Discovery of a Novel Linear-in-k Spin Splitting for Holes in the 2D GaAs/AlAs System

Jun-Wei Luo,[1] Athanasios N. Chantis,[2] Mark van Schilfgaarde,[3] Gabriel Bester,[4] and Alex Zunger[1]

[1]*National Renewable Energy Laboratory, Golden, Colorado 80401, USA*
[2]*Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[3]*Arizona State University, Tempe, Arizona 85287, USA*
[4]*Max Planck Institute for Solid State Research, D-70569 Stuttgart, Germany*

(Received 1 October 2009; published 10 February 2010)

The spin-orbit interaction generally leads to spin splitting (SS) of electron and hole energy states in solids, a splitting that is characterized by a scaling with the wave vector $\mathbf{k}$. Whereas for 3D bulk zinc blende solids the electron (heavy-hole) SS exhibits a cubic (linear) scaling with k, in 2D quantum wells, the electron (heavy-hole) SS is currently believed to have a mostly linear (cubic) scaling. Such expectations are based on using a small 3D envelope function basis set to describe 2D physics. By treating instead the 2D system explicitly as a system in its own right, we discover a large linear scaling of hole states in 2D. This scaling emerges from coupling of hole bands that would be unsuspected by the standard model that judges coupling by energy proximity. This discovery of a linear Dresselhaus k scaling for holes in 2D implies a different understanding of hole physics in low dimensions.

DOI: 10.1103/PhysRevLett.104.066405 PACS numbers: 71.70.Ej, 71.15.−m, 73.21.Fg

Spin-orbit interaction causes the energy levels of 3D bulk-solids [1] and 2D quantum wells [2] to exhibit a zero-field spin splitting for sufficiently low-symmetry states. Attention has recently focused on spin of holes in 2D quantum wells (QWs) (because of their spin-Hall effect [3]) and in 0D quantum dots (because of the highly coherent hole spin [4] and unusually long hole spin lifetime [5,6]). On the theoretical side, the long-standing tradition [1,2,7–9] has been to describe hole or electron spin physics in low-dimensional ($<$3D) nanostructures by an expansion in a rather small basis of 3D bulk envelope functions, using effective-mass approaches. In general, when a basis set is restricted, the resolution of the expansion is limited. Low-resolution expansions can be ''far sighted'' [10] in that the actual atomistic symmetry of the 0D, 1D, or 2D object [11] is replaced by a fictitious higher symmetry, thus missing important degeneracy-splitting and interband coupling effects. The farsightedness can be cured by systematically increasing the basis set [12] or by introducing *ad hoc* terms in the Hamiltonian intended to lower the symmetry [13,14]. Both modifications come at the expense of introducing more parameters that are not calculable by the envelope function theory itself. Indeed, in the standard model for spin splitting (SS) of nanostructures [15–17], one uses a phenomenological Hamiltonian where one needs to decide at the outset which 3D bands couple in 2D by the spin-orbit interaction (SOI), rather than have the theory force such realization upon us. The potential of missing important physical interactions not selected to be present in the model Hamiltonian can be substantial [10].

The current state of the art for the hole states in 2D is illustrated by the work of Bulaev and Loss [15]. Starting from a bulk 3D Hamiltonian restricted to heavy-hole (HH) and light-hole (LH) bands (''4×4''), they have derived an effective 2×2 Hamiltonian for the 2D heavy-hole (hh0) subband, demonstrating an exact cancellation of the linear-in-k (Dresselhaus) terms [15]. This result (implying a pure, uncoupled hh0 state in low dimensions) has been used in numerous theories of hole spin in 2D QW and 0D quantum-dot systems, including an estimation of the hole spin-relaxation time [5], demonstration of intrinsic hole spin-Hall effect [18], and other hole spin related phenomena [19–21]. We adopt instead a different approach in which the 2D nanostructures is viewed as a system in its own right, rather than express it in terms of a preselected basis drawn from a reference 3D system. We do so by solving the 2D band structure using explicitly the microscopic potential of the 2D system under consideration, thus freeing us from the need to judge at the outset which selected 3D bands (e.g., 4×4 in Ref. [15]) will couple in 2D. The results show an hitherto unsuspected linear term for holes. This discovery of a linear Dresselhaus k scaling for holes in 2D implies a different understanding of hole-physics in low dimensions.

The central point of the approach utilized here is that the 3D and 2D systems are each described by their own microscopic Hamiltonian which is solved in basis sets whose sole property is that it produces a converged solution to the system at hand. Physical understanding is not sacrificed as it emerges later from the ex post facto analysis of the final result in the language of basis set expansion.

We use a rather general microscopic Hamiltonian in the ''GW representation''

$$H_{GW} = -\frac{1}{2}\nabla^2 + H_{\mathrm{so}} + [V_{\mathrm{ext}} + V_H + \Sigma], \qquad (1)$$

where V_{ext} is the electron-ion potential, V_H is the interelectronic Hartree potential of the specific (3D or 2D) system, $\Sigma = iG^0W$ is the self-energy with W being the screened Coulomb interaction, and $G^0 = 1/(\omega - H^0 \pm i\epsilon)$

is the Green's function of a noninteracting Hamiltonian H^0. The latter is determined self-consistently so as to minimize the difference, within the GW approximation, between the many-body H_{GW} and H^0 [22]. (Thus, exchange-correlation effects are included within the QSGW theory, not LDA). This QSGW approach predicts accurate energy bands for a wide range of materials [23], including the Dresselhaus splitting in bulk GaAs [22,24].

The approach described above is computationally intensive and can be readily applied only to rather small nanostructures. Thus, for computational expediency, when considering larger period quantum wells [e.g., $(\mathrm{GaAs})_n/(\mathrm{AlAs})_n$ with $n \gg 2$], we will map both the small-n behavior and the $n = \infty$ (bulk) QSGW solutions of Eq. (1) to a screened pseudopotential Hamiltonian that captures the former limits yet can be readily applied to orders of magnitude larger systems (10^6 atoms were demonstrated in Ref. [10]):

$$H_{\mathrm{PP}} = -\frac{1}{2}\nabla^2 + H_{\mathrm{so}} + \sum_{n,\alpha,j} v_\alpha(\mathbf{r} - \mathbf{R}_n - \mathbf{d}_{\alpha,j}). \qquad (2)$$

Here, the external (V_{ext}) plus screened ($V_H + \Sigma$) terms of Eq. (1) are described by a superposition of atom-centered functions v_α (where $\mathbf{d}_{\alpha,j}$ is the position of atom j of type α in the n-th cell $\mathbf{R}_n$). They can be constrained to fit approximately yet accurately the QSGW calculated SS of bulk solids ($n = \infty$) [25] and of low-period ($n \sim 2$) 2D QWs. In addition, they reproduce the bulk band gaps and electron and hole effective-mass tensors, as well as band offsets [10,11,25]. The spin splitting Δ_i of band i obtained by the direct calculation of Eqs. (1) and (2) will be fitted to the conventional form $\Delta_i = \alpha_i k + \gamma_i k^3$ for $i =$ electrons (e) or holes (h), allowing both linear and cubic terms to be present.

Results of the many-body multiband calculation.—For 3D bulk GaAs, Fig. 1 shows our results along (110) direction in the Brillouin zone, demonstrating the linear and cubic SS of the three lowest conduction bands (CB1, CB2, and CB3) and three highest valence bands (HH, LH, and SO). We find that the SS of all bands has a cubic term $\gamma_i^{(3D)}$, but only HH, LH, and CB3 with angular momentum $J = 3/2$ have linear term $\alpha_i^{(3D)}$. The screened pseudopotential solution of Eq. (2) gives similar results to QSGW (see inset to Fig. 1). The exception is that Eq. (2) gives $\alpha_i^{(3D)} = 0$ for all bands because this approach is coreless and this term results from coupling to the core states [26]. Fortunately, according to $\mathbf{k}\cdot\mathbf{p}$ description, in 2D α_i^{3D} leads to a term of small magnitude and independent on period n; such small contributions can be safely neglected for small period.

For 2D $(\mathrm{GaAs})_2/(\mathrm{AlAs})_2$ superlattice, the SS of electrons obtained by the atomistic multiband approach agrees well with $\mathbf{k}\cdot\mathbf{p}$ [9] in that the linear term $\alpha_e^{(2D)} \propto \alpha_e^{(3D)}/d^2$ originates from the folding-in of 3D bulk cubic term $\alpha_e^{(3D)}$ due to the confinement to a well of width d. This is not the case for holes. The SS of 2D hole subbands (hh0, lh0, and hh1) is presented in Fig. 2 along the (100) direction. Both atomistic multiband methods [Eqs. (1) and (2)] show: (i) a linear scaling of SS for all three hole subbands including hh0. This is in striking contrast to a cubic-only scaling of hh0 in the model-Hamiltonian derived by Bulaev and Loss [15]. (ii) The linear term dominates the SS at small k; $\alpha_{\mathrm{hh0}}^{(2D)}$, $\alpha_{\mathrm{hh1}}^{(2D)}$, and $\alpha_{\mathrm{lh0}}^{(2D)}$ are comparable. Our results of Figs. 2 and 3 demonstrate a linear-in-k term for 2D hh0 holes, independent of any $\mathbf{k}\cdot\mathbf{p}$ modeling. In what follows, we will attempt to discuss ex post facto some of the features of the predicted linear term in the language of $\mathbf{k}\cdot\mathbf{p}$.

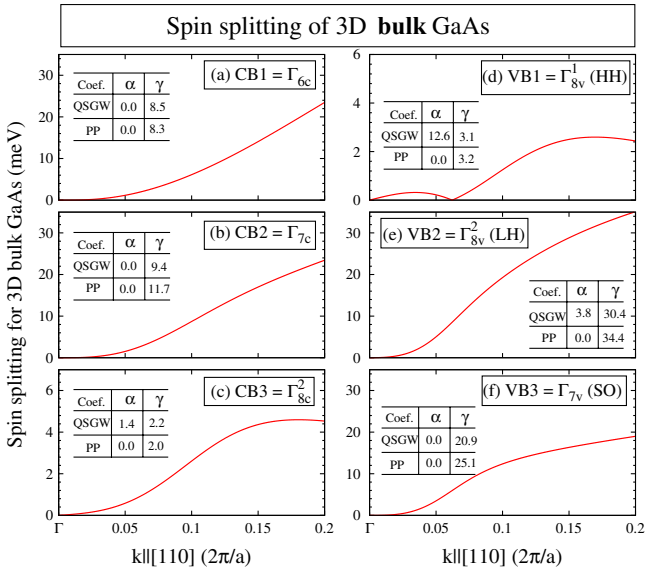

FIG. 1 (color online). QSGW predicted SS of 3D bulk GaAs for the three lowest conduction bands (a) CB1, (b) CB2, and (c) CB3, and three highest valence bands (d) VB1, (e) VB2, and (f) VB3. Dresselhaus constants α_i in meV Å and γ_i in eV Å^3 predicted by QSGW and screened pseudopotential fit to QSGW, respectively, are given for each band in the inset. We give the absolute value of α and γ.

The standard 2×2 model Hamiltonian for 2D does not explain the results. Bulaev and Loss [15] have shown that starting from a 4×4 basis in 3D, there is no linear terms for hh0 in the 2D model Hamiltonian. However, they did not include the 3D relativistic cubic terms of Γ_{8v} bands in their derivation. Rashba and Sherman [17] demonstrated earlier that the 3D relativistic cubic terms can give rise to 2D linear term for hh0 subbands [Eq. (8) in Ref. [17]]. We tested this idea calculating the 2D SS using a pseudopotential which was constructed to fit the Rashba-Sherman 3D band structure including the relativistic cubic terms. The results of the pseudopotential calculation for 2D (in which all bands are allowed to couple) are compared with the Rashba-Sherman 2D model (in which the 2D hh0 band is uncoupled). Figure 3 shows that for sufficiently large superlattice periods n (for which the 2D model of Ref. [17] is applicable), the model recovers only a small fraction of the multiband results for linear SS of 2D hh0.

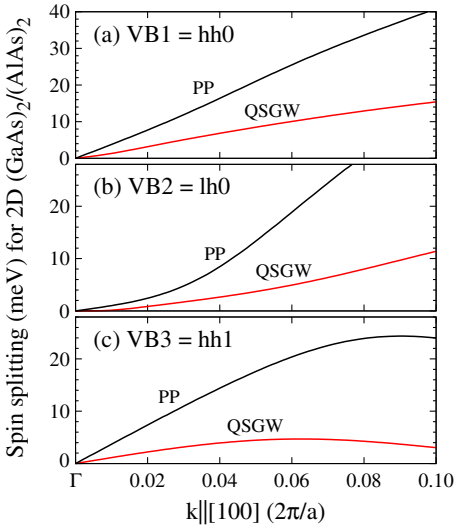

FIG. 2 (color online). QS*GW* predicted SS of 2D $(GaAs)_2/(AlAs)_2$ superlattice for the three highest valence subbands (a) hh0, (b) lh0, and (c) hh1. The corresponding SS calculated by screened pseudopotential fit to model Hamiltonian result [17] is also given in black line for each subband.

For smaller periods ($n \leq 20$), we report in Fig. 3(a) a clear nonmonotonic period dependence of both $\alpha_{hh0}^{(2D)}$ and $\alpha_{lh0}^{(2D)}$, with $\alpha_{hh0}^{(2D)}$ to $\alpha_{lh0}^{(2D)}$ ratios varying from $\sim$10 to $<$1 as the period decreases from 50 to 2 ML. This is in sharp contrast with the predictions of the model Hamiltonian [17] which predicts a monotonic increase of linear terms and a ratio of $\alpha_{hh0}^{(2D)}$ to $\alpha_{lh0}^{(2D)}$ which is independent of d.

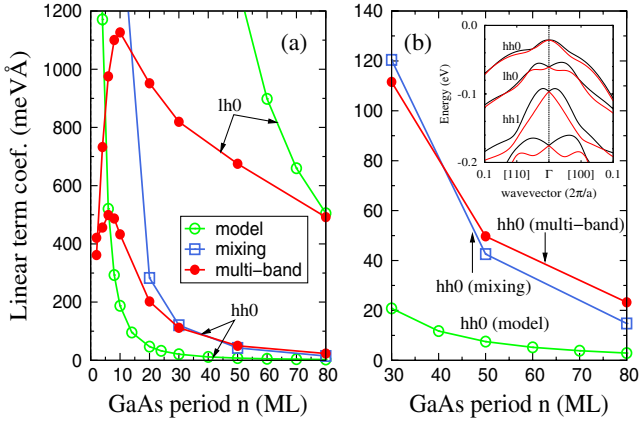

FIG. 3 (color online). (a) Comparison of linear SS coefficients $\alpha_{hh0}^{(2D)}$ and $\alpha_{lh0}^{(2D)}$ calculated from present pseudopotential multiband approach (solid circles), as well as by the Rashba-Sherman model Hamiltonian (open circles), and mixing approximation [Eq. (4); open squares] for 2D $(GaAs)_n/(AlAs)_{20}$ QWs. (b) Expanded scale for long-period QW's highlighting the comparison of hh0 subbands for the direct calculation vs model Hamiltonian and mixing approximation. Inset shows energy dispersion of valence subbands for 2D $(GaAs)_{20}/(AlAs)_{20}$ calculated by atomistic pseudopotential approach. Each orbit subband has two spin subands.

Sources and magnitude of coupling between 3D bands.—The quantitative and qualitative disagreements of atomistic multiband calculation with the standard model Hamiltonian approach suggest possible undiscovered SOI terms. Such terms due to symmetry lowering down from 3D T_d bulk symmetry to 2D D_{2d} QW symmetry could originate from coupling of 3D bands via the 2D potential and SOI (e.g., a linear-in-k dependence can be generated by symmetry lowering via strain [26]). Such coupling is signaled by the distinctly nonparabolic 2D energy dispersion curves manifesting clear anticrossings between neighboring subbands [see multiband calculation in the inset to Fig. 3(b)]. This inset shows also that the lh0 level lies between the hh0 and hh1 levels in 2D which is true for all multiband-calculated GaAs periods. Clearly, the coupling of 3D states and its effects on SS of 2D bands can not be ignored.

Indeed, zero coupling between 3D HH and LH states at $k_{\|} = 0$ is often assumed, although the 2D potential and SOI introduce such couplings. If one starts from the 3D 8×8 Kane model, neglecting the relativistic terms, which we find to be small, and reduces it from 3D to 2D, one gets a Hamiltonian with no HH-LH coupling at all, thus no linear term. The reason is that this restricted $\mathbf{k} \cdot \mathbf{p}$ basis set does not "see" the real symmetry of the subband in 2D [10]. In such model Hamiltonian approaches [15,17], the hh0, hh1, ..., wave functions near the zone-center in 2D all derive from a single bulk state $|HH\rangle$ and similarly, all lh0, lh1 ..., wave functions in 2D derive from a single bulk state $|LH\rangle$. A better approach than this "band decoupled" model Hamiltonian [15,17] approximation would allow the 2D state hh0 to derive from a few bulk states. In such a "mixing of decoupled states" approximation,

$$\Psi_{hh0}^{(2D)} = w_{hh0}^{(2D)}(HH)|HH\rangle + w_{hh0}^{(2D)}(LH)|LH\rangle + \cdots, \quad (3)$$

where $w_{hh0}^{(2D)}(\lambda)$ is the percent weight of 3D state $\lambda =$ HH, LH, ... in the 2D state $\Psi_{hh0}^{(2D)}$. In the band decoupled model, $w_{hh0}^{(2D)}(\lambda) \equiv 0$ for $\lambda \neq$ HH. We have calculated the weights by numerical projection of the 2D pseudopotential wave functions and show them in Fig. 4. The interband coupling is large even at the zone center. We see that whereas for long periods (GaAs thickness $\geq$ 20 ML) the mixing is small (i.e., $\Psi_{hh0}^{(2D)}$ is made of 90% $|HH\rangle$ and 5% $|LH\rangle$), for shorter periods, the mixing increases (the LH content of $\Psi_{hh0}^{(2D)}$ increases to 10–20%).

As expected, the interband mixing will approach zero as the period n goes to infinity, but the convergence is very slow and nonmonotonic. This nonmonotonic mixing is evident in Fig. 4 by the increase of 3D-HH character in the 2D lh0 state as the period increases from 50 to 80 ML. The reason is that the magnitude of the mixing is inversely proportional to the QW period and to the energy splitting of 2D hh, lh, and so subbands. But the latter splitting is also inversely proportional to the QW period. This nonmonotonicly enhanced mixing of LH and HH into the 2D hh0 as

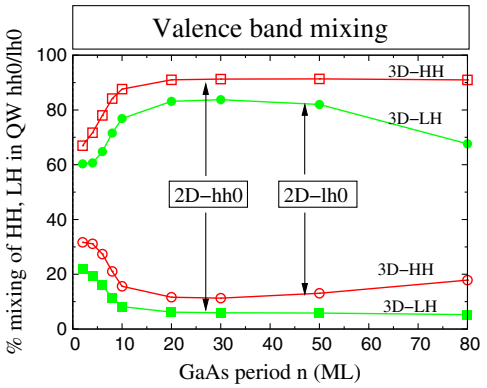

FIG. 4 (color online). Percent of bulk HH (open symbols) and LH (filled symbols) bands mixed into the 2D hh0 (squares) and lh0 (circles) subbands at zone-center calculated by numerical projection. The results are almost same in a small in-plane k range [e.g., $k < 0.01(2\pi/a)$]. Note that the sum of bulk HH and LH content is less than 100%.

the QW period is reduced signals the breakdown of the model-Hamiltonian-thinking that neglects such mixing on the ground that the energy splitting of hh0-lh0 must be larger than that of hh0-hh1 for sufficient small periods [15].

Analysis of linear term.—The linear coefficient of 2D hh0 SS can be written in terms of the weights in Eq. (3) in a model of "mixing of decoupled states" as

$$\alpha_{\mathrm{hh0}}^{(2\mathrm{D})} = -w_{\mathrm{hh0}}^{(2\mathrm{D})}(\mathrm{HH})\tilde{\alpha}_{\mathrm{hh0}}^{(2\mathrm{D})} + w_{\mathrm{hh0}}^{(2\mathrm{D})}(\mathrm{LH})\tilde{\alpha}_{\mathrm{lh0}}^{(2\mathrm{D})} + \cdots, \quad (4)$$

where $\tilde{\alpha}_{\mathrm{hh0}}^{(2\mathrm{D})}(\tilde{\alpha}_{\mathrm{lh0}}^{(2\mathrm{D})})$ is the contribution of a single bulk HH (or LH) band to linear SS of 2D hh0 (1h0), which had been derived by Rashba and Sherman [17] (the negative sign accounts for band repulsion effect). The result of the first two terms in Eq. (4) is shown as open squares in Fig. 3 and is compared with the multiband-calculated $\alpha_{\mathrm{hh0}}^{(2\mathrm{D})}$ (solid circles). We see that the mixing of decoupled states [Eq. (4)] gives a much better approximation to the full calculation than the model Hamiltonian treating one decoupled band at a time (open circles in Fig. 3) {see note (i) in supplementary material [27]}. Going from 3D to 2D entails simultaneously (i) a symmetry lowering (from T_d to D_{2d}), as well as (ii) an increase in HH-LH splitting due to different confinement of HH and LH. One might think that whereas (i) will encourage a linear term (due to additional HH-LH coupling at lower symmetry), effect (ii) might discourage it because it presumably reduces HH-LH mixing which is argued to be the cause of the linear term. The traditional approach is to assume that the second effect dominates, and the coupling effects will be small (can be treated perturbatively [15], or neglected altogether [16]). However, realistic calculations disprove this intuitive guess. Figure 4 shows the degree of HH-LH coupling in a realistic QW demonstrating that whereas this mixing is rather small in 2D hh0 (5% 3D-LH character for $n \geq 20$), the ensuing effect on SS is nevertheless very large, since $\tilde{\alpha}_{\mathrm{lh0}}^{(2\mathrm{D})}$ is significantly larger than $\tilde{\alpha}_{\mathrm{hh0}}^{(2\mathrm{D})}$. Thus, the mixing of bulk bands leads to a large linear SS of 2D hh0, and is unsuspected by the standard model that judges coupling by energy proximity {see note (ii) in supplementary material [27]}.

The emergence of a large linear term for Dresselhaus hole SS in 2D nanostructures suggests (i) the dominance of Dresselhaus over Rashba SOI (having a cubic term as its lowest order term) [15], (ii) a larger spin-Hall effect [18], and (iii) an explanation of the observed large optical anisotropy [14]. The occurrence of a larger SS of hh0 corresponding to HH-LH coupling leads to a shorter hole spin-relaxation time in 2D QWs [6] from the D'Yakonov and Perel (DP) mechanism [7].

A. Z. thanks E. Rashba and D. Loss for helpful discussions on this subject. Research at NREL supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, under Contract No. DE-AC36-08GO28308. M. v S. was supported by ONR, Project No. N00014-07-1-0479 and NSF No. QMHP-0802216.

[1] G. Dresselhaus, Phys. Rev. **100**, 580 (1955).
[2] Y. A. Bychkov *et al.*, J. Phys. C **17**, 6039 (1984).
[3] S. Murakami *et al.*, Science **301**, 1348 (2003).
[4] D. Brunner *et al.*, Science **325**, 70 (2009).
[5] D. Heiss *et al.*, Phys. Rev. B **76**, 241306(R) (2007).
[6] I. A. Yugova *et al.*, Phys. Rev. Lett. **102**, 167402 (2009).
[7] *Optical Orientation*, edited by B. P. Zakharchenya and F. Meier (North-Holland, Amsterdam, 1984).
[8] J. W. Wang *et al.*, Appl. Phys. Lett. **92**, 012106 (2008).
[9] R. Eppenga *et al.*, Phys. Rev. B **37**, 10923(R) (1988).
[10] A. Zunger, Phys. Status Solidi A **190**, 467 (2002).
[11] R. Magri *et al.*, Phys. Rev. B **64**, 081305(R) (2001).
[12] S. Richard *et al.*, Phys. Rev. B **70**, 235204 (2004).
[13] E. L. Ivchenko, *et al.*, Phys. Rev. B **54**, 5852 (1996).
[14] B. A. Foreman, Phys. Rev. Lett. **86**, 2641 (2001), and references therein.
[15] D. V. Bulaev and D. Loss, Phys. Rev. Lett. **95**, 076805 (2005).
[16] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer, New York, 2003).
[17] E. I. Rashba and E. Ya. Sherman, Phys. Lett. A **129**, 175 (1988).
[18] B. A. Bernevig *et al.*, Phys. Rev. Lett. **95**, 016801 (2005).
[19] M. F. Borunda *et al.*, Phys. Rev. Lett. **99**, 066604 (2007).
[20] D. Stepanenko*et al.*, Phys. Rev. B **79**, 235301 (2009).
[21] O. Olendski *et al.*, Phys. Rev. B **77**, 125338 (2008).
[22] A. N. Chantis*et al.*, Phys. Rev. Lett. **96**, 086405 (2006).
[23] M. van Schilfgaarde *et al.*, Phys. Rev. Lett. **96**, 226402 (2006).
[24] J. J. Krich *et al.*, Phys. Rev. Lett. **98**, 226802 (2007).
[25] J. W. Luo *et al.*, Phys. Rev. Lett. **102**, 056405 (2009).
[26] M. Cardona *et al.*, Phys. Rev. B **38**, 1806 (1988).
[27] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevLett.104.066405 for detail derivation of model Hamiltonian of 3D and 2D semiconductors and two notes.