Theory of intervalley Coulomb interactions in monolayer transition-metal dichalcogenides

Hanan Dery

Phys. Rev. B **94**, 075421 — Published 15 August 2016

DOI: 10.1103/PhysRevB.94.075421

Theory of intervalley Coulomb interactions in monolayer transition-metal dichalcogenides

Hanan Dery[1,2,*]

[1]*Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14627, USA*
[2]*Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA*

Exciton optical transitions in transition-metal dichalcogenides offer unique opportunities to study rich many-body physics. Recent experiments in monolayer WSe_2 and WS_2 have shown that while the low-temperature photoluminescence from neutral excitons and three-body complexes is suppressed in the presence of elevated electron densities or strong photoexcitation, new dominant peaks emerge in the low-energy side of the spectrum. I present a theory that elucidates the nature of these optical transitions showing the role of the intervalley Coulomb interaction. After deriving a compact dynamical form for the Coulomb potential, I calculate the self-energy of electrons due to their interaction with this potential. For electrons in the upper valleys of the spin-split conduction band, the self energy includes a moderate redshift due to exchange, and most importantly, a correlation-induced virtual state in the band-gap. The latter sheds light on the origin of the luminescence in monolayer WSe_2 and WS_2 in the presence of pronounced many-body interactions.

PACS numbers: 71.45.Gm 71.10.-w 71.35.-y 78.55.-m

I. INTRODUCTION

Monolayer transition-metal dichalcogenides (ML-TMDs) have recently sparked wide interest due to their *d*-band semiconducting behavior and spin-valley coupling.[1–9] The markedly strong optical absorbance in these atomic monolayers and their compatibility with flexible substrates can enable the next generation of ultrathin photonic and optoelectronic devices.[10–18] The reduced dielectric screening of Coulomb interactions and formation of tightly bound excitons in ML-TMDs,[19–28] allow us to study many-body interactions through the behavior of excitons in a far wider range of temperatures and background plasma densities compared with typical gated semiconductor quantum wells.[29–33]

The motivation for this work comes from the observation of unique photoluminescence (PL) peaks that emerge in ML-WX_2, where X={S, Se}.[34–38] In the presence of large electron or hole densities, the PL of neutral excitons and three-body complexes decays due to screening.[29,32,39,40] However, recent PL experiments found that a new peak emerges in the low-energy side of the spectrum in ML-WX_2 when the gate-induced electron density is large.[34,35] The supplemental material includes animated evolution of the measured PL in ML-WSe_2 and ML-$MoSe_2$ when the gate voltage is continuously changed.[41] These measurements, kindly provided by Aaron M. Jones and Xiaodong Xu, show that while the many-body peak dominates the PL of ML-WSe_2 at elevated electron densities, it never appears in ML-$MoSe_2$.[41] Relative to other peaks, it shows a strong redshift when increasing the electron density.[34] You *et al.* have observed that the bi-exciton peak in strongly photoexcitated ML-WSe_2 emerges in the same spectral region.[36] Shang *et al.* reported of similar patterns in ML-WS_2.[35]

To date, there are no models that could explain why these peaks appear in the PL ML-WX_2 but not in that of ML-MoX_2,[42–44] or why they are not suppressed by screening as one would expect at elevated electron densities. The only available models deal with neutral excitons or few-body complexes.[21–28] The emergence of unique peaks in the PL of strongly photoexcited or electron-rich samples indicates the signature of many-body effects.[45] To deal with the difference between ML-MoX_2 and ML-WX_2 and the fact that this behavior is not observed in hole-rich samples, one should also consider the subtle change in their optical transitions. The photexcitation involves transitions from the top of the valence band to the lower (upper) valleys in the spin-split conduction band of ML- MoX_2 (WX_2).[46,47] Given that all other properties are similar, this subtlety is a key difference.

The main contribution of this work is the finding of particular intervalley Coulomb interactions in ML-TMDs that emerge at elevated electron densities. I show that the electron's self-energy in the upper valleys of the spin-split conduction band has a correlation-induced virtual state in the band gap, thereby affecting photoexcited excitons in ML-WX_2 but not in ML-MoX_2. As will be argued, inclusion of the intervalley Coulomb interaction provides a self-consistent explanation for the optical properties in ML-WX_2 when subjected to strong photoexcitation or elevated electron densities.

Before embarking on the theory, I emphasize that the intervalley Coulomb interaction in TMDs is not negligible compared with the intravalley one in the presence of elevated charge densities. Quantitatively, it can be seen by inspecting the ratio between intervalley and intravalley Coulomb interactions in the static screening limit,[48]

$$\frac{V_s(\mathbf{q}\rightarrow\pm\mathbf{K}_0,\omega=0)}{V_s(\mathbf{q}\rightarrow 0,\omega=0)}\approx\frac{1}{K_0}\frac{2g_v}{a_B}\,. \tag{1}$$

$\mathbf{K}_0$ is the crystal momentum connecting the $\mathbf{K}$ and

$\mathbf{K}'$ points of the hexagonal 2D Brillouin zone, where $K_0 = 4\pi/3a$ and $a \sim 0.32$ nm is the in-plane sublattice constant. $2g_v/a_B$ denotes the Thomas-Fermi screening wavenumber relevant for the intravalley interaction at elevated charge densities and low temperatures, where $g_v = 2$ is the valley degeneracy and $a_B = \hbar^2\epsilon_r/me^2$ is the electron Bohr radius.[49] Plugging typical values for the effective mass and dielectric constant in ML-TMDs provides $a_B \sim 0.5$ nm ($m = 0.5m_0$ and $\epsilon_r = 5$). Since the Bohr radius extends over very few lattice constants, the intravalley Coulomb interaction does not overwhelm the intervalley one at elevated charge densities. As important, intervalley plasmon modes are gapped in ML-TMDs due to the spin splitting of the K-point. The splitting magnitude is about 20 - 30 meV in the conduction bands of ML- $MoSe_2$, WS_2 and WSe_2.[50–53] This attribute allows one to differentiate their signature from that of the gapless intravalley 2D plasmons, $\omega_{q\to 0} \sim 0$,[29] in contrast with the case of graphene studied by Tudorovskiy and Mikhailov.[54] Below I focus on intervalley plasmons and quantify their salient signatures on the self-energy of electrons.

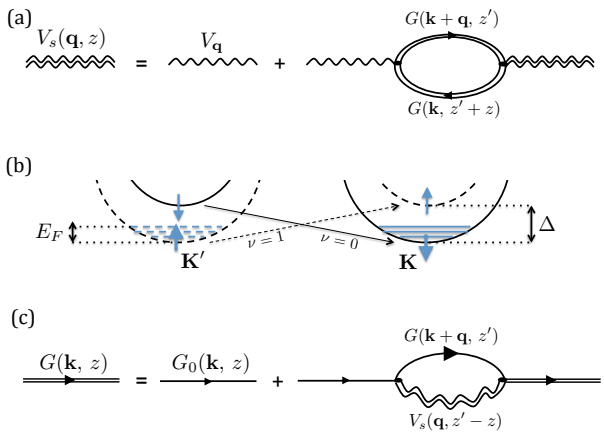

FIG. 1: (a) Dyson equation for the screened Coulomb potential. The basic polarization bubble diagram includes intravalley or intervalley processes depending on whether its two propagator lines are from the same or opposite valleys. (b) The valley diagram in the conduction band of ML-TMDs. The dashed/solid lines denote valleys populated with spin-up/down electrons. The arrows represent the spin-dependent intervalley excitations. Δ and ε_F denote the splitting and Fermi energies. (c) The electron Green's function. The screened potential in the self-energy diagram includes both intravalley and intervalley processes.

II. INTERVALLEY PLASMONS IN ML-TMDS

The plasmon modes are found from singularities in the dynamically screened Coulomb potential, $V_s(\mathbf{q},\omega) = V_\mathbf{q}/\epsilon(\mathbf{q},\omega)$, where $V_\mathbf{q}$ is the bare 2D Coulomb potential and $\epsilon(\mathbf{q},\omega)$ is the longitudinal dynamic dielectric function.[29,55,56] Figure 1(a) shows the diagram representation of the Dyson equation for $V_s(\mathbf{q},\omega)$ when using the random phase approximation and neglecting vertex corrections.[33] Intravalley or intervalley processes are represented by the basic polarization bubble when the two propagators are from the same or opposite valleys, respectively. To account for intervalley processes ($qa \sim 4\pi/3$), the well-known Lindhard formula for $\epsilon(\mathbf{q},\omega)$ in long wavelengths ($qa \ll 1$) is recast as a matrix, where plasmon modes are found from its determinant, $|\bar{\bar{\epsilon}}(\mathbf{q},\omega)| = 0$.[54] The matrix elements represent umklapp processes due to atomic-scale local fields,[57,58]

$$\epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},z) = \delta_{\mathbf{G},\mathbf{G}'} - V_{\mathbf{q}+\mathbf{G}}\sum_{\mathbf{k},\nu}\frac{f(\varepsilon_\mathbf{k}) - f(\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}+\Delta)}{(-1)^\nu z - (\Delta+\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}-\varepsilon_\mathbf{k})} \times\langle\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{G}')\mathbf{r}}|\mathbf{k}\rangle\langle\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}}|\mathbf{k}+\mathbf{q}\rangle\,. \quad (2)$$

The sum has two terms, $\nu = \{0,1\}$, coming from the two spin configurations that contribute to intervalley excitations, as shown in Fig. 1(b). $\mathbf{G}$ and $\mathbf{G}'$ are reciprocal lattice vectors, $z = \hbar\omega + i\delta$, $\bar{\mathbf{q}} = \mathbf{q} - \mathbf{K}_0$ ($\bar{q}a \ll 1$), and Δ is the K-point spin splitting. $f(\varepsilon_\mathbf{k})$ and $f(\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}+\Delta)$ are Fermi-Dirac distributions in the lower and upper valleys of the spin-split conduction band, respectively. Here, $\mathbf{k}$ and $\mathbf{k}+\bar{\mathbf{q}}$ are taken with respect to their valley edge.

Finding the plasmon energy dispersion is greatly simplified by invoking two approximations. The first one is the assumption of a rigid energy shift of the conduction band due to the electron-electron interaction, which allows one to replace the energy difference of interacting electrons in (2), $\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}-\varepsilon_\mathbf{k}$, with that of free electrons. This approximation was found to be well justified in the case of intravalley plasma excitations,[59] and it circumvents the need to *a priori* assume a zeroth-order bubble diagram (i.e, using single-line rather than double-line propagators in Fig. 1(a)). The second approximation is motivated by the dominant contribution of the d_{z^2} orbital in the conduction band,[2,60] which leads to the following estimate of the matrix elements in (2),

$$\langle\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}|\mathbf{k}\rangle \sim \frac{8 - G^2r_d^2}{\left(4+G^2r_d^2\right)^{3/2}}\,. \quad (3)$$

r_d is the effective radius of the d_{z^2} orbital in the transition-metal atom. This expression is derived from integration of $\frac{5}{16\pi}\int_0^\infty d^3\mathbf{r}e^{-2r/r_d}(3\cos\theta^2-1)^2e^{iGr\cos\phi}$. Using the facts that $r_d \sim 1\,\mathring{A}$ and $G = 4\pi\ell/\sqrt{3}a$ where ℓ is an integer, the expression has a dominant contribution at $\ell = 0$, reducing the problem of finding plasmon modes to that of $\Re e\{\epsilon_{0,0}(\mathbf{q},z)\} = 0$. Assuming parabolic energy dispersion and zero net spin polarization, the intervalley plasmon energies are (see Appendix A)

$$\hbar\omega_\mathbf{q} = \Delta + \varepsilon_{\bar{\mathbf{q}}}\left(1+3\alpha^{-1}\right) + \tfrac{1}{3}\alpha\varepsilon_F\,, \quad (4)$$

where $\alpha = (K_0a_B)^{-1}$ and ε_F is the Fermi energy. These plasmons can freely propagate when $\varepsilon_{\bar{\mathbf{q}}} < \frac{1}{9}\alpha^2\varepsilon_F$ (Appendix A). That is, they are not Landau-damped due

to electron-pair excitations in a small region of nearly perfect intervalley transitions ($\mathbf{q} \to \mathbf{K}_0$). While I have used zero-temperature analysis to derive (4), the result should remain valid for $\Delta > \varepsilon_F \gg k_B T$, at which $\hbar\omega_{\mathbf{q}} \sim \Delta$. These are often typical conditions for all but ML-MoS_2 in which the spin-splitting is minute ($\Delta \sim$ 4 meV).[61] Finally, I use the single-plasmon pole (SPP) approximation[29,59,62,63] to simplify the form of the intervalley screened potential (see Appendix B),

$$V_s(\mathbf{q}=\mathbf{K}_0+\bar{\mathbf{q}},\omega) \approx V_{\mathbf{K}_0}\left(1+\frac{f_{iv}}{(\omega+i\delta)^2-\omega_{\mathbf{q}}^2}\right), \quad (5)$$

where the residue value is $f_{iv} \approx 4\alpha\varepsilon_F\Delta/3\hbar^2$, revealing linear dependence on electron density ($f_{iv} \propto \varepsilon_F \propto n$).

III. THE ELECTRON SELF-ENERGY

The compact SPP spectral form allows one to readily identify the salient features in the electron's self energy. Performing finite-temperature Green's function analysis of the zeroth-order diagram, shown in Fig. 1(c), the self-energy due to the intervalley interaction follows

$$\Sigma(\mathbf{k},z)=-\frac{3V_{\mathbf{K}_0}}{\beta}\sum_{\bar{\mathbf{q}},z'}\left[G_0(\bar{\mathbf{q}},z')+\frac{\hbar^2 f_{iv}G_0(\bar{\mathbf{q}},z')}{(z-z')^2-(\hbar\omega_{\mathbf{q}})^2}\right], \quad (6)$$

where $\beta^{-1} = k_B T$ and G_0 is the free-electron Green's function. The first and second terms in the sum correspond to exchange and correlation, respectively. To quantify their contributions, the sum over Matsubara frequencies (z') is replaced with contour integration.[33] The exchange self energies due to the intervalley interaction in the lower and upper valleys are, respectively,

$$\Sigma_x^l(\mathbf{k}) = 0 \quad , \qquad \Sigma_x^u(\mathbf{k}) \approx -\frac{2\pi n e^2}{g_v\epsilon_r K_0}\,. \quad (7)$$

The energy in a lower valley is not renormalized due to the vanishing electron population in the opposite upper valley, $f(\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}+\Delta) \to 0$. Conversely, the nonzero energy renormalization in an upper valley comes from electron filling in the opposite lower valley. Overall, the exchange effect is small and largely wavevector-independent resulting in a rigid redshift of the upper valleys by about 1 meV per n=10^{12} cm^{-2} in the lower valleys.

Coulomb correlations in the self energy are calculated by repeating the analysis for the second term in (6). The correlation term in the lower valleys follows

$$\Sigma_c^l(\mathbf{k},z-\mu) = \frac{\alpha^2\varepsilon_F}{\pi}\frac{\hbar^2}{m}\int d^2\bar{q}\frac{g(-\omega_{\mathbf{k}-\mathbf{q}})}{z-\varepsilon_{\bar{\mathbf{q}}}-\Delta-\hbar\omega_{\mathbf{k}-\mathbf{q}}}, \quad (8)$$

and in the upper valleys,

$$\Sigma_c^u(\mathbf{k},z-\mu) = \Sigma_c^l(\mathbf{k},z+\Delta-\mu) + \frac{\alpha^2\varepsilon_F}{\pi}\frac{\hbar^2}{m}\int d^2\bar{q}\frac{2\hbar\omega_{\mathbf{k}-\mathbf{q}}f(\varepsilon_{\bar{\mathbf{q}}})}{(z-\varepsilon_{\bar{\mathbf{q}}})^2-(\hbar\omega_{\mathbf{k}-\mathbf{q}})^2}\,, \quad (9)$$

where μ is the chemical potential and $g(-\omega_{\bar{q}}) \to 1$ is the Bose-Einstein distribution ($\Delta \gg k_B T$). Σ_c^u is affected by plasmon emission, denoted by the first line in (9), and electron filling in the opposite lower valley, denoted by the second line in (9). In the following, the integration is limited to the region of free plasmon propagation, $|\mathbf{k}-\bar{\mathbf{q}}| < \alpha k_F/3$ where k_F is the Fermi wavevector. Using the fact that $\alpha^2/9 \ll 1$, the principal value of (8) is singular at $z = \varepsilon_{\mathbf{k}} + 2\Delta - \mu + i\delta$. At these energies, plasmon emission to the opposite valley is enabled. The renormalization of electron energies in the lower valley, resolved from $z = \varepsilon_{\mathbf{k}}$ in (8), is negligible. On the other hand, the energy renormalization in the upper valleys, resolved from $z = \Delta + \varepsilon_{\mathbf{k}}$ in (9), is non-negligible in the immediate vicinity of $k_0^2 = (1+2\alpha/3)k_F^2$ at which the integral is singular, as shown in Fig. 2(a). The imaginary part, shown in Fig. 2(b), corresponds to the rate of plasmon emission. It vanishes for $k < k_0$ and decays to zero from about $-\alpha^3\varepsilon_F$ for $k > k_0$. The second effect, coming from the second line in (9), yields additional logarithmic singularity. As shown in Fig. 2(c), a double resonance spectral feature emerges for states in the bottom of the upper valleys at about $z \sim \varepsilon_{\mathbf{k}} - \Delta$ (i.e., within the band-gap). It can be understood as plasmon mediated virtual transition due to electron filling in the opposite lower valley. Note that the singularities in Σ_c^u, shown in Figs. 2(a) and (c), are robust; they withstand stringent integration over the small region of free plasmon propagation.

Before discussing the implication of these results, a point on the role of the intravalley Coulomb interaction is in place. Most relevant is the dependence of the spin splitting on population of the lower valleys,

$$\Delta = \Delta_0 + \Sigma_x^u - \Sigma_0\,, \quad (10)$$

where $\Delta_0 = \Delta_{so} + \Delta_{e-h}$ is the splitting induced by spin-orbit coupling[2,60,64] and electron-hole exchange.[65–67]. Σ_x^u is the aforementioned small redshift of the upper valleys due to intervalley-induced exchange. Σ_0 is the rigid redshift of the populated lower valleys in ML-TMDs due to the intravalley Coulomb interaction.[33,68] This redshift is calculated by repeating the above procedures with using the appropriate SPP form in the long range limit:[29] replacing the residue $f_{iv} \to \omega_{\mathrm{pl},q}^2$ and pole energy $\hbar^2\omega_{\mathbf{q}}^2 \to \hbar^2\omega_{\mathrm{pl},q}^2(1+q/\kappa) + C\varepsilon_q^2$, where $\omega_{\mathrm{pl},q} \simeq \sqrt{2\pi e^2 nq/m\epsilon_r}$ is the plasma frequency, κ is the Thomas-Fermi screening wavenumber, and $C\varepsilon_q^2$ reflects the role of pair excitations ($C \sim 1$).[33,59] Schmidt-Rink and Ell found that $\Sigma_0 \simeq -C_1 E_0(a_B^2 n)^{1/3}$,[30] where E_0 is the effective Rydberg energy and C_1 is a constant that depends on the integration cutoff (see Appendix C). The dashed line in Fig. 2(d) shows this redshift when using $\Delta_0 = 20$ meV and C_1=0.6. Another important parameter is $\Delta - \varepsilon_F$, which denotes the energy spacing between the conduction edge of the upper valleys and the Fermi level in the lower valleys (see Fig. 1(b)). The symbols denote numerical results (Appendix C). As shown by the solid line in Fig. 2(d), $\Delta - \varepsilon_F$ is mostly governed by Δ at low densities due to the redshift of the lower valleys and by

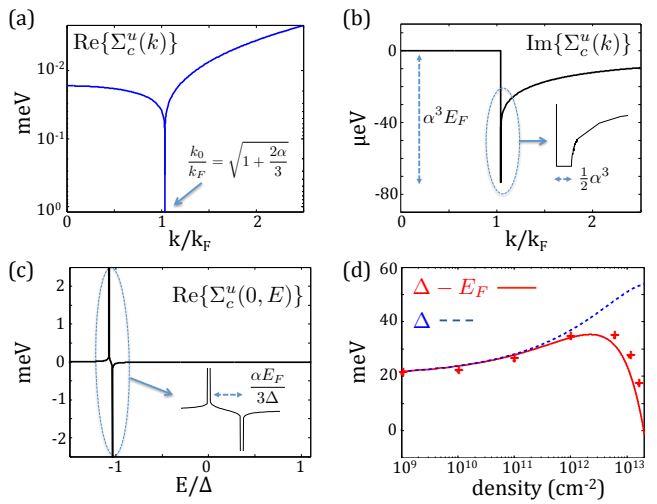

FIG. 2: (a)-(c): Effects of intervalley Coulomb correlations on the self energy of electrons in the upper valleys for n=8×10^{12} cm^{-2} in the lower valleys and α=0.115. (a)/(b) Real/Imaginary parts of the renormalization energy, $\Sigma^u_c(k) \equiv \Sigma^u_c(k, \Delta+\varepsilon_{\mathbf{k}}-\mu+\imath\delta)$. The real part is singular slightly above k_F, at which the imaginary part is finite reaching $-\alpha^3\varepsilon_F$ over a small region as shown in the inset. (c) $\Sigma^u_c(\mathbf{k}=0, E)$, where $E=0$ denotes the edge of the lower valleys. The double-resonance feature in the band gap, spaced by $\sim \alpha\varepsilon_F/3$, is located about Δ (2Δ) below the edge of the lower (upper) valleys. (d) Renormalized conduction-band splitting (Δ) and of $\Delta-\varepsilon_F$ as a function of electron density.

ε_F at large densities (i.e., competition between $n^{1/3}$ and n). Population of the upper valleys, $\Delta < \varepsilon_F$, starts at $n \sim 10^{13}$ cm^{-2}.

IV. DISCUSSION AND CONCLUSION

Putting these pieces together, the intervalley Coulomb interaction offers a self consistent explanation for recent experiments in ML-WX_2.[34,35] In these materials, bright direct excitons are formed from states in the upper valleys.[46,47] At elevated electron densities, the self energies of these states include a resonance in the band-gap due to the intervalley Coulomb correlations, as shown in Fig. 2(c). Optical transitions in ML-WX_2 can therefore be mediated by shaking up the Fermi sea via creation of intervalley plasma excitations. Importantly, the attraction to holes is not weakened by screening from the background electrons due to the shortwave and fast oscillation of these excitations ($K_0 > \kappa$ and $\Delta/\hbar > \omega_{pl,q}$). Their signature in the optical spectrum can be resolved due to the spin splitting in the conduction band, contrary to the gapless intravalley plasma excitations in quantum wells.[31] Roughly, the emerged PL peak should appear $\sim 2\Delta$ below that of neutral excitons, formed by the free states in the upper valleys. Since the redshift of the lower valleys is much stronger than that of the upper ones (Σ_0 vs Σ^u_x in (10)), the emerged peak redshifts with increasing electron density. Such behavior is not observed in ML-MoX_2 since its excitons are formed from states in the lower valleys. As can be seen from (8), the self-energy of an electron in the lower valleys has a resonance in the continuum of states at about 2Δ above the edge of the conduction band. Contrary to the band-gap resonance of electrons in the upper valleys, a resonance in the continuum of states is metastable and cannot be spectrally resolved from the self-energy of energetic free electrons with whom it is likely to experience ultrafast elastic scattering (leading to destructive interference). The net result is that the signature of intervalley plasmons do not show up in the exciton spectrum of ML-MoX_2, which similar to the case of typical semiconductor quantum wells, can only be affected by intravalley plasma excitations.

The same physical picture is offered to explain the emerged PL peaks in strongly photoexcited ML-WX_2.[35–38] Here, intervalley excitations can take place between the electron components of direct and indirect bright excitons while the holes are 'spectators'. In ML-TMDs, the exciton band structure is comprised of a direct and indirect branches.[46,64,69] The direct exciton branch is associated with electrons and holes from the same region in the Brillouin zone (e.g., both from the K-point valley), and the indirect exciton branch with pairs from opposite valleys. In addition, these branches are split to optically active (bright) and inactive (dark) excitons depending on their spin configuration. The direct-bright energy branch is located below that of the direct-dark in ML-MoX_2, and above it in ML-WX_2.[46] The indirect-bright and indirect-dark branches have the opposite order. As a result, photogenerated excitons in the direct-bright branch can undergo intervalley energy relaxation to the indirect-bright branch only in ML-WX_2. Intervalley excitations between direct and indirect bright excitons are therefore viable in strongly photo-excited ML-WX_2, supporting the emergence of the many-body resonance peak.[70] Clearly, the spin-split conduction band and the valley degree of freedom renders the physics intriguing compared with that of conventional biexciton luminescence.

In conclusion, I have presented a theory for the intervalley Coulomb interaction in monolayer transition-metal dichalogenides, finding the energy dispersion of the resulting shortwave plasmons and the effect on the self-energy of electrons. The energy dispersion is gapped due to the spin-splitting of the energy bands in the K and K' points of the Brillouin zone, and the gap energy increases due to the redshift of the lower valleys induced by the intravalley (long-range) Coulomb interaction. Importantly, states in the upper valleys are affected by intervalley Coulomb correlations through the emergence of a resonance in the band-gap at elevated electron (or exciton) densities. This result is central to the difference in the luminescence properties of tungsten-based and molybdenum-based compounds. In addition to providing a self-consistent picture to explain experi-

mental findings, the presented theory should lead to further investigations. These include: (1) studying excitons using the Bethe-Salpeter equation in which the electron Green's function is dressed by the intervalley Coulomb interactions. (2) Studying spin selective intervalley interactions by polarizing the background electrons via optical valley orientation or by application of a large magnetic field in the regime where the Zeeman energy is larger than the thermal energy. (3) Studying plasmon-phonon coupling, making use of the fact that Δ increases with population of the lower valleys. This increase can lead to resonance crossing with zone-edge phonons, providing alternative explanation for the observed splitting of the charged exciton in tungsten-based compounds,[74] which so far was attributed to electron-hole exchange.

Appendix A: Intervalley plasmon energy dispersion

As mentioned in the discussion of (4), the finding of plasmon dispersion reduces to that of $\Re e\{\epsilon_{0,0}(\mathbf{q},z)\}=0$. Figure 3 shows the six possible intervalley transitions from the K' valley to the K valley for $\mathbf{K_0}=K_0\hat{x}$. Only two of the six transitions, $i'\rightarrow i$, are limited to the first Brillouin zone ($i=\{1,6\}$), while the other four are umklapp processes (i.e., $G\neq 0$ for $i=\{2-5\}$). Thus, in transforming the sum in the dielectric function (2), one should factor the sum by 1/3. The reduced plasmon equation is then written as,

$$\epsilon_{0,0}(\mathbf{q},\omega) \;=\; 0 \;=\; 1-\tfrac{1}{3}V_{K_0}\sum_{\mathbf{k}} f_{\mathbf{k}} \qquad \text{(A1)}$$

$$\left[\frac{1}{\hbar\omega_{\mathbf{q}}-(\Delta+\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}-\varepsilon_{\mathbf{k}})}-\frac{1}{\hbar\omega_{\mathbf{q}}+(\Delta+\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}-\varepsilon_{\mathbf{k}})}\right],$$

where $\mathbf{q}=\mathbf{K}_0+\bar{\mathbf{q}}$ and $V_{K_0}=2\pi e^2/A\epsilon_r K_0$. The first and second terms in square brackets imply that if $\hbar\omega_{\mathbf{q}}$ is a plasmon mode then $-\hbar\omega_{\mathbf{q}}$ should be one as well. Given that the plasmon modes should appear close to Δ, only one of the two terms in the sum will provide a dominant contribution. I continue with the first term and assume zero temperature and no net spin polarization ($f_{\mathbf{k}}=1$ for $k<k_F$). Considering parabolic dispersion, the sum can be transformed into the following implicit equation,

$$\begin{aligned} 1 \;&=\; \frac{\alpha}{6\pi}\int_0^{\varepsilon_F} d\varepsilon \int_0^{2\pi} d\theta \frac{1}{\varepsilon_\omega+2\sqrt{\varepsilon\varepsilon_{\bar{q}}}\cos\theta} \\ &=\; \frac{\alpha}{6\varepsilon_{\bar{q}}}\left[\varepsilon_\omega-\sqrt{\varepsilon_\omega^2-4\varepsilon_F\varepsilon_{\bar{q}}}\right], \end{aligned} \qquad \text{(A2)}$$

where $\alpha=3ame^2/4\pi\hbar^2\epsilon_r$ and $\varepsilon_\omega=\hbar\omega_{\mathbf{q}}-\varepsilon_{\bar{q}}-\Delta$. The solution of this equation yields that intervalley plasmons can propagate freely (i.e., ε_ω has real values) when

$$\varepsilon_{\bar{q}} < \frac{\alpha^2}{9}\varepsilon_F\,, \qquad \text{(A3)}$$

in which the plasmon energy dispersion follows

$$\hbar\omega_{\mathbf{q}}=\Delta+\varepsilon_{\bar{\mathbf{q}}}\left(1+3\alpha^{-1}\right)+\tfrac{1}{3}\alpha\varepsilon_F\,. \qquad \text{(A4)}$$

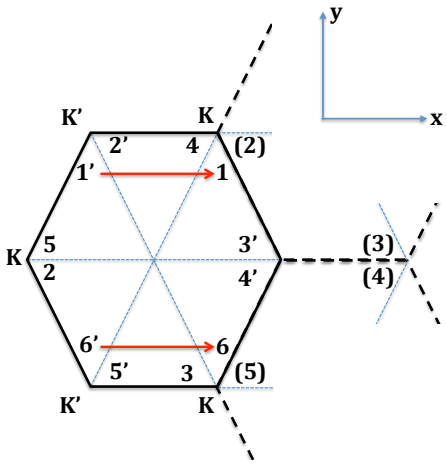

FIG. 3: Intervalley scattering for $\mathbf{K_0}=K_0\hat{x}$.

The small region of free propagation reinforces the fact that the second term in (A1) can be neglected.

Appendix B: Single-plasmon pole approximation

By satisfying the conductivity and f-sum rules, one can replace the cumbersome expression for the longitudinal dynamic dielectric function, $\epsilon_{0,0}(\mathbf{q},z)$, with a simplified single-plasmon pole expression.[33] For the intervalley case, the simplified form of the dielectric function follows from

$$\frac{V_{\mathbf{K}_0}}{\tilde{\epsilon}(q,\omega)} \;\approx\; V_{\mathbf{K}_0}\left(1+\frac{f_{iv}}{(\omega+i\delta)^2-\omega_q^2}\right), \qquad \text{(B1)}$$

where

$$\tilde{\epsilon}(q,\omega) \;=\; 1-\frac{f_{iv}}{(\omega+i\delta)^2-(\omega_q^2-f_{iv})}. \qquad \text{(B2)}$$

To satisfy the conductivity sum rule we should compare the SPP and random-phase approximation forms,

$$\int_0^\infty d\omega\omega \mathrm{Im}\{\epsilon_{0,0}(\mathbf{q},\omega)\}=\int_0^\infty d\omega\omega\mathrm{Im}\{\tilde{\epsilon}(q,\omega)\}. \qquad \text{(B3)}$$

Substituting (A1) and (B2) in left and right hand sides, respectively, and using Dirac identity one gets,

$$\frac{V_{K_0}}{3\hbar^2}\sum_{\mathbf{k}}(\Delta+\varepsilon_{\mathbf{k}+\bar{\mathbf{q}}}-\varepsilon_{\mathbf{k}})f_{\mathbf{k}}=\frac{f_{iv}}{2}. \qquad \text{(B4)}$$

Assuming parabolic dispersion, the residue is readily resolved,

$$f_{iv}=\frac{4\alpha\varepsilon_F(\Delta+\varepsilon_{\bar{\mathbf{q}}})}{3\hbar^2}\approx\frac{4\alpha\varepsilon_F\Delta}{3\hbar^2} \qquad \text{(B5)}$$

Appendix C: Intravalley band-gap renormalization

The redshift of the lower spin-split conduction band is calculated using Eqs. (6)-(8) in Ref. [30]. This calculation seems to overestimate the redshift compared with the one measured from the shift of the spectral peaks in the PL of ML-TMDs. The reason for this discrepancy stems from the treatment of the ultraviolet singularity (i.e., short-wave limit). Specifically, the correlation integral for long-range Coulomb excitations (intravalley) converges due to the phenomenological insertion of the particle-hole excitation to deal with short wavelengths (see Eq. (8) in Ref. [30]) . This insertion is not as problematic at relatively low densities applicable to typical semiconductor quantum wells. Schmitt-Rink and Ell found that the redshift (Σ_0 in (10) of the main text) behaves close to $\Sigma_0 \simeq -C_1 E_0 (a_B^2 n)^{1/3}$, where E_0 is the effective Rydberg energy and $C_1 = 3.1$ for exciton plasma when the electron and hole masses are equal (or $C_1 = 1.55$ for a single plasma component relevant for our discussion).

To better match the experimental result in ML-TMDs, I chose an alternative approach in which the integration cutoff is limited to $2k_F$ (or, equivalently, $4\varepsilon_F$), due to the change in the behavior of screening at larger values (see discussion in Sec. II.C in Ref. [49]). This choice yields a moderate redshift due to intravalley exchange and correlation, as shown in Fig. 2(d) of the main text. Similarly, this choice amounts to changing the value of C_1 from 1.55 to 0.6, as shown by the dashed line in Fig. 2(d).

Appendix D: Hole-rich systems

If the upper spin-split valence band is populated with holes, then intervalley plasmons should also play a role for type B excitons (formed by transitions from the lower spin-split valence band). However, their detection could be elusive since the transfer rate of excitons from type B to type A is ultrafast (much faster than the exciton recombination lifetime), leaving no time for the intervalley Coulomb excitation to dress the excitons. In this view, the smallness of Δ in the conduction band renders intervalley Coulomb excitations relevant in electron-rich ML-TMDs. For example, if $\Delta - \varepsilon_F$ is smaller than the energy of zone-edge phonons, a bottleneck in the energy relaxation of excitons is created, leaving enough time for intevalley Coulomb excitations to dress the excitons. This fact can be recognized from the comparable PL intensities in MoX_2 and WX_2,[71–73] implying that the recombination lifetime is comparable or shorter than the intervalley energy relaxation (otherwise the PL intensity in WX_2 should be much weaker). Another support is provided by the opposite-sign contributions of the spin-orbit coupling and electron-hole exchange, which overall diminish the magnitude of Δ in the conduction band.[67]

Acknowledgments

I am indebted for Aaron M. Jones and Xiaodong Xu for providing the animated PL measurements in the supplemental material and for many useful discussions. I am grateful to Jie Shan and Kin Fai Mak for discussing and sharing valuable experimental results prior to their publication. I am grateful to Refik Kortan for encouraging me to pursue this topic. This work was supported by the Department of Energy under Contract No. DE-SC0014349, National Science Foundation under Contract No. DMR-1503601, and the Defense Threat Reduction Agency under Contract No. HDTRA1-13-1-0013.

* hanan.dery@rochester.edu

[1] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, Nat. Nanotechnol. **6**, 147 (2011).
[2] D. Xiao, G.-B. Liu, W. Feng, X. Xu, and W. Yao, Phys. Rev. Lett. **108**, 196802 (2012).
[3] H. Zeng, J. Dai, W. Yao, D. Xiao, and X. Cui, Nat. Nanotechnol. **7**, 490 (2012).
[4] K. F. Mak, K. L. He, J. Shan, and T. F. Heinz, Nat. Nanotechnol. **7**, 494 (2012).
[5] T. Cao, G. Wang, W. Han, H. Ye, C. Zhu, J. Shi, Q. Niu, P. Tan, E. Wang, B. Liu, and J. Feng, Nat. Commun. **3**, 887 (2012).
[6] X. Xu, W. Yao, D. Xiao, and T. F. Heinz, Nat. Phys. **10**, 343 (2014).
[7] K. F. Mak, K. L. McGill, J. Park, P. L. McEuen, Science **344**, 1489 (2014).
[8] L. Yang, N. A. Sinitsyn, W. Chen, J. Yuan, J. Lou, and S. A. Crooker, Nat. Phys. **11**, 830-834 (2015).
[9] E. J. Bushong, Y. Luo, K. M. McCreary, M. J. Newburger, S. Singh, B. T. Jonker, and R. K. Kawakami, arXiv:1602.03568
[10] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C.-Y. Chim, G. Galli, and F. Wang, Nano Lett. **10**, 1271 (2010).
[11] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Phys. Rev. Lett. **105**, 136805 (2010).
[12] T. Korn, S. Heydrich, M. Hirmer, J. Schmutzler, and C. Schuller, Appl. Phys. Lett. **99**, 102109 (2011).
[13] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, M. S. Strano, Nat. Nanotechnol. **7**, 699 (2012).
[14] L. Britnell, R. M. Ribeiro, A. Eckmann, R. Jalil, B. D. Belle, A. Mishchenko, Y.-J. Kim, R. V. Gorbachev, T. Georgiou, S. V. Morozov, A. N. Grigorenko, A. K. Geim, C. Casiraghi, A. H. Castro Neto, and K. S. Novoselov, Science **340**, 1311 (2013).
[15] A. K. Geim and I. V. Grigorieva, Nature **499**, 419 (2013).
[16] F. Withers, O. Del Pozo-Zamudio, A. Mishchenko, A. P. Rooney, A. Gholinia, K. Watanabe, T. Taniguchi, S. J. Haigh, A. K. Geim, A. I. Tartakovskii, and K. S. Novoselov, Nat. Mater. **14**, 301 (2015).
[17] P. Li and I. Appelbaum, Phys. Rev. B **92**, 195129 (2015).
[18] A. T. Hanbicki, K. M. McCreary, G. Kioseoglou, M. Currie, C. S. Hellberg, A. L. Friedman, and B. T. Jonker, AIP

Advances **6**, 055804 (2016).
[19] L. V. Keldysh, JETP Lett. **29**, 658 (1979).
[20] P. Cudazzo, I. V. Tokatly, and A. Rubio, Phys. Rev. B **84**, 085406 (2011).
[21] T. Cheiwchanchamnangij and W. R. L. Lambrecht, Phys. Rev. B **85**, 205302 (2012).
[22] A. Ramasubramaniam, Phys. Rev. B **86**, 115409 (2012).
[23] K. F. Mak, K. He, C. Lee, G. H. Lee, J. Hone, T. F. Heinz, and J. Shan, Nat. Mater. **12**, 207 (2013).
[24] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, Phys. Rev. B **88**, 045318 (2013).
[25] A. Chernikov, T. C. Berkelbach, H. M. Hill, A. Rigosi, Y. Li, O. B. Aslan, D. R. Reichman, M. S. Hybertsen, and T. F. Heinz, Phys. Rev. Lett. **113**, 076802 (2014).
[26] A. Thilagam, J. Appl. Phys. **116**, 053523 (2014).
[27] D. K. Zhang, D. W. Kidd, K. Varga, Nano Lett. **15**, 7002 (2015).
[28] B. Ganchev, N. Drummond, I. Aleiner, and V. Falko, Phys. Rev. Lett. **114**, 107401 (2015).
[29] H. Haug and S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 3rd ed. (World Scientific, Singapore, 1994).
[30] S. Schmitt-Rink and C. Ell, J. Lumin. **30**, 585 (1985).
[31] R. Sooryakumar, D. S. Chemla, A. Pinczuk, A. C. Gossard and W. Wiegmann, L. J. Sham, Solid State Commun. **54**, 859 (1985)
[32] S. Schmitt-Rink, C. Ell, and H. Haug, Phys. Rev. B **33**, 1183 (1986).
[33] H. Haug and S. Schmitt-Rink, Prog. Quant. Electr. **9**, 3 (1984).
[34] A. M. Jones, H. Yu, N. J. Ghimire, S. Wu, G. Aivazian, J. S. Ross, B. Zhao, J. Yan, D. G. Mandrus, D. Xiao, W. Yao, and X. Xu, Nat. Nano. **8**, 634 (2013).
[35] J. Shang, X. Shen, C. Cong, N. Peimyoo, B. Cao, M. Eginligil, and T. Yu, ACS Nano **9**, 647 (2015). Figure 4(b) in this reference shows that the many-body PL peak in WS_2 emerges at low temperatures around 1.99 eV under weak excitation power and elevated electron densities (gate voltage of ∼60 V). At about 100 times stronger excitation power, shown in Fig. 4(d) of this reference, the exciton-exciton peak appears in the same spectral region, showing weaker dependence on gate voltage.
[36] Y. You, X.-X. Zhang, T. C. Berkelbach, M. S. Hybertsen, D. R. Reichman, and T. F. Heinz, Nat. Phys. **11**, 477 (2015).
[37] G. Plechinger, P. Nagler, J. Kraus, N. Paradiso, C. Strunk, C. Schller, and T. Korn, Phys. Status Solidi RRL **9**, 457 (2015).
[38] M. S. Kim, S. J. Yun, Y. Lee, C. Seo, G. H. Han, K. K. Kim, Y. H. Lee, and J. Kim, ACS Nano **10**, 2399 (2016).
[39] B. Scharf, Z. Wang, D. Van Tuan, K. F. Mak, J. Shan, I. Zutic, and H. Dery, arXiv:1604.00068
[40] J. P. Löwenau, S. Schmitt-Rink, and H. Haug, Phys. Rev. Lett. **49**, 1511(1982).
[41] See Supplemental Material for animations of the PL evolution in ML-WSe_2 and ML-$MoSe_2$ when changing the gate voltage. Details on the fabrication and measurement of the gated samples are found in Refs. [34] and [43].
[42] D. MacNeill, C. Heikes, K.-F. Mak, Z. Anderson, A. Kormányos, V. Zólyomi, J. Park, and D. C. Ralph, Phys. Rev. Lett. **114**, 037401 (2015).
[43] J. S. Ross, S. Wu, H. Yu, N. J. Ghimire, A. M. Jones, G. Aivazian, J. Yan, D. G. Mandrus, D. Xiao, W. Yao, X. Xu, Nat. Commun. **4**, 1474 (2013).
[44] Note that while bi-excitons were observed in the PL of WX_2 but not of MoX_2 under comparable photoexcitation powers, their presence in MoX_2 was resolved in ultrafast transient spectroscopy. See, e.g., E. J. Sie, A. J. Frenzel, Y.-H. Lee, J. Kong, and N. Gedik, Phys. Rev. B **92**, 125417 (2015); N. Kumar, Q. Cui, F. Ceballos, D. He, Y. Wang, and H. Zhao, Phys. Rev. B **89**, 125427 (2014); C. Mai, A. Barrette, Y. Yu, Y. G. Semenov, K. W. Kim, L. Cao, and K. Gundogdu, Nano Lett. **14**, 202 (2014).
[45] R. Schmidt, G. Berghauser, R. Schneider, M. Selig, P. Tonndorf, E. Malic, A. Knorr, S. Michaelis de Vasconcellos, and R. Bratschitsch, Nano Lett. **16** 2945 (2016).
[46] H. Dery and Y. Song, Phys. Rev. B **92**, 125431 (2015).
[47] X.-X. Zhang, Y. You, S. Yang, F. Zhao, and T. F. Heinz, Phys. Rev. Lett. **115**, 257403 (2015).
[48] For the bare potential, I have used the conventional 2D form, $V_{\mathbf{q}} = 2\pi e^2/A\epsilon_r q$, where A is the quantization area. Note that the geometrical correction to the dielectric screening introduced by Keldysh is relevant when $qd \ll 1$ ($d \sim 1$ nm is the monolayer thickness including the van der Waals gap).[19] This condition breaks for $q = K_0$, at which Eq. (1) in Ref. [19] conforms to the conventional 2D form. As for intravalley Coulomb interaction in the limit $q \to 0$, the dominant contribution comes from the inverse dependence on the Thomas-Fermi screening wavenumber at elevated densities.
[49] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[50] K. Kośmider and J. Fernández-Rossier, Phys. Rev. B **87**, 075451 (2013).
[51] G.-B. Liu, W.-Y. Shan, Y. Yao, W. Yao, and D. Xiao, Phys. Rev. B **88**, 085433 (2013).
[52] K. Kośmider, J. W. González, and J. Fernández-Rossier, Phys. Rev. B **88**, 245436 (2013).
[53] A. Kormányos, G. Burkard, M. Gmitra, J. Fabian, V. Zólyomi, N. D. Drummond, V. Fal'ko, 2D Mater. **2**, 022001 (2015).
[54] T. Tudorovskiy and S. A. Mikhailov, Phys. Rev. B **82**, 073411 (2010).
[55] D. Pines and D. Bohm, Phys. Rev. **85**, 338 (1952).
[56] P. Nozières and D. Pines, Phys. Rev. **109**, 741 (1958).
[57] S. L. Adler, Phys. Rev. **126**, 413 (1962).
[58] N. Wiser, Phys. Rev. **129**, 62 (1963).
[59] A. W. Overhauser, Phys. Rev. B **3**, 1888 (1971).
[60] Z. Y. Zhu, Y. C. Cheng, and U. Schwingenschlögl, Phys. Rev. B **84**, 153402 (2011).
[61] T. Cheiwchanchamnangij, W. R. L. Lambrecht, Y. Song and H. Dery, Phys. Rev. B **88**, 155404 (2013).
[62] T. M. Rice, Nuovo Cimento **23B**, 226 (1974).
[63] R. Zimmermann and M. Rösler, Phys. Stat. Solidi B **75**, 633 (1976).
[64] Y. Song and H. Dery, Phys. Rev. Lett. **111**, 026601 (2013).
[65] M. Z. Maialle, E. A. de Andrada e Silva, and L. J. Sham, Phys. Rev. B **47**, 15776 (1993).
[66] H. Yu, G.-B. Liu, P. Gong, X. Xu, and W. Yao, Nat. Commun. **5**, 3876 (2014).
[67] J. P. Echeverry, B. Urbaszek, T. Amand, X. Marie, and I. C. Gerber, Phys. Rev. B **93**, 121107 (2016).
[68] The intravalley Coulomb interaction largely keeps the unpopulated upper valleys intact, similar to the case of excited states in quantum wells. See, e.g., R. Zimmermann, E. H. Böttcher, K. Kirstaedted, and D. Bimberg, Superlattices Microstruct. **7**, 433 (1990); J. C. Ryan and T. L. Reinecke, Phys. Rev. B **47**, 9615 (1993).

[69] F. Wu, F. Qu, and A. H. MacDonald, Phys. Rev. B **91**, 075310 (2015).

[70] The population of dark excitons can be neglected by assuming slow spin relaxation compared with the recombination lifetime. This assumption is justified given that the measured spin relaxation at low temperatures[8,9] is about three order of magnitude longer that the measured recombination lifetime.[71–73] In addition, one difference from the elevated electron density case is that the redshift due to long-range excitations (Σ_0) should be reevaluated if the background electron density is smaller than the exciton density.[33]

[71] G. Wang, L. Bouet, D. Lagarde, M. Vidal, A. Balocchi, T. Amand, X. Marie, and B. Urbaszek, Phys. Rev. B **90**, 075413 (2014).

[72] G. Wang, E. Palleau, T. Amand, S. Tongay, X. Marie, and B. Urbaszek, Appl. Phys. Lett. **106**, 112101 (2015).

[73] C. Robert, D. Lagarde, F. Cadiz, G. Wang, B. Lassagne, T. Amand, A. Balocchi, P. Renucci, S. Tongay, B. Urbaszek, and X. Marie, Phys. Rev. B **93**, 205423 (2016).

[74] A. M. Jones, H. Yu, J. R. Schaibley, J. Yan, D. G. Mandrus, T. Taniguchi, K. Watanabe, H. Dery, W. Yao, and X. Xu, Nat. Phys. **12**, 323 (2016).