plane
$[G_{\text{incoh}}\theta(eV_b - \hbar\omega_{\mathbf{q}_0})]$	LA	TA	LO	TO	ZA and ZO
$\mathcal{H}_{\text{inter}}$ Eq. (104)	$(\kappa_F \ell_{\text{ZPM}})^2 \sin^2\theta$	$(\kappa_F \ell_{\text{ZPM}})^2 \cos^2\theta$	$(\kappa_F \ell_{\text{ZPM}})^2 \sin^2\theta$	$(\kappa_F \ell_{\text{ZPM}})^2 \cos^2\theta$	$(\kappa_F a)^2 \left(\ell_{\text{ZMP}} \frac{\partial \ln w}{\partial d}\right)^2$
$\mathcal{H}_{\text{intra}}$ Eq. (116)	$(\kappa_F \ell_{\text{ZPM}})^2$	$(\kappa_F \ell_{\text{ZPM}})^2$	$\left(\frac{\kappa_F \ell_{\text{ZPM}}}{q_0 a}\right)^2$	$\left(\frac{\kappa_F \ell_{\text{ZPM}}}{q_0 a}\right)^2$	Higher order

that

$$\frac{d^2 I}{dV_b^2}\bigg|_{\text{intra}} \sim G_{\text{incoh}} \left(\frac{\kappa_F \ell_{\text{ZPM}}}{q_0 a}\right)^2 \delta\left(V_b - \hbar\omega_{r,\mathbf{q}_0}/e\right). \quad (16)$$

At long wavelengths, this dominates over the contribution of $\mathcal{H}_{\text{inter}}$ in Eqs. (12) and (13). The full result for the contribution of $\mathcal{H}_{\text{intra}}$ for both acoustic and optical phonons is given in Eq. (116).

We summarize these results for the inelastic contributions to the tunneling current in Table I. We conclude this section by addressing the limit of small twist angles θ (but sufficiently large that elastic scattering can be neglected). For acoustic modes, $q_0 \sim \theta$ and $\omega_{r,\mathbf{q}_0} \sim q_0$, which implies $\ell_{\text{ZPM}} \sim 1/\sqrt{q_0}$. For overall charge neutrality, the threshold condition $eV_b = \hbar\omega_{q_0}$ implies that $\kappa_F \sim q_0$, while κ_F approaches a nonzero constant for small q_0 away from overall charge neutrality. We thus find that for acoustic modes, $d^2 I/dV_b^2$ diverges as $1/q_0$ at small twist angles away from charge neutrality. For optical modes in the limit $q_0 \to 0$, κ_F and ℓ_{ZPM} approach nonzero constants, so that one finds a $1/q_0^2$ divergence both at and away from charge neutrality. The divergence is cut off by the condition for the twist angle in Eq. (5), which corresponds to $q_0 \gtrsim w/v_D$. A further limitation specific to optical phonons is set by Eq. (106).

E. Scattering picture

The current can be estimated using a scattering approach, which gives insight beyond the Fermi-golden-rule calculations presented in subsequent sections. We consider a scattering geometry, in which electrons in the tip impinge on the tip-sample junction in the y direction, with channels defined by wave vectors k_x (Fig. 6). When normalizing the incoming and outgoing states to unit flux in the y direction, the tunneling current is

$$I = \frac{e}{h} \sum_{k_x k_x'} \int dE\, dE'\, T_{k_x k_x'}(E, E') f_\mu(E)[1 - f_{-\mu}(E')], \quad (17)$$

where we assume $V_b > 0$ and zero temperature. Here, $T_{k_x k_x'}(E, E')$ is the probability density that an electron in channel k_x of the tip impinging on the junction at energy E is scattered into channel k_x' of the sample at energy E'.

For elastic tunneling, we can approximate

$$T_{k_x k_x'}(E, E') \sim \frac{w^2}{\hbar^2 |v_y v_y'|} \left|\int_\Omega d\mathbf{r} \frac{e^{i\mathbf{k}\cdot\mathbf{r} - i\mathbf{k}'\cdot\mathbf{r}}}{L_x}\right|^2 \delta(E - E') \quad (18)$$

to lowest order in the tunneling amplitude (Born approximation). Here, we defined the mode velocities $v_y = (1/\hbar)(\partial E/\partial k_y)$ and $v_y' = (1/\hbar)(\partial E'/\partial k_y')$ of tip and sample, respectively. Focusing on order-of-magnitude estimates for small twist angles, we have suppressed the sublattice structure of the graphene wave functions and consider only momentum-conserving tunneling at the K point, i.e., we neglect umklapp processes. Evaluating the integral and taking the limits of
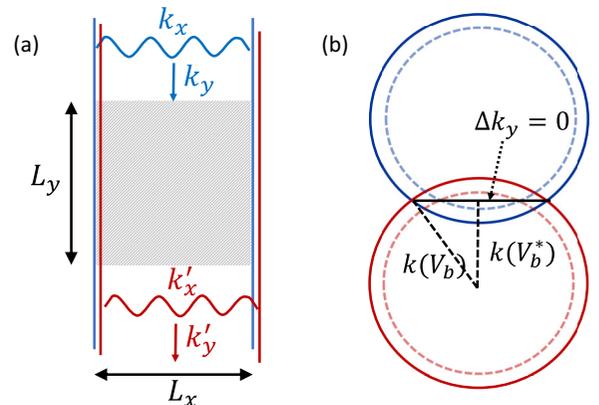
FIG. 6. (a) Scattering geometry of the QTM junction between tip (blue) and sample (red). Plane waves in the tip layer with wave vector k_y in channel k_x impinge on the tunneling contact (black shaded area) and are transmitted into plane waves with momentum k_y' in channel k_x' in the sample. (b) Constant-energy lines of the dispersions of tip and sample for bias voltage V_b close to the characteristic bias voltage V_b^* (full colored lines: energy $E = eV_b$; dashed colored lines: $eV_b > E > eV_b^*$). Tunneling occurs at the crossing points of the dispersions of tip (blue) and sample (red) along the line $\Delta k_y = 0$ (black). The length $2k_{\text{max}}$ of this line at a given bias voltage V_b can be obtained from the indicated right triangle.

large L_x and L_y gives

$$T_{k_x k_x'}(E, E') \sim \frac{w^2}{\hbar^2 |v_y v_y'|} \delta_{k_x, k_x'} L_y \delta(k_y - k_y') \delta(E - E'), \quad (19)$$

which makes the momentum-conserving nature of tunneling explicit. Here, $k_y = k_y(k_x, E)$ and $k_y' = k_y'(k_x', E')$ are determined by the electron dispersions in tip and sample, respectively. Inserting $T_{k_x k_x'}(E, E')$ into the expression for the current and performing the energy integrals, we find

$$I \sim \int_{\Delta k_y = 0} dk_x \frac{e w^2 \Omega}{\hbar^2 |v_y - v_y'|} \theta(k_{max} - |k_x|). \quad (20)$$

For convenience, we temporarily measure k_x from the line connecting the Dirac points of tip and sample. The integral is over the line defined by $\Delta k_y = k_y - k_y' = 0$ and we used that $\partial \Delta k_y / \partial E = 1/\hbar v_y - 1/\hbar v_y'$.

For voltages close to V_b^*, the dispersions of tip and sample are illustrated in Fig. 6(b). The condition $\Delta k_y = 0$ enforces $k_y = k_y' = 0$, so that $v_y' = -v_y \simeq v_D$. Moreover, one reads off $\hbar v_D k_{max} = [(eV_b/2)^2 - (eV_b^*/2)^2]^{1/2}$. Inserting these relations into Eq. (20), we recover the parametric dependencies in Eq. (8).

For bias voltages near V_b^{**}, the dispersions of tip and sample are illustrated in Fig. 4(d). Approximate nesting of the Fermi circles of tip and sample implies that v_y is close to v_y', leading to a strongly enhanced current. Nesting occurs for any energy E within the bias window, with the current dominated by energies $|E| > e\phi(V_b^{**})$. For these energies, one readily estimates

$$\frac{|v_y - v_y'|}{v_D} \sim \left(\frac{e\phi(V_b^{**})}{|E|}\right)\left(\frac{\hbar v_D k_x}{E}\right)^2, \quad (21)$$

accounting for the fact that $v_y - v_y'$ is nonzero due to $e\phi(V_b^{**})$ and increases from zero symmetrically about $k_x = 0$. Almost nesting crossing points of the dispersions of tip and sample exist only for $\phi < \phi(V_b^{**})$ (and thus $V_b < V_b^{**}$). At energy E, these crossing points satisfy

$$\frac{\phi - \phi(V_b^{**})}{\phi(V_b^{**})} \sim \left(\frac{\hbar v_D k_x}{E}\right)^2, \quad (22)$$

so that

$$k_{max} \sim \frac{eV_b}{\hbar v_D} \left(\frac{|\phi - \phi(V_b^{**})|}{\phi(V_b^{**})}\right)^{1/2}. \quad (23)$$

These relations combined with Eq. (20) reproduce the parametric dependencies in Eq. (9).

The scattering approach is readily extended to inelastic current flow between tip and sample. Here, we consider the contribution of the interlayer electron-phonon coupling. The contribution of the intralayer electron-phonon coupling can be obtained by replacing the relevant electron-phonon coupling strength along the lines sketched in Sec. II D above.

The plane-wave factor $\frac{1}{\sqrt{\Omega}} e^{i\mathbf{Q} \cdot \mathbf{r}}$ of the phonon introduces the phonon wave vector $\mathbf{Q}$ into the momentum-conservation factors. Moreover, we account for the phonon energy $\hbar\omega_\mathbf{Q}$ in the energy balance and use that in addition to the interlayer tunneling amplitude w, the characteristic strength of the interlayer electron-phonon coupling is controlled by the ratio ℓ_{ZPM}/a of the atomic zero-point motion ℓ_{ZPM} and the

graphene bond length a (see Sec. V for a detailed discussion). We can then estimate

$$T_{k_x k_x'}(E, E') \sim \frac{1}{N} \sum_\mathbf{Q} \frac{w^2}{\hbar^2 v_y v_y'} \left(\frac{\ell_{ZPM}}{a}\right)^2$$
$$\times \delta_{k_x, k_x' + Q_x} L_y \delta(k_y - k_y' - Q_y) \delta(E - E' - \hbar\omega_\mathbf{Q}), \quad (24)$$

with $N = \Omega/\Omega_{uc}$ being the number of unit cells of area Ω_{uc}.

For sufficiently large twist angle $q_0 \gg \kappa_F$, we approximate the phonon frequency $\omega_\mathbf{Q} \simeq \omega_{\mathbf{q}_0}$ [Fig. 2(d)]. We can then perform the sum over $\mathbf{Q}$ to obtain

$$T_{k_x k_x'}(E, E') \sim \frac{L_y^2}{N} \frac{w^2}{\hbar^2 v_y v_y'} \left(\frac{\ell_{ZPM}}{a}\right)^2 \delta(E - E' - \hbar\omega_{\mathbf{q}_0}). \quad (25)$$

For bias voltages close to $eV_b = \hbar\omega_{\mathbf{q}_0}$, we have

$$\frac{d}{dV_b} f_\mu(E)[1 - f_{-\mu}(E')]\delta(E - E' - \hbar\omega_{\mathbf{q}_0})$$
$$\simeq e\theta(eV_b - \hbar\omega_{\mathbf{q}_0})\delta(E - \mu)\delta(E' + \mu). \quad (26)$$

The δ functions constrain the energies of the initial and final states to the Fermi circles of tip and sample. Thus, we can approximate $|v_y v_y'| \sim v_D^2$ and the sums over k_x and k_x' each contribute a factor of the order of the number of channels, $2\kappa_F L_x$. Collecting factors into Eq. (17), this reproduces Eqs. (12) and (13) up to the twist-angle dependence, which we dropped in the estimate of the electron-phonon coupling.

III. TWISTED GRAPHENE LAYERS: ELECTRONIC STATES

For completeness and for fixing notation, we briefly review some elements of the electronic properties of graphene [39] and of tunneling between twisted layers [40]. We also include a discussion of the electrostatics of graphene-graphene junctions in the QTM [15].

A. Graphene

Each of the two graphene layers is described by a tight-binding Hamiltonian

$$H = -t^{\|} \sum_\mathbf{R} \sum_{j=1}^3 \{|\mathbf{R}\rangle \langle \mathbf{R} + \mathbf{e}_j| + |\mathbf{R} + \mathbf{e}_j\rangle \langle \mathbf{R}|\}, \quad (27)$$

where the $\mathbf{e}_j$ denote the three bond vectors

$$\mathbf{e}_1 = a\begin{pmatrix} 0 \\ 1 \end{pmatrix}; \quad \mathbf{e}_{2/3} = a\begin{pmatrix} \mp\sqrt{3}/2 \\ -1/2 \end{pmatrix} \quad (28)$$

emanating from an A site to the three nearest-neighbor B sites (Fig. 7). The sum is over the sites $\mathbf{R}$ of sublattice A. In the Bloch basis defined through

$$|\mathbf{k}, \alpha\rangle = \frac{1}{\sqrt{N}} \sum_\mathbf{R} e^{i\mathbf{k} \cdot (\mathbf{R} + \boldsymbol{\tau}_\alpha)} |\mathbf{R} + \boldsymbol{\tau}_\alpha\rangle \quad (29)$$

($\alpha = A, B$ denotes the sublattice, $\boldsymbol{\tau}_A = 0$, and $\boldsymbol{\tau}_B = \mathbf{e}_1$), the Hamiltonian takes the form

$$H_\mathbf{k} = \begin{pmatrix} 0 & -t^{\|} \sum_j e^{i\mathbf{k} \cdot \mathbf{e}_j} \\ -t^{\|} \sum_j e^{-i\mathbf{k} \cdot \mathbf{e}_j} & 0 \end{pmatrix}. \quad (30)$$
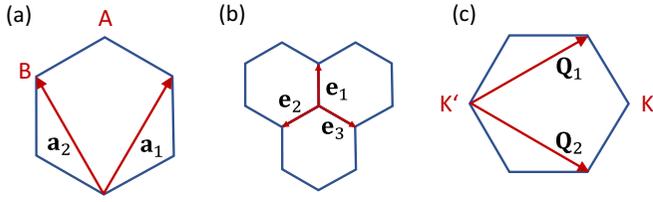
FIG. 7. Honeycomb lattice of graphene layers. (a) A and B sublattices and lattice vectors $\mathbf{a}_{1/2}$ of direct lattice. (b) Bond vectors $\mathbf{e}_j$ of direct lattice. (c) Brillouin zone with $\mathbf{K}$ and $\mathbf{K}'$ points and reciprocal lattice vectors $\mathbf{Q}_{1/2}$. We also occasionally use $\mathbf{Q}_0 = \mathbf{0}$.

The band structure

$$E_{\mathbf{k},\pm} = \pm t^{\parallel}\sqrt{A_{\mathbf{k}}^2 + B_{\mathbf{k}}^2} \tag{31}$$

has valence $(-)$ and conduction $(+)$ bands with the lattice-periodic part $|\mathbf{k}, \pm\rangle = (u_{A,\mathbf{k},\pm}, u_{B,\mathbf{k},\pm})^T$ of the Bloch functions given by

$$u_{A,\mathbf{k},\pm} = \frac{1}{\sqrt{2}}; \quad u_{B,\mathbf{k},\pm} = \mp\frac{1}{\sqrt{2}}e^{-i\gamma_{\mathbf{k}}}. \tag{32}$$

Here, we define $A_{\mathbf{k}} = \sum_j \cos(\mathbf{k} \cdot \mathbf{e}_j)$ and $B_{\mathbf{k}} = \sum_j \sin(\mathbf{k} \cdot \mathbf{e}_j)$ as well as the phase $\gamma_{\mathbf{k}} = \arctan(B_{\mathbf{k}}/A_{\mathbf{k}})$. For graphene lattice vectors $\mathbf{a}_{1/2} = a[\pm\sqrt{3}/2, 3/2]$, the reciprocal lattice is spanned by the vectors (see Fig. 7)

$$\mathbf{Q}_{1/2} = \frac{4\pi}{3a}\begin{pmatrix} \sqrt{3}/2 \\ \pm 1/2 \end{pmatrix}. \tag{33}$$

We occasionally find it useful to define $\mathbf{Q}_0 = \mathbf{0}$.

We measure wave vectors $\mathbf{k}$ from the Γ point. For states close to the K point at $\mathbf{K} = \frac{4\pi}{3a}(1/\sqrt{3}, 0)$ or the K' point at $-\mathbf{K}$, the electron dispersion simplifies to the Dirac form $E_{\boldsymbol{\kappa},\pm} = \pm\hbar v_D|\boldsymbol{\kappa}|$ (with $v_D = 3t^{\parallel}a/2\hbar$), where $\boldsymbol{\kappa}$ is measured from the K or K' point, respectively. The phase $\gamma_{\mathbf{k}}$ becomes

$$\gamma_{\mathbf{k}} = \pi - \arctan(\kappa_y/\kappa_x) \tag{34}$$

(K point) and $\gamma_{\mathbf{k}} = \arctan(\kappa_y/\kappa_x)$ (K' point).

B. Tunneling between twisted layers

We describe tunneling between the twisted graphene layers on tip (layer 1, unprimed) and sample (layer 2, primed) following Bistritzer and MacDonald [40]. Starting with the Bernal-stacked configuration [see Fig. 2(a)], the sites of the A sublattice of the twisted layers (denoted $\mathbf{R}$ and $\mathbf{R}'$, respectively) are related by

$$\mathbf{R}' = D(\theta)(\mathbf{R} - \mathbf{e}_1) + \mathbf{d}. \tag{35}$$

Here, $D(\theta)$ is a rotation matrix involving the twist angle θ and $\mathbf{d}$ a relative shift of the rotated lattices. Note that with our conventions, the vector $\mathbf{K}'$ denotes the location of the K point of the primed (sample) layer.

Tunneling between the layers is described by the matrix elements

$$\langle \mathbf{R} + \boldsymbol{\tau}_\alpha | H_T | \mathbf{R}' + \boldsymbol{\tau}'_\beta \rangle = t^{\perp}(\mathbf{R} + \boldsymbol{\tau}_\alpha - \mathbf{R}' - \boldsymbol{\tau}'_\beta) \tag{36}$$

of the (first-quantized) tunneling Hamiltonian H_T, where $t^{\perp}(\mathbf{r})$ is assumed to be only a function of the distance of the sites projected into the graphene plane.

We consider the tunneling matrix elements

$$T_{\mathbf{kp}'}^{\alpha\beta} = \langle \mathbf{k}\alpha | H_T | \mathbf{p}'\beta \rangle \tag{37}$$

between states $|\mathbf{k}\alpha\rangle$ with momentum $\mathbf{k}$ and sublattice α in the tip [Eq. (29)] and states

$$|\mathbf{p}'\beta\rangle = \frac{1}{\sqrt{N}}\sum_{\mathbf{R}'} e^{i\mathbf{p}'\cdot(\mathbf{R}'+\boldsymbol{\tau}'_\beta)}|\mathbf{R}' + \boldsymbol{\tau}'_\beta\rangle \tag{38}$$

with momentum $\mathbf{p}'$ and sublattice β in the sample. The vector $\boldsymbol{\tau}'_\beta$ is rotated relative to $\boldsymbol{\tau}_\beta$ by the twist angle θ. Inserting definitions and expanding

$$t^{\perp}(\mathbf{r}) = \frac{1}{\Omega}\sum_{\mathbf{q}} t_{\mathbf{q}}^{\perp} e^{i\mathbf{q}\cdot\mathbf{r}} \tag{39}$$

into a Fourier series (note that the sum over $\mathbf{q}$ is not restricted to the Brillouin zone), one finds

$$T_{\mathbf{kp}'}^{\alpha\beta} = \frac{1}{\Omega}\sum_{\mathbf{q}}\frac{1}{N}\sum_{\mathbf{R}}\sum_{\mathbf{R}'} t_{\mathbf{q}}^{\perp}$$
$$\times e^{i\mathbf{q}\cdot(\mathbf{R}+\boldsymbol{\tau}_\alpha-\mathbf{R}'-\boldsymbol{\tau}'_\beta)}e^{-i\mathbf{k}\cdot(\mathbf{R}+\boldsymbol{\tau}_\alpha)}e^{i\mathbf{p}'\cdot(\mathbf{R}'+\boldsymbol{\tau}'_\beta)}. \tag{40}$$

We transform the sums over $\mathbf{R}$ and $\mathbf{R}'$ into sums over reciprocal lattice vectors $\mathbf{G}$ and $\mathbf{G}'$ of the two layers using

$$\sum_{\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}} = N\sum_{\mathbf{G}} \delta_{\mathbf{q},\mathbf{G}} \tag{41}$$

and

$$\sum_{\mathbf{R}'} e^{-i\mathbf{q}\cdot\mathbf{R}'} = N\sum_{\mathbf{G}} e^{-i\mathbf{G}'\cdot(-\mathbf{e}'_1+\mathbf{d})}\delta_{\mathbf{q},\mathbf{G}'}. \tag{42}$$

Unlike the sum over $\mathbf{R}$, the sum over $\mathbf{R}'$ generally does not include a term with $\mathbf{R} = 0$, resulting in the phase factor on the right-hand side. This yields [40]

$$T_{\mathbf{kp}'}^{\alpha\beta} = \sum_{\mathbf{G}_1}\sum_{\mathbf{G}_2}\frac{t_{\mathbf{k}+\mathbf{G}_1}^{\perp}}{\Omega_{\text{uc}}}e^{i\mathbf{G}_1\cdot\boldsymbol{\tau}_\alpha-i\mathbf{G}_2\cdot(\boldsymbol{\tau}_\beta-\mathbf{e}_1)-i\mathbf{G}'_2\cdot\mathbf{d}}\delta_{\mathbf{k}+\mathbf{G}_1,\mathbf{p}'+\mathbf{G}'_2}. \tag{43}$$

With momenta measured relative to the Γ point, translation invariance enforces that tunneling conserve crystal momentum modulo reciprocal lattice vectors of the two layers.

Unlike for elastic tunneling, inelastic tunneling at larger twist angles involves virtual states far from the Fermi energy. However, matrix elements of H_T always involve one momentum close to the K points of one of the layers. This can be used to simplify the tunneling matrix element since $t_{\mathbf{q}}$ decays rapidly on the scale of the Brillouin zone [40]. Assuming for definiteness that $\mathbf{k}$ is close to the K point, one retains only those terms in the sum over $\mathbf{G}_1$, in which the momenta in $t_{\mathbf{k}+\mathbf{G}_1}^{\perp}$ have the smallest magnitude, i.e., the three contributions with $\mathbf{G}_1 = 0$, $\mathbf{G}_1 = -\mathbf{Q}_1$, and $\mathbf{G}_1 = -\mathbf{Q}_2$ corresponding to vectors $\mathbf{k} - \mathbf{Q}_j$ located near the three equivalent K points in the hexagonal Brillouin zone. Using that momentum conservation effectively restricts $\mathbf{G}_1 = \mathbf{G}_2$ for relevant twist angles, this gives

$$T_{\mathbf{kp}'}^{\alpha\beta} \simeq w\sum_{j=0}^{2} e^{-i\mathbf{Q}_j\cdot\boldsymbol{\tau}_\alpha+i\mathbf{Q}_j\cdot(\boldsymbol{\tau}_\beta-\mathbf{e}_1)+i\mathbf{Q}'_j\cdot\mathbf{d}}\delta_{\mathbf{k}-\mathbf{Q}_j,\mathbf{p}'-\mathbf{Q}'_j}, \tag{44}$$

where $w = t_{\mathbf{K}}^{\perp}/\Omega_{\mathrm{uc}}$. Explicit evaluation of the exponentials gives

$$T_{\mathbf{kp'}}^{\alpha\beta} \simeq \sum_{j=0}^{2} T_j^{\alpha\beta} e^{i\mathbf{Q}_j \cdot \mathbf{d}} \delta_{\mathbf{k}-\mathbf{Q}_j, \mathbf{p'}-\mathbf{Q}_j'} \tag{45}$$

with the matrices

$$T_0 = w\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad T_{1/2} = w\begin{pmatrix} e^{\mp i\zeta} & 1 \\ e^{\pm i\zeta} & e^{\mp i\zeta} \end{pmatrix} \tag{46}$$

in sublattice space. Here, we used the abbreviation $\zeta = 2\pi/3$. We note that the same expression holds when $\mathbf{p'}$ is close to the K point of the sample, but $\mathbf{k}$ is possibly further from the tip's K point.

C. Matrix elements

Calculations of the tunneling current by Fermi's golden rule require the matrix elements of the tunneling Hamiltonian between eigenstates of the upper and lower layers,

$$T_{\mathbf{kp'}}^{ss'} = \langle \mathbf{k}, s | H_T | \mathbf{p'}, s' \rangle. \tag{47}$$

Here, $s, s' = \pm$ enumerate the valence and conduction bands of the top and bottom graphene layers, respectively. According to Eq. (45), we can write

$$T_{\mathbf{kp'}}^{ss'} \simeq \sum_{j=0}^{2} T_{\mathbf{kp'};j}^{ss'} e^{i\mathbf{Q}_j \cdot \mathbf{d}} \delta_{\mathbf{k}-\mathbf{Q}_j, \mathbf{p'}-\mathbf{Q}_j'} \tag{48}$$

with

$$T_{\mathbf{kp'};j}^{ss'} = \langle \mathbf{k}, s | T_j | \mathbf{p'}, s' \rangle. \tag{49}$$

Equation (46) gives

$$\langle u | T_j | u' \rangle = w e^{-ij\zeta} (u_A + e^{ij\zeta} u_B)^* (u_A' + e^{ij\zeta} u_B'), \tag{50}$$

where we used $e^{3i\zeta} = 1$. Interestingly, the matrix elements factorize into independent contributions of the two spinors. Using the explicit Bloch spinors in Eq. (32), this becomes

$$T_{\mathbf{kp'};j}^{ss'} = \frac{w}{2} e^{-ij\zeta} [1 - s e^{-i(\gamma_{\mathbf{k}} - j\zeta)}]^* [1 - s' e^{-i(\gamma_{\mathbf{p'}}' - j\zeta)}]. \tag{51}$$

We note that the phases $\gamma_{\mathbf{k}}$ and $\gamma_{\mathbf{p'}}'$ are defined in terms of the bond vectors of tip and sample, respectively. We can also express the matrix elements for momenta $\mathbf{k}$ and $\mathbf{p'}$ close to the Dirac points. Writing $\mathbf{k} = \mathbf{K} + \boldsymbol{\kappa}$ and $\mathbf{p'} = \mathbf{K'} + \boldsymbol{\pi'}$, specifying to the K point, and using Eq. (34), we find

$$T_{\boldsymbol{\kappa}\boldsymbol{\pi'};j}^{ss'} = \frac{w e^{-ij\zeta}}{2} [1 + s e^{i(\theta_{\boldsymbol{\kappa}} - \frac{\theta}{2} + j\zeta)}][1 + s' e^{i(\theta_{\boldsymbol{\pi'}} + \frac{\theta}{2} + j\zeta)}]. \tag{52}$$

Note that here we have defined the angles $\theta_{\boldsymbol{\kappa}}$ and $\theta_{\boldsymbol{\pi'}}$ in a global coordinate system, relative to which the tip and sample layers are rotated by $\pm\theta/2$.

D. Electrostatics

We review the electrostatics of the QTM contact [15]. We assume a configuration (see Fig. 1), in which a bias voltage V_b is applied between tip and sample. The electron densities n_{T} and n_{S} of tip and sample are further controlled by gate voltages $V_{\mathrm{TG}} = V_G + V_D$ and $V_{\mathrm{BG}} = V_G - V_D$ applied to the top and bottom gates. Here, we defined the symmetrized and

antisymmetrized gate voltages $V_{\mathrm{G}} = \frac{1}{2}(V_{\mathrm{TG}} + V_{\mathrm{BG}})$ and $V_{\mathrm{D}} = V_{\mathrm{TG}} - V_{\mathrm{BG}}$.

The gate electrodes are assumed to have a high density of states (large quantum capacitance), so that their chemical potentials are independent of the applied voltages. We set $\mu_{\mathrm{TG}} \simeq \mu_{\mathrm{BG}} \simeq 0$. Consequently, the gate voltages control their electric potentials,

$$e(V_{\mathrm{G}} + V_{\mathrm{D}}) = e\phi_{\mathrm{TG}}; \quad e(V_{\mathrm{G}} - V_{\mathrm{D}}) = e\phi_{\mathrm{BG}}. \tag{53}$$

Similarly, the bias voltage V_b controls the electrochemical potentials (encompassing the chemical potentials μ and the electrostatic potentials ϕ) of tip (T) and sample (S),

$$\pm \frac{eV_b}{2} = \mu_{\mathrm{T/S}} + e\phi_{\mathrm{T/S}}. \tag{54}$$

The chemical potentials are related to the electron densities of tip and sample through

$$n_{\mathrm{T/S}} = N_f \frac{\mu_{\mathrm{T/S}}^2}{4\pi \hbar^2 v_D^2} \operatorname{sgn}\mu_{\mathrm{T/S}}. \tag{55}$$

Here, $N_f = 4$ is the number of flavors and we assume that the graphene dispersions can be approximated as linear for relevant densities.

Electrostatics relates the potentials and electron densities through

$$e(\phi_{\mathrm{TG}} - \phi_{\mathrm{T}}) = -\frac{e^2 d_g}{\epsilon_{\mathrm{hBN}}\epsilon_0} n_{\mathrm{TG}}, \tag{56}$$

$$e(\phi_{\mathrm{T}} - \phi_{\mathrm{S}}) = \frac{e^2 d}{\epsilon \epsilon_0}(n_{\mathrm{S}} + n_{\mathrm{BG}}) = -\frac{e^2 d}{\epsilon \epsilon_0}(n_{\mathrm{T}} + n_{\mathrm{TG}}), \tag{57}$$

$$e(\phi_{\mathrm{S}} - \phi_{\mathrm{BG}}) = \frac{e^2 d_g}{\epsilon_{\mathrm{hBN}}\epsilon_0} n_{\mathrm{BG}}. \tag{58}$$

Here, d_g denotes the thickness of the gate dielectrics (dielectric constant ϵ_{hBN}) and d the distance of tip and sample (dielectric constant ϵ). Finally, the relation

$$n_{\mathrm{T}} + n_{\mathrm{S}} = -(n_{\mathrm{TG}} + n_{\mathrm{BG}}) \tag{59}$$

is imposed by overall charge neutrality.

We first solve these equations for the charges on tip and sample. The gate voltage V_{G} controls the overall charge density in tip and sample,

$$2eV_{\mathrm{G}} = \frac{e^2 d_g}{\epsilon_{\mathrm{hBN}}\epsilon_0}(n_{\mathrm{T}} + n_{\mathrm{S}}) - (\mu_{\mathrm{T}} + \mu_{\mathrm{S}}). \tag{60}$$

Assuming that the screening lengths are small compared to d_g in gate electrodes as well as tip and sample, we can further neglect the chemical potential shifts, so that

$$2eV_{\mathrm{G}} = \frac{e^2 d_g}{\epsilon_{\mathrm{hBN}}\epsilon_0}(n_{\mathrm{T}} + n_{\mathrm{S}}). \tag{61}$$

Correspondingly, the difference in electron densities on tip and sample is controlled by the bias voltage V_b in conjunction with the displacement field V_{D},

$$\frac{\epsilon_{\mathrm{hBN}}\epsilon_0}{e^2 d_g} 2eV_{\mathrm{D}} - \frac{2\epsilon\epsilon_0}{e^2 d} eV_b = \frac{2\epsilon\epsilon_0}{e^2 d}(\mu_{\mathrm{T}} - \mu_{\mathrm{S}}) + (n_{\mathrm{T}} - n_{\mathrm{S}}). \tag{62}$$

Here, we assumed that $d \ll d_g$. Equation (55) can now be used to extract $n_{\mathrm{T/S}}$ as well as $\mu_{\mathrm{T/S}}$. This in turn yields $\phi_{\mathrm{T/S}}$ with

Eq. (54). We note that the setup in Fig. 1 admits independent control of the chemical potentials μ_T and μ_S as well as the potential difference $\phi_T - \phi_S$.

In our analytical calculations, we choose $V_G = V_D = 0$ and a small quantum capacitance. Then, tip and sample are overall charge neutral, $n_T + n_S = 0$, and have opposite chemical potentials, $\mu_T = -\mu_S = \mu$. For a linear graphene dispersion, the density of states per flavor at the (bias-dependent) Fermi energy is $\nu(\mu) = \mu/(2\pi\hbar^2 v_D^2)$ for both layers and the difference $\phi = \phi_T - \phi_S$ in electric potentials is

$$e\phi = \frac{e^2 d}{\epsilon \epsilon_0} n_T = (q_{TF} d)\mu \ll \mu, \qquad (63)$$

where $q_{TF} = (e^2/2\epsilon\epsilon_0)N_f \nu(\mu)$ is the Thomas-Fermi wave vector. Thus, $eV_b = 2\mu + e\phi \simeq 2\mu$. At the same time, $e\phi$ gives a relative shift of the Dirac points in tip and sample. This leads to the bias regimes sketched in Fig. 4.

The electrostatics of the contact region ($d \ll d_g$) differs from that far from the contact ($d \gg d_g$), where the gate charges directly control the electron densities of tip and sample, $n_T \simeq -n_{TG}$ and $n_S \simeq -n_{BG}$. For certain parameters, this difference in electrostatics induces *pn* junctions in the sample, which enclose the contact region. This leads to Fabry-Perot–type resonances within the sample, which affect the measured tunneling currents. We do not consider this experimental issue in the following, as it can be avoided in phonon spectroscopy by a judicious choice of parameters.

IV. ELASTIC TUNNELING

We begin our discussion of the tunneling conductance of QTM junctions by considering elastic tunneling. For small twist angles, the Dirac cones of tip and sample layers intersect at small bias voltages and for the same valley, as illustrated in Fig. 2(c). In practice, one limits the strong tunnel coupling in this limit by separating tip and sample by a few atomic layers of a transition metal dichalcogenide (e.g., WSe$_2$) [15].

Complementing the scattering approach sketched in Sec. II, we evaluate the current from tip to sample using Fermi's golden rule,

$$I = \frac{2\pi e N_f}{\hbar} \sum_{s,s'} \sum_{\mathbf{k}} \sum_{\mathbf{p'}} |\langle \mathbf{k}, s | H_T | \mathbf{p'}, s' \rangle|^2$$
$$\times \delta\left(E_{\mathbf{k},s}^{(T)} + e\phi_T - E_{\mathbf{p'},s'}^{(S)} - e\phi_S\right)$$
$$\times \left[f_{\mu_T}\left(E_{\mathbf{k},s}^{(T)}\right) - f_{\mu_S}\left(E_{\mathbf{p'},s'}^{(S)}\right)\right]. \qquad (64)$$

For overall charge neutrality and small quantum capacitance (see Sec. III D), the tunneling current becomes

$$I = \frac{2\pi e N_f}{\hbar} \sum_{s,s'} \sum_{\mathbf{k}} \sum_{\mathbf{p'}} |\langle \mathbf{k}, s | H_T | \mathbf{p'}, s' \rangle|^2$$
$$\times \delta\left(E_{\mathbf{k},s}^{(T)} + e\phi - E_{\mathbf{p'},s'}^{(S)}\right)$$
$$\times \left[f_{eV_b/2}\left(E_{\mathbf{k},s}^{(T)}\right) - f_{-eV_b/2}\left(E_{\mathbf{p'},s'}^{(S)}\right)\right]. \qquad (65)$$

In describing elastic scattering at small twist angles, it is advantageous to measure momenta from the K points of tip and sample. With $\mathbf{k} = \mathbf{K} + \boldsymbol{\kappa}$ and $\mathbf{p'} = \mathbf{K'} + \boldsymbol{\pi'}$ as well as

Eq. (48) for the tunneling matrix elements, we find

$$I = \frac{2\pi e N_f}{\hbar} \sum_{s,s'} \sum_{\boldsymbol{\kappa}} \sum_{\boldsymbol{\pi'}} \sum_{j} |T_{\boldsymbol{\kappa},\boldsymbol{\pi'};j}^{ss'}|^2 \delta_{\boldsymbol{\pi'}-\boldsymbol{\kappa},\mathbf{q}_j}$$
$$\times \delta(E_{\boldsymbol{\kappa},s} + e\phi - E_{\boldsymbol{\pi'},s'})$$
$$\times \left[f_{eV_b/2}(E_{\boldsymbol{\kappa},s}) - f_{-eV_b/2}(E_{\boldsymbol{\pi'},s'})\right]. \qquad (66)$$

We dropped the layer superscripts on the dispersions, which are identical provided that trigonal warping can be neglected.

One expects structure in the differential conductance dI/dV_b at two characteristic voltages [15] (see Fig. 4). As the bias voltage increases, the tunneling current onsets at

$$eV_b^* = \hbar v_D q_j = 2\hbar v_D |\mathbf{K}| \sin\frac{\theta}{2}, \qquad (67)$$

which depends linearly on small twist angles. At a larger bias voltage [see Eq. (63) and Fig. 4]

$$V_b^{**} = \frac{2V_b^*}{q_{TF} d} \gg V_b^*, \qquad (68)$$

the potential difference ϕ leads to nesting of the Dirac cones of tip and sample. This occurs when

$$e\phi(V_b^{**}) = \hbar v_D q_j = 2\hbar v_D |\mathbf{K}| \sin\frac{\theta}{2}. \qquad (69)$$

The linear dependence of V_b^* on (small) θ implies that V_b^{**} has a leading square-root dependence. Retaining the electric potential in the relation $eV_b = 2\mu + e\phi$ yields a subleading linear term.

A. Voltages of order V_b^*

We first consider voltages of order V_b^*, such that the electrostatic potential difference ϕ is negligible. The energy δ function imposes $s = s'$, so that

$$I = \frac{2\pi e N_f}{\hbar} \sum_{j} \sum_{s} \sum_{\boldsymbol{\kappa}} |T_{\boldsymbol{\kappa},\boldsymbol{\kappa}+\mathbf{q}_j;j}^{ss}|^2 \delta(E_{\boldsymbol{\kappa},s} - E_{\boldsymbol{\kappa}+\mathbf{q}_j,s})$$
$$\times \left[f_{eV_b/2}(E_{\boldsymbol{\kappa},s}) - f_{-eV_b/2}(E_{\boldsymbol{\kappa}+\mathbf{q}_j,s})\right]. \qquad (70)$$

It is convenient to rewrite this as

$$I = \frac{2\pi e N_f}{\hbar} \sum_{j} \sum_{\boldsymbol{\kappa}} \sum_{s} |T_{\boldsymbol{\kappa},\boldsymbol{\kappa}+\mathbf{q}_j;j}^{ss}|^2$$
$$\times |E_{\boldsymbol{\kappa},s} + E_{\boldsymbol{\kappa}+\mathbf{q}_j,s}| \delta(E_{\boldsymbol{\kappa},s}^2 - E_{\boldsymbol{\kappa}+\mathbf{q}_j,s}^2)$$
$$\times \left[f_{eV_b/2}(E_{\boldsymbol{\kappa},s}) - f_{-eV_b/2}(E_{\boldsymbol{\kappa}+\mathbf{q}_j,s})\right]. \qquad (71)$$

We pass to the differential conductance at zero temperature. The derivative of the first Fermi function places $E_{\boldsymbol{\kappa},s}$ and thus also $E_{\boldsymbol{\kappa}+\mathbf{q}_j,s}$ at the chemical potential $\mu = eV_b/2$ with $s = +$. Corresponding results hold for the derivative of the second Fermi function, which yields

$$\frac{dI}{dV_b} = \frac{2\pi e^3 N_f V_b}{\hbar} \sum_{j} \sum_{\boldsymbol{\kappa}} \sum_{s} |T_{\boldsymbol{\kappa},\boldsymbol{\kappa}+\mathbf{q}_j;j}^{ss}|^2$$
$$\times \delta(E_{\boldsymbol{\kappa},+}^2 - E_{\boldsymbol{\kappa}+\mathbf{q}_j,+}^2) \delta\left(E_{\boldsymbol{\kappa},+} - \frac{eV_b}{2}\right). \qquad (72)$$

The first δ function enforcing energy conservation has the argument

$$E_{\kappa,+}^2 - E_{\kappa+\mathbf{q}_j,+}^2 = \hbar^2 v_D^2 q_j (2\kappa \cos\theta_j + q_j), \qquad (73)$$

where $\theta_j = \angle(\kappa, \mathbf{q}_j)$. Thus, the δ function imposes $\theta_j \simeq \pi$ for bias voltages near the threshold, which implies $\theta_\kappa + j\zeta = -\pi/2$ and $\theta_{\kappa+\mathbf{q}_j} + j\zeta = \pi/2$. Evaluating the matrix element in Eq. (52) for this case gives $|T_{\kappa,\kappa+\mathbf{q}_j;j}^{ss}|^2 = w^2$ for small twist angles, which is independent of j and s. Thus, we find

$$\frac{dI}{dV_b} \simeq \frac{12\pi e^3 N_f w^2 V_b}{\hbar} \Omega \nu(eV_b/2)$$
$$\times \frac{1}{\hbar^2 v_D^2 q_0} \int \frac{d\theta_0}{2\pi} \delta(2\kappa_F \cos\theta_0 + q_0). \qquad (74)$$

Performing the angular integral yields the result given in Eq. (8) of Sec. II.

B. Voltages of order V_b^{**}

In the vicinity of V_b^{**}, we need to retain the electrostatic potential in Eq. (66). We write the potential as $\phi = \phi(V_b^{**}) + \delta\phi$ and evaluate the current for small $\delta\phi$. For small quantum capacitance $q_{\text{TF}}d \ll 1$, the chemical potential $\mu = eV_b/2$ is much larger than ϕ, so that we retain only the contributions of regions I and III in Fig. 4(c), where tunneling satisfies $s = s'$. Regions I and III give identical contributions by particle-hole symmetry. Moreover, the sum over j gives a factor of 3 in view of C_3 symmetry. Thus, we have

$$I = \frac{12\pi e N_f}{\hbar} \sum_\kappa |T_{\kappa,\kappa+\mathbf{q}_0;0}^{++}|^2 \delta(E_{\kappa,+} + e\phi - E_{\kappa+\mathbf{q}_0,+})$$
$$\times [f_{eV_b/2}(E_{\kappa,+}) - f_{-eV_b/2}(E_{\kappa+\mathbf{q}_0,+})]. \qquad (75)$$

The Fermi-function factor is nonzero and equal to unity as long as $0 < \hbar v_D \kappa < eV_b/2$.

For small $\delta\phi$, we can approximate

$$E_{\kappa,+} + e\phi - E_{\kappa+\mathbf{q}_0,+} \simeq e\delta\phi + \hbar v_D \frac{\kappa q_0(1 - \cos\theta_0)}{\kappa + q_0}. \quad (76)$$

We thus have $\theta_0 \ll 1$. Moreover, we observe that $\theta_\kappa \simeq \theta_{\kappa+\mathbf{q}_0} \simeq \pi/2$ in region I. According to Eq. (52), we can now approximate the matrix element as $|T_{\kappa,\kappa-\mathbf{q}_0;0}^{++}|^2 \simeq w^2$ for small twist angles. Thus, we find

$$I = \frac{3e N_f w^2 \Omega}{\pi \hbar} \int_0^{\mu/(\hbar v_D)} d\kappa\, \kappa \int d\theta_0\, \delta\left(e\delta\phi + \frac{\hbar v_D \kappa q_0}{2(\kappa + q_0)} \theta_0^2\right). \tag{77}$$

We note that the integral over κ is dominated by the upper limit, so that we can approximate $\kappa + q_0 \simeq \kappa$ in the argument of the δ function. Evaluating the remaining integrals yields the result in Eq. (9) of Sec. II.

V. PHONONS AND ELECTRON-PHONON COUPLING

A. Mode expansion

We consider the phonon modes of tip or sample layer, neglecting mechanical coupling between the graphene layers. We expand the atomic displacements $\mathbf{u}(\mathbf{R} + \boldsymbol{\tau}_\alpha)$ of each of the layers, including both in-plane and out-of-plane components, into phonon modes (annihilation operator $b_{r,\mathbf{Q}}$) enumerated by their momenta $\mathbf{Q}$ and mode index r:

$$\mathbf{u}(\mathbf{R} + \boldsymbol{\tau}_\alpha, t) = \frac{1}{\sqrt{N}} \sum_\mathbf{Q} \sum_r \boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha \frac{1}{\sqrt{2M\omega_{r,\mathbf{Q}}}}$$
$$\times e^{i\mathbf{Q}\cdot(\mathbf{R}+\boldsymbol{\tau}_\alpha)}(b_{r,\mathbf{Q}} e^{-i\omega_{r,\mathbf{Q}}t} + b_{r,-\mathbf{Q}}^\dagger e^{i\omega_{r,\mathbf{Q}}t}). \tag{78}$$

Here, M is the mass of the unit cell and the momentum sum is restricted to the Brillouin zone. The polarization vectors $\boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha$ denote the mode displacement of sublattice α, which we normalize according to

$$\sum_\alpha M_\alpha (\boldsymbol{\epsilon}_{r,\mathbf{q}}^\alpha)^* \cdot \boldsymbol{\epsilon}_{r',\mathbf{q}}^\alpha = M\delta_{r,r'} \tag{79}$$

with the M_α denoting the mass of the atom on sublattice α (i.e., $M_\alpha = M/2$ for graphene layers). We keep the mode expansion of the atomic displacements general throughout this paper to facilitate generalization from the case of graphene to other types of layers on tip and sample.

For graphene, the polarization vectors can be chosen real due to inversion symmetry. The polarization vectors satisfy the general relation

$$[\boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha]^* = \boldsymbol{\epsilon}_{r,-\mathbf{Q}}^\alpha, \tag{80}$$

while inversion symmetry implies $\boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha = \boldsymbol{\epsilon}_{r,-\mathbf{Q}}^\alpha$. At long wavelengths, the phonon modes include longitudinal and transverse acoustic modes, longitudinal and transverse optical modes, as well as a flexural mode. The quadratic dispersion of the flexural mode is cut off at a nonzero frequency at long wavelengths due to the coupling between the graphene layers and the substrates, even in the absence of mechanical coupling between tip and sample.

B. Electron-phonon coupling

Electron-phonon coupling arises from changes in the bond lengths of the graphene layers associated with the atomic displacements as well as from the deformation potential. Changes in the intralayer bond lengths result in the electron-phonon coupling of individual graphene layers. At long wavelengths, this intralayer coupling can be incorporated in the Dirac description as a gauge field. The atomic displacements also modify the interlayer tunneling discussed in Sec. III B, leading to interlayer electron-phonon coupling. The electron-phonon coupling originating from the deformation potential gives a separate contribution to the intralayer coupling.

1. Intralayer electron-phonon coupling

We consider the intralayer electron-phonon due to changes in the bond lengths for one of the layers. The atomic displacements modify the electronic tight-binding Hamiltonian of the graphene layer by

$$H_{\text{intra}} = -\sum_\mathbf{R} \sum_{i=1}^3 \delta t_{\mathbf{R},\mathbf{R}+\mathbf{e}_i}^\parallel |\mathbf{R}\rangle \langle \mathbf{R} + \mathbf{e}_i| + \text{H.c.} \tag{81}$$

Assuming that the hopping amplitude $t^\parallel$ depends only on the interatomic distance, one finds

$$\delta t^{\parallel}_{\mathbf{R},\mathbf{R}+\mathbf{e}_i} \simeq \beta \hat{\mathbf{e}}_i \cdot [\mathbf{u}(\mathbf{R} + \mathbf{e}_i) - \mathbf{u}(\mathbf{R})] \quad (82)$$

to linear order in the atomic displacements $\mathbf{u}$. Here, $\hat{\mathbf{e}}_j$ denotes the unit vector in the direction of $\mathbf{e}_j$ and $\beta = \partial t^\parallel / \partial a$ the derivative of the hopping amplitude with respect to the bond length. Notice that the intralayer coupling is limited

to in-plane phonons. Coupling to flexural phonons appears only in quadratic order in the atomic displacements and can be neglected since there is coupling to flexural modes already at linear order in the interlayer electron-phonon coupling.

The matrix elements of the electron-phonon coupling can be obtained by inserting the mode expansion in Eq. (78) into $H_{\text{intra}} |\mathbf{k}, A\rangle$. This yields

$$H_{\text{intra}} |\mathbf{k}, A\rangle = -\frac{1}{N} \sum_{\mathbf{R}} \sum_{j} \sum_{\mathbf{Q}} \sum_{r} \frac{\beta}{\sqrt{2M\omega_{r,\mathbf{Q}}}} \hat{\mathbf{e}}_j \cdot [\epsilon^B_{r,\mathbf{Q}} e^{i\mathbf{Q}\cdot\mathbf{e}_j} - \epsilon^A_{r,\mathbf{Q}}](b_{r,\mathbf{Q}} e^{-i\omega_{r,\mathbf{Q}}t} + b^\dagger_{r,-\mathbf{Q}} e^{i\omega_{r,\mathbf{Q}}t}) e^{i(\mathbf{k}+\mathbf{Q})\cdot\mathbf{R}} |\mathbf{R} + \mathbf{e}_j\rangle. \quad (83)$$

Evaluating the sum over $\mathbf{R}$ yields

$$H_{\text{intra}} |\mathbf{k}, A\rangle = -\frac{1}{\sqrt{N}} \sum_{j} \sum_{\mathbf{Q}} \sum_{r} \frac{\beta}{\sqrt{2M\omega_{r,\mathbf{Q}}}} \hat{\mathbf{e}}_j \cdot [\epsilon^B_{r,\mathbf{Q}} - \epsilon^A_{r,\mathbf{Q}} e^{-i\mathbf{Q}\cdot\mathbf{e}_j}](b_{r,\mathbf{Q}} e^{-i\omega_{r,\mathbf{Q}}t} + b^\dagger_{r,-\mathbf{Q}} e^{i\omega_{r,\mathbf{Q}}t}) e^{-i\mathbf{k}\cdot\mathbf{e}_j} |\mathbf{k} + \mathbf{Q}, B\rangle \quad (84)$$

with $|\mathbf{k} + \mathbf{Q}, B\rangle$ interpreted as the state, for which $\mathbf{k} + \mathbf{Q}$ is folded back into the Brillouin zone. Here, we have used the identity $\frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{p}\cdot\mathbf{R}} |\mathbf{R} + \mathbf{e}_j\rangle = e^{-i\mathbf{p}\cdot\mathbf{e}_j} |\mathbf{p}, B\rangle$, which can be checked by direct calculation. Similarly, one finds

$$H_{\text{intra}} |\mathbf{k}, B\rangle = -\frac{1}{\sqrt{N}} \sum_{j} \sum_{\mathbf{Q}} \sum_{r} \frac{\beta}{\sqrt{2M\omega_{r,\mathbf{Q}}}} \hat{\mathbf{e}}_j \cdot [\epsilon^B_{r,\mathbf{Q}} e^{i\mathbf{Q}\cdot\mathbf{e}_j} - \epsilon^A_{r,\mathbf{Q}}](b_{r,\mathbf{Q}} e^{-i\omega_{r,\mathbf{Q}}t} + b^\dagger_{r,-\mathbf{Q}} e^{i\omega_{r,\mathbf{Q}}t}) e^{i\mathbf{k}\cdot\mathbf{e}_j} |\mathbf{k} + \mathbf{Q}, A\rangle. \quad (85)$$

As a result, the intralayer electron-phonon interaction takes the form

$$H_{\text{intra}} = \sum_{\mathbf{Q}} \sum_{r} (b_{r,-\mathbf{Q}} e^{-i\omega_{r,-\mathbf{Q}}t} + b^\dagger_{r,\mathbf{Q}} e^{i\omega_{r,\mathbf{Q}}t}) \sum_{\mathbf{k}} \{|\mathbf{k} - \mathbf{Q}, B\rangle M^r_{\mathbf{k}-\mathbf{Q},B;\mathbf{k};A} \langle \mathbf{k}, A| + |\mathbf{k} - \mathbf{Q}, A\rangle [M^r_{\mathbf{k},B;\mathbf{k}-\mathbf{Q};A}]^* \langle \mathbf{k}, B|\} \quad (86)$$

with the electron-phonon matrix element

$$M^r_{\mathbf{k}-\mathbf{Q},B;\mathbf{k},A} = \frac{1}{\sqrt{N}} \sum_{j} \frac{\beta}{\sqrt{2M\omega_{r,\mathbf{Q}}}}$$
$$\times \hat{\mathbf{e}}_j \cdot [\epsilon^A_{r,\mathbf{Q}} e^{-i(\mathbf{k}-\mathbf{Q})\cdot\mathbf{e}_j} - \epsilon^B_{r,\mathbf{Q}} e^{-i\mathbf{k}\cdot\mathbf{e}_j}]. \quad (87)$$

For acoustic phonons, the term in square brackets vanishes linearly in $\mathbf{Q}$, while the denominator vanishes only as $|\mathbf{Q}|^{1/2}$. Thus, the intralayer electron-phonon coupling vanishes in the long-wavelength limit. For optical phonons, the coupling approaches a constant in the long-wavelength limit.

2. Interlayer electron-phonon coupling

The interlayer electron-phonon coupling arises from modifications in the interlayer distances between atoms in the two layers. Importantly, this contribution to the electron-phonon coupling cannot be obtained starting with the continuum model of twisted bilayer graphene, even at long phonon

wavelengths. The reason is that as we will see, the dominant contribution to the coupling arises from phonon-induced changes of phases of the tunneling matrix element, which depend on the large momenta of the K points as measured from the Γ point.

The interlayer tunneling amplitude is a function of the in-plane and out-of-plane distances of the atoms in the two layers. Including the modification of the distances by the atomic displacements, we have

$$t^\perp(|\mathbf{R} + \boldsymbol{\tau}_\alpha - \mathbf{R}' - \boldsymbol{\tau}'_\beta + \mathbf{u}_\parallel - \mathbf{u}'_\parallel|, d + u_\perp - u'_\perp)$$
$$= \frac{1}{\Omega} \sum_{\mathbf{q}} t^\perp_{\mathbf{q}}(d + u_\perp - u'_\perp) e^{i\mathbf{q}\cdot(\mathbf{R}+\boldsymbol{\tau}_\alpha-\mathbf{R}'-\boldsymbol{\tau}'_\beta+\mathbf{u}_\parallel-\mathbf{u}'_\parallel)}. \quad (88)$$

Here, we used the shorthands $\mathbf{u} = \mathbf{u}(\mathbf{R} + \boldsymbol{\tau}_\alpha)$ and $\mathbf{u}' = \mathbf{u}'(\mathbf{R}' + \boldsymbol{\tau}'_\alpha)$ and denote the equilibrium interlayer distance by d. Expanding to linear order in the displacements yields

$$t^\perp(|\mathbf{R} + \boldsymbol{\tau}_\alpha - \mathbf{R}' - \boldsymbol{\tau}'_\beta + \mathbf{u}_\parallel - \mathbf{u}'_\parallel|, d + u_\perp - u'_\perp) \simeq \frac{1}{\Omega} \sum_{\mathbf{q}} t^\perp_{\mathbf{q}}(d) e^{i\mathbf{q}\cdot(\mathbf{R}+\boldsymbol{\tau}_\alpha-\mathbf{R}'-\boldsymbol{\tau}'_\beta)} \left\{1 + i\mathbf{q}\cdot(\mathbf{u}_\parallel - \mathbf{u}'_\parallel) + \frac{\partial \ln t^\perp_{\mathbf{q}}(d)}{\partial d}(u_\perp - u'_\perp)\right\}. \quad (89)$$

The first term in the curly brackets is just the interlayer tunneling discussed in Sec. III B [40]. The contributions of the second and third terms, collectively denoted $\delta t^\perp$, encapsulate the electron-phonon coupling

$$H_{\text{inter}} = \sum_{\mathbf{R}} \sum_{\mathbf{R}'} \sum_{\alpha,\beta} |\mathbf{R} + \boldsymbol{\tau}_\alpha\rangle \delta t^\perp(|\mathbf{R} + \boldsymbol{\tau}_\alpha - \mathbf{R}' - \boldsymbol{\tau}'_\beta + \mathbf{u}_\parallel - \mathbf{u}'_\parallel|, d + u_\perp - u'_\perp) \langle \mathbf{R}' + \boldsymbol{\tau}'_\beta| + \text{H.c.} \quad (90)$$

to in-plane ($\mathbf{u}_\parallel$) and out-of-plane ($u_\perp$) phonons.

The derivation of matrix elements $\langle \mathbf{k}, \alpha | H_{\text{inter}} | \mathbf{p}', \beta \rangle$ closely follows the discussion of interlayer tunneling in Sec. III B. Using the mode expansion in Eq. (78) of the atomic displacements, one finds the following rule for shifting the electronic momenta in the expressions for the tunneling matrix element: Scattering from phonons in the tip (wave vector $\mathbf{Q}$) can be accounted for by shifting the outgoing momentum according as $\mathbf{k} \to \mathbf{k} - \mathbf{Q}$ and leaves the incoming momentum $\mathbf{p}'$ invariant. Scattering from phonons in the sample (wave vector $\mathbf{Q}'$) can be accounted for by shifting the incoming momentum $\mathbf{p}' \to \mathbf{p}' + \mathbf{Q}'$ and leaves the outgoing momentum $\mathbf{k}$ invariant. This yields

$$\langle \mathbf{k}, \alpha | H_{\text{inter}} | \mathbf{p}', \beta \rangle = \sum_{\mathbf{Q}} \sum_r \langle \mathbf{k}, \alpha | H_{\text{inter}} | \mathbf{p}', \beta; r, \mathbf{Q} \rangle \left(b_{r,\mathbf{Q}} e^{-i\omega_{r,\mathbf{Q}} t} + b_{r,-\mathbf{Q}}^\dagger e^{i\omega_{r,\mathbf{Q}} t} \right) - (\mathbf{Q} \to \mathbf{Q}'), \tag{91}$$

where [cf. Eq. (43)]

$$\langle \mathbf{k}, \alpha | H_{\text{inter}} | \mathbf{p}', \beta; r, \mathbf{Q} \rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{G}_1} \sum_{\mathbf{G}_2} e^{i\mathbf{G}_1 \cdot \boldsymbol{\tau}_\alpha - i\mathbf{G}_2 \cdot (\boldsymbol{\tau}_\beta - \mathbf{e}_1) - i\mathbf{G}_2' \cdot \mathbf{d}} \delta_{\mathbf{k}-\mathbf{Q}+\mathbf{G}_1, \mathbf{p}'+\mathbf{G}_2'}$$

$$\times \frac{1}{\sqrt{2M\omega_{r,\mathbf{Q}}}} \left\{ \frac{t_{\mathbf{p}'+\mathbf{G}_2'}^\perp}{\Omega_{\text{uc}}} i(\mathbf{p}' + \mathbf{G}_2') \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha + \frac{1}{\Omega_{\text{uc}}} \frac{\partial t_{\mathbf{p}'+\mathbf{G}_2'}^\perp(d)}{\partial d} \hat{\mathbf{z}} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha \right\} \tag{92}$$

and

$$\langle \mathbf{k}, \alpha | H_{\text{inter}} | \mathbf{p}', \beta; r, \mathbf{Q}' \rangle = -\frac{1}{\sqrt{N}} \sum_{\mathbf{G}_1} \sum_{\mathbf{G}_2} e^{i\mathbf{G}_1 \cdot \boldsymbol{\tau}_\alpha - i\mathbf{G}_2 \cdot (\boldsymbol{\tau}_\beta - \mathbf{e}_1) - i\mathbf{G}_2' \cdot \mathbf{d}} \delta_{\mathbf{k}+\mathbf{G}_1, \mathbf{p}'+\mathbf{Q}'+\mathbf{G}_2'}$$

$$\times \frac{1}{\sqrt{2M\omega_{r,\mathbf{Q}'}}} \left\{ \frac{t_{\mathbf{k}+\mathbf{G}_1}^\perp}{\Omega_{\text{uc}}} i(\mathbf{k} + \mathbf{G}_1) \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}'}^\beta + \frac{1}{\Omega_{\text{uc}}} \frac{\partial t_{\mathbf{k}+\mathbf{G}_1}^\perp(d)}{\partial d} \hat{\mathbf{z}} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}'}^\beta \right\}. \tag{93}$$

The first terms in the curly brackets in Eqs. (92) and (93) describe coupling to in-plane phonons, while the second terms describe coupling to flexural phonons.

The contribution of H_{inter} to the inelastic tunneling current can be accounted for in first-order perturbation theory. Consequently, the electronic momenta $\mathbf{k}$ and $\mathbf{p}'$ are close to the K points of tip and sample. In this case, we can further simplify the matrix elements. As discussed in Sec. III B, the dominant terms are then given by $\mathbf{G}_1 = 0$, $\mathbf{G}_1 = -\mathbf{Q}_1$, and $\mathbf{G}_1 = -\mathbf{Q}_2$ with $\mathbf{G}_1 = \mathbf{G}_2$. Moreover, except for the smallest twist angles, we can neglect the distance of $\mathbf{k}$ and $\mathbf{p}'$ from the respective K points relative to the distance q_j between the K points of tip and sample. With these approximations, the Kronecker-δ imposing momentum conservation in Eq. (92) simplifies as $\delta_{\mathbf{k}-\mathbf{Q}+\mathbf{G}_1, \mathbf{p}'+\mathbf{G}_2'} \to \delta_{\mathbf{K}-\mathbf{Q}-\mathbf{Q}_j, \mathbf{K}'-\mathbf{Q}_j'} = \delta_{\mathbf{Q},\mathbf{q}_j}$, so that the phonon wave vector equals a connection vector between the Dirac points of tip and sample [and analogously for Eq. (93)]. Thus, we find

$$\langle \mathbf{k}, \alpha | H_{\text{inter}} | \mathbf{p}', \beta; r, \mathbf{Q} \rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^2 T_j^{\alpha\beta} e^{i\mathbf{Q}_j' \cdot \mathbf{d}} \delta_{\mathbf{Q},\mathbf{q}_j} \frac{1}{\sqrt{2M\omega_{r,\mathbf{Q}}}} \left\{ i(\mathbf{K}' - \mathbf{Q}_j') \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha + \frac{\tilde{w}}{w} \hat{\mathbf{z}} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}}^\alpha \right\}, \tag{94}$$

$$\langle \mathbf{k}, \alpha | H_{\text{inter}} | \mathbf{p}', \beta; r, \mathbf{Q}' \rangle = -\frac{1}{\sqrt{N}} \sum_{j=0}^2 T_j^{\alpha\beta} e^{i\mathbf{Q}_j' \cdot \mathbf{d}} \delta_{\mathbf{Q}',\mathbf{q}_j} \frac{1}{\sqrt{2M\omega_{r,\mathbf{Q}'}}} \left\{ i(\mathbf{K} - \mathbf{Q}_j) \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}'}^\beta + \frac{\tilde{w}}{w} \hat{\mathbf{z}} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}'}^\beta \right\}. \tag{95}$$

Here, we defined $\tilde{w} = \frac{\partial w}{\partial d}$. (Note that $\tilde{w}$ and w have different dimensions.) Several comments are in order concerning these results in the limit of small twist angles. (i) The vectors $\mathbf{q}_j$ and hence the phonon wave vector $\mathbf{Q}$ are approximately perpendicular to $\mathbf{K} - \mathbf{Q}_j$ as well as $\mathbf{K}' - \mathbf{Q}_j'$. Thus, as a consequence of the scalar product, the coupling to in-plane phonons is predominantly to transverse phonon modes for the phonon vectors probed in QTM experiments. (ii) The coupling to transverse acoustic modes diverges as $|\mathbf{Q}|^{-1/2}$ at small phonon wave vectors $\mathbf{Q}$. (iii) The coupling to the transverse acoustic modes is effectively to layer-antisymmetric phonons, also known as the phason mode.

As mentioned in the beginning of this section, the relevant momenta entering Eqs. (94) and (95) are the large momenta of the K points rather than the phonon wave vector. It is for this reason that even in the long-wavelength limit, this coupling cannot be obtained by adding phonon displacements to the continuum model of twisted bilayer graphene. Instead, one has to follow the derivation of the continuum model after taking the phonon displacements into account.

The underlying reason is that the interlayer coupling arises from phase factors associated with the interlayer tunneling. We use this observation in the Appendix A to show that by means of a gauge transformation, the interlayer coupling can be brought into a form, which is analogous to the intralayer coupling. As it appears in this section, the contribution of the interlayer coupling to the inelastic tunneling current can be accounted for in a first-order golden rule calculation. In the transformed form, it must be treated in second order. Remarkably, in the transformed form, the intralayer and interlayer

couplings are related by the replacement

$$\frac{\partial t^{\parallel}}{\partial a} \leftrightarrow it_{\parallel}\mathbf{K}. \tag{96}$$

This shows that one expects both contributions to the inelastic tunneling current to have corresponding parametric dependencies, as borne out by the explicit calculations (see Table I).

3. Deformation-potential coupling

For long-wavelength phonons, the deformation potential

$$V(\mathbf{r}) = -D\nabla \cdot \mathbf{u} \tag{97}$$

gives an additional contribution to the electron-phonon coupling of longitudinal acoustic phonons. This can be compared with the intralayer electron-phonon coupling in Eqs. (81) and (82) due to changes in the bond length. In contrast to the deformation potential, the latter is of comparable magnitude for transverse and longitudinal acoustic phonons. Apart from this difference, the magnitudes of the deformation-potential and the gauge couplings are related by the replacement

$$D \leftrightarrow \frac{\partial t^{\parallel}}{\partial \ln a}. \tag{98}$$

We can use this correspondence to obtain the additional contribution of the deformation potential to the inelastic tunneling current from the result for the intralayer gauge coupling.

We also note that the deformation coupling locally shifts the chemical potential, inducing changes of the charge density. These charge fluctuations will be screened by electron-electron interactions, effectively reducing the strength of the deformation potential. We take D to be the renormalized coupling.

VI. PHONON SPECTROSCOPY

A. Inelastic-tunneling current

We are now in a position to compute the inelastic tunneling current to the leading orders in tip-sample tunneling and electron-phonon coupling. Inelastic electron tunneling in conjunction with phonon emission emerges from the interlayer electron-phonon coupling $\mathcal{H}_{\text{inter}}$ [see Eqs. (94) and (95)] as well as the intralayer electron-phonon coupling $\mathcal{H}_{\text{intra}}$ in conjunction with interlayer tunneling. While the first contribution can be captured by Fermi's golden rule in lowest order, the second requires a higher-order calculation.

Keeping the calculation general at first, we assume a nonzero matrix element $\langle \mathbf{p}', s'; r, \mathbf{Q} | \mathcal{T}_{\text{inel}} | \mathbf{k}, s \rangle$ of the inelastic contribution to the $\mathcal{T}$ matrix without specifying to a particular process. We assume zero temperature and $eV_b > 0$. Then, tunneling is unidirectional from tip to sample and only phonon emission contributes. We also specify to charge neutrality and small quantum capacitance, so that $\phi_{\text{T}} = \phi_{\text{S}}$ and $\mu = eV_b/2$ at bias voltages corresponding to typical phonon frequencies. With these assumptions, the phonon contribution δI to the tunneling current from tip to sample is given by

$$\delta I = \frac{2\pi e N_f}{\hbar} \sum_{\mathbf{Q}} \sum_r \sum_{\mathbf{k}, \mathbf{p}'} \sum_{s,s'} |\langle \mathbf{p}', s'; r, \mathbf{Q} | \mathcal{T}_{\text{inel}} | \mathbf{k}, s \rangle|^2 \delta(E_{\mathbf{p}',s'} + \hbar\omega_{r,\mathbf{Q}} - E_{\mathbf{k},s}) f_\mu(E_{\mathbf{k},s})[1 - f_{-\mu}(E_{\mathbf{p}',s'})] + (\mathbf{Q} \to \mathbf{Q}'). \tag{99}$$

Passing to the differential conductance in the limit of zero temperature using $\mu = eV_b/2$ gives

$$\frac{d\delta I}{dV_b} = \frac{2\pi e^2 N_f}{2\hbar} \sum_{\mathbf{Q}} \sum_r \sum_{\mathbf{k}, \mathbf{p}'} \sum_{s,s'} \theta(eV_b - \hbar\omega_{r,\mathbf{Q}}) |\langle \mathbf{p}', s'; r, \mathbf{Q} | \mathcal{T}_{\text{inel}} | \mathbf{k}, s \rangle|^2$$

$$\times \{\delta(E_{\mathbf{p}',s'} + \hbar\omega_{r,\mathbf{Q}} - \mu)\delta(E_{\mathbf{k},s} - \mu) + \delta(-\mu + \hbar\omega_{r,\mathbf{Q}} - E_{\mathbf{k},s})\delta(E_{\mathbf{p}',s'} + \mu)\} + (\mathbf{Q} \to \mathbf{Q}'). \tag{100}$$

The singular contribution to the second derivative $d^2\delta I/dV_b^2$ arises from the derivative of the threshold factor $\theta(eV_b - \hbar\omega_{r,\mathbf{Q}})$. The resulting δ function enforces that the bias voltage match the phonon energy. Using this constraint, the two terms in curly brackets become equal, and we obtain

$$\frac{d^2\delta I}{dV_b^2} = \frac{2\pi e^3 N_f}{\hbar} \sum_{\mathbf{Q}} \sum_r \delta(eV_b - \hbar\omega_{r,\mathbf{Q}}) \sum_{\mathbf{k}, \mathbf{p}'} |\langle \mathbf{p}', -; r, \mathbf{Q} | \mathcal{T}_{\text{inel}} | \mathbf{k}, + \rangle|^2 \delta(E_{\mathbf{p}',-} + \mu)\delta(E_{\mathbf{k},+} - \mu) + (\mathbf{Q} \to \mathbf{Q}'). \tag{101}$$

Here, we also used that tunneling at the threshold is from $s = +$ to $s' = -$. Due to the δ functions, the initial and final electron states are located at the Fermi energies of tip and sample, respectively. Thus, we find

$$\frac{d^2\delta I}{dV_b^2} = \frac{2\pi e^3 N_f \Omega^2}{\hbar} \nu(\mu)\nu(-\mu) \sum_r \sum_{\mathbf{Q}} \delta(eV_b - \hbar\omega_{r,\mathbf{Q}}) \int \frac{d\theta_{\mathbf{k}}}{2\pi} \frac{d\theta_{\mathbf{p}'}}{2\pi} |\langle \mathbf{p}', -; r, \mathbf{Q} | \mathcal{T}_{\text{inel}} | \mathbf{k}, + \rangle|^2 + (\mathbf{Q} \to \mathbf{Q}'), \tag{102}$$

which depends on Fermi-circle averages of the $\mathcal{T}$-matrix element.

B. Interlayer electron-phonon coupling: First-order perturbation theory

We first consider the contribution δI_{inter} of the interlayer electron-phonon coupling $\mathcal{T}_{\text{inel}} \to \mathcal{H}_{\text{inter}}$ in Eq. (102). We can readily perform the angular averages in Eq. (102) using Eqs. (94), (95), and (52). By C_3 symmetry, the sum over j implicit in the matrix

elements can be accounted for by a factor of 3 and we obtain

$$\frac{d^2\delta I_{\text{inter}}}{dV_b^2} = \frac{6\pi\, e^3 N_f \Omega^2}{\hbar N} \sum_r \delta(eV_b - \hbar\omega_{r,\mathbf{q}_0}) \frac{\nu(\mu)\nu(-\mu)}{2M\omega_{r,\mathbf{q}_0}} \sum_\alpha \left\{ \left| iw\mathbf{K}' \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}=\mathbf{q}_0}^\alpha + \tilde{w}\hat{\mathbf{z}} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}=\mathbf{q}_0}^\alpha \right|^2 + \left| iw\mathbf{K} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}'=\mathbf{q}_0}^\alpha + \tilde{w}\hat{\mathbf{z}} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}'=\mathbf{q}_0}^\alpha \right|^2 \right\}.$$

$$(103)$$

Here, we also used that by symmetry, $\omega_{r,\mathbf{q}_0}$ is the same for the tip and sample layers. Finally, we note that the contributions of phonon emissions in tip and sample are identical in magnitude and use the fact that the polarization vectors are real. This yields the result

$$\frac{d^2\delta I_{\text{inter}}}{dV_b^2} = G_{\text{incoh}} \sum_r \delta(V_b - \hbar\omega_{r,\mathbf{q}_0}/e) \frac{8\pi^2(\kappa_F \ell_{r,\mathbf{q}_0})^2}{\sqrt{3}}$$

$$\times \sum_\alpha \left\{ \left(\hat{\mathbf{K}}' \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}=\mathbf{q}_0}^\alpha\right)^2 + \frac{\tilde{w}^2}{w^2|\mathbf{K}|^2} \left(\hat{\mathbf{z}} \cdot \boldsymbol{\epsilon}_{r,\mathbf{Q}=\mathbf{q}_0}^\alpha\right)^2 \right\}.$$

$$(104)$$

Here, we have rewritten the prefactor by introducing the length $\ell_{r,\mathbf{q}_0} = \sqrt{\hbar/M\omega_{r,\mathbf{q}_0}}$ characterizing the contribution of phonon mode r with wave vector $\mathbf{q}_0$ to the amplitude of the zero-point motion of the atoms and using $|\mathbf{K}|^2\Omega_{\text{uc}} = 8\pi^2/(3\sqrt{3})$. We note that in Sec. II D as well as Table I, we use the less specific notation ℓ_{ZPM} for $\ell_{r,\mathbf{q}_0}$.

Each phonon mode contributes a δ-function peak to $d^2\delta I/dV_b^2$ at $eV_b = \hbar\omega_{r,\mathbf{q}_0}$. Several comments are in order: (i) At small twist angles, transverse phonons contribute more strongly to the inelastic tunneling current as the vectors $\mathbf{K}$ and $\mathbf{K}'$ are nearly orthogonal to the phonon wave vector $\mathbf{q}_0$. (ii) As the twist angle decreases and the phonon wave vector $q_0 \to 0$, the strength of the electron-phonon coupling diverges for acoustic phonons. Concurrently, the Fermi wave vector κ_F decreases. On balance, we have $(\kappa_F \ell_r)^2 \sim \omega_{r,\mathbf{q}_0}$, so that the strength of the phonon resonance decreases as the twist angle becomes smaller. We note, however, that this is specific to the case of overall charge neutrality. Away from charge neutrality, the Fermi wave vector remains nonzero at zero bias and one finds a singular enhancement at small twist angles. The cutoff at small twist angles is discussed at the end of Sec. II D. (iii) For the same reason, acoustic phonons contribute more weakly than optical phonons at overall charge neutrality, but stronger away from charge neutrality provided that the phonon frequencies are small compared to the zero-bias chemical potential. (iv) Similar considerations apply to the out-of-plane ZA and ZO modes. The lower-frequency ZA mode is expected to be weaker than the ZO mode at charge neutrality, but stronger away from it. We also note that the ZA mode is gapped at small twist angles, cutting off the electron-phonon coupling.

In deriving $\frac{d^2\delta I}{dV_b^2}$, we neglected the electronic momenta measured from the respective K points, relative to q_0 [see

Eqs. (94) and (95) and the preceding discussion]. Within this approximation, all inelastic tunneling events involve phonons with momenta $\mathbf{q}_j$ and $\frac{d^2\delta I}{dV_b^2}$ becomes a δ peak as a function of bias voltage. Retaining the small electronic momenta effectively broadens the δ peak. The phonon momenta are now equal to $\mathbf{q}_j$ only up to wave vectors of order κ_F. For acoustic modes, we can estimate the change in phonon frequency as $\Delta\omega_r \sim \hbar c\kappa_F$ with c the speed of sound. Thus, we find

$$\frac{\Delta\omega_r}{\omega_{r,\mathbf{q}_0}} \sim \frac{c}{v_D} \ll 1 \qquad (105)$$

for the relative broadening of the phonon peak. We note that this approximation also implies a limitation on the twist angle due to the requirement $\kappa_F \ll q_0 \approx |\mathbf{K}|\theta$. For an optical mode (for which this condition is more stringent), this implies

$$\theta \gg \frac{c}{v_D}. \qquad (106)$$

Here, we estimated the frequency of the optical phonon as $c|\mathbf{K}|$.

C. Intralayer electron-phonon coupling: Second-order perturbation theory

The contribution of the intralayer electron-phonon interaction is obtained by $\mathcal{T}_{\text{inel}} \to \mathcal{H}_T \mathcal{G}_0 \mathcal{H}_{\text{intra}} + \mathcal{H}_{\text{intra}} \mathcal{G}_0 \mathcal{H}_T$ in Eq. (102). This is written in second quantization (as indicated by calligraphic symbols) to keep proper track of the relative exchange phase of the two contributions.

The relevant processes are illustrated in Fig. 8, for a bias voltage at threshold ($eV_b = \hbar\omega_{r,\mathbf{q}_j}$). Electron tunneling is from the chemical potential in the tip (conduction band) to the chemical potential in the sample (valence band). We specify to electron wave vectors in the vicinity of $\mathbf{K}$ and $\mathbf{K}'$. Figures 8(a) and 8(c) describe phonon emission in the tip, 8(b) and 8(d) in the sample. By symmetry, both sets of processes contribute equally. For definiteness, we specify to emission of a phonon in the tip.

In the process shown in Fig. 8(a), tunneling follows the initial phonon emission, so that the process contributes to $\mathcal{H}_T \mathcal{G}_0 \mathcal{H}_{\text{intra}}$. Similarly, in the process shown in Fig. 8(c), phonon emission follows the initial tunneling, so that the process contributes to $\mathcal{H}_{\text{intra}} \mathcal{G}_0 \mathcal{H}_T$. These processes are described by the matrix elements

$$\langle \mathbf{p}', -; r, \mathbf{Q} | H_T G_0 H_{\text{intra}} | \mathbf{k}, + \rangle = \sum_{\mathbf{r}} \frac{\langle \mathbf{p}', -; r, \mathbf{Q} | H_T | \mathbf{r}, +; r, \mathbf{Q} \rangle \langle \mathbf{r}, +; r, \mathbf{Q} | H_{\text{intra}} | \mathbf{k}, + \rangle}{E_{\mathbf{k},+} - E_{\mathbf{r},s} - \hbar\omega_{r,\mathbf{Q}}} \qquad (107)$$
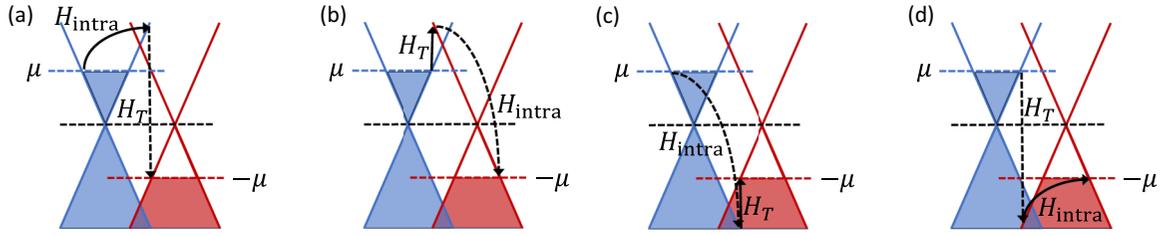
FIG. 8. Second-order processes contributing to the inelastic tunneling current involving the intralayer electron-phonon coupling. (a), (b) Processes with intermediate electrons in the conduction band. (c), (d) Processes with intermediate valence holes. All processes are shown at the threshold voltage $eV_b = \hbar\omega_{r,\mathbf{q}_j}$, with full (dashed) black lines indicating the first (second) process. Notice that due to the difference between electronic and phononic velocities, actual phonon energies (as well as μ) are much smaller than represented here for reasons of readability. This also implies that the electronic momenta measured from the respective Dirac points are much smaller than the distance between the Dirac points of tip and sample.

and

$$\langle \mathbf{p}', -; r, \mathbf{Q}|H_{\text{intra}}G_0 H_T|\mathbf{k}, +\rangle = \sum_{\mathbf{r}} \frac{\langle \mathbf{p}', -; r, \mathbf{Q}|H_{\text{intra}}|\mathbf{p}', -; \overline{\mathbf{r}, -}; \mathbf{k}, +\rangle \langle \mathbf{p}', -; \overline{\mathbf{r}, -}; \mathbf{k}, +|H_T|\mathbf{k}, +\rangle}{E_{\mathbf{r},s} - E'_{\mathbf{p}',-}} \tag{108}$$

of the $\mathcal{T}$ matrix, respectively. Here, we denote holes by overlines. In this section, we also denote energies of tip (sample) without (with) prime. Energy and momentum conservation demand that $E_{\mathbf{k},+} = E'_{\mathbf{p}',-} + \hbar\omega_{\mathbf{Q}}$ as well as $\mathbf{Q} = \mathbf{k} - \mathbf{p}' \simeq \mathbf{q}_j$. The three contributions to tunneling are related by C_3 symmetry and incoherent, so that we focus on the $j = 0$ contribution, which leaves the momentum unchanged, i.e., $\mathbf{r} = \mathbf{p}'$ in both processes.

The processes in Figs. 8(a) and 8(c) contribute with a relative statistical minus sign. This follows as they involve the operator products $c^\dagger_{\mathbf{p},-}c_{\mathbf{r},+}c^\dagger_{\mathbf{r},+}c_{\mathbf{k},+}$ and $c^\dagger_{\mathbf{r},-}c_{\mathbf{k},+}c^\dagger_{\mathbf{p},-}c_{\mathbf{r},-}$, respectively. Anticommuting the operators in the second product turns it into $-c^\dagger_{\mathbf{p},-}c^\dagger_{\mathbf{r},-}c_{\mathbf{r},-}c_{\mathbf{k},+}$. The relative minus sign follows by noting that $c_{\mathbf{r},+}c^\dagger_{\mathbf{r},+}$ and $c^\dagger_{\mathbf{r},-}c_{\mathbf{r},-}$ both reduce to unity when acting on the ground state.

The energy denominators of the two processes in Eqs. (107) and (108) are equal to each other up to corrections which are small in the ratio of the phonon and electron velocities. As can also be seen by inspection of Figs. 8(a) and 8(c), the energy denominators differ by the phonon energy $\hbar\omega_{r,\mathbf{Q}}$, which can be neglected in leading order. Thus, they can be approximated by $E_{\mathbf{k},+} - E_{\mathbf{p}',+} \simeq \hbar v_D q_0$ provided that the linearized Dirac spectrum accurately describes the intermediate state at momentum $\mathbf{r} = \mathbf{p}'$.

Using Eq. (51), we find

$$\langle \mathbf{p}', -; r, \mathbf{Q}|H_T|\mathbf{r} = \mathbf{p}', +; r, \mathbf{Q}\rangle = \frac{w}{2}[1 + e^{i\gamma_{\mathbf{p}'}}][1 - e^{-i\gamma'_{\mathbf{p}'}}] \tag{109}$$

and

$$\langle \mathbf{p}', -; \overline{\mathbf{r} = \mathbf{p}', -}; \mathbf{k}, +|H_T|\mathbf{k}, +\rangle = \frac{w}{2}[1 + e^{i\gamma_{\mathbf{p}'}}][1 + e^{-i\gamma'_{\mathbf{p}'}}] \tag{110}$$

for the matrix elements of the tunneling Hamiltonian. Note that here, $\gamma_{\mathbf{p}'}$ and $\gamma'_{\mathbf{p}'}$ are evaluated using the bond vectors of the respective layer. To compute the matrix elements of H_{intra}, we use that it has only off-diagonal matrix elements in sublattice space. We can approximate the electron momenta as $\mathbf{k} \simeq \mathbf{K}$ and $\mathbf{p}' = \mathbf{r} \simeq \mathbf{K}'$. Using Eqs. (32) and (86), we then have

$$\langle \mathbf{r} = \mathbf{p}', +; r, \mathbf{Q}|H_{\text{intra}}|\mathbf{k}, +\rangle$$
$$\simeq -\delta_{\mathbf{Q},\mathbf{q}_0}\frac{1}{2}\left\{ M^r_{\mathbf{K}',B;\mathbf{K},A}e^{i\gamma_{\mathbf{p}'}} + \left[M^r_{\mathbf{K},B;\mathbf{K}',A}\right]^* e^{-i\gamma_{\mathbf{k}}} \right\} \tag{111}$$

and

$$\langle \mathbf{p}', -; r, \mathbf{Q}|H_{\text{intra}}|\mathbf{p}', -; \overline{\mathbf{r} = \mathbf{p}', -}; \mathbf{k}, +\rangle$$
$$\simeq \delta_{\mathbf{Q},\mathbf{q}_0}\frac{1}{2}\left\{ M^r_{\mathbf{K}',B;\mathbf{K},A}e^{i\gamma_{\mathbf{p}'}} - \left[M^r_{\mathbf{K},B;\mathbf{K}',A}\right]^* e^{-i\gamma_{\mathbf{k}}} \right\}. \tag{112}$$

Adding the amplitudes accounting for their relative minus sign, we find

$$\langle \mathbf{p}', -; r, \mathbf{Q}|H_{\text{intra}}G_0 H_T - H_T G_0 H_{\text{intra}}|\mathbf{k}, +\rangle$$
$$= -\delta_{\mathbf{Q},\mathbf{q}_0}\frac{w}{4\hbar v_D q_0}[1 + e^{i\gamma_{\mathbf{p}'}}]\left\{[1 - e^{-i\gamma'_{\mathbf{p}'}}]\left(M^r_{\mathbf{K}',B;\mathbf{K},A}e^{i\gamma_{\mathbf{p}'}} + \left[M^r_{\mathbf{K},B;\mathbf{K}',A}\right]^* e^{-i\gamma_{\mathbf{k}}}\right)\right.$$
$$\left. + [1 + e^{-i\gamma'_{\mathbf{p}'}}]\left(M^r_{\mathbf{K}',B;\mathbf{K},A}e^{i\gamma_{\mathbf{p}'}} - \left[M^r_{\mathbf{K},B;\mathbf{K}',A}\right]^* e^{-i\gamma_{\mathbf{k}}}\right)\right\}$$
$$= -\delta_{\mathbf{Q},\mathbf{q}_0}\frac{w}{2\hbar v_D q_0}[1 + e^{i\gamma_{\mathbf{p}'}}]\left\{M^r_{\mathbf{K}',B;\mathbf{K},A}e^{i\gamma_{\mathbf{p}'}} - \left[M^r_{\mathbf{K},B;\mathbf{K}',A}\right]^* e^{-i\gamma_{\mathbf{k}} - i\gamma'_{\mathbf{p}'}}\right\}$$
$$= \delta_{\mathbf{Q},\mathbf{q}_0}\frac{w}{2\hbar v_D q_0}[1 - e^{-i(\theta_{\pi'} - \theta/2)}]\left\{M^r_{\mathbf{K}',B;\mathbf{K},A}e^{-i(\theta_{\pi'} - \theta/2)} + \left[M^r_{\mathbf{K},B;\mathbf{K}',A}\right]^* e^{i\theta_{\kappa} + i\theta'_{\pi'}}\right\}. \tag{113}$$

In the last step, we have approximated the phases in the Dirac approximation. Taking the absolute value, averaging over the Fermi circles, and noting that as a result of the average over θ_{κ}, the two terms contribute incoherently gives

$$\int \frac{d\theta_{\mathbf{k}}}{2\pi} \frac{d\theta_{\mathbf{p}'}}{2\pi} |\langle \mathbf{p}', -; r, \mathbf{Q} | \mathcal{T}_{\text{inel}} | \mathbf{k}, + \rangle|^2 \to \frac{w^2}{2(\hbar v_D q_0)^2} \left\{ \left| M^r_{\mathbf{K}',B;\mathbf{K},A} \right|^2 + \left| M^r_{\mathbf{K},B;\mathbf{K}',A} \right|^2 \right\}. \tag{114}$$

Inserting this into Eq. (102) and accounting for phonon emission in the substrate, we find

$$\frac{d^2 \delta I_{\text{intra}}}{dV_b^2} = G_{\text{incoh}} \sum_r \delta(V_b - \hbar\omega_{r,\mathbf{q}_0}/e) \frac{\kappa_F^2 \Omega}{(\hbar v_D q_0)^2} \left\{ \left| M^r_{\mathbf{K}',B;\mathbf{K},A} \right|^2 + \left| M^r_{\mathbf{K},B;\mathbf{K}',A} \right|^2 \right\}. \tag{115}$$

Both matrix elements contribute equally, so that we find the result

$$\frac{d^2 \delta I_{\text{intra}}}{dV_b^2} = G_{\text{incoh}} \sum_r \delta\left(V_b - \hbar\omega_{r,\mathbf{q}_0}/e\right) \left(\kappa_F \ell_{r,\mathbf{q}_0}\right)^2 \frac{\beta^2 \Omega_{\text{uc}}}{(\hbar v_D q_0)^2} \left| \sum_j \left[\boldsymbol{\epsilon}^A_{r,\mathbf{q}_0} e^{-i\mathbf{q}_0 \cdot \mathbf{e}_j} - \boldsymbol{\epsilon}^B_{r,\mathbf{q}_0} \right] \right|^2. \tag{116}$$

The behavior in the limit of small twist angle, i.e., $\mathbf{q}_0 \to 0$, depends on the type of phonon mode. For acoustic phonons, the dependencies on q_0 of the last two factors on the right-hand side compensate to yield a constant. At charge neutrality, $(\kappa_F \ell_r)^2 \sim \omega_{r,\mathbf{q}_0}$ implies that the overall expression reduces linearly with decreasing $q_0 \to 0$. Away from charge neutrality, the peak grows as $1/q_0$. For optical phonons, the last factor as well as $(\kappa_F \ell_r)^2$ remain constant for $q_0 \to 0$. Thus, the peak grows as $1/q_0^2$ as the twist angle decreases both at and away from charge neutrality.

We end this section by remarking that making the replacement derived in the Appendix A, one can reproduce the contribution of acoustic phonons due to the interlayer electron-phonon interaction from Eq. (116) for the intralayer coupling.

VII. CONCLUSIONS

While STM probes local tunneling in real space, the QTM exploits tunneling which is local in reciprocal space. This makes the QTM ideally suited to probing momentum-dependent dispersions. Elastic tunneling between twisted graphene layers requires voltages above a twist-angle-dependent threshold. This opens a window at small voltages, in which the tunneling current in a QTM is purely inelastic and gives access to collective-mode dispersions along certain lines in momentum space. In this paper, we illustrated this modality by developing a comprehensive and analytical theory for phonon spectroscopy in twisted graphene-graphene junctions. We showed that intralayer and interlayer electron-phonon couplings give contributions to the inelastic tunneling current, and that the dominant contribution can come from various processes for a specific phonon mode, as summarized in Table I. Phonon spectroscopy relies on (i) small Fermi surfaces of the tip and sample layers and (ii) the Fermi velocity being large compared to the phonon velocity. These conditions are also satisfied by other van der Waals materials. Thus, graphene-graphene junctions should be considered as a model system illustrating a more broadly applicable technique. It is also conceivable that tip layers with different arrangements of Fermi pockets give access to phonon dispersions and electron-phonon couplings along additional directions in momentum space.

The intralayer electron-phonon coupling has two contributions at long wavelengths. Phonon-induced modifications in the bond lengths can be described as a gauge field within the effective Dirac description of the graphene band structure. The conventional deformation potential associated with compressions and dilations of the lattice enters the Dirac equation as a scalar potential. It has been difficult to reliably extract corresponding coupling constants from experiment. Provided that the contribution of intralayer electron-phonon coupling contributes significantly in QTM measurements, we find that for small twist angles, the gauge coupling is much stronger for transverse acoustic modes, while the deformation potential couples only to longitudinal acoustic modes. Thus, QTM measurements may well resolve this issue in addition to providing direct access to coupling constants.

Our considerations were limited to situations in which tunneling preserves the valley. In practice, rotating the tip by $\pi/3$ merely replaces the tip's K valley by the K' valley. This makes the twist-angle-dependent tunneling invariant under $\pi/3$ rotations. Moreover, every phonon mode contributes twice at a particular twist angle, due to scattering between identical as well as opposite valleys. Both contributions become symmetric at a twist angle of $\pi/6$, making this a symmetry point of the measured spectra.

Our analytical considerations focused on zero temperature, where only phonon emission contributes to the inelastic tunneling current. At finite temperatures, also phonon absorption contributes. We also did not consider higher-order umklapp processes. These do affect elastic scattering at specific angles, making their contribution to inelastic tunneling an interesting question. Both effects can be included by straightforward extension of the calculations presented here.

Our considerations further neglected the finite size of the tunneling contact. A finite contact area limits the accuracy of momentum conservation in the tunneling process. This enables a background elastic-tunneling current at all angles. The magnitude of this background depends sensitively on the contact shape. We can evaluate the tunneling amplitude for a finite contact area following the discussion of the scattering picture in Sec. II E. While a rectangular contact would

have independent Lorentzian-type broadenings of the x and y components of the momentum, these components are coupled in more general situations. For instance, one would expect just a single Lorentzian broadening for a weakly disordered contact. The background current is dominated by the tail of these broadening functions, which yields different behaviors in these two cases.

In addition to phonon dispersions, QTM measurements also give access to momentum-resolved electron-phonon couplings. In view of the multiple inelastic tunneling processes, extracting this information from experiment must rely on theoretical results of the kind that we provide in this paper. Corresponding results promise to shed light on the nature of superconductivity and the linear temperature dependence of the resistivity in magic-angle twisted bilayer graphene. Phonon spectroscopy based on QTM measurements may finally bring us closer to understanding the origin of these phenomena, which may or may not originate from electron-phonon coupling.

APPENDIX: INTERLAYER ELECTRON-PHONON COUPLING

It may seem accidental that for acoustic phonons, the contributions of first- and second-order perturbation theory are of the same order. Here, we rationalize this result by showing that the two kinds of electron-phonon couplings can be made to look similar by means of a gauge transformation.

Focusing on the contribution of T_0, the interlayer tunneling can be written in the continuum limit as

$$\mathcal{H}_T = \int d\mathbf{r}\, \psi_{t,\alpha}^\dagger(\mathbf{r}) T_0^{\alpha\beta}(\mathbf{r}) e^{i\mathbf{K}\cdot[\mathbf{u}(\mathbf{r}+\boldsymbol{\tau}_\alpha)-\mathbf{u}'(\mathbf{r}+\boldsymbol{\tau}_\beta)]} \psi_{b,\beta}(\mathbf{r}) + \text{H.c.} \tag{A1}$$

The phonon displacements of the two layers contribute a phase to the interlayer tunneling, akin to a vector potential introducing a Peierls phase. This phase can be eliminated by a gauge transformation

$$\psi_{b,\beta}(\mathbf{r}) \rightarrow \psi_{b,\beta}(\mathbf{r}) e^{i\mathbf{K}\cdot\mathbf{u}'(\mathbf{r}+\boldsymbol{\tau}_\beta)} \tag{A2}$$

and an analogous transformation for the upper layer. This is implemented by the unitary transformation

$$\mathcal{U} = \exp\left\{ -i\sum_\alpha \int d\mathbf{r}\, \mathbf{K}\cdot[\mathbf{u}(\mathbf{r}+\boldsymbol{\tau}_\alpha)\psi_{t,\alpha}^\dagger(\mathbf{r})\psi_{t,\alpha}(\mathbf{r}) - \mathbf{u}'(\mathbf{r}+\boldsymbol{\tau}_\alpha)\psi_{b,\alpha}^\dagger(\mathbf{r})\psi_{b,\alpha}(\mathbf{r})] \right\}. \tag{A3}$$

Apart from eliminating the phase factor including the phonon displacements in H_T, this transforms the electronic Hamiltonian of the graphene layers. (An additional contribution due to the transformation of the phonon Hamiltonian is small in the parameter c/v_D.) We revert to a tight-binding description for convenience. Then, the unitary transformation becomes

$$\mathcal{U} = \exp\left\{ -i\sum_\alpha \sum_\mathbf{R} \mathbf{K}\cdot[\mathbf{u}(\mathbf{R}+\boldsymbol{\tau}_\alpha)c_t^\dagger(\mathbf{R}+\boldsymbol{\tau}_\alpha)c_t(\mathbf{R}+\boldsymbol{\tau}_\alpha) - \mathbf{u}'(\mathbf{R}'+\boldsymbol{\tau}_\alpha')c_b^\dagger(\mathbf{R}'+\boldsymbol{\tau}_\alpha')c_b(\mathbf{R}'+\boldsymbol{\tau}_\alpha')] \right\}. \tag{A4}$$

The graphene Hamiltonian, say for the top layer, transforms into

$$\mathcal{H} = -t^\parallel \sum_\mathbf{R} \sum_{\mathbf{e}_j} e^{i\mathbf{K}\cdot[\mathbf{u}(\mathbf{R})-\mathbf{u}(\mathbf{R}+\mathbf{e}_j)]} c_t^\dagger(\mathbf{R}+\mathbf{e}_j)c_t(\mathbf{R}) + \text{H.c.} \tag{A5}$$

Expanding the exponential to linear order in the atomic displacements, we find that the resulting electron-phonon coupling takes essentially the same form as the intralayer electron-phonon coupling of the top layer, except for the fact that it is phase shifted by $\pi/2$ due to the factor of i. The calculation for the bottom layer proceeds by complete analogy and generates a corresponding electron phonon coupling for the bottom layer.

With this form of the interlayer electron-phonon coupling, we readily see that the results for intralayer and interlayer electron-phonon coupling are related by the replacement

$$\frac{\partial t^\parallel}{\partial a} \leftrightarrow it^\parallel \mathbf{K}. \tag{A6}$$

One can check that making this replacement in Eq. (116) indeed reproduces the contribution of in-plane phonons to Eq. (104) for the interlayer electron-phonon coupling.

The vectorial nature of the interlayer coupling implies that it is strongly twist-angle dependent, with coupling to

transverse phonons dominating over longitudinal phonons. This relation also makes it explicit that apart from this difference in their twist-angle dependence, the two types of electron-phonon couplings give contributions of the same order, when making the (approximate) replacement $\partial t^{\parallel}/\partial a \to t^{\parallel}/a$. In particular, we expect that the parametric dependencies on electron filling and phonon wave vector are identical in both cases. Moreover, the relative phase shift of $\pi/2$ implies that there are no interference terms between the two contributions.

It is also interesting to note that the dependence of the electron-phonon coupling on the phonon momentum $\mathbf{Q}$ is seemingly different before and after the gauge transformation.

Focusing on acoustic phonons, the coupling before the gauge transformation diverges as $1/\sqrt{Q}$ due to the linear dispersion. This dependence emerges from the zero-point amplitude of the phonon displacements. In contrast, after the transformation, there is an additional factor of Q due to the difference in phonon displacements of neighboring graphene sites. Thus, the coupling behaves as $\sqrt{Q}$. This shows that the coupling is consistent with the expectation that it vanishes in the long-wavelength limit corresponding to near-uniform relative shifts of the top and bottom layers. Of course, in keeping with gauge invariance, physical results are identical in both cases, with the singular Q dependence emerging from the energy denominator after the gauge transformation.

[1] 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 behaviour at half-filling in magic-angle graphene superlattices, Nature (London) **556**, 80 (2018).

[2] 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).

[3] E. Y. Andrei, D. K. Efetov, P. Jarillo-Herrero, A. H. MacDonald, K. F. Mak, T. Senthil, E. Tutuc, A. Yazdani, and A. F. Young, The marvels of moiré materials, Nat. Rev. Mater. **6**, 201 (2021).

[4] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, F. von Oppen, K. Watanabe, T. Taniguchi, and S. Nadj-Perge, Electronic correlations in twisted bilayer graphene near the magic angle, Nat. Phys. **15**, 1174 (2019).

[5] A. Kerelsky, L. J. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, A. Rubio, and A. N. Pasupathy, Maximized electron interactions at the magic angle in twisted bilayer graphene, Nature (London) **572**, 95 (2019).

[6] Y. Xie, B. Lian, B. Jäck, X. Liu, C.-L. Chiu, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, Spectroscopic signatures of many-body correlations in magic-angle twisted bilayer graphene, Nature (London) **572**, 101 (2019).

[7] Y. Jiang, X. Lai, K. Watanabe, T. Taniguchi, K. Haule, J. Mao, and E. Y. Andrei, Charge order and broken rotational symmetry in magic-angle twisted bilayer graphene, Nature (London) **573**, 91 (2019).

[8] M. Oh, K. P. Nuckolls, D. Wong, R. L. Lee, X. Liu, K. Watanabe, T. Taniguchi, and A. Yazdani, Evidence for unconventional superconductivity in twisted bilayer graphene, Nature (London) **600**, 240 (2021).

[9] K. P. Nuckolls, R. L. Lee, M. Oh, D. Wong, T. Soejima, J. P. Hong, D. Călugăru, J. Herzog-Arbeitman, B. A. Bernevig, K. Watanabe, T. Taniguchi, N. Regnault, M. P. Zaletel, and A. Yazdani, Quantum textures of the many-body wavefunctions in magic-angle graphene, Nature (London) **620**, 525 (2023).

[10] U. Zondiner, A. Rozen, D. Rodan-Legrain, Y. Cao, R. Queiroz, T. Taniguchi, K. Watanabe, Y. Oreg, F. von Oppen, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, Cascade of phase transitions and Dirac revivals in magic-angle graphene, Nature (London) **582**, 203 (2020).

[11] D. Wong, K. P. Nuckolls, M. Oh, B. Lian, Y. Xie, S. Jeon, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, Cascade of electronic transitions in magic-angle twisted bilayer graphene, Nature (London) **582**, 198 (2020).

[12] A. Rozen, J. M. Park, U. Zondiner, Y. Cao, D. Rodan-Legrain, T. Taniguchi, K. Watanabe, Y. Oreg, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, Entropic evidence for a Pomeranchuk effect in magic-angle graphene, Nature (London) **592**, 214 (2021).

[13] Y. Xie, A. T. Pierce, J. M. Park, D. E. Parker, E. Khalaf, P. Ledwith, Y. Cao, S. H. Lee, S. Chen, P. R. Forrester, K. Watanabe, T. Taniguchi, A. Vishwanath, P. Jarillo-Herrero, and A. Yacoby, Fractional Chern insulators in magic-angle twisted bilayer graphene, Nature (London) **600**, 439 (2021).

[14] S. Grover, M. Bocarsly, A. Uri, P. Stepanov, G. Di Battista, I. Roy, J. Xiao, A. Y. Meltzer, Y. Myasoedov, K. Pareek, K. Watanabe, T. Taniguchi, B. Yan, A. Stern, E. Berg, D. K. Efetov, and E. Zeldov, Chern mosaic and Berry-curvature magnetism in magic-angle graphene, Nat. Phys. **18**, 885 (2022).

[15] A. Inbar, J. Birkbeck, J. Xiao, T. Taniguchi, K. Watanabe, B. Yan, Y. Oreg, A. Stern, E. Berg, and S. Ilani, The quantum twisting microscope, Nature (London) **614**, 682 (2023).

[16] J. Xiao, Y. Vituri, and E. Berg, Probing the order parameter symmetry of two-dimensional superconductors by twisted Josephson interferometry, Phys. Rev. B **108**, 094520 (2023).

[17] V. Peri, S. Ilani, P. A. Lee, and G. Refael, Probing quantum spin liquids with a quantum twisting microscope, Phys. Rev. B **109**, 035127 (2024).

[18] F. Pichler, W. Kadow, C. Kuhlenkamp, and M. Knap, Probing magnetism in moiré heterostructures with quantum twisting microscopes, Phys. Rev. B **110**, 045116 (2024).

[19] J. Birkbeck, J. Xiao, A. Inbar, T. Taniguchi, K. Watanabe, E. Berg, L. Glazman, F. Guinea, F. von Oppen, and S. Ilani, Measuring phonon dispersion and electron-phason coupling in twisted bilayer graphene with a cryogenic quantum twisting microscope, arXiv:2407.13404.

[20] F. Wu, A. H. MacDonald, and I. Martin, Theory of phonon-mediated superconductivity in twisted bilayer graphene, Phys. Rev. Lett. **121**, 257001 (2018).

[21] Y. W. Choi and H. J. Choi, Strong electron-phonon coupling, electron-hole asymmetry, and nonadiabaticity in magic-angle twisted bilayer graphene, Phys. Rev. B **98**, 241412(R) (2018).

[22] B. Lian, Z. Wang, and B. A. Bernevig, Twisted bilayer graphene: A phonon-driven superconductor, Phys. Rev. Lett. **122**, 257002 (2019).

[23] F. Wu, E. Hwang, and S. Das Sarma, Phonon-induced giant linear-in-t resistivity in magic angle twisted bilayer graphene: Ordinary strangeness and exotic superconductivity, Phys. Rev. B **99**, 165112 (2019).

[24] M. Angeli, E. Tosatti, and M. Fabrizio, Valley Jahn-Teller effect in twisted bilayer graphene, Phys. Rev. X **9**, 041010 (2019).

[25] T. Cea and F. Guinea, Coulomb interaction, phonons, and superconductivity in twisted bilayer graphene, Proc. Natl. Acad. Sci. USA **118**, e2107874118 (2021).

[26] C. Lewandowski, D. Chowdhury, and J. Ruhman, Pairing in magic-angle twisted bilayer graphene: Role of phonon and plasmon umklapp, Phys. Rev. B **103**, 235401 (2021).

[27] Y.-Z. Chou, F. Wu, J. D. Sau, and S. Das Sarma, Acoustic-phonon-mediated superconductivity in bernal bilayer graphene, Phys. Rev. B **105**, L100503 (2022).

[28] C. Chen, K. P. Nuckolls, S. Ding, W. Miao, D. Wong, M. Oh, R. L. Lee, S. He, C. Peng, D. Pei, Y. Li, S. Zhang, J. Liu, Z. Liu, C. Jozwiak, A. Bostwick, E. Rotenberg, C. Li, X. Han, D. Pan *et al.*, Strong inter-valley electron-phonon coupling in magic-angle twisted bilayer graphene, arXiv:2303.14903.

[29] C.-X. Liu, Y. Chen, A. Yazdani, and B. A. Bernevig, Electron–k-phonon interaction in twisted bilayer graphene, Phys. Rev. B **110**, 045133 (2024).

[30] G. Sharma, I. Yudhistira, N. Chakraborty, D. Y. H. Ho, M. M. A. Ezzi, M. S. Fuhrer, G. Vignale, and S. Adam, Carrier transport theory for twisted bilayer graphene in the metallic regime, Nat. Commun. **12**, 5737 (2021).

[31] H. Ishizuka, A. Fahimniya, F. Guinea, and L. Levitov, Purcell-like enhancement of electron–phonon interactions in long-period superlattices: Linear-temperature resistivity and cooling power, Nano Lett. **21**, 7465 (2021).

[32] S. Das Sarma and F. Wu, Strange metallicity of moiré twisted bilayer graphene, Phys. Rev. Res. **4**, 033061 (2022).

[33] S. A. Hartnoll and A. P. Mackenzie, Colloquium: Planckian dissipation in metals, Rev. Mod. Phys. **94**, 041002 (2022).

[34] H. Ishizuka and L. Levitov, Wide-range T^2 resistivity and umklapp scattering in moiré graphene, New J. Phys. **24**, 052001 (2022).

[35] A. Aydin, J. Keski-Rahkonen, and E. J. Heller, Quantum acoustics unravels planckian resistivity, Proc. Natl. Acad. Sci. USA **121**, e2404853121 (2024).

[36] S. M. Davis, Y.-Z. Chou, F. Wu, and S. Das Sarma, Phonon-limited resistivity of multilayer graphene systems, Phys. Rev. B **107**, 045426 (2023).

[37] H. Ochoa and R. M. Fernandes, Extended linear-in-T resistivity due to electron-phason scattering in moiré superlattices, Phys. Rev. B **108**, 075168 (2023).

[38] H. Suzuura and T. Ando, Phonons and electron-phonon scattering in carbon nanotubes, Phys. Rev. B **65**, 235412 (2002).

[39] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, The electronic properties of graphene, Rev. Mod. Phys. **81**, 109 (2009).

[40] R. Bistritzer and A. H. MacDonald, Moiré bands in twisted double-layer graphene, Proc. Natl. Acad. Sci. USA **108**, 12233 (2011).

[41] F. von Oppen, F. Guinea, and E. Mariani, Synthetic electric fields and phonon damping in carbon nanotubes and graphene, Phys. Rev. B **80**, 075420 (2009).

[42] H. Ochoa, Moiré-pattern fluctuations and electron-phason coupling in twisted bilayer graphene, Phys. Rev. B **100**, 155426 (2019).

[43] M. Koshino and N. N. T. Nam, Effective continuum model for relaxed twisted bilayer graphene and moiré electron-phonon interaction, Phys. Rev. B **101**, 195425 (2020).