Collective magnetic excitations in AA- and AB-stacked graphene bilayers

M. Sánchez Sánchez [1,2] G. Gómez-Santos,[2] and T. Stauber[1,3,*]

[1]*Materials Science Factory, Instituto de Ciencia de Materiales de Madrid, CSIC, E-28049 Madrid, Spain*

[2]*Departamento de Física de la Materia Condensada, Instituto Nicolás Cabrera and Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, E-28049 Madrid, Spain*

[3]*Theoretical Physics III, Center for Electronic Correlations and Magnetism, Institute of Physics, University of Augsburg, D-86135 Augsburg, Germany*

(Received 23 July 2021; revised 1 December 2021; accepted 2 December 2021; published 10 December 2021)

We discuss novel transverse plasmon polaritons that are hosted by AA- and AB-stacked bilayer graphene due to perfect nesting. They are composed of oscillating counterflow currents in between the layers, giving a clear interpretation of these collective modes as magnetic excitations carrying magnetic moments parallel to the planes. For AA-stacked bilayer graphene, these modes can reach zero frequency at the neutrality point and we thus predict a symmetry-broken ground state leading to in-plane orbital ferromagnetism. Even though it could be hard to detect them in real solid-state devices, these novel magnetic plasmons should be observable in artificial setups such as optical lattices. Also, our results might be relevant for magic angle twisted bilayer graphene samples, as their electronic properties are mostly determined by confined AA-stacked regions.

DOI: 10.1103/PhysRevB.104.245412

I. INTRODUCTION

Plasmonics in graphene-based materials has attracted much attention since the wavelength of light can be reduced by several orders of magnitude governed by the inverse fine-structure constant $\alpha^{-1} \sim 137$ [1–7]. This strong confinement further gives rise to various applications such as perfect absorption [8], the sensoring of biochemical molecules [9], plasmon-induced transparency [10], and plasmon-enhanced chirality [11]. Also, other two-dimensional systems have shown interesting plasmonic properties such as molybdenum disulfide [12], black phosphorous [13], and general van der Waals materials [14].

Plasmonics is usually based on longitudinal or transverse-magnetic plasmons which consist of collective density oscillations collinear to the propagation direction. These excitations are strongly confined due to the enhanced Coulomb coupling between the electromagnetic field and the charge carrier density. On the other hand, transverse or transverse-electric plasmons can also exist in graphene and consist of collective current oscillations that are perpendicular to the propagation direction [15]. These excitations are closely pinned to the light cone and thus weakly confined, which allows for their detection in the Otto configuration [16,17]. Negative refractive index environments can help to enhance their confinement [18,19], and magnetically biased graphene-ferrite structures give rise to nonreciprocal plasmons [20]. Also, in AB-stacked bilayer graphene the transverse modes are considerably more confined than in its monolayer counterpart [21]. The nonlinear response of transverse plasmons is analyzed in Ref. [22].

Transverse plasmons in single-layer graphene arise from interband transitions close to the absorption threshold and it is difficult to visualize their magnetic character [23–25]. This might further be a reason for their fragile nature, i.e., their ultrahigh refractive index sensitivity [26].

Here, we study the antisymmetric transverse mode of AA- or AB-stacked graphene bilayers, which has not been mentioned in the literature so far [27]. They are considerably more stable, as they arise from a resonant "magnetic" absorption, absent in the uncoupled bilayers discussed in Ref. [28]. Further, the antisymmetric counterflow leads to perfect screening

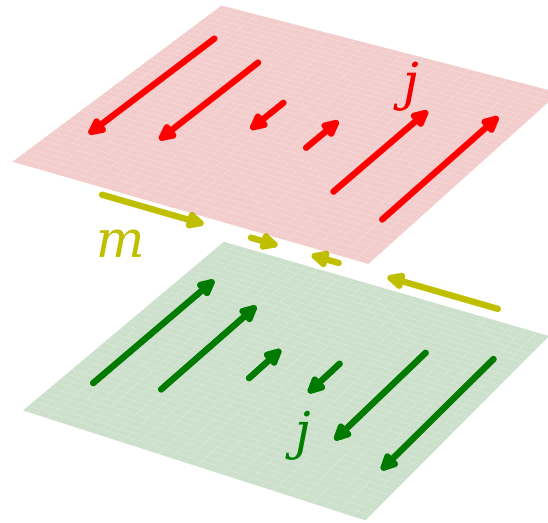

FIG. 1. Sketch of the transverse antisymmetric plasmons in bilayer graphene. The loop currents, j, create an in-plane magnetic moment, m, between the layers. In particular, we are interested in the long-wavelength limit $q \to 0$, where a uniform, in-plane, oscillating magnetic moment emerges out of transverse (chargeless) currents which are opposite in each layer.

*tobias.stauber@csic.es

outside the sample for $qa \to 0$ such that the plasmon dispersion does not depend on the two external, possibly different dielectric environments, contributing to additional stability. Finally, these antisymmetric modes have a clear physical interpretation and correspond to loop currents around the layers resulting in an in-plane magnetic moment density (see Fig. 1). Remarkably, for AA-stacked graphene these modes can also exist at zero frequency and we thus predict a magnetic instability and spontaneous symmetry breaking at zero temperature, giving rise to orbital in-plane ferromagnetism.

The paper is organized as follows: in Sec. II, we discuss the electromagnetic response in a bilayer and define the plasmonic excitations within the random-phase approximation (RPA). We also comment on the effect of an inhomogeneous dielectric environment for the transverse (magnetic) modes. In Sec. III, we discuss the magnetic response of two typical graphene bilayers that are AA and AB stacked and give numerical estimates for the resonant frequency. We finally discuss the stability in AA-stacked bilayer graphene in Sec. IV and close with conclusions.

II. ELECTROMAGNETIC RESPONSE IN BILAYERS

Following Ref. [29], the most general (bare) in-plane current response in isotropic (and nonchiral) bilayer systems can be written as

$$\begin{pmatrix} j_\nu^1 \\ j_\nu^2 \end{pmatrix} = -\begin{pmatrix} \chi_{11} & \chi_{12} \\ \chi_{21} & \chi_{22} \end{pmatrix}\begin{pmatrix} A_\nu^1 \\ A_\nu^2 \end{pmatrix}, \tag{1}$$

where A_ν^ℓ denotes the gauge field acting on layer $\ell = 1, 2$ and $\nu = l, t$ defines the longitudinal and transverse channels, which are decoupled. Further, $\chi_{\ell,\ell'} = \langle\langle j_\nu^\ell, j_\nu^{\ell'}\rangle\rangle$ is the current-current response with respect to layers ℓ and ℓ'. Note that the polarization subscript ν is suppressed in $\chi_{n,m}$ since both polarizations coincide in the $q \to 0$ limit.

In a homogeneous dielectric environment, the propagators of the gauge fields with in-plane momentum q and frequency ω are given by

$$\mathcal{D}_\nu = d_\nu \begin{pmatrix} 1 & e^{-q'a} \\ e^{-q'a} & 1 \end{pmatrix}, \tag{2}$$

with a the distance between layers, $q' = \sqrt{q^2 - \epsilon\mu\omega^2/c^2}$, $d_l = \frac{q'}{2\epsilon_0\omega^2}$, and $d_t = -\frac{1}{2\epsilon_0 c^2 q'}$ [30]. Time-reversal symmetry sets $\chi_{12} = \chi_{21}$, and in the following, we further assume that the exchange of layers is also a symmetry leading to $\chi_{11} = \chi_{22}$. The collective excitations (plasmons) are then defined in the RPA by [31]

$$\det(\mathbf{1} - \chi\mathcal{D}_\nu) = 0\,, \tag{3}$$

where $\chi = \chi_{\ell,\ell'}$ denotes the 2×2 response matrix.

In terms of bonding and antibonding modes, we have $\chi_\pm = \chi_{11} \pm \chi_{12}$ and $d_{\nu\pm} = d_\nu(1 \pm \exp(-q'a))$, and plasmons are the zeros of the effective dielectric functions:

$$1 - d_{\nu\pm}\chi_\pm = 0. \tag{4}$$

The transverse field propagator is purely imaginary in the $qa \to 0$ limit, causing any current excitation leak out to the propagating electromagnetic field for a single layer. This is avoided in the double layer by the combination $1 - \exp(-2q'a)$ in d_{t-}, making the effective propagator for the antisymmetric mode real to lowest order in q. Leaking to the electromagnetic field thus only occurs in higher orders. Of course, this effective real magnetic coupling is still very small, but the condition for the self-sustained collective modes is guaranteed by the divergent nature of the current response χ_- near sharp features of the spectrum as we show below.

The response function χ_- refers to the counterflow that leads to an in-plane magnetic moment. The intrinsic excitations are thus collective magnetic dipole oscillations or magnetic plasmons—contrary to conventional plasmons, which are collective electric dipole oscillations. Nevertheless, the resonance condition, (4), is usually not fulfilled due to the weak magnetic coupling and a true magnetic plasmon cannot be formed.

In the next section we show that in typical graphene bilayers with AA and AB stacking, the resonance condition can always be reached due to "perfect nesting" close to a frequency ω_0, where the real part of χ_- diverges as $(\omega - \omega_0)^{-1}$. For extremely clean samples, collective modes are thus guaranteed by the divergent nature of the current response χ_- near sharp features of the spectrum and we predict genuine magnetic excitations close to the resonance condition.

Interestingly, in an inhomogeneous environment, i.e., if the bilayer is surrounded by two dielectrics with permittivities ϵ_1 and ϵ_2 and permeabilities μ_1 and μ_2, the final result for the antisymmetric modes does not change in the long-wavelength limit due to perfect screening. The antisymmetric eigenvalue of $\mathbf{1} - \chi\mathcal{D}_\mathbf{t}$ then reads (see Appendix B)

$$\lambda^- = 1 + \frac{\mu_0 a}{2}(\chi_{11} - \chi_{12})(1 + O(q'a))\,. \tag{5}$$

The magnetic mode, alternatively defined by $\mathrm{Re}\lambda^- = 0$, is therefore essentially independent of the surrounding dielectric media. On the other hand, the symmetric plasmonic mode does depend on the environment, and the existence of excitations closely pinned to the light cone will strongly depend on the refractive index difference, $n_1 - n_2$ [26].

III. GRAPHENE BILAYERS AND CURRENT RESPONSE

The calculation of χ_{11} and χ_{12} only requires the knowledge of the underlying Hamiltonian. We consider the general low-energy Hamiltonian of graphene bilayers [32–34],

$$H = \sum_{\boldsymbol{k}} H_{\boldsymbol{k}}, \quad \text{with} \quad H_{\boldsymbol{k}} = \begin{pmatrix} H_{\boldsymbol{k}}^0 & U^\dagger \\ U & H_{\boldsymbol{k}}^0 \end{pmatrix}, \tag{6}$$

where $H_{\boldsymbol{k}}^0 = \hbar v_F \boldsymbol{\sigma}\cdot\boldsymbol{k}$ is the single-layer-graphene Dirac Hamiltonian acting on states of Bloch momentum $\boldsymbol{k}$. The 2×2 matrix U denotes the interlayer coupling, characterized by the interlayer hopping matrix element $t \sim 0.33$ eV.

We are interested in the generalized current-current response given by the Kubo formula [31],

$$\chi_{\ell,\ell'}^{i,i'} = \frac{e^2 g_s g_v}{A}\sum_{n,m}\frac{f_n - f_m}{\hbar\omega - (\epsilon_m - \epsilon_n) + i\eta} v_{n;m}^{i,\ell} v_{m;n}^{i',\ell'}\,, \tag{7}$$

where $i, i' = x, y$ denote the spatial direction, $\ell, \ell' = 1, 2$ the different layers, and n, m the set of quantum numbers; $g_s = 2$ and $g_v = 2$ are the spin and valley degeneracy, respectively.

The charge of the electron is e, f_n is the Fermi-Dirac distribution, and $v_{n;m} = \langle n|\hat{v}|m\rangle$ are the matrix elements of the general velocity operator $\hat{v}$, which in the local approximation ($q = 0$) couples states with the same momentum, $\boldsymbol{k}$.

The "magnetic" velocity can be defined by the counterflow between the two layers and the magnetic moment is thus proportional to this velocity and the area (distance) between the two layers. For the external gauge field in the i direction, the magnetic response is thus defined by $\chi_- = \chi^{ii}_{11} - \chi^{ii}_{12}$, but we suppress the superindex in the following.

The magnetic susceptibility to an in-plane magnetic field can be obtained from the magnetic response. If $m_\parallel$ and $B_\parallel$ are the moment and field, then writing $m_\parallel = \chi_\parallel B_\parallel$, one has $2\chi_\parallel = -a^2\chi_-$. Note that the emergence of a magnetic mode requires paramagnetism for which $\chi_\parallel$ becomes positive.

A. AA-stacked bilayer graphene

With $U = t\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$, we describe AA-stacked graphene, whose optical conductivity has been discussed in Ref. [35] and its plasmonic properties in Ref. [36]. The eigenstates are the bonding and antibonding states of single-layer graphene. The eigenenergies are given by $E_{k,n_1,n_2} = n_1 t + n_2\hbar v_F k$ and the eigenvectors by $\psi_{\boldsymbol{k}}^{n_1,n_2} = \frac{1}{2}(n_1, n_1 n_2 e^{i\varphi_{\boldsymbol{k}}}, 1, n_2 e^{i\varphi_{\boldsymbol{k}}})^T$, where $n_1, n_2 = \pm$ and $\varphi_{\boldsymbol{k}}$ denotes the angle that $\boldsymbol{k}$ forms with the x axis. The velocity operator of layer ℓ is given by $\hat{\mathbf{v}}_\ell = \boldsymbol{\sigma} \otimes \mathbf{1}_\ell$, where $\mathbf{1}_\ell$ performs a projection onto states of sublayer ℓ and $\boldsymbol{\sigma}$ are the Pauli matrices of the pseudospin variables.

The "electric"/"magnetic" excitations couple to the total/counterflow current $\hat{\mathbf{j}}_\pm = -e\hat{\mathbf{v}}_\pm = -e(\hat{\mathbf{v}}_1 \pm \hat{\mathbf{v}}_2)$ and the relevant matrix element is

$$v^{\pm}_{\boldsymbol{k},n_1,n_2;\boldsymbol{k},m_1,m_2} = (n_1 m_1 \pm 1)\frac{v_F}{4}\left(m_2 e^{i\varphi_{\boldsymbol{k}}} + n_2 e^{-i\varphi_{\boldsymbol{k}}}\right). \quad (8)$$

This means that only transitions with $n_1 m_1 = -1$ yield a finite contribution to the imaginary part of the magnetic response. Interestingly, for $\hbar\omega = 2t$ there is a perfect matching that gives rise to a delta function and that is only present in the "magnetic absorption" (see Fig. 5, Appendix A). The resonant contribution of the magnetic response $2\chi_- = \langle\langle j^x_- j^x_-\rangle\rangle$ at $T = 0$ reads

$$\mathrm{Im}\chi^{\mathrm{nested}}_-(\omega) = -\frac{e^2 g_s g_v}{16\hbar^2}\delta(\hbar\omega - 2t) \times [2(\mu^2 + t^2)\theta(t - |\mu|) + 4t|\mu|\theta(|\mu| - t)]. \quad (9)$$

The real part is given by the Kramers-Kronig relation

$$\mathrm{Re}\chi_-(\omega) = \frac{1}{\pi}\int_0^\infty d\omega' \mathrm{Im}\chi_-(\omega')\frac{2\omega'}{(\omega')^2 - \omega^2}, \quad (10)$$

and the main contribution from the perfect nesting condition is thus given by

$$\mathrm{Re}\chi^{\mathrm{nested}}_- = -\frac{1}{\pi}\frac{e^2 g_s g_v}{8\hbar^2}\frac{2t}{(2t)^2 - (\hbar\omega)^2} \times \left[2(\mu^2 + t^2)\theta(t - |\mu|) + 4t|\mu|\theta(|\mu| - t)\right]. \quad (11)$$

In Fig. 2, we show the real and imaginary parts of the full magnetic response as discussed in Appendix A for $\mu = 0$ and $\mu = t$. The delta function of the imaginary part of the analytical solution (full line) is broadened by the same value ($\eta = 0.1t$) as used in the numerical solution (small circles). The singularity in the real part of the analytical solution, however, is plotted for the clean system which is algebraic for both chemical potentials at $\hbar\omega = 2t$. For $\mu = 0$, there is a logarithmic singularity at $\omega = 0$ which is responsible for the symmetry-broken ground state.

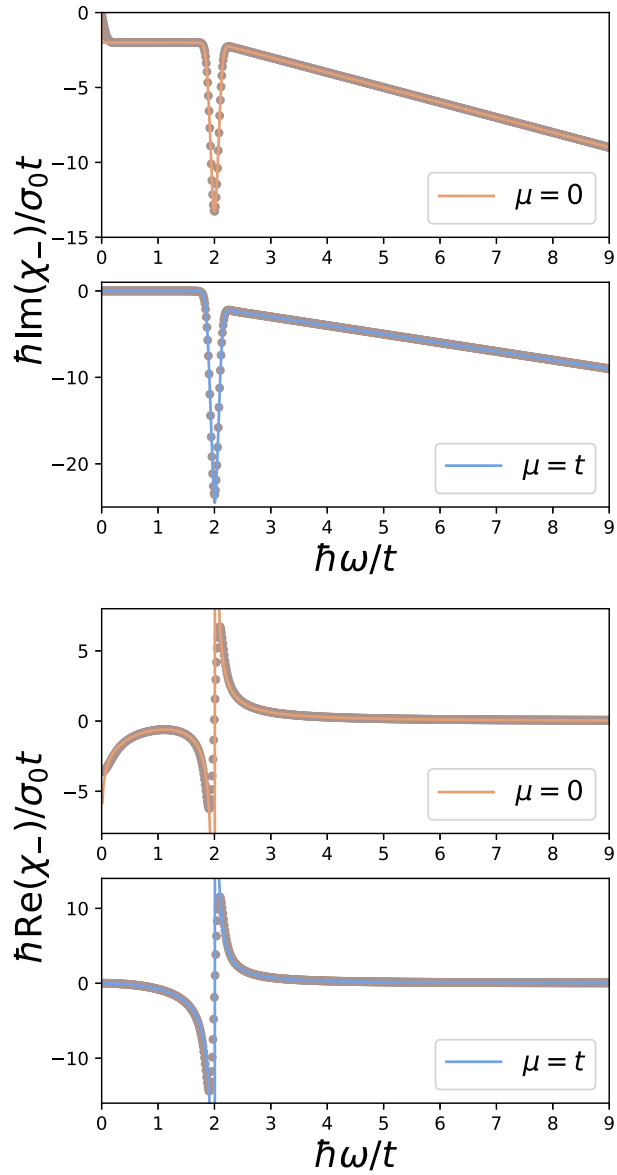

FIG. 2. Magnetic response function for AA-stacked bilayer graphene as a function of the frequency, for $\mu = 0$ and $\mu = t$. Top: Imaginary part of the magnetic response, $\mathrm{Im}\chi_-$, in units of $\sigma_0 t/\hbar$, where $\sigma_0 = e^2/4\hbar$ is the universal conductivity of graphene. Bottom: Real part of the magnetic response function, $\mathrm{Re}\chi_-$, in units of $\sigma_0 t/\hbar$. The small circles are the results of the numerical computations, and the solid lines are the formulas derived in Appendix A. The width of the delta function has been tuned to match the numerical results for which $\eta = 0.1t$ was used.

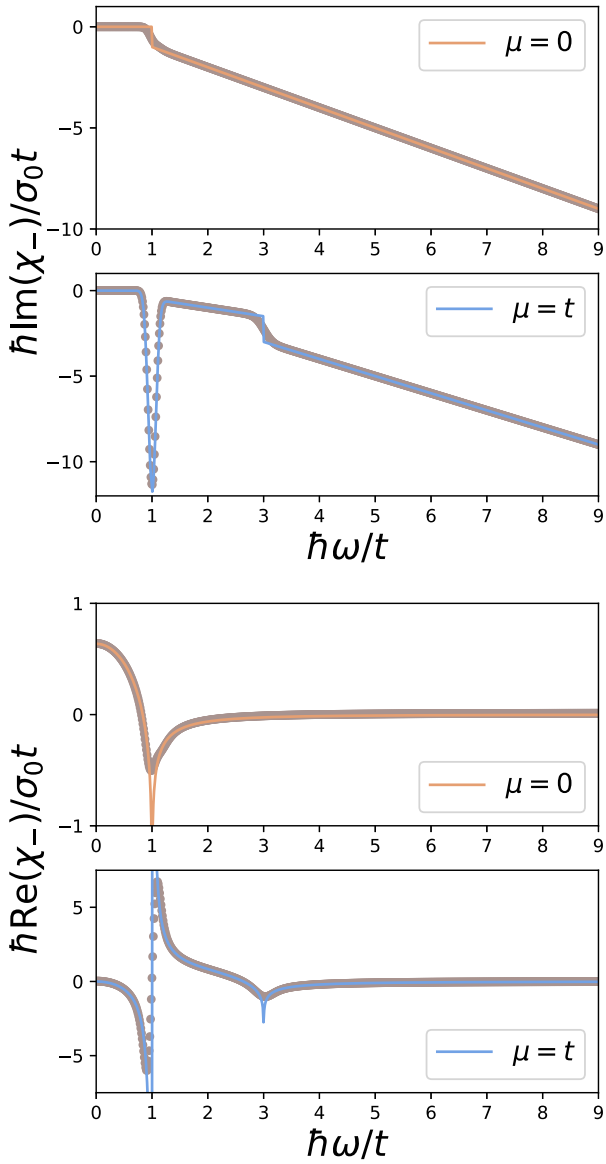

FIG. 3. Same as Fig. 2, but for AB-stacked bilayer graphene.

B. AB-stacked bilayer graphene

With $U = t\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, we describe AB-stacked graphene, whose optical conductivity has been discussed in Ref. [37] and its plasmonic properties in Ref. [38]. The eigenenergies are now given by $E_{k,n_1,n_2} = \frac{1}{2}(n_1 t + n_2 \xi_k)$ with $n_1, n_2 = \pm$ and $\xi_k = \sqrt{(2\hbar v_F k)^2 + t^2}$. The eigenvectors are

$$\psi_{\boldsymbol{k}}^{+,\pm} = N\left(\tfrac{t-\xi_k}{2\hbar v_F k}, \mp e^{i\varphi_{\boldsymbol{k}}}, e^{-i\varphi_{\boldsymbol{k}}}, \mp\tfrac{t-\xi_k}{2\hbar v_F k}\right)^T,$$

$$\psi_{\boldsymbol{k}}^{-,\pm} = N\left(1, \mp\tfrac{t-\xi_k}{2\hbar v_F k} e^{i\varphi_{\boldsymbol{k}}}, -\tfrac{t-\xi_k}{2\hbar v_F k} e^{-i\varphi_{\boldsymbol{k}}}, \pm 1\right)^T,$$

where $\varphi_{\boldsymbol{k}}$ denotes the angle that $\boldsymbol{k}$ forms with the x axis and $N = (2(1 + (\frac{t-\xi_k}{2\hbar v_F k})^2))^{-\frac{1}{2}}$. The velocity operator of layer ℓ is again given by $\hat{\mathbf{v}}_\ell = \boldsymbol{\sigma} \otimes \mathbf{1}_\ell$ and the matrix elements of the current counterflow read $v^-_{\boldsymbol{k}++,\boldsymbol{k}+-} = iv_F \sin(\varphi_{\boldsymbol{k}})$, $v^-_{\boldsymbol{k}++,\boldsymbol{k}-+} = -v_F \cos(\varphi_{\boldsymbol{k}})$, $v^-_{\boldsymbol{k}+-,\boldsymbol{k}--} = v_F \cos(\varphi_{\boldsymbol{k}})$, $v^-_{\boldsymbol{k}+-,\boldsymbol{k}--} = iv_F \sin(\varphi_{\boldsymbol{k}})$, and 0 otherwise. Again, there is a perfect nesting condition, this time at $\hbar\omega = t$. This will also give rise to collective magnetic excitations and the resonant contribution of the magnetic response reads

$$\chi_-^{\text{nested}}(\omega) = -\frac{e^2 g_s g_v}{32\pi\hbar^2}\left[\frac{2t}{t^2 - (\hbar\omega)^2} + i\pi\delta(\hbar\omega - t)\right] \times [2|\mu|(|\mu| + t)\theta(t - |\mu|) + 4|\mu| t\theta(|\mu| - t)]\,. \tag{12}$$

However, contrary to the AA-stacked bilayer, perfect nesting also occurs in the optical absorption of AB-stacked graphene and allows the existence of high-energy electronic or charged plasmons with frequency $\omega \approx t/\hbar$ [38].

In Fig. 3, we show the real and imaginary parts of the full magnetic response as discussed in Appendix A for $\mu = 0$ and $\mu = t$. The delta function of the imaginary part of the analytical solution (solid line) is broadened by the same value ($\eta = 0.1t$) as used in the numerical solution (small circles). The singularity in the real part of the analytical solution, however, is plotted for the clean system, and for $\mu = 0$ this singularity is logarithmic, whereas for $\mu = t$ it is algebraic.

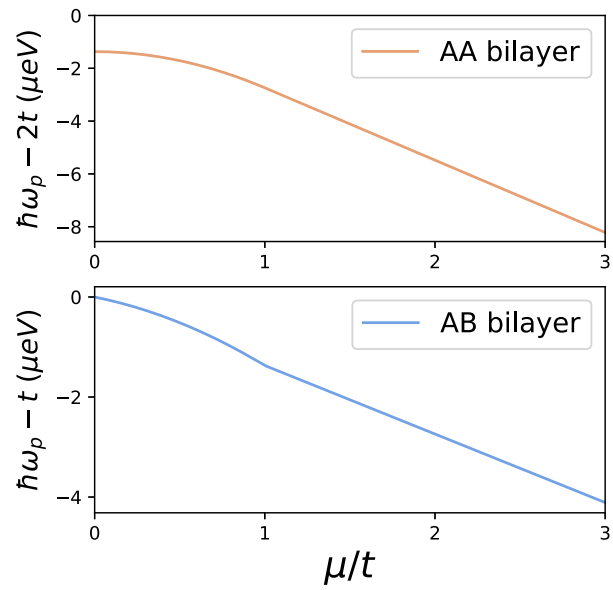

FIG. 4. Frequency of the magnetic plasmons ($\hbar\omega_p$) for AA and AB bilayer graphene as a function of the Fermi level. We set $t = 0.33$ eV. In both cases, we plot the difference between the plasmon frequencies and the nesting frequency.

C. Numerical estimates

As discussed above, the transverse antisymmetric photonic propagator is given by $d_{t-} = -(1 - e^{-q'a})/2\epsilon_0 c^2 q' \approx -a/2\epsilon_0 c^2$, and plasmonic excitations obey $1 - d_{t-}\chi_- = 0$. Therefore, only a negative real response, $\mathrm{Re}\chi_-(\omega) < 0$, allows for solutions. True dissipationless plasmons further demand $\mathrm{Im}\chi_-(\omega) = 0$, and a finite value of $\mathrm{Im}\chi_-(\omega)$ yields damped plasmons with plasmon frequency ω_p given by $1 - d_{t-}\mathrm{Re}\chi_-(\omega_p) = 0$, and inverse lifetime $\gamma = \mathrm{Im}\chi_-(\omega_p)/\frac{\partial}{\partial\omega}\mathrm{Re}\chi_-(\omega_p)$ [39].

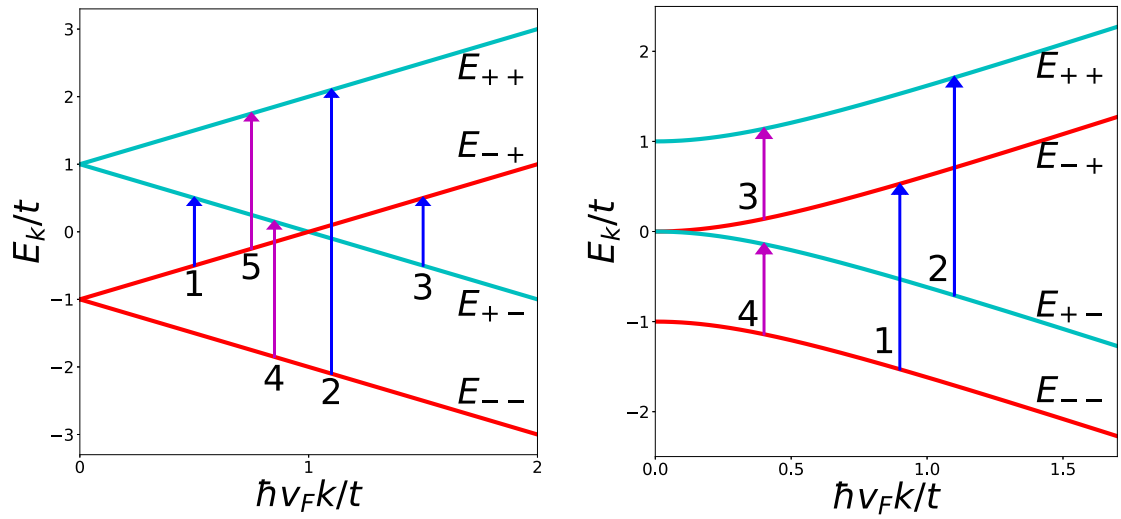

FIG. 5. Band structure of AA-stacked (left) and AB-stacked (right) bilayer graphene, and allowed transitions of the current counterflow operator. Bonding states ($n_1 = -1$) are depicted in red, and antibonding states ($n_1 = +1$) in blue. Note the perfect nesting of transitions 4 and 5 in AA-stacked and transitions 3 and 4 in AB-stacked graphene.

Setting the doping level to $\mu = t$, the real part of the magnetic current response near the nesting frequency reads

$$\mathrm{Re}\chi_{-}^{\mathrm{AA}}(\omega) = \frac{16\sigma_0 t}{\pi\hbar}\frac{1}{(\hbar\omega/t)^2 - 4}\,, \tag{13}$$

$$\mathrm{Re}\chi_{-}^{\mathrm{AB}}(\omega) = \frac{4\sigma_0 t}{\pi\hbar}\frac{1}{(\hbar\omega/t)^2 - 1}\,, \tag{14}$$

where $\sigma_0 = e^2/4\hbar$ is the universal conductivity of graphene [40].

With the fine-structure constant $\alpha = e^2/4\pi\epsilon_0\hbar c \sim 1/137$, $\hbar c = 1973$ eV Å, $a = 3.4$ Å, and $t = 0.33$ eV, we thus have the following plasmonic resonances:

$$\hbar\omega_p^{\mathrm{AA}} = \sqrt{1 - \alpha\frac{2at}{\hbar c}}\,2t = \sqrt{1 - 8.3\times10^{-6}}\,2t\,, \tag{15}$$

$$\hbar\omega_p^{\mathrm{AB}} = \sqrt{1 - \alpha\frac{2at}{\hbar c}}\,t = \sqrt{1 - 8.3\times10^{-6}}\,t\,. \tag{16}$$

The plasmon frequencies are thus separated from the resonant or nesting energy by 2.7 and 1.4 μeV for AA- and AB-stacked bilayers, respectively, which obviously calls for ultraclean samples. In Fig. 4, we show its dependence on the chemical potential.

Since the broadening of the delta function is of the order of the free-mean path, the pole is normally superposed by the dissipative delta function in the imaginary part. So, even though the predicted modes are not likely to be observed in actual bilayer graphene samples, these excitations might be observable in artificial systems such as cold atoms. For example, if we set the separation between the layers to be $a = 100$ nm, then we get $2t - \hbar\omega_p = 0.8$ meV (0.4 meV) for AA (AB) bilayers, which should be observable.

IV. MAGNETIC INSTABILITY IN NEUTRAL AA-STACKED BILAYER GRAPHENE

The antisymmetric response of AA-stacked bilayer graphene at the neutrality point $\mu = 0$ is logarithmically diverging at $\omega = 0$ [see Eq. (A11), Appendix A]. This hints at a mode condensation or Condon instability as discussed in Refs. [41,42]. The relevant equilibrium response function requires the order of limit to be $\omega \to 0$ and then $q \to 0$, opposite to the one calculated here. The difference between the two

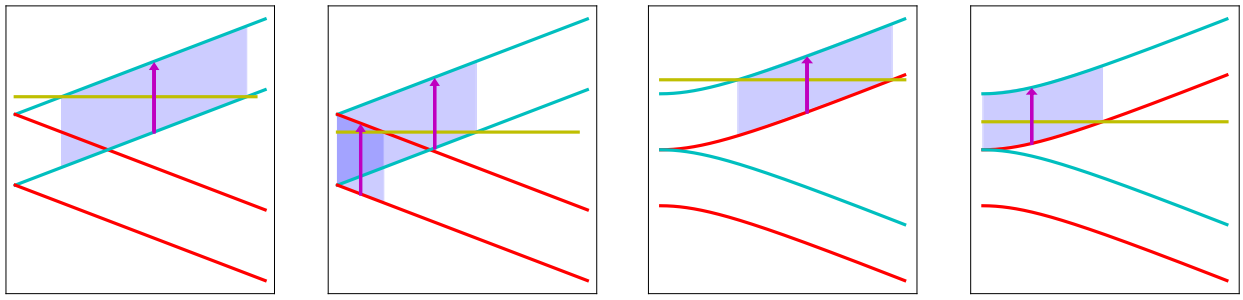

FIG. 6. Nested transitions that contribute to the resonance of the magnetic response in AA- and AB-stacked bilayer graphene, for $\mu > t$ and $\mu < t$. Note that at charge neutrality the weight of the resonance becomes zero for AB graphene, whereas it is always nonzero in the AA bilayer.

limit orders is a contact term related to the density of states, a finite magnitude that cannot compete with the divergence [43]. Therefore, we can use our previous results for the ground-state stability analysis.

Now we estimate the critical temperature. We start from the RPA instability $d_{t,-}\chi_- = 1$, which can be thought of as a Stoner criterion. A proper analysis would require the calculation of a finite-temperature response. On physical grounds, the latter can be inferred from the replacement $\hbar\omega \to k_B T$. For $\mu = 0$ and $\omega \to 0$, we have $\chi_- = t\frac{e^2}{\pi\hbar^2}\ln\frac{2t}{k_B T}$ (see Appendix A). We thus obtain for the critical temperature

$$k_B T_c = 2t \exp\left(-\frac{3}{4\alpha}\frac{t_0}{t}\frac{a_0}{a}\frac{c}{v_F}\right), \tag{17}$$

where $t_0 \sim 2.7$ eV is the in-plane hopping amplitude and $a_0 = 1.42$ Å the carbon-carbon distance.

The critical temperature is virtually zero, however, it may increase after renormalization of the single-particle parameters in twisted bilayer graphene, where the electron density is concentrated in the regions of local AA stacking. In this context, we mention that orbital ferromagnetism has recently been observed for filling fraction $n = 3/4$ of the first conduction band of twisted bilayer graphene [44].

Our observation also links with a recent publication stating that the divergent paramagnetic response in twisted bilayer graphene can lead to a permanent in-plane magnetic moment [45]. Note that this is not possible in one layer with a perpendicular magnetic moment [42]. Also, note that this instability is different from the one due to the short-ranged Hubbard interaction [46–48]. Finally, due to perfect screening, this mode cannot be enhanced by a superconducting cavity [41] (see Appendix C).

V. CONCLUSIONS

In this work, we have discussed the magnetic response of AA- and AB-stacked bilayer graphene. For both stacking forms, we find a resonance due to perfect nesting which gives rise to an algebraic divergence in the real part of the magnetic instability. Within the RPA, this gives rise to an oscillating in-plane magnetic moment that might be observable in ultraclean samples. In AA-stacked samples, the ground state is further given by a symmetry-broken state leading to in-plane orbital ferromagnetism. By identifying in the electronic structure of simple models the key features for appearance, as we do here, we thus offer physical guidance for its search in other systems.

The magnetic excitations/instabilities are composed of counterflow currents in each layer, giving rise to a magnetic moment parallel to the planes, thus providing a clear intuition of the magnetic nature, in contrast to the transverse excitations in a single layer. Also, they do not depend on the external dielectric environment including optical cavities which should help to stabilize them, thus paving the way to novel plasmonics that is not limited by charged impurities. Finally, these excitations might play a role in magic angle twisted bilayer samples, as their electronic properties are mostly determined by confined AA-stacked regions.

ACKNOWLEDGMENTS

This work was supported by Spain's MICINN under Grants No. FIS2017-82260-P, No. PGC2018-096955-B-C42, No. PID2020-113164GB-I00, and No. CEX2018-000805-M as well as by the CSIC Research Platform on Quantum Technologies PTI-001. The access to computational resources of Centro de Supercomputación de Galicia (CESGA) is also gratefully acknowledged.

APPENDIX A: MAGNETIC RESPONSE IN AA- AND AB-STACKED BILAYER GRAPHENE

1. AA-stacked bilayer graphene

In Fig. 5, we represent and label the allowed transitions of the magnetic current operator for AA and AB bilayers. In the $T = 0$ limit, the Fermi distribution becomes the step function $\theta(\mu - \epsilon)$, and taking $\eta \to 0$ in Eq. (7) allows for an easy evaluation of the imaginary part of the response function. The nonresonant transitions yield the nonresonant or not nested part of the response function:

$$\begin{aligned}\mathrm{Im}\chi_-^{\text{not nested}}(\omega) = &-\pi\frac{e^2 g_s g_v}{2A}\frac{A}{(2\pi)^2}\left(\pi v_F^2\right)\int kdk\Bigg\{\delta(\hbar\omega - (2t - 2\hbar v_F k))\theta\left(\frac{t-|\mu|}{\hbar v_F} - k\right)\\ &+\delta(\hbar\omega - (2t + 2\hbar v_F k))\theta\left(k - \frac{|\mu| - t}{\hbar v_F}\right) + \delta(\hbar\omega - (-2t + 2\hbar v_F k))\theta\left(k - \frac{|\mu| + t}{\hbar v_F}\right)\Bigg\}.\end{aligned} \tag{A1}$$

The prefactor πv_F^2 comes from the angular integral of $\frac{v_F^2}{4}(2 + 2\cos(2\phi_k))$. Above, the first, second, and third terms are the contributions of transitions 1, 2, and 3, respectively. Performing the integrals gives

$$\begin{aligned}\mathrm{Im}\chi_-^{\text{not nested}}(\omega) = &-\frac{e^2 v_F^2}{2}\Bigg\{\frac{2t - \hbar\omega}{(2\hbar v_F)^2}\theta(\hbar\omega - 2|\mu|)\theta(2t - \hbar\omega) + \frac{\hbar\omega - 2t}{(2\hbar v_F)^2}\theta(\hbar\omega - 2|\mu|)\theta(\hbar\omega - 2t)\\ &+\frac{2t + \hbar\omega}{(2\hbar v_F)^2}\theta(\hbar\omega - 2|\mu|)\theta(2t + \hbar\omega)\Bigg\}.\end{aligned} \tag{A2}$$

For the resonant transitions, the imaginary part becomes

$$\mathrm{Im}\chi_-^{\mathrm{nested}}(\omega) = -\pi\frac{e^2 g_s g_v}{2A}\frac{A}{(2\pi)^2}(\pi v_F^2)\int kdk\Big\{\delta(\hbar\omega-2t)\theta\Big(k-\frac{-(t+\mu)}{\hbar v_F}\Big)\theta\Big(\frac{t-\mu}{\hbar v_F}-k\Big) + \delta(\hbar\omega-2t)\theta\Big(k-\frac{\mu-t}{\hbar v_F}\Big)\theta\Big(\frac{\mu+t}{\hbar v_F}-k\Big)\Big\}. \tag{A3}$$

The first and second terms are the contribution of transitions 4 and 5, respectively. Equivalently,

$$\mathrm{Im}\chi_-^{\mathrm{nested}}(\omega) = -\frac{e^2 g_s g_v}{8}\frac{v_F^2}{(2\hbar v_F)^2}\delta(\hbar\omega-2t)\{4(\mu^2+t^2)\theta(t-|\mu|)+8|\mu|t\theta(|\mu|-t)\}. \tag{A4}$$

The total response is the sum of the nonresonant and the resonant parts:

$$\mathrm{Im}\chi_-(\omega) = -\frac{e^2 g_s g_v}{16\hbar^2}\{[\hbar\omega\theta(\hbar\omega-2t)+2t\theta(2t-\hbar\omega)]\theta(\hbar\omega-2|\mu|)+[2(\mu^2+t^2)\theta(t-|\mu|)+4t|\mu|\theta(|\mu|-t)]\delta(\hbar\omega-2t)\}. \tag{A5}$$

In terms of the dimensionless energy, $x=\frac{\hbar\omega}{t}$, and the universal conductivity of graphene, $\sigma_0=\frac{e^2}{4\hbar}$ [40], yields with $g_s = g_v = 2$

$$\mathrm{Im}\chi_-(x) = -\frac{\sigma_0 t}{\hbar}\left\{x\left[\theta(x-2)+\frac{2}{x}\theta(2-x)\right]\theta(x-2|\mu|/t)+\left[2\left(\frac{\mu^2}{t^2}+1\right)\theta(1-|\mu|/t)+\frac{4|\mu|}{t}\theta(|\mu|/t-1)\right]\delta(x-2)\right\}. \tag{A6}$$

For the real part we use the Kramers-Kronig relation

$$\mathrm{Re}\chi_-(\omega) = \frac{2}{\pi}\mathcal{P}\int_0^\infty d\omega'\mathrm{Im}\chi_-(\omega')\frac{\omega'}{\omega'^2-\omega^2}, \tag{A7}$$

where $\mathcal{P}$ denotes the Cauchy principal value.

The delta function can be integrated directly, which gives the contribution of the resonant transitions to the real part presented in the text (see Fig. 6),

$$\mathrm{Re}\chi_-^{\mathrm{nested}}(\omega) = \frac{-e^2 g_s g_v}{8\pi\hbar^2}\frac{2t}{(2t)^2-(\hbar\omega)^2}\Big[2(\mu^2+t^2)\theta(t-|\mu|)+4t|\mu|\theta(|\mu|-t)\Big]. \tag{A8}$$

For the nonresonant transitions, we perform the integrals

$$\begin{aligned}\mathrm{Re}\chi_-^{\mathrm{not\ nested}}(\omega) =& \frac{-e^2 g_s g_v}{8\pi\hbar^2}\mathcal{P}\int_{2|\mu|/\hbar}^\infty d\omega'\frac{\omega'}{\omega'^2-\omega^2}\Big[\hbar\omega'\theta(\hbar\omega'-2t)+2t\theta(2t-\hbar\omega')\Big]\\ =& \frac{-e^2 g_s g_v}{8\pi\hbar^2}\left\{\hbar\omega\mathcal{P}\int_{\max(2|\mu|/\hbar\omega,2t/\hbar\omega)}^\infty dy\frac{y^2}{y^2-1}+t\mathcal{P}\int_{(2|\mu|/\hbar\omega)^2}^{(2t/\hbar\omega)^2}\frac{dz}{z-1}\theta(t-|\mu|)\right\},\end{aligned} \tag{A9}$$

with $y=\omega'/\omega$ and $z=(\omega'/\omega)^2$. The first integral diverges in the Dirac cone approximation, so we proceed by separating the divergent part $y^2/(y^2-1) = 1+1/(y^2-1)$ and integrate up to a frequency cutoff Λ. With this scheme, the first term of the real part gives

$$\hbar\omega\mathcal{P}\int_{\max(2|\mu|/\hbar\omega,2t/\hbar\omega)}^\infty dx\frac{x^2}{x^2-1} = \hbar\Lambda-\max(2t,2|\mu|)-\frac{\hbar\omega}{2}\log\left(\left|\frac{\hbar\omega-\max(2t,2|\mu|)}{\hbar\omega+\max(2t,2|\mu|)}\right|\right), \tag{A10}$$

and the second term can easily be integrated. The total real part of χ_- is thus

$$\mathrm{Re}\chi_-(\omega) = \frac{-e^2 g_s g_v}{8\pi\hbar^2}\left\{-\max(2t,2|\mu|)-\frac{\hbar\omega}{2}\log\left(\left|\frac{\hbar\omega-\max(2t,2|\mu|)}{\hbar\omega+\max(2t,2|\mu|)}\right|\right)+t\log\left(\left|\frac{(2t)^2-(\hbar\omega)^2}{(2|\mu|)^2-(\hbar\omega)^2}\right|\right)\theta(t-|\mu|) + \frac{2t}{(2t)^2-(\hbar\omega)^2}\Big[2(\mu^2+t^2)\theta(t-|\mu|)+4t|\mu|\theta(|\mu|-t)\Big]\right\}. \tag{A11}$$

We have removed the part proportional to Λ above, because the constant contribution of the response due to the diamagnetic currents cancels exactly the cutoff dependent term, as required by gauge invariance.

Again, with $g_s, g_v = 2$ and in terms of $x = \frac{\hbar\omega}{t}$ and $\sigma_0 = \frac{e^2}{4\hbar}$, the real part reads

$$\mathrm{Re}\chi_-(x) = -\frac{\sigma_0 t}{\pi\hbar}\Bigg\{ -4\max(1,|\mu|/t) - x\log\left(\left|\frac{x-2\max(1,|\mu|/t)}{x+2\max(1,|\mu|/t)}\right|\right) + 2\log\left(\left|\frac{1-(x/2)^2}{(|\mu|/t)^2-(x/2)^2}\right|\right)\theta(1-|\mu|/t) + \frac{8}{4-x^2}\left[\left(1+\frac{\mu^2}{t^2}\right)\theta(1-|\mu|/t) + \frac{2|\mu|}{t}\theta(|\mu|/t-1)\right]\Bigg\}. \tag{A12}$$

2. AB-stacked bilayer graphene

Similarly to AA-stacked graphene, the imaginary part of the magnetic response can be decomposed into a sum of nonresonant and resonant transitions. The nonresonant or not nested part (excitations 1 and 2 in Fig. 5) reads

$$\mathrm{Im}\chi_-^{\text{not nested}}(\omega) = -\pi\frac{e^2 g_s g_v}{2A}\frac{A}{(2\pi)^2}\left(\pi e^2 v_F^2\right)\int kdk\{\delta(\hbar\omega-\xi_k)\theta(\xi_k-2|\mu+t/2|) + \delta(\hbar\omega-\xi_k)\theta(\xi_k-2|\mu-t/2|)\}. \tag{A13}$$

The prefactor has the same origin as in the AA-stacked case. In this case, πv_F^2 is the angular integral of $v_F^2\cos^2(\phi_k)$ or $v_F^2\sin^2(\phi_k)$, depending on the transition. Performing the integral yields

$$\mathrm{Im}\chi_-^{\text{not nested}}(\omega) = -\frac{e^2 g_s g_v}{8}v_F^2\left\{\frac{\hbar\omega}{(2\hbar v_F)^2}[\theta(\hbar\omega-2|\mu+t/2|)+\theta(\hbar\omega-2|\mu-t/2|)]\theta(\hbar\omega-t)\right\}. \tag{A14}$$

On the other hand, the nested transitions 3 and 4 in Fig. 5 contribute with

$$\mathrm{Im}\chi_-^{\text{nested}}(\omega) = -\frac{e^2 g_s g_v}{8}v_F^2\delta(\hbar\omega-t)\int kdk\{\theta(\xi_k-(2\mu-t))\theta(2\mu+t-\xi_k)+\theta(\xi_k-(-2\mu-t))\theta(-2\mu+t-\xi_k)\}. \tag{A15}$$

Keeping in mind that $\xi_k = \sqrt{t^2+(2\hbar v_F k)^2}$, the resonant part reads

$$\mathrm{Im}\chi_-^{\text{nested}}(\omega) = \frac{-e^2 g_s g_v}{32\hbar^2}\{[2|\mu|(|\mu|+t)\theta(t-|\mu|)+4|\mu|t\theta(|\mu|-t)]\delta(\hbar\omega-t)\}. \tag{A16}$$

Finally, the total imaginary part is the sum of both contributions, which yields

$$\mathrm{Im}\chi_-(\omega) = \frac{-e^2 g_s g_v}{32\hbar^2}\{\hbar\omega[\theta(\hbar\omega-2|\mu+t/2|)+\theta(\hbar\omega-2|\mu-t/2|)]\theta(\hbar\omega-t) + [2|\mu|(|\mu|+t)\theta(t-|\mu|)+4|\mu|t\theta(|\mu|-t)]\delta(\hbar\omega-t)\}. \tag{A17}$$

Using the dimensionless variable $x = \frac{\hbar\omega}{t}$ and the constant $\sigma_0 = \frac{e^2}{4\hbar}$, we get with $g_s = g_v = 2$

$$\mathrm{Im}\chi_-(x) = -\frac{\sigma_0 t}{2\hbar}\Bigg\{x\Big[\theta(x-2|\mu/t+1/2|)+\theta(x-2|\mu/t-1/2|)\Big]\theta(x-1) + \left[\frac{2|\mu|}{t}\left(\frac{|\mu|}{t}+1\right)\theta(1-|\mu|/t)+\frac{4|\mu|}{t}\theta(|\mu|/t-1)\right]\delta(x-1)\Bigg\}. \tag{A18}$$

Again, we calculate the real part via the Kramers-Kronig relations. The delta function gives

$$\mathrm{Re}\chi_-^{\text{nested}}(\omega) = -\frac{e^2 g_s g_v}{16\pi\hbar^2}\Big[2|\mu|(|\mu|+t)\theta(t-|\mu|)+4|\mu|t\theta(|\mu|-t)\Big]\frac{t}{t^2-(\hbar\omega)^2}. \tag{A19}$$

The nonresonant part can be calculated in a way similar to the AA-stacked case and gives

$$\mathrm{Re}\chi_-^{\text{not nested}}(\omega) = -\frac{e^2 g_s g_v}{16\pi\hbar^2}\Bigg\{2\hbar\Lambda - \Big(\max(t,2|\mu-t/2|)+\max(t,2|\mu+t/2|)\Big) + \frac{\hbar\omega}{2}\log\left(\left|\frac{\max(t,2|\mu-t/2|)-\hbar\omega}{\max(t,2|\mu-t/2|)+\hbar\omega}\right|\right) + \frac{\hbar\omega}{2}\log\left(\left|\frac{\max(t,2|\mu+t/2|)-\hbar\omega}{\max(t,2|\mu+t/2|)+\hbar\omega}\right|\right)\Bigg\}. \tag{A20}$$

After canceling the cutoff term by the diamagnetic contribution, the total real part of the response function then reads

$$\mathrm{Re}\chi_-(\omega) = -\frac{e^2 g_s g_v}{16\pi\hbar^2}\Bigg\{ -(\max(t, 2|\mu - t/2|) + \max(t, 2|\mu + t/2|)) + \frac{\hbar\omega}{2}\log\left(\left|\frac{\max(t, 2|\mu - t/2|) - \hbar\omega}{\max(t, 2|\mu - t/2|) + \hbar\omega}\right|\right) + \frac{\hbar\omega}{2}\log\left(\left|\frac{\max(t, 2|\mu + t/2|) - \hbar\omega}{\max(t, 2|\mu + t/2|) + \hbar\omega}\right|\right) + \left[2|\mu|(|\mu| + t)\theta(t - |\mu|) + 4|\mu|t\theta(|\mu| - t)\right]\frac{t}{t^2 - (\hbar\omega)^2}\Bigg\}. \tag{A21}$$

In dimensionless variables this yields

$$\mathrm{Re}\chi_-(x) = -\frac{\sigma_0 t}{\pi\hbar}\Bigg\{ -\Big(\max(1, 2|\mu/t - 1/2|) + \max(1, 2|\mu/t + 1/2|)\Big) + \frac{x}{2}\log\left(\left|\frac{\max(1, 2|\mu/t - 1/2|) - x}{\max(1, 2|\mu/t - 1/2|) + x}\right|\right) + \frac{x}{2}\log\left(\left|\frac{\max(1, 2|\mu/t + 1/2|) - x}{\max(1, 2|\mu/t + 1/2|) + x}\right|\right) + \left[\frac{2|\mu|}{t}\left(\frac{|\mu|}{t} + 1\right)\theta(1 - |\mu|/t) + \frac{4|\mu|}{t}\theta(|\mu|/t - 1)\right]\frac{1}{1 - x^2}\Bigg\}. \tag{A22}$$

APPENDIX B: PLASMONS IN INHOMOGENEOUS ENVIRONMENTS

The transverse photonic Green's function when the bilayer lies between two dielectrics is given by [30]

$$\mathcal{D}_{\mathbf{t}} = -\frac{\mu_1\mu_2\mu_0 q'}{N_t}\begin{pmatrix}\cosh(q'a) + \frac{q'_2}{\mu_2 q'}\sinh(q'a) & 1\\ 1 & \cosh(q'a) + \frac{q'_1}{\mu_1 q'}\sinh(q'a)\end{pmatrix}, \tag{B1}$$

with $N_t = q'(\mu_2 q'_1 + \mu_1 q'_2)\cosh(q'a) + (q'_1 q'_2 + \mu_1\mu_2 q'^2)\sinh(q'a)$, and $q'_i = \sqrt{q^2 - \epsilon_i\mu_i(\omega/c)^2}$. In between the two graphene layers, we assume a vacuum with wave number $q' = \sqrt{q^2 - (\omega/c)^2}$. Plasmons are the solutions of $\det(\mathbf{1} - \chi\mathcal{D}_{\mathbf{t}}) = 0$. The zero eigenvalues of $\mathbf{1} - \chi\mathcal{D}_{\mathbf{t}}$ are the symmetric and antisymmetric plasmonic excitations. Let us calculate these. Expanding in powers of the $q'_i a$'s, the Green's function becomes

$$\mathcal{D}_{\mathbf{t}} = -\frac{\mu_1\mu_2\mu_0 a}{\mu_1 q'_2 a + \mu_2 q'_1 a}\left(1 - \frac{(q'_1 a)(q'_2 a)}{\mu_1 q'_2 a + \mu_2 q'_1 a} - \frac{\mu_1\mu_2(q'a)^2}{\mu_1 q'_2 a + \mu_2 q'_1 a}\right)\begin{pmatrix}1 + \frac{q'_2 a}{\mu_2} + \frac{(q'a)^2}{2} & 1\\ 1 & 1 + \frac{q'_2 a}{\mu_1} + \frac{(q'a)^2}{2}\end{pmatrix}. \tag{B2}$$

Hence, the eigenvalues (λ') of the matrix

$$\begin{pmatrix}\chi_{11} & \chi_{12}\\ \chi_{12} & \chi_{11}\end{pmatrix}\begin{pmatrix}1 + q'_2 a/\mu_2 + (q'a)^2/2 & 1\\ 1 & 1 + q'_2 a/\mu_1 + (q'a)^2/2\end{pmatrix} \tag{B3}$$

are related to the eigenvalues (λ) of $\mathbf{1} - \chi\mathcal{D}_{\mathbf{t}}$ by

$$\lambda = 1 + \frac{\mu_1\mu_2\mu_0 a}{\mu_1 q'_2 a + \mu_2 q'_1 a}\left(1 - \frac{(q'_1 a)(q'_2 a)}{\mu_1 q'_2 a + \mu_2 q'_1 a} - \frac{\mu_1\mu_2(q'a)^2}{\mu_1 q'_2 a + \mu_2 q'_1 a}\right)\lambda'. \tag{B4}$$

Up to order $(q'_i a)^2$, these solutions read

$$\lambda' = (\chi_{11} + \chi_{12}) + \chi_{11}\left(\frac{(q'a)^2}{2} + \frac{1}{2}\left(\frac{q'_1 a}{\mu_1} + \frac{q'_2 a}{\mu_2}\right)\right) \pm\left\{(\chi_{11} + \chi_{12}) + \chi_{12}\left(\frac{(q'a)^2}{2} + \frac{1}{2}\left(\frac{q'_1 a}{\mu_1} + \frac{q'_2 a}{\mu_2}\right)\right) + \frac{\chi_{11} - \chi_{12}}{8}\left(\frac{q'_1 a}{\mu_1} - \frac{q'_2 a}{\mu_2}\right)^2\right\}. \tag{B5}$$

Therefore, the plasmonic resonances of the system are the zeros of the following eigenvalues:

$$\lambda^- = 1 + \frac{\mu_0 a(\chi_{11} - \chi_{12})}{2}\left(1 - \frac{1}{4}\left(\frac{q'_2 a}{\mu_2} + \frac{q'_1 a}{\mu_1}\right) + O((q'_i a)^2)\right), \tag{B6}$$

$$\lambda^+ = 1 + \frac{\mu_0 a}{q'_1 a/\mu_1 + q'_2 a/\mu_2}\Bigg\{(\chi_{11} + \chi_{12})\left(2 + \left(\frac{(q'_1 a/\mu_1 - q'_2 a/\mu_2)^2 - (2q'a)^2}{2(q'_1 a/\mu_1 + q'_2 a/\mu_2)}\right) + 2\left(\frac{(q'_1 a/\mu_1)(q'_2 a/\mu_2) + (q'a)^2}{q'1_a/\mu_2 + q'_2 a/\mu_2}\right)^2 - \left(\frac{2(q'a)^2 + (q'_1 a/\mu_1)(q'_2 a/\mu_2)}{2}\right) + \frac{\chi_{11} - \chi_{12}}{8}(q'_1 a/\mu_1 - q'_2 a/\mu_2)^2\Bigg\} + O((q'_i a)^2). \tag{B7}$$

One thing to note is that this approximation breaks down if the dielectric media 1 and 2 are the same for $q_1' = 0$. In this case, when approaching the slower light cone, $q_1'a$ and $q_2'a$ tend to 0 while $q'a$ tends to a constant, and the terms with $(q'a)^2/(q_1'a/\mu_1 + q_2'a/\mu_2)$ are no longer small and, in fact, become infinite. However, the divergence kicks off at very small values of $q_1'a$ and we can trust the formula for essentially the entire dynamic range of q, as we will see.

A simple calculation tells us that $(q'a)^2/(2q_1'a/\mu_1) \sim 1$ when $qa \sim \omega a/c_1 + \mu_1^2/8((c/c_1)^2 - 1)^2(c_1/c)(\omega a/c)^3$ or, equivalently, $q_1'a \sim \mu_1/2((c/c_1)^2 - 1)(\omega a/c)^2$(where c_1 is the speed of light in medium 1). Taking the frequency to be of the order of the interlayer coupling, $\hbar\omega = 0.5$ eV, we have $\omega a/c \sim 10^{-3}$, $qa \sim \omega a/c_1 + 10^{-9}$, and $q_1'a \sim 10^{-6}$. On the other hand, setting the momentum scale to, say, $q_0 = G/100 \sim 3 \times 10^8$ m^{-1}, where G is the modulus of the reciprocal lattice vector of graphene and the interlayer distance to $a = 3.4$ Å, we have that our approximation is valid up to $q_1'/q_0 \sim 10^{-5}$, or $q/q_0 \sim 10^{-8}$ to the right of the light cone of the dielectric.

Setting $q_1'a, q_2'a = 0$ in Eq. (B1), the eigenvalues at the dielectric light cone are $\lambda^+ = 1 + 2\mu_0 a(\chi_{11} + \chi_{12})/(q'a)^2$ and $\lambda^- = 1 - \mu_0 a(\chi_{11} - \chi_{12})/2$. In conclusion, the symmetric eigenvalue remains bounded and we can trust Eq. (B6) even at $q_1' = 0$.

On the other hand, if the dielectric media are not the same, the quantity $(q'a)^2/(q_1'a/\mu_1) + (q_2'a/\mu_2))$ will not blow up, and for refractive indices $|n_1 - n_2| \gtrsim 10^{-6}$ the approximation is good all the way up to $q_1' = 0$ (1 being the slower medium).

APPENDIX C: INFLUENCE OF AN OPTICAL CAVITY

Let us consider the bilayer placed inside an optical cavity, with boundaries at $z = \pm L/2$. In this setting, the Green's function or photonic propagator for the transverse fields of a source located at z' takes the form

$$D_l(z, z') = \frac{\mu_0}{q'} \frac{\sinh(q'(z^> - L/2))\sinh(q'(z^< + L/2))}{\sinh(q'L)} \tag{C1}$$

for $-L/2 < z < L/2$ and $D_l(z, z') = 0$ otherwise [where $z^< = \min(z, z')$ and $z^> = \max(z, z')$]. If the bilayer is located at $z = \pm a/2$, the Green's function of the double-layer structure is

$$\mathcal{D}_\mathbf{t} = \frac{\mu_0}{q'\sinh(q'L)} \begin{pmatrix} \sinh((a-L)/2)\sinh((a+L)/2) & -\sinh^2((L-a)/2) \\ -\sinh^2((L-a)/2) & \sinh((a-L)/2)\sinh((a+L)/2) \end{pmatrix}. \tag{C2}$$

Transverse plasmons are obtained after the in-phase and out-of-phase (counterflow) combinations:

$$1 + \frac{\mu_0 a\chi_+}{q'a} \frac{\cosh(a/2)\sinh((L-a)/2)}{\cosh(L/2)} = 0, \tag{C3}$$

$$1 + \frac{\mu_0 a\chi_-}{q'a} \frac{\sinh(a/2)\sinh((L-a)/2)}{\sinh(L/2)} = 0. \tag{C4}$$

The form factors are increasing functions of L, so the light-matter coupling is larger when there is no cavity. Hence, the presence of the cavity does not favor the presence of these elusive excitations, as pointed out in Ref. [41]. This is due to the "perfect" screening of the magnetic excitations.

[1] J. Chen, M. Badioli, P. Alonso-Gonzalez, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenovic, A. Centeno, A. Pesquera, P. Godignon, A. Zurutuza Elorza, N. Camara, F. J. G. de Abajo, R. Hillenbrand, and F. H. L. Koppens, Nature (London) **487**, 77 (2012).

[2] 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. C. Neto, C. N. Lau, F. Keilmann, and D. N. Basov, Nature (London) **487**, 82 (2012).

[3] A. N. Grigorenko, M. Polini, and K. S. Novoselov, Nat. Photon. **6**, 749 (2012).

[4] Y. V. Bludov, A. Ferreira, N. M. R. Peres, and M. I. Vasilevskiy, Int. J. Mod. Phys. B **27**, 1341001 (2013).

[5] T. Stauber, J. Phys.: Condens. Matter **26**, 123201 (2014).

[6] P. A. D. Gonçalves and N. M. R. Peres, *An Introduction to Graphene Plasmonics* (World Scientific, Singapore, 2016).

[7] T. Low, A. Chaves, J. D. Caldwell, A. Kumar, N. X. Fang, P. Avouris, T. F. Heinz, F. Guinea, L. Martin-Moreno, and F. Koppens, Nat. Mater. **16**, 182 (2017).

[8] F. H. L. Koppens, D. E. Chang, and F. J. Garcia de Abajo, Nano Lett. **11**, 3370 (2011).

[9] D. Rodrigo, O. Limaj, D. Janner, D. Etezadi, F. J. García de Abajo, V. Pruneri, and H. Altug, Science **349**, 165 (2015).

[10] M. A. Baqir, P. K. Choudhury, A. Farmani, T. Younas, J. Arshad, A. Mir, and S. Karimi, IEEE Photon. J. **11**, 1 (2019).

[11] T. Stauber, T. Low, and G. Gómez-Santos, Nano Lett. **20**, 8711 (2020).

[12] A. Scholz, T. Stauber, and J. Schliemann, Phys. Rev. B **88**, 035135 (2013).

[13] T. Low and P. Avouris, ACS Nano **8**, 1086 (2014).

[14] D. N. Basov, M. M. Fogler, and F. J. García de Abajo, Science **354**, aag1992 (2016).

[15] S. A. Mikhailov and K. Ziegler, Phys. Rev. Lett. **99**, 016803 (2007).

[16] F. Ramos-Mendieta, J. A. Hernández-López, and M. Palomino-Ovando, AIP Adv. **4**, 067125 (2014).

[17] S. G. Menabde, D. R. Mason, E. E. Kornev, C. Lee, and N. Park, Sci. Rep. **6**, 21523 (2016).

[18] X. Zhang, H. Hu, X. Lin, L. Shen, B. Zhang, and H. Chen, npj 2D Mater. Appl. **4**, 25 (2020).
[19] A. Reserbat-Plantey, I. Epstein, I. Torre, A. T. Costa, P. A. D. Gonçalves, N. A. Mortensen, M. Polini, J. C. W. Song, N. M. R. Peres, and F. H. L. Koppens, ACS Photon. **8**, 85 (2021).
[20] N. Chamanara and C. Caloz, Phys. Rev. B **94**, 075413 (2016).
[21] M. Jablan, H. Buljan, and M. Soljavcić, Opt. Express **19**, 11236 (2011).
[22] V. Andreeva, M. Luskin, and D. Margetis, Phys. Rev. B **98**, 195407 (2018).
[23] A. Principi, M. Polini, and G. Vignale, Phys. Rev. B **80**, 075418 (2009).
[24] T. Stauber and G. Gómez-Santos, Phys. Rev. B **82**, 155412 (2010).
[25] A. Gutiérrez-Rubio, T. Stauber, and F. Guinea, J. Opt. **15**, 114005.
[26] O. V. Kotov, M. A. Kol'chenko, and Y. E. Lozovik, Opt. Express **21**, 13533 (2013).
[27] D. N. Basov, A. Asenjo-Garcia, P. J. Schuck, X. Zhu, and A. Rubio, Nanophotonics **10**, 549 (2021).
[28] T. Stauber and G. Gómez-Santos, Phys. Rev. B **85**, 075410 (2012).
[29] T. Stauber, T. Low, and G. Gómez-Santos, Phys. Rev. Lett. **120**, 046801 (2018).
[30] T. Stauber and G. Gómez-Santos, New J. Phys. **14**, 105018 (2012).
[31] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, UK, 2005).
[32] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
[33] E. McCann and M. Koshino, Rep. Prog. Phys. **76**, 056503 (2013).
[34] A. Rozhkov, A. Sboychakov, A. Rakhmanov, and F. Nori, Phys. Rep. **648**, 1 (2016).
[35] C. J. Tabert and E. J. Nicol, Phys. Rev. B **86**, 075439 (2012).
[36] R. Roldán and L. Brey, Phys. Rev. B **88**, 115420 (2013).
[37] D. S. L. Abergel and V. I. Fal'ko, Phys. Rev. B **75**, 155430 (2007).
[38] T. Low, F. Guinea, H. Yan, F. Xia, and P. Avouris, Phys. Rev. Lett. **112**, 116801 (2014).
[39] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, New J. Phys. **8**, 318 (2006).
[40] R. R. Nair, P. Blake, A. N. Grigorenko, K. S. Novoselov, T. J. Booth, T. Stauber, N. M. R. Peres, and A. K. Geim, Science **320**, 1308 (2008).
[41] P. Nataf, T. Champel, G. Blatter, and D. M. Basko, Phys. Rev. Lett. **123**, 207402 (2019).
[42] G. M. Andolina, F. M. D. Pellegrino, V. Giovannetti, A. H. MacDonald, and M. Polini, Phys. Rev. B **102**, 125137 (2020).
[43] T. Stauber, T. Low, and G. Gómez-Santos, Phys. Rev. B **98**, 195414 (2018).
[44] A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. A. Kastner, and D. Goldhaber-Gordon, Nano Lett. **21**, 4299 (2021).
[45] D. Guerci, P. Simon, and C. Mora, Phys. Rev. B **103**, 224436 (2021).
[46] A. L. Rakhmanov, A. V. Rozhkov, A. O. Sboychakov, and F. Nori, Phys. Rev. Lett. **109**, 206801 (2012).
[47] L. Brey and H. A. Fertig, Phys. Rev. B **87**, 115411 (2013).
[48] D. S. de la Peña, M. M. Scherer, and C. Honerkamp, Ann. Phys. **526**, 366 (2014).