


Spin-orbit interaction enabled electronic Raman scattering from charge collective modesSurajit Sarkar,¹ Alexander Lee,² Girsh Blumberg^{2,3} and Saurabh Maiti^{1,4}¹*Department of Physics, Concordia University, Montreal QC H4B 1R6, Canada*²*Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854, USA*³*National Institute of Chemical Physics and Biophysics, 12618 Tallinn, Estonia*⁴*Centre for Research in Molecular Modelling, Concordia University, Montreal QC H4B 1R6, Canada* (Received 9 October 2023; revised 4 January 2024; accepted 5 January 2024; published 25 January 2024)

Electronic Raman scattering couples to the charge excitations in the system, including the plasmons. However, the plasmon response has a spectral weight of $\sim q^2$, where q , the momentum transferred by light, is small. In this paper, we show that in inversion-symmetry broken systems where Rashba type spin-orbit coupling affects the states at the Fermi energy (which is a known low-energy effect) as well as the transition elements to other states (a high-energy effect), there is an additional coupling of the plasmons to the Raman vertex, even at zero momentum transfer, that results in a spectral weight that is proportional to the spin-orbit coupling. The high-energy effect is due to the breaking of SU(2) spin invariance in the spin-flip transitions to the intermediate state. We present a theory for this coupling near the resonant regime of Raman scattering and show that in giant Rashba systems it can dominate over the conventional q^2 weighted coupling. We also provide experimental support along with a symmetry based justification for this spin-mediated coupling by identifying a prominent c -axis plasmon peak in the fully symmetric A_1 channel of the resonant Raman spectrum of the giant Rashba material BiTeI. This new coupling could lead to ways of manipulating coherent charge excitations in inversion-broken systems. This process is also relevant for spectroscopic studies in ultrafast spectroscopies, certain driven Floquet systems and topologically nontrivial phases of matter where strong inversion-breaking spin-orbit coupling plays a role.

DOI: [10.1103/PhysRevB.109.035160](https://doi.org/10.1103/PhysRevB.109.035160)**I. INTRODUCTION**

Raman scattering has long been used to study charge fluctuations in crystalline systems [1–5]. The most commonly probed Raman excitations are optical phonon modes in solids and other vibrational modes of molecules. Electronic Raman spectroscopy (eRS) that studies the charge dynamics of purely electronic degrees of freedom is relatively more challenging. In single band systems, the response is weighted by q^2 , where \mathbf{q} is the momentum transferred by light scattering and usually corresponds to the long wavelength ($q \rightarrow 0$) regime in solids. In multiband systems where the response could be viewed as a sum of independent contributions from each band, the same problem persists. In the literature this q^2 suppression has been attributed to charge conservation and/or to Coulomb screening. To overcome the smallness of the response due to q^2 , one might turn towards *resonant* eRS where the incoming laser light is made to resonate between states at the Fermi level and those at a higher or lower energy band (see Fig. 1). While this resonance boosts the overall signal for the electronic excitations near the Fermi surface [4], the strength is still difficult to detect.

Updating the above description is one of the focuses of this paper. What is known is that the eRS response from a spin-degenerate single-band system consists of the incoherent excitations from the two-particle continuum of particle-hole (ph) excitations and coherent excitations from a collective mode, the plasmon, at energies *above* the ph continuum. Both these excitations couple to eRS in the fully symmetric channel (which is captured by the parallel-polarization set-up of

an eRS experiment) and are weighted by q^2 . This channel couples to excitations that preserve all the crystal symmetries. Resonant eRS can amplify signals from both these contributions [6–8]. In the presence of spin splitting of the bands, which can be brought about due to the Zeeman effect in an external field or via the Rashba effect in inversion broken systems, the interband excitations also couple to eRS, but without the q^2 factor. However, these excitations involve spin flips and are only accessible in the antisymmetric channel (which is odd under swapping the incident and scattered polarizations and can be captured in the cross-polarization set-up of the experiment) [9]. The response in the fully symmetric channel, however, is still weighted by q^2 even in the presence of spin splitting.

In this paper we provide evidence for a novel coupling channel to plasmonic excitations in systems with broken inversion symmetry. Such systems have large spin-orbit coupling (SOC) of orbital origin, which break the SU(2) spin invariance. In the presence of such a SOC, there are two effects that arise: (i) the usual spin splitting in the conduction band where the Fermi level lies; and (ii) the SOC induced spin-dependent modification of the interband transitions to and from the Fermi level to the intermediate states. Earlier studies on the effects of SOC on collective modes and their eRS signatures mostly focused on (i) which led to novel responses in the spin sector in the form of new chiral-spin waves [9–14], but (ii) was not explored to our knowledge, which is what we do here. We substantiate our theory with support from the experimental observation of a prominent collective mode peak in the polar semiconductor, BiTeI, in

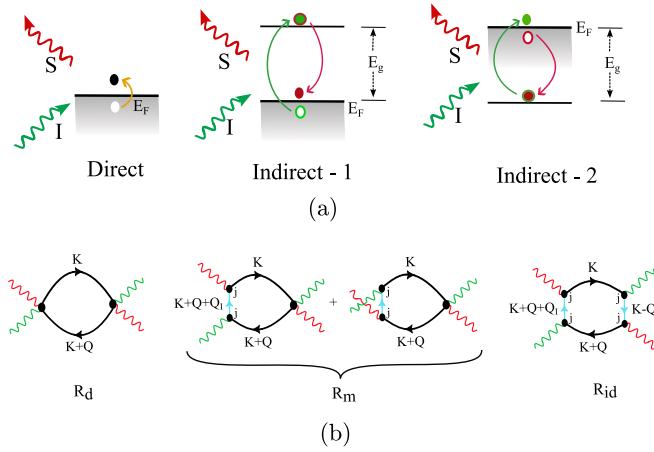


FIG. 1. (a) The direct and indirect Raman processes. (b) The response consists of contributions from the direct (R_d), mixed (R_m), and indirect (R_{id}) bubbles. The blue line indicates propagation in the intermediate state whereas the black lines are the particle and hole propagation near the Fermi level.

the fully symmetric channel of the resonant eRS spectrum. The collective mode frequency and its carrier concentration dependence ascertain that this mode is the c -axis plasmon of the system. BiTeI is a giant Rashba material with the SOC parameter ~ 4 eV \AA and is about two orders of magnitude larger [15–17] than those in 2D electron gases in heterostructures where it is usually induced by gating and/or structural asymmetry of the interfaces.

In particular, we show that in resonant eRS, the orbital Rashba SOC between the conduction band states and the intermediate states (a high-energy effect) generates an additional coupling to the plasmons, or any charge excitations in general, via a spin-charge susceptibility of the system, which is also known to be induced (a low-energy effect) due to Rashba SOC [18]. The spectral weight in this additional channel does not require q to be finite but is controlled by a factor E_{SOC}^2/E_F^2 , where E_{SOC} is the SOC energy, which is $\sim \lambda_{\text{SOC}} k_F$ (where λ_{SOC} is the spin-orbit coupling constant with units of velocity and k_F is the Fermi momentum), and E_F is the Fermi energy. In usual semiconductors, this ratio is small ($\sim 10^{-2}$) (see e.g., Table I of Ref. [19]) and hence the effect was likely never observed. However, in giant Rashba systems including BiTeI, $E_{\text{SOC}}^2/E_F^2 \sim 1$ and thus would have a visible effect. This high-energy effect is attributed to the breaking of SU(2) spin invariance in the spin-flip transitions between the Fermi level and intermediate bands. In fact, with regards to BiTeI, we argue that the direct observation of a c -axis plasmon in eRS is made possible due to a momentum-spin locking along the c axis, which would result in an out-of-plane canting of the spin states and a c -axis Drude-weight renormalization.

The discussion around coupling of charge fluctuations to eRS needs some consideration. For this reason, we organize the text as follows: In Sec. II we summarize the current understanding of the eRS spectrum involving the charge and spin degrees of freedom, the q^2 suppression, and discuss the various effects of SOC that have been investigated using eRS. In Sec. III, for completeness, we review the formulation of the resonant Raman vertex and arrive at some known results.

In Sec. IV we formulate, on general grounds, the origin of the non- q^2 coupling to the plasmon in SOC systems. In Sec. V we provide an explicit example (SOC in a Dirac system) for the high-energy effect demonstrating the presence of the matrix elements that are responsible for the enhancement of the plasmon pole due to the spin-charge coupling. In Sec. VI we discuss the theoretical and experimental results for BiTeI, present some symmetry considerations, followed by some other predictions in regards to BiTeI. Finally, we conclude in Sec. VII where we contemplate the implications of this result to other systems. In the Appendices we present details of the derivation of the resonant vertex, computation of the eRS for various polarizations, and the details of the toy models used in this paper.

II. ERS FROM CHARGE AND SPIN EXCITATIONS

The electronic response to light can be modeled as a correlation between Raman vertices computed at the momentum \mathbf{q} and frequency Ω transferred to the system [20]. The Raman vertices are density-like vertices dressed with factors arising from the direct (popularly called nonresonant) and indirect (which can be preresonant or resonant) Raman processes (see Fig. 1).

(a) *The q^2 factor.* In the nonresonant case, the entire Raman vertex (with both direct and indirect processes) is approximated by an effective-mass vertex, which can be reduced to the usual density vertex but modulated with form factors consistent with the irreducible representations of the lattice. In this approximation, if there are poles in the associated correlation function, they manifest themselves as $W_q^2/(\Omega^2 - \Omega_q^2)$, where $\Omega_q^2 = v^2 q^2$ for acoustic modes and $\Omega_q^2 = \Omega_0^2 + O(q^2)$ for optical modes, with v being the mode velocity, Ω_0 being the mode mass, and W_q representing the spectral weight. In the noninteracting case, $W_q \propto \Omega_q$ and thus, the acoustic modes have a spectral weight $\propto q^2$. This is why in optical experiments, where $aq \ll 1$ (a being the lattice constant of the material), acoustic modes do not have noticeable spectral weights.

In the presence of interactions, the pole at Ω_q is renormalized to Ω_q^* , while W_q could be renormalized differently. For acoustic modes in charge neutral systems, short range interactions renormalize v (which is usually related to the Fermi velocity), thereby changing the correlation function to $q^2/(\Omega^2 - v^{*2}q^2)$. This is ultimately what is responsible for zero sound in liquid He-3 [21,22]. In charged systems, however, long range interactions in the form of the unscreened part of the Coulomb interaction are also present, and are typically accounted for within the random-phase approximation (RPA) [23]. The Coulomb renormalization due to RPA has two important consequences for eRS: it affects the Raman response only in the fully symmetric channel; and it shifts the acoustic poles to either optical bulk plasmons in 3D or to “super-acoustic” ($\Omega_q \sim q^{1/2}$) ones in 2D, without affecting W_q^2 factor, which remains $\propto q^2$. Thus, the bulk plasmon in 3D systems, although technically an optical mode, still couples to the Raman process in the acoustic sense. In some sense, the plasmon is an acoustic mode that is singularly renormalized due to the Coulomb interaction. This renormalization is similar to the Anderson’s mechanism in superconductors where

the Goldstone mode expected from the spontaneously broken symmetry gets lifted to the bulk plasmon [24].

It is sometimes presented that the q^2 factor arises from charge conservation. This association is derived from the fact that for the fully symmetric channel in the nonresonant limit the Raman vertex reduces to the density vertex. In this case, the Raman correlation functions capture the density fluctuations. In the limit of $\mathbf{q} \rightarrow 0$ and finite Ω one is then probing the changes in total density, which is conserved. Hence, the correlation function and the corresponding response is zero. This argument also holds in the presence of interactions as charge conservation is still preserved. However, this argument cannot be directly extended to the resonant eRS as the Raman vertex does not reduce to the density vertex. Nevertheless, the response can still be proportional to q^2 as will be shown below.

(b) *Finite charge response in eRS.* In single band, spin-degenerate systems, only acoustic modes (which includes plasmons) are possible and hence $\{\Omega_q, W_q\} \propto q^2$. However, if there could be two bands or if the spin degeneracy could be lifted, the interband transitions would allow for modes of the form $\Omega_q^2 = \Omega_0^2 + \mathcal{O}(q^2)$ (along with the two-particle continuum of incoherent excitations). The eRS spectrum in such cases would not suffer from the q^2 suppression discussed above. Historically, the eRS observable interband effects could only be observed in either semiconductors in magnetic field (transition between Landau levels due to the orbital effect of the field [25] and between the spin-split bands due to the Zeeman effect [10]), or in superconductors [26,27] (excitations in the particle-hole symmetric Bogoliubov bands). Interband excitations in most other materials are usually of the order of eV, which is beyond the typical energy scales for the Raman shifts, and hence not observed. Note that the *total* charge is still conserved in multiband systems, but we get a response that is not suppressed by q^2 and electronic correlations often play an important role in causing the spectral weight to be observable. This is already known and understood in multiband superconductors as the Leggett response [27–30]. Thus, the association of charge conservation and q^2 -suppressed eRS is only of significance in single-band systems (or systems modelled as copies of single band systems) for the fully symmetric Raman channel.

In this context, it is relevant to discuss the result of Ref. [31], where there was no consideration of SOC and it showed that accounting for the resonant processes boosted the signal from the ph continuum of excitations in addition to the plasmon. This is relevant as it seemed to explain the results of experiments (in 1D systems) with two comparable peaks, where one corresponded to the plasmon and the other was attributed to the ph excitation [8]. The conventional theory involving RPA failed to explain the comparable weights of the ph excitations and the plasmon. While Ref. [31] provided an explanation for the comparable weights of those excitations, it is important to note that one needed finite q to get the response itself as the overall result still has to be proportional to q^2 . As we will demonstrate, our results for the eRS in an inversion-symmetry broken system remains finite even in the $q \rightarrow 0$ limit, and hence it really provides a coupling channel to the plasmon. It is also relevant to reiterate here that we are only interested in purely electronic mechanisms to couple to plasmons. See Ref. [32] that describes other “direct” means

of plasmon observation, which still require aide of phonons or some incoherent superposition of many q excitations.

(c) *Finite spin response in eRS.* When one accounts for the indirect terms in the Raman vertex, one is able to also couple to the spin-flip excitations in the cross-polarization set-up of the eRS, creating an effective spin vertex (e.g., in III-V semiconductors [1,4,5] and perovskite semiconductors [33]). This effective spin vertex was later reformulated for a square-lattice Hubbard model wherein it was pointed out that the indirect Raman processes could be used to extract the excitations in the antisymmetric channel [34,35] (which only contained information about the spin-flip excitations in the system). This property is what allowed the study of spin collective modes in strongly correlated systems (magnons) [20], in semiconductors (silin mode, at the Larmour frequency, and other spin-flip excitations) [10–13], in topological insulators (chiral-spin modes) [14], and in the heavy fermion superconductor URu₂Si₂ (chiral density wave) [36–38].

(d) *Previous works on the effect of SOC on resonant eRS.* Of the two effects of SOC on eRS discussed in the introduction, the theory for effect (i) which leads to new collective mode behavior is well understood [14,18,39–42]. The effect on resonant eRS has also been investigated theoretically but only in the context of the usual spin coupling of the Raman vertex. That is, the models were such that the coupling of light to the spin vertex was not modified but the spin fluctuations in the system were [43,44]. In such treatments, the resonant processes only trivially enhanced the spin susceptibility via the factor $\sim \frac{1}{(\Omega - \Omega_I)^2 + \Gamma^2}$, which was derived for a Kane model [5,45] for III-V semiconductors in cubic systems. Other studies that accounted for indirect processes in the eRS were formulated for graphene (without SOC) theoretically [46,47] and experimentally [48]. But none of the studies explored the effect (ii) mentioned in the Introduction.

III. THE RESONANT RAMAN SCATTERING CROSS SECTION

We shall model the eRS response in equilibrium where the cross-section per unit solid angle $d\mathcal{O}$ per unit scattered energy $d\Omega_S$ is given by [20,34,35,49]

$$\frac{d^2\sigma}{d\mathcal{O}d\Omega_S} = r_0^2 \frac{\Omega_S}{\Omega_I} [1 + n_B(\Omega)] R(\Omega, \mathbf{q}), \quad (1)$$

where r_0 is the classical radius of the electron, $\Omega_{I,S}$ are the frequencies of incident and scattered light with the Raman shift $\Omega = \Omega_I - \Omega_S$, $n_B(\Omega)$ is the Bose-Einstein distribution function, and $R(\Omega, \mathbf{q})$, which has dimensions of density of states, is computed from the branch cuts of the analytic continuation of [50]

$$\chi(Q) = \int_0^\beta d\tau e^{i\Omega_m\tau} \langle T_\tau \hat{\gamma}_\mathbf{q}(\tau) \hat{\gamma}_{-\mathbf{q}}(0) \rangle, \quad (2)$$

where $Q \equiv (i\Omega_m, \mathbf{q})$ with Ω_m being the bosonic Matsubara frequency, $\beta = 1/k_B T$, and $\hat{\gamma}_\mathbf{q}(\tau)$ is the Heisenberg evolution of $\hat{\gamma}_\mathbf{q}$ in imaginary time. Specifically, $R(\Omega, \mathbf{q}) = \text{Im}\chi(\Omega + i\eta, \mathbf{q})$. Further, $\hat{\gamma}_\mathbf{q} = \sum_{\mathbf{k}} \gamma_{\mathbf{k}}^{nn'} \hat{c}_{n,\mathbf{k}+\mathbf{q}/2}^\dagger \hat{c}_{n',\mathbf{k}-\mathbf{q}/2}$, where

(repeated indices are summed)

$$\begin{aligned} \gamma_{\mathbf{k}}^{nn'} &= \epsilon_{\alpha}^I \gamma_{\alpha\beta}^{nn'}(\mathbf{k}) \epsilon_{\beta}^S, \\ \gamma_{\alpha\beta}^{nn'}(\mathbf{k}) &= \underbrace{\delta_{\alpha\beta} \delta_{nn'}}_{\text{direct(d)}} + m_e \sum_m \underbrace{\left(\frac{[j_{\beta}^S]_{nm} [j_{\alpha}^I]_{mn'}}{\epsilon_n - \epsilon_m + \Omega_I - i\Gamma} + \frac{[j_{\alpha}^I]_{nm} [j_{\beta}^S]_{mn'}}{\epsilon_n - \epsilon_m - \Omega_S - i\Gamma} \right)}_{\text{indirect(id)}}. \end{aligned} \quad (3)$$

The $\hat{c}_{n,\mathbf{k}}^{\dagger}$, $\hat{c}_{n,\mathbf{k}}$ correspond to the creation and annihilation operators for states in band n with momentum \mathbf{k} . The direct and indirect labels refer to the processes shown in Fig. 1, $\alpha, \beta \in \{x, y, z\}$ (the spatial components of the polarization vectors ϵ^I and ϵ^S). ϵ_n are the eigenstates of the system, and the labels $n, n', m \in$ Hilbert space of the system in the eigenbasis. Γ encapsulates the lifetime effects of the intermediate states. The operator $\hat{j}_{\alpha} \equiv \frac{1}{e} \frac{\partial \hat{H}}{\partial A_{\alpha}} |_{\mathbf{A}=0}$, where \mathbf{A} is the vector potential,¹ and m_e is the effective electron mass for Hamiltonians with parabolic dispersion and the bare electron mass for those with linear dispersion (in which case there would also be no direct term). The two terms in Eq. (3) lead to three types of terms in $R(\Omega, \mathbf{q})$ [50],

$$R(\Omega, \mathbf{q}) = R_d(\Omega, \mathbf{q}) + R_m(\Omega, \mathbf{q}) + R_{id}(\Omega, \mathbf{q}), \quad (4)$$

where $R_d(\Omega, \mathbf{q})$ is the contribution from the correlations between the direct processes ($\sim \langle \gamma_d \gamma_{id} \rangle$), $R_m(\Omega, \mathbf{q})$ is the mixed contribution $\sim \langle \gamma_d \gamma_{id} \rangle$, and $R_{id}(\Omega, \mathbf{q})$ is from $\langle \gamma_{id} \gamma_{id} \rangle$.

Note. The nonresonant contribution arises from *both* direct and indirect processes while the resonant contribution arises only from indirect processes. In a field theory sense, the contributions from indirect processes can be further classified into on-shell and off-shell contributions. The former refers to transitions between the real energy states of the system, while the latter are virtual processes. These virtual processes could be interband or intraband in nature. In the absence of any on-shell contributions, the effects of off-shell contributions become relevant (see e.g., [46,47,50] where such terms were considered in nonsemiconducting systems). In resonant eRS one usually focuses on the on-shell contributions. We refer the reader to Appendix A for a more thorough discussion.

In this paper, we focus on the on-shell contribution from the R_{id} term. This is formally achieved by approximating $\gamma_{\alpha\beta}^{nn'} \approx m_e \sum_m \frac{[j_{\beta}^S]_{nm} [j_{\alpha}^I]_{mn'}}{\epsilon_n - \epsilon_m + \Omega_I - i\Gamma} + \frac{[j_{\alpha}^I]_{nm} [j_{\beta}^S]_{mn'}}{\epsilon_n - \epsilon_m - \Omega_S - i\Gamma}$. The electronic structure is usually such that resonance is only possible in either the term containing Ω_I (resonance with bands above Fermi level) or the one containing Ω_S (resonance with bands below). Without loss of generality, we retain the former (with Ω_I). Further, if the gap E_g between the Fermi surface states and those in the intermediate band is large, the dispersion of the intermediate states around the k_F wavevectors could be

¹Note that this derivative is carried out not in the eigenbasis, but in the basis where the \mathbf{A} is covariantly introduced to the system. At the level of the tight-binding model, this is usually done via Peierl's substitution. It can also be done by $\mathbf{k} \rightarrow \mathbf{k} + e\mathbf{A}$ in effective Hamiltonians. It is not guaranteed that the different prescriptions would agree.

ignored. We can then write $\gamma_{\alpha\beta}^{nn'} \approx m_e \sum_m \frac{[j_{\beta}^S]_{nm} [j_{\alpha}^I]_{mn'}}{\Omega_I - E_g - i\Gamma}$. In this approximation,

$$\begin{aligned} \chi(Q) &\approx -\Pi_{\alpha\beta;\gamma\delta}(Q) \epsilon_{\alpha}^I \epsilon_{\beta}^{S*} \epsilon_{\gamma}^{I*} \epsilon_{\delta}^S, \\ \Pi_{\alpha\beta;\gamma\delta}(Q) &\equiv \int_K \text{Tr}[\hat{\gamma}_{\alpha\beta}(\mathbf{k}) \hat{G}_K \hat{\gamma}_{\gamma\delta}(\mathbf{k}) \hat{G}_{K+Q}], \end{aligned} \quad (5)$$

where $K \equiv (i\omega_n, \mathbf{k})$ with ω_n being the fermionic Matsubara frequency, and $\int_K \equiv T \sum_n \int_{\mathbf{k}}$. The hat represents a matrix structure in the subspace containing only the Fermi surface states, and \hat{G}_K represents the Green's function in the same subspace. Note that the transition operators \hat{j}_{α} are defined in the Hilbert space made of both the Fermi surface states and the intermediate states. However, the resonance condition allows us to pick the dominant contributions, which effectively factor out the Hilbert subspace associated with the intermediate states from those at the Fermi level. The steps demonstrating this are detailed in Appendix B.

a. Interaction renormalization. In general, interactions renormalize the three terms $R_{d,m,id}$ differently, and the renormalization also varies with the choice of polarizations of the incoming and scattered light. Furthermore, the short-range part of the interaction predominantly leads to vertex corrections (which are captured in the ladder approximation in a diagrammatic approach), while the long-range part due to the unscreened part of Coulomb interaction leads to additional contributions, which are captured via the random-phase-approximation (RPA) (see Fig. 2). This

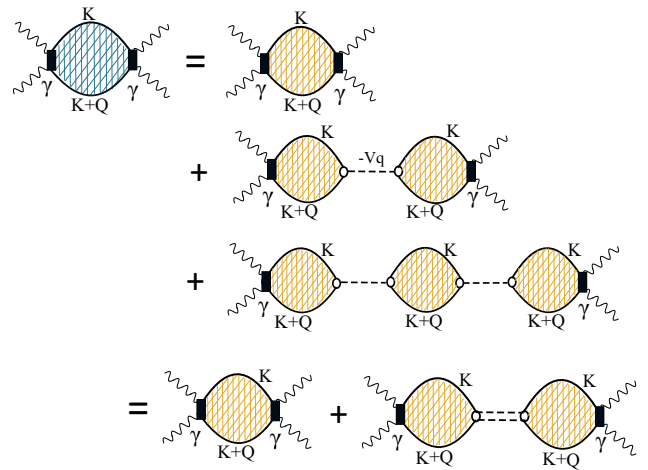


FIG. 2. Renormalization from the unscreened Coulomb interaction V_q in the RPA scheme. The shaded yellow bubbles indicate vertex corrections, which are ignored in this work. The doubled line is the RPA-renormalized Coulomb interaction.

leads to

$$\Pi_{\alpha\beta;\gamma\delta}(\mathcal{Q}) = \Pi_{\alpha\beta;\gamma\delta}^V(\mathcal{Q}) + \frac{\Pi_{\alpha\beta;0}^V(\mathcal{Q})V_{\mathbf{q}}\Pi_{0;\gamma\delta}^V(\mathcal{Q})}{1 - V_{\mathbf{q}}\Pi_{00}^V(\mathcal{Q})}, \quad (6)$$

where Π^V denotes the vertex corrected form of the bare Π (which are described below), and $V_{\mathbf{q}} = e^2/\epsilon_0q^2$ (the Fourier transform of the Coulomb interaction in 3D). Π^V leads to either spin collective modes [18] or other acoustic modes. If we wish to focus on the plasmons, we can drop the vertex corrections in all the terms in Eq. (6), and can identify the various Π 's as

$$\begin{aligned} \Pi_{\alpha\beta;0}(\mathcal{Q}) &= \int_K \text{Tr}[\hat{\gamma}_{\alpha\beta}\hat{G}(K)\hat{\sigma}_0\hat{G}(K+\mathcal{Q})], \\ \Pi_{0,\alpha\beta}(\mathcal{Q}) &= \int_K \text{Tr}[\hat{\sigma}_0\hat{G}(K)\hat{\gamma}_{\alpha\beta}\hat{G}(K+\mathcal{Q})], \\ \Pi_{ij}(\mathcal{Q}) &= \int_K \text{Tr}[\hat{\sigma}_i\hat{G}(K)\hat{\sigma}_j\hat{G}(K+\mathcal{Q})], \end{aligned} \quad (7)$$

where $\hat{\sigma}_{1,2,3}$ are the spin vertices and $\hat{\sigma}_0$ is the Coulomb vertex in the reduced Hilbert space of states at the Fermi surface. $\Pi_{ij}(\mathcal{Q})$ is usually referred to as the generalized susceptibility of a system. The denominator in Eq. (6) is the same one that appears in the renormalization of the charge susceptibility and contains the plasmon pole of the system. Note that the plasmon pole couples to the Raman response via $\Pi_{\alpha\beta;0}^V$ and $\Pi_{0;\gamma\delta}^V$, which are sensitive to the choice of polarization vectors of the eRS experiment. Since we wish to work with frequencies close to the plasmon pole, we could further write

$$\begin{aligned} \Pi_{\alpha\beta;\gamma\delta}(\mathcal{Q}) &\approx \frac{\Pi_{\alpha\beta;0}(\mathcal{Q})V_{\mathbf{q}}\Pi_{0;\gamma\delta}(\mathcal{Q})}{1 - V_{\mathbf{q}}\Pi_{00}(\mathcal{Q})}, \\ &= \frac{e^2}{\epsilon_0q^2} \frac{\Pi_{\alpha\beta;0}(\mathcal{Q})\Pi_{0;\gamma\delta}(\mathcal{Q})}{1 + \frac{\Omega_{\text{pl}}^2}{\Omega_m^2}}, \end{aligned} \quad (8)$$

where the plasma frequency is given by $\Omega_{\text{pl}}^2 = e^2v_F v_F^2/3\epsilon_0$. In semiconductors the ϵ_0 should be replaced by ϵ_∞ to account for the background charges. Here we have used the fact that $\Pi_{00}(\mathcal{Q}) = -v_F v_F^2 q^2/3\Omega_m^2$, where v_F is the density of states at the Fermi level and v_F is the Fermi velocity (which are to be seen as Fermi surface averages in anisotropic systems).

b. Known results. For the plasmon pole to show up in the response, one needs $\Pi_{\alpha\beta;0}$ and $\Pi_{0;\gamma\delta}$ to be nonzero. Diagrammatically, this object is a bubble with the $\alpha\beta$ vertex on one side and the charged vertex on the other. We may call this the *polarization-charge bubble*. This is a central object in our calculations and is composed of (see Fig. 3) three propagators: the first one reflects propagation in the intermediate state after the photon excitation, while the other two reflect the propagation of the created particle-hole pair at the Fermi level after photon emission. This bubble is unique to the contribution in the symmetric channel of eRS. We summarize the computation of this bubble in Appendix C but the end result is that

$$\Pi_{\alpha\beta;0}(\mathcal{Q}) \approx \frac{m_e \mathcal{F}_i^{\beta\alpha} \Pi_{i0}(\mathcal{Q})}{\Omega_I - E_g - i\Gamma}, \quad (9)$$

where $i \in \{0, 1, 2, 3\}$, and $\mathcal{F}_i^{\beta\alpha}$ is a *polarization factor* that couples the charge-charge ($i=0$) and spin-charge ($i=1, 2, 3$) susceptibilities to the Raman vertex and has units

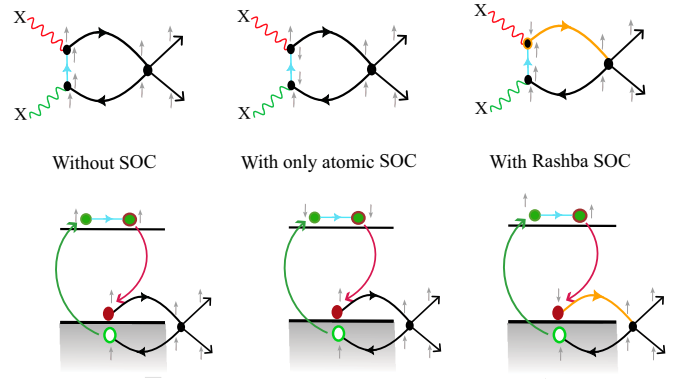


FIG. 3. Schematic view of the polarization-charge bubble, which represents the final state renormalizations by particle-hole propagation near the Fermi level in the fully symmetric channel of eRS. In the cases without SOC and with atomic SOC the final state is dressed by the usual charge-like bubble (black propagators) with a weight of q^2 . Broken inversion contributes additional spin-flip processes (as yellow vertices and propagators) that provide a stronger coupling to the plasmons in the fully symmetric channel.

of v_F^2 . It is easy to show that $\Pi_{0;\gamma\delta}(\mathcal{Q})$, within the same approximations, has $i\Gamma \rightarrow -i\Gamma$. The computation of the polarization factor is sensitive to the choice of polarizations of light and is central to our problem. In the parallel polarization set-up, $\alpha\beta = XX$ (or YY), and $\mathcal{F}_0^{XX} \neq 0$, while $\mathcal{F}_i^{XX} = 0$ for $i \in \{1, 2, 3\}$. Thus, the charge-charge susceptibility [$\Pi_{00}(\mathcal{Q})$], which is renormalized by the unscreened Coulomb interaction, is picked up in the parallel polarization set-up. On the other hand, for a cross-polarization set-up where $\alpha\beta = XY$, we have $\mathcal{F}_0^{XY} = 0$, while some of $\mathcal{F}_i^{XY} \neq 0$ for $i \in \{1, 2, 3\}$. Thus, the cross-polarization set-up couples to the spin-charge susceptibility of the system. The XY set-up can also couple to the spin-spin susceptibility in the non-RPA channel via vertex corrections, but those diagrams are presently not in consideration and have been a subject of previous studies [43,44]. Further, the spin-charge susceptibility itself is zero when there is spin-charge separation. This is why we can usually only couple to the charge-charge susceptibility [$\Pi_{00}(\mathcal{Q})$] in the XX set-up leading to the conventional result,

$$\Pi_{\alpha\alpha;\alpha\alpha}(\mathcal{Q}) = \frac{v_F^2 q^2}{\Omega_m^2 + \Omega_{\text{pl}}^2} \frac{v_F (m_e \mathcal{F}_0^{\alpha\alpha})^2}{(\Omega_I - E_g)^2 + \Gamma^2}, \quad (10)$$

where α could be X or Y. The circular polarizations of eRS can be constructed out of polarization vectors $\epsilon^X \pm i\epsilon^Y$. Despite the enhancement from the interband resonance, the plasmon pole in parallel polarization setup is still weighted by q^2 .

If we were to account for spin-splitting of the Fermi surface states either due to the Zeeman effect or due to the Rashba effect, the charge susceptibility would still be $\sim q^2$ [18] and thus would not really affect the result in Eq. (10). The presence of Rashba SOC, however, modifies the spin-charge susceptibility and \mathcal{F} in a characteristic manner. Formulating this is the subject of the next section. This modification will not be present when the spin splitting is induced by the Zeeman effect.

IV. NOVEL EFFECT DUE TO SOC

It is clear from Eq. (8) that the plasmon coupling to the Raman vertex is controlled by the polarization-charge bubble $\Pi_{\alpha\beta;0}(Q)$, which near resonance can be approximated by Eq. (9). Thus, the computation of the polarization factors and the generalized susceptibilities is crucial. From Appendix C we learn that the term $j_\beta j_\alpha$ in the Raman vertex [Eq. (3)] is only evaluated between states where the Fermi level lies (say the v states) and intermediate states (say the c states), and can be expressed, on general grounds, as

$$[j_{\beta,vc}j_{\alpha,cv}] = \mathcal{F}_i^{\beta\alpha}(\mathbf{k})\hat{\sigma}_i. \quad (11)$$

This is where the polarization factors are introduced. It will be more instructive to actually arrive at this form by modeling the matrix elements of the transition elements $j_{\alpha,cv}$ themselves. Without loss of generality, we can say

$$\hat{j}_{\alpha,cv} = p_0^\alpha \hat{\sigma}_0 + p_1^\alpha \hat{\sigma}_1 + p_2^\alpha \hat{\sigma}_2 + p_3^\alpha \hat{\sigma}_3, \quad (12)$$

where p_0^α corresponds to the average of the two spin nonflip transitions, p_3^α corresponds to the differences in the two spin nonflip transitions, p_1^α and p_2^α correspond to the average and difference of the spin-flip transitions. This leads to

$$\begin{aligned} \hat{j}_{\beta,vc}\hat{j}_{\alpha,cv} &= \underbrace{(p_0^\beta p_0^\alpha + \vec{p}^\beta \cdot \vec{p}^\alpha)}_{\mathcal{F}_0} \hat{\sigma}_0 + \underbrace{([p_0^\beta \vec{p}^\alpha + p_0^\alpha \vec{p}^\beta] + i\vec{p}^\beta \times \vec{p}^\alpha)}_{\vec{\mathcal{F}}} \cdot \hat{\sigma} \\ &= \mathcal{F}_0^{\beta\alpha} \hat{\sigma}_0 + \mathcal{F}_1^{\beta\alpha} \hat{\sigma}_1 + \mathcal{F}_2^{\beta\alpha} \hat{\sigma}_2 + \mathcal{F}_3^{\beta\alpha} \hat{\sigma}_3 \\ &\equiv \mathcal{F}_i^{\beta\alpha} \hat{\sigma}_i. \end{aligned} \quad (13)$$

In systems with inversion and time-reversal symmetry, parity is conserved and the SU(2) invariance of the spins is maintained, which leads to the transition elements being such that either only p_0^α is present (when there is no atomic SOC splitting), or only $p_{1,2,3}^\alpha$ are present and equal (in the presence of atomic SOC splitting). This means that $\mathcal{F}_0^{\beta\alpha} = p_0^\beta p_0^\alpha$ or $\vec{p}^\beta \cdot \vec{p}^\alpha$, and $\mathcal{F}_i^{\beta\alpha} = i\vec{p}^\beta \times \vec{p}^\alpha$. Observe that if $\alpha = \beta$ only \mathcal{F}_0 component contributes and couples the Raman vertex to the charge-charge susceptibility of the system near resonance. Whereas if $\alpha \neq \beta$, then \mathcal{F}_i components are nonzero and couples the Raman vertex to the spin-charge susceptibilities of the system. However, in systems with inversion, these susceptibilities are zero and the plasmon pole can only couple to the Raman vertex through the \mathcal{F}_0 factor and hence the charge-charge susceptibility, leading to Eq. (10).

In inversion-broken systems, parity is broken, which results in the following changes: (a) SU(2) invariance is broken, which no longer constraints p_0^α and p_i^α as it did in the conventional case; (b) it allows for the coupling of spin and charge degrees of freedom, which makes the spin-charge susceptibilities nonzero. The effect (a) leads to the important result that $\mathcal{F}_i^{\beta\alpha}$ now is no longer just the cross product of the \vec{p} factors but also contains the $p_0 \vec{p}$ terms as in Eq. (13). This means that if $\alpha = \beta$ and the cross product vanishes, there is still a contribution from the mixing of the p_0 and p_i terms, which was prevented when SU(2) invariance was preserved. This is an effect that happens at large energy scales. The effect (b) is a low-energy effect that was explicitly shown in a 2D metal with

Rashba type SOC [18] and is expected to hold in general. In fact, for a system with a spin-orbit velocity of λ_{SOC} , the form is expected to be $\Pi_{i0}(Q) \approx v_F \lambda_{\text{SOC}} q / \Omega_m$ for $i \in \{1, 2, 3\}$ (the form for Π_{00} is still q^2 / Ω_m^2). These two effects lead to the fact that an effective spin vertex is created in the *parallel* polarization geometry, which couples to the spin-charge susceptibility of the system, which is also rendered nonzero due to SOC.

Plugging the form for Π_{i0} we just discussed into Eq. (8) we get

$$\begin{aligned} \Pi_{\alpha\beta;\gamma\delta}(Q) &\approx \frac{e^2}{\epsilon_0 q^2} \frac{m_e^2}{1 + \frac{\Omega_{\text{pl}}^2}{\Omega_m^2}} \frac{\mathcal{F}_i^{\beta\alpha} \Pi_{i0}(Q) \mathcal{F}_j^{\delta\gamma} \Pi_{j0}(Q)}{(\Omega_I - E_g)^2 + \Gamma^2} \\ &\xrightarrow{XX} \underbrace{\frac{\lambda_{\text{SOC}}^2}{v_F^2}}_{\text{SOC effect}} \underbrace{\frac{\Omega_{\text{pl}}^2}{\Omega_m^2 + \Omega_{\text{pl}}^2}}_{\text{plasmon pole}} \underbrace{\frac{v_F (m_e \mathcal{F}_0^{XX})^2}{(E_g - \Omega_I)^2 + \Gamma^2}}_{\text{Res. enhancement}} \\ &\quad + \mathcal{O}\left(\frac{v_F^2 q^2}{\Omega_m^2 + \Omega_{\text{pl}}^2}\right). \end{aligned} \quad (14)$$

The regular q^2 weighted plasmon pole that arises from $\Pi_{00} \sim q^2$ contribution is still present, but appears as a correction to the SOC induced term. Comparing Eqs. (10) and (14) we see that the ratio of the SOC induced term to the regular q^2 -weighted term is $\sim [\lambda_{\text{SOC}} \Omega_{\text{pl}} / v_F^2 q]^2 \equiv r^2$. This is the main result of this article. Thus, if we want the SOC induced response to dominate we would need to increase r . In usual semiconductors (used in heterostructure quantum wells) we have $v_F \sim 10^4 - 10^5$ m/s, $\lambda_{\text{SOC}} \sim 10^3$ m/s for the Rashba parameters of about 10 meV Å and $\Omega_{\text{pl}} \sim 100$ meV [51]. These numbers lead to $\lambda_{\text{SOC}} / v_F \sim 10^{-2}$, and $v_F q \sim 0.1 - 1$ meV (where q is the wavenumber of light). This leads to $r \sim 1-10$. However, for giant Rashba systems $\lambda_{\text{SOC}} / v_F \sim 1$ leading to $r \sim 100$. Since the response is $\propto r^2$ one can have up to four orders of magnitude increase in the effect. It should be asserted that the coupling to the spin-charge susceptibility is a general result that is a consequence of inversion breaking SOC, but the particular form of the eRS response in the above form is valid only near the plasmon pole.

Note that to get this effect, it is important to have a spin-dependent correction to the transition elements j_{cv} . It is not sufficient to have spin-splitting in the intermediate states and the Fermi surface states. The SOC-based modification of j_{cv} is expected in the presence of inversion breaking SOC of the Rashba type that couples spin and momentum. This spin-momentum coupling is essential because the transition element is proportional to the interband component of the current operator. A spin structure to the current operator can only be introduced in the presence of Rashba type SOC. Zeeman-like spin splitting will not achieve this. In fact, in the next section, we demonstrate this precisely for a simple toy model. Finally, we note that a Rashba coupling can also be externally introduced by gating 2D systems. But this effect usually manifests itself in the Rashba coupling of the quasiparticles at the Fermi surface. The intermediate states are unlikely to get coupled to the Fermi surface states. However, in the orbital Rashba systems, the Rashba coupling emerges from the inversion breaking in the unit cell and thus couples multiple bands.

V. TOY MODEL DEMONSTRATING THE HIGH-ENERGY EFFECT: SPIN SENSITIVITY OF THE TRANSITION ELEMENT

We now demonstrate a case where the proposed structure of $\mathcal{F}_i^{\alpha\beta}$ is induced due to Rashba type SOC. Consider the Dirac system: a hypothetical model of doped graphene on a transition metal dichalcogenide such that the chemical potential is in one of the valence bands. The Hamiltonian is given by

$$\hat{H} = v_F(\tau_z \hat{\sigma}_0 \hat{\sigma}_x k_x + \hat{\sigma}_0 \hat{\sigma}_y k_y) + \Delta \hat{\sigma}_0 \hat{\sigma}_z + \frac{\Delta_R}{2}(\tau_z \hat{\sigma}_y \hat{\sigma}_x - \hat{\sigma}_x \hat{\sigma}_y) + \frac{\Delta_Z}{2} \tau_z \hat{\sigma}_z \hat{\sigma}_0, \quad (15)$$

where Δ is the charge gap, Δ_R is the Rashba SOC energy and Δ_Z is the valley Zeeman SOC energy [52–54], $\hat{\sigma}$, $\hat{\sigma}$ are the spin and sublattice matrices, and $\tau_z = \pm 1$ marks the valley index. This Hamiltonian is written in the sublattice basis $\{a_\uparrow, a_\downarrow, b_\uparrow, b_\downarrow\}$, where a, b represent the two atoms of the unit cell. We refer to this as hypothetical because Δ , in real graphene, is too small to resonate with visible light. However, we can artificially let $\Delta \sim$ frequency of visible light. We choose this model for its analytical tractability and the fact that it allows for dipole-active transitions between the two bands with intraband and interband spin splitting. Finally, since we will be interested in the resonant terms, the analysis will not be affected by the lack of the direct process in Dirac systems.

Let us first ignore the valley Zeeman coupling ($\Delta_Z = 0$). In the resulting Hamiltonian we can perform $\mathbf{k} \rightarrow \mathbf{k} + e\mathbf{A}$, which leads to the definition of current to be $\hat{\mathbf{j}} = v_F \tau_z \hat{\sigma}_0 \hat{\sigma}$ in the a - b basis. To proceed with our calculations, we need to move into the eigenbasis denoted by $\{c_\uparrow, c_\downarrow, v_\uparrow, v_\downarrow\}$. However, instead of solving the problem exactly, it is sufficient to tackle the problem perturbatively in Δ_R . Upon transforming $\hat{\mathbf{j}}$ to the eigenbasis, we show in Appendix D that $\hat{\mathbf{j}}$ acquires a spin dependence in the interband sector.

To get an intuitive understanding of how this dependence arises, consider an intermediate basis where there is no SOC. This basis was used in Ref. [55]. In this basis the Hamiltonian (with SOC) takes the form

$$H = \begin{pmatrix} H^{cc} & H^{cv} \\ H^{vc} & H^{vv} \end{pmatrix}, \quad (16)$$

where

$$\begin{aligned} H^{cc} &= \hat{\sigma}_0 \xi_k + \frac{\Delta_R v_F}{2 \xi_k} (\mathbf{k} \times \hat{\mathbf{s}}) \cdot \hat{\mathbf{z}}, \\ H^{vv} &= -\hat{\sigma}_0 \xi_k - \frac{\Delta_R v_F}{2 \xi_k} (\mathbf{k} \times \hat{\mathbf{s}}) \cdot \hat{\mathbf{z}}, \\ H^{cv} &= \frac{\Delta_R}{2} \left(\frac{i \xi_k k_x - \Delta k_y}{k \xi_k} \hat{s}_x + \frac{i \xi_k k_y + \Delta k_x}{k \xi_k} \hat{s}_y \right), \\ H^{vc} &= -\frac{\Delta_R}{2} \left(\frac{i \xi_k k_x + \Delta k_y}{k \xi_k} \hat{s}_x + \frac{i \xi_k k_y - \Delta k_x}{k \xi_k} \hat{s}_y \right). \end{aligned} \quad (17)$$

Here $\xi_k = \sqrt{v_F^2 k^2 + \Delta^2}$. In the absence of Δ_R , H is block diagonal. Δ_R induces a correction to the diagonal terms, which results in the usual intraband Rashba SOC (which leads to the modification of the low-energy behavior) and also

induces block-off-diagonal terms (H^{cv}, H^{vc}), which leads to an interband Rashba SOC (this is what we call the high-energy effect). In this model, both interband and intraband Rashba couplings are the same Δ_R . However, when we allow for renormalizations from interactions and other higher order interband processes, the interband Δ_R will be renormalized differently from the intraband Δ_R . To reflect this, we shall use Δ_R to denote the intraband Rashba coupling and δ_R to denote the interband coupling. This changes the prefactors of H_{cv} and H_{vc} to δ_R . Because of $H_{cv,vc}$, which is \mathbf{k} and spin dependent, the resulting dipole transition element (the interband current operator) also acquires spin dependence. This coupling does not arise for a Zeeman effect and hence Rashba type SOC is essential to get this effect.

Returning back to the eigenbasis, the current operator can be computed to leading order in SOC (both Δ_R and δ_R) as

$$\hat{\mathbf{j}} = \begin{pmatrix} \mathcal{J}_{cc} & \mathcal{J}_{cv} \\ \mathcal{J}_{vc} & \mathcal{J}_{vv} \end{pmatrix}.$$

The explicit forms of the \mathcal{J} 's are listed in Appendix D. We can then use Eq. (13) to infer that (to linear order in $\{\Delta_R, \delta_R\}$)

$$\begin{aligned} \frac{\mathcal{F}_0^{XX}}{v_F^2} &= \frac{\mathcal{F}_0^{YY}}{v_F^2} = \frac{v_F^2 k^2 + 2\Delta^2}{2\xi_k^2}, & \frac{\mathcal{F}_0^{XY}}{v_F^2} &= \frac{i\tau_z \Delta}{\xi_k}, \\ \frac{\mathcal{F}_3^{XX}}{v_F^2} &= \frac{\mathcal{F}_3^{YY}}{v_F^2} = \frac{-v_F k \Delta^2 \delta_R}{2\xi_k^4}, & \frac{\mathcal{F}_3^{XY}}{v_F^2} &= \frac{-i\tau_z v_F k \Delta \delta_R}{2\xi_k^3}, \end{aligned} \quad (18)$$

and k will ultimately be restricted to the Fermi surface. Notice that if $\delta_R \rightarrow 0$, then only \mathcal{F}_0 survives, which couples the Raman vertex to the charge-charge susceptibility leading to the conventional result. However, in the presence of δ_R , only the contribution (upon summing over valleys) from \mathcal{F}_3^{XX} and \mathcal{F}_3^{YY} survive but not from \mathcal{F}_3^{XY} , indicating that this SOC induced contribution is only present in the XX/YY set up, but not in the XY set-up. To estimate the size of this high-energy effect, note that $k \rightarrow k_F$ and that the SOC induced \mathcal{F}_i 's are $\sim E_F \Delta \delta_R / [\max(E_F, \Delta)]^3$. In semiconductors, $E_F < \Delta$, and hence $\mathcal{F}_i \sim E_F \delta_R / \Delta^2$.

Next, if we now only consider the valley Zeeman term ($\Delta_R = 0, \Delta_Z \neq 0$), we note that $H_{cv,vc} = 0$ (Appendix D). This leads to $\mathcal{J}_{cv,vc} = 0$. Since there is no spin-dependent modification of the $\hat{\mathbf{j}}$ operators in the interband sector, we trivially get that $\mathcal{F}_0^{\beta\alpha}$ are unchanged (as expected), and $\mathcal{F}_i^{\beta\alpha} = 0$ for $i \in \{1, 2, 3\}$, indicating that the plasmon pole would not couple to the Raman vertex. These explicit calculations of the spectrum in the two scenarios with $\Delta_R = 0$ and $\Delta_Z = 0$ validate the assertions made in Sec. IV. While we have explicitly demonstrated that Rashba SOC leads to a spin-flip vertex in parallel polarization (and hence in the fully symmetric channel of the eRS), in this specific example of the 2D Dirac system, $\mathcal{F}_3^{\beta\alpha}$ was found to be nonzero. This means that the appropriate susceptibility that couples to the resonant Raman vertex is Π_{30} . In a strict 2D Dirac system, this component of the spin-charge susceptibility is zero unless a second layer (or multiple layers) is included. The physics of inducing a spin-flip vertex does not change, but the spin-charge susceptibility along the c -axis can now be probed. This motivates the study

of a 3D system with large Rashba SOC. This is precisely what we do next.

VI. RESONANT ERS IN BiTeI

We begin with a theoretical analysis for BiTeI, which has one of the largest known bulk Rashba coupling factors of any metallic system with the Rashba parameter α_R estimated to be in the range $\approx 3.7\text{eV \AA}$ to 4.0eV \AA [15,18,56–58], satisfying the requirements of the first factor in Eq. (14). To investigate the plasmon coupling, we need to evaluate $\Pi_{\alpha\beta;\gamma\delta}(Q)$. As noted in the previous section, the form of \mathcal{J}_{cv} is essential for the effect of interest. To compute the transition elements in BiTeI, one needs knowledge of the wavefunctions of the intermediate states. Although there is no model that provides information about the higher energy states, there exists an effective low-energy model for the conduction band electrons. We can proceed to model the intermediate states phenomenologically.

Consider the following low-energy continuum model for BiTeI:

$$\hat{H} = \left(\frac{k_1^2 + k_2^2}{2m_1} + \frac{k_3^2}{2m_3} \right) \hat{\sigma}_0 + \alpha_R (\vec{\sigma} \times \vec{k})_3 + \lambda \hat{\sigma}_3 k_3, \quad (19)$$

where m_1 and m_3 are the in-plane and out-of-plane masses, α_R is the in-plane Rashba SOC in BiTeI, and λ is a SOC for the out-of-plane direction. Here 1,2 refer to the x, y axes, and 3 refers to the z axis. This model, without the λ term, was previously used to study possible spin-collective modes [18]; however, this resulted in the expectation value of the spin for its eigenstates being strictly in-plane. The λ term is added² to account for the canting of the spins in the z direction as computed in Ref. [59]. The Matsubara Green's function for the noninteracting system is then given by

$$\hat{G}_K(i\omega_m) = \sum_{s=\pm} \frac{1}{i\omega_m - \varepsilon_{\mathbf{k}}^s + \mu} \hat{\Omega}_{\mathbf{k}}^s, \quad (20)$$

$$\hat{\Omega}_{\mathbf{k}}^s \equiv \frac{1}{2} \left[\hat{\sigma}_0 + s \left(\hat{\sigma}_1 \frac{\alpha_R k_2}{D_k} - \hat{\sigma}_2 \frac{\alpha_R k_1}{D_k} + \hat{\sigma}_3 \frac{\lambda k_3}{D_k} \right) \right]$$

where $D_k = \sqrt{\alpha_R^2 k_{\parallel}^2 + \lambda^2 k_3^2}$, $k_{\parallel}^2 = k_1^2 + k_2^2$, and $\varepsilon_{\mathbf{k}}^{\pm} = \frac{k_{\parallel}^2}{2m_1} \pm D_k$.

Using $\hat{G}_K(i\omega_m)$, we can compute $\Pi_{i0}(Q)$ using Eq. (7) as (we have carried out the Matsubara sums)

$$\Pi_{i0}(Q) = \sum_{s,s'=\pm} \int_{\mathbf{k}} \mathcal{N}_{s,s'}^i \left[\frac{n_F(\varepsilon_{\mathbf{k}}^s) - n_F(\varepsilon_{\mathbf{k}+\mathbf{q}}^{s'})}{i\Omega_m + \varepsilon_{\mathbf{k}}^s - \varepsilon_{\mathbf{k}+\mathbf{q}}^{s'}} \right], \quad (21)$$

$$\mathcal{N}_{s,s'}^i \equiv \text{Tr}[\sigma_i \hat{\Omega}_{\mathbf{k}}^s \sigma_0 \hat{\Omega}_{\mathbf{k}+\mathbf{q}}^{s'}],$$

²The correct form of this term is of the type $\bar{\lambda} k_3 f(k_1^2, k_2^2)$ such that along the Γ -A line of the Brillouin zone there is no SOC splitting. However, due to the chemical potential lying below the Dirac point, the Fermi surface has a donut topology, so that we will always be dealing with states away from the Γ -A line. We absorb all such dependency into a constant λ . What is essential is that it needs to be odd in k_3 and even in other momenta.

where $n_F(\varepsilon)$ is the Fermi distribution function. For BiTeI, we cannot compute $\mathcal{F}_i^{\beta\alpha}$ as we do not have the information about $\hat{\mathbf{j}}$. However, from Secs. IV and V we know that in the presence of orbital Rashba SOC, a spin dependence arises in the interband transition elements such that $\mathcal{F}_i^{XX} \neq 0$. This leads to the coupling to the susceptibilities Π_{i0} . In line with known experiments in BiTeI [32], we restrict the chemical potential (μ) to lie in the lower band $\varepsilon_{\mathbf{k}}^-$ and the momentum transfer from the light to the z direction. Next, we need Π_{i0} s to leading order in q_3 . Evaluating this we get (see Appendix E)

$$\Pi_{10}(Q) = \Pi_{20}(Q) = 0$$

$$\Pi_{30}(Q) = -\lambda q_3 \int_{\mathbf{k}} \frac{[-\partial_{\varepsilon} n_F(\varepsilon^-)] k_3 \partial_{k_3} \varepsilon^-}{D_k i\Omega_m} + \alpha^2 \lambda q_3 \int_{\mathbf{k}} \frac{[n_F(\varepsilon^+) - n_F(\varepsilon^-)]}{D_k^3} \frac{i\Omega_m k_{\parallel}^2}{\Omega_m^2 + 4D_k^2}. \quad (22)$$

As ascertained in Sec. IV, the spin-charge susceptibility $\Pi_{30} \propto \lambda q_3$. These can now be plugged into Eq. (9) to get

$$\Pi_{\alpha\beta;0}(Q) \approx \frac{\sqrt{m_{\alpha} m_{\beta}} \mathcal{F}_i^{\beta\alpha} \Pi_{i0}(Q)}{\Omega_I - E_g - i\Gamma}$$

$$= \frac{\sqrt{m_{\alpha} m_{\beta}} \mathcal{F}_3^{\beta\alpha} \Pi_{30}(Q)}{\Omega_I - E_g - i\Gamma}. \quad (23)$$

Here m_{α} refers to the mass in the α direction. Since our polarizations will be in plane, $m_{\alpha} = m_{\beta} = m_1$. This leads to

$$\Pi_{\alpha\beta;\beta\alpha}(Q) \approx \frac{e^2}{\epsilon_{\infty} q^2} \frac{\Pi_{30}^2(Q)}{(\Omega_I - E_g)^2 + \Gamma^2} \frac{(m_1 \mathcal{F}_3^{\beta\alpha})^2}{1 - V_{\mathbf{q}} \Pi_{00}(Q)},$$

$$\xrightarrow{q=q_3 \hat{z}} \frac{e^2}{\epsilon_{\infty} q_3^2} \frac{\Pi_{30}^2(Q)}{(\Omega_I - E_g)^2 + \Gamma^2} \frac{(m_1 \mathcal{F}_3^{\beta\alpha})^2}{1 - V_{\mathbf{q}} \Pi_{00}(Q)}, \quad (24)$$

where $\Pi_{00}(Q)$ is calculated as (see Appendix E)

$$\Pi_{00}(Q) = q_3^2 \int_{\mathbf{k}} \frac{[-\partial_{\varepsilon} n_F(\varepsilon^-)] (\partial_{k_3} \varepsilon^-)^2}{-\Omega_m^2} + \lambda^2 q_3^2 \int_{\mathbf{k}} \frac{[n_F(\varepsilon^+) - n_F(\varepsilon^-)]}{D_k^3} \frac{\alpha^2 k_{\parallel}^2}{\Omega_m^2 + 4D_k^2}. \quad (25)$$

As expected, it is $\propto q_3^2$, which counters the q_3^2 from $V_{\mathbf{q}}$. Further, the q_3^2 in the denominator of Eq. (24) is also countered by the q_3^2 in Π_{30}^2 , leading to a finite response even if $q_3 \rightarrow 0$. Observe that, and this is expected from the results discussed in Sec. V, the result is proportional to the SOC component λ , which accounts for canting of the spins out-of-plane. This is necessary to couple the spin degree of freedom to the q_3 dispersion. Had we chosen to work with an in-plane momentum (q_1 or q_2) then we would not need this canting effect (see Appendix E). However, we will be interested in the q_3 component as this is the scenario that will correspond to the eRS setup we will use to study BiTeI where the momentum transferred by light will be along the c axis.

To calculate the location of the plasmon pole itself we would need the solution to the equation $V_{\mathbf{q}} \Pi_{00}(Q) = 1$, which

is the condition for zero of the denominator of Eq. (8). This is computed in Appendix F and the result in real frequencies is

$$\Omega_{\text{pl}}^2 = \Omega_0^2 \frac{m_1^*}{\sqrt{m_3^*}} f_{\text{pl}}(\tilde{\alpha}, \tilde{\lambda}), \quad (26)$$

where $\Omega_0^2 \equiv n_0 e^2 / m_0 \epsilon_\infty$ with $n_0 \equiv \frac{(2m_0|\mu|)^{3/2}}{6\pi^2}$ being the number density of single spin parabolic electrons with a chemical potential $|\mu|$; $m_1^* \equiv m_1/m_0$, $m_3^* \equiv m_3/m_0$ (with m_0 being the bare electron mass); and f_{pl} is a dimensionless function of dimensionless variables $\tilde{\alpha} \equiv \alpha_R \sqrt{2m_1/|\mu|}$ and $\tilde{\lambda} \equiv \lambda \sqrt{2m_3/|\mu|}$, such that in the limit $\tilde{\alpha} \rightarrow 0$, $\tilde{\lambda} \rightarrow 0$, we get $f_{\text{pl}} \rightarrow 1$.

The function f_{pl} only weakly depends on Ω and hence can be treated as Ω independent for the estimation of the plasma frequency. The full numerical calculation that we use shortly will find the true pole in the eRS response function. However, as is clear from Eq. (25), the function f_{pl} consists of interband *spin-flip* transitions between the Rashba subbands with weight $\propto \lambda^2 q_3^2$. Consequently, this leads to a continuum of spin-flip excitations with the same weight. This means that the plasmon that would result in the q_3 direction would likely be damped by this continuum of spin-flip excitations. This is certainly not new. Even in 2DEG, the 2D plasmon is damped by the spin-flip continuum [60]. But this was the plasmon that dispersed in the plane. The same remained true in the 3D-in-plane plasmon that was considered in Ref. [18] (also a model for BiTeI). In that paper, however, the canting effect was not modelled and this led to the fact that the c -axis plasmon was neither renormalized by SOC, nor affected by the spin-flip continuum. What is new here is that in general for a system belonging to a polar group, the 3-component of spin can couple to the 3-component of momentum leading to the c -axis plasmon being renormalized by SOC and also be damped by the continuum. This is what is accounted for by the λ term in the model considered above.

A. Resonant eRS experiment

To test our theoretical analysis of BiTeI we took Raman measurements of an electronic collective mode in the fully symmetric channel under the following conditions. The incident light enters the samples parallel to the z axis of the crystal with light polarized along the optically equivalent x and y axes. The single crystals of BiTeI that were grown using the vertical Bridgman technique had concentrations of elemental I sufficient to ensure the chemical potential lay below the Dirac point (for characterization details, see [32]). The Raman spectra were taken in multiple polarization geometries to perform algebraic decomposition of the spectra into the different symmetry channels. Calculations of the Raman spectra used a model band structure of BiTeI shown in Fig. 4.

In Fig. 5(a), we show the collected eRS spectrum in the symmetric channel for temperatures ranging from 10 K to 300 K around Raman shifts near the collective mode energy of ≈ 74 meV. In Fig. 5(b), we plot the calculated temperature dependence of the eRS for the XX polarization for the following choice of parameters that are within the ranges reported in the literature: $\alpha_R = 3.82$ eV \AA [15,18,56–58], $\lambda = 0.35$ eV \AA , $m_1 = 0.14 m_0$ [18,57], $m_3 = 0.91 m_0$ [61], $\epsilon_\infty = 11.9 \epsilon_0$ [62].

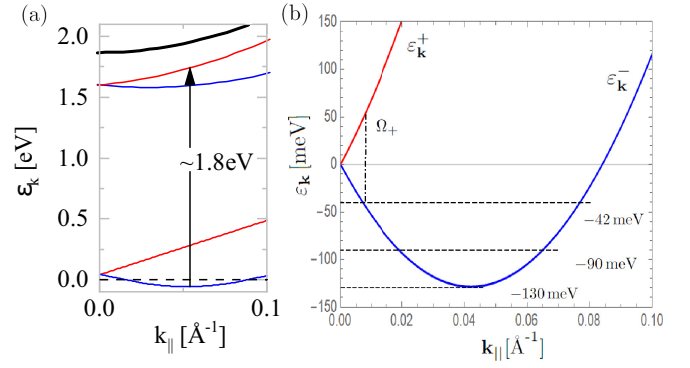


FIG. 4. (a) Low-energy (where the chemical potential [dashed line] lies) and high-energy bands of BiTeI plotted against k_{\parallel} at $k_{\perp} = \pi/c$. The colors represent the spin-split bands and the black band at ≈ 2 eV represents a spin-degenerate band. (b) A cartoon of the band structure of BiTeI with the zero fixed at the Dirac point. The chemical potential is varied between -90 meV to -42 meV in the samples we investigated [32].

This leaves us with two additional parameters for the above plot. One is the chemical potential μ , which has been chosen to be ≈ -42 meV [the minimum of band is at $\epsilon_{\text{min}}^- \approx -130$ meV, see Fig. 4(b)]. All energy measurements are rela-

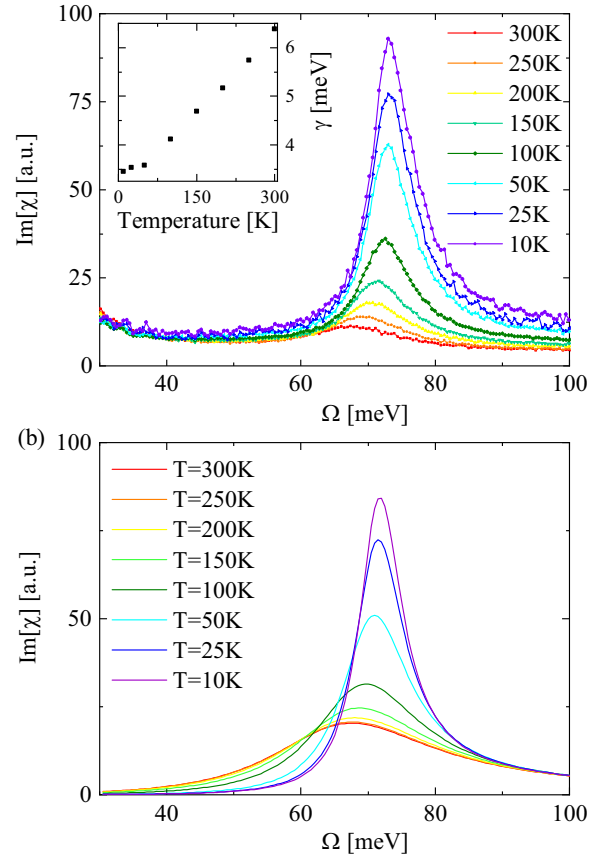


FIG. 5. (a) The temperature dependence of the resonant eRS spectrum of the collective mode in the fully symmetric A_1 channel of BiTeI. The inset shows the corresponding half width at half maxima. (b) The theoretical calculation of eRS for the parameters mentioned in the text.

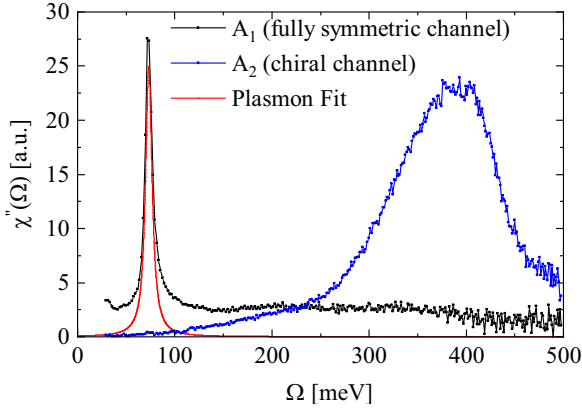


FIG. 6. The response at 10 K over extended Raman shifts in the A_1 (fully symmetric) and A_2 (chiral) irrep decompositions of the collected eRS spectrum. The red line is the calculation from Fig. 5(b). The A_2 response shows a spin-flip continuum starting from ≈ 100 meV, while the A_1 response shows a much weaker feature in that frequency range.

tive to the Dirac point in the spin-split spectrum. This choice of μ leads to an onset of spin flip transitions at $\Omega_+ \approx 90$ meV consistent with the result in Ref. [9]. The other parameter is the damping rate, which includes contributions from iodine vacancies, the nearby Rashba spin-flip continuum of excitations, finite temperature, and electron-electron interactions such as in Refs. [63–65]. To model these, we simply introduce a scattering rate that reproduced the mode width of 7 meV at 10 K and the plots at other temperatures evolved accordingly.

We saw in Eq. (25) that due to the λ parameter, we expect to also pick up contribution from the continuum of spin-flip excitations, albeit weakly. In Fig. 6, we show the eRS in the symmetric channel [which is the A_1 irreducible representation (irrep) of the point group C_{3v} relevant for BiTeI] over extended range of frequencies. The continuum is evident by a weak feature in the black line in Fig. 6. For comparison, we also show the continuum obtained in the A_2 irrep of C_{3v} , which couples directly to spin-flip excitations and hence shows a much stronger response. The onset of the spin-flip continuum in the channels belonging to different irreps is at the same value, but the spectral weight of the continuum in the response in A_2 is controlled predominantly by α_R , whereas that in A_1 is controlled by λ and hence is smaller. From the A_2 response we deduce that the onset of the spin-flip continuum is ≈ 100 meV. Having presented the theory we now understand that this coupling of charge fluctuations to the spin ones is due to Rashba SOC. However, since the plasmon is a fully symmetric excitation of the system, it does not show up in the A_2 channel. We also present the calculated plasmon response in the XX polarization in red [taken from Fig. 5(b)], where we see that the spectral weight of the plasmon completely overshadows the already weak continuum. We had explicitly demonstrated in Sec. V that this coupling is not present in the cross-polarization setup, i.e., $\mathcal{F}^{XY} = 0$, consistent with the data in the A_2 channel.

To verify that we are indeed observing the plasmon peak, we tracked the collective mode frequency for different samples grown from the same batch with light variations in the

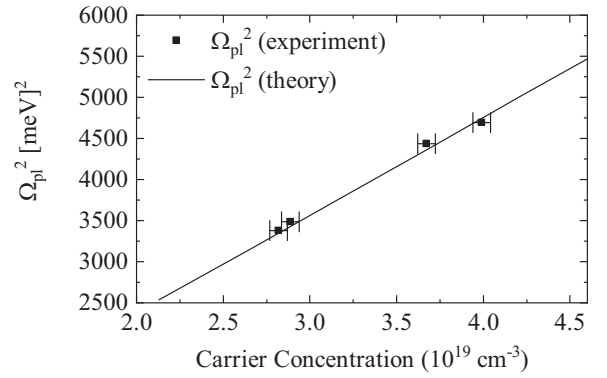


FIG. 7. The collective mode peak position (Ω_{pl}) vs carrier concentration deduced from Hall measurements for various samples [32] (data points) and the theoretical plot of the same for the parameters given in the text. The theoretical number density was estimated from the Fermi surface volume. The linear scaling of Ω_{pl}^2 with the carrier density suggests that it is a plasmon.

chemical potential induced by changes in iodine deficiencies. In Fig. 7, we plot the square of the energy of the electronic collective mode as a function of carrier concentration (as determined from Hall measurements [32]) and compare it against the calculation of the renormalized plasma frequency obtained from Eq. (26) for number densities determined as below (we restrict ourselves to the case where the chemical potential only lies in the lower Rashba sub-band),

$$\begin{aligned} n_{\text{BiTeI}} &= \int_{-\infty}^{\mu} dE \int_{\mathbf{k}} \delta(E - \varepsilon_{\mathbf{k}}^-) \\ &= n_0 m_1^* \sqrt{m_3^*} f_n(\tilde{\alpha}, \tilde{\lambda}), \end{aligned} \quad (27)$$

with

$$f_n(\tilde{\alpha}, \tilde{\lambda}) \equiv \frac{3}{2} \int_{-\infty}^{\tilde{\mu}} d\tilde{E} \int_0^{\infty} d\tilde{k}_3 \int_0^{\infty} dy \delta(\tilde{E} - \tilde{\varepsilon}_{\sqrt{y}, \tilde{k}_3}^-), \quad (28)$$

where $\tilde{\mu} \equiv \mu/|\mu|$, $\tilde{E} \equiv E/|\mu|$, $\tilde{k}_i \equiv \hbar k_i / \sqrt{2m_i|\mu|}$ and y enters after setting $k_{\parallel}^2 = y$. The function $f_n(\tilde{\alpha}, \tilde{\lambda})$ is such that $f_n(0, 0) = 1$. To remain below the Dirac point of BiTeI, we need to set $\tilde{\mu} = -1$. From Eqs. (26) and (27) the slope of Ω_{pl}^2 vs n_{BiTeI} is

$$s \equiv \frac{e^2}{m_3 \epsilon_{\infty}} \frac{f_{\text{pl}}(\tilde{\alpha}, \tilde{\lambda})}{f_n(\tilde{\alpha}, \tilde{\lambda})}.$$

With the above choice of parameters, the theoretical estimate aligns well with the experiment. It is interesting to note that experimentally we deduce the number density from the Hall measurements whereas theoretically, we calculate the number density from the Fermi-surface volume. The observed agreement suggests that the effective one-band Hall-effect theory still seems to apply insofar as to determine the carrier concentration. This justifies the Hall analysis done in Ref. [32].

To validate the requirement of resonance for the spin-mediated coupling to plasmons, we changed the incident photon energy from 1.55–2.60 eV to scan above and below the resonance with the spin-split bands shown in Fig. 4. We plot the resonant Raman profile in Fig. 8 against the optical absorption data. The latter captures all the dipole-allowed

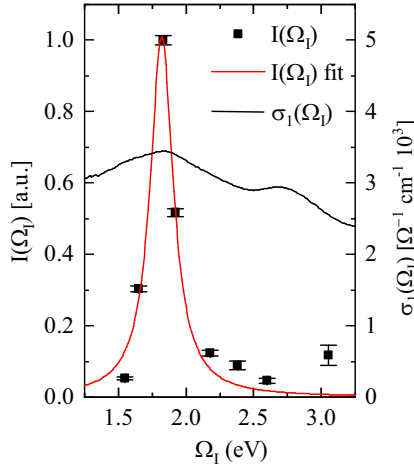


FIG. 8. Resonant Raman excitation profile of the collective mode $[I(\Omega_I)]$ as a function of the incident laser energy showing sharp feature at 1.83 eV. The absorption spectrum $\sigma_1(\Omega)$ is also plotted to demonstrate that there are indeed dipole active transitions at that excitation energy.

transitions indicating the presence of interband transitions. We see that there is a broad absorption peak around 1.7–2.2 eV and then another around 2.7–3 eV. The first one corresponds to the gap shown in Fig. 4, which has spin-split bands at ≈ 1.8 eV and a spin-degenerate band ≈ 2.1 eV. Observe that the excitation profile only resonates with the bands around 1.8 eV. Since the excitation profile is obtained by integrating across the collective mode spectral weight, we can infer that the collective mode is only excited while the resonance is between the spin-split bands and not the spin-degenerate ones. This is relevant because, as stated earlier, we need SOC in the transition elements for our effect to materialize. And such a coupling ensures SOC splitting in *both* bands. The lack of such a splitting for the band at ≈ 2.1 eV [in Fig. 4(a)] suggests a lack of such coupling and hence lack of weight in the excitation profile. This in turn suggests that the polarization factor \mathcal{F}_0^{XX} for the band at $\Omega_I = 1.83$ eV is likely finite, while that at $\Omega_I = 2.1$ eV is negligible.

B. Symmetry considerations of the observed *c*-axis plasmon in BiTeI

Being a polar self-doped metal, the planar properties of BiTeI are dictated by the C_{3v} point group. A unique property of C_{nv} groups is that the dipole field z belongs to the A_1 representation, along with the usual quadrupolar fields $x^2 + y^2, z^2$. The quadrupolar form is another way of seeing why any collective excitation in the A_1 channel only ends up coupling quadratically in q . The presence of the z -dipole term in A_1 allows for the q_3 to be present. Since plasmon is a longitudinal mode, this would imply that the excitation would correspond to the *c*-axis plasmon.³

³We inform the reader of our decision to simultaneously use 3-, z , and c . The first one is useful for theoretical modeling, the second is the convention reported in the character tables and the third is prevalent

While the possibility of the *c*-axis plasmon showing up is evident, it still remains to be explained why SOC is necessary. The way this plays out is this: Without SOC, the only dipolar field (z) in the system could arise from finite q_3 . For the A_1 excitation, the plasmon mode appears as q_3^2 as we need the result to be comprised of a z^2 contribution. Of course, this physics was reflected in the calculations that took the form of $q^2 \cdot q^2 / q^2$, where the numerators are from the polarization-charge bubbles and the denominator is from the Coulomb interaction renormalization. In the presence of SOC of the type λ , the dipole field (z) is provided by λ itself resulting in a λ^2 contribution in the A_1 response. In the calculations this appeared as $\lambda q \cdot \lambda q / q^2$ leading to the λ^2 coupling. Note that in both cases, you need to break inversion for the plasmon to couple to the Raman vertex. In the former case the inversion breaking was due to finite q_3 , but in the presence of SOC, this was intrinsically present due to λ .

C. SOC induced Drude weight renormalization along the *c*-axis

It is instructive to theoretically investigate the behavior of the slope s introduced above. Observe that in the absence of SOC, since $f_{pl} = 1 = f_n$, s measures $1/m_3\epsilon_\infty$ irrespective of the carrier concentration and the in-plane masses. Further, if $\lambda = 0$, s remains unaffected even when we change α_R . This is demonstrated in Fig. 9(a). This does not mean that the Ω_{pl} and n are unaffected, it simply means that they change in a proportional manner as shown by the constant slope. However, in the presence of λ , the slope evolves and this is plotted in Fig. 9(b).

The meaning behind the sensitivity of s to a SOC parameter can be understood from the following. Consider a spinless free electron gas (with possibly different masses along the ab axis and c axis). Its conductivity in the α th direction is given by $\sigma_{\alpha\alpha}(\Omega) = \frac{n_0 e^2}{m_\alpha} \frac{\tau}{1 - i\Omega\tau}$, where τ is the scattering lifetime and m_α is the mass in the α th direction. Let us now define two quantities: (i) $\mathcal{D} \equiv \lim_{\Omega\tau \rightarrow \infty} \Omega \text{Im}[\sigma(\Omega)]$, which is $n_0 e^2 / m_\alpha$; and (ii) $W \equiv \frac{2}{\pi} \int_0^\infty d\Omega \text{Re}[\sigma(\Omega)]$, which also evaluates to $n_0 e^2 / m_\alpha$. The latter is nothing but the optical sum rule. The relation between the integrated optical weight W and n_0, m_α remains the same irrespective of the details of the system (even if we include spin and even SOC) The quantity \mathcal{D} , however, represents the spectral weight carried by the free carriers in the system and is referred to as the Drude weight. In systems without SOC, $\mathcal{D} = W$ and is thus protected by the optical sum rule value preventing any renormalization. This protection remains valid even in the presence of spin splitting induced by a Zeeman field. However, in the presence of spin splitting due to SOC, while the relation involving W still holds (irrespective of how n is renormalized), the relation involving \mathcal{D} no longer holds and in fact is renormalized downward, meaning that the Drude weight decreases. The lost spectral weight is recovered at higher energies pertaining to the interband spin-flip transitions, thereby restoring the sum rule. This is reminiscent of the spectral weight rearrangement in a BCS superconductor from the δ peak in optical conductivity to the finite-energy bump

lent in crystallography. We choose them according to the relevant context, as is standard practice.

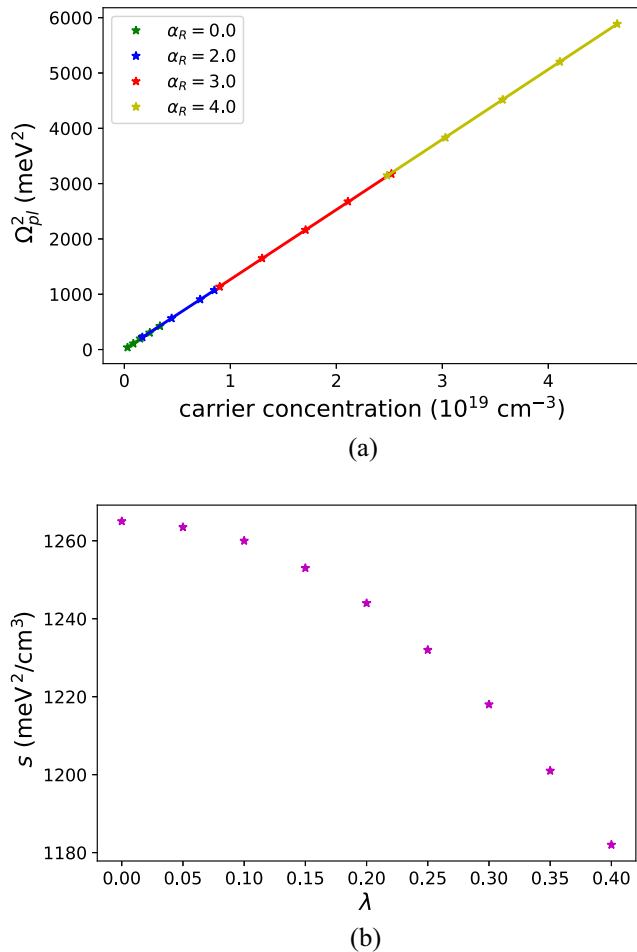


FIG. 9. (a) The square of the plasmon energy Ω_{pl}^2 is plotted against n_{BiTeI} for different values of α_R and $\lambda = 0$. The fixed slope for different α_R demonstrates a lack of Drude-weight renormalization. (b) Sensitivity of the slope s to λ (in units of eV Å). The reduction of the slope indicates downward renormalization of the Drude weight.

at twice the gap energy, as we increase disorder [66] and has been reported as the “color change” effect [67,68].

Since the slope s is derived from the plasma frequency, which is a response from the free carriers, s actually tracks the Drude weight \mathcal{D} . Further, since $s \sim \mathcal{D}/n$ and $W \sim n$, s could also be seen as a measure of \mathcal{D}/W . This should be constant without SOC and should dip downwards in the presence of SOC. In 2D Rashba systems [18,69], this is exactly what happens: SOC drains the spectral weight away from \mathcal{D} to higher energies, while still satisfying the sum rule. However, observe from Fig. 9(a) that s is invariant under changes to the Rashba parameter α_R when $\lambda = 0$. This can be understood from the result in Ref. [18], which demonstrated that the c -axis plasmon is not renormalized by SOC in the absence of λ . As soon as we introduce λ [see Fig. 9(b)], we observe that s becomes sensitive to the SOC constant indicating that the Drude weight is renormalized (downwards). The lost spectral weight from the Drude weight is recovered in the continuum introduced due to λ . This is consistent with the observation that our observed c -axis plasmon is damped from the continuum, the presence of which would necessarily imply a Drude

weight renormalization. While our modeling and interpretation imply renormalization along the c axis of BiTeI, verifying this directly is beyond the scope of an eRS experiment.

VII. CONCLUSIONS

It has long been known from group theory considerations that the eRS in the fully symmetric channel can couple to the plasmon excitations. But even in resonant eRS, it does so only at finite momentum transfer q and the spectral weight of the plasmon scales as q^2 . In this paper we have demonstrated that in resonant eRS of light in systems with broken inversion, the presence of Rashba-type SOC induces an additional coupling to plasmons at zero-momentum transfer. This happens due to two effects: first, as a high-energy effect, the fully symmetric channel develops an effective spin vertex, which is characterized by a polarization factor $\mathcal{F}_i^{\beta\alpha}$ that couples to the spin-charge susceptibility of the system; second, as a low-energy effect, the spin-charge susceptibility (Π_{i0}) in the system itself becomes nonzero due to Rashba type SOC. The development of the effective spin vertex happens under resonant conditions and requires Rashba coupling of the spins involved in the resonant interband transitions. This effect is not present if the spins are merely split in the two bands. The resulting coupling to plasmons ($\mathcal{F}_i \Pi_{i0}$) is particularly significant in giant Rashba systems such as BiTeI.

While the emergence of spin-charge susceptibilities was known, showing the presence of the effective spin vertex in the fully symmetric channel is the main contribution of this paper. We developed the general theoretical formalism to capture this effect and explicitly demonstrated the existence of this vertex in a toy Dirac system with SOC where we had access to the wavefunctions of all the relevant states in the system. For BiTeI, although we did not have access to the wavefunction of the intermediate states, we demonstrated all the necessary ingredients to be present to observe this effect. We then verified these ideas by performing resonant eRS on various BiTeI samples demonstrating a charge sector collective mode that scales just like a plasmon with respect to the carrier density. Assuming a simple toy model for BiTeI conduction band we were also able to reproduce the temperature dependence and the laser excitation profile dependence as seen in the experiment. Within the context of BiTeI, we also argued for the presence of an out-of-plane canting of spins that is ultimately responsible for the observation of the plasmon mode in eRS under normal incidence. We additionally predicted a Drude-weight renormalization along the c axis of BiTeI.

This high-energy effect of SOC, which primarily arises due to the breaking of SU(2) invariance in the spin-flip interband transitions, not only allows plasmons to couple to optical probes but also opens up a route to explore the interplay of spin and charge degrees of freedom in semiconductors. In fact, it is quite possible that the observation of many charge collective modes, which are expected to be weighted by q^2 but still seen through Raman could be so due to the enhancement from the SOC in such materials. Apart from BiTeI, this can be relevant to other C_{nv} systems like transition-metal dichalcogenides and multilayer graphene. This coupling channel should be further investigated to study its effect on nonequilibrium ultrafast spectroscopies, an emerging field, as

it provides an additional mechanism for energy relaxation for charge and spin excitations alike. This coupling channel could also be exploited to use charge-driven collective modes to manipulate spin. Finally, it is encouraging to note that the spectroscopic landscape has expanded tremendously due to considerations of effects of spin, which has been largely underappreciated for a long time. We believe that the physics we outline here will be relevant to studying the interplay of spin-orbit coupled electrons either with chiral phonons, which can also couple to eRS [70], with excitations in strongly SOC Floquet driven systems [71], and to other metallic topologically nontrivial systems that often require strong SOC.

ACKNOWLEDGMENTS

We thank Dmitrii Maslov for the discussions. The work at Concordia University (S.S. and S.M.) was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Grant No. RGPIN-2019-05486. S.S. was also supported by the Horizon Postdoctoral fellowship from Concordia University. The spectroscopic work conducted at Rutgers (A.L. and G.B.) was supported by the NSF Grant No. DMR-2105001. The work at NICPB (G.B.) was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant Agreement No. 885413.

APPENDIX A: DISCUSSION ON THE RAMAN PROCESSES AND THEIR CONTRIBUTION TO THE eRS SPECTRUM

a. The eRS response from the direct term. In Eq. (3) if we only keep the direct term, the final response $R_d(\Omega, \mathbf{q})$ can be obtained from the analytic continuation of

$$\chi_d(Q) = - \int_K \text{Tr}[\hat{\gamma}_d \hat{G}_K \hat{\gamma}_d \hat{G}_{K+Q}]. \quad (\text{A1})$$

At this stage, the hat represents matrices in the full Hilbert space of states at the Fermi surface as well as those in the intermediate states. The analytic continuation technique has been detailed in Ref. [50] (although it was done for a one-band system, the technique outlined there still applies to a multi-band system, provided the prescription to couple the vector potential is appropriately identified). The frequency sum in $\int_K GG$ leads to the Lindhard form $\frac{n(\epsilon_{\mathbf{k}+\mathbf{q}}) - n(\epsilon_{\mathbf{k}})}{i\Omega_m + \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}}$, which restricts the $\int_{\mathbf{k}}$ to states along the Fermi surface (at $\mathbf{q} \rightarrow 0$). Thus $\chi_d(Q)$ only picks up the contributions from the states at the Fermi surface.

b. The eRS response from the mixed terms. A mixed term comprises one vertex depicting the direct process and the other the indirect process. The term of the type $\sum_m \frac{[j_\beta^S]_{nm} [j_\alpha^I]_{m'n'}}{\Omega_I - E_\alpha - i\Gamma}$ in the vertex is not conducive to a field theory treatment. However, this term can be seen as an on-shell contribution of the propagator \hat{G} of the system (defined in the full Hilbert space). This allows the following generalization:

$$\sum_m [j_\beta^S]_{nm} \left[\frac{1}{\epsilon_n + \Omega_I - \epsilon_m - i\Gamma} \right]_m [j_\alpha^I]_{m'n'} \\ = \sum_{mm'} [j_\beta^S]_{nm} \left[\frac{\delta_{mm'}}{\epsilon_n + \Omega_I - \epsilon_m - i\Gamma} \right] [j_\alpha^I]_{m'n'} \quad (\text{A2})$$

$$\rightarrow [j_\beta^S \hat{G}(\epsilon_n + \Omega_I) j_\alpha^I]_{mm'}. \quad (\text{A3})$$

Here we have used that $\hat{G}(Q) = \sum_m \frac{|m\rangle\langle m|}{i\omega - \epsilon_m - i\Gamma_m}$ such that on-shell condition (which would lead to a pole) would be met when ω (translated to real frequencies) is $\sim \epsilon_m$. Thus, in the presence of an external laser field from Ω_I , the on-shell contribution selectively comes from the intermediate states. Even if the intermediate states are incoherent or absent, there still remains the off-shell contribution that can produce a finite response from the mixed terms. In common parlance, these are referred to as “virtual” processes. The choice of Ω_I is what ultimately picks, which type (on- or off-shell) virtual process would contribute the most. This promotion of the vertex to the Green's function is what ultimately allows one to draw the diagrams as in Fig. 1. Thus, the response $R_m(\Omega, \mathbf{q})$ is obtained from analytic continuation of terms such as

$$\chi_m^{(1)}(Q) = \int_K \text{Tr}[\hat{\gamma}_d \hat{G}_{K+Q} j^I \hat{G}_{K+Q+Q_I} j^S \hat{G}_K]. \quad (\text{A4})$$

Here $Q_I \equiv (\Omega_m^I, \mathbf{q}_I)$, where Ω_m^I is a Matsubara frequency that would be continued to the incoming laser frequency Ω_I and \mathbf{q}_I is the momentum of the incoming photon. The superscript (1) denotes that this is just one of the terms. Other terms arise from interchanging the vertices with direct and indirect processes, and also interchanging the direct process containing Ω_I with that containing Ω_S [where we would get $j_\alpha^I \hat{G}(\epsilon_n - \Omega_S) j_\beta^S$]. We do not list all the terms here as we do not need them in this paper. It is important to emphasize that the above calculation can be carried out in any basis as it is a trace, but then $\hat{\mathbf{j}}$ would have to be rotated to that basis from the original basis where $\hat{\mathbf{j}} = \frac{\partial \hat{H}}{\partial \mathbf{A}}$.

c. The eRS response from indirect terms. Like in mixed terms, the two vertices corresponding to the indirect processes are again composed of the on-shell and off-shell contributions. Near resonance, picking the on-shell contributions within the same approximations as above, $R_{id}(\Omega, \mathbf{q})$ can be computed from terms like

$$\chi_{id}^{(1)}(Q) = \int_K \text{Tr}[j^S \hat{G}_{K+Q} j^I \hat{G}_{K+Q+Q_I} j^S \hat{G}_K j^I \hat{G}_{K-Q_I}]. \quad (\text{A5})$$

APPENDIX B: HILBERT SPACE FACTORING OF THE RAMAN VERTEX DUE TO RESONANCE

Although we are only interested in the resonant terms, the Coulomb renormalization happens through the term $\Pi_{\alpha\beta;0}$, as mentioned in the main text, and this mimics the structure of the mixed terms with $\hat{\gamma}_d \rightarrow 1$. It will be instructive to understand the evaluation of one of the mixed terms to see how the condition of resonance allows us to factor the Hilbert space into those involving the Fermi surface states and those involving intermediate states. In particular, consider the evaluation of

$$L(Q) = \int_K \text{Tr}[j^I \hat{G}_{K+Q+Q_I} j^S \hat{G}_K \hat{G}_{K+Q}]. \quad (\text{B1})$$

Let us now specify the system in some orbital basis (with components a, b where the vector potential would be coupled) as

$$H = \begin{pmatrix} \mathcal{H}_{aa} & \mathcal{H}_{ab} \\ \mathcal{H}_{ba} & \mathcal{H}_{bb} \end{pmatrix}, \quad (\text{B2})$$

where each of the component above has (2×2) internal spin structure. We can now introduce an inversion-breaking SOC (whose energy scale can be denoted by E_{SOC}). For convenience, let us switch to a basis (which we can call the c and v basis) that would be diagonal in the absence of SOC and look like

$$H = \begin{pmatrix} \mathcal{H}_{cc} & 0 \\ 0 & \mathcal{H}_{vv} \end{pmatrix}. \quad (\text{B3})$$

Upon adding SOC, this matrix would acquire off-diagonal entries $\mathcal{H}_{cv}, \mathcal{H}_{vc} \sim \mathcal{O}(E_{\text{SOC}}/E_g)$, where E_g is the gap between the c states and the v states. We could then split Green's functions (in the cv basis) into

$$\hat{G}_K = \begin{pmatrix} \mathcal{G}^{cc} & 0 \\ 0 & \mathcal{G}^{vv} \end{pmatrix} + \mathcal{O}\left(\frac{E_{\text{SOC}}}{E_g}\right). \quad (\text{B4})$$

The transition elements $j^{S,I}$ can also be written in the cv basis as

$$\hat{j}^x = \begin{pmatrix} \mathcal{J}_{cc}^x & \mathcal{J}_{cv}^x \\ \mathcal{J}_{vc}^x & \mathcal{J}_{vv}^x \end{pmatrix}, \quad x \in \{I, S\}.$$

Plugging these forms into Eq. (B1) and evaluating the trace we get

$$\begin{aligned} L(Q) &= \int_K [\mathcal{J}_{cc}^I \mathcal{G}_{K+Q,Q}^{cc} \mathcal{J}_{cc}^S] \mathcal{G}_K^{cc} \mathcal{G}_{K+Q}^{cc} \\ &+ \int_K [\mathcal{J}_{cv}^I \mathcal{G}_{K+Q,Q}^{vv} \mathcal{J}_{vc}^S] \mathcal{G}_K^{cc} \mathcal{G}_{K+Q}^{cc} \\ &+ \int_K [\mathcal{J}_{vc}^I \mathcal{G}_{K+Q,Q}^{cc} \mathcal{J}_{cv}^S] \mathcal{G}_K^{vv} \mathcal{G}_{K+Q}^{vv} \\ &+ \int_K [\mathcal{J}_{vv}^I \mathcal{G}_{K+Q,Q}^{vv} \mathcal{J}_{vv}^S] \mathcal{G}_K^{vv} \mathcal{G}_{K+Q}^{vv}. \end{aligned} \quad (\text{B5})$$

To proceed further, we can also assume a general structure for the 2×2 components of $(G)_K^{ii}$ as

$$\mathcal{G}_K^{ii} = \sum_{s=\pm} \frac{1}{i\omega_m - \varepsilon_k^s} [\mathcal{M}_k^i]_{2 \times 2}, \quad i \in \{c, v\}. \quad (\text{B6})$$

The $s = \pm$ denotes the two flavors resulting from the spin degree of freedom. The exact form of \mathcal{M}_k^i is explicitly dependent on the system of interest and the type of the perturbation. Performing the fermionic Matsubara sums we get

$$\begin{aligned} L(Q) &= \sum_{s,s',s''=\pm} \int_{\vec{k}} \frac{[\mathcal{J}_{cc}^I \mathcal{M}_{\vec{k}'}^{c,s''} \mathcal{J}_{cc}^S \mathcal{M}_{\vec{k}}^{c,s} \mathcal{M}_{\vec{k}+\vec{q}}^{c,s'}]}{i\Omega_m + \varepsilon_{\vec{k}}^{c,s} - \varepsilon_{\vec{k}+\vec{q}}^{c,s'}} \left[\frac{n_F(\varepsilon_{\vec{k}}^{c,s}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''})}{i\Omega_m + i\Omega_I + \varepsilon_{\vec{k}}^{c,s} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''}} - \frac{n_F(\varepsilon_{\vec{k}+\vec{q}}^{c,s'}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''})}{i\Omega_I + \varepsilon_{\vec{k}+\vec{q}}^{c,s'} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''}} \right] \\ &+ \int_{\vec{k}} \frac{[\mathcal{J}_{cv}^I \mathcal{M}_{\vec{k}'}^{v,s''} \mathcal{J}_{vc}^S \mathcal{M}_{\vec{k}}^{c,s} \mathcal{M}_{\vec{k}+\vec{q}}^{c,s'}]}{i\Omega_m + \varepsilon_{\vec{k}}^{c,s} - \varepsilon_{\vec{k}+\vec{q}}^{c,s'}} \left[\frac{n_F(\varepsilon_{\vec{k}}^{c,s}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''})}{i\Omega_m + i\Omega_I + \varepsilon_{\vec{k}}^{c,s} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''}} - \frac{n_F(\varepsilon_{\vec{k}+\vec{q}}^{c,s'}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''})}{i\Omega_I + \varepsilon_{\vec{k}+\vec{q}}^{c,s'} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''}} \right] \\ &+ \int_{\vec{k}} \frac{[\mathcal{J}_{vc}^I \mathcal{M}_{\vec{k}'}^{c,s''} \mathcal{J}_{cv}^S \mathcal{M}_{\vec{k}}^{v,s} \mathcal{M}_{\vec{k}+\vec{q}}^{v,s'}]}{i\Omega_m + \varepsilon_{\vec{k}}^{v,s} - \varepsilon_{\vec{k}+\vec{q}}^{v,s'}} \left[\frac{n_F(\varepsilon_{\vec{k}}^{v,s}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''})}{i\Omega_m + i\Omega_I + \varepsilon_{\vec{k}}^{v,s} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''}} - \frac{n_F(\varepsilon_{\vec{k}+\vec{q}}^{v,s'}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''})}{i\Omega_I + \varepsilon_{\vec{k}+\vec{q}}^{v,s'} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''}} \right] \\ &+ \int_{\vec{k}} \frac{[\mathcal{J}_{vv}^I \mathcal{M}_{\vec{k}'}^{v,s''} \mathcal{J}_{vv}^S \mathcal{M}_{\vec{k}}^{v,s} \mathcal{M}_{\vec{k}+\vec{q}}^{v,s'}]}{i\Omega_m + \varepsilon_{\vec{k}}^{v,s} - \varepsilon_{\vec{k}+\vec{q}}^{v,s'}} \left[\frac{n_F(\varepsilon_{\vec{k}}^{v,s}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''})}{i\Omega_m + i\Omega_I + \varepsilon_{\vec{k}}^{v,s} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''}} - \frac{n_F(\varepsilon_{\vec{k}+\vec{q}}^{v,s'}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''})}{i\Omega_I + \varepsilon_{\vec{k}+\vec{q}}^{v,s'} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''}} \right]. \end{aligned} \quad (\text{B7})$$

Here $n_F(\varepsilon)$ is the Fermi function. For definiteness, let us now choose the chemical potential (μ) to lie in the v band. We will also assume $\varepsilon_{\vec{k}}^c - \varepsilon_{\vec{k}}^v \sim E_g > 0$. Since SOC is a perturbation, we will have $\varepsilon_{\vec{k}}^{c,s} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{c,s''} \sim E_{\text{SOC}}$, and $\varepsilon_{\vec{k}}^{v,s} - \varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''} \sim E_{\text{SOC}}$. Resonance condition would imply that $\Omega_I \sim E_g$, and we will be interested in Raman shifts $\Omega \sim E_{\text{SOC}}$. With the various variables tuned to this regime, the fourth term in Eq. (B7) gives

$$\begin{aligned} L(Q)|_{4^{\text{th}}\text{term}} &= \sum_{s,s',s''=\pm} \int_{\vec{k}} \frac{[\mathcal{J}_{vv}^I \mathcal{M}_{\vec{k}'}^{v,s''} \mathcal{J}_{vv}^S \mathcal{M}_{\vec{k}}^{v,s} \mathcal{M}_{\vec{k}+\vec{q}}^{v,s'}]}{i\Omega_m + \varepsilon_{\vec{k}}^{v,s} - \varepsilon_{\vec{k}+\vec{q}}^{v,s'}} \left[\frac{n_F(\varepsilon_{\vec{k}}^{v,s}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''})}{i\Omega_I} - \frac{n_F(\varepsilon_{\vec{k}+\vec{q}}^{v,s'}) - n_F(\varepsilon_{\vec{k}+\vec{q}+\vec{q}}^{v,s''})}{i\Omega_I} \right] + \mathcal{O}\left(\frac{E_{\text{SOC}}}{E_g}\right) \\ &= \sum_{s,s',s''=\pm} \int_{\vec{k}} \frac{[\mathcal{J}_{vv}^I \mathcal{M}_{\vec{k}'}^{v,s''} \mathcal{J}_{vv}^S \mathcal{M}_{\vec{k}}^{v,s} \mathcal{M}_{\vec{k}+\vec{q}}^{v,s'}]}{i\Omega_I} \left[\frac{n_F(\varepsilon_{\vec{k}}^{v,s}) - n_F(\varepsilon_{\vec{k}+\vec{q}}^{v,s'})}{i\Omega_m + \varepsilon_{\vec{k}}^{v,s} - \varepsilon_{\vec{k}+\vec{q}}^{v,s'}} \right] + \mathcal{O}\left(\frac{E_{\text{SOC}}}{E_g}\right). \end{aligned} \quad (\text{B8})$$

Notice the Linhard term in the square braces in the last line of Eq. (B8). Because μ is assumed to lie in the v band, this term would contribute to the final result due to finite weight from the ph excitations. The first term for $L(Q)$ in Eq. (B7) would also result in a similar expression but with $v \leftrightarrow c$. But the $n_F \rightarrow 0$ for the c bands as those are empty (since $\varepsilon_c > \varepsilon_v$). These terms of Eq. (B8), as is evident, are not resonantly enhanced: they only have a factor of Ω_I in the denominator. Following similar steps

for the second and third terms of $L(Q)$, we are led to

$$L(Q)|_{2+3} = \sum_{s,s',s''=\pm} \int_{\bar{k}} \frac{[\mathcal{J}_{cv}^I \mathcal{M}_{\bar{k}'}^{v s''} \mathcal{J}_{vc}^S \mathcal{M}_{\bar{k}}^{c s} \mathcal{M}_{\bar{k}+\bar{q}}^{c s'}]}{i\Omega_m + \varepsilon_{\bar{k}}^{c s} - \varepsilon_{\bar{k}+\bar{q}}^{c s'}} \left[\frac{n_F(\varepsilon_{\bar{k}}^{c s}) - n_F(\varepsilon_{\bar{k}+\bar{q}+\bar{q}}^{v s''})}{i\Omega_l + E_g} - \frac{n_F(\varepsilon_{\bar{k}+\bar{q}}^{c s'}) - n_F(\varepsilon_{\bar{k}+\bar{q}+\bar{q}}^{v s''})}{i\Omega_l + E_g} \right] \\ + \int_{\bar{k}} \frac{[\mathcal{J}_{vc}^I \mathcal{M}_{\bar{k}'}^{c s''} \mathcal{J}_{cv}^S \mathcal{M}_{\bar{k}}^{v s} \mathcal{M}_{\bar{k}+\bar{q}}^{v s'}]}{i\Omega_m + \varepsilon_{\bar{k}}^{v s} - \varepsilon_{\bar{k}+\bar{q}}^{v s'}} \left[\frac{n_F(\varepsilon_{\bar{k}}^{v s}) - n_F(\varepsilon_{\bar{k}+\bar{q}+\bar{q}}^{c s''})}{i\Omega_m + i\Omega_l - E_g} - \frac{n_F(\varepsilon_{\bar{k}+\bar{q}}^{v s'}) - n_F(\varepsilon_{\bar{k}+\bar{q}+\bar{q}}^{c s''})}{i\Omega_l - E_g} \right] + \mathcal{O}\left(\frac{E_{\text{SOC}}}{E_g}\right). \quad (\text{B9})$$

Close to the resonance condition, the dominant contribution to the Raman bubble arises from the second term in Eq. (B9). Thus,

$$L(Q) \approx \sum_{s,s',s''=\pm} \int_{\bar{k}} \frac{[\mathcal{J}_{vc}^I \mathcal{M}_{\bar{k}'}^{c s''} \mathcal{J}_{cv}^S \mathcal{M}_{\bar{k}}^{v s} \mathcal{M}_{\bar{k}+\bar{q}}^{v s'}]}{i\Omega_m + \varepsilon_{\bar{k}}^{v s} - \varepsilon_{\bar{k}+\bar{q}}^{v s'}} \left[\frac{n_F(\varepsilon_{\bar{k}}^{v s}) - n_F(\varepsilon_{\bar{k}+\bar{q}+\bar{q}}^{c s''})}{i\Omega_m + i\Omega_l - E_g} - \frac{n_F(\varepsilon_{\bar{k}+\bar{q}}^{v s'}) - n_F(\varepsilon_{\bar{k}+\bar{q}+\bar{q}}^{c s''})}{i\Omega_l - E_g} \right]. \quad (\text{B10})$$

We expect the intermediate states to be broad and thus have a lifetime $\sim 1/\Gamma$. If we are in a regime such that $E_{\text{SOC}} \ll \Gamma$, we can further ignore the Ω_m in the denominator above to get (after continuing Ω_l to real frequencies)

$$L(Q) \approx \sum_{s,s',s''=\pm} \int_{\bar{k}} \frac{[\mathcal{J}_{vc}^I \mathcal{M}_{\bar{k}'}^{c s''} \mathcal{J}_{cv}^S \mathcal{M}_{\bar{k}}^{v s} \mathcal{M}_{\bar{k}+\bar{q}}^{v s'}]}{\Omega_l - E_g + i\Gamma} \left[\frac{n_F(\varepsilon_{\bar{k}}^{v s}) - n_F(\varepsilon_{\bar{k}+\bar{q}}^{v s'})}{i\Omega_m + \varepsilon_{\bar{k}}^{v s} - \varepsilon_{\bar{k}+\bar{q}}^{v s'}} \right]. \quad (\text{B11})$$

The above expression implies that under the resonant condition, the bubble involving a direct and an indirect vertex reduces to a form containing the Linhard factor created out of states near the Fermi level, the resonance enhancement factor, and the precise form associated with the transition from the states near the Fermi level to the intermediate state and back (\mathcal{J}_{vc} and \mathcal{J}_{cv}). The main thing to note is that the original Hilbert space involving both the c and v states is factored into a particle-hole contribution from the subspace of states near the Fermi surface and a transition factor associated with c - v transitions weighted by the resonance factor. This factorization allows us to decouple the light-excitation processes (of the order of the laser frequency) from the quasiparticle excitations (of the order of the Raman shifts). There are other terms of the order of $\mathcal{O}(E_{\text{SOC}}/E_g)$ and $\mathcal{O}(E_{\text{SOC}}/\Gamma)$, which are ignored here, but they will only provide corrections to the physical effects captured by the resonant enhancement term. It should be stated that the form is well expected and has been known since the study of resonant Raman in semiconductors. However, the transition element was only discussed for III-V semiconductors where it was modelled as a constant, which is usually referred to as the Kane parameter [45]. In the above derivation, we have access to the precise form of these transition elements.

Finally, we note that the above form was obtained for a form of the Green's function as in Eq. (B4). In general, we would have the block-off-diagonal terms \mathcal{G}_{cv} to be nonzero (they are induced by SOC after all). An explicit evaluation of $L(Q)$ in such a case would involve 32 terms. Four of the relevant ones are accounted for above, and the other terms look like

$$T_1 : \int_K [\mathcal{J}_{vv}^I \mathcal{G}_{K+Q,Q}^{vc} \mathcal{J}_{cv}^S] \mathcal{G}_K^{vv} \mathcal{G}_{K+Q}^{vv}, \\ T_2 : \int_K [\mathcal{J}_{vv}^I \mathcal{G}_{K+Q,Q}^{vv} \mathcal{J}_{vc}^S] \mathcal{G}_K^{cv} \mathcal{G}_{K+Q}^{vv}. \quad (\text{B12})$$

In the T_1 family of terms, the low-energy quasiparticle response comes from well-defined poles $\mathcal{G}_K^{vv} \mathcal{G}_{K+Q}^{vv}$. But note that there is no $c \rightarrow v$ transition due to the incoming light as \mathcal{J}^I is associated with the vv index. It is the Green's function that propagates an electron from one band to the other. This adversely affects the resonance enhancement that we were able to get with the other four terms we considered above. Further, a T_1 type of term that is allowed in general due to the existence of \mathcal{G}_{cv} term, however, is also suppressed by the gap E_g between the c and the v bands. In the T_2 type of terms the low-energy quasiparticle response is already incoherent due to the $\mathcal{G}_K^{cv} \mathcal{G}_{K+Q}^{vv}$ structure. Thus, near the resonance, only the four terms that we considered are the relevant processes. Even though not all the four terms contributed in the Eq. (B11), they would do so if the resonance was considered with bands below the Fermi level. But the T_1 and T_2 type of terms still would not contribute.

APPENDIX C: COMPUTATION OF POLARIZATION-CHARGE BUBBLE

As is evident from Eq. (8), to couple the plasmon pole to the Raman response, we need, what we call the polarization-charge bubble $\Pi_{\alpha\beta,0}$. This bubble is a trace over the structure of $L(Q)$ computed above with \mathcal{J} acquiring the indices α, β . Let us try to evaluate this expression. First, note that the s'' dependence in the Fermi functions drops out under the approximations we made above. Then, by properties of Green's functions, $\sum_{s'} \mathcal{M}_{\bar{k}}^{c s'} = 1$. This leads to the expression

$$L(Q) \approx \frac{1}{\Omega_l - E_g + i\Gamma} \sum_{s,s'} \int_{\bar{k}} [\mathcal{J}_{vc}^I \mathcal{J}_{cv}^S] [\mathcal{M}_{\bar{k}}^{v s} \mathcal{M}_{\bar{k}+\bar{q}}^{v s'}] \\ \times \left[\frac{n_F(\varepsilon_{\bar{k}}^{v s}) - n_F(\varepsilon_{\bar{k}+\bar{q}}^{v s'})}{i\Omega_m + \varepsilon_{\bar{k}}^{v s} - \varepsilon_{\bar{k}+\bar{q}}^{v s'}} \right]. \quad (\text{C1})$$

It should be kept in mind that the matrix $[\mathcal{J}_{vc}^I \mathcal{J}_{cv}^S]$ is \vec{k} -dependent and hence cannot be factored out. However, without the $\mathcal{J} \mathcal{J}$, the expression would simply be exactly what arises in the computation of the charge bubble $\int_K \text{Tr}[\hat{G}_K \hat{G}_{K+Q}]$ prior to the trace operation. In fact, since $\mathcal{J} \mathcal{J}$ is a 2×2 matrix, we can write

$$[\mathcal{J}_{\beta,vc}^I \mathcal{J}_{\alpha,cv}^S] = \mathcal{F}_i^{\beta\alpha}(\vec{k}) \hat{\sigma}_i. \quad (\text{C2})$$

This allows us to write

$$L(Q) \approx \frac{1}{\Omega_I - E_g + i\Gamma} \sum_{s,s'} \int_{\vec{k}} \mathcal{F}_i^{\beta\alpha}(\vec{k}) [\hat{\sigma}_i \mathcal{M}_{\vec{k}}^{v_s} \mathcal{M}_{\vec{k}+\vec{q}}^{v_{s'}}] \times \left[\frac{n_F(\varepsilon_{\vec{k}}^{v_s}) - n_F(\varepsilon_{\vec{k}+\vec{q}}^{v_{s'}})}{i\Omega_m + \varepsilon_{\vec{k}}^{v_s} - \varepsilon_{\vec{k}+\vec{q}}^{v_{s'}}} \right]. \quad (\text{C3})$$

If $\mathcal{F}^{\beta\alpha}(\vec{k})$ has an s -wave component denoted by $\mathcal{F}^{\beta\alpha}$, then we arrive at

$$L(Q) \approx \frac{\mathcal{F}_i^{\beta\alpha}}{\Omega_I - E_g + i\Gamma} \sum_{s,s'} \int_{\vec{k}} [\hat{\sigma}_i \mathcal{M}_{\vec{k}}^{v_s} \mathcal{M}_{\vec{k}+\vec{q}}^{v_{s'}}] \times \left[\frac{n_F(\varepsilon_{\vec{k}}^{v_s}) - n_F(\varepsilon_{\vec{k}+\vec{q}}^{v_{s'}})}{i\Omega_m + \varepsilon_{\vec{k}}^{v_s} - \varepsilon_{\vec{k}+\vec{q}}^{v_{s'}}} \right] + \text{h.h.} \\ = \frac{\mathcal{F}_i^{\beta\alpha}}{\Omega_I - E_g + i\Gamma} \int_K [\hat{\sigma}_i \hat{G}_K \hat{\sigma}_0 \hat{G}_{K+Q}] + \text{h.h.} \quad (\text{C4})$$

The h.h. stands for higher harmonics, which usually integrate to zero for most systems. Upon taking trace we get

$$\Pi_{\alpha\beta;0}(Q) \approx m_e \text{Tr}[L(Q)] \\ \approx \frac{\mathcal{F}_i^{\beta\alpha} \Pi_{i0}(Q)}{\Omega_I - E_g + i\Gamma}. \quad (\text{C5})$$

This is the form used in the main text. In the case of III-V semiconductors, $\vec{p}^\alpha \epsilon^\alpha \rightarrow \mathcal{P}\vec{\epsilon}$. Thus, following the discussion after Eq. (13) in the main text, we can write $\mathcal{F}_i^{\beta\alpha} \epsilon_i^\alpha \epsilon_S^\beta \rightarrow \mathcal{P}(\vec{\epsilon}_I \times \vec{\epsilon}_S)_i$, restoring the conventional result.

APPENDIX D: CALCULATION DETAILS FOR SOC DIRAC SYSTEM

Starting from Eq. (15), we note that $\hat{\mathbf{j}} = v_F(\tau_z \hat{\sigma}_x \hat{x} + \hat{\sigma}_y \hat{y})$. We move to the eigenbasis using the transformation

$$\hat{\mathbf{j}}^{(cv)} = M^\dagger \hat{\mathbf{j}}^{(ab)} M, \quad (\text{D1})$$

where the matrix M is formed by column juxtaposition of the eigenvectors of H . In the absence of Δ_R and Δ_Z , the transformed current becomes

$$\mathcal{J}_{x,cc}^0 = v_F \frac{v_F k}{\xi_k} \cos \theta \hat{\sigma}_0, \\ \mathcal{J}_{x,cv}^0 = v_F \left(i\tau_z \sin \theta + \frac{\Delta}{\xi_k} \cos \theta \right) \hat{\sigma}_0, \\ \mathcal{J}_{x,vv}^0 = -\mathcal{J}_{x,cc}^0, \\ \mathcal{J}_{x,vc}^0 = (\mathcal{J}_{x,cv}^0)^\dagger. \quad (\text{D2})$$

Similarly, the various components of j_y current elements are

$$\mathcal{J}_{y,cc}^0 = v_F \frac{v_F k}{\xi_k} \sin \theta \hat{\sigma}_0, \\ \mathcal{J}_{y,cv}^0 = v_F \left(-i\tau_z \cos \theta + \frac{\Delta}{\xi_k} \sin \theta \right) \hat{\sigma}_0, \\ \mathcal{J}_{y,vv}^0 = -\mathcal{J}_{y,cc}^0, \\ \mathcal{J}_{y,vc}^0 = (\mathcal{J}_{y,cv}^0)^\dagger. \quad (\text{D3})$$

Here θ is the azimuthal angle of \vec{k} . Note the presence of the interband components, $\mathcal{J}_{y,vc}$, that is responsible for optical interband transitions (as in graphene). However, there is no spin splitting in these blocks. When we include Δ_R in the Hamiltonian, the M -transformed current now becomes

$$\mathcal{J}_{x,cc} = \mathcal{J}_{x,cc}^0 + v_F \frac{\Delta_R}{2\xi_k} \left(\frac{\Delta^2}{\xi_k^2} \cos \theta \hat{\sigma}_z - \sin \theta \hat{\sigma}_y \right), \\ \mathcal{J}_{x,cv} = \mathcal{J}_{x,cv}^0 - v_F \frac{\Delta_R}{2\xi_k} \\ \times \left(\frac{v_F k \Delta}{\xi_k^2} \cos \theta \hat{\sigma}_z + \left[\frac{\Delta}{v_F k} \sin \theta - i \frac{\tau_z \xi_k}{v_F k} \cos \theta \right] \hat{\sigma}_y \right), \\ \mathcal{J}_{x,vv} = -\mathcal{J}_{x,cc}, \\ \mathcal{J}_{x,vc} = (\mathcal{J}_{x,cv})^\dagger. \quad (\text{D4})$$

Similarly, the various components of j_y current elements are

$$\mathcal{J}_{y,cc} = \mathcal{J}_{y,cc}^0 + v_F \frac{\Delta_R}{2\xi_k} \left(\frac{\Delta^2}{\xi_k^2} \sin \theta \hat{\sigma}_z + \cos \theta \hat{\sigma}_y \right), \\ \mathcal{J}_{y,cv}^0 = \mathcal{J}_{y,cv}^0 - v_F \frac{\Delta_R}{2\xi_k} \\ \times \left(\frac{v_F k \Delta}{\xi_k^2} \sin \theta \hat{\sigma}_z - \left[\frac{\Delta}{v_F k} \cos \theta + i \frac{\tau_z \xi_k}{v_F k} \sin \theta \right] \hat{\sigma}_y \right), \\ \mathcal{J}_{y,vv} = -\mathcal{J}_{y,cc}, \\ \mathcal{J}_{y,vc} = (\mathcal{J}_{y,cv})^\dagger. \quad (\text{D5})$$

Note that the $\mathcal{J}_{cv} = \mathcal{J}_{vc}^\dagger$ blocks are now spin dependent. Also, as discussed in the main text, in the prefactors of the interband \mathcal{J}'_{cv} s, the Δ'_R s get promoted to the interband δ_R . Using this form for $\hat{\mathbf{j}}$ and Eq. (13), we arrive at the following forms for $\mathcal{F}_i^{\alpha\beta}(\vec{k})$,

$$\mathcal{F}_0^{XX} = \frac{v_F^2 \Delta^2}{2\xi_k^2} + \frac{v_F^4 k^2 \sin^2 \theta}{2\xi_k^2}, \\ \mathcal{F}_0^{YY} = \frac{v_F^2 \Delta^2}{2\xi_k^2} + \frac{v_F^4 k^2 \cos^2 \theta}{2\xi_k^2}, \\ \mathcal{F}_0^{XY} = \frac{i\tau_z v_F^2 \Delta}{\xi_k} - \frac{v_F^4 k^2 \sin 2\theta}{2\xi_k^2}, \\ \mathcal{F}_1^{\beta\alpha} = 0, \\ \mathcal{F}_2^{XX} = \frac{\delta_R v_F^3 k \sin 2\theta}{2\xi_k^2}, \\ \mathcal{F}_2^{YY} = -\frac{\delta_R v_F^3 k \sin 2\theta}{2\xi_k^2},$$

$$\begin{aligned}
 \mathcal{F}_2^{XY} &= -\frac{\delta_R v_F^3 k \cos 2\theta}{2\xi_k^2}, \\
 \mathcal{F}_3^{XX} &= -\frac{\delta_R v_F^3 k \Delta^2 \cos^2 \theta}{\xi_k^4}, \\
 \mathcal{F}_3^{YY} &= -\frac{\delta_R v_F^3 k \Delta^2 \sin^2 \theta}{\xi_k^4}, \\
 \mathcal{F}_3^{XY} &= -\frac{i\tau_z \delta_R v_F^3 k \Delta}{2\xi_k^3} - \frac{\delta_R v_F^3 k \Delta^2 \sin 2\theta}{2\xi_k^4}. \quad (\text{D6})
 \end{aligned}$$

In the main text, we only keep the s -wave component of the above terms. The other angle-dependent terms evaluate to zero in the bubble of interest and are expected to be small even when nonresonant terms are accounted for.

When we consider only the valley Zeeman term, the blocks of the Hamiltonian takes the form

$$\begin{aligned}
 H^{cc} &= \hat{s}_0 \xi_k + \frac{\Delta_Z}{2} \hat{s}_z, & H^{vv} &= -\hat{s}_0 \xi_k + \frac{\Delta_Z}{2} \hat{s}_z, \\
 H^{cv} &= 0, & H^{vc} &= 0. \quad (\text{D7})
 \end{aligned}$$

Since there is no \vec{k} -dependent correction, the expression for the current remains the same as Eqs. (D2) and (D3) where $\mathcal{J}_{cv} = 0 = \mathcal{J}_{vc}$. Likewise $\mathcal{F}_0^{\alpha\beta}$ remains the same as in Eq. (D6) and $\mathcal{F}_i^{\alpha\beta} = 0$ for $i = 1, 2, 3$.

APPENDIX E: CALCULATION DETAILS FOR BiTeI

In this section we present calculation details of $\Pi_{i0}(Q)$ for $i \in \{0, 1, 2, 3\}$. First, we observe that all the integrals involved in the computation of Π_{i0} are convergent in \mathbf{k} integration. We shall perform a change of variable of $\vec{k} \rightarrow \vec{k} - \vec{q}/2$. This makes the calculation a lot simpler. For brevity, let us define

$$L_{s,s'}(\vec{k}, \vec{q}) \equiv \frac{n_F(\varepsilon_{\vec{k}-\vec{q}/2}^s) - n_F(\varepsilon_{\vec{k}+\vec{q}/2}^{s'})}{i\Omega_m + \varepsilon_{\vec{k}-\vec{q}/2}^s - \varepsilon_{\vec{k}+\vec{q}/2}^{s'}}. \quad (\text{E1})$$

Then we have

$$\Pi_{i0}(Q) = \sum_{ss'} \int_{\vec{k}} \mathcal{N}_{s,s'}^i(\vec{k}, \vec{q}) L_{s,s'}(\vec{k}, \vec{q}). \quad (\text{E2})$$

1. Out of plane $\Pi_{i0}(Q)$

Since we are interested in small q_3 , to leading orders in it we have

$$\begin{aligned}
 L_{s,s}(\vec{k}, \vec{q}) &= -\partial_{\varepsilon_k^s} n_F(\varepsilon_k^s) \left[\frac{\partial_{k_3} \varepsilon_k^s q_3}{i\Omega_m} - \frac{(\partial_{k_3} \varepsilon_k^s)^2 q_3^2}{\Omega_m^2} \right] + \mathcal{O}(q_3^3), \\
 L_{s,s'}(\vec{k}, \vec{q}) &= \left[\frac{n_F(\varepsilon_k^s) - n_F(\varepsilon_k^{s'})}{i\Omega_m + (s-s')D_k} \right] + \mathcal{O}(q_3). \quad (\text{E3})
 \end{aligned}$$

We have retained different orders in different terms so as to ensure that we get the appropriate leading order terms in the full calculation of Π_{i0} . Next we calculate the various $\mathcal{N}_{s,s'}^i$

factors to leading order in q_3 ,

$$\begin{aligned}
 \mathcal{N}_{++}^0 &= 1 + \mathcal{O}(q_3^2), \\
 \mathcal{N}_{--}^0 &= \mathcal{N}_{++}^0, \\
 \mathcal{N}_{+-}^0 &= \frac{\alpha^2 k_{\parallel}^2 \lambda^2 q_3^2}{4D_k^4} + \mathcal{O}(q_3^3), \\
 \mathcal{N}_{-+}^0 &= \mathcal{N}_{+-}^0, \\
 \mathcal{N}_{++}^1 &= \frac{\alpha k_2}{D_k} - i \frac{\alpha k_1 \lambda q_3}{2(D_k)^2} + \mathcal{O}(q_3^2), \\
 \mathcal{N}_{--}^1 &= -(\mathcal{N}_{++}^1)^*, \\
 \mathcal{N}_{+-}^1 &= \frac{i\alpha k_1 (D_k) + \alpha \lambda k_2 k_3}{2(D_k)^3} \lambda q_3 + \mathcal{O}(q_3^3), \\
 \mathcal{N}_{-+}^1 &= -(\mathcal{N}_{+-}^1)^*, \\
 \mathcal{N}_{ss'}^2 &= \mathcal{N}_{ss'}^1 (k_1 \leftrightarrow k_2), \\
 \mathcal{N}_{++}^3 &= \frac{\lambda k_3}{(D_k)} + \mathcal{O}(q_3^2), \\
 \mathcal{N}_{--}^3 &= -\mathcal{N}_{++}^3, \\
 \mathcal{N}_{+-}^3 &= -\frac{\alpha^2 k_{\parallel}^2}{2(D_k)^3} \lambda q_3 + \mathcal{O}(q_3^3), \\
 \mathcal{N}_{-+}^3 &= -\mathcal{N}_{+-}^3. \quad (\text{E4})
 \end{aligned}$$

To calculate $\Pi_{00}(Q)$, the leading order contribution is $\mathcal{O}(q_3^2)$ and it is given by (below we have assumed that our chemical potential is in the lower band $\varepsilon_{\vec{k}}^-$)

$$\begin{aligned}
 \Pi_{00}(Q) &= q_3^2 \int_{\vec{k}} [-\partial_{\varepsilon^-} n_F(\varepsilon^-)] \frac{(\partial_{k_3} \varepsilon^-)^2}{-\Omega_m^2} \\
 &+ \lambda^2 q_3^2 \int_{\vec{k}} \frac{[n_F(\varepsilon^+) - n_F(\varepsilon^-)]}{D_k^3} \frac{\alpha^2 k_{\parallel}^2}{\Omega_m^2 + 4D_k^2}. \quad (\text{E5})
 \end{aligned}$$

Notice that this is like the single-band contribution plus a SOC-induced interband correction. Notice also that this correction would be absent if $\lambda = 0$. To calculate $\Pi_{10}(Q)$ observe that the inter- and intraband terms are odd functions of k_1 and k_2 . The L_{ss} function only has a term that is odd in k_3 . Thus, $\Pi_{10}(Q)$ vanishes for $\vec{q} = (0, 0, q_3)$. Since $\Pi_{20}(Q)$ simply interchanges $k_1 \leftrightarrow k_2$, it also vanishes. Finally, to calculate $\Pi_{30}(Q)$ we only need contribution to $\mathcal{O}(q_3)$. Using the $\mathcal{N}_{ss'}$ and the $L_{ss'}$ functions above, we arrive at

$$\begin{aligned}
 \Pi_{30}(Q) &= -\lambda q_3 \int_{\vec{k}} \frac{[-\partial_{\varepsilon^-} n_F(\varepsilon^-)]}{D_k} \frac{k_3 \partial_{k_3} \varepsilon^-}{i\Omega_m} \\
 &+ \alpha^2 \lambda q_3 \int_{\vec{k}} \frac{[n_F(\varepsilon^+) - n_F(\varepsilon^-)]}{D_k^3} \frac{i\Omega_m k_{\parallel}^2}{\Omega_m^2 + 4D_k^2}. \quad (\text{E6})
 \end{aligned}$$

2. In plane $\Pi_{i0}(\mathbf{Q})$

For completeness, we can also look at small in-plane momentum transfer q_1 . To leading orders in q_1 we have

$$L_{s,s}(\vec{k}, \vec{q}) = -\partial_{\varepsilon_k^s} n_F(\varepsilon_k^s) \left[\frac{\partial_{k_1} \varepsilon_k^s q_1}{i\Omega_m} - \frac{(\partial_{k_1} \varepsilon_k^s)^2 q_1^2}{\Omega_m^2} \right] + \mathcal{O}(q_1^3),$$

$$L_{s,s'}(\vec{k}, \vec{q}) = \left[\frac{n_F(\varepsilon_k^s) - n_F(\varepsilon_k^{s'})}{i\Omega_m + (s - s')D_k} \right] + \mathcal{O}(q_1). \quad (\text{E7})$$

Next we calculate the various $\mathcal{N}_{ss'}^i$ factors to leading order in q_1 ,

$$\begin{aligned} \mathcal{N}_{++}^0 &= 1 + \mathcal{O}(q_1^2), \\ \mathcal{N}_{--}^0 &= \mathcal{N}_{++}^0, \\ \mathcal{N}_{+-}^0 &= \frac{\alpha^2(\alpha^2 k_2^2 + \lambda^2 k_3^2) q_1^2}{4D_k^4} + \mathcal{O}(q_1^3), \\ \mathcal{N}_{-+}^0 &= \mathcal{N}_{+-}^0, \\ \mathcal{N}_{++}^1 &= \frac{\alpha k_2}{D_k} + i \frac{\alpha k_3 \lambda q_1}{2(D_k)^2} + \mathcal{O}(q_1^2), \\ \mathcal{N}_{--}^1 &= -(\mathcal{N}_{++}^1)^*, \\ \mathcal{N}_{+-}^1 &= \frac{-i\alpha \lambda k_3 (D_k) + \alpha^3 k_1 k_2}{2(D_k)^3} q_1 + \mathcal{O}(q_1^3), \\ \mathcal{N}_{-+}^1 &= -(\mathcal{N}_{+-}^1)^*, \\ \mathcal{N}_{++}^2 &= -\frac{\alpha k_1}{D_k} + \mathcal{O}(q_1^2), \\ \mathcal{N}_{--}^2 &= -\mathcal{N}_{++}^2, \\ \mathcal{N}_{+-}^2 &= \frac{\alpha(\alpha^2 k_2^2 + \lambda^2 k_3^2) q_1}{2D_k^3} + \mathcal{O}(q_1^3), \\ \mathcal{N}_{-+}^2 &= -\mathcal{N}_{+-}^2, \\ \mathcal{N}_{++}^3 &= \frac{\lambda k_3}{(D_k)} - i \frac{\alpha^2 k_2 q_1}{2(D_k)^2} + \mathcal{O}(q_1^2), \\ \mathcal{N}_{--}^3 &= -(\mathcal{N}_{++}^3)^*, \\ \mathcal{N}_{+-}^3 &= \frac{\alpha^2}{2} \left(\frac{\lambda k_1 k_3}{2(D_k)^3} + \frac{ik_2}{2(D_k)^2} \right) q_1 + \mathcal{O}(q_1^3), \\ \mathcal{N}_{-+}^3 &= -(\mathcal{N}_{+-}^3)^*. \end{aligned} \quad (\text{E8})$$

These can be used to calculate $\Pi_{00}(\mathbf{Q})$ to the leading order contribution to $\mathcal{O}(q_1^2)$ and it is given by

$$\begin{aligned} \Pi_{00}(\mathbf{Q}) &= q_1^2 \int_{\vec{k}} [-\partial_{\varepsilon^-} n_F(\varepsilon^-)] \frac{(\partial_{k_1} \varepsilon^-)^2}{-\Omega_m^2} \\ &\quad + \alpha^2 q_1^2 \int_{\vec{k}} \frac{[n_F(\varepsilon^+) - n_F(\varepsilon^-)]}{D_k^3} \frac{\alpha^2 k_2^2 + \lambda^2 k_3^2}{\Omega_m^2 + 4D_k^2}. \end{aligned} \quad (\text{E9})$$

For $\Pi_{10}(\mathbf{Q})$ and $\Pi_{30}(\mathbf{Q})$ observe that the inter- and intra-band terms are odd functions of k_2 and k_3 . The L_{ss} function only has a term that is odd in k_1 . Thus, they vanish for $\vec{q} = (q_1, 0, 0)$. Finally, to calculate $\Pi_{20}(\mathbf{Q})$ we only need contribution to $\mathcal{O}(q_1)$. Using the $\mathcal{N}_{ss'}$ and the $L_{ss'}$ functions above, we arrive at

$$\begin{aligned} \Pi_{20}(\mathbf{Q}) &= -\alpha q_1 \int_{\vec{k}} \frac{[-\partial_{\varepsilon^-} n_F(\varepsilon^-)]}{D_k} \frac{k_1 \partial_{k_1} \varepsilon^-}{i\Omega_m} \\ &\quad + \alpha q_1 \int_{\vec{k}} \frac{[n_F(\varepsilon^+) - n_F(\varepsilon^-)]}{D_k^3} \frac{i\Omega_m(\alpha^2 k_2^2 + \lambda^2 k_3^2)}{\Omega_m^2 + 4D_k^2}. \end{aligned} \quad (\text{E10})$$

As expected, the charge-charge susceptibility has a correction $\propto \alpha^2$ and the spin-charge susceptibility is $\propto \alpha$. In this case we do not need any canting effect (from the λ term).

APPENDIX F: DERIVATION OF PLASMA FREQUENCY

We need to evaluate $\Pi_{00}(\mathbf{q}, \Omega + i0^+)$. It is instructive to rescale the integration variables and integrands to dimensionless variables as such, $\tilde{\varepsilon} \equiv \varepsilon/|\mu|$, $\tilde{k}_i \equiv k_i/\sqrt{2m_i|\mu|}$ ($i \in \{1, 2, 3\}$), $\tilde{T} = T/|\mu|$, $\tilde{\alpha} \equiv \alpha_R \sqrt{2m_1/|\mu|}$, and $\tilde{\lambda} \equiv \lambda \sqrt{2m_3/|\mu|}$, where μ is the chemical potential relative to the Dirac point. This leads to

$$V_{\mathbf{q}} \Pi_{00}(\mathbf{q}, \Omega) = \frac{e^2}{\epsilon_{\infty}} \frac{\sqrt{(2m_1|\mu|)(2m_1|\mu|)(2m_3|\mu|)}}{(2m_q|\mu|)} \frac{1}{\mu} \left[\frac{1}{\tilde{q}^2} \tilde{\Pi}_{00} \right],$$

where m_q is the effective mass along the direction of \mathbf{q} , and $\tilde{\Pi}_{00}$ is the same as Π_{00} [Eq. (E5)], but expressed in terms of dimensionless variable $\tilde{\varepsilon}$, \tilde{T} , $\tilde{\alpha}$, and $\tilde{\lambda}$. We can rewrite the above expression as

$$V_{\mathbf{q}} \Pi_{00} = \frac{\Omega_0^2}{\Omega^2} \frac{\sqrt{m_1^* m_1^* m_3^*}}{m_q^*} \underbrace{\left[3\pi^2 \frac{\tilde{\Omega}^2}{\tilde{q}^2} \tilde{\Pi}_{00} \right]}_{f_{\text{pl}}(\tilde{\alpha}, \tilde{\lambda})},$$

where $m_i^* = m_i/m_0$ is the effective mass. Since in our problem, \mathbf{q} is along the 3-axis, we have $m_q^* = m_3^*$. Solving for $V_{\mathbf{q}} \Pi_{00} = 1$ we arrive at the Eq. (26).

- [1] G. Abstreiter, M. Cardona, and A. Pinczuk, Light scattering by free carrier excitations in semiconductors, in *Light Scattering in Solids IV: Electronics Scattering, Spin Effects, SERS, and Morphic Effects*, edited by M. Cardona and G. Güntherodt (Springer, Berlin, 1984), pp. 5–150.
- [2] W. Hayes and R. Loudon, *Scattering of Light by Crystals*, Dover Books on Physics (Courier Corporation, New York, 2012).

- [3] G. Placzek, *The Rayleigh and Raman Scattering* (Lawrence Radiation Laboratory, Berkeley, 1959), Vol. 526.
- [4] B. H. Bairamov, I. P. Ipatova, and V. A. Voitenko, *Phys. Rep.* **229**, 221 (1993).
- [5] E. Ivchenko, *Optical Spectroscopy of Semiconductor Nanostructures* (Springer, New York, 2004).
- [6] A. Pinczuk, L. Brillson, E. Burstein, and E. Anastassakis, *Phys. Rev. Lett.* **27**, 317 (1971).

- [7] A. Pinczuk, S. Schmitt-Rink, G. Danan, J. P. Valladares, L. N. Pfeiffer, and K. W. West, *Phys. Rev. Lett.* **63**, 1633 (1989).
- [8] A. R. Goñi, A. Pinczuk, J. S. Weiner, J. M. Calleja, B. S. Dennis, L. N. Pfeiffer, and K. W. West, *Phys. Rev. Lett.* **67**, 3298 (1991).
- [9] A. C. Lee, B. Peng, K. Du, H.-H. Kung, B. Monserrat, S. W. Cheong, C. J. Won, and G. Blumberg, *Phys. Rev. B* **105**, L161105 (2022).
- [10] F. Perez, *Phys. Rev. B* **79**, 045306 (2009).
- [11] F. Baboux, F. Perez, C. A. Ullrich, I. D'Amico, G. Karczewski, and T. Wojtowicz, *Phys. Rev. B* **87**, 121303(R) (2013).
- [12] F. Baboux, F. Perez, C. A. Ullrich, G. Karczewski, and T. Wojtowicz, *Phys. Rev. B* **92**, 125307 (2015).
- [13] F. Perez, F. Baboux, C. A. Ullrich, I. D'Amico, G. Vignale, G. Karczewski, and T. Wojtowicz, *Phys. Rev. Lett.* **117**, 137204 (2016).
- [14] H.-H. Kung, S. Maiti, X. Wang, S.-W. Cheong, D. L. Maslov, and G. Blumberg, *Phys. Rev. Lett.* **119**, 136802 (2017).
- [15] K. Ishizaka, M. S. Bahramy, H. Murakawa, M. Sakano, T. Shimojima, T. Sonobe, K. Koizumi, S. Shin, H. Miyahara, A. Kimura *et al.*, *Nat. Mater.* **10**, 521 (2011).
- [16] Y. L. Chen, M. Kanou, Z. K. Liu, H. J. Zhang, J. A. Sobota, D. Leuenberger, S. K. Mo, B. Zhou, S.-L. Yang, P. S. Kirchmann *et al.*, *Nat. Phys.* **9**, 704 (2013).
- [17] M. Sakano, M. S. Bahramy, A. Katayama, T. Shimojima, H. Murakawa, Y. Kaneko, W. Malaeb, S. Shin, K. Ono, H. Kumigashira, R. Arita, N. Nagaosa, H. Y. Hwang, Y. Tokura, and K. Ishizaka, *Phys. Rev. Lett.* **110**, 107204 (2013).
- [18] S. Maiti, V. Zyuzin, and D. L. Maslov, *Phys. Rev. B* **91**, 035106 (2015).
- [19] S. Maiti, M. Imran, and D. L. Maslov, *Phys. Rev. B* **93**, 045134 (2016).
- [20] T. P. Devereaux and R. Hackl, *Rev. Mod. Phys.* **79**, 175 (2007).
- [21] L. Landau and E. Lifshitz, *Statistical Physics*, Part 2, Vol. 9 (Pergamon Press, London, 1980).
- [22] C. H. Aldrich, C. J. Pethick, and D. Pines, *Phys. Rev. Lett.* **37**, 845 (1976).
- [23] D. Bohm and D. Pines, *Phys. Rev.* **92**, 609 (1953).
- [24] P. W. Anderson, *Phys. Rev.* **130**, 439 (1963).
- [25] A. Pinczuk, B. S. Dennis, D. Heiman, C. Kallin, L. Brey, C. Tejedor, S. Schmitt-Rink, L. N. Pfeiffer, and K. W. West, *Phys. Rev. Lett.* **68**, 3623 (1992).
- [26] M. V. Klein, in *Lasers and Applications*, edited by W. O. N. Guimaraes, C.-T. Lin, and A. Mooradian (Springer, Berlin, 1981), pp. 45–54.
- [27] M. V. Klein, *Phys. Rev. B* **82**, 014507 (2010).
- [28] G. Blumberg, A. Mialitsin, B. S. Dennis, M. V. Klein, N. D. Zhigadlo, and J. Karpinski, *Phys. Rev. Lett.* **99**, 227002 (2007).
- [29] T. Cea and L. Benfatto, *Phys. Rev. B* **94**, 064512 (2016).
- [30] S. Maiti, A. V. Chubukov, and P. J. Hirschfeld, *Phys. Rev. B* **96**, 014503 (2017).
- [31] S. Das Sarma and D.-W. Wang, *Phys. Rev. Lett.* **83**, 816 (1999).
- [32] A. C. Lee, S. Sarkar, K. Du, H.-H. Kung, C. J. Won, K. Wang, S.-W. Cheong, S. Maiti, and G. Blumberg, following paper, *Phys. Rev. B* **109**, L041111 (2024).
- [33] A. V. Rodina and E. L. Ivchenko, *Phys. Rev. B* **106**, 245202 (2022).
- [34] B. S. Shastry and B. I. Shraiman, *Phys. Rev. Lett.* **65**, 1068 (1990).
- [35] B. S. Shastry and B. I. Shraiman, *Int. J. Mod. Phys. B* **05**, 365 (1991).
- [36] J. Buhot, M.-A. Méasson, Y. Gallais, M. Cazayous, A. Sacuto, G. Lapertot, and D. Aoki, *Phys. Rev. Lett.* **113**, 266405 (2014).
- [37] H.-H. Kung, R. E. Baumbach, E. D. Bauer, V. K. Thorsmølle, W.-L. Zhang, K. Haule, J. A. Mydosh, and G. Blumberg, *Science* **347**, 1339 (2015).
- [38] H.-H. Kung, S. Ran, N. Kanchanavatee, V. Krapivin, A. Lee, J. A. Mydosh, K. Haule, M. B. Maple, and G. Blumberg, *Phys. Rev. Lett.* **117**, 227601 (2016).
- [39] A. Shekhter, M. Khodas, and A. M. Finkel'stein, *Phys. Rev. B* **71**, 165329 (2005).
- [40] A. Ashrafi and D. L. Maslov, *Phys. Rev. Lett.* **109**, 227201 (2012).
- [41] S. Raghu, S. B. Chung, X.-L. Qi, and S.-C. Zhang, *Phys. Rev. Lett.* **104**, 116401 (2010).
- [42] F. Baboux, F. Perez, C. A. Ullrich, I. D'Amico, J. Gómez, and M. Bernard, *Phys. Rev. Lett.* **109**, 166401 (2012).
- [43] R. Z. Vitlina, L. I. Magarill, and A. V. Chaplik, *JETP Lett.* **95**, 253 (2012).
- [44] S. Maiti and D. L. Maslov, *Phys. Rev. B* **95**, 134425 (2017).
- [45] E. O. Kane, *J. Phys. Chem. Solids* **1**, 249 (1957).
- [46] O. Kashuba and V. I. Fal'ko, *Phys. Rev. B* **80**, 241404(R) (2009).
- [47] E. Heller, Y. Yang, L. Kocia, W. Chen, S. Fang, M. Borunda, and E. Kaxiras, *ACS Nano* **10**, 2803 (2016).
- [48] E. Ricciardi, M.-A. Méasson, M. Cazayous, A. Sacuto, and Y. Gallais, *Phys. Rev. Lett.* **116**, 066805 (2016).
- [49] M. Cardona, *Light Scattering in Solids I* (Springer-Verlag, Berlin, 1983).
- [50] A. M. Shvaika, O. Vorobyov, J. K. Freericks, and T. P. Devereaux, *Phys. Rev. B* **71**, 045120 (2005).
- [51] D. Bercioux and P. Lucignano, *Rep. Prog. Phys.* **78**, 106001 (2015).
- [52] Z. Wang, D.-K. Ki, H. Chen, H. Berger, A. H. MacDonald, and A. F. Morpurgo, *Nat. Commun.* **6**, 8339 EP (2015).
- [53] A. W. Cummings, J. H. Garcia, J. Fabian, and S. Roche, *Phys. Rev. Lett.* **119**, 206601 (2017).
- [54] J. H. Garcia, M. Vila, A. W. Cummings, and S. Roche, *Chem. Soc. Rev.* **47**, 3359 (2018).
- [55] A. Kumar, S. Maiti, and D. L. Maslov, *Phys. Rev. B* **104**, 155138 (2021).
- [56] M. Sakano, J. Miyawaki, A. Chainani, Y. Takata, T. Sonobe, T. Shimojima, M. Oura, S. Shin, M. S. Bahramy, R. Arita, N. Nagaosa, H. Murakawa, Y. Kaneko, Y. Tokura, and K. Ishizaka, *Phys. Rev. B* **86**, 085204 (2012).
- [57] D. VanGennep, S. Maiti, D. Graf, S. W. Tozer, C. Martin, H. Berger, D. L. Maslov, and J. J. Hamlin, *J. Phys.: Condens. Matter* **26**, 342202 (2014).
- [58] L. Cai, C. Yu, W. Zhao, Y. Li, H. Feng, H.-A. Zhou, L. Wang, X. Zhang, Y. Zhang, Y. Shi, J. Zhang, L. Yang, and W. Jiang, *Nano Lett.* **22**, 7441 (2022).
- [59] H. Maaß, H. Bentmann, C. Seibel, C. Tusche, S. V. Ereameev, T. R. F. Peixoto, O. E. Tereshchenko, K. A. Kokh, E. V. Chulkov, J. Kirschner, and F. Reinert, *Nat. Commun.* **7**, 11621 (2016).
- [60] M. Pletyukhov and V. Gritsev, *Phys. Rev. B* **74**, 045307 (2006).
- [61] J. S. Lee, G. A. H. Schober, M. S. Bahramy, H. Murakawa, Y. Onose, R. Arita, N. Nagaosa, and Y. Tokura, *Phys. Rev. Lett.* **107**, 117401 (2011).

- [62] S. Schwalbe, R. Wirnata, R. Starke, G. A. H. Schober, and J. Kortus, *Phys. Rev. B* **94**, 205130 (2016).
- [63] J. Gavoret, P. Nozières, B. Roulet, and M. Combescot, *J. Phys. France* **30**, 987 (1969).
- [64] P. Sharma, A. Principi, and D. L. Maslov, *Phys. Rev. B* **104**, 045142 (2021).
- [65] A. P. Goyal, P. Sharma, and D. L. Maslov, *Ann. Phys.* **456**, 169355 (2023).
- [66] H. Chen, *Phys. Rev. Lett.* **71**, 2304 (1993).
- [67] M. V. Klein and G. Blumberg, *Science* **283**, 42 (1999),.
- [68] D. Basov, S. Woods, A. Katz, E. Singley, R. Dynes, M. Xu, D. Hinks, C. Homes, and M. Strongin, *Science* **283**, 49 (1999).
- [69] A. Agarwal, S. Chesi, T. Jungwirth, J. Sinova, G. Vignale, and M. Polini, *Phys. Rev. B* **83**, 115135 (2011).
- [70] S. Parlak, S. Ghosh, and I. Garate, *Phys. Rev. B* **107**, 104308 (2023).
- [71] N. Arakawa and K. Yonemitsu, *Phys. Rev. B* **103**, L100408 (2021).