Intrinsically Undamped Plasmon Modes in Narrow Electron Bands

Cyprian Lewandowski, Leonid Levitov
Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA02139, USA

We argue that surface plasmons in moiré graphene feature an interesting regime in which Landau damping (dissipation via electron-hole pair excitation) is completely quenched. This surprising behavior is made possible by strong coupling in narrow-band systems, characterized by large values of the “fine structure” constant $\alpha = e^2/\hbar c v_F$. Dissipation quenching occurs when dispersing plasmon modes rise above the particle-hole continuum, extending into the forbidden energy gap which is completely free of particle-hole excitations. The effect is predicted to be prominent in magic twist angle moiré graphene, where flat bands feature $\alpha \gg 1$. The absence of Landau damping enhances optical coherence, leading to speckle-like interference and other striking signatures in moiré graphene plasmonics that are directly accessible in ongoing near-field imaging experiments.

Landau damping, a process by which collective mode decays into electron-hole pairs, is often regarded as an inherent attribute of graphene plasmon excitations. Here we predict extinction of this dissipation mechanism in materials with narrow electron bands such as twisted bilayer graphene (TBG). Intrinsically undamped plasmons in narrow-band materials arise due to large fine structure parameter values $\alpha = e^2/\hbar c v_F$: strong interactions push plasmon dispersion into the energy gap above the particle-hole (p-h) continuum, as illustrated in Fig.1. As discussed below, in this region plasmons are decoupled from particle-hole pair excitations. Dissipation quenching, which is a surprising manifestation of strong interactions, is a robust effect, present in all narrow electron bands, our analysis will focus on TBG flat bands, a system of high current interest in which undamped plasmons can be directly probed.

The new behavior is illustrated in Fig.1 showing plasmon mode for a narrow-band model that mimics the key features of the TBG band as described below. Mode dispersion (red line) is conventional at energies less than the bandwidth, $\omega \lesssim W$. At lowest energies, plasmon mode is positioned outside the p-h continuum; this suppresses the $T = 0$ Landau damping, but does not protect the mode from decaying into p-h excitations through disorder scattering or from the conventional $T > 0$ Landau damping. At higher energies, $\omega \sim 2E_F$ (marked by arrow in Fig.1), the mode plunges into p-h continuum and is Landau-damped at $2E_F \lesssim \omega \lesssim 2W$ even at $T = 0$. However, a dramatic change occurs once the mode rises above the p-h continuum. In the forbidden gap region, $\omega > 2W$, it becomes damping-free, since at these energies there are no free e-h pairs into which plasmon could decay. This behavior is manifest in the

![Graph showing plasmon behavior](image-url)
$T$ dependence of the resonances, which are washed out with increasing temperature at $\omega \lesssim W$ but remain sharp at $\omega > W$ even at $T \sim E_F$ (Fig. 1 panels b and c).

As we will see, mode dispersion has a square-root form characteristic of two-dimensional (2D) plasmons.\(^2\)\(^2\)\(^3\)

$$\omega_p(q) = \sqrt{\beta_q q},$$

(1)

with a weak $q$ dependence in $\beta_q$, see \[^{11}\]. This expression, however, is valid not just at low energies $0 < \omega < W$ but also at higher energies, $\omega \gg W$, where the mode is undamped. While the dispersion in (1) is of the conventional 2D plasmon form, we emphasize that here it takes on a new role, as it describes the plasmon mode at frequencies much higher than the carrier bandwidth, extending to

$$\omega_p \sim \sqrt{\alpha W} \gg W, \quad \alpha \sim 20 - 30,$$

(2)

where we use typical $\alpha$ values for flat bands in magic-angle moiré graphene. Another difference from conventional plasmons is that the dispersion in (1) is not limited to long wavelengths. Indeed, as illustrated Fig. 1, it extends to wavenumbers which are on the order of the superlattice Brillouin zone width.

Furthermore, the wavelengths of the undamped plasmons are only 2-3 times greater than the moiré superlattice period. Such short wavelengths are of great interest for plasmonics and are within resolution of the state-of-the-art scanning near-field microscopy techniques\[^{4, 5}\] (currently as good as 10nm \[^{29, 30}\]). In addition to measuring plasmon dispersion, these techniques can be used to directly visualize the qualitative change in the damping character and strength. Enhanced optical coherence will manifest itself in striking speckle-like interference, as illustrated in Fig. 1 panel d) and Fig. 2.

Indeed, because of the absence of Landau damping at the energies of interest, $\omega > W$, and also because these energies are smaller than carbon optical phonon energies, the dominant dissipation mechanism will be elastic scattering by disorder. At low energies, where plasmon mode coexists with p-h continuum, disorder scattering mostly assists Landau damping, allowing plasmons to decay into p-h pairs by passing some of their momentum to the lattice. However, at the energies above p-h continuum, $\omega > W$, where decay into e-h pairs is quenched, disorder will lead to predominantly elastic scattering among plasmon excitations. Such scattering preserves optical coherence and is expected to produce speckle patterns in spatial near-field images, as illustrated in Fig. 1.

To model this behavior we consider the signal $S(r)$, excited by the scanning tip and measured at the same location. Monochromatic plasmon excitation at energy $E$ is scattered by impurities or defects, and, upon returning to the tip, produces signal

$$S(r) = J_0 \int d^2 r' G_E(r-r')\eta(r')G_E(r' - r),$$

(3)

where $\eta(r)$ is the disorder potential, $J_0$ is excitation amplitude, and $G_E(r)$ is the Green’s function of the plasmon excitation (see supplemental information). The spatial signal, Fig. 2, exhibits a characteristic speckle pattern familiar from laser physics. In graphene plasmonics, speckle-like interference provides a direct manifestation of optical coherence enhancement in the absence of Landau damping. Consequently, the Fourier transform of the image, $S_k = \int d^2 r S(r)e^{-ikr}$, yields power spectrum $|S_k|^2$ that features a ring-like structure; the ring of radius $k = 2q_0$, where $q_0$ is the plasmon excitation wavenumber (see Fig. 1 inset). Simple calculation, described in supplemental information, predicts a sharp peak at the ring:

$$|S_k|^2 \sim \frac{|\eta_k|^2}{|k^2 - 4(q_0 - i\delta)^2|},$$

(4)

where $\delta$ is a parameter characterizing extrinsic damping due to phonon scattering and other inelastic processes. In the fully coherent regime ($\delta = 0$) the quantity $|S_k|^2$ exhibits a power law singularity at the ring, $k = 2q_0$. As the amount of incoherent scattering increases, the peak is gradually washed out. This behavior is illustrated in Fig. 2.

We note that recent work\[^{15}\] analyzed interband plasmon excitations in TBG dominated by polarization of the bands above the flat band, which are distinct from the flat-band plasmons analyzed here. Recent experiment\[^{10}\]...
reported observation of plasmons in TBG, however their appeal for constructing intrinsically protected collective modes remained unnoticed in graphene literature. Also, plasmons in narrow bands were analyzed in the context of high-$T_c$ superconductivity [23], finding that plasmon mode can rise above the flat band. However, in cuprates, unlike moiré graphene, the narrow band is not separated from higher bands by a forbidden energy gap, and thus the modes such as those studied in Ref. [23] will plunge into a higher band before acquiring an undamped character. Our prediction of the intrinsically undamped plasmons in TBG and their potential subsequent observation therefore constitute an uncharted territory.

Next, we present analysis of the hexagonal-lattice toy model that mimics the key features of Landau-damped and intrinsically undamped modes in TBG. The hexagonal-lattice tight binding model possesses the same symmetry and the same number of sub-bands as the flat band in TBG. We match the energy and length scales by choosing the width of a single band $W$ and the hexagonal lattice period $L_M$ identical to the parameters in TBG: $W = 3.75$ meV and $L_M = a/2\sin(\theta/2)$ is the moiré superlattice periodicity. For the magic angle value $\theta = 1.05$, using carbon spacing $a = 0.246$ nm, this gives $L_M = 13.4$ nm. To ensure that a unit cell of the toy model accommodates 4 electrons just as the moiré cell does in TBG, we make the toy model fourfold degenerate. Compare to plasmons for the actual TBG model, presented below, will help us to identify the features which are general, as well as those which are a specific property of TBG.

Our nearest-neighbor tight-binding Hamiltonian is

$$H_{\text{toy}} = \begin{pmatrix} 0 & h_k \\ h_k^* & 0 \end{pmatrix}, \quad h_k = \frac{W}{3} \sum_{e_j} e^{i k \cdot e_j},$$

with the hopping matrix element $W/3$ to nearest neighbors at positions $e_j = (\cos(2\pi j/3), \sin(2\pi j/3))L_M/\sqrt{3}$, $j = 0, 1, 2$. Here $W$ is the bandwidth measured from Dirac point and the nearest neighbor distance $L_M/\sqrt{3}$ is chosen such that the lattice period of the hexagonal toy model matches the moiré superlattice period. Corresponding energies $E_{s,k}$ and eigenstates $\Psi_{s,k}$ are then

$$E_{s,k} = s |h_k|, \quad \Psi_{s,k} = \frac{1}{\sqrt{2}} \begin{pmatrix} s e^{i \varphi_k} \\ 1 \end{pmatrix},$$

where $\varphi_k = \arg h_k$ and the band index $s = \pm$ labels the conduction and valence band.

Plasmons can be obtained from the nodes of the complex dielectric function, describing the dynamical response of a material to an outside electric perturbation,

$$\varepsilon(\omega, q) = 1 - V_q \Pi(\omega, q).$$

Here $V_q = 2\pi e^2/\kappa q$ is the Coulomb interaction in a medium with a background dielectric constant $\kappa$ and $\Pi(\omega, q)$ is the electron polarization function. Within the random phase approximation, this quantity is given by

$$\Pi(\omega, q) = 4 \sum_{k, s, s'} \frac{(f_{s,k+q} - f_{s',k}) F_{s,k+q,k}^{s,s'}}{E_{s,k+q} - E_{s',k} - \omega - i0^+}.$$  (8)

Here summation $\sum_k$ denotes integration over the Brillouin zone (BZ), the indices $s, s'$ run over the electron bands and the factor of 4 in front of the summation accounts for the fourfold degeneracy of the toy model. Here $f_{s,k}$ is the equilibrium distribution $1/(e^{\beta(E_{s,k} - E_F)} + 1)$, and $F_{s,k+q,k}^{s,s'}$ describes band coherence factors. For our toy model

$$F_{k+q,k}^{s,s'} = |\langle \Psi_{s,k+q} | \Psi_{s',k} \rangle|^2 = \frac{1 + s s' \cos(\varphi_k + \varphi_k)}{2}.$$  (9)

where $\Psi_{s,k}$ are pseudospinors given in [23].

As we now show, an analytic expression for plasmon dispersion can be obtained, describing both the Landau-damped and the undamped cases in a unified way. We first rewrite (8) by performing a standard replacement $k + q \rightarrow -k$ in the term containing $f_{s,k+q}$ followed by $-k - q \rightarrow k + q, k$ justified by the $k \rightarrow -k$ time-reversal symmetry. This gives

$$\Pi(\omega, q) = \sum_{k, s, s'} \frac{f_{s',k+q}^{s,s'} (E_{s',k+q} - E_{s,k+q})}{(E_{s,k+q} - E_{s',k})^2 - (\omega + i0)^2}.$$  (10)

The behavior of this expression at small $q$, which will be of interest for us, can be found in a closed form. In the small-$q$ limit the coherence factors behave as

$$F_{k+q,k}^{s=s'} \approx 1, \quad F_{k+q,k}^{s=-s'} \approx \frac{1}{4} (q \cdot \nabla_k \varphi_k)^2.$$  (11)

The values $O(1)$ for intraband transitions and $O(q^2)$ for interband transitions might suggest that the polarization function is dominated by the intraband transitions. However, as we now show, the interband and intraband contributions are of the same order of magnitude.

Indeed, the intraband contributions, $s = s'$, can be rewritten by noting that upon integration over $k$ only the even-$k$ part of series expansion $E_{s,k+q} - E_{s,k}$ survives, giving $\Pi_1(\omega, q) \approx \sum_{k,s} f_{s,k} (E_{s,k+q} + E_{s,k} - 2E_{s,k})$. Expanding in small $q$, we have

$$\Pi_1(\omega, q) \approx \frac{4}{\omega^2} \sum_{k,s} f_{s,k} (q \cdot \nabla_k)^2 E_{s,k}.$$  (12)

As a sanity check, for parabolic band $E_{s,k} = k^2/2m$ this yields the familiar “classical acceleration” result $\Pi(\omega, q) = \frac{q^2}{\omega^2 m}$. 

$$\varepsilon(\omega, q) = 1 - V_q \Pi(\omega, q).$$
The interband contributions, \( s = -s' \), can be simplified by noting that \( E_{s',k+q} \approx -E_{s',k} \), giving

\[
\Pi_2(\omega, q) \approx \sum_{k,s} f_{s,k} \left( \frac{E_{s,k}(q \cdot \nabla k^s k)}{4E_{s,k}^2 - (\omega + i0)^2} \right)^2. \tag{13}
\]

As a sanity check, at \( T = 0 \) the imaginary part of \( \Pi_2 \), describing interband transitions, is nonzero only for \( 2E_F < \omega < 2W \), as expected. The real part of \( \Pi_2 \) is negative at all \( \omega \) and positive at large \( \omega \) because the valence band contribution dominates over that of the conduction band.

Plasmon dispersion \( \omega_p \) is given by the solution of the equation \( \varepsilon(\omega, q) = 0 \) with \( \Pi = \Pi_1 + \Pi_2 \). Comparing the \( \omega \) dependence of \( \Pi_1 \) and \( \Pi_2 \), we see that at small frequencies, \( \omega < 2E_F \), the intraband contribution \( \Pi_1 \) dominates. This gives the dispersion in \( \omega \) with

\[
\beta_q = \beta_0 + \beta_1 q + O(q^2) \tag{14}
\]

where the leading term \( \beta_0 = 4\alpha v_F E_F / \hbar \) originates from \( \Pi_1 \) (see supplemental information) and the subleading \( q \)-dependent contribution is due to \( \Pi_2 \). Negative sign of \( \Pi_2 \) translates into \( \beta_1 < 0 \), softening the dispersion at low frequencies. This behavior, which holds the limit \( \omega < 2E_F \), agrees with Refs. [1] [2] [22].

In the same manner we can obtain the dispersion at high frequencies, \( \omega > 2W \) (the intrinsically undamped regime). The analysis is again simplified by noting that, since \( \alpha = e^2 / \hbar c v_F \gg 1 \), the relevant values of \( q \) are small compared to the Brillouin zone size, and thus the small-\( q \) limit considered above is sufficient to describe this new behavior. Taking both the intraband and interband contributions in the asymptotic form \( \Pi_1 = \lambda_1 q^2 / \omega^2 \), \( \Pi_2 = \lambda_2 q^2 / \omega^2 \) where \( \lambda_1 \approx 2E_F / \hbar^2 \pi \), \( \lambda_2 \approx 2(W - E_F) / \hbar^2 \pi \) (see supplemental information), yields \( \omega_p \) with \( \beta = \beta_0 + \beta_1 = \frac{2\pi e^2}{\hbar}(\lambda_1 + \lambda_2) \). The first term is identical to \( \beta_0 \) found at low frequencies, the second term is of a positive sign, \( \lambda_2 > 0 \), describing hardening of the plasmon dispersion due to interband transitions.

In the undamped regime, plasmon frequency peaks at \( q \) values on the order of Brillouin zone scale. The peak value of \( \omega_p \), given in \( \Box \), can be found by estimating the energy differences \( E_{s,k+q} - E_{s',k} \) in \( \Box \) as \( W \) and noting that the coherence band factor for large \( q \) is in general non-vanishing and of order 1. This gives for the practically interesting case of \( E_F \sim W \) the result \( \omega_p \sim \sqrt{W} \), which agrees with the dispersion \( \omega_p = \sqrt{2q} = 2\sqrt{\alpha v_F W q / \hbar} \) provided that \( h c v_F q \) saturates at \( W \). Indeed the estimated values of \( \beta_0, \beta \) as compared to the fitted curve in Fig.\( \Box \) (see supplemental information) indicate that \( \omega_p = \sqrt{2\beta q} \) relation from \( \Box \) is a good approximation for the plasmon dispersion at both small and large \( q \).

The dielectric function of the two-band toy model faithfully reproduces all of the qualitative features expected for the TBG bandstructure. However, we find that, despite matching the bandwidth \( W \) and lattice period to those of TBG, the resulting plasmon dispersion extends to much higher energies than those that will be found below for the actual TBG bandstructure. This is simply because the two-band model does not account for the effects of interband polarization of higher electron bands, which renormalize the dielectric constant down and soften the plasmon dispersion. We account for this in the toy model case by rescaling the effective fine structure constant such that the resulting plasmon dispersion is comparable in magnitude to the TBG result. Specifically in the Fig.\( \Box \) we use an effective background dielectric constant \( \kappa = 12.12 \), which is four times larger than the dielectric constant \( \kappa = 3.03 \) corresponding to an air/TBG/hBN heterostructure.

Next, we turn to the analysis of plasmons in TBG flat bands at an experimentally relevant magic angle value \( \theta = 1.05^\circ \). To accurately describe the TBG bandstructure and eigenstates we employ the effective continuum Hamiltonian \( H_{TBG} \) introduced in Ref. [31]. The full discussion of the bandstructure details can be found in the supplemental material; here we only discuss two relevant energy scales: flat-band bandwidth \( W \), and the gap \( \Delta \) between the flat-bands and the rest of the bandstructure. With regard to \( W \) value we note that, technically, the bandwidth of the flat-bands, as predicted by the continuum mode \( H_{TBG} \), is on the order of \( W \approx 3.75 \) meV. However, the effective bandwidth relevant for the interband and intraband excitations is expected to be closer to \( W \approx 2 \) meV. This is because a vast majority of the bandstructure lies below 2 meV scale. Furthermore, the states with energies outside \( -2 \) meV \(< E < 2 \) meV are small-\( k \) and, therefore, their contributions to polarization function, [13], [12], evaluated at small \( q \) are strongly suppressed. Similarly the bandgap as predicted by the continuum model is \( \Delta \approx 11.75 \) meV, but the actual gap is still a subject of debate [32].

The definition of the polarization function for the TBG continuum model is identical to the one of the tight binding toy model, [8], with exception of requiring the valley and spin degrees of freedom to be accounted for explicitly, a larger number of electron bands and different coherence factors. The first two issues are resolved by promoting the band indices \( s, s' \) in [8] to composite labels \( n, m \), which range over all electron bands, spins \( \sigma \) and valleys \( \xi \) making the additional factor of 4 in front of [8] unnecessary. Instead of the toy model coherence factors we use the TBG coherence band factors \( F_{k+q,k}^m \), which are given by

\[
F_{k+q,k}^m = \left( \int d^2 r \Psi_{n,k+q,r}^\dagger e^{i q \cdot r} \Psi_{m,k,r} \right)^2, \tag{15}
\]

where \( \Psi_{n,k}(r) \) are the Bloch wavefunctions for momentum \( k \) and band/valley/spin composite label \( n \), which diagonalize the continuum Hamiltonian (see supplemen-
pairs of indices \( n, m \) such that (on average) we expect \( \omega^2 > (E_{n,k+q} - E_{n,k})^2 \) to hold, and pairs of indices \( n, m \) where the opposite \( \omega^2 < (E_{n,k+q} - E_{n,k})^2 \) is true. To leading order in \( \omega \) we can therefore rewrite the dielectric function as

\[
\varepsilon(\omega, q) \approx 1 + A(q) - \frac{B(q)}{\omega^2},
\]

where we defined two auxiliary functions:

\[
A(q) = \frac{8\pi e^2}{\kappa q} \sum_{k,n,m} \left| m \sum_{k,n,m} F_{km}^{nm}(E_{n,k+q} - E_{m,k}) \right|
\]

and

\[
B(q) = \frac{8\pi e^2}{\kappa q} \sum_{k,n,m} \left| m \sum_{k,n,m} F_{km}^{nm}(E_{n,k+q} - E_{m,k}) \right|
\]

Here the band summations \( \sum' \) and \( \sum'' \) run over bands such that \( \omega^2 > (E_{n,k+q} - E_{n,k})^2 \) and \( \omega^2 < (E_{n,k+q} - E_{n,k})^2 \), respectively. E.g. at large momenta, as seen in Fig.3, the plasmon mode lies in the gap between the flat and non-flat bands and hence the \( B(q) \) summation extends only over the flat bands, whereas the summation in \( A(q) \) includes all of the remaining combinations of band indices. This allows us to write a closed form expression for the plasmon dispersion as

\[
\omega_p^2 \approx \frac{B(q)}{\Gamma + A(q)},
\]

which must hold for both small and large \( q \). We consider these two limits separately.

At small \( q \) the matrix element of the Bloch wavefunctions, just as in the toy model case, favours the overlap between states from the same band. At the same time there are fewer states in the \( A(q) \) satisfying the condition \( \omega^2 > (E_{n,k+q} - E_{n,k})^2 \) and hence \( A(q) \) vanishes for small \( q \). This amounts to the plasmon dispersion \( \omega_p \) from [19] reducing to \( \omega_p^2 \approx B(q) \) and, by comparison with [12], we similarly expect a conventional 2D plasmon dispersion \( \omega_p = \sqrt{\beta q} q \) with \( \beta_q \) given by the series from [14]. As we see in Fig.3 the \( \omega_p^2 = \sqrt{\beta q} q \) dispersion is a valid description only at very small \( q \), as compared to the Fig.1, which can be traced back to higher bands softening the plasmon dispersion through the \( A(q) \) term in [19].

To determine how high the plasmon mode rises above the p-h continuum, we consider large \( q \) values comparable to the reciprocal lattice vector. The arguments similar to those in the toy model show that, since \( \alpha \gg 1 \), we have \( A(q) \gg 1 \). The dependence on the \( e^2/\kappa q \) ratio therefore cancels between the \( A(q) \) and \( B(q) \) functions resulting in the value of the plasmon dispersion \( \omega_p \approx \sqrt{B(q)/A(q)} \sim \sqrt{W\Delta} \approx 6.6 \text{ meV} \) being dictated only by the continuum model’s bandstructure parameters. This lack of explicit dependence on \( q \) suggests that, once the doping is such that \( \alpha \gg 1 \), the large-\( q \) value of
$\hbar \omega_p \approx \sqrt{W \Delta}$ becomes insensitive to doping (and hence Fermi velocity). This behavior is different than that in the toy model where $\omega_p \sim \sqrt{W}$ at large $q$. The relatively more week dependence on $\alpha$ in the TBG case is due to interband polarization involving higher bands, which significantly alters the effective dielectric constant. The weak $q$ dependence at large $q$ is in agreement with the properties of interband plasmons described in Ref. [15].

Before closing, we note that suppressing damping has always been central to the quest for tightly-confined loss-few low-loss surface plasmon excitations. An early approach utilized surface EM modes traveling at the edge of an air/metal boundary [23], in which dissipation is low because most of the EM field resides outside the metal, however the field confinement scale is on the order of optical wavelength. Next came surface plasmons propagating in high-mobility 2D GaAs-based electron gases and graphene-hBN heterostructures [27], which can provide deep-subwavelength confinement [3]. However, plasmons in these systems are prone to a variety of dissipation mechanisms, with Landau damping usually regarded as the one that sets the fundamental limit on possible plasmon wavelengths and corresponding lifetimes. The possibility to overcome this fundamental limitation in narrow band systems such as moiré graphene, discussed above, opens a new line of investigation in graphene plasmonics. Damping-free plasmons can enable novel interference phenomena, dissipationless photon-matter coupling, and other interesting behaviors. It is also widely expected that low-dissipation plasmons can lead to unique applications for photon-based Quantum Information Processing [11]. Furthermore, reduced damping has more immediate consequences, as it translates into enhanced optical coherence that can be directly probed by scanning near-field microscopy, as discussed above, providing a clear signature of the intrinsically undamped collective modes.

We thank Ali Fahimnia for useful discussions. This work was supported, in part, by the STC Center for Integrated Quantum Materials, NSF Grant No. DMR-1231319; and by Army Research Office Grant W911NF-18-1-0116.

[1] B. Wunsch, T. Stauber, F. Sols, F. Guinea, Dynamical polarization of graphene at finite doping, New J. Phys. 8, 318 (2006).
[2] E. H. Hwang, S. Das Sarma, Dielectric function, screening, and plasmons in two-dimensional graphene. Phys. Rev. B 75, 205408 (2007).
[3] M. Jablani, H. Buljan, M. Soljacic, Plasmonics in graphene at infrared frequencies. Phys. Rev. B 80, 245435 (2009).
[4] J. Chen, M. Badioli, P. Alonso-Gonzalez, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenovic, A. Centeno, A. Pesquera, P. Godignon, A. Zurutuza, N. Camara, F. J. Garcia De Abajo, R. Hillenbrand, F. H. L. Koppen, Optical Nano-Imaging of Gate-Tunable Graphene Plasmons, Nature 487, 77 (2012).
[5] Z. Fei, A. S. Rodin, G. O. Andreev, W. Bao, A. S. McLeod, M. Wagner, L. M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, M. M. Fogler, A. H. Castro Neto, C. N. Lau, F. Keilmann, D. N. Basov, Gate-Tuning of Graphene Plasmons Revealed by Infrared Nano-Imaging, Nature 487, 82 (2012).
[6] E. J. Mele, Commensuration and interlayer coherence in twisted bilayer graphene, Phys. Rev. B 81, 161405(R) (2010).
[7] P. San-Jose, J. González, F. Guinea, Non-Abelian gauge potentials in graphene bilayers. Phys. Rev. Lett. 108:216802 (2012).
[8] R. Bistritzer, A. H. MacDonald, Moiré bands in twisted double-layer graphene, PNAS 108 (30) 12233-12237 (2011).
[9] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Graphene Bilayer with a Twist: Electronic Structure, Phys. Rev. Lett. 99, 256802 (2007).
[10] S. Fang, and E. Kaxiras, Electronic structure theory of weakly interacting bilayers, Phys. Rev. B 93, 235153, (2016).
[11] M. Gullans, D. E. Chang, F. H. L. Koppen, F. J. Garcia de Abajo, and M. D. Lukin, Single-Photon Nonlinear Optics with Graphene Plasmons, Phys. Rev. Lett. 111, 247401 (2013).
[12] 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 556, 80 EP (2018).
[13] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Unconventional superconductivity in magic-angle graphene superlattices, Nature 556, 43 EP (2018).
[14] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Tuning superconductivity in twisted bilayer graphene, Science 363, 1059 (2019).
[15] T. Stauber and H.Kohler, Quasi-Flat Plasmonic Bands in Twisted Bilayer Graphene, Nano Lett. 16, 6844 (2016).
[16] F. Hu, S. R. Das, Y. Luan, T. F. Chung, Y. P. Chen, and Z. Fei, Real-Space Imaging of the Tailored Plasmons in Twisted Bilayer Graphene, Phys. Rev. Lett. 119, 247402 (2017).
[17] N. K. Emani, T. F. Chung, X. Ni, A. V. Kildishev, Y. P. Chen, and A. Boltasseva, Electrically tunable damping of plasmonic resonances with graphene, Nano Letters 12, 5202 (2012).
[18] T. Low and P. Avouris, Graphene Plasmonics for Terahertz to Mid-Infrared Applications, ACS Nano 8, 1086 (2014).
[19] A. Principi, G. Vignale, M. Carrega, and M. Polini, Bulk and shear viscosities of the two-dimensional electron liquid in a doped graphene sheet, Phys.Rev. B 88, 195405 (2013).
[20] M. Polini, R. Asgari, G. Borghi, Y. Barlas, T. Perez-Barneva, and A. H. MacDonald, Plasmons and the spectral function of graphene, Phys. Rev. B 77, 081411(2008).
[21] H. Yan, T. Low, W. Zhu, Y. Wu, M. Freitag, X. Li, J. Garcia , P. Avouris, and F. Xia, Damping pathways of mid-infrared plasmons in graphene nanostructures, Na-
A. Principi, M. Polini, and G. Vignale, Linear response
M. Cohen, R. Shavit, and Z. Zalevsky, Observing Optical
T. Langer, J. Baringhaus, H. Pfünir, H. W. Schumacher,
H. Duan, A. I. Fernandez-Dominguez, M. Bosman, S. A.
E. A. Pashitskii, Plasmon mechanism of high-
A. Woessner, M. B. Lundeberg, Y. Gao, A. Princi
G. X. Ni, A. S. McLeod, Z. Sun, L. Wang, L. Xiong, K.
A. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K.
SPATIAL SPECKLE PATTERNS IN NEAR-FIELD

A. K. W. Yang, Nanoplasmonics: Classical
H. L. Koppens, Highly confined low-
S. L. Tomarken, Y. Cao, A. Demir, K. Watanabe, T.
As argued in the main text, we can estimate the

This expression, which is a convolution of two functions, will generate a product under Fourier transform.

For purposes of Fig[1]1 and Fig[2]2-d we evaluate the above convolution numerically by using the convolution theorem. That is, we first perform a fast Fourier transform of both terms individually, multiply them together and then carry out an inverse transform. The ring-like distribution of spectral weight in $k$ space, shown in the insets of the Fig[1]1 and Fig[2]2-d, is the intermediate step of this process. An analytic expression for this quantity can be determined by evaluating the Fourier transform of the signal $S(r)$

$$S_k = \int d^2r e^{-ik \cdot r} S(r). \quad (S.4)$$

As expected by the convolution theorem the expression factorizes into two separate integrals

$$S_k = \int d^2r' \eta(r') e^{-ik \cdot r'} \int d^2r \frac{e^{-ik \cdot r + i2q_0|r|} e^{-2\delta|x|}}{2\pi|\mathbf{r} - \mathbf{r}'|}, \quad (S.5)$$

where the first integral is nothing but the Fourier harmonic of $\eta(x)$ and in the second integral we performed a variable change $\mathbf{r} - \mathbf{r}' \rightarrow \mathbf{r}$. To evaluate the integral over $d^2\mathbf{r}$ we first integrate over $|\mathbf{r}|$ and then perform angular integration using the identity

$$\int_0^{2\pi} d\theta \frac{1}{a + b \cos \theta} = \frac{2\pi}{\sqrt{a^2 - b^2}}. \quad (S.6)$$

After plugging $a = k$, $b = 2k_0 + 2\delta i$ this gives [4] of the main text.
Here we estimate $\lambda_1$ and $\lambda_2$ coefficients discussed in the main text for the toy model. Let us start with the $\lambda_2$ coefficient corresponding to the intraband transitions. In the small $q$ the interband coherence factor from \cite{ citesomeotherref } is non-zero only in proximity of the $K$ and $K'$ points. Near these points a linear dispersion $E_{s,k} = sv_F k$ is a good approximation of the actual bandstructure. In that limit, the small $q$ interband coherence band factor from \cite{ citesomeotherref } becomes

$$ F_{k+q,k}^{s,s'} \approx -\frac{1}{4} (q \cdot \nabla_k \varphi_k)^2 \approx \frac{q^2}{4 k^2} \sin^2 \theta, \quad (S.7) $$

where $\theta$ is the angle between $k$ and $q$. Consequently the $\Pi_2(\omega, q)$ is given by:

$$ \Pi_2(\omega, q) = -\frac{8q^2}{\omega^2} \sum_{k,s} f_{s,k} sv_F k \frac{\sin^2 \theta}{k^2}, \quad (S.8) $$

In the above we used the linear dispersion approximation $E_{s,k} = sv_F k$ for the whole band and accounted for the $K$ and $K'$ points through an additional factor of 2. This gives

$$ \Pi_2(\omega, q) \approx -\frac{2E_F}{\pi} \frac{q^2}{\omega^2} + \frac{2W}{\pi} \frac{q^2}{\omega^2} = \frac{2}{\pi} (W - E_F) \frac{q^2}{\omega^2}, \quad (S.9) $$

with the first term originating from the conduction band and the second term from the valence band contribution. This results in

$$ \lambda_2 = 2(W - E_F)/\pi. \quad (S.10) $$

We now proceed to estimate the $\lambda_1$. Let us place the Fermi energy in the conduction band, which immediately implies that the intraband contribution to the polarization function is non-vanishing only in the conduction band. To see this, let us go back to the \cite{ citesomeotherref }, which for $s = s' = -1$ and small $q$ is

$$ \Pi_1(\omega, q) \approx -\frac{8}{\omega^2} \sum_k f_{-1,k} (E_{-1,k} - E_{-1,k+q}) \quad (S.11) $$

$$ = -\frac{8}{\omega^2} \sum_k (E_{-1,k} - E_{-1,k+q}) = 0 \quad (S.12) $$

since $f_{-1,k} = 1$ for all $k$ in the valence band. It is therefore sufficient to just focus on the partially filled band. To be consistent with the $\lambda_2$ estimation we replace the dispersion energy as $E_{1,k} = v_F k$. The intraband contribution to the polarization function $\Pi_1(\omega, q)$ is then

$$ \Pi_1(\omega, q) \approx \frac{8q^2}{\omega^2} \sum_k f_{1,k} v_F \frac{\sin^2 \theta}{k} = \frac{2}{\pi} E_F \frac{q^2}{\omega^2}, \quad (S.13) $$

giving

$$ \lambda_1 = 2E_F/\pi. \quad (S.14) $$

As argued in the paragraph with \cite{ citesomeotherref }, the above result is exactly the same in the $\omega < 2E_F$ regime and, therefore,

$$ \beta_0 = 4\alpha v_F E_F. \quad (S.15) $$

Going back to the $\omega > 2E_F$ limit and using $\lambda_1, \lambda_2$ derived above gives a plasmon dispersion $\omega_p = \sqrt{\beta_0 q}$, which in the $\omega < 2W$, $\omega_p = 2\sqrt{\alpha v_F W q}$ by a factor $\sqrt{W/E_F}$. We thus see that at small $\omega < 2E_F$ the plasmon dispersion goes as $\omega_p = 2\sqrt{\alpha v_F E_F q}$ and hardens at high energies $\omega > 2W$, $\omega_p = 2\sqrt{\alpha v_F W q}$ by a factor $\sqrt{W/E_F}$. The last energy region missing from the above analysis of the plasmon dispersion is that of the interband p-h continuum, $2E_F < \omega < 2W$. Working again in the small $q$ limit, we find the $\lambda$s pointed out earlier, only the interband contribution to the polarization function $\Pi_2(\omega, q)$ develops an imaginary part, whilst the intraband polarization function $\Pi_1(\omega, q)$ is simply given by the \cite{ citesomeotherref }. To determine the form of $\Pi_2(\omega, q)$ in the interband p-h continuum energy range, let us approximate the coherence factor as in \cite{ citesomeotherref } giving

$$ \Pi_2(\omega, q) \approx \frac{8}{\omega^2} \sum_k f_{s,k} sv_F k \frac{\sin^2 \theta}{k^2}, \quad (S.17) $$

where we used the linear approximation to the energy dispersion $E_{s,k} = sv_F k$ for the whole band and again accounted for the $K$ and $K'$ points with an additional factor of 2. After carrying out the integrations we arrive at:

$$ \Pi_2(\omega, q) \approx -i \frac{2q^2}{\omega} \Theta(\omega - 2E_F) \Theta(\omega - 2W) \quad (S.18) $$

Here $\Theta \xi$ is the Heaviside function guaranteeing that the imaginary part is non-zero only in the $2E_F < \omega < 2W$ region. The dielectric function in the interband continuum region is therefore

$$ \varepsilon(\omega, q) = 1 - \beta_0 \frac{q}{\omega^2} + i \frac{\beta_0 \pi q}{E_F} \frac{q}{\omega}, \quad (S.19) $$

which shows that the collective mode $\omega_p$ in the $2E_F < \omega_p < 2W$ region

$$ \omega_p \approx \sqrt{\beta_0 q} - i \frac{\pi \beta_0}{2E_F} q \quad (S.20) $$

is a $\sqrt{\beta_0 q}$ dispersing plasmon, which gets rapidly damped by pair production. We finish the discussion of the collective modes by comparing the analytically predicted dispersion with the numerical result seen in Fig \cite{ citesomeotherfig }. While the simulated dispersion closely follows the square-root dependence $\omega_p \propto \sqrt{q}$,
the agreement between the simulation and $\omega_p = \sqrt{\beta q} q$ dispersion becomes nearly perfect if the two first terms $\beta_0$ and $\beta_1$ from the series expansion in (10) are used for a fitting. Although the terms $\beta_0$ and $\beta_1$ could in principle be computed by carrying out an expansion of the polarization function in (10) in powers of $q$ and then evaluating the resulting integrals numerically, we instead treat $\beta$ and $\beta_1$ as free parameters and fit them to the simulated dispersion. This approach gives

$$\beta_0 = 0.96 \times 10^3 \text{ meV}^2 \text{ nm},$$

$$\beta_1 = -10^3 \text{ meV}^2 \text{ nm}^2.$$  \hfill (S.21)

The best-fit $\beta_0$ value is close to $\beta_0 = 4\alpha v_F E_F \approx 0.86 \times 10^3 \text{ meV}^2 \text{ nm}$ predicted from (S.15). We also see that, since $\beta_1$ is negative, the plasmon dispersion is indeed softened by interband polarization, in agreement with the argument given in the main text [see (14)].

**TWISTED BILAYER GRAPHENE - DETAILS OF THE MODEL**

For description of the TBG’s bandstructure and eigenstates we employ the effective continuum Hamiltonian introduced in Ref.[31], together with its associated notation and numerical values of simulation parameters. The effective continuum Hamiltonian for one valley $\xi = -1, 1$ takes the form

$$H^{(\xi)} = \begin{pmatrix} H_1 & U^1 \\ U & H_2 \end{pmatrix}$$  \hfill (S.22)

in the basis of $(A_1, B_1, A_2, B_2)$ sites and is spin $\sigma$ degenerate. The matrices $H_l$ $(l = 1, 2)$ correspond to the intralayer Hamiltonian of the layer $l$ and are explicitly given as

$$H_l = -\frac{\hbar v}{a} \begin{pmatrix} 0 & e^{-i\xi q/2} + \frac{4\pi}{3} \\ e^{i\xi q/2} + \frac{4\pi}{3} & 0 \end{pmatrix},$$  \hfill (S.23)

where $k_{\pm} = k_x \pm ik_y$ and $k_x, k_y$ are lattice momenta. The BZ is defined as in the inset of the Fig.3a, with the two reciprocal lattice vectors being

$$G_1^M = -\frac{2\pi}{\sqrt{3} L_M} \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}, \quad G_2^M = \frac{4\pi}{\sqrt{3} L_M} \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$  \hfill (S.24)

with the moiré real space lattice constant $L_M = a/2 \sin(\theta/2)$. Matrix $U$ is the effective interlayer coupling given by:

$$U = \begin{pmatrix} u & u' \\ u' & u \end{pmatrix} + \begin{pmatrix} u & u' \nu - \xi \\ u' \nu - \xi & u \end{pmatrix} e^{i\xi G_1^M \cdot r} + \begin{pmatrix} u & u' \nu - \xi \\ u' \nu - \xi & u \end{pmatrix} e^{i\xi (G_1^M + G_2^M) \cdot r},$$  \hfill (S.25)

where $\nu = e^{i2\pi/3}$. We take the energy scale as $\hbar v/a = 2.1354 \text{ eV}$ and the lattice constant $a = 0.246 \text{ nm}$. The interlayer coupling terms $u$ and $u'$ are taken as $u = 0.0797 \text{ eV}$ and $u' = 0.0975 \text{ eV}$. The Bloch wavefunction for a valley $\xi$ is taken as

$$\Psi_{\xi,n,k}^X(r) = \sum_G C_{\xi,n,k}^X(G) e^{i(k+G) \cdot r}$$  \hfill (S.26)

with $X$ corresponding to each of the spinor components $X = A_1, B_1, A_2, B_2$. The band index is labelled by $n$ and $k$ is the Bloch wave vector in the Moiré Brillouin zone. Here $G$ runs over all possible integer combinations of the reciprocal lattice vectors, $G = m_1 G_1^M + m_2 G_2^M$ with integer $m_1$ and $m_2$. In practice, it is sufficient that for a given $k$ the integers $m_1$ and $m_2$ in total cover only around $\sim 60$ different combinations of the $G_1, G_2$ reciprocal lattice wave vectors. This simplification is due to the cutoff procedure utilized in Ref.[31].