APS physics

CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

# Excitons without effective mass: Biased bilayer graphene

Pengke Li (□□□) and Ian Appelbaum

Phys. Rev. B **99**, 035429 — Published 22 January 2019

DOI: 10.1103/PhysRevB.99.035429

# Excitons without effective mass: biased bilayer graphene

Pengke Li (李鹏科)* and Ian Appelbaum†
*Department of Physics, U. of Maryland, College Park, MD 20742*

Understanding the dynamics of excitons in two dimensional semiconductors requires a theory that incorporates the essential physics distinct from their three-dimensional counterparts. In addition to the modified dielectric environment, single-particle states with strongly non-parabolic dispersion appear in many two-dimensional band structures, so that "effective mass" is ill-defined. Focusing on electrostatically-biased bilayer graphene as an example where quartic (and higher) dispersion terms are necessary, we present a semi-analytic theory used to investigate the properties of ground and excited excitonic states. This includes determination of relative oscillator strengths and magnetic moments ($g$-factors) which can be directly compared to recent experimental measurements.

Analytic solution of the electron Schrödinger equation with the attractive Coulomb potential, yielding the Rydberg spectrum of the hydrogen atom, was among the first – and still monumental – achievements of quantum mechanics beginning nearly one hundred years ago. Despite its nominal origin in atomic physics, this problem is also very relevant to the solid-state, as a nearly identical mathematical formulation determines the interaction of charge carriers with immobile shallow impurities,[1] and also the electrostatic interaction between electrons and holes themselves, resulting in their mutually bound state: *excitons*, somewhat analogous to positronium.[2] The presence of these excitons can be indirectly observed in experiments, e.g. optical absorption or photoconduction spectroscopy, as resonances at energy just below the interband excitation edge (see Fig. 1).

The 'envelope approximation' often used to model physical attributes of these examples assumes that the effect of absorbing the periodic lattice potential into quasiparticle dispersion only modifies the effective mass, and the lowest-order parabolic relationship between (quasi)momentum and energy remains. However, parabolic dispersion is by no means the only possible outcome endowed by a periodic potential. Especially in two dimensional electronic materials, where weak inter-subband $k \cdot p$ matrix elements suppress otherwise strong band repulsion across a forbidden gap, nonparabolic 'Mexican hat' or 'caldera'-shaped bands are quite common.[3,4] As shown in Fig. 1, the extrema of these dispersions are indeed approximately quadratic in the radial $k$-direction, but completely flat (ignoring higher order warping from remote bands) in the orthogonal azimuthal direction, yielding a divergent density of states. Such unfamiliar behavior departs considerably from the hydrogen atom problem and cannot be captured by simple mass renormalization. A generalized theory that captures such nontraditional quasiparticle dispersion is necessary to promote the study of unusual excitonic behavior in 2D van der Waals systems.

Motivated by recent experimental measurements of field-tunable exciton spectrum in biased bilayer graphene (BBG),[5] where both electron and hole have nearly symmetrical 'caldera' dispersion, we present a general variational theory for the (screened) Coulomb problem in two dimensions when quasiparticle dispersion cannot be captured solely by a single lowest-order parabola $\propto k^2$. This theory allows the calculation of exciton wavefunctions and spectra, oscillator strengths, and valley-dependent orbital magnetic moments in a transparent way not dependent on opaque numerical schemes such as density-functional theory (DFT).[6]

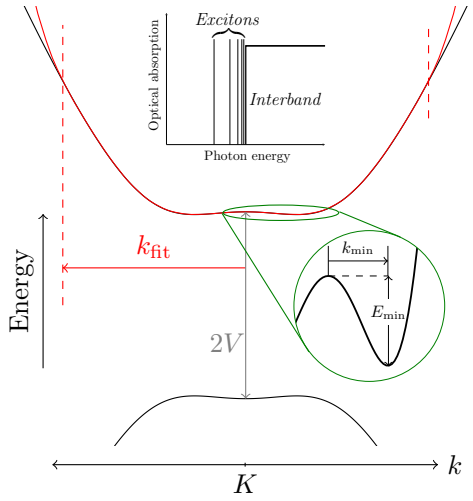


FIG. 1. Single-particle electronic structure of biased bilayer graphene near the $K(K')$-point under bias $2V = 100$ meV shown in black. Red line is least-squares 8th-order polynomial fit within the vertical dashed lines. Magnified region (circled in green) emphasizes the nonparabolic dispersion and energetic depth of band-edge extrema. Inset above shows schematic optical absorption spectrum close to the interband transition threshold, with a manifold of discrete exciton states at lower energy.

In general, a two-particle exciton wavefunction can be viewed as the superposition of direct products of electron and hole quasiparticle states in momentum space, weighted by an envelope function. As a result, an exact evaluation of the exciton binding energy through the field-theoretic Bethe-Salpeter equation[7] using quasiparticle states from DFT is computationally demanding, and any physical insight into the problem would be obscured behind the numerical details. Our theory focuses on the dominant contributions so that, instead of pursuing ab-

solute precision of the binding energy, it reveals insight into the fundamental exciton physics. For BBG with an analytic Hamiltonian, our theory is especially important to explain excitonic evolution under electronic structure tuning via external electrical gate bias.

In the following, we describe the appropriate Hamiltonian accounting for arbitrary nonparabolic dispersion and attractive interaction between electron- and hole-like quasiparticles. Then, we give analytic expressions for its matrix elements in a generalized Gaussian basis, amenable to variational energy minimization. Implementing this formalism to the specific case of bilayer graphene, we incorporate the perpendicular electric field-dependence of both band structure and Coulomb screening length to obtain envelope functions for both the ground and first excited state. Perturbation theory applied to a tight-binding description then allows the calculation of analytic expressions for oscillator strength and magnetic field-induced valley splitting, which can both be quantitatively compared to experimental observations. In order to maintain continuity in our discussion of the underlying physics, we defer the lengthy technical details of some derivations to the Supplemental Material (SM)[8].

In this theory, the quasiparticle kinetic energy necessarily acquires additional terms (quartic $\propto k^4$ and so on) in higher order, appearing in the effective Hamiltonian through canonical substitution $\boldsymbol{k} \to -i\nabla$[9,10] giving $H = -A_1\nabla^2 + A_2\nabla^4 - A_3\nabla^6 + \cdots + V(r)$. The coefficients $A_\xi$ of all salient orders can be calculated via least-squares fitting (rather than the unsuitable Taylor expansion; see SM) over a test range including the dispersion extrema ($k_{\text{fit}}$ as shown in Fig. 1). Importantly, the size of this range in reciprocal space must be consistent with the exciton wavefunction radius in real space, which is readily calculated from our theory below.

The presence of nonparabolic terms in the kinetic energy complicates the usual reduction of the two-particle problem to a separable system of relative and center-of-mass coordinates. Furthermore, in the rotationally-invariant caldera dispersion, 'mass' is not well defined along the azimuthal direction. However, relative position $r = r_e - r_h$ and total momentum $P = p_e + p_h$ are still meaningful quantities. As detailed in SM, when both electron and hole have identical dispersions as is nearly the case in BBG, use of canonically conjugate variables $p = (p_e - p_h)/2$ and $R = (r_e + r_h)/2$ allow the two-particle effective (classical) Hamiltonian to be written up to quartic order as

$$\left[A_1(\frac{1}{2}P^2 + 2p^2) + A_2(\frac{1}{8}P^4 + 3p^2P^2 + 2p^4)\dots\right] + V(r)$$
$$\xrightarrow{P=0} 2\left[A_1p^2 + A_2p^4\dots\right] + V(r). \quad (1)$$

Unlike the usual parabolic kinetic energy case, it is not possible to eliminate all terms that mix momenta $p$ and $P$, so full separation into decoupled equations of motion fails here; in general, the free exciton dispersion will be nonparabolic *and* the exciton wavefunction in relative coordinate $\psi(r)$ will depend on total momentum $P$. However, negligible photon momentum requires $P \sim 0$ for analysis of behavior under optical excitation, which is our focus.[11]

When the wavefunction is confined to two dimensions, the electrostatic interaction is modified, as initially derived by Rytova[12] and later rediscovered by Keldysh.[13] There are two asymptotic limits as elaborated by Cudazzo *et al.*[14]: at large relative distances, the potential behaves like the usual Coulomb interaction, but close to the origin it diverges only logarithmically. A screening length $r_0$, determined by the 2D polarizability, separates these two limiting behaviors and is an important ingredient in our calculation.

By considering photon-induced transition rate and Kramers-Kronig relations in the usual way (see SM), the 2D screening length is generically given by

$$r_0 = \frac{q^2\hbar^2}{4\pi^2\epsilon_0 m_0^2}\sum_{\text{c,v}}\int\frac{|P_{\text{cv}}|^2}{E_{\text{cv}}^3}d^2k, \quad (2)$$

where $q$ is fundamental charge, $\epsilon_0$ is the vacuum permittivity, and $m_0$ is the free electron mass. In addition, the generally $\boldsymbol{k}$-dependent terms in the integrand are the momentum matrix element $P_{\text{cv}}$ and the energy gap $E_{\text{cv}}$, with the latter indicating an inverse relationship with $r_0$, which further affects the binding energy.

Our full two-particle Hamiltonian, consisting of nonparabolic kinetic energy operators and the Rytova-Keldysh form of electron-hole interaction, is not amenable to analytic diagonalization, so a variational method is applied. First of all, in this quasi-rotationally invariant system, the centrifugal term of the Laplacian ($\frac{1}{r^2}\frac{d^2}{d\phi^2} \to -\frac{m^2}{r^2}$) demands that the wavefunction behave like $r^{|m|}$ for small $r$, where $m$ is the angular momentum quantum number. Using a modified stretched exponential trial function $r^{|m|}\exp[-(r/b)^\beta]\exp(im\phi)$, we find that the expectation value of nonparabolic terms ($\nabla^4$ and higher) requires $\beta \geq 2$ to avoid divergence. Values of $\beta$ significantly greater than 2 would cause a sharp wavefunction suppression at large distance and is disfavored by the Coulombic potential asymptotics[15].

By choosing the $\beta = 2$ Gaussian trial envelope wavefunction, we can calculate matrix elements of kinetic energy operators to arbitrary order with

$$\langle\psi_i|(-i\nabla)^{2\xi}|\psi_j\rangle = \pi 4^\xi(\xi+m)!\frac{(b_i^2b_j^2)^{m+1}}{(b_i^2+b_j^2)^{\xi+m+1}}, \quad (3)$$

where $\xi = 0, 1, \dots$ indexes powers of the Laplacian. When normalized by the $\xi = 0$ inner product, this yields a single-particle variational kinetic energy (for $b_i = b_j = b$) of

$$K_m = \sum_{\xi=1} A_\xi 2^\xi\frac{(\xi+m)!}{m!}b^{-2\xi}. \quad (4)$$

Evaluating the expectation value of the potential energy requires deeper analysis. Here, we find the integral

representation provided by Cudazzo *et al.*[14] especially useful, where the Rytova-Keldysh potential is due to a fictitious charge density distributed normal to the plane $q\delta(r)\frac{e^{-|z|/r_0}}{2r_0}$. As detailed in SM, normalized diagonal matrix elements in the $m = 0$ Gaussian basis can be analytically calculated by inverting the order of integration over $r$ and $z$, yielding a generic 2D potential energy

$$U_0 = -\frac{q^2}{8\pi\epsilon_0\bar{\epsilon}r_0}\left[e^{-x^2}\left(\pi\mathrm{erfi}\left(x\right) - \mathrm{Ei}\left(x^2\right)\right)\right], \quad (5)$$

where $x^2 = \frac{b^2}{8r_0^2}$ and $\bar{\epsilon}$ is the relative permittivity of the surrounding medium. Here, $\mathrm{Ei}(x)$ is the exponential integral function and $\mathrm{erfi}(x)$ is the imaginary error function.

An analytic expression for the Rytova-Keldysh potential matrix element with $m = 1$ is given in SM. For this and higher quantum numbers, the kinetic energy expectation values in Eq. (4) monotonically increase, whereas the potential energy tends to decrease, leading to steadily larger envelope wavefunctions and shallower binding energy.

Having presented the basic elements of our generic approach, we now focus on the realistic example of excitons in the BBG system, whose low energy electronic structure is captured by the four coupled $p_z$ orbitals of both atomic layers, each of which contains two carbon sublattices, $A$ and $B$. We follow the notation of McCann and Koshino,[16] using the basis ordering $\{A_1, B_1, A_2, B_2\}$ and write the $4 \times 4$ tight-binding effective Hamiltonian at the $K$-point as $H_0 + H_1 + H_2$. $H_0$ is the nearest-neighbor $p_z$-orbital Hamiltonian accounting for lowest-order intra-/inter-layer coupling with hopping parameters $\gamma_0 = 3$ eV and $\gamma_1 = 0.4$ eV, respectively, and electric-field biasing with on-site energy $\pm V$. This dominant term determines the eigenstates and captures the gross structure of the electron/hole dispersion $E = \pm\left[\frac{\gamma_1^2}{2} + \frac{3}{4}(a\gamma_0 k)^2 + V^2 - \frac{1}{2}\sqrt{\gamma_1^4 + 3(\gamma_1^2 + 4V^2)(a\gamma_0 k)^2}\right]^{\frac{1}{2}}$ (with $a = 2.46$ Å the lattice constant), and is used to extract the polynomial coefficients $A_\xi$ used in Eq. (4). Additional terms $H_1$ and $H_2$ have only a minor effect on the energy dispersion and the eigenstates, but are essential perturbations to include in understanding the exciton oscillator strength and orbital magnetic moment. The former reflects next-nearest-neighbor interlayer "skew" coupling $\gamma_3 = 0.3$ eV between non-dimer sites, resulting in trigonal warping of the bands. The remaining $H_2$ is responsible for the electron-hole dispersion asymmetry, including the dimer/nondimer on-site asymmetry energy $\Delta' \approx 0.02$ eV and the skew interlayer coupling $\gamma_4 = 0.14$ eV between a non-dimer and a dimer sites. Full matrix expressions for the Hamiltonian are given in SM.

The simplicity of $H_0$ allows analytic evaluation of the momentum matrix element between the conduction and valence bands as (see detailed calculation in SM)

$$|\langle P_{\mathrm{cv}}(k,\phi)\rangle|^2 = \frac{9\gamma_1^2(a\gamma_0 k)^2(V^2\cos^2\phi + E^2\sin^2\phi)}{4[\gamma_1^4 + 4(\gamma_1^2 + 4V^2)(a\gamma_0 k)^2]E^2}\left(\frac{m_0 a\gamma_0}{\hbar}\right)^2, \quad (6)$$

where $\phi$ is the angle between quasimomentum $\boldsymbol{k}$ and the photon polarization (chosen as parallel to the $x$-axis). This result is notable for the absence of optical coupling across the fundamental band gap at the $K$-point ($k = 0$).[17] It is often the case that symmetry is responsible for vanishing matrix elements, but here no such constraint exists. As we will show, symmetry-allowed terms in perturbation $H_1$ are responsible for nonzero interband optical coupling and a bright $m = 0$ exciton.

With $P_{\mathrm{cv}}$, the integration for the screening length $r_0$ in Eq. (2) is straightforward but yields a cumbersome expression (see SM). Graphically, however, it is a featureless

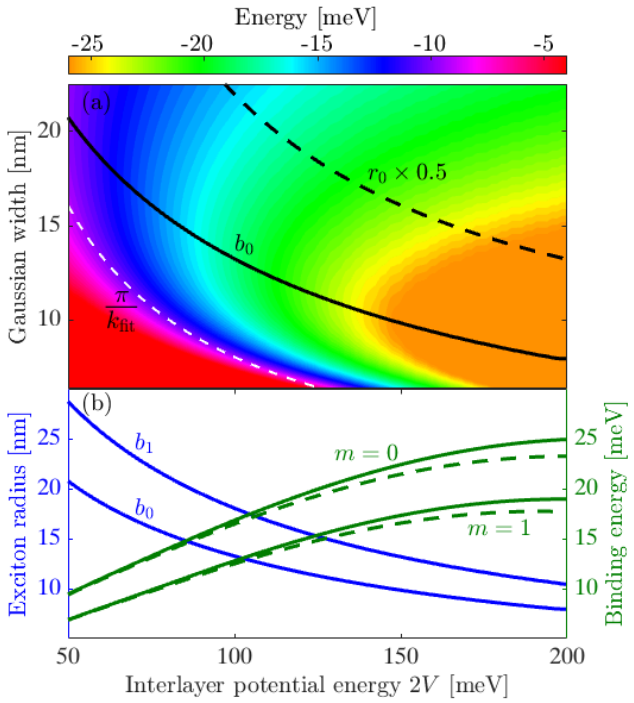


FIG. 2. (a) Energy landscape of a single $m = 0$ Gaussian trial function for the ground-state exciton envelope function of BBG, using Eqs. (4) and (5). Kinetic energy polynomial coefficients $A_\xi$, where $\xi = 1, 2, 3, 4$, are determined by least-squares fitting the exact dispersion between our choice of $\pm k_{\mathrm{fit}}$, giving a real-space lengthscale shown as a dashed white line. The energetic minima, indicating variational optimum lengthscale $b_0 > \pi/k_{\mathrm{fit}}$, is shown as a solid black line. Black dashed line presents half of the 2D screening length $r_0 > b_0$. (b) Variational optimum lengthscale shown in blue (left axis) for both ground-state ($b_0$) and $m = 1$ excited state ($b_1$). Binding energy in green (right axis) accounts for reduction in bound-continuum energy due to $E_{\mathrm{min}}$ (see Fig. 1). Dashed lines are energies of the single trial function, which are improved by the lowest generalized eigenvalue of the problem with an optimized 5-function basis (see text).

curve, as shown by the dashed black line in Fig. 2(a); this length scale should be compared to the exciton Gaussian width discussed below. Clearly, with increasing gate bias (larger $|V|$), $r_0$ decreases mainly due to the increased band gap.

The transcendental functions in Eq. (5) with $r_0$ as an input require a numerical minimization of the total energy $K_0 + U_0$ to find the optimum value of variational parameter $b$. In Fig. 2(a) we plot the energy of the $m = 0$ Gaussian exciton for $\bar{\epsilon} = 4$ (appropriate for BN encapsulation) as a function of electric field bias, and indicate the length scale $b_0$ that minimizes it with a solid curve. The dashed white curve is the equivalent length scale determined by the reciprocal of the polynomial fitting region $\pi/k_{\min}$, showing consistency with our initial assumptions.

This exciton size variation with bias field is reproduced in Fig. 2(b) as a solid blue line, along with the equivalent result for $m = 1$. Both indicate increased confinement with gate bias, consistent with increasing variational binding energies (dashed green lines) of both excitons using a single trial wavefunction. To improve upon the single-function variational binding energies, we augment the basis with four additional functions of the same form but with optimized exponentially-spaced length scales[18] and solve for the lowest generalized eigenvalue of $\langle i|H|j\rangle\Psi = E\langle i|j\rangle\Psi$, using Eq. (3) and a generalization of Eq. (5) where $2/b^2 \to (1/b_i^2 + 1/b_j^2)$. Binding energies calculated in this way (solid green lines) can typically be improved by only less than a few percent, indicating the suitability of the chosen Gaussian-type basis for this problem. The difference between the two exciton binding energies (several meV) and its gate bias dependence agree with the experimentally-measured value.[5]

Our envelope wavefunctions can now be used to examine the exciton "brightness", by evaluating the oscillator strength $f_m^{\rm x} \propto |\int \Phi_m(k) P_{\rm cv} d^2k|^2/E_{\rm x}$,[19] where $\Phi_m(k)$ is the Fourier transform of exciton envelope function, and $E_{\rm x}$ is the excitation energy of the exciton. Here we employ the Löwdin partitioning method to reduce the full $4 \times 4$ Hamiltonian to a $2 \times 2$ matrix in the non-dimer $\{A_1, B_2\}$ basis that captures the two gap-edge bands.[16] Considering only the dominant term $H_0$, the eigenstates of this two-level system are

$$|{\rm c}\rangle = \begin{bmatrix} \cos\frac{\eta}{2} \\ e^{2i\phi}\sin\frac{\eta}{2} \end{bmatrix}, \quad \text{and} \quad |{\rm v}\rangle = \begin{bmatrix} -e^{-2i\phi}\sin\frac{\eta}{2} \\ \cos\frac{\eta}{2} \end{bmatrix}, \quad (7)$$

where $\eta \approx 3(a\gamma_0)^2k^2/4V\gamma_1$ (which vanishes at the $K$-point). We must emphasize here that, to maintain the adiabaticity of the wavefunction through the $K$-point, $\boldsymbol{k}$-dependent phase factors $e^{\pm 2i\phi}$ *should not* be assigned arbitrarily among the components of the states[6], which is crucial in determining the exciton optical selection rules (see SM). In this band basis, the matrix element of the momentum operator $\frac{m_0}{\hbar}\nabla_k(H_0 + H_1)$ is

$$P_{\rm cv} \approx \frac{m_0}{2\hbar}\left[\sqrt{3}a\gamma_3 - \frac{3a^2\gamma_0^2\gamma_1}{\gamma_1^2 + V^2}ke^{-i\phi}\right]. \quad (8)$$

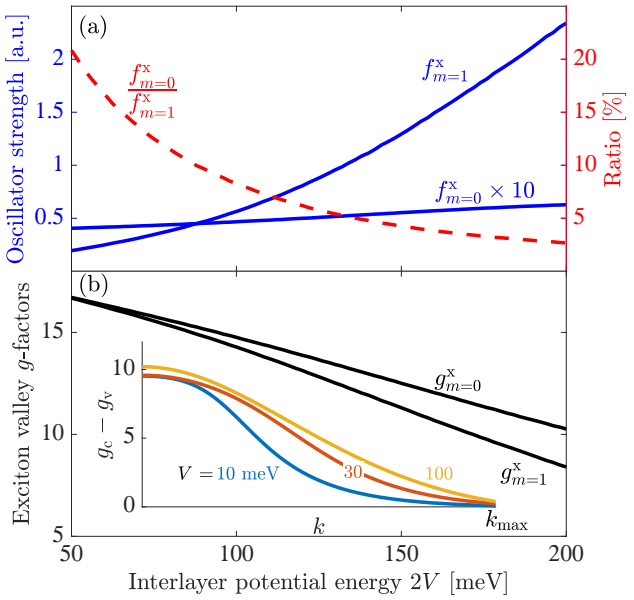


FIG. 3. (a) Relative oscillator strengths of the $m = 0$ and $m = 1$ excitons (blue curves) and their ratio (red dashed curve), as a function of the gate bias energy $2V$. (b) Valley $g$-factors of the $m = 0$ and $m = 1$ excitons. Inset: $k$-dependent $g$-factor differences of the conduction and valence bands, under bias conditions $V = 10$, 30 and 100 meV. Here, the full scale of $k$ is normalized according to $\frac{\sqrt{3}}{2}a\gamma_0 k_{\rm max} = \frac{\gamma_1}{2}$, so that $k_{\rm max}$ is less than 2% from the $K$ point to the $\Gamma$ point.

The two bracketed terms play different roles due to their parity. Specifically, the first ($k$-independent) and the second ($k$-linear) terms are relevant to the $\Phi_{m=0}$ and $\Phi_{m=1}$ envelope wavefunctions, respectively, to produce nonvanishing azimuthal integration of $\Phi_m P_{\rm cv}$. Importantly, electromagnetic coupling of the $m = 0$ exciton ground state depends crucially on the coupling parameter $\gamma_3$. Both oscillator strengths increase as a function of gate bias, as shown in Fig. 3(a). Since $\gamma_3 \approx 0.1\gamma_0$, $f^{\rm x}_{m=0}$ is one order of magnitude smaller than $f^{\rm x}_{m=1}$, even though the single particle excitation of the latter is of higher order in $k$. At large gate bias when both excitons share similar $E_{\rm x}$, the ratio of their oscillator strengths can be estimated solely from integration of $\Phi_m P_{\rm cv}$ (see SM),

$$\frac{\left(\frac{\sqrt{6\pi}m_0a\gamma_3}{\hbar b_0}\right)^2}{\left(\frac{12\sqrt{\pi}m_0a^2\gamma_0^2\gamma_1}{\hbar b_1^2(\gamma_1^2+V^2)}\right)^2} = \left(\frac{1}{2\sqrt{6}}\frac{b_1^2}{b_0 a}\frac{(\gamma_1^2+V^2)\gamma_3}{\gamma_0^2\gamma_1}\right)^2. \quad (9)$$

For example, with a gate bias $2V = 100$ meV, using the variational values [see Fig. 2(b)] $b_0 \approx 13$ nm and $b_1 \approx 18$ nm, Eq. (9) gives a ratio of $\sim 8\%$ that matches well with experimental observation.[5]

Lastly, we examine the exciton magnetic susceptibility. Similar to positronium, the angular momentum of the envelope function has diminished contribution to the magnetic moment,[2] due to the similar dispersion but opposite charge of the electron and hole. On the other hand, the difference between conduction and valence quasi-particle

orbital $g$-factors can contribute to the magnetic susceptibility through the Bloch part of the exciton wavefunction. Indeed, electron-hole asymmetry is induced by $H_2$, resulting in an exciton *valley* $g$-factor due to the opposite magnetic moments at time-reversed $K$ and $K'$ valleys.

The orbital magnetic moment of a quasiparticle state[20]

$$g_n(\boldsymbol{k})\mu_B = i\frac{e\hbar}{2m_0^2}\sum_{\ell\neq n}\frac{\boldsymbol{P}_{n\ell}(\boldsymbol{k})\times\boldsymbol{P}_{\ell n}(\boldsymbol{k})}{E_n(\boldsymbol{k})-E_\ell(\boldsymbol{k})} \qquad (10)$$

is identical for the two bands in a generic two level system, so we return to the full $4\times4$ Hamiltonian and treat $H_2$ perturbatively. The difference between $g$-factors of the electron and hole states is analytic (see SM) at the $K$-point,

$$g_{\rm v}-g_{\rm c}\approx\frac{3m_0a^2}{\hbar^2\gamma_1^2}\left[\gamma_0^2\Delta'\left(1+\frac{4V^2}{\gamma_1^2}\right)+2\gamma_1\gamma_0\gamma_4\right]\approx 10, \quad (11)$$

composed of two contributions within square brackets. The first one $\propto\Delta'$ is due to the dimer-nondimer onsite asymmetry resulting in different energy denominators for conduction and valence bands in Eq. (10). The remaining part is more dominant, involving interference between the $\gamma_0$- and $\gamma_4$-dependent matrix elements in the momentum operator, as evident by their product. As a result, the difference between $g_{\rm v}-g_{\rm c}$ at $K$ and $K'$ is $\sim 20$. The $k$-dependence of $g_{\rm v}-g_{\rm c}$ is shown in the inset of Fig. 3(b) under three different bias fields. As expected, the energy denominators between the gap edge bands and remote bands in Eq. (10) increase as $k^2$ and quickly suppress the value of $g_{\rm v}(k)-g_{\rm c}(k)$ at large $k$.

The exciton valley $g$-factors contributed by the Bloch wave part are calculated (see SM) by

$$g_m^{\rm x} = 2\int|\Phi_m(k)|^2[g_{\rm v}(\boldsymbol{k})-g_{\rm c}(\boldsymbol{k})]d^2\boldsymbol{k} \qquad (12)$$

for both the $m=0$ and $m=1$ excitons, and presented in Fig. 3(b) as a function of the gate bias. As $V$ increases, excitons are more confined with smaller radii and larger $k$-space distributions of their envelope wavefunctions, which reduce the valley $g$-factors. $g_{m=1}^{\rm x}$ decreases faster than $g_{m=0}^{\rm x}$ since $\Psi_{m=1}$ is linear in $k$ and further suppresses the contribution around $k=0$. Note that the Bloch wave contributions to both exciton $g$-factors do not closely match the experimentally observed large $g$-factor $\sim 20$ for $m=0$ and a negligible magnetic susceptibility for $m=1$ excitons.[5] In that experiment, broadband excitation of a relatively high density of excitons and free carriers may push the system into a strong correlation regime and cause significant deviation from the expected behavior of an isolated exciton. This extension to our theory, however, is beyond the scope of discussion in this Letter.

We end by emphasizing the generality of Eqs. (2), (3), (4) and (5) applied to excitons in an arbitrary two-dimensional semiconductor with approximately rotationally-invariant nonparabolic bands, such as the 'caldera'-like valence band in $D_{3h}$ three-six-enes $Ga_{1-x}In_xS_ySe_{1-y}$.[4] Other deviations from parabolic dispersion abound, including Rashba spin-split bands[10,21,22] and anisotropic examples of recent interest such as in the valence band of phosphorene[3,23] or the 'camel-back' valence band in 3D bulk tellurium,[24] for which our matrix element expressions can be appropriately modified; finite thickness corrections due to multilayer geometry can also be incorporated[25].

* pengke@umd.edu
† appelbaum@physics.umd.edu

[1] A. K. Ramdas and S. Rodriguez, Rep. Prog. Phys. **44**, 1297 (1981).
[2] M. Deutsch, Proc. Am. Acad. Arts and Sci. **82**, 331 (1953).
[3] P. Li and I. Appelbaum, Phys. Rev. B **90**, 115439 (2014).
[4] P. Li and I. Appelbaum, Phys. Rev. B **92**, 195129 (2015).
[5] L. Ju, L. Wang, T. Cao, T. Taniguchi, K. Watanabe, S. G. Louie, F. Rana, J. Park, J. Hone, F. Wang, and P. L. McEuen, Science **358**, 907 (2017).
[6] C.-H. Park and S. G. Louie, Nano Lett. **10**, 426 (2010).
[7] E. E. Salpeter and H. A. Bethe, Phys. Rev. **84**, 1232 (1951).
[8] See Supplemental Material at [URL will be inserted by publisher] for a PDF of detailed derivations.
[9] P. Bhardwaj and N. Das, Int. J. Eng. Sci. Res. Tech. **5**, 289 (2016).
[10] B. Skinner, Phys. Rev. B **93**, 235110 (2016).
[11] V. V. Cheianov, I. L. Aleiner, and V. I. Fal'ko, Phys. Rev. Lett. **109**, 106801 (2012).
[12] N. S. Rytova, Moscow U. Phys. Bull. **3**, 30 (1967), arxiv:cond-mat/1806.00976.
[13] L. Keldysh, JETP Lett. **29**, 658 (1979).
[14] P. Cudazzo, I. V. Tokatly, and A. Rubio, Phys. Rev. B **84**, 085406 (2011).
[15] B. Skinner, B. I. Shklovskii, and M. B. Voloshin, Phys. Rev. B **89**, 041405 (2014).
[16] E. McCann and M. Koshino, Rep. Prog. Phys. **76**, 056503 (2013).
[17] W. Yao, D. Xiao, and Q. Niu, Phys. Rev. B **77**, 235406 (2008).
[18] R. Ditchfield, W. J. Hehre, and J. A. Pople, J. Chem. Phys. **52**, 5001 (1970).
[19] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 2010).
[20] M. Dresselhaus, G. Dresselhaus, and A. Jorio, in *Group Theory: Application to the Physics of Condensed Matter* (Springer, Berlin, 2007) p. 411.
[21] E. Rashba and V. Sheka, Fiz. Tverd. Tela: Collected Papers **2**, 162 (1959).
[22] G. Bihlmayer, O. Rader, and R. Winkler, New J. Phys. **17**, 050202 (2015).
[23] E. Prada, J. V. Alvarez, K. L. Narasimha-Acharya, F. J. Bailen, and J. J. Palacios, Phys. Rev. B **91**, 245421 (2015).
[24] P. Li and I. Appelbaum, Phys. Rev. B **97**, 201402 (2018).

[25] D. Van Tuan, M. Yang, and H. Dery, Phys. Rev. B **98**, 125308 (2018).